The Journal of Animal & Plant Sciences, 27(6): 2017, Page: 1806-1815
ISSN: 1018-7081

ISOLATION, SEQUENCING AND *IN SILICO* CHARACTERIZATION OF PROMOTER REGION OF *HSPA6* GENE FROM ARABIAN CAMEL (*CAMELUS DROMEDARIES*)

M. Elrobh[1]*, M. Alanazi [1], Z. Abduljaleel[2,3], W. El-Huneidi[4] and W. Khan[4]*

[1]Genomic Research Chair Unit, Department of Biochemistry, College of Science, King Saud University, PO Box 2455, Riyadh 11451, Saudi Arabia; [2]Department of Medical Genetics, Faculty of Medicine, Umm Al-Qura University, P.O. Box 715, Makkah 21955, Saudi Arabia; [3]Science and Technology Unit, Umm Al Qura University, P.O. Box 715, Makkah 21955, Saudi Arabia; [4]Department of Basic Sciences, College of Science and Health Professions, King Saud Bin Abdulaziz University for Health Sciences, PO Box 3660, Riyadh 11426, Saudi Arabia

*Corresponding author's email: melrobh@ksu.edu.sa, khanmuh@ksau-hs.edu.sa

ABSTRACT

Heat-shock proteins, divided in numerous families, play an important role in cellular processes. The human Hsp70 (PDB ID 1HJO) performs multiple steps including protein folding, trafficking, remodeling and degradation. In addition, the Hsp70 is also involved in cell signaling, and tumor repression. We have previously sequenced the full-length cDNA encoding the putative stress induced heat shock protein HSPA6 from Arabian camel (NCBI Gene Bank accession number HQ214118.1). The sequence analysis of *HSPA6* gene revealed a 1932bp-long open reading frame encoding 643 amino acids. In this work, we extended our study to explore the mechanism that manages the *HSPA6* gene expression by isolating and characterizing the flanking 5' region using online tools, expected to harbor the regulatory elements. The straight-walk method was used to isolate the upstream genomic DNA sequences of the *HSPA6*. The isolated fragment has a length of 1417bp and after verifying the sequence, it was deposited in GenBank under accession number JX888464. The sequence analysis revealed the presence of several heat shock elements (HSEs) that act as binding sites for the major heat shock factor. These HSE sites were dispersed in the isolated region and had consensus sequences identical to those found in eukaryotes. Other core promoter elements such as potential transcriptional start site for the *HSPA6* gene were found including the canonical TATA-box, DPE, MTE, BRE, CpG island, palindromic and repeated sequences. Taken together, this is the first report on the promoter region of the *HSPA6* gene from Arabian camel which will enhance our understanding of how such regulatory *cis*-elements are engaged in regulating the gene expression.

Keywords: genome walk, heat shock elements, *HSPA6* gene, *Camelus dromedaries*, *in silico*.

INTRODUCTION

The Arabian camel (*Camelus dromedaries*) is one of those rare animals that can survive long periods without water under severe desert environments such as high daytime temperatures and severe cold night temperatures. However, the molecular basis that underlines characteristics is still not very clear. In an attempt to start investigating components that could play a role in contributing to such characteristics, we focused on one of the cellular chaperone components; the heat shock protein A6 (*HSPA6*, also called *HSP70B'*) gene from Arabian camel. The *HSPA6* gene is peculiar in the sense that it is not expressed continuously like other heat shock genes at basal level and also does not exist in all mammals (Noonan, Place, Giardina, and Hightower, 2007). We have previously shown and isolated the cDNA of this gene from *Camelus dromedaries* and identified the gene sequences (NCBI accession number HQ214118) (Elrobh *et al*., 2011). The absence of basal transcriptional level of this gene in humans was suggested to be due to the absence of the main core promoter elements; CAAT and TATA boxes (Leung, Rajendran, Monfries, Hall, and Lim, 1990). The deletion mutation studies of *HSP70* gene unveiled the significance of certain sequences in the upstream promoter region such as CTGGAATATTCCCG which also matched to 12 out of the 14 sequences that were found in Drosophila heat shock genes (Wu, Kingston, and Morimoto, 1986). Another specific feature of human *HSP70* promoter region is the presence of a metal-response element sequence (CGNCCCGG) located at -107 bases upstream of the transcription initiation point that is responsible for cadmium associated gene induction (Wada, Taniguchi, and Okano, 2007). Hence, the cadmium cytotoxicity can be detected at the cellular level in a novel biosensor vector based system. Interestingly, Wada *et al* (Wada *et al*., 2007) found a response sequence that is located upstream of the human *HSPA6* gene and cloned it in tandem order in a reporter vector to use it as a sensitive tool to detect cytotoxicity. The promoter region of human *HSPA6* gene represents an interesting area in the field of biotechnology. Rohmer and his colleagues have used this promoter region as a tool to design effective adenovirus transfer vectors for therapeutic applications (Rohmer, Mainka, Knippertz, Hesse, and Nettelbeck, 2008). Hence,

identifying promoter sequences of camel *HSPA6* will allow us to better understand the gene regulation and may provide information for developing additional thermo-sensing vectors. This study deals with the isolation and characterization of the flanking 5' region, expected to harbor the regulatory elements. Furthermore, the study demonstrated the occurrence of several heat shock elements (HSEs) as binding sites for the major heat shock factor including other core promoter elements like transcriptional start site for the *HSPA6*, canonical TATA-box, DPE, MTE, BRE, CpG island, palindromic and repeated sequences. Overall, this is the first published work on the promoter region of the Arabian camel *HSPA6* gene and other core promoter elements which will enrich our understanding of regulatory *cis*-elements that are involved in modulating the gene expression.

MATERIALS AND METHODS

Isolation of genomic DNA: Camel liver tissues were collected from the National Slaughter House in Riyadh, Saudi Arabia. The samples were transferred in portable liquid nitrogen to the lab where they were kept at -80ºC. Genomic DNA was extracted using Qiagen DNeasy Blood and Tissue Kit according to the manufacturer's protocol to get 20ng/µl.

