ISSN 0233-7657
Biopolymers and Cell. 2015. Vol. 31. N 1. P. 63–71
doi: http://dx.doi.org/10.7124/bc.0008CE

UDC 577.214.6

Bioinformatics analysis of cis-regulatory elements in *Mbl1* and *Mbl2* genes in *Rattus norvegicus*

V. S. Bondarenko[1, 2], M.Yu Obolenska[1]

[1] Institute of Molecular Biology and Genetics, NAS, Ukraine;
Zabolotnoho Str. 150, Kyiv, Ukraine, 03680

[2] ESC «Institute of Biology», Taras Shevchenko National University
64/13, Volodymyrska Str. Kyiv, Ukraine, 01601
m.obolenska@gmail.com

Aim. To identify and characterize with the help of bioinformatics the transcription factors binding sites in promoters of *Mbl1* and *Mbl2* genes, encoding mannose binding lectins in *Rattus norvegicus*. ***Methods.*** Bioinformatics, MatInspector software. The position weight matrices of transcription factor binding sites were obtained from the Matrix Family Library Version 9.0. Within the frame of the program we selected the binding sites, the cognate transcription factors of which are specifically expressed in the liver and immune cells, and have passed the filter for conservation after comparison with the binding sites in orthologous genes in *Mus musculus*. ***Results.*** The promoters of both genes share the binding sites for the members of the four common families of transcription factors (HNF, homeodomain transcription factors, GRs and ETS factors). The promoters of *Mbl1* and *Mbl2* gene possess correspondingly additional binding sites for the members of six (AP1 related factors, Ccaat/Enhancer Binding Proteins, FOX, p53, NFAT, ISGF3) and four (cAMP-responsive element binding proteins, heat shock factors, Nf-κB/c-rel and TATA binding protein) families of transcription factors. The *Mbl1* specific transcription factors are mainly involved in the regulation of differentiation, development, metabolic homeostasis, organogenesis and cell cycle. Unlike them the *Mbl2* specific transcription factors are more prone to mediate a stress-response. ***Conclusion.*** The variety of transcription factors potentially involved in regulation of the *Mbl1* and *Mbl2* transcription argue for these genes involvement in various cellular processes with specific role of each gene. The obtained results provide the basis for the task-oriented wet-lab bench experiments on their regulation.

Keywords: mannose binding lectins, position weight matrices, transcription factor binding sites.

Introduction

Mannose-binding lectin (Mbl) is a recognition molecule in the lectin pathway of a complement cascade. It binds directly to carbohydrate patterns on the surface of bacterial, fungal, parasitic cells, certain viruses and apoptotic, senescent, injured and transformed host cells and maintains tolerance to self-antigens of healthy host cells [1–4]. The bound Mbl attracts Mbl-associated serine proteases (MASPs) and initiates the lectin pathway of a complement system [5, 6]. There are three pathways of a complement cascade – lectin, classical and alternative. Their initial steps are somewhat different but all three of them converge on C3 convertase, C4b2b. The subsequent complement cascade results in eventual opsonisation of particles and their engulfment by phagocytes, lysis of pathogens, attraction of phagocytic cells through chemotaxis, release of inflammatory peptides and cytokines [4, 7–9]. Lectin pathway unlike other complement pathways is crucial for the prompt answer for pathogen invaders and altered host cells till the acquired (adaptive) immune system gains momentum.

The largest amount of Mbl is produced in the liver and secreted to the circulation. Underproduction of Mbl increases the susceptibility to the infection but increases the resistance to malaria whereas overproduction is injurious during sepsis, shock, ischemia/reperfusion and rejection of transplanted solid organs, etc. [8, 10–12]. The genetic differences in both structural and regulatory parts of the human gene define the altered response of Mbl to stress [13]. Thereby Mbl is considered as a perspective target in critical care and other fields of medicine. Several drugs are already designed and implemented in clinical practice [14].

In our previous study we have predicted the presence of interferon stimulated response element (ISRE) in the promoter of *Mbl1* gene in *Rattus norvegicus* and hence the *Mbl1* gene as a previously unknown gene of potential response to IFNα. Shortly the search was conducted in a rat genome by the conservation-aided transcription factor binding sites finder (http://biomed.org.ua/ COTRASIF/about.html) and the obtained genes were classified for experimentally confirmed and not confirmed IFNα targets [15–17]. In accord with this prediction the Mbl insufficiency is associated with the resistance of patients with hepatitis C to the IFNα treatment [18]. Both, MBL and IFNα, are the members of innate immunity system and all these evidences argue for their more tight cooperation. The regulation of expression of the *Mbl* genes is scarcely investigated.

