Journal of Bioinformatics and Sequence Analysis Vol. 3(3), pp. 37-62, March 2011
Available online at http://www.academicjournals.org/JBSA
ISSN 2141-2464 ©2011 Academic Journals

*Full Length Research Paper*

# Polymorphisms in the transcription factor binding sites of host genes influences evolutionary susceptibility to falciparum malaria

**Tabish Qidwai[1], Prabhash K. Pandey[2], Sangram Singh[2] and Farrukh Jamal[2]***

[1]Department of Biochemistry, University of Allahabad, Allahabad, U. P., India.
[2]Department of Biochemistry, Dr. Ram Manohar Lohia Avadh University, Faizabad-224001, U.P., India.

Accepted 22 February, 2011

**Gene expression at the level of transcription is regulated by a set of transcription factors (TFs) that recognizes *cis* elements. We accessed the human promoters from eukaryotic promoter database. These sequences have been run in P-match tool. MEME software has been used for detection of conserved sequences in the promoter region. All the predicted known TFs and their binding sites along with weight matrices were collected from TRANSFAC database under vertebrate TFs category. P-match tool combines pattern matching and weight matrix approaches thus providing higher accuracy of recognition than each of the methods alone. P-Match is closely interconnected with the TRANSFAC® database. Using results of extensive tests of recognition accuracy, we selected three sets of optimized cut-off values that minimize either false negatives or false positives, or the sum of both errors. In this report, we focus on those polymorphisms of transcription factor binding sites (TFBS) in the regulatory region of host genes and hypothesize that these variation increases the susceptibility/resistance to a particular disease by alteration of gene product in the cell. Therefore, we have concluded that 124 promoter polymorphisms in the 9 genes involved in malaria pathogenesis play important role in susceptibility to falciparum malaria.**

**Key words:** Single-nucleotide polymorphism (SNP), transcription factors, transcription factors binding sites, evolution, infectious diseases.

## INTRODUCTION

Although, human cells have the same set of genome but cellular differentiation gives rise to different types of cells and tissues. The formation of different types of tissues from the same genetic material requires coordinated and controlled expression of the genes in the genome. This process involves interplay of a large number of factors involved in gene expression including enhancers, silencers and gene methylation (Frisch et al., 2002; Robertson et al., 2002; Werner, 1999).

Specific nuclear factors called transcription factors binds with selected DNA sequences called transcription factor binding site. The expression of a gene is primarily regulated by transcription factors that interact with regulatory *cis*-elements on DNA sequences. The identification of functional regulatory elements can be done by computer searching using software (s) which can predict TF binding sites using position weight matrices (PWMs) that represent positional base frequencies of collected experimentally determined transcription factor binding sites (TFBS) (Murakami et al., 2004).

Promoter sequences are generally present upstream from the transcription start site but in some cases, it overlaps with first exon of the gene. So promoters are considered as the processors of transcriptional regulation. TF binding site detection is an important tool for the study of transcriptional regulation of gene expression. Zhu et al. (2002) reported the computational modeling of promoter organization as a tool to study the transcriptional regulation. Pipel et al. (2001) used the computational approaches based on DNA array data to study genome wide transcriptional regulation.

---

*Corresponding author. E-mail: farrukhrmlau@gmail.com, journal.farrukh@gmail.com. Fax: +91-05278-246330.

The human genome has more than 2500 genes, which are regulated spatially, temporally and show dramatic diversity (Xueping et al., 2006). All these genes are controlled by less than 2000 transcription factors. Individual genes have binding site for multiple transcription factors. These transcriptional factors bind and work in combination to control the individual genes. This is termed as combinatorial gene regulation, a common process of gene regulation in higher eukaryotes including humans. Xeuping et al. (2006) reported the computational analysis of tissue specific combinatorial gene regulation. They described the interaction of known human TF to their binding sites and predicted the TF pairs that may co-regulate the gene expression. Wasserman and Fickett (1998) employed the concept of a regulatory module (cluster of TFBS) to predict muscle- and liver-specific regulatory regions. In more complex systems, the functional transcription factor binding sites within promoters are organized hierarchically (Klingenhoff et al., 1999, 2002; Krivan and Wasserman, 2001).

In this paper, we used P- match tool for identification of transcription factor binding sites and regulatory analysis of variations in enhancers (RAVEN) tool to predict the polymorphisms in the enhancer region of genes involved in malaria pathogenesis. In addition, the promoters of housekeeping genes (constitutive expression) and regulatory genes have been compared and percent conservation has been calculated to understand their significance during the course of evolution.

## METHODOLOGY

### Prediction of transcription factor binding sites in human promoters

We accessed the human promoters from eukaryotic promoter database. Promoter sequences are taken from -1000 to +100. These sequences have been run in P-match tool. All the predicted known TFs and their binding sites along with weight matrices were collected from TRANSFAC database (www.gene-regulation.de) under vertebrate TFs category. P-match tool combines pattern matching and weight matrix approaches thus providing higher accuracy of recognition than each of the methods alone. P-Match is closely interconnected with the TRANSFAC® database. In particular, P-match uses the matrix library as well as sets of aligned known TF-binding sites collected in TRANSFAC® and therefore provides the possibility to search for a large variety of different TF binding sites. Using results of extensive tests of recognition accuracy, we selected three sets of optimized cut-off values that minimize either false negatives or false positives, or the sum of both errors. A public version of the P-Match tool is available at http://www.pubmedcentral.nih.gov/redirect3.cgi?&&.

### Prediction of conserved motifs in human promoters

MEME software has been used for detection of conserved sequences in the promoter region. We accessed the promoter sequences of different organisms from transcription regulatory element database (TRED) and NCBI (National Center for Biotechnology Information). A set of promoter sequences of different organisms run through MEME software.

### Prediction of DNA sequence variation in transcription factor binding sites

We used RAVEN (Regulatory Analysis of Variations in Enhancers) web tool for prediction of SNPs extending in the region -3000 to 500 base pairs, affecting the TFBS (Identification of transcription factor binding sites). RAVEN system is a web tool and is available at http://www.cisreg.ca. According to the keywords entered, search engine gives a list of human gene. By clicking the gene of interest, genome location of gene is displayed by the software. Selection of the genomic regions from -3000 to 500 bp gives results in graphical and in tabulated forms. In the result view, we have option for analysis of SNPs with a particular transcription factor or the entire transcription factor, JASPAR, an open-access database for eukaryotic transcription factor binding profiles.

## RESULTS

Table 1 shows TF, their binding site and weight matrices in TRANSFAC database. We have accessed some promoter sequences from NCBI and formed a set of promoters for the same gene in different organisms. These promoter sequences were analyzed through MEME software and results are shown in Table 3. Predicted polymorphisms in transcription factor binding sites are given in Table 2. Table 4 shows comparison of conservation patterns in the promoter region of some genes.

## DISCUSSION

The aim of this study was to predict the role of specific sequences involved in the regulation of gene expression of host genes involved in malaria pathogenesis. Interaction of transcription factor to transcription factor binding site mediate the transcriptional response to specific signal transduction pathways, cell type-specific expression and events central to developmental regulation (Stormo, 2000). Several transcription factor binding sites in host genes are present (Table 1). A given promoter module may show a robust stimulus specific response in one tissue but in a second cell type, may not be functional. This can result, for example, from a different complement or concentration of specific transcription factors or from selective signaling events further upstream in an activation pathway. Promoter modules can also exhibit cooperative protein binding and often include one binding element that represents a poor binding site for a specific transcription factor. Through cooperative effects, the stronger binding protein partner can stabilize the binding of the weaker partner, but the loss of either binding site abolishes its function. This highlights an important aspect of transcriptional regulation, namely, that a weak binding site embedded in the correct context can be functionally as important as a

**Table 1.** Showing TF their binding site and weight matrices in TRANSFAC-DATABASE.

| Matrix identifier | Position (strand) | Core match | Matrix match | Sequence (always the (+)-strand is shown) | Factor name |
|---|---|---|---|---|---|
| **IL-1B (Interleukin-1 beta)** | | | | | |
| V$NFKB_Q6 | 108 (-) | 1.000 | 0.976 | ttggaaaGTCCCag | NF-kappaB |
| V$NFKB_C | 109 (-) | 1.000 | 0.997 | tggaaaGTCCCa | NF-kappaB |
| V$CREL_01 | 110 (-) | 1.000 | 0.990 | GGAAAgtccc | c-Rel |
| V$NFKAPPAB_01 | 110 (-) | 0.986 | 0.992 | ggaaaGTCCC | NF-kappaB |
| V$HAND1E47_01 | 792 (-) | 1.000 | 0.961 | gtttCCAGAccctgga | Hand1/E47 |
| V$OCT1_Q6 | 1306 (+) | 0.888 | 0.912 | actcatGGAAAtgat | Oct-1 |
| V$AP1_Q4 | 1706 (+) | 1.000 | 0.990 | agTGACTcaag | AP-1 |
| V$CREB_Q2 | 2055 (-) | 1.000 | 0.983 | tgttaCGTCAgg | CREB |
| V$CREBP1_Q2 | 2055 (-) | 1.000 | 0.979 | tgttaCGTCAgg | CRE-BP1 |
| V$VBP_01 | 2056 (+) | 1.000 | 0.974 | gTTACGtcag | VBP |
| V$AP1_Q4 | 2165 (-) | 1.000 | 0.991 | gatgAGTCAct | AP-1 |
| V$EVI1_04 | 2892 (+) | 0.842 | 0.874 | agaaaagaaAAGAAa | Evi-1 |
| V$OCT1_Q6 | 3007 (-) | 1.000 | 0.894 | ttaaTTTGCaaagtg | Oct-1 |
| V$COMP1_01 | 4580 (+) | 1.000 | 0.838 | cacaacGATTGtcaggaaaacaat | COMP1 |
| V$OCT1_Q6 | 4605 (-) | 1.000 | 0.944 | cataTTTGCatggtg | Oct-1 |
| | | | | | |
| **IL-1A (Interleukin-1 alpha)** | | | | | |
| V$HAND1E47_01 | 390 (+) | 1.000 | 0.962 | cctggtgTCTGGatcc | Hand1/E47 |
| V$NKX25_01 | 780 (+) | 1.000 | 1.000 | tcAAGTG | Nkx2-5 |
| V$SOX9_B1 | 861 (+) | 1.000 | 0.992 | aaggaACAATggca | SOX-9 |
| V$AP1_Q4 | 929 (-) | 1.000 | 0.982 | cttcAGTCAcc | AP-1 |
| V$CREL_01 | 1097 (-) | 1.000 | 0.982 | GGAAAgaccc | c-Rel |
| V$AP1_Q4 | 1565 (+) | 1.000 | 0.996 | agTGACTcatt | AP-1 |
| V$HNF1_C | 1947 (-) | 0.829 | 0.862 | ttggggaattaTTTACa | HNF-1 |
| V$HNF1_C | 1949 (+) | 0.771 | 0.858 | gGGGAAttatttacaac | HNF-1 |
| V$AP1_Q4 | 2206 (+) | 1.000 | 0.982 | agTGACTgaag | AP-1 |
| V$OCT1_Q6 | 2219 (-) | 0.883 | 0.915 | atcaTTAGCattgtc | Oct-1 |
| V$NKX25_01 | 2771 (+) | 1.000 | 1.000 | tcAAGTG | Nkx2-5 |
| | | | | | |
| **IL1R1 (Interleukin-1 receptor-1)** | | | | | |
| V$NKX25_01 | 115 (-) | 1.000 | 1.000 | CACTTga | Nkx2-5 |
| V$OCT1_Q6 | 196 (-) | 0.909 | 0.892 | cccaTTTACatgggt | Oct-1 |
| V$AP1_Q4 | 441 (+) | 1.000 | 0.991 | ggTGACTaatc | AP-1 |
| V$FOXD3_01 | 1225 (-) | 1.000 | 0.966 | aaacaAACAAac | FOXD3 |
| V$FOXD3_01 | 1229 (-) | 1.000 | 0.966 | aaacaAACAAac | FOXD3 |
| V$FOXD3_01 | 1233 (-) | 1.000 | 0.966 | aaacaAACAAac | FOXD3 |
| V$FOXD3_01 | 1237 (-) | 1.000 | 0.966 | aaacaAACAAac | FOXD3 |
| V$FOXD3_01 | 1241 (-) | 1.000 | 0.966 | aaacaAACAAac | FOXD3 |
| V$FOXD3_01 | 1245 (-) | 1.000 | 0.966 | aaacaAACAAac | FOXD3 |
| V$FOXD3_01 | 1249 (-) | 1.000 | 0.966 | aaacaAACAAac | FOXD3 |
| V$HNF1_01 | 1504 (+) | 1.000 | 0.933 | ggTTAATatttggga | HNF-1 |
| V$CREL_01 | 2356 (+) | 0.973 | 0.983 | ggggaATTCC | c-Rel |
| V$COMP1_01 | 2926 (+) | 1.000 | 0.805 | ggatgtGATTGtcacataaagccg | COMP1 |
| V$ELK1_02 | 2975 (+) | 1.000 | 0.990 | gcctcCGGAAgtgg | Elk-1 |
| V$CETS1P54_01 | 2978 (+) | 1.000 | 0.995 | tCCGGAagtg | c-Ets-1(p54) |
| V$PAX4_01 | 3465 (-) | 0.881 | 0.835 | tcatccccaccCCTCAactct | Pax-4 |
| V$HAND1E47_01 | 3563 (-) | 1.000 | 0.963 | atctCCAGAccccagt | Hand1/E47 |
| V$COMP1_01 | 4086 (+) | 0.856 | 0.801 | tggaagAATTGacttccagaccca | COMP1 |
| V$HAND1E47_01 | 4097 (-) | 1.000 | 0.970 | acttCCAGAcccagga | Hand1/E47 |