Cloning and sequencing of the 5' region: The Straight Walk Kit (Bex, Japan) was used to clone the 5' region of *HSPA6* (Tsuchiya, Kameya, and Nakamura, 2009). Briefly, the extracted genomic DNA was digested with six different restriction enzymes to produce separate libraries of fragments. The restriction enzymes used were *Avr*II, *Bam*HI, *Bcl*I, *Nhe*I, *Spe*I and *Xba*I. Then, each library was ligated to specific adaptor sequences provided by the Kit. Next, two rounds of PCR reactions were set up for each library to isolate the 5' region. In the primary PCR reaction, the primers used were the Kit specific forward primer WP-1 (that will anneal with the adaptor sequence); and the *HSPA6* reverse primer; SP1, (5' CTAGGATCTCTACCCGGCCGTGCTG 3') that corresponds to nucleotide coordinates 206-230 in the *HSPA6* gene (Accession number HQ214118). The total reaction volume of the primary PCR was 50µl containing: 10 pmol of each primer, 2 ng of genomic camel DNA, 5 U/µl of Kit *Taq*-Plus, 5 µl of 10×PCR buffer, 2 mM dNTPs and 25 mM $MgSO_4$ under the following conditions: first denaturation at 94ºC for 2 min followed by 35 cycles at 94ºC for 30 sec, 65ºC for 30 sec and 68ºC for 5 min. In the second nested PCR, the product of the first PCR was diluted 100 times and 1 µl was used as a template with 10 pmol of the nested primers (forward kit primer is WP-2 which anneal to the adaptor sequences and the reverse *HSPA6* primer is SP2 (5' GCCTATGGCCACTTCCTTTGCGG 3') which anneal to nucleotide coordinates of 144-166 of *HSPA6* gene under the same conditions as stated above for the primary PCR reaction but for 30 cycles. Results of PCR amplification were run on 1.5% agarose gel. The purified fragment was then cloned in TOPO TA cloning vector (Invitrogen) and transformed into DH5α cells using standard molecular biology techniques. Sanger sequencing was performed using vector specific primers.

Isolation of camel HSPA6 5' region: The PCR-adaptor genome walk method was used to obtain the upstream sequences of camel *HSPA6* gene. Camel genomic DNA was digested with different restriction enzymes (*Bam*HI, *Bcl*I, *Spe*I and *Nhe*I, *Xba*I and *Avr*II) to produce six patches of fragments. The products were then ligated to known flanking sequences. In order to isolate the specific 5' region, two rounds of PCR reactions were performed using primer pair specific to the flanking sequences and *HSPA6* gene in each round. In the primary PCR (data not shown), the libraries of restricted fragments were used as templates to isolate large fragments. In the second nested PCR (Fig. 1), the product of the primary PCR was diluted and used as a template. In the absence of both primers, all reactions give a smear band indicating that camel genomic is digestible by the different enzymes used. Additionally, it seems that the 5' upstream region could be isolated by the primary PCR step alone because a band of size 1.5kb (lane 1) is visible when using amplification from *Bam*HI restriction. However, the band fails to be reproduced as shown in lane 2 where it fainted. In another set of control reactions, we used only the Kit specific primer WP-2. The results are shown in lanes 4, 8, 12, 16 and 20. A vague band is visible when using the control primer WP-2 alone in case of *Bam*HI, *Bcl*I and *Spe*I (lanes 4, 8 and 12, respectively) while the remaining three enzyme reactions did not show this band. The results of PCR amplification using two nested primers are shown in lanes 2, 6, 10, 14, 18 and 22. Only a strongly amplified band is obtained in case of the restriction enzymes *Bcl*I, *Spe*I and *Avr*II. However, multiple bands were also observed when using *Bcl*I and *Spe*I restriction amplification (lanes 6 and 10). The restriction amplification of *Avr*II showed a clear single amplified band of size 1417 bp, hence, this fragment was cloned and sequenced. The sequence of this fragment is deposited in GenBank under accession number JX888464.

Sequence Analysis: DNA complexity graph and scanning for HSE sequence motifs were done using "CLC Genomics Workbench" (v8.5.1 for Mac). Transcription start site was predicted using "Neural Network Promoter Prediction" (available online at: http://www.fruitfly.org/seq_tools/promoter.html)(Reese, 2001). Core promoter sequences (BRE, DPE, TATA box and MTE) were deduced using "Position Weight Matrices" (available online at: http://www.bioinformatics.org/yapp/cgi-

bin/yapp_intro.cgi). The analysis of CpG island was performed using "EMBOSS Newcpgreport" (available online at: http://www.ebi.ac.uk/Tools/seqstats/emboss_newcpgreport/)(Rice, Longden, and Bleasby, 2000). Repeated sequences were scanned using "MREP" (available online at: http://mreps.univ-mlv.fr/)(Kolpakov, Bana, and Kucherov, 2003) with resolution set to 0. Palindromic sequence was generated using "Palindromic sequences finder" (available online at: http://www.biophp.org/minitools/find_palindromes/demo.php). Results of analysis from online tools were used in "CLC Genomics Workbench" to generate Fig. 2. Dot plot graphs were done using Dotlet(available online at: http://myhits.isb-sib.ch/cgi-bin/dotlet)(Junier and Pagni, 2000). Identification of putative transcription factor binding sites was done using "GPMiner" (available online at http://gpminer.mbc.nctu.edu.tw/index.php)(Lee, Chang, Hsu, Chang, and Shien, 2012). Consensus HSE was generated using Weblogo (available online at: http://weblogo.berkeley.edu/logo.cgi)(Crooks, Hon, Chandonia, and Brenner, 2004).

RESULTS AND DISCUSSION

Characterization of the 5' flanking region: We previously published the full length of the camel heat shock protein 6 (HSPA6) mRNA (Genbank accession number HQ214118)(Elrobh *et al*., 2011). In order to have a complete picture about the sequences of the *HSPA6* gene and its predicted promoter region, we fused the sequence isolated in this study with the first 136bp of HSPA6 mRNA to generate a total of 1553bp. We selected the first 136bp because it covers the starting methionine in HSPA6.