Prior to start the wet-lab bench experiments we decided to analyze in more details the promoter of *Mbl1* gene and to include into analysis the *Mbl2* gene as the general Mbl activity in rats is provided by two genes.

The bioinformatics analysis of cis-regulatory elements in the promoters of *Mbl1* and *Mbl2* genes has

Table 1. **Transcription factors binding sites found in *Mbl1* promoter**

Family	Transcription factor name	Position from -to	Strand	Sequence
MAF and AP1 related factors	V-maf musculoaponeurotic fibrosarcoma oncogene homolog K (half site)	-426; -402	-	tcaaacaTGCTtactgactac
Ccaat/Enhancer Binding Proteins	CCAAT/enhancer binding protein beta	**-920; -906**	-	accatttaGAAAtat
	Ccaat/Enhancer Binding Protein	-716; -702	-	gggatttgGTAAgta
Glucocorticoid responsive and related elements	Androgene receptor binding site, IR3 sites	-485; -467	+	tcgGAACacagtgtctccg
Human and murine ETS1 factors	Spi-B transcription factor (Spi-1/PU.1 related)	-120; -100	-	tgtaattgGGAAgctgtccag
	ETV4 (Ets variant gene 4)	-15; 5	+	attgaggAGGAagctaccaga
	GA binding protein	**22; 42**	-	ggagcagtGGAAgagaaactg
Fork head domain factors	FHXA and FHXB	-340; -324	+	cccaagATAAcatttgc
	FOXP1	-310; -294	-	ttgtaaaAACAatcaat
	Fork head related activator-2 (FOXF2)	-307; -291	-	ttattgTAAAaacaatc
Hepatic Nuclear Factors	Hepatic Nuclear Factor 1	-880; -864	+	aactggtaGTTAttctg
	Hepatic Nuclear Factor 6		-	aaaaacAATCaatagag
	CUT-homeodomain transcription factor Onecut-2	-314; -298		
Homeodomain transcription factors	NOBOX	-114; -96	+	agcttcccAATTacaacag
	NOBOX	-115; -95	-	gctgttgTAATtgggaagc
Interferon regulatory factors	Interferon regulatory factor 6 (ISGF)	**44; 64**	+	ttactttgacGAAAccctagt
Nuclear factor of activated T-cells (NFAT)	Nuclear factor of activated T-cells	**-965; -926**	+	tatttGGAAatagagtaga
	Nuclear factor of activated T-cells 5	-764; -746	+	gtgGGAAaccagcaggcta
Tumor suppressor p53	p53 tumor suppressor	-583; -517	+	catgtagtcagtaagCATGtttgaa

respectively revealed the transcription factors binding sites (TFBSs) for the members of ten and eight partly overlapping families of transcription factors. The variety of transcription factors potentially involved in the regulation of expression of the *Mbl1* and *Mbl2* genes argue for the genes participation in various cellular activities. The specific transcription factors characteristic for each gene point to differential regulation of their expression. During evolution the *Mbl1* and *Mbl2* genes in *Rattus norvegicus* have co-opted the ancient DNA binding motifs and acquired the «younger» ones to perform their versatile functions.

Materials and Methods

The sequences of 1100 bp (from –1000 to + 100 bp), containing adjacent 5'UTR of *Mbl1* and *Mbl2* genes of *Rattus norvegicus* (Gene IDs: 24548 and 64668) and *Mus musculus* (Gene IDs: 17194 and 17195) were chosen for the search for TFBS and putative transcription factors regulating the *Mbl* genes expression via these TFBS. The promoter analysis was conducted with Genomatix software tools (http:// www.genomatix.de/index.html).

The search for TFBSs was undertaken with the help of the program MatInspector that is based on the usage of position weight matrices (PWM) representing the complete nucleotide occurrence probabilities and information content for each position in the sequence [19, 20]. The selection of TFBSs was initialized at recommended threshold > 0.85 for the core similarity and at the «optimized» one for the family matrix similarity. We used the special options and selected those individual matches from each family that have the scores of matrix and core similarities > 0.8 and were associated with TFs «expressed in liver», «expressed in immune system» and «ubiquitous» given that the liver successfully combines the role of a «biochemical laboratory» with the role of a major organ of innate immunity. In a physiologically relevant state, it eliminates the antigens derived from the gastrointestinal tract, aging and transformed cells and reveals tolerance to their continuous presence [21]. The resulting list (List #1) contained matrix families with individual binding sites in correspondence with the cell specificity of TFs, that potentially regulate gene expression via these binding sites.