**Table 1.** Contd.

| | | | | | |
|---|---|---|---|---|---|
| V$MYOD_01 | 4271 (+) | 1.000 | 0.971 | ggtCAGGTgctg | MyoD |
| V$CDPCR1_01 | 5066 (+) | 0.929 | 0.925 | aATTGAtggg | CDP CR1 |
| V$COMP1_01 | 5192 (+) | 0.786 | 0.802 | gttattGATGGccattaaaaggcc | COMP1 |
| V$CDPCR1_01 | 5194 (+) | 0.929 | 0.929 | tATTGAtggc | CDP CR1 |
| V$USF_C | 5604 (+) | 1.000 | 1.000 | cCACGTgc | USF |
| | | | | | |
| **IL1RN (Interleukin-1 receptor antagonist)** | | | | | |
| V$COMP1_01 | 467 (-) | 0.914 | 0.801 | agcctccttttgcCAATGttgcaa | COMP1 |
| V$OCT1_Q6 | 822 (+) | 1.000 | 0.951 | gaaaatGCAAAttga | Oct-1 |
| V$PAX4_01 | 907 (-) | 0.979 | 0.842 | tctgccccaagCCTGAggccc | Pax-4 |
| V$PAX6_01 | 1812 (-) | 0.792 | 0.845 | aaatagtcaagCTTAAagaag | Pax-6 |
| V$EVI1_06 | 1914 (+) | 1.000 | 1.000 | ACAAGataa | Evi-1 |
| V$OCT1_06 | 2144 (+) | 0.982 | 0.971 | ctaatTGAGAtgta | Oct-1 |
| V$COMP1_01 | 2175 (-) | 0.842 | 1.000 | gctttttgtatgtCAATCgtacct | COMP1 |
| V$ELK1_01 | 2542 (+) | 0.955 | 1.000 | gtgacaGGAAGtgcag | Elk-1 |
| V$EVI1_04 | 2594 (+) | 0.842 | 0.862 | agttatgagAAGTTa | Evi-1 |
| V$PAX4_01 | 2749 (+) | 0.979 | 0.857 | tggatTCAGGagtcctcagtt | Pax-4 |
| V$GATA3_03 | 3264 (+) | 0.981 | 0.982 | acaGATATta | GATA-3 |
| V$HAND1E47_01 | 3454 (+) | 1.000 | 0.961 | agctgagTCTGGaagc | Hand1/E47 |
| V$IK1_01 | 4305 (+) | 1.000 | 0.989 | acctGGGAAtgcc | Ik-1 |
| V$COMP1_01 | 4326 (+) | 1.000 | 0.803 | tgtgctGATTGacttgtagtcccc | COMP1 |
| V$HNF1_C | 4444 (+) | 0.942 | 0.852 | gGCTAAtacttaaataa | HNF-1 |
| V$ELK1_01 | 4886 (+) | 1.000 | 0.922 | agaacaGGAAGtgaag | Elk-1 |
| V$CREL_01 | 4954 (+) | 1.000 | 1.000 | ggggcTTTCC | c-Rel |
| V$OCT1_Q6 | 5109 (+) | 1.000 | 0.891 | ccttctGCAAAtgag | Oct-1 |
| V$ELK1_02 | 5296 (-) | 1.000 | 0.987 | atgcTTCCGgtgag | Elk-1 |
| V$HNF4_01 | 5743 (-) | 1.000 | 0.931 | acaaggcCTTTGcccttgc | HNF-4 |
| V$AP1_Q4 | 6194 (-) | 1.000 | 0.995 | cctgAGTCAcc | AP-1 |
| V$VMYB_01 | 741 (+) | 1.000 | 0.960 | attAACGGct | v-Myb |
| V$HNF1_C | 938 (-) | 0.771 | 0.843 | ttggccattcaTTCCCc | HNF-1 |
| V$PAX4_01 | 1038 (-) | 1.000 | 0.847 | ggaaggcaaggCATGAaccca | Pax-4 |
| V$ELK1_02 | 1355 (-) | 1.000 | 0.991 | cgacTTCCGgcctc | Elk-1 |
| V$CETS1P54_01 | 1356 (-) | 1.000 | 0.995 | gactTCCGGc | c-Ets-1(p54) |
| V$HAND1E47_01 | 1386 (-) | 1.000 | 0.965 | agttCCAGAttcctgt | Hand1/E47 |
| V$OCT1_Q6 | 1472 (-) | 0.893 | 0.922 | ttgaTTTTCattcac | Oct-1 |
| V$NKX25_01 | 1555 (+) | 1.000 | 1.000 | tcAAGTG | Nkx2-5 |
| V$FOXJ2_02 | 1775 (-) | 1.000 | 0.970 | caaaatATTATtgg | FOXJ2 |
| V$GRE_C | 2666 (+) | 0.978 | 0.951 | ggcacaggatGTCCTg | GR |
| V$VMAF_01 | 2709 (+) | 1.000 | 0.896 | aactGCTGAcgttggagct | v-Maf |
| V$PAX5_01 | 2813 (-) | 0.988 | 0.911 | tttctcgctatgctgCCCTGccctttgc | BSAP |
| V$NKX25_01 | 3229 (-) | 1.000 | 1.000 | CACTTga | Nkx2-5 |
| V$PAX4_01 | 3442 (+) | 0.979 | 0.847 | gcaggTCAGGggtgcctgcca | Pax-4 |
| V$PAX6_01 | 4255 (+) | 0.842 | 0.856 | cccctTCATGgttgacttcag | Pax-6 |
| V$NKX25_01 | 4738 (-) | 1.000 | 1.000 | CACTTga | Nkx2-5 |
| V$NKX25_01 | 5096 (+) | 1.000 | 1.000 | tcAAGTG | Nkx2-5 |
| V$HNF1_C | 5363 (-) | 1.000 | 0.884 | atagcaactcaTTAACa | HNF-1 |
| V$NFY_Q6 | 5559 (-) | 1.000 | 0.988 | gttATTGGtta | NF-Y |
| V$NKX25_01 | 6048 (-) | 1.000 | 1.000 | CACTTga | Nkx2-5 |
| V$OCT1_Q6 | 6065 (+) | 1.000 | 0.903 | agcagtGCAAAtaat | Oct-1 |
| V$OCT1_Q6 | 6115 (+) | 0.893 | 0.926 | cagaatGAAAAtcag | Oct-1 |
| V$COMP1_01 | 6826 (-) | 1.000 | 0.823 | tatgctttctagaCAATCttagcc | COMP1 |
| V$COMP1_01 | 7233 (-) | 1.000 | 0.818 | tggaactcgtggcCAATCttcctg | COMP1 |

**Table 1.** Contd.

| | | | | | |
|---|---|---|---|---|---|
| V$OCT1_Q6 | 7977 (+) | 0.909 | 0.906 | ggttatGTAAAttcc | Oct-1 |
| V$ELK1_01 | 8082 (-) | 1.000 | 0.931 | gtgggCTTCCtgtgtt | Elk-1 |
| V$HNF1_C | 8094 (+) | 0.853 | 0.841 | tGTTATttaatgaaccg | HNF-1 |
| V$HNF4_01 | 8287 (-) | 1.000 | 0.924 | atttgtgCTTTGttcttca | HNF-4 |
| V$HNF4_01 | 9432 (+) | 0.915 | 0.904 | cactggcTAAAGatcaaaa | HNF-4 |
| | | | | | |
| **Il-3 (Interleukin-5)** | | | | | |
| V$PAX4_01 | 355 (-) | 0.979 | 0.833 | tgccactcactCCTGAccctg | Pax-4 |
| V$HAND1E47_01 | 696 (+) | 1.000 | 0.962 | ggctgtgTCTGGctgt | Hand1/E47 |
| V$CREL_01 | 815 (-) | 1.000 | 0.987 | GGAAAaccca | c-Rel |
| V$HNF4_01 | 975 (-) | 0.883 | 0.896 | ctatgtcCTTTCtccccca | HNF-4 |
| V$USF_C | 1056 (-) | 1.000 | 1.000 | gcACGTGg | USF |
| V$PAX4_01 | 1348 (+) | 0.986 | 0.881 | tgaggTCAAGagtttgagacc | Pax-4 |
| V$PAX4_01 | 1470 (-) | 0.986 | 0.837 | cacaaaaatcgCTTGAaccca | Pax-4 |
| V$SOX9_B1 | 1764 (-) | 1.000 | 0.995 | acccATTGTtccct | SOX-9 |
| V$CREL_01 | 1808 (+) | 1.000 | 1.000 | ggggaTTTCC | c-Rel |
| V$AP1_Q4 | 2124 (+) | 1.000 | 0.988 | ggTGACTgagc | AP-1 |
| V$CREL_01 | 2194 (-) | 1.000 | 0.989 | GGAAAatccc | c-Rel |
| V$VMAF_01 | 2317 (+) | 1.000 | 0.892 | cactGCTGAcaaagggacc | v-Maf |
| V$EVI1_04 | 2453 (+) | 0.842 | 0.823 | taataaaatAAAATa | Evi-1 |
| V$HNF4_01 | 3152 (-) | 1.000 | 0.923 | gcaggtaCTTTGacccgga | HNF-4 |
| V$PAX4_01 | 3940 (-) | 1.000 | 0.847 | gagcaagaaccCATGAgtgcc | Pax-4 |
| V$PAX4_01 | 4248 (-) | 0.788 | 0.833 | ggcccagcacgGGTGAgggcg | Pax-4 |
| V$COMP1_01 | 4950 (+) | 1.000 | 0.832 | gccctgGATTGaggccaacagatg | COMP1 |
| V$AP1_Q4 | 5562 (+) | 1.000 | 0.987 | agTGACTgagg | AP-1 |
| V$ARP1_01 | 6117 (-) | 1.000 | 0.927 | tcggtcCAAGGgctca | ARP-1 |
| V$MYOD_01 | 7355 (-) | 1.000 | 0.992 | caccACCTGtct | MyoD |
| V$PAX4_01 | 9156 (-) | 0.986 | 0.846 | ctttctcctccCTTGAcctca | Pax-4 |
| | | | | | |
| **IL-5 (Interleukin-5)** | | | | | |
| V$PAX4_01 | 867 (-) | 0.986 | 0.837 | caggagaatcgCTTGAaccca | Pax-4 |
| V$OCT1_Q6 | 1057 (-) | 0.883 | 0.903 | ttcaTTAGCataagg | Oct-1 |
| V$HNF1_C | 1400 (+) | 0.853 | 0.862 | tGTTATttattcaatta | HNF-1 |
| V$FOXD3_01 | 1764 (+) | 1.000 | 0.992 | aaTTGTTtattt | FOXD3 |
| V$HFH1_01 | 1764 (+) | 1.000 | 0.997 | aattGTTTAttt | HFH-1 |
| V$PAX4_01 | 2013 (+) | 0.979 | 0.877 | cgaggTCAGGagtttgagacc | Pax-4 |
| V$CP2_01 | 2594 (-) | 0.961 | 0.976 | CTGGAtagggc | CP2 |
| V$HNF1_C | 2897 (+) | 0.852 | 0.904 | aGTTAGttatttaacaa | HNF-1 |
| V$CDPCR1_01 | 2958 (-) | 1.000 | 0.934 | gagaTCGATt | CDP CR1 |
| V$HNF1_01 | 3179 (-) | 1.000 | 0.933 | aaacaatcATTAAct | HNF-1 |
| V$HNF3B_01 | 3209 (+) | 1.000 | 0.966 | attatTATTTattta | HNF-3beta |
| V$FOXD3_01 | 3211 (+) | 0.948 | 0.955 | taTTATTtattt | FOXD3 |
| V$STAF_02 | 3496 (+) | 0.938 | 0.926 | ctgACCCAtcatgcctcgccg | Staf |
| V$SOX9_B1 | 3617 (-) | 1.000 | 0.990 | tcccATTGTtaaca | SOX-9 |
| V$CEBP_C | 3670 (-) | 0.985 | 0.953 | ggtacATTACtaaactcc | C/EBP |
| V$HNF4_01 | 3944 (+) | 0.915 | 0.891 | acagggaTAAAGttccttt | HNF-4 |
| V$EVI1_04 | 3955 (-) | 0.772 | 0.844 | gTTCCTttcttatca | Evi-1 |
| V$HNF1_C | 4219 (-) | 1.000 | 0.871 | gattaaaatcaTTAACt | HNF-1 |
| V$HNF1_01 | 4221 (-) | 1.000 | 0.936 | ttaaaatcATTAAct | HNF-1 |
| V$FOXJ2_02 | 4793 (-) | 1.000 | 0.925 | tgaaatATTATtca | FOXJ2 |
| V$FOXD3_01 | 5151 (-) | 0.996 | 0.956 | aattaAACATtt | FOXD3 |
| V$FOXD3_01 | 5365 (+) | 1.000 | 0.960 | ttTTGTTtgttt | FOXD3 |