The core promoter elements predicted by position weight matrix are shown in Fig. 2. The sequence numbered starting from the 5' end indicating the presence of TATA-box having TGGATAAAAAGC sequence was predicted to be at position 1389-1400 bp. The motif ten element (MTE) containing GAAGCGGAGCGA sequence was located at position 1406-1417bp. The downstream promoter element (DPE) was shown to be situated with the sequence AGACG at 1201-1205bp. The sole B recognition element (BRE) was deduced to be located at 1505-1511bp showing CGACGCC sequence. The initiator (INR) sequence CCAGTCC, was found to be predicted at position 1482-1488bp. Other features of the sequence were also deduced using the appropriate tools mentioned in the methodology section such as a 287 bp long CpG island located at 1151-1437 bp with a 64% CG content. The longest palindrome sequence (CAGATCTG) was identified at 1419-1426bp site. An eleven base repeat of A was present at 974-984bp. The transcription start site (TSS) corresponding to the fist adenine in the sequence (HQ214118) was labeled TSS1 (position 1418). The deduced TSS2 was generated using Neural network promoter prediction and was located at position 1421bp. The complexity plot (Fig. 3) of the total 15534bp shows a region of low complexity around position 980bp from the 5' end. The transcription factor (TF) binding sites analysis revealed that several TFs were present in both strands (Fig. 4). It was observed that the nuclear factor of activated T-cells (NFAT) was predicted to be more frequent than other TFs in the top strand. The dot plot analysis revealed that various regions of the sequence contain different motif sequences that include local complementarity to itself to make a stem-and-loop structure around the sequence position #89 (Fig. 5A), #390 (Fig. 5B) and #1303 (Fig. 5E); whereas repeated regions were observed at position #223 (Fig. 5A) along with a low complexity region at position #983. The scanning for the putative heat shock elements (HSEs) using the consensus HSE sequence (nTTCnnGAAnnTTCn) revealed 18 sites at different locations (Fig. 2 and Table 1) showing the highest match to consensus at 1245-1259bp position with sequence cTACcgGAAccTTCt. The Weblogo server was used to generate a consensus HSE from the 18 predicted sites and the results exhibited that camel HSPA6 gene promoter was identical to the canonical HSE elements and the consensus sequence deduced was nTTCnnGAAnnTTCn (Fig. 6).

The Arabian camel is a unique mammal that can survive stressful environments that are normally detrimental to most organisms including humans. However, there are no adequate literatures available to understand the molecular mechanisms underlying this physiology. It is anticipated that the camel may have an efficient mechanism that can tolerate stressful conditions (such as the desert's high temperature which can reach up to 55^{o}C) and preserve its capability of usual protein synthesis machinery from being affected (Al Ghumlas, Abdel Gader, Hussein, Al Haidary, and White, 2008; Tefera, 2004; Ulmasov, Karaev, Lyashko, and Evgen'ev, 1993). The work presented here is in continuation of our previous work where one of the members of heat shock proteins HSPA6 mRNA was isolated (Genbank accession number HQ214118) and characterized from Arabian camel (Elrobh *et al*., 2011). This gene in particular is not present in all vertebrates such as mice (Parsian *et al*., 2000; Ramirez, Stamatis, Shmukler, and Aneskievich, 2015), while in goats and humans, its expression fluctuates (Banerjee *et al*., 2014; Ramirez *et al*., 2015). Its presence in camel evokes the question about how it is induced, and investigating the promoter region of *HSPA6* from the Arabian camel could be of significant biotechnological value. In addition most of the housekeeping genes have well defined structural regulatory elements (Maston, Evans, and Green, 2006). Their relative orientation to each other and the distance

between them is essential for the activity of that promoter (Weingarten-Gabbayand Segal, 2014). In this study, we isolated the 5' flanking region of camel *HSPA6* gene, where the isolated region showed a composition of typical eukaryotic promoters but in different orientation relative to each other. The core promoter element, the TATA-box (Juven-Gershon and Kadonaga, 2010) which binds RNA polymerase II is predicted to be located at the upstream of the transcription start site #1 (TSS1) with the sequence GATAAA. The distance between the second adenine in TATA-box and TSS1 was 26 nucleotides. The previously reported motif ten element (MTE) site (Lim *et al*., 2004) was located downstream of the initiator site (INR) (Kugel and Goodrich, 2017) in TATA-box-INR-MTE order, however in the camel *HSPA6* promoter, it was located upstream of the INR site where the TATA-box was located 4bp upstream of the MTE. The MTE has been reported to confer basal transcriptional activity and compensate for TATA-less promoters (Kadonaga, 2012). The presence of both elements (TATA-box and MTE) indicates that the HSPA6 expression is regulated constitutively. The transcription factor IID recognized the INR sequence as well as another core promoter region called the downstream promoter element (DPE). The presence of DPE sequence upstream of the TATA-box rather than the expected canonical downstream location in the camel *HSPA6*, suggests that it could be considered an enhancer element or that the DNA is looped to facilitate the contribution of DPE sequence to the transcriptional machinery. The close proximity of the CpG Island to the TSS also suggests that it could accommodate another enhancer sequence within it. The presence of palindromic sequence at the start of the TSS shows that it may act as *cis*-element for a transcription factor to regulate the *HSPA6* expression (Krawczyk, Thurow, Niggeweg, and Gatz, 2002). The binding site for ribosomes which is responsible for initiating translation (the initiator site, INR) was located 71 bases from the starter methionine as well as upstream of the BRE sequence. The transcription factor IIB has been shown to bind to specific 7 nucleotides located immediately upstream of the TATA-box (Lagrange, Kapanidis, Tang, Reinberg, and Ebright, 1998) called BRE. However, this BRE site in the camel was located in a peculiar place where it was further downstream of the TATA-box between the INR site and the starting methionine, indicating a possible regulatory role (Roy and Singer, 2015; Yang *et al*., 2011). Regarding the heat shock element (HSE) which is important in responding to stress responses like high temperature (Akerfelt, Morimoto, and Sistonen, 2010; Morano, Grant, and Moye-Rowley, 2012), the isolated 1417bp fragment showed 17 HSE sites for binding the transcriptional regulator heat shock factor (HSF), with a single HSE site overlapping the INR site in the mRNA region. Among these 17 HSE sites, the highest consensus-matching HSE sequence (93% identical) is located at the 1245-1259bp position within the CpG Island and was probably the core site for HSF binding and regulation (Enokiand Sakurai, 2011). A string of adenine repeats around position 980 from the 5' end was confirmed using both the complexity graph and Dotlet program. These adenine repeats may be used as an additional signal to the heat shock response via antisense pathway (Pandey, Mandal, Jha, and Mukerji, 2011) because of the low DNA melting temperature in this region. Overall, in this study a total of six core promoter elements TATA-box, BRE, MTE, DPE, INR and HSE were deduced from the 5' upstream sequence of camel *HSPA6* gene that may be responsible for controlling the expression. An extra palindromic site was found to overlap the transcription start site, suggesting another factor that might play a role in this region. Hence, the expression of *HPSA6* is under tight regulatory control elements and further *in vitro* studies are needed to evaluate the role of each element.