According to the assumption that transcriptional regulation of close orthologous genes has been evolutionarily maintained to control specific gene expression patterns we conducted the search for conservative TFBS by comparing the *Mbl1* and *Mbl2* promoters of *Rattus norvegicus* with those of *Mus musculus.* The DiAlign TF program of GEMS launcher (http://www.genomatix.de/cgi-bin//gems/ launch.pl?s =8257 ac78f55b0c8449b1858b1db438 d6; GE MS=1;TASK= dialign_TF) was used to align pairwise the sequences and to display TFBSs matches within alignment. By default, the TFBSs are considered as conservative if they are identical at least for 85 % and located at the same position within the alignment. This search was made at the core similarity ≥ 0.85 and at the optimized threshold for the matrix similarity. The List #2 contained the evolutionary conservative matches in promoters of both genes irrespective of their cell/tissue specificity. The intersecting of lists #1 and #2 yields liver- and immune specific evolutionary conserved TFBSs of the *Mbl1* and *Mbl2* genes in *Rattus norvegicus.*

Results and Discussion

The specific pattern of transcription factors bound to their cognate *cis*-acting regulatory DNA sequences interacts with the transcriptional machinery and enables selective gene expression. The *in silico* search for TFBSs in the promoters of *Mbl1* and *Mbl2* genes has revealed the cognate sites for the members of ten and eight transcription factors families respectively (Tables 1, 2). The disclosure of the set of TFBSs in the *Mbl* genes gives an idea which transcription factors might regulate the expression of these genes, and accordingly in which TF-mediated different orchestras the gene of interest may be a player. Whether it plays or not and in which orchestra – it is a question. The answer depends on the context – the kind of stimulus, cell and species specificity, stage of development etc.

According to the set of TFBSs, the *Mbl1* and *MBL2* genes in the liver may be regulated by the transcription factors referring to four families shared by both genes – HNF, homeodomain transcription factors, GRs and ETS factors (Table 3). The mem-

bers of these families and their disposition and collocation in both promoters are characteristic for each gene. They are mainly responsive for the regulation of differentiation, development, organogenesis, hematopoesis, proliferation of immune cells [22–25] (see Supplement for detailed information about TFs at http://dx.doi.org/10.7124/bc.0008CE).

Six families of TFs are characteristic for the *Mbl1* promoter (AP1 related factors, Ccaat/Enhancer Binding Proteins, FOX, p53, NFAT, ISGF3). Among them two pairs of TFBSs are localized in close vicinity – NFAT and C/EBP (906–920 bp and 926–965 bp) and ETS and ISGF factors (22–42 bp and 42–46 bp). Such disposition of binding sites makes the basis for the binding of composite regulatory elements, NFAT-C/EBP and ETS-ISGF. The specified NFAT-C/EBP pairs of TFs are already known as functionally active composite regulators in several genes e.g. in peroxisome proliferator-activated receptor-gamma 2 gene, insulin-like growth factor 2, angiotensin-converting enzyme homolog, and transcription factor POU4F3 genes [26]. Specific transcription factors containing ETS domains readily synergize with the interferon regulated factors making the composite regulators [27]. This notion supports the plausible responsiveness of *Mbl1* to IFNα. Four families of TFs are typical for the *Mbl2* gene (cAMP-responsive element binding proteins, Heat shock factors, Nf-κB/c-rel and TATA binding protein).

Comparing the *Mbl1* and *Mbl2* specific sets of TFBSs the general functional difference between

Table 2. **Transcription factors binding sites found in *Mbl2* promoter**

Family	Transcription factor name	Position in promoter	Strand	Sequence
cAMP-responsive element binding proteins	E4BP4, bZIP domain, transcriptional repressor	-550; -530	+	acaccttaagGTAAgagaacc
Glucocorticoid responsive and related elements	Androgene receptor binding site, IR3 sites	-984; -966	-	gaataactaccaGTTCtca
Heat shock factors	Heat shock factor 1	-129; -117	+	ggtaacctttcatgatttTTCTtac
	Heat shock factor 2	-120; -96	-	caccttgacgtaAGAAaaatcatga
Human and murine ETS1 factors	Spi-B transcription factor (Spi-1/PU.1 related)	-908; -888	+	atgaaattGGAAgacagtatc
	SPI-1 proto-oncogene; hematopoietic transcription factor PU.1	-301; -281	+	tgtgtagaGGAAgtgtgcgca
	v-ets erythroblastosis virus E26 oncogene homolog	-243; -223	+	gtaagcaaGGAAattgacatg
	Ets - family member ELF-2 (NERF1a)	48; 68	-	aggggcaaGGAAgagtctctt
Hepatic Nuclear Factors 1	Hepatocyte nuclear factor 1 alpha (Tcf-1)	-826; -810	-	tctaagaaGTTAaatgt
Homeodomain transcription factors	Brain specific homeobox	-871; -853	+	catatgatAATTattgcat
	Brain specific homeobox	-870; -852	-	tatgcaatAATTatcatat
	Muscle segment homeo box 2, homologue of Drosophila (HOX 8)	-186; -168	-	actttgCTAAttctcatga
	H6 homeodomain HMX3/Nkx5.1 transcription factor	-179; -161	+	aattagcaAAGTggatgtc
Nuclear factor kappa B/c-rel	c-Rel	-466; -452	-	tcaggcttTTCCtct
Vertebrate TATA binding protein factor	C-type LTR TATA box	-158; -142	-	cagtatgTATGtaccaa