**Table 1.** Contd.

| | | | | | |
|---|---|---|---|---|---|
| V$FOXD3_01 | 5369 (+) | 1.000 | 0.966 | gtTTGTTtgttt | FOXD3 |
| V$FOXD3_01 | 5373 (+) | 1.000 | 0.966 | gtTTGTTtgttt | FOXD3 |
| V$RFX1_01 | 5587 (+) | 1.000 | 0.989 | gtgtcacctgGCAACca | RFX1 |
| V$HNF4_01 | 6145 (-) | 1.000 | 0.908 | aagagcaCTTTGaccaagc | HNF-4 |
| V$NKX25_01 | 6266 (+) | 1.000 | 1.000 | tcAAGTG | Nkx2-5 |
| V$GRE_C | 6496 (+) | 1.000 | 0.967 | ggtaccatctGTTCTt | GR |
| V$TAL1ALPHAE47_01 | 6496 (-) | 1.000 | 0.993 | ggtacCATCTgttctt | Tal-1alpha/E47 |
| V$TAL1BETAE47_01 | 6496 (-) | 1.000 | 0.996 | ggtaccATCTGttctt | Tal-1beta/E47 |
| V$PAX4_01 | 6516 (-) | 0.979 | 0.833 | ctcaagacaggCCTGAagtca | Pax-4 |
| V$PAX4_01 | 6529 (+) | 0.979 | 0.852 | tgaagTCAGGcttctaggctg | Pax-4 |
| V$CAAT_01 | 6929 (+) | 1.000 | 0.989 | gagagCCAATca | CCAAT box |
| V$AP1_Q4 | 7252 (-) | 1.000 | 0.992 | ttttAGTCAct | AP-1 |
| V$FOXD3_01 | 7422 (-) | 0.996 | 0.986 | aaaaaAACATtc | FOXD3 |
| V$OCT1_Q6 | 7533 (+) | 1.000 | 0.951 | agaaatGCAAAtgtg | Oct-1 |
| **IRF1 (Interferon regulatory factor-1)** | | | | | |
| V$PAX4_01 | 351 (-) | 0.986 | 0.874 | tagagtccactCTTGAgcaca | Pax-4 |
| V$VMYB_01 | 1028 (+) | 1.000 | 0.986 | aatAACGGct | v-Myb |
| V$PAX4_01 | 1218 (+) | 1.000 | 0.867 | ggcggTCATGctttctcccac | Pax-4 |
| V$HNF1_C | 1975 (-) | 0.942 | 0.904 | gtgggtatttaTTAGCt | HNF-1 |
| V$IK1_01 | 2297 (-) | 1.000 | 0.983 | gggaTTCCCaggt | Ik-1 |
| V$AP1_Q4 | 2308 (-) | 1.000 | 0.990 | gttgAGTCAcc | AP-1 |
| V$STAT_01 | 3658 (+) | 1.000 | 1.000 | TTCCCggaa | STATx |
| V$VMAF_01 | 4216 (+) | 1.000 | 0.932 | agctGCTGAcccggctgca | v-Maf |
| V$PAX4_01 | 5269 (+) | 0.986 | 0.837 | tgcctTCAAGcttttggatgc | Pax-4 |
| V$CEBP_C | 5304 (+) | 1.000 | 0.984 | tgtgttctGCAATtcaca | C/EBP |
| V$PAX4_01 | 0 (+) | 0.979 | 0.866 | tgaggTCAGGagttggagaac | Pax-4 |
| V$PAX4_01 | 6878 (+) | 1.000 | 0.848 | aaggtTCATGcgttcctgttc | Pax-4 |
| V$OCT1_Q6 | 7691 (+) | 0.909 | 0.939 | gtaaatGTAAAtgac | Oct-1 |
| V$OCT1_Q6 | 7801 (+) | 0.888 | 0.927 | aaaaatGGAAAtaac | Oct-1 |
| V$EVI1_04 | 8365 (-) | 0.772 | 0.832 | tTTCCTctcctagct | Evi-1 |
| V$CREL_01 | 9302 (+) | 1.000 | 0.988 | tgggaTTTCC | c-Rel |
| **NOS2A (Inducible form of nitric oxide synthase)** | | | | | |
| V$PAX6_01 | 398 (-) | 0.925 | 0.878 | actcactcattCGTCAatggt | Pax-6 |
| V$CDPCR1_01 | 727 (-) | 0.929 | 0.929 | gccaTCAATt | CDP CR1 |
| V$USF_Q6 | 817 (-) | 1.000 | 0.974 | gccaCGTGAa | USF |
| V$BRN2_01 | 901 (+) | 1.000 | 0.996 | aacatccAAAATaagc | Brn-2 |
| V$USF_Q6 | 1115 (-) | 1.000 | 0.993 | ggcaCGTGAc | USF |
| V$ELK1_02 | 1574 (-) | 1.000 | 0.982 | ttccTTCCGgcatg | Elk-1 |
| V$HNF4_01 | 1819 (-) | 0.915 | 0.902 | ctttgccCTTTAaacagtt | HNF-4 |
| V$MYOD_01 | 2256 (+) | 1.000 | 0.971 | gggCAGGTgctg | MyoD |
| V$VMYB_01 | 2438 (-) | 1.000 | 0.995 | tgCCGTTatt | v-Myb |
| V$HAND1E47_01and1/E47 | 2533 (-) | 1.000 | 0.974 | agagCCAGAccccacc | |
| V$EVI1_04 | 3021 (+) | 1.000 | 0.875 | ggtaaagaaAAGATa | Evi-1 |
| V$NKX25_01 | 3269 (+) | 1.000 | 1.000 | tcAAGTG | Nkx2-5 |
| V$E47_01 | 3468 (+) | 1.000 | 0.995 | actgCAGGTgttcac | E47 |
| V$COMP1_01 | 4638 (+) | 0.786 | 0.801 | ggggagGCTTGacaagaaacgagg | COMP1 |
| V$CP2_01 | 4696 (+) | 0.974 | 0.972 | gcaccaGCCAG | CP2 |
| V$NKX25_01 | 4916 (-) | 1.000 | 1.000 | CACTTga | Nkx2-5 |
| V$USF_Q6 | 5047 (-) | 1.000 | 0.997 | gacaCGTGAc | USF |
| V$USF_Q6 | 5047 (+) | 0.973 | 0.976 | gACACGtgac | USF |