Table 1. List of heat shock elements deduced, their position and matching percent accuracy to the consensus sequence (in Red).

#	Position	Sequence	% Accuracy
		n**TTC**nn**GAA**nn**TTC**n	
1	90..104	g**TT**Aca**GAA**gaC**T**Tt	80
2	113..127	t**TTC**ttCCA**g**t**TC**Ca	80
3	135..149	a**T**CTct**GAA**aaG**TC**t	80
4	188..202	aCA**C**ca**GAA**aaC**TC**a	80
5	339..353	tG**T**Aag**GAA**tt**TAC**t	80
6	495..509	tG**TC**acCA**C**ag**TTC**a	80
7	505..519	g**TTC**atCTGca**TTC**t	80
8	569..583	cG**TC**tcC**AA**gcC**TC**a	80
9	819..833	cC**T**Aca**GAA**aa**TGC**c	80
10	912..926	t**TTC**gaAGA**t**a**TT**Gc	80
11	934..948	t**TTC**aa**GA**CccC**TC**c	86
12	1075..1089	a**TTC**ttC**AA**ctAA**C**g	80
13	1106..1120	tC**TT**gg**GAA**gt**TAC**t	80
14	1171..1185	aAG**C**tg**GAA**at**TTC**t	86
15	1245..1259	c**TAC**cg**GAA**cc**TTC**t	93
16	1337..1351	gCG**C**tg**GAA**gg**TTC**g	86
17	1378..1392	tC**TC**cg**GAA**ac**T**GGa	80
18	1485..1499	g**TCC**ct**GA**Ggc**TC**Ct	80

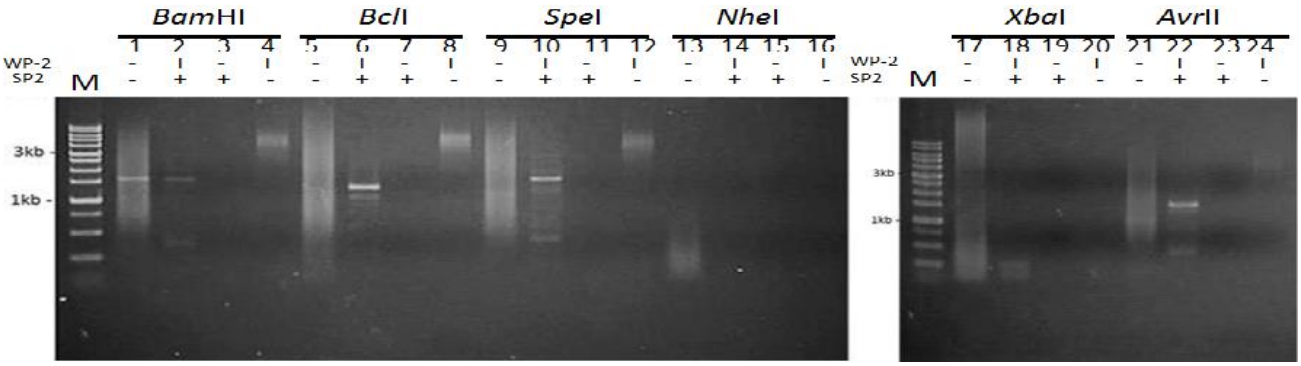

Fig 1. Gel electrophoresis of the second nested PCR products was carried out after Straight Walk experiment. The restriction enzymes used are shown and lane M is the 1kb DNA ladder marker (Gene Craft) with locations of 1kb and 3kb denoted. In each case of the restriction enzyme used, the PCR reaction was tested either in the presence (+) or absence (-) of individual primers WP-2 or SP2 as a control experiment.