their transcription factors may be noted. In the *Mbl1* gene they are mainly responsible for the regulation of differentiation, organogenesis, metabolic homeostasis (AP-1, C/EBP, FOX) [28–30], T-cell differentiation and self-tolerance (NFAT) [31], regulation of cell cycle (p53) [32] and responsiveness to IFNα (ISGF3). Unlike them the *Mbl2* specific transcription factors are more prone to stress response. Heat shock factors are essential for all organisms to survive the exposures to acute stress [33]. The cAMP-response element binding proteins possess kinase inducible element(s) in their transactivation domain that makes them susceptible to the modification by phosphorylation in response to a diverse array of stimuli [34]. Nf-κB is found in almost all animal cell types and is activated in cellular responses to stimuli such as stress, cytokines, free radicals, ultraviolet irradiation, oxidized low density lipoproteins, and bacterial or viral antigens [35, 36]. TATA box subsumes the *Mbl2* gene under the general category of TATA-containing genes with their specific characteristics.

TBP protein binds TATA box in core promoter of genes. Together with RNA polymerase II and general transcription factors it forms the pre-initiation complex. TATA-containing genes depend more strongly on the SAGA complex (Spt-Ada-Gcn5-Acetyltransferase) whereas TATA-less genes – on the TFIID complex-dominated TBP binding [37]. TATA-containing genes are characterized by a propensity for being subtelomeric, expressed at extremely high or low levels, stress-induced, and under evolutionary selective pressure [38, 65]. The TATA box usually evolves as a gain process in result of gene duplication. The duplicated genes are enriched in TATA-containing genes [38]. The TATA box occurs in approximately 10.7 and 11.2 % of the protein encoding genes respectively in the mice and human genomes [39].

The *Mbl1* and *Mbl2* genes are paralogues arising from the gene duplication. According to our calculations, the nonself BlastP hit has an E-value less than an E-value cutoff of $1.0 \cdot 10^{-20}$ that supports the mentioned notion. Like a typical TATA containing gene, *Mbl2* is localized in the subtelomeric region of chromosome 1 in the rat genome (see NCBI Map viewer, annotation release 105). Also the *Mbl2* gene was/is under evolutionary selective pressure as Old World monkeys still have both genes whereas chimpanzees like humans have only one functionally active gene *MBL2* homologous to the rat *Mbl2* gene. The *Mbl1* isoform in rodents and some primates is homologous to the human *MBL1* pseudogene with low level expression of truncated protein [40]. Lynch and Conery [41] have suggested that the duplication of genes is a relatively frequent event in evolution and typically these duplicated genes are lost because of the lack of selective pressure to maintain both copies.

Therefore on the basis of carried out analysis we may suggest that the *Mbl1* and *Mbl2* genes may differentially respond to the intra- and extracellular factors with *Mbl2* more prone to the stress-induced response. Our preliminary data have revealed that the *Mbl2* gene is induced to substantially greater extent than the *Mbl1* gene in the liver of rats treated with IFNα. The presence of Nf-κB binding site in promoter of the *Mbl2* gene may define the *Mbl2* gene responsiveness to IFNα via the PI3 kinase pathway [42].

Complement has long been appreciated as a rapid and local immune surveillance system. However, new research has ascribed many new functions of complement that extend far beyond host defense and inflammatory processes [reviewed in 8, 43 and 44]. The variety of TFBSs in the promoters of both genes partly explains this versatility and provides the basis for new associations between the lectin complement system and other systems in organism that may be ahead.