**Table 1.** Contd.

| | | | | | |
|---|---|---|---|---|---|
| V$CREL_01 | 5644 (-) | 1.000 | 0.987 | GGAAAgccca | c-Rel |
| V$MYOD_01 | 5724 (-) | 1.000 | 0.971 | caccACCTGact | MyoD |
| V$EVI1_04 | 6481 (+) | 0.930 | 0.838 | ggttaagagAGGATt | Evi-1 |
| V$OCT1_02 | 6521 (-) | 0.992 | 0.989 | attgtGAATAttcta | Oct-1 |
| V$EVI1_04 | 7145 (+) | 1.000 | 0.835 | caagatgagAAGATa | Evi-1 |
| V$EVI1_04 | 7329 (-) | 0.772 | 0.844 | tTTCCTttcctttct | Evi-1 |
| V$EVI1_04 | 7334 (-) | 0.772 | 0.849 | tTTCCTttcttttcc | Evi-1 |
| V$EVI1_04 | 7339 (-) | 0.842 | 0.869 | tTTCTTttcctttcc | Evi-1 |
| V$EVI1_04 | 7344 (-) | 0.772 | 0.844 | tTTCCTttcctttcc | Evi-1 |
| V$EVI1_04 | 7349 (-) | 0.772 | 0.844 | tTTCCTttcctttcc | Evi-1 |
| V$EVI1_04 | 7354 (-) | 0.772 | 0.844 | tTTCCTttcctttcc | Evi-1 |
| V$EVI1_04 | 7359 (-) | 0.772 | 0.844 | tTTCCTttcctttcc | Evi-1 |
| V$EVI1_04 | 7364 (-) | 0.772 | 0.844 | tTTCCTttcctttcc | Evi-1 |
| V$EVI1_04 | 7369 (-) | 0.772 | 0.844 | tTTCCTttcctttcc | Evi-1 |
| V$EVI1_04 | 7374 (-) | 0.772 | 0.844 | tTTCCTttcctttcc | Evi-1 |
| V$EVI1_04 | 7379 (-) | 0.772 | 0.844 | tTTCCTttcctttcc | Evi-1 |
| V$EVI1_04 | 7384 (-) | 0.772 | 0.844 | tTTCCTttcctttcc | Evi-1 |
| V$EVI1_04 | 7389 (-) | 0.772 | 0.844 | tTTCCTttcctttcc | Evi-1 |
| V$EVI1_04 | 7394 (-) | 0.772 | 0.844 | tTTCCTttcctttcc | Evi-1 |
| V$HAND1E47_01 | 7489 (-) | 1.000 | 0.967 | actcCCAGActcaatt | Hand1/E47 |
| V$PAX4_01 | 7956 (+) | 0.986 | 0.845 | tgtccTCAAGcctggtgcttg | Pax-4 |
| V$YY1_02 | 8502 (+) | 1.000 | 0.959 | gcccgGCCATcttgtcactt | YY1 |
| V$PAX6_01 | 8595 (-) | 0.818 | 0.825 | ttgagctgaagCCTGAaggga | Pax-6 |
| V$PAX4_01 | 8831 (+) | 0.986 | 0.863 | tgagaTCAAGggtgacttttt | Pax-4 |
| V$NKX25_01 | 9026 (-) | 1.000 | 1.000 | CACTTga | Nkx2-5 |
| V$OCT1_Q6 | 9320 (-) | 0.888 | 0.908 | atcaTTTCCattata | Oct-1 |
| V$NKX25_01 | 9467 (+) | 1.000 | 1.000 | tcAAGTG | Nkx2-5 |
| **TGF-beta (Transforming growth factor beat-1)** | | | | | |
| V$PAX4_01 | 31 (+) | 0.879 | 0.833 | tggcgTGACGcggaaggcggg | Pax-4 |
| V$CREL_01 | 91 (+) | 1.000 | 1.000 | ggggcTTTCC | c-Rel |
| V$PAX4_01 | 185 (+) | 1.000 | 0.861 | ggcttTCATGggttggcaatt | Pax-4 |
| V$OCT1_Q6 | 545 (-) | 0.888 | 0.914 | ctcaTTTCCatcagt | Oct-1 |
| V$HNF1_01 | 1010 (-) | 1.000 | 0.939 | atttaaatATTAAct | HNF-1 |
| V$HNF4_01 | 1040 (-) | 1.000 | 0.931 | tggagacCTTTGgactcga | HNF-4 |
| V$PAX4_01 | 1272 (+) | 1.000 | 0.869 | agttgTCATGagtactaaacg | Pax-4 |
| V$PAX4_01 | 1462 (-) | 0.979 | 0.831 | atgtcttcaggCCTGAgtcct | Pax-4 |
| V$PAX4_01 | 1463 (+) | 0.979 | 0.844 | tgtctTCAGGcctgagtccta | Pax-4 |
| V$PAX4_01 | 1879 (+) | 0.986 | 0.837 | tgggtTCAAGcgattctccta | Pax-4 |
| V$PAX4_01 | 2171 (+) | 0.986 | 0.837 | tgggtTCAAGcgattctccta | Pax-4 |
| V$PAX4_01 | 3141 (+) | 0.979 | 0.877 | tgaggTCAGGagtttgagacc | Pax-4 |
| V$PAX4_01 | 3266 (-) | 0.986 | 0.837 | caggagaatcgCTTGAaccca | Pax-4 |
| V$EVI1_04 | 3372 (+) | 0.842 | 0.829 | aaataaaatAAAATa | Evi-1 |
| V$FOXJ2_01 | 3382 (+) | 1.000 | 0.994 | aaataaatAAACAgataa | FOXJ2 |
| V$AP1_Q4 | 3418 (+) | 1.000 | 0.993 | ggTGACTcacc | AP-1 |
| V$NKX25_01 | 3469 (-) | 1.000 | 1.000 | CACTTga | Nkx2-5 |
| V$EVI1_04 | 3752 (+) | 0.907 | 0.825 | tgaaaagaaATGATg | Evi-1 |
| V$NKX25_01 | 3823 (-) | 1.000 | 1.000 | CACTTga | Nkx2-5 |
| V$HNF4_01 | 3906 (+) | 1.000 | 0.922 | gccagtgCAAAGgccctga | HNF-4 |
| V$USF_Q6 | 4019 (-) | 0.987 | 0.972 | gccaCGTGGg | USF |
| V$NKX25_01 | 4261 (+) | 1.000 | 1.000 | tcAAGTG | Nkx2-5 |
| V$GATA1_02 | 4347 (+) | 1.000 | 0.992 | gtagAGATAgggtt | GATA-1 |

**Table 1.** Contd.

| | | | | | |
|---|---|---|---|---|---|
| V$PAX4_01 | 4514 (-) | 0.979 | 0.887 | ggtctcgaaccCCTGAcctca | Pax-4 |
| V$NKX25_01 | 4532 (+) | 1.000 | 1.000 | tcAAGTG | Nkx2-5 |
| V$PAX4_01 | 4650 (-) | 0.979 | 0.877 | ggtctcaaactCCTGAcctca | Pax-4 |
| V$AP1_Q4 | 4715 (-) | 1.000 | 0.992 | cgtgAGTCAct | AP-1 |
| V$PAX6_01 | 5347 (+) | 0.842 | 0.823 | cccatTCATGgattaatgggt | Pax-6 |
| V$PAX4_01 | 5877 (+) | 0.979 | 0.888 | tcctgTCAGGcgtgcgggagg | Pax-4 |
| V$GATA1_02 | 5918 (+) | 1.000 | 0.991 | ccgcAGATAgggga | GATA-1 |
| V$LMO2COM_02 | 5921 (+) | 1.000 | 1.000 | cAGATAggg | Lmo2 |
| V$ARNT_01 | 7155 (-) | 1.000 | 0.991 | acagcCACGTgccacc | Arnt |
| V$USF_Q6 | 7158 (+) | 0.987 | 0.983 | gCCACGtgcc | USF |
| V$USF_C | 7159 (+) | 1.000 | 1.000 | cCACGTgc | USF |
| V$PAX4_01 | 7237 (-) | 0.979 | 0.878 | ggtctcgaactCCTGAcctca | Pax-4 |
| V$HNF4_01 | 8601 (-) | 0.915 | 0.897 | ggttggaCTTTAtactgag | HNF-4 |
| V$GATA1_06 | 8699 (+) | 1.000 | 1.000 | ataGATAAga | GATA-1 |
| V$USF_Q6 | 8987 (+) | 0.987 | 0.972 | cCCACGtggc | USF |
| V$PAX4_01 | 9251 (+) | 0.986 | 0.839 | agtggTCAAGagcacagactc | Pax-4 |
| V$LMO2COM_02 | 9357 (-) | 1.000 | 1.000 | cccTATCTg | Lmo2 |
| V$CP2_01 | 9909 (-) | 1.000 | 0.988 | CTGGGtggggc | CP2 |
| **(IFNAR1) Interferon –alpha receptor-1** | | | | | |
| V$EVI1_04 | 16 (+) | 0.842 | 0.845 | agatatggcAAAATa | Evi-1 |
| V$HNF1_C | 20 (-) | 1.000 | 0.841 | atggcaaaataTTAACg | HNF-1 |
| V$XFD2_01 | 693 (-) | 1.000 | 1.000 | tcaatgTTTATaat | XFD-2 |
| V$FOXJ2_02 | 935 (-) | 1.000 | 0.940 | ataagtATTATttt | FOXJ2 |
| V$PAX6_01 | 1066 (-) | 0.960 | 0.817 | aatcgatgaatCGTAAagtga | Pax-6 |
| V$COMP1_01 | 1120 (+) | 1.000 | 0.829 | tggataGATTGcctaaaaatcatt | COMP1 |
| V$AP1_Q4 | 1227 (+) | 1.000 | 0.992 | ggTGACTcacg | AP-1 |
| V$CREL_01 | 1320 (-) | 1.000 | 0.988 | GGAAAtccca | c-Rel |
| V$PAX4_01 | 1405 (-) | 0.986 | 0.837 | caggagaatcgCTTGAacccg | Pax-4 |
| V$OCT1_02 | 2367 (+) | 0.992 | 0.989 | tggaaTATTCacgta | Oct-1 |
| V$HLF_01 | 4248 (-) | 0.939 | 0.938 | attatGTAAT | HLF |
| V$HNF1_C | 4258 (+) | 0.942 | 0.861 | tGCTAAttaattgcttg | HNF-1 |

strong binding site. Additional transcription elements working in concert with modules appear to fine-tune promoter regulation (Malin et al., 2008; Wasserman and Fickett, 1998). Promoters that are constitutively expressed often referred to as housekeeping genes have upstream sequence elements that are recognized by ubiquitous activators. Promoters that are expressed only in certain times or place have sequence elements that require activators that are available only at those times or places.

Different types of variations in the genome alter this interaction of TF to TFBS, thereby effecting the gene expression. Two alterations interfere with gene regulation viz., epigenetic changes and DNA sequence variations and may thus be responsible for the final outcome of disease (Jones and Stephen, 2002). The presence of polymorphisms in coding region affects the activity of protein, while in regulatory region affect the expression levels of genes and thus play important roles in the pathogenesis of many complex diseases. Nucleotide variations in the TFBS may alter the DNA *cis*-elements and protein (TF) interaction and expression of the gene. Promoter polymorphisms affects the level of gene product thus susceptibility to disease (Teeranaipong et al., 2008; Sinha, et al., 2008). A study from Japan reported association of the -1031 (rs1799964), -863A (1800630) and -857 polymorphisms with increased reporter gene expression and increased concanavalin A-stimulated TNF production from peripheral blood mononuclear cells (Higuchi et al., 1998). Predicted124 SNPs in TFBS of host genes involved in malaria pathogenesis might alters the level of gene product by influencing the binding of transcription factor to transcription factor binding site (Table 2).

## Falciparum malaria and selection

In response to disease pressure, genome selects those variations/ polymorphisms (SNP) that are beneficial, providing resistance against the disease. The best

**Table 2.** Single nucleotide polymorphism lies in transcription factor binding sites of genes involved in malaria pathogenesis. Predicted polymorphisms in conserved regions are more likely to influence the regulation of gene.