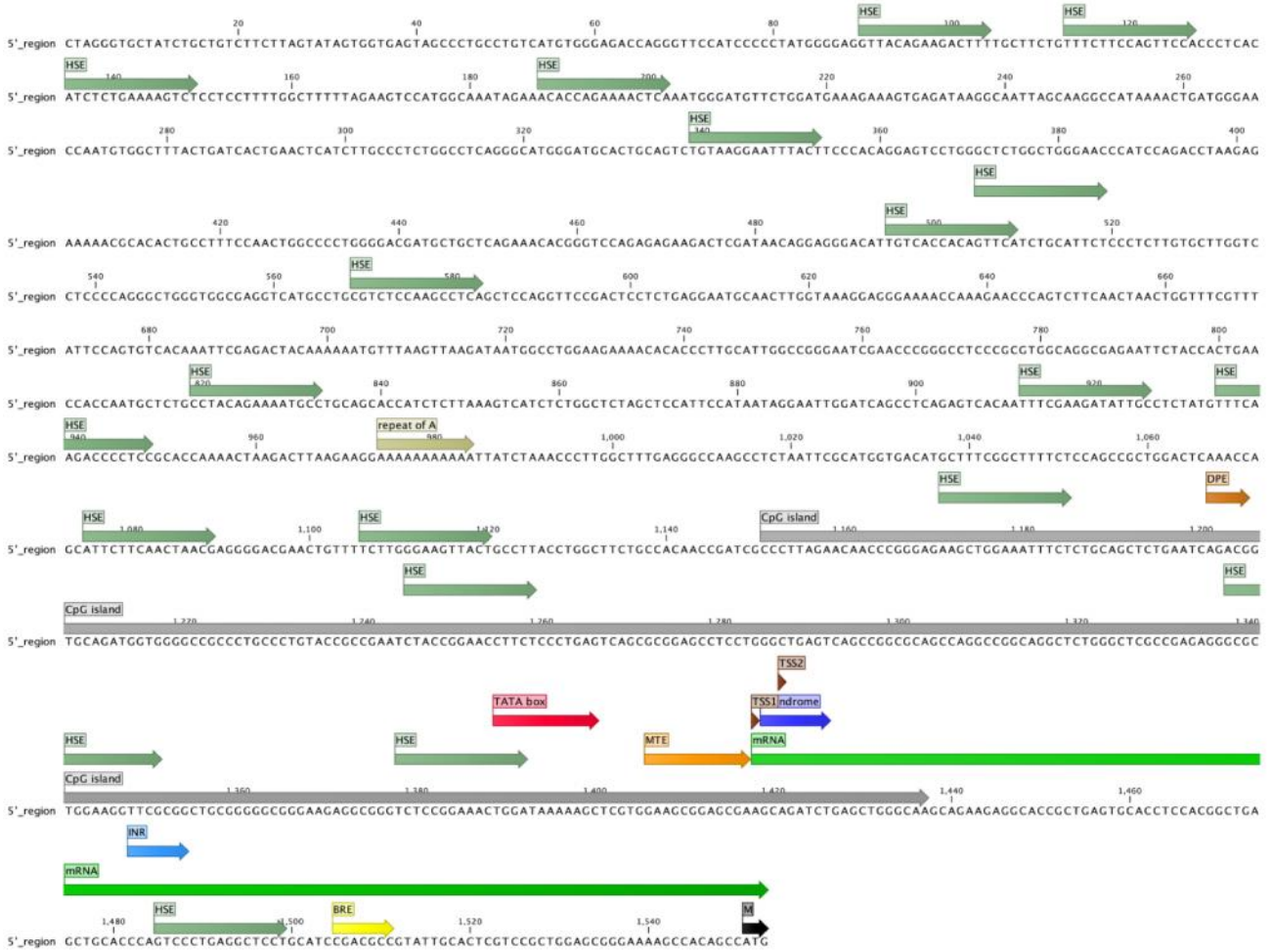

Fig 2. Sequence analysis of the 5' upstream region of *HSPA6* gene region fused with the first 136bp of mRNA to generate 1553bp fragment. All motif and predicted regulatory elements are shown above the sequence in a little flag. The anticipated heat shock element (HSE) is shown in green, the TATA-box in red, the motif ten element (MTE) in orange, the IIB recognition element (BRE) in yellow, the downstream promoter element (DPE) in brown, the CpG island in gray, the palindromic sequence in dark blue, the repeat of A in pale gray, the transcription start site (TSS) is also shown in dark brown, the 136bp from the previously published mRNA for HSPA6 is shown in pale green and the starting methionine in black.