The lectin pathway of the complement cascade is an evolutionary ancient form of immune system. It

Table 3. **The distribution of TFBSs between promoters of *Mbl1* and *Mbl2* genes**

Character of distribution	Families of transcription factors
Common	Hepatocyte nuclear factors, Homeodomain TFs, Glucocorticoid receptors, ETS factors
Mbl1 – specific	AP1 related factors, Ccaat/Enhancer Binding Proteins, FOX, p53, NFAT, ISGF3
Mbl2 – specific	cAMP-responsive element binding proteins, Heat shock, NF- κB/c-rel, TATA-binding protein

emerges in the early Cambrian in Tunicates [4, 8] and developed in parallel with the transcription factors predating complement appearance, developed simultaneously with it and emerged later during evolution process. We have screened the literature to check whether both genes differ in the «age» of their TFs, particularly their DNA binding motifs. The «youngest» motifs refer to the TFs common for both genes. ETS domain was detected in Porifera [45]; homeodomain – in gastropod mollusks referring to late Cambrian [46]; glucocorticoid receptor – in Chondrichthyes or cartilaginous fishes which appeared about 395 million years ago, during the middle Devonian [47].

Most of DNA binding motifs of TFs that might regulate the expression of *Mbl1* gene are more ancient except NFAT and IFN system. bZip motif inherent to the AP1 related factors and Ccaat/Enhancer Binding Proteins was detected in the period prior to the divergence of the metazoa and fungi that is long before the emergence of lectin pathway [48]. The p53 superfamily predates animal evolution and first appears in unicellular Flagellates. In the invertebrate models amenable to genetic analysis , the p53 superfamily members mainly act in apoptosis regulation in response to genotoxic agents and do not have overt developmental functions [49]. The appearance of the winged-helix FOX motif refers to the same period [50]. NFAT domain was detected much later at Cephalochordata, jawless or agnathan fishes as the earliest known members of Vertebrata that appeared during the Ordovician Period (510–439 Mya) [51]. The IFN system originated in early vertebrates (ca. 385 millions years ago in the Devonian Period) and is conserved in all tetrapods as well as in fishes but not in *Tunicates* [52, 53].

The DNA binding motifs of TFs that might regulate expression of the *Mbl2* gene arose also early in evolution. About bZip motif of cAMP-responsive element binding proteins see above. Heat shock factors and chaperons originate in evolution approximately 3.5 billion years ago, because they are present in archaebacteria as well as in bacteria [54]. One of the major destructive stresses faced by all cells was the problem of reactive oxygen following the achievement of high levels of atmospheric oxygen approximately 2.2 bya. The nuclear transcription factor Nf-κB is induced by oxidative stress and functions to protect diverse cells from apoptotic events. Defense systems containing homologous elements, in fact, are widely distributed among plants, protozoans, echinoderms, protostomes, lower vertebrates and mammals [54].

Therefore, the *Mbl1* and *Mbl2* genes in *Rattus norvegicus* co-opted the ancient DNA binding motifs and acquired the «younger» ones to perform their versatile functions.

Conclusion

The variety of transcription factors potentially involved in regulation of the *Mbl1* and *Mbl2* transcription argue for these genes involvement in various cellular processes with specific role of each gene. According to the TFBSs specific for each of genes both genes might be differentially regulated – *Mbl1* by TFs regulating mainly the processes of differentiation, organogenesis, cell cycle, homeostasis, and *Mbl2* — by stress-induced factors. The *Mbl1* and *Mbl2* genes in *Rattus norvegicus* co-opted the ancient DNA binding motifs and acquired the «younge» ones to perform their versatile functions.

REFERENCES

1. *Turner MW.* Mannose-binding lectin: the pluripotent molecule of the innate immune system. *Immunol Today*. 1996; **17**(11):532–40.
2. *Takahashi K, Yamanaka S*. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006;**126**(4):663–76.
3. *Stuart LM, Takahashi K, Shi L, Savill J, Ezekowitz RA*. Mannose-binding lectin-deficient mice display defective apoptotic cell clearance but no autoimmune phenotype. *J Immunol*. 2005;**174**(6):3220–6.
4. *Fujita T, Endo Y, Nonaka M*. Primitive complement system – recognition and activation. *Mol Immunol*. 2004;**41** (2–3): 103–11.
5. *Matsushita M, Fujita T.* Activation of the classical complement pathway by mannose-binding protein in association with a novel C1s-like serine protease. *J Exp Med*. 1992;**176** (6):1497–502.
6. *Schwaeble W, Dahl MR, Thiel S, Stover C, Jensenius JC*. The mannan-binding lectin-associated serine proteases (MASPs) and MAp19: four components of the lectin pathway activa-