| SNP ID | Position | TF | Std | Start | End | Allele 1 | Score 1 (%) | Allele 2 | Score 2 (%) | Diff | Con? |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **IL-1 (Interleukin-1)** | | | | | | | | | | | |
| rs6542093 | 113244767 | E74A | - | 113244762 | 113244768 | CTGGCAG | 2.967 (69.2) | CCGGCAG | 6.627 (83.7) | -3.66 | No |
| rs6542093 | 113244767 | NRF-2 | - | 113244760 | 113244769 | ACTGGCAGGC | 5.006 (71.7) | ACCGGCAGGC | 8.338 (82.4) | -3.332 | No |
| rs6542093 | 113244767 | SOX17 | + | 113244764 | 113244772 | GCCAGTGTC | 6.273 (82.6) | GCCGGTGTC | 1.778 (68.1) | 4.495 | No |
| rs6542093 | 113244767 | Snail | + | 113244765 | 113244770 | CCAGTG | 2.989 (72.3) | CCGGTG | 6.063 (83.3) | -3.074 | No |
| rs6542094 | 113244856 | E74A | + | 113244855 | 113244861 | AAGGAAG | 7.540 (87.3) | AGGGAAG | 4.991 (77.2) | 2.549 | No |
| rs10169524 | 113245103 | Broad-complex_1 | + | 113245101 | 113245114 | ATAAAAAGAAAAAT | 9.120 (83.2) | ATGAAAAGAAAAAT | 7.404 (78.7) | 1.716 | No |
| rs10169524 | 113245103 | Broad-complex_4 | + | 113245102 | 113245112 | TAAAAAGAAAA | 7.405 (82.1) | TGAAAAGAAAA | 4.492 (72.8) | 2.913 | No |
| rs10169524 | 113245103 | HMG-IY | + | 113245099 | 113245114 | TGATAAAAAGAAAAAT | 10.306 (86.3) | TGATGAAAAGAAAAAT | 8.706 (82.7) | 1.6 | No |
| rs10169524 | 113245103 | Hunchback | + | 113245099 | 113245108 | TGATAAAAAG | 7.917 (84.9) | TGATGAAAAG | 4.109 (73.2) | 3.808 | No |
| rs10169524 | 113245103 | Hunchback | + | 113245103 | 113245112 | AAAAAGAAAA | 5.356 (77.0) | GAAAAGAAAA | 7.525 (83.7) | -2.169 | No |
| rs10169524 | 113245103 | SQUA | + | 113245099 | 113245112 | TGATAAAAAGAAAA | 8.817 (83.8) | TGATGAAAAGAAAA | 4.664 (75.2) | 4.153 | No |
| rs10169524 | 113245103 | SQUA | + | 113245101 | 113245114 | ATAAAAAGAAAAAT | 8.112 (82.3) | ATGAAAAGAAAAAT | 5.872 (77.7) | 2.24 | No |
| rs10169524 | 113245103 | TBP | + | 113245099 | 113245113 | TGATAAAAAGAAAAA | 6.008 (80.7) | TGATGAAAAGAAAAA | 0.832 (70.4) | 5.176 | No |
| rs6542095 | 113245414 | CF2-II | + | 113245414 | 113245423 | CTATAGataa | 7.044 (82.3) | ttatCTATAA | 5.512 (78.5) | 1.532 | No |
| rs6542095 | 113245414 | MEF2 | - | 113245414 | 113245423 | ttatCTATAG | 9.782 (84.9) | ttatCTATAA | 7.067 (79.2) | 2.715 | No |
| rs10539616 | 113245450 | AGL3 | - | 113245441 | 113245450 | Catataataa | 1.274 (69.6) | CAATATatat | 7.050 (83.1) | -5.776 | No |
| rs10539616 | 113245450 | Broad-complex_1 | + | 113245444 | 113245457 | ttatatGTCATATA | 4.079 (69.9) | GTATATGACAATAT | 9.219 (83.5) | -5.14 | No |
| rs10539616 | 113245450 | Broad-complex_1 | - | 113245445 | 113245458 | GTATATGACatata | 7.666 (79.4) | GTATATGACAATAT | 9.219 (83.5) | -1.553 | No |
| rs10539616 | 113245450 | CF2-II | + | 113245444 | 113245453 | ttatatGTCA | 6.180 (80.1) | ATATatataa | 12.441 (95.6) | -6.261 | No |
| rs10539616 | 113245450 | CF2-II | - | 113245444 | 113245453 | TGACatataa | 2.537 (71.1) | ATATatataa | 12.441 (95.6) | -9.904 | No |
| rs10539616 | 113245450 | CF2-II | - | 113245449 | 113245458 | GTATATGACa | 5.446 (78.3) | ATATatataa | 12.441 (95.6) | -6.995 | No |
| rs10539616 | 113245450 | E4BP4 | - | 113245441 | 113245451 | ACatataataa | 3.464 (70.2) | ATatataataa | 7.675 (80.9) | -4.211 | No |
| rs10539616 | 113245450 | E4BP4 | + | 113245444 | 113245454 | ttatatGTCAT | 3.725 (70.8) | ATatataataa | 7.675 (80.9) | -3.95 | No |
| rs10539616 | 113245450 | E4BP4 | - | 113245446 | 113245456 | ATATGACatat | 2.652 (68.1) | ATatataataa | 7.675 (80.9) | -5.023 | No |
| rs10539616 | 113245450 | E4BP4 | - | 113245448 | 113245458 | GTATATGACat | 3.725 (70.8) | ATatataataa | 7.675 (80.9) | -3.95 | No |
| rs10539616 | 113245450 | HFH-3 | + | 113245442 | 113245453 | tattatatGTCA | 2.046 (68.5) | tattatatATAT | 6.838 (80.1) | -4.792 | No |
| rs10539616 | 113245450 | RORalfa-1 | + | 113245444 | 113245453 | ttatatGTCA | 7.555 (81.2) | tATATTGTCA | 2.536 (67.3) | 5.019 | No |
| rs10539616 | 113245450 | SOX17 | + | 113245444 | 113245452 | ttatatGTC | 1.308 (66.6) | tATATTGTC | 10.360 (95.8) | -9.052 | No |
| rs10539616 | 113245450 | SQUA | - | 113245440 | 113245453 | TGACatataataat | 4.217 (74.3) | ATATatataataat | 10.072 (86.3) | -5.855 | No |
| rs10539616 | 113245450 | SQUA | + | 113245441 | 113245454 | ttattatatGTCAT | 2.525 (70.8) | ATATatataataat | 10.072 (86.3) | -7.547 | No |
| rs10539616 | 113245450 | bZIP910 | - | 113245448 | 113245454 | ATGACat | 7.070 (85.2) | ATGACAA | 2.446 (70.4) | 4.624 | No |
| rs10539616 | 113245450 | TBP | - | 113245437 | 113245451 | ACatataataattta | 2.619 (73.9) | ATATatataataatt | 6.078 (80.8) | -3.459 | No |
| rs10539616 | 113245450 | TBP | - | 113245444 | 113245458 | GTATATGACatataa | 2.177 (73.1) | ATATatataataatt | 6.078 (80.8) | -3.901 | No |
| rs11677416 | 113245471 | AGL3 | + | 113245471 | 113245480 | CTTTTAAAAG | 6.120 (80.9) | TTTTTAAAAG | 1.686 (70.6) | 4.434 | No |
| rs11677416 | 113245471 | AGL3 | - | 113245471 | 113245480 | CTTTTAAAAG | 6.120 (80.9) | TTTTTAAAAG | 1.686 (70.6) | 4.434 | No |
| rs11677416 | 113245471 | Athb-1 | + | 113245465 | 113245472 | CATTCACT | 2.314 (66.9) | CATTCATT | 6.650 (80.6) | -4.336 | No |
| rs11677416 | 113245471 | Broad-complex_4 | - | 113245467 | 113245477 | TTAAAAGTGAA | 5.421 (75.8) | TTAAAAATGAA | 7.220 (81.5) | -1.799 | No |
| rs11677416 | 113245471 | COUP-TF | - | 113245470 | 113245483 | TGACTTTTAAAAGT | 6.613 (75.5) | TGACTTTTAAAAAT | 8.727 (80.1) | -2.114 | No |

**Table 2.** Contd.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs11677416 | 113245471 | Sox-5 | - | 113245470 | 113245476 | TAAAAGT | 2.350 (68.1) | TAAAAAT | 6.624 (84.1) | -4.274 | No |
| rs3783560 | 113246025 | Broad-complex_1 | - | 113246016 | 113246029 | ttaaTataaaaata | 8.628 (81.9) | ttaaCataaaaata | 5.458 (73.5) | 3.17 | No |
| rs3783560 | 113246025 | Broad-complex_4 | - | 113246016 | 113246026 | aTataaaaata | 6.868 (80.4) | aCataaaaata | 4.775 (73.7) | 2.093 | No |
| rs3783560 | 113246025 | Broad-complex_4 | - | 113246018 | 113246028 | taaTataaaaa | 7.571 (82.7) | taaCataaaaa | 4.375 (72.4) | 3.196 | No |
| rs3783560 | 113246025 | HFH-3 | - | 113246023 | 113246034 | atatattaaTat | 7.252 (81.1) | atatattaaCat | 3.063 (71.0) | 4.189 | No |
| rs3783560 | 113246025 | Hunchback | - | 113246017 | 113246026 | aTataaaaat | 6.162 (79.5) | aCataaaaat | 9.332 (89.2) | -3.17 | No |
| rs3783560 | 113246025 | SQUA | - | 113246017 | 113246030 | attaaTataaaaat | 8.084 (82.2) | atttttatGttaat | 4.918 (75.7) | 3.166 | No |
| rs3783560 | 113246025 | SQUA | - | 113246019 | 113246032 | atattaaTataaaa | 8.086 (82.2) | atattaaCataaaa | 4.056 (73.9) | 4.03 | No |
| rs3783560 | 113246025 | TBP | - | 113246012 | 113246026 | aTataaaaataattc | 9.213 (87.0) | aCataaaaataattc | 6.589 (81.8) | 2.624 | No |
| rs1040193 | 113246330 | Broad-complex_4 | - | 113246325 | 113246335 | aaggaGaaaaa | 3.679 (70.2) | aaggaAaaaaa | 7.111 (81.2) | -3.432 | Yes |
| rs1040193 | 113246330 | E74A | - | 113246329 | 113246335 | aaggaGa | 2.540 (67.5) | aaggaAa | 6.669 (83.8) | -4.129 | Yes |
| rs1040193 | 113246330 | HMG-IY | - | 113246321 | 113246336 | gaaggaGaaaaagagc | 6.036 (76.6) | gaaggaAaaaaagagc | 8.059 (81.2) | -2.023 | Yes |
| rs1040193 | 113246330 | HMG-IY | - | 113246325 | 113246340 | agtggaaggaGaaaaa | 5.536 (75.4) | agtggaaggaAaaaaa | 7.743 (80.5) | -2.207 | No |
| rs1040193 | 113246330 | Hunchback | - | 113246325 | 113246334 | aggaGaaaaa | 3.413 (71.0) | aggaAaaaaa | 7.221 (82.7) | -3.808 | Yes |
| rs1040193 | 113246330 | Hunchback | - | 113246326 | 113246335 | aaggaGaaaa | 3.034 (69.8) | aaggaAaaaa | 7.121 (82.4) | -4.087 | Yes |
| rs1040193 | 113246330 | Hunchback | - | 113246327 | 113246336 | gaaggaGaaa | 2.396 (67.9) | gaaggaAaaa | 6.484 (80.5) | -4.088 | Yes |
| rs1516789 | 113247255 | E74A | - | 113247254 | 113247260 | cgggaGg | 2.967 (69.2) | cgggaAg | 7.095 (85.5) | -4.128 | No |
| rs3783555 | 113247324 | Broad-complex_4 | + | 113247320 | 113247330 | tacaAaaaaaa | 7.111 (81.2) | tacaGaaaaaa | 5.312 (75.4) | 1.799 | No |
| rs3783555 | 113247324 | HFH-2 | - | 113247321 | 113247332 | atttttttTtgt | 7.084 (81.1) | atttttttCtgt | 4.083 (75.1) | 3.001 | No |
| rs3783555 | 113247324 | HFH-2 | - | 113247323 | 113247334 | tcatttttttTt | 7.353 (81.6) | tcatttttttCt | 3.864 (74.7) | 3.489 | No |
| rs3783555 | 113247324 | Hunchback | + | 113247319 | 113247328 | ctacaAaaaa | 7.899 (84.8) | ctacaGaaaa | 3.811 (72.2) | 4.088 | No |
| rs3783555 | 113247324 | Hunchback | + | 113247320 | 113247329 | tacaAaaaaa | 8.706 (87.3) | tacaGaaaaa | 4.899 (75.6) | 3.807 | No |
| rs3783555 | 113247324 | SQUA | + | 113247319 | 113247332 | ctacaAaaaaaaat | 8.043 (82.2) | ctacaGaaaaaaat | 4.297 (74.4) | 3.746 | No |
| rs3783555 | 113247324 | SQUA | + | 113247324 | 113247337 | Aaaaaaaatgaaaa | 7.022 (80.1) | Gaaaaaaatgaaaa | 4.638 (75.1) | 2.384 | No |
| rs3783555 | 113247324 | TBP | + | 113247319 | 113247333 | ctacaAaaaaaaatg | 6.787 (82.2) | ctacaGaaaaaaatg | 0.997 (70.7) | 5.79 | No |
| rs1516790 | 113247522 | HLF | + | 113247511 | 113247522 | GTATGTGTAATA | 3.001 (70.3) | CATTACACATAC | 7.247 (81.4) | -4.246 | No |
| rs1516790 | 113247522 | TEF-1 | - | 113247513 | 113247524 | CATATTACACAT | 5.448 (72.9) | CACATTACACAT | 9.341 (82.8) | -3.893 | No |
| rs3783592 | 113247718 | Athb-1 | + | 113247715 | 113247722 | ACATTATT | 7.756 (84.1) | AAACTATG | 3.770 (71.5) | 3.986 | No |
| rs3783592 | 113247718 | HFH-1 | + | 113247716 | 113247726 | CATTATTCATT | 6.517 (76.9) | CATAGTTTATT | 12.868 (93.2) | -6.351 | No |
| rs3783592 | 113247718 | HFH-2 | - | 113247712 | 113247723 | GAATAATGTGTC | 2.298 (71.5) | CATAGTTTATTC | 7.928 (82.8) | -5.63 | No |
| rs3783592 | 113247718 | HFH-2 | + | 113247716 | 113247727 | CATTATTCATTG | 3.259 (73.5) | CATAGTTTATTC | 7.928 (82.8) | -4.669 | No |
| rs3783592 | 113247718 | Sox-5 | + | 113247713 | 113247719 | ACACATT | 2.413 (68.4) | AAACTAT | 6.563 (83.8) | -4.15 | No |
| rs3783592 | 113247718 | Staf | + | 113247709 | 113247728 | TCTGACACATTATTCATTGC | 10.260 (80.0) | TCTGACACATAGTTTATTCA | 4.843 (70.8) | 5.417 | No |
| rs3783591 | 113247745 | AGL3 | + | 113247736 | 113247745 | CTAAATAAAA | 4.306 (76.7) | CTAAATAAAG | 9.505 (88.8) | -5.199 | No |
| rs3783591 | 113247745 | AGL3 | - | 113247736 | 113247745 | TTTTATTTAG | 3.948 (75.9) | CTAAATAAAG | 9.505 (88.8) | -5.557 | No |
| rs3783591 | 113247745 | HFH-3 | - | 113247734 | 113247745 | TTTTATTTAGTT | 11.465 (91.3) | CTTTATTTAGTT | 9.214 (85.8) | 2.251 | No |
| rs3783591 | 113247745 | HNF-3beta | - | 113247734 | 113247745 | TTTTATTTAGTT | 6.836 (79.9) | CTTTATTTAGTT | 8.820 (85.0) | -1.984 | No |
| rs3783591 | 113247745 | MEF2 | + | 113247736 | 113247745 | CTAAATAAAA | 4.372 (73.6) | CTTTATTTAG | 9.128 (83.5) | -4.756 | No |
| rs3783591 | 113247745 | MEF2 | - | 113247736 | 113247745 | TTTTATTTAG | 6.584 (78.2) | CTTTATTTAG | 9.128 (83.5) | -2.544 | No |
| rs3783553 | 113247947 | Chop-cEBP | - | 113247938 | 113247949 | AATTGAAACAAG | 3.701 (71.3) | TCATTCAATTCC | 7.325 (80.4) | -3.624 | Yes |
| rs3783553 | 113247947 | Chop-cEBP | - | 113247944 | 113247955 | AGGTGGAATTGA | 3.109 (69.9) | TCATTCAATTCC | 7.325 (80.4) | -4.216 | Yes |