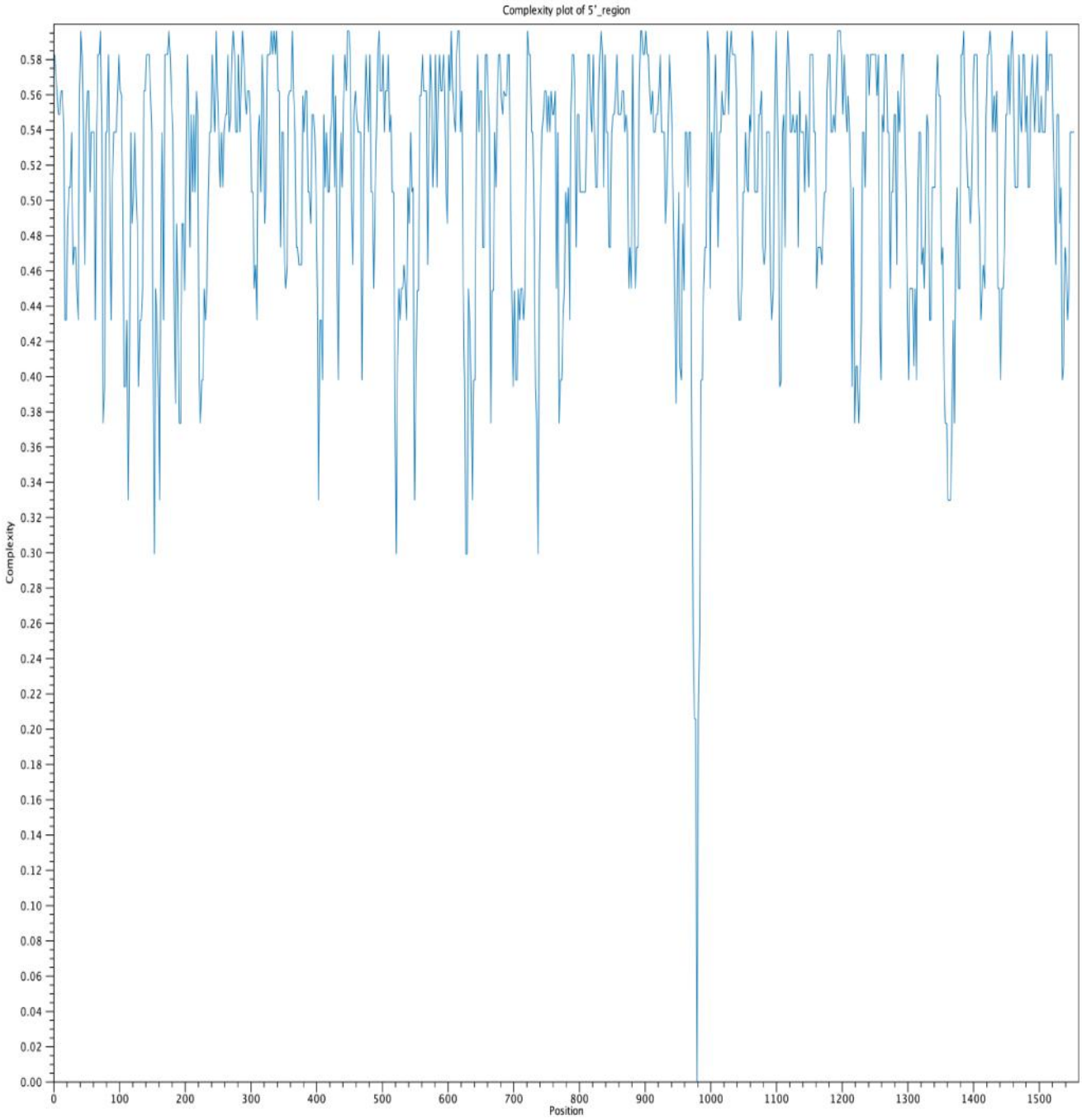

Fig 3. Complexity plot of the 1553bp fragment. The position of the nucleotides are on the x-axis while the relative complexity is shown on the y-axis.

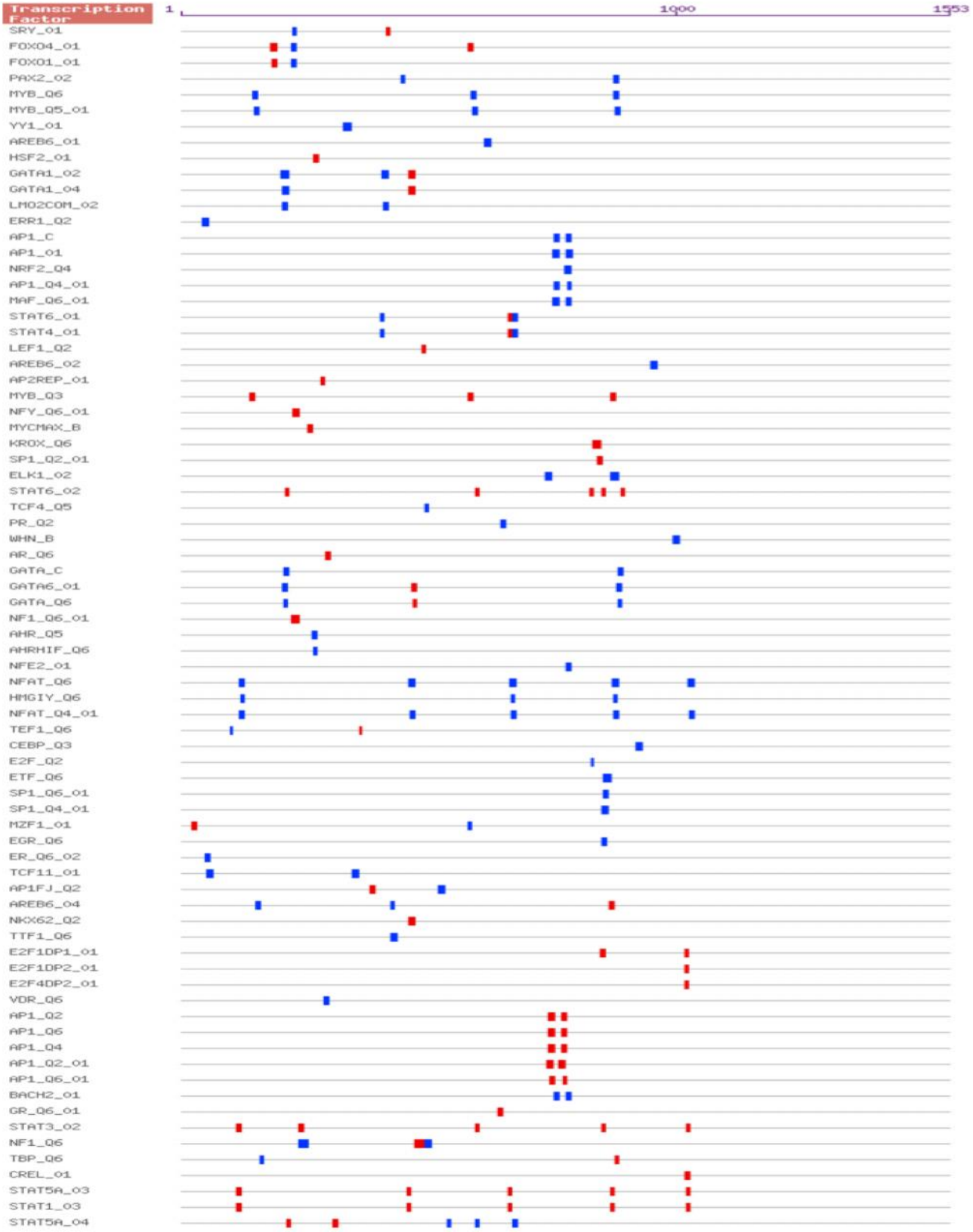

Fig 4. A schematic representation of the various predicted transcription factors (TFs) on the 5' region. The blue bars indicate the location of these TF on the top strand, while the red barsshowlocations at the bottom strand.