tion complex encoded by two genes. *Immunobiology*. 2002; **205**(4–5):455–66.
7. *Takahashi K, Chang WC, Takahashi M, et al.* Mannose-binding lectin and its associated proteases (MASPs) mediate coagulation and its deficiency is a risk factor in developing complications from infection, including disseminated intravascular coagulation. *Immunobiology*. 2011; **216**(1–2): 96–102.
8. *Ricklin D, Hajishengallis G, Yang K, Lambris JD.* Complement: a key system for immune surveillance and homeostasis. *Nat Immunol*. 2010;**11**(9):785–97.
9. *Takahashi K, Ip WE, Michelow IC, Ezekowitz RA*. The mannose-binding lectin: a prototypic pattern recognition molecule. *Curr Opin Immunol*. 2006;**18**(1):16–23.
10. *Farrar CA, Sacks SH*. Mechanisms of rejection: role of complement. *Curr Opin Organ Transplant*. 2014;**19**(1):8–13.
11. *Huber-Lang M, Kovtun A, Ignatius A*. The role of complement in trauma and fracture healing. *Semin Immunol*. 2013; **25**(1):73–8.
12. *Jha AN, Sundaravadivel P, Singh VK, et al.* MBL2 variations and malaria susceptibility in Indian populations. *Infect Immun*. 2014;**82**(1):52–61.
13. *Herpers BL, Endeman H, de Jong BA, et al.* Acute-phase responsiveness of mannose-binding lectin in community-acquired pneumonia is highly dependent upon MBL2 genotypes. *Clin Exp Immunol*. 2009;**156**(3):488–94.
14. *Larghi EL, Kaufman TS*. Modulators of complement activation: a patent review (2008–2013). *Expert Opin Ther Pat*. 2014;**24**(6):665–86.
15. *Tokovenko BT, El'skaya AV, Obolenskaya MYu*. *In silico* approach to study and functionally analyze interferon regulated genes. *Biopolym Cell*. 2007;**23**(4):368–75.
16. *Tokovenko B, Golda R, Protas O, Obolenskaya M, El'skaya A*. COTRASIF: conservation-aided transcription-factor-binding site finder. *Nucleic Acids Res*. 2009;**37**(7):e49.
17. *Drahushchenko OO, Tokovenko BT, Obolens'ka MIu*. [Initial analysis of the results of the genome-wide search for the genes responsive to interferon alpha]. *Ukr Biokhim Zh*. 2010; **82**(1):82–9.
18. *Matsushita M, Hijikata M, Matsushita M, Ohta Y, Mishiro S*. Association of mannose-binding lectin gene haplotype LXPA and LYPB with interferon-resistant hepatitis C virus infection in Japanese patients. *J Hepatol*. 1998;**29**(5): 695–700.
19. *Quandt K, Frech K, Karas H, Wingender E, Werner T*. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* 1995;**23**(23):4878–84.
20. *Cartharius K, Frech K, Grote K, et al.* MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics*. 2005;**21**(13):2933–42.
21. *Crispe IN*. The liver as a lymphoid organ. *Annu Rev Immunol*. 2009;**27**:147–63.
22. *Schrem H, Klempnauer J, Borlak J*. Liver-enriched transcription factors in liver function and development. Part I: the hepatocyte nuclear factor network and liver-specific gene expression. *Pharmacol Rev*. 2002;**54**(1):129–58.
23. *Bürglin TR*. Homeodomain subtypes and functional diversity. *Subcell Biochem*. 2011;**52**:95–122.
24. *Kassel O, Herrlich P*. Crosstalk between the glucocorticoid receptor and other transcription factors: molecular aspects. *Mol Cell Endocrinol*. 2007;**275**(1–2):13–29.
25. *Garrett-Sinha LA*. Review of Ets1 structure, function, and roles in immunity. *Cell Mol Life Sci*. 2013;**70**(18):3375–90.
26. *Yang TT, Chow CW*. Transcription cooperation by NFAT. C/EBP composite enhancer complex. *J Biol Chem*. 2003; **278** (18):15874–85.
27. *Marecki S, Riendeau CJ, Liang MD, Fenton MJ*. PU.1 and multiple IFN regulatory factor proteins synergize to mediate transcriptional activation of the human IL-1 beta gene. *J Immunol*. 2001;**166**(11):6829–38.
28. *Shaulian E, Karin M*. AP-1 in cell proliferation and survival. *Oncogene*. 2001;**20**(19):2390–400.
29. *Schrem H, Klempnauer J, Borlak J*. Liver-enriched transcription factors in liver function and development. Part II: the C/EBPs and D site-binding protein in cell cycle control, carcinogenesis, circadian gene regulation, liver regeneration, apoptosis, and liver-specific gene regulation. *Pharmacol Rev*. 2004;**56**(2):291–330.
30. *Le Lay J, Kaestner KH.* The Fox genes in the liver: from organogenesis to functional integration. *Physiol Rev*. 2010; **90**(1):1–22.
31. *Macian F*. NFAT proteins: key regulators of T-cell development and function. *Nat Rev Immunol*. 2005;**5**(6): 472–84.
32. *Vaughan C, Pearsall I, Yeudall A, Deb SP, Deb S*. p53: its mutations and their impact on transcription. *Subcell Biochem*. 2014;**85**:71–90.
33. *Akerfelt M, Morimoto RI, Sistonen L*. Heat shock factors: integrators of cell stress, development and lifespan. *Nat Rev Mol Cell Biol*. 2010;**11**(8):545–55.
34. *Shaywitz AJ, Greenberg ME*. CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals. *Annu Rev Biochem*. 1999;**68**:821–61.
35. *Gilmore TD.* Introduction to NF-kappaB: players, pathways, perspectives. *Oncogene*. 2006;**25**(51):6680–4.
36. *Hayden MS, West AP, Ghosh S*. NF-kappaB and the immune response. *Oncogene*. 2006;**25**(51):6758–80.
37. *Basehoar AD, Zanton SJ, Pugh BF*. Identification and distinct regulation of yeast TATA box-containing genes. *Cell*. 2004;**116**(5):699–709.
38. *Zou Y, Huang W, Gu Z, Gu X*. Predominant gain of promoter TATA box after gene duplication associated with stress responses. *Mol Biol Evol*. 2011;**28**(10):2893–904.
39. *Han HW, Bae SH, Jung YH, Kim JH, Moon J*. Genome-wide characterization of the relationship between essential and TATA-containing genes. *FEBS Lett*. 2013;**587**(5): 444–51.