**Table 2.** Contd.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs16347 | 113247951 | Chop-cEBP | - | 113247944 | 113247955 | AGGTGGAATTGA | 3.109 (69.9) | TCATTCAATTCC | 7.325 (80.4) | -4.216 | Yes |
| rs3783553 | 113247947 | Irf-1 | - | 113247940 | 113247951 | GGAATTGAAACA | 8.779 (81.1) | GGAATTGAATGA | 3.552 (69.4) | 5.227 | Yes |
| rs16347 | 113247951 | Irf-1 | - | 113247940 | 113247951 | GGAATTGAAACA | 8.779 (81.1) | GGAATTGAATGA | 3.552 (69.4) | 5.227 | Yes |
| rs3783553 | 113247947 | Chop-cEBP | + | 113247944 | 113247955 | AATTGAAACAAG | 3.701 (71.3) | TCATTCAATTCC | 7.325 (80.4) | -3.624 | Yes |
| rs16347 | 113247951 | Chop-cEBP | + | 113247944 | 113247955 | AATTGAAACAAG | 3.701 (71.3) | TCATTCAATTCC | 7.325 (80.4) | -3.624 | Yes |
| rs16347 | 113247951 | Hunchback | + | 113247949 | 113247958 | AAACAAGAAT | 3.712 (71.9) | GAAACAAGAA | 6.484 (80.5) | -2.772 | No |
| rs3783590 | 113247956 | Hunchback | + | 113247949 | 113247958 | AAACAAGAAT | 3.712 (71.9) | GAAACAAGAA | 6.484 (80.5) | -2.772 | No |
| rs3783553 | 113247947 | Irf-1 | + | 113247942 | 113247953 | GGAATTGAAACA | 8.779 (81.1) | GGAATTGAATGA | 3.552 (69.4) | 5.227 | Yes |
| rs16347 | 113247951 | Irf-1 | + | 113247942 | 113247953 | GGAATTGAAACA | 8.779 (81.1) | GGAATTGAATGA | 3.552 (69.4) | 5.227 | Yes |
| rs3783553 | 113247947 | Sox-5 | - | 113247945 | 113247951 | TTTCAAT | 3.288 (71.6) | AAACAAG | 6.563 (83.8) | -3.275 | Yes |
| rs16347 | 113247951 | Sox-5 | - | 113247945 | 113247951 | TTTCAAT | 3.288 (71.6) | AAACAAG | 6.563 (83.8) | -3.275 | Yes |
| rs3783590 | 113247956 | HFH-2 | + | 113247952 | 113247963 | GCATACTTGTTT | 7.259 (81.4) | GCATTCTTGTTT | 4.922 (76.8) | 2.337 | No |
| rs3783590 | 113247956 | HFH-3 | + | 113247952 | 113247963 | GCATACTTGTTT | 9.604 (86.8) | GCATTCTTGTTT | 6.234 (78.6) | 3.37 | No |
| rs3783590 | 113247956 | HNF-3beta | + | 113247952 | 113247963 | GCATACTTGTTT | 8.731 (84.8) | GCATTCTTGTTT | 5.186 (75.7) | 3.545 | No |
| rs3783590 | 113247956 | Hunchback | - | 113247955 | 113247964 | GAAACAAGTA | 2.484 (68.1) | GAAACAAGAA | 6.484 (80.5) | -4 | No |
| rs3783590 | 113247956 | Sox-5 | - | 113247954 | 113247960 | CAAGTAT | 1.465 (64.8) | CAAGAAT | 5.801 (81.0) | -4.336 | No |
| rs16347 | 113247951 | c-REL | + | 113247949 | 113247958 | TGGGCATACT | 4.576 (74.3) | TGGGCATTCT | 6.979 (81.7) | -2.403 | No |
| rs3783590 | 113247956 | c-REL | + | 113247949 | 113247958 | TGGGCATACT | 4.576 (74.3) | TGGGCATTCT | 6.979 (81.7) | -2.403 | No |
| rs3783589 | 113248186 | Brachyury | + | 113248184 | 113248194 | CTAAGTGTGAC | 11.785 (85.5) | CTGAGTGTGAC | 7.068 (75.8) | 4.717 | No |
| **IL-1R (Interleukin-1 receptor)** | | | | | | | | | | | |
| rs6758647 | 102215391 | CFI-USP | + | 102215387 | 102215396 | GGGGACAGGG | 8.380 (81.3) | GGGGTCAGGG | 13.061 (93.4) | -4.681 | No |
| rs4141633 | 102215556 | FREAC-4 | - | 102215555 | 102215562 | GTAAGCGG | 3.949 (71.2) | GTAAGCAG | 8.188 (85.5) | -4.239 | No |
| rs4141632 | 102215657 | E4BP4 | - | 102215653 | 102215663 | TTACGGGAGCC | 3.155 (69.4) | TTACGGAAGCC | 7.490 (80.5) | -4.335 | No |
| rs4141632 | 102215657 | E4BP4 | + | 102215656 | 102215666 | TCCCGTAAGGT | 4.597 (73.1) | TTCCGTAAGGT | 8.808 (83.8) | -4.211 | No |
| rs4141632 | 102215657 | E74A | - | 102215655 | 102215661 | ACGGGAG | 4.523 (75.4) | ACGGAAG | 8.651 (91.7) | -4.128 | No |
| rs4141632 | 102215657 | c-REL | + | 102215650 | 102215659 | GGAGGCTCCC | 2.778 (68.8) | GGAGGCTTCC | 6.737 (80.9) | -3.959 | No |
| rs4141632 | 102215657 | c-REL | - | 102215650 | 102215659 | GGGAGCCTCC | 4.938 (75.4) | GGAGGCTTCC | 6.737 (80.9) | -1.799 | No |
| rs4141631 | 102215734 | bZIP910 | + | 102215728 | 102215734 | CTGACTC | 2.446 (70.4) | CTGACTT | 7.070 (85.2) | -4.624 | No |
| rs4141631 | 102215734 | c-FOS | + | 102215728 | 102215735 | CTGACTCT | 5.981 (80.9) | CTGACTTT | 2.288 (68.6) | 3.693 | No |
| rs4141631 | 102215734 | c-FOS | - | 102215729 | 102215736 | GAGAGTCA | 7.560 (86.1) | GAAAGTCA | 3.429 (72.4) | 4.131 | No |
| rs6708048 | 102216173 | HMG-IY | + | 102216169 | 102216184 | ACAGAAATGGCATCAC | 9.430 (84.3) | ACAGTAATGGCATCAC | 5.757 (75.9) | 3.673 | No |
| rs6708048 | 102216173 | SOX17 | - | 102216168 | 102216176 | ATTTCTGTG | 2.332 (69.9) | ATTACTGTG | 6.046 (81.9) | -3.714 | No |
| rs6708048 | 102216173 | SQUA | + | 102216169 | 102216182 | ACAGAAATGGCATC | 9.353 (84.9) | ACAGTAATGGCATC | 7.811 (81.7) | 1.542 | No |
| rs6708048 | 102216173 | ATHB5 | - | 102216170 | 102216178 | CCATTTCTG | 3.684 (71.4) | CCATTACTG | 8.346 (83.1) | -4.662 | No |
| rs2871448 | 102216405 | AGL3 | + | 102216403 | 102216412 | CTCTTAATGC | 2.497 (72.5) | GCATTAAAAG | 5.769 (80.1) | -3.272 | No |
| rs2871448 | 102216405 | Athb-1 | + | 102216398 | 102216405 | GAATTCTC | 2.753 (68.3) | GAATTCTT | 7.031 (81.8) | -4.278 | No |
| rs2871448 | 102216405 | Hunchback | - | 102216403 | 102216412 | GCATTAAGAG | 6.542 (80.6) | GCATTAAAAG | 8.864 (87.8) | -2.322 | No |
| rs2871448 | 102216405 | NF-kappaB | - | 102216396 | 102216405 | GAGAATTCCT | 6.814 (80.1) | AGGAATTCTT | 2.968 (70.7) | 3.846 | No |
| rs2871448 | 102216405 | SQUA | - | 102216399 | 102216412 | GCATTAAGAGAATT | 6.588 (79.2) | GCATTAAAAGAATT | 8.571 (83.2) | -1.983 | No |
| rs2871448 | 102216405 | Sox-5 | - | 102216400 | 102216406 | AGAGAAT | 3.288 (71.6) | AAAGAAT | 6.624 (84.1) | -3.336 | No |