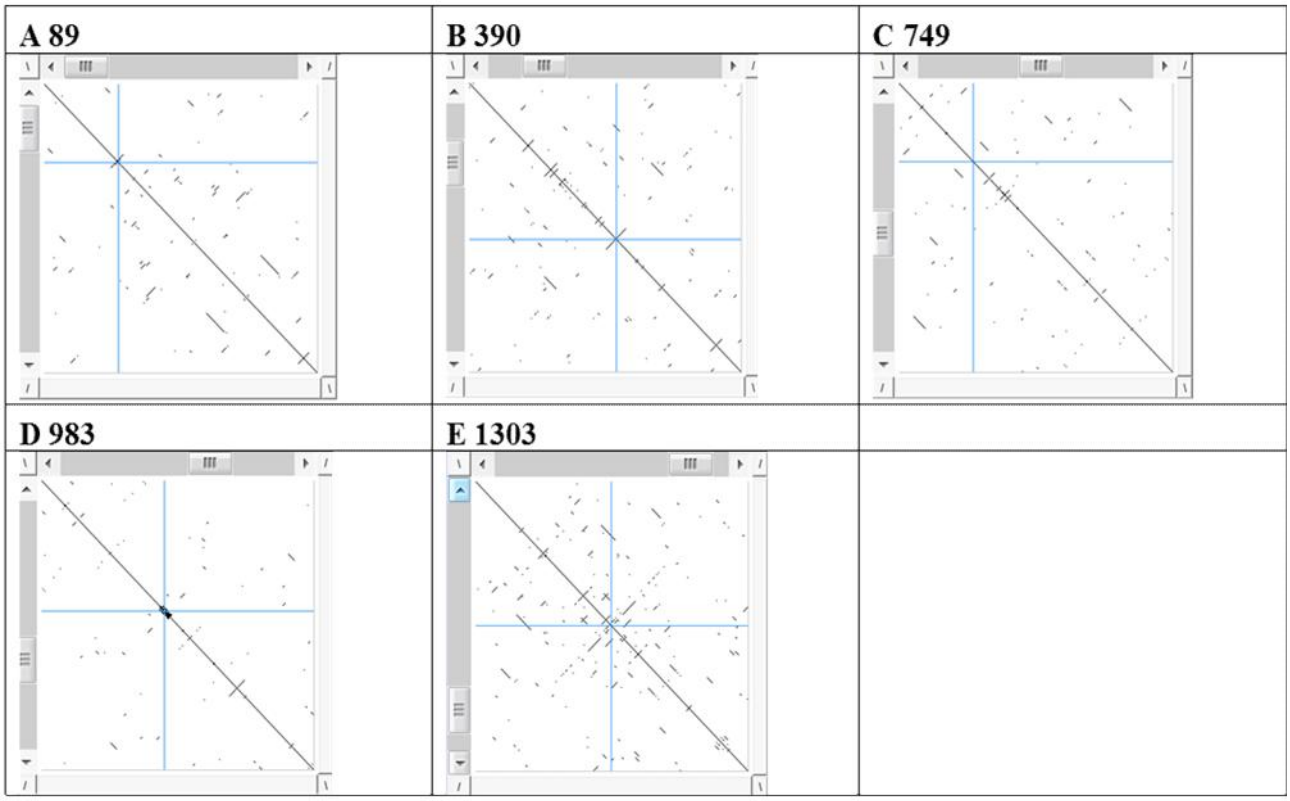

Fig 5. Dot plot analysis of the deduced promoter region. Different windows show various sequence locations on the 1553bp fragment. The windows were generated using default settings, however, the background was adjusted to make lines black on a white background.

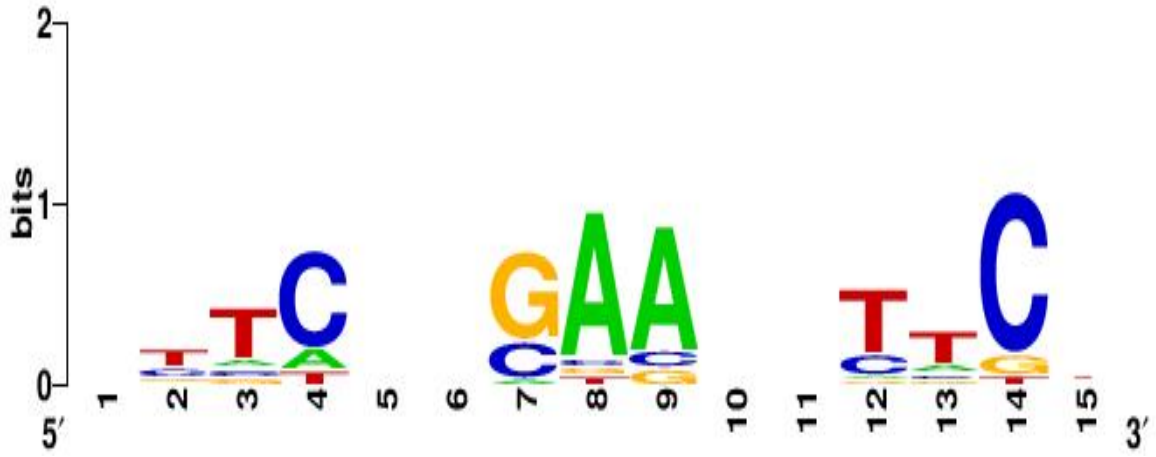

Fig 6. Weblogo analysis of the HSE sequence sites extracted from the full length 1553bp fragment.

Conclusions: The promoter region of any gene contains the necessary nucleotide complements that regulate the activity of that gene whether on or off according to the environmental stimuli. It also harbors the information that identifies the regulation specificity of that particular gene. Hence, studying and identifying the nucleotide composition of the promoter region of *HSPA6* gene is an integral part of knowing the gene itself. In this work, we present for the first time the isolation and characterization of *HSPA6* gene promoter region from the Arabian camel. The 5' upstream region sequences of the *HSPA6* gene were submitted to NCBI GenBank (accession number JX888464.1). We found that various promoter elements are predicted to be involved in the regulation of HSPA6

including the prototype heat shock transcription factor (HSF).

Acknowledgments: The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding the work through the research group project No: RGP-VPP-200.

REFERENCES

Akerfelt, M., R. I., Morimoto and L. Sistonen, (2010). Heat shock factors: integrators of cell stress, development and lifespan. Nat Rev Mol Cell Biol, 11(8): 545-555.

Al Ghumlas, A. K., A. G. Abdel Gader, M. F. Hussein, A. Al Haidary, and J. G. White (2008). Effects of heat on camel platelet structure and function-a comparative study with humans. Platelets, 19(3), 163-171.

Banerjee, D., R. C. Upadhyay, U. B. Chaudhary, R. Kumar, S. Singh and Ashutosh, (2014). Seasonal variation in expression pattern of genes under HSP70 : Seasonal variation in expression pattern of genes under HSP70 family in heat- and cold-adapted goats (Capra hircus). Cell Stress Chaperones, 19(3), 401-408.