40. *Guo N, Mogues T, Weremowicz S, Morton CC, Sastry KN*. The human ortholog of rhesus mannose-binding protein-A gene is an expressed pseudogene that localizes to chromosome 10. *Mamm Genome*. 1998;**9**(3):246–9.
41. *Lynch M, Conery JS*. The evolutionary fate and consequences of duplicate genes. *Science*. 2000;**290**(5494): 1151–5.
42. *Platanias LC*. Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nat Rev Immunol*. 2005;5(5): 375–86.
43. *Rutkowski MJ, Sughrue ME, Kane AJ, Ahn BJ, Fang S, Parsa AT*. The complement cascade as a mediator of tissue growth and regeneration. *Inflamm Res.* 2010;**59**(11): 897–905.
44. *Rutkowski MJ, Sughrue ME, Kane AJ, Mills SA, Fang S, Parsa AT*. Complement and the central nervous system: emerging roles in development, protection and regeneration. *Immunol Cell Biol*. 2010;**88**(8):781–6.
45. *Degnan BM, Degnan SM, Naganuma T, Morse DE*. The ets multigene family is conserved throughout the Metazoa. *Nucleic Acids Res*. 1993;**21**(15):3479–84.
46. *Degnan BM, Morse DE*. Identification of eight homeobox-containing transcripts expressed during larval development and at metamorphosis in the gastropod mollusc Haliotis rufescens. *Mol Mar Biol Biotechnol*. 1993;**2**(1): 1–9.
47. *Baker ME, Funder JW, Kattoula SR*. Evolution of hormone selectivity in glucocorticoid and mineralocorticoid receptors. *J Steroid Biochem Mol Biol*. 2013;**137**:57–70.
48. *Amoutzias GD, Veron AS, Weiner J 3rd, et al*. One billion years of bZIP transcription factor evolution: conservation and change in dimerization and DNA-binding site specificity. *Mol Biol Evol*. 2007;**24**(3):827–35.
49. *Rutkowski R, Hofmann K, Gartner A*. Phylogeny and function of the invertebrate p53 superfamily. *Cold Spring Harb Perspect Biol*. 2010;**2**(7):a001131.
50. *Shimeld SM, Degnan B, Luke GN*. Evolutionary genomics of the Fox genes: origin of gene families and the ancestry of gene clusters. *Genomics*. 2010;**95**(5):256–60.
51. *Song X, Hu J, Jin P, Chen L, Ma F.* Identification and evolution of an NFAT gene involving Branchiostoma belcheri innate immunity. *Genomics*. 2013;**102**(4):355–62.
52. *Langevin C, Aleksejeva E, Passoni G, Palha N, Levraud JP, Boudinot P*. The antiviral innate immune response in fish: evolution and conservation of the IFN system. *J Mol Biol*. 2013;**425**(24):4904–20.
53. *Xu L, Yang L, Liu W*. Distinct evolution process among type I interferon in mammals. *Protein Cell*. 2013;**4**(5):383–92.
54. *Marchalonis JJ, Whitfield GK, Schluter SF*. Rapid evolutionary emergence of the combinatorial recognition repertoire. *Integr Comp Biol*. 2003;**43**(2):347–59.