**Table 2.** Contd.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs2871448 | 102216405 | TBP | - | 102216398 | 102216412 | GCATTAAGAGAATTC | 3.815 (76.3) | GCATTAAAAGAATTC | 6.029 (80.7) | -2.214 | No |
| rs956730 | 102216634 | HMG-IY | - | 102216620 | 102216635 | TTAGAAACTCAAGAAT | 8.383 (81.9) | TCAGAAACTCAAGAAT | 5.984 (76.4) | 2.399 | No |
| rs2234650 | 102216845 | Chop-cEBP | - | 102216841 | 102216852 | GTCTGCAGCTCC | 3.949 (71.9) | GTCTGCAACTCC | 7.824 (81.6) | -3.875 | No |
| rs2234650 | 102216845 | Myf | + | 102216839 | 102216850 | AGGGAGCTGCAG | 11.907 (90.6) | AGGGAGTTGCAG | 7.907 (81.5) | 4 | No |
| rs2234650 | 102216845 | Thing1-E47 | - | 102216844 | 102216853 | AGTCTGCAGC | 4.827 (76.6) | AGTCTGCAAC | 6.824 (82.5) | -1.997 | No |
| rs2234650 | 102216845 | Spz1 | + | 102216841 | 102216851 | GGAGCTGCAGA | 4.042 (72.2) | GGAGTTGCAGA | 7.691 (81.7) | -3.649 | No |
| rs2234652 | 102217009 | Broad-complex_4 | - | 102217004 | 102217014 | TTTTAGActta | 3.555 (69.8) | TTTTAAActta | 6.987 (80.8) | -3.432 | No |
| rs2234652 | 102217009 | COUP-TF | - | 102216998 | 102217011 | TAGActtaaaaccc | 6.515 (75.3) | TAAActtaaaaccc | 9.278 (81.2) | -2.763 | No |
| rs2234652 | 102217009 | Hunchback | - | 102217000 | 102217009 | GActtaaaac | 7.180 (82.6) | AActtaaaac | 5.010 (75.9) | 2.17 | No |
| rs2234652 | 102217009 | Hunchback | + | 102217007 | 102217016 | gTCTAAAAAT | 7.010 (82.1) | gTTTAAAAAT | 5.010 (75.9) | 2 | No |
| rs10168222 | 102217186 | Broad-complex_1 | - | 102217176 | 102217189 | tttTaaaaaaaatg | 6.580 (76.5) | tttAaaaaaaaatg | 8.727 (82.2) | -2.147 | No |
| rs10168222 | 102217186 | Broad-complex_4 | - | 102217178 | 102217188 | ttTaaaaaaaa | 6.291 (78.6) | ttAaaaaaaaa | 8.384 (85.3) | -2.093 | No |
| rs10168222 | 102217186 | Broad-complex_4 | - | 102217179 | 102217189 | tttTaaaaaaa | 8.091 (84.3) | tttAaaaaaaa | 6.291 (78.6) | 1.8 | No |
| rs10168222 | 102217186 | Broad-complex_4 | - | 102217180 | 102217190 | ttttTaaaaaa | 4.894 (74.1) | ttttAaaaaaa | 8.091 (84.3) | -3.197 | No |
| rs10168222 | 102217186 | Broad-complex_4 | + | 102217182 | 102217192 | ttttAaaaaac | 6.927 (80.6) | ttttTaaaaac | 3.731 (70.4) | 3.196 | No |
| rs10168222 | 102217186 | HFH-2 | + | 102217175 | 102217186 | tcattttttttA | 2.932 (72.8) | tcattttttttT | 7.353 (81.6) | -4.421 | No |
| rs10168222 | 102217186 | Hunchback | - | 102217178 | 102217187 | tTaaaaaaaa | 7.221 (82.7) | tAaaaaaaaa | 10.028 (91.4) | -2.807 | No |
| rs10168222 | 102217186 | Hunchback | - | 102217179 | 102217188 | ttTaaaaaaa | 3.899 (72.5) | ttAaaaaaaa | 7.221 (82.7) | -3.322 | No |
| rs10168222 | 102217186 | TBP | - | 102217175 | 102217189 | tttTaaaaaaaatga | 5.858 (80.4) | tttAaaaaaaaatga | 0.280 (69.3) | 5.578 | No |
| rs10168222 | 102217186 | TBP | - | 102217176 | 102217190 | ttttTaaaaaaaatg | 2.871 (74.4) | ttttAaaaaaaaatg | 6.232 (81.1) | -3.361 | No |
| rs10168222 | 102217186 | TBP | + | 102217182 | 102217196 | ttttAaaaaactgtt | 6.090 (80.8) | ttttTaaaaactgtt | 2.729 (74.2) | 3.361 | No |
| rs1800919 | 102217811 | E74A | + | 102217805 | 102217811 | ccggagA | 5.756 (80.2) | ccggagC | 4.185 (74.0) | 1.571 | No |
| **IL-1RN (Interleukin-1 receptor antagonist)** | | | | | | | | | | | |
| rs315920 | 113589249 | NRF-2 | - | 113589246 | 113589255 | AATGGAGGAG | 5.186 (72.3) | AATGGAAGAG | 8.720 (83.6) | -3.534 | No |
| rs4251954 | 113589830 | AML-1 | - | 113589825 | 113589833 | catGtggtt | 9.467 (91.1) | catAtggtt | 4.823 (77.9) | 4.644 | No |
| rs4251954 | 113589830 | ARNT | + | 113589828 | 113589833 | caCatg | 6.112 (83.1) | caTatg | 1.872 (66.2) | 4.24 | No |
| rs4251954 | 113589830 | ARNT | - | 113589828 | 113589833 | catGtg | 6.112 (83.1) | caTatg | 1.872 (66.2) | 4.24 | No |
| rs4251954 | 113589830 | Max | + | 113589825 | 113589834 | aaccaCatgc | 9.213 (88.2) | aaccaTatgc | 5.084 (75.6) | 4.129 | No |
| rs4251954 | 113589830 | Myc-Max | - | 113589826 | 113589836 | tggcatGtggt | 9.949 (85.0) | tggcatAtggt | 6.648 (76.7) | 3.301 | No |
| rs4251954 | 113589830 | Snail | - | 113589828 | 113589833 | catGtg | 6.136 (83.5) | caTatg | 2.669 (71.1) | 3.467 | No |
| rs4251954 | 113589830 | Sox-5 | + | 113589824 | 113589830 | caaccaC | 1.405 (64.6) | caaccaT | 5.740 (80.8) | -4.335 | No |
| rs4251954 | 113589830 | Tal1beta-E47S | + | 113589824 | 113589835 | caaccaCatgcc | 5.355 (76.1) | caaccaTatgcc | 9.018 (84.1) | -3.663 | No |
| rs4251954 | 113589830 | USF | - | 113589827 | 113589833 | catGtgg | 8.367 (89.2) | catAtgg | 4.687 (76.6) | 3.68 | No |
| rs4251954 | 113589830 | n-MYC | + | 113589828 | 113589833 | caCatg | 7.418 (88.4) | caTatg | 3.703 (74.5) | 3.715 | No |
| rs4251954 | 113589830 | n-MYC | - | 113589828 | 113589833 | catGtg | 6.819 (86.2) | caTatg | 3.703 (74.5) | 3.116 | No |
| rs4251955 | 113590102 | HFH-2 | - | 113590100 | 113590111 | AAATGTCAATTT | 7.780 (82.5) | AAATGTCAACTT | 5.745 (78.4) | 2.035 | No |
| rs4251955 | 113590102 | SQUA | + | 113590097 | 113590110 | ACAAAATTGACATT | 7.437 (80.9) | ACAAAGTTGACATT | 3.691 (73.2) | 3.746 | No |
| rs4251956 | 113590240 | E74A | + | 113590239 | 113590245 | GCGGAAG | 8.651 (91.7) | GTGGAAG | 4.991 (77.2) | 3.66 | No |
| rs4251956 | 113590240 | NRF-2 | + | 113590238 | 113590247 | GGCGGAAGAG | 10.283 (88.6) | GGTGGAAGAG | 6.951 (77.9) | 3.332 | No |
| rs4251956 | 113590240 | Snail | + | 113590236 | 113590241 | AAGGCG | 1.346 (66.4) | AAGGTG | 6.063 (83.3) | -4.717 | No |

**Table 2.** Contd.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs4251958 | 113590307 | RREB-1 | - | 113590298 | 113590317 | CCCCCCACCCGCCGCCCCGC | 12.018 (80.6) | CCCCCCACCCCCCGCCCCGC | 15.257 (85.6) | -3.239 | No |
| rs4251958 | 113590307 | RREB-1 | - | 113590301 | 113590320 | AAGCCCCCCACCCGCCGCCC | 10.062 (77.6) | AAGCCCCCCACCCCCCGCCC | 11.963 (80.5) | -1.901 | No |
| rs4251958 | 113590307 | Snail | + | 113590307 | 113590312 | CGGGTG | 6.063 (83.3) | GGGGTG | 2.088 (69.1) | 3.975 | No |
| rs4251959 | 113590353 | AGL3 | + | 113590351 | 113590360 | CCAATTCAGG | 5.912 (80.4) | CCCATTCAGG | 2.256 (71.9) | 3.656 | No |
| rs4251959 | 113590353 | AGL3 | - | 113590351 | 113590360 | CCTGAATTGG | 6.094 (80.9) | CCCATTCAGG | 2.256 (71.9) | 3.838 | No |
| rs4251959 | 113590353 | AML-1 | - | 113590347 | 113590355 | ATTGGGGTG | 5.689 (80.4) | ATGGGGGTG | 1.203 (67.6) | 4.486 | No |
| rs4251959 | 113590353 | RREB-1 | + | 113590339 | 113590358 | GCCCCTCCCACCCCAATTCA | 9.994 (77.5) | GCCCCTCCCACCCCCATTCA | 11.791 (80.2) | -1.797 | No |
| rs4251960 | 113590408 | Thing1-E47 | + | 113590400 | 113590409 | GGTCTGCAAG | 7.910 (85.7) | GGTCTGCAGG | 5.913 (79.8) | 1.997 | No |
| rs4251960 | 113590408 | Spz1 | + | 113590407 | 113590417 | AAGGTTAGAGG | 6.932 (79.8) | AGGGTTAGAGG | 10.442 (88.9) | -3.51 | No |
| rs4251961 | 113590698 | SAP-1 | - | 113590695 | 113590703 | ACCTGGTGC | 3.301 (68.6) | ACCTGATGC | 7.540 (80.5) | -4.239 | No |
| rs4251961 | 113590698 | Snail | - | 113590696 | 113590701 | CTGGTG | 6.770 (85.8) | CTGATG | 3.303 (73.4) | 3.467 | No |
| rs4251962 | 113590873 | FREAC-4 | - | 113590869 | 113590876 | GCAGACAA | 3.862 (70.9) | GCAAACAA | 8.102 (85.2) | -4.24 | No |
| rs4251962 | 113590873 | HFH-2 | + | 113590867 | 113590878 | AGTTGTCTGCTC | 2.857 (72.7) | AGTTGTTTGCTC | 7.656 (82.2) | -4.799 | No |
| rs4251962 | 113590873 | HNF-3beta | + | 113590867 | 113590878 | AGTTGTCTGCTC | 6.227 (78.4) | AGTTGTTTGCTC | 9.295 (86.2) | -3.068 | No |
| rs4251962 | 113590873 | Hunchback | - | 113590869 | 113590878 | GAGCAGACAA | 2.297 (67.6) | GAGCAAACAA | 6.384 (80.2) | -4.087 | No |
| rs4251962 | 113590873 | Sox-5 | - | 113590868 | 113590874 | AGACAAC | 3.227 (71.4) | AAACAAC | 6.563 (83.8) | -3.336 | No |
| rs4251963 | 113590924 | HMG-IY | + | 113590922 | 113590937 | CTACCAACGGCAAAAA | 9.067 (83.5) | CTCCCAACGGCAAAAA | 6.707 (78.1) | 2.36 | No |
| rs4251964 | 113590925 | HMG-IY | + | 113590922 | 113590937 | CTACCAACGGCAAAAA | 9.067 (83.5) | CTCCCAACGGCAAAAA | 6.707 (78.1) | 2.36 | No |
| rs4251965 | 113590926 | HMG-IY | + | 113590922 | 113590937 | CTACCAACGGCAAAAA | 9.067 (83.5) | CTCCCAACGGCAAAAA | 6.707 (78.1) | 2.36 | No |
| rs4251966 | 113591023 | AGL3 | + | 113591023 | 113591032 | CCAGGTATAG | 6.967 (82.9) | TCAGGTATAG | 2.533 (72.6) | 4.434 | No |
| rs4251966 | 113591023 | CFI-USP | + | 113591019 | 113591028 | GAGGCCAGGT | 3.277 (68.1) | GAGGTCAGGT | 7.958 (80.2) | -4.681 | No |
| rs4251966 | 113591023 | COUP-TF | - | 113591012 | 113591025 | TGGCCTCTGCCGTG | 8.503 (79.6) | TGACCTCTGCCGTG | 11.265 (85.5) | -2.762 | No |
| rs4251966 | 113591023 | RORalfa-1 | + | 113591016 | 113591025 | GCAGAGGCCA | 2.958 (68.5) | GCAGAGGTCA | 7.350 (80.6) | -4.392 | No |
| rs4251966 | 113591023 | Thing1-E47 | + | 113591021 | 113591030 | GGCCAGGTAT | 4.087 (74.4) | GGTCAGGTAT | 7.008 (83.0) | -2.921 | No |
| rs4251966 | 113591023 | bZIP910 | - | 113591020 | 113591026 | CTGGCCT | 2.446 (70.4) | CTGACCT | 7.070 (85.2) | -4.624 | No |
| rs4251967 | 113591199 | NF-kappaB | + | 113591198 | 113591207 | aCggcttccc | 4.172 (73.7) | aGggcttccc | 8.853 (85.0) | -4.681 | No |
| rs11677397 | 113591312 | Broad-complex_1 | - | 113591306 | 113591319 | caagttaGcaaata | 6.524 (76.4) | caagttaAcaaata | 8.671 (82.0) | -2.147 | No |
| rs11677397 | 113591312 | Broad-complex_4 | - | 113591308 | 113591318 | aagttaGcaaa | 6.476 (79.2) | aagttaAcaaa | 8.275 (84.9) | -1.799 | No |
| rs11677397 | 113591312 | FREAC-4 | - | 113591309 | 113591316 | gttaGcaa | 4.998 (74.7) | gttaAcaa | 8.172 (85.4) | -3.174 | No |
| rs11677397 | 113591312 | HFH-2 | + | 113591303 | 113591314 | ttttatttgCta | 6.016 (79.0) | ttttatttgTta | 8.051 (83.0) | -2.035 | No |
| rs11677397 | 113591312 | HFH-2 | + | 113591307 | 113591318 | atttgCtaactt | 4.674 (76.3) | atttgTtaactt | 9.503 (85.9) | -4.829 | No |
| rs11677397 | 113591312 | HFH-3 | + | 113591303 | 113591314 | ttttatttgCta | 6.055 (78.2) | ttttatttgTta | 10.244 (88.3) | -4.189 | No |
| rs11677397 | 113591312 | HMG-IY | - | 113591299 | 113591314 | taGcaaataaaataac | 9.882 (85.4) | taAcaaataaaataac | 12.242 (90.8) | -2.36 | No |
| rs11677397 | 113591312 | HMG-IY | - | 113591307 | 113591322 | aaacaagttaGcaaat | 5.794 (76.0) | aaacaagttaAcaaat | 8.001 (81.1) | -2.207 | No |
| rs11677397 | 113591312 | HNF-1 | - | 113591304 | 113591317 | agttaGcaaataaa | 6.631 (74.7) | agttaAcaaataaa | 9.629 (80.8) | -2.998 | No |
| rs11677397 | 113591312 | HNF-3beta | + | 113591307 | 113591318 | atttgCtaactt | 4.523 (74.0) | atttgTtaactt | 7.591 (81.9) | -3.068 | No |
| rs11677397 | 113591312 | Hunchback | - | 113591303 | 113591312 | Gcaaataaaa | 7.888 (84.8) | Acaaataaaa | 5.718 (78.1) | 2.17 | No |
| rs11677397 | 113591312 | SQUA | - | 113591299 | 113591312 | Gcaaataaaataac | 7.405 (80.8) | Acaaataaaataac | 9.789 (85.8) | -2.384 | No |
| rs11677397 | 113591312 | Sox-5 | - | 113591308 | 113591314 | taGcaaa | 3.164 (71.2) | taAcaaa | 6.563 (83.8) | -3.399 | No |
| rs4251968 | 113591448 | TEF-1 | + | 113591441 | 113591452 | CACATGCATGAG | 6.924 (76.7) | CACATGCCTGAG | 10.701 (86.3) | -3.777 | No |
| rs4251968 | 113591448 | TEF-1 | - | 113591441 | 113591452 | CTCATGCATGTG | 3.846 (68.8) | CACATGCCTGAG | 10.701 (86.3) | -6.855 | No |