Crooks, G. E., G. Hon, J. M. Chandonia and S. E. Brenner (2004). WebLogo: a sequence logo generator. Genome Res, 14(6), 1188-1190.

Elrobh, M. S., M. S. Alanazi, W. Khan, Z. Abduljaleel, A. Al-Amri and M. D. Bazzi (2011). Molecular cloning and characterization of cDNA encoding a putative stress-induced heat-shock protein from Camelus dromedarius. Int J Mol Sci, 12(7), 4214-4236.

Enoki, Y., and H. Sakurai (2011). Diversity in DNA recognition by heat shock transcription factors (HSFs) from model organisms. FEBS Lett, 585(9), 1293-1298.

Junier, T., and M. Pagni (2000). Dotlet: diagonal plots in a web browser. Bioinformatics, 16(2), 178-179.

Juven-Gershon, T., and J. T. Kadonaga (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev Biol, 339(2), 225-229.

Kadonaga, J. T. (2012). Perspectives on the RNA polymerase II core promoter. Wiley Interdiscip Rev Dev Biol, 1(1), 40-51.

Kolpakov, R., G. Bana and G. Kucherov (2003). mreps: Efficient and flexible detection of tandem repeats in DNA. Nucleic Acids Res, 31(13), 3672-3678.

Krawczyk, S., C. Thurow, R. Niggeweg and C. Gatz (2002). Analysis of the spacing between the two palindromes of activation sequence-1 with respect to binding to different TGA factors and transcriptional activation potential. Nucleic Acids Res, 30(3), 775-781.

Kugel, J. F., and J. A. Goodrich (2017). Finding the start site: redefining the human initiator element. Genes Dev, 31(1), 1-2.

Lagrange, T., A. N. Kapanidis, H. Tang, D. Reinberg, and R. H. Ebright. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes Dev, 12(1), 34-44.

Lee, T. Y., W. C. Chang, J. B. Hsu, T. H. Chang and D. M. Shien. (2012). GPMiner: an integrated system for mining combinatorial cis-regulatory elements in mammalian gene group. BMC Genomics, 13 Suppl 1, S3.

Leung, T. K., M. Y. Rajendran, C. Monfries, C. Hall and L. Lim (1990). The human heat-shock protein family. Expression of a novel heat-inducible HSP70 (HSP70B') and isolation of its cDNA and genomic DNA. Biochem J, 267(1), 125-132.

Lim, C. Y., B. Santoso, T. Boulay, E. Dong, U. Ohler and J. T. Kadonaga (2004). The MTE, a new core promoter element for transcription by RNA polymerase II. Genes Dev, 18(13), 1606-1617.

Maston, G. A., S. K. Evans and M. R. Green (2006). Transcriptional regulatory elements in the human genome. Annu Rev Genomics Hum Genet, 7, 29-59.

Morano, K. A., C. M. Grant and W. S. Moye-Rowley. (2012). The response to heat shock and oxidative stress in Saccharomyces cerevisiae. Genetics, 190(4), 1157-1195.

Noonan, E. J., R. F. Place, C. Giardina, and I. E. Hightower (2007). Hsp70B' regulation and function. Cell Stress Chaperones, 12(4), 393-402.

Pandey, R., A. K. Mandal, V. Jha, and M. Mukerji (2011). Heat shock factor binding in Alu repeats expands its involvement in stress through an antisense mechanism. Genome Biol, 12(11), R117.

Parsian, A. J., J. E. Sheren, T. Y. Tao, P. C. Goswami, R. Malyapa, R. Van Rheeden (2000). The human Hsp70B gene at the HSPA7 locus of chromosome 1 is transcribed but non-functional. Biochim Biophys Acta, 1494(1-2), 201-205.

Ramirez, V. P., M. Stamatis, A. Shmukler and B. J. Aneskievich (2015). Basal and stress-inducible expression of HSPA6 in human keratinocytes is regulated by negative and positive promoter regions. Cell Stress Chaperones, 20(1), 95-107.

Reese, M. G. (2001). Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome. Comput Chem, 26(1), 51-56.

Rice, P., I. Longden, and A. Bleasby (2000). EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet, 16(6), 276-277.

Rohmer, S., A. Mainka, I. Knippertz, A. Hesse and D. M. Nettelbeck (2008). Insulated hsp70B' promoter: stringent heat-inducible activity in replication-deficient, but not replication-competent adenoviruses. J Gene Med, 10(4), 340-354.

Roy, A. L., and D. S. Singer (2015). Core promoters in transcription: old problem, new insights. Trends Biochem Sci, 40(3), 165-171.

Tefera, M. (2004). Observations on the clinical examination of the camel (Camelus dromedarius) in the field. Trop Anim Health Prod, 36(5), 435-449.

Tsuchiya, T., N. Kameya and I. Nakamura (2009). Straight walk: a modified method of ligation-mediated genome walking for plant species with large genomes. Anal Biochem, 388(1), 158-160.

Ulmasov, H. A., K. K. Karaev, V. N. Lyashko and M. B. Evgen'ev (1993). Heat-shock response in camel (Camelus dromedarius) blood cells and adaptation to hyperthermia. Comp Biochem Physiol B, 106(4), 867-872.

Wada, K., A. Taniguchi and T. Okano (2007). Highly sensitive detection of cytotoxicity using a modified HSP70B' promoter. Biotechnol Bioeng, 97(4): 871-876.

Weingarten-Gabbay, S., and E. Segal (2014). The grammar of transcriptional regulation. Hum Genet, 133(6), 701-711.

Wu, B. J., R. E. Kingston and R. I. Morimoto (1986). Human HSP70 promoter contains at least two distinct regulatory domains. Proc Natl Acad Sci U S A, 83(3): 629-633.

Yang, M. Q., K.Laflamme, V. Gotea, C. H. Joiner, N. E. Seidel, C. Wong (2011). Genome-wide detection of a TFIID localization element from an initial human disease mutation. Nucleic Acids Res, 39(6), 2175-2187.