В. С. Бондаренко, М. Ю. Оболенська

Біоінформатичний аналіз цис-регуляторних елементів в генах *Mbl1* і *Mbl2 Rattus norvegicus*

Мета. Провести біоінформатичне дослідження промоторів генів *Mbl1* і *Mbl2 Rattus norvegicus*, що кодують манозозв'зуючий лектини, на наявність та функціональну специфічність сайтів зв'язування транскрипційних факторів. ***Методи.*** Обидві послідовності промоторів довжиною 1100 п.н. проаналізовано в програмі MatInspector за використання позиційно-вагових матриць з бази данних Matrix Family Library Version 9.0. В межах програми відібрано сайти зв'язування для транскрипційних факторів, які специфічно експресуються в печінці і клітинах імунної системи, а також, що пройшли тест на консервативність через порівняння з сайтами зв'язування транскрипційних факторів в промоторах генів — ортологів *Mus musculus*. ***Результати.*** В промоторах гена *Mbl1* і гена *Mbl2* наявні сайти для чотирьох родин транскрипційних факторів (HNF, homeodomain transcription factors, GRs and ETS factors). В промоторі гена *Mbl1* виявлено також сайти для представників шістьох (AP1 related factors, Ccaat/Enhancer Binding Proteins, FOX, p53, NFAT, ISGF3), а в гені *Mbl2* — для чотирьох (cAMP-responsive element binding proteins, heat shock factors, Nf-κB/c-rel and TATA binding protein) інших родин транскрипційних факторів. Характерні для гена *Mbl1* транскрипційні фактори переважно задіяні в регуляції диференціювання, розвитку, підтримання гомеостазу, клітинного циклу. Через характерні для гена *Mbl2* сайти транскрипційні фактори можуть опосередковувати переважно відповідь на чинники, які індукують стрес. ***Висновок.*** Різноманітність транскрипційних факторів, які потенційно можуть регулювати транскрипцію обох генів, вказує на можливу участь генів в різних клітинних процесах зі специфічною функцією кожного з генів. Отримані результати є підставою для цілеспрямованого експериментального дослідження регуляції транскрипції зазначених генів.

Ключові слова: манозо-зв'язувальні лектини, позиційно-вагові матриці, сайти зв'язування транскрипційних факторів.

В. С. Бондаренко, М. Ю. Оболенская

Биоинформатический анализ цис-регуляторных элементов в генах *Mbl1* и *Mbl2 Rattus norvegicus*

Цель. Провести биоинформатический анализ промоторов генов *Mbl1* и *Mbl2 Rattus norvegicus*, которые кодируют маннозосвязывающие лектины, на наличие и функциональную специфичность сайтов связывания транскрипционных факторов. ***Методы.*** Обе последовательности промоторов длиной 1100

п.н. проанализированы с помощью программы MatInspector с использованием позицонно-весовых матриц из базы данных Matrix Family Library Version 9.0. С помощью программы отобраны сайты связывания для транскрипционных факторов, которые специфически экспрессируются в печени и клетках иммунной системы, а также которые прошли тест на консервативность после сравнения с сайтами связывания в промоторах генов-ортологов *Mus musculus*. ***Результаты.*** В промоторах каждого из двух генов найдены сайты для представителей четырех семейств транскрипционных факторов (HNF, homeodomain transcription factors, GRs and ETS factors). Кроме того в гене *Mbl1* выявлены сайты для представителей шести других семейств транскрипционных факторов (AP1 related factors, Ccaat/Enhancer Binding Proteins, FOX, p53, NFAT, ISGF3), а в гене *Mbl2* — для четырех семейств (cAMP-responsive element binding proteins, heat shock factors, Nf-κB/c-rel and TATA binding protein). Транскрипционные факторы, характерные для гена *Mbl1,* преимущественно участвуют в регуляции процессов дифференцировки, развития, клеточного цикла и поддержании гомеостаза. Транскрипционные факторы, которые могут связываться с промотором гена *Mbl2,* опосредуют ответ преимущественно на действие факторов, индуцирующих стресс. ***Выводы.*** Разнообразие транскрипционных факторов, которые потенциально могут регулировать транскрипцию обоих генов, указывает на возможное участие генов в различных процессах со специфической функцией каждого из генов. Полученные результаты являются базисом для целенаправленного экспериментального исследования регуляции транскрипции обоих генов.

Ключевые слова: маннозосвязывающие лектины, позиционно-весовые матрицы, сайты связывания транскрипционных факторов.

Received 26.11.2014