**Table 2.** Contd.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| rs2234678 | 113591796 | AML-1 | - | 113591790 | 113591798 | TCTGAGGTC | 7.747 (86.2) | TCCGAGGTC | 4.461 (76.9) | 3.286 | No |
| rs2234678 | 113591796 | Dorsal_2 | - | 113591796 | 113591805 | GAGGTCTTCT | 5.445 (74.2) | GAGGTCTTCC | 8.989 (84.5) | -3.544 | No |
| rs2234678 | 113591796 | E74A | + | 113591794 | 113591800 | TCAGAAG | 3.945 (73.1) | TCGGAAG | 8.074 (89.4) | -4.129 | No |
| rs2234678 | 113591796 | NRF-2 | + | 113591793 | 113591802 | CTCAGAAGAC | 4.612 (70.4) | CTCGGAAGAC | 8.146 (81.8) | -3.534 | No |
| rs2234678 | 113591796 | bZIP910 | - | 113591791 | 113591797 | CTGAGGT | 7.070 (85.2) | CCGAGGT | 2.446 (70.4) | 4.624 | No |
| rs4251970 | 113592004 | SU_h | + | 113592002 | 113592017 | CCATGGGAGACCATGC | 5.737 (71.1) | CCGTGGGAGACCATGC | 9.508 (80.1) | -3.771 | No |
| rs4252038 | 113592123 | Dorsal_1 | + | 113592122 | 113592133 | GAGTCATTTTCC | 4.151 (71.4) | GGGTCATTTTCC | 7.990 (81.0) | -3.839 | No |
| rs4252038 | 113592123 | RORalfa-1 | + | 113592118 | 113592127 | GTGAGAGTCA | 4.232 (72.0) | GTGAGGGTCA | 8.624 (84.1) | -4.392 | No |
| rs4252038 | 113592123 | c-FOS | + | 113592120 | 113592127 | GAGAGTCA | 7.560 (86.1) | GAGGGTCA | 3.429 (72.4) | 4.131 | No |
| rs4252038 | 113592123 | c-FOS | - | 113592121 | 113592128 | ATGACTCT | 6.960 (84.1) | ATGACCCT | 3.846 (73.8) | 3.114 | No |
| **CR1 (Complement receptor-1)** | | | | | | | | | | | |
| rs11807805 | 204054992 | AGL3 | + | 204054986 | 204054995 | cgatatCtgg | 4.995 (78.3) | ccaCatatcg | 7.404 (83.9) | -2.409 | No |
| rs11807805 | 204054992 | AML-1 | + | 204054989 | 204054997 | tatCtgggc | 1.600 (68.8) | tatGtgggc | 6.245 (81.9) | -4.645 | No |
| rs11807805 | 204054992 | SU_h | + | 204054990 | 204055005 | atCtgggcaaagagca | 5.679 (71.0) | atGtgggcaaagagca | 9.449 (80.0) | -3.77 | No |
| rs9429780 | 204055585 | AML-1 | - | 204055583 | 204055591 | tttgctGtc | 6.955 (83.9) | tttgctCtc | 2.273 (70.7) | 4.682 | No |
| rs9429781 | 204055617 | E74A | + | 204055614 | 204055620 | cagGaat | 7.203 (86.0) | cagTaat | 3.074 (69.6) | 4.129 | No |
| rs9429781 | 204055617 | HFH-2 | + | 204055617 | 204055628 | Gaatgttgtttt | 14.310 (95.5) | Taatgttgtttt | 12.640 (92.2) | 1.67 | No |
| rs9429781 | 204055617 | TEF-1 | - | 204055612 | 204055623 | aacattCctggt | 11.517 (88.4) | aacattActggt | 7.624 (78.5) | 3.893 | No |
| rs9429941 | 204055698 | Broad-complex_4 | + | 204055692 | 204055702 | ttataaAgaca | 6.987 (80.8) | ttataaTgaca | 3.791 (70.6) | 3.196 | No |
| rs9429941 | 204055698 | TBP | - | 204055685 | 204055699 | cTttataatcaccat | 6.337 (81.3) | cAttataatcaccat | 2.430 (73.6) | 3.907 | No |
| rs9429941 | 204055698 | TBP | + | 204055692 | 204055706 | ttataaAgacagagt | 8.906 (86.4) | ttataaTgacagagt | 3.845 (76.4) | 5.061 | No |
| rs12078329 | 204055892 | Athb-1 | + | 204055886 | 204055893 | aaatcaCt | 4.960 (75.3) | aaatcaTt | 9.297 (88.9) | -4.337 | No |
| rs12078329 | 204055892 | Athb-1 | - | 204055887 | 204055894 | caGtgatt | 4.328 (73.3) | caAtgatt | 8.720 (87.1) | -4.392 | No |
| rs12078329 | 204055892 | SOX17 | + | 204055888 | 204055896 | atcaCtgag | 2.498 (70.4) | atcaTtgag | 6.212 (82.4) | -3.714 | No |
| rs12078329 | 204055892 | SOX17 | - | 204055888 | 204055896 | ctcaGtgat | 0.946 (65.4) | atcaTtgag | 6.212 (82.4) | -5.266 | No |
| rs12078329 | 204055892 | ATHB5 | + | 204055886 | 204055894 | aaatcaCtg | 2.881 (69.3) | caAtgattt | 10.084 (87.4) | -7.203 | No |
| rs12078329 | 204055892 | ATHB5 | - | 204055886 | 204055894 | caGtgattt | 5.421 (75.7) | caAtgattt | 10.084 (87.4) | -4.663 | No |
| rs11117911 | 204056380 | E74A | + | 204056377 | 204056383 | GAGAAAG | 3.411 (71.0) | GAGGAAG | 7.540 (87.3) | -4.129 | No |
| rs11117913 | 204056630 | AGL3 | + | 204056625 | 204056634 | CCATACTTCT | 2.315 (72.1) | CCATATTTCT | 5.786 (80.1) | -3.471 | No |
| rs11117913 | 204056630 | Broad-complex_4 | - | 204056629 | 204056639 | AAATAAGAAGT | 4.024 (71.3) | AAATAAGAAAT | 7.220 (81.5) | -3.196 | No |
| rs11117913 | 204056630 | SQUA | - | 204056629 | 204056642 | CTGAAATAAGAAGT | 7.450 (80.9) | CTGAAATAAGAAAT | 9.806 (85.8) | -2.356 | No |
| rs7525160 | 204056809 | Dorsal_1 | - | 204056808 | 204056819 | GGGTTTTTCTGT | 6.163 (76.4) | GGGTTTTTCTCT | 8.543 (82.4) | -2.38 | No |
| rs7525160 | 204056809 | Hunchback | + | 204056808 | 204056817 | ACAGAAAAAC | 6.747 (81.3) | AGAGAAAAAC | 5.162 (76.4) | 1.585 | No |
| rs7525160 | 204056809 | MEF2 | + | 204056800 | 204056809 | TTATTTTAAC | 5.099 (75.1) | TTATTTTAAG | 9.868 (85.0) | -4.769 | No |
| rs12069288 | 204056859 | COUP-TF | - | 204056853 | 204056866 | TGTCCTTTTGACAT | 8.844 (80.3) | TGTCCTTCTGACAT | 5.005 (72.1) | 3.839 | No |
| rs12069288 | 204056859 | Thing1-E47 | - | 204056853 | 204056862 | CTTTTGACAT | 2.949 (71.0) | CTTCTGACAT | 6.593 (81.8) | -3.644 | No |
| rs12069288 | 204056859 | bZIP910 | - | 204056853 | 204056859 | TTGACAT | 3.592 (74.1) | CTGACAT | 7.070 (85.2) | -3.478 | No |
| rs9429942 | 204057025 | CF2-II | + | 204057018 | 204057027 | acatatgCat | 6.294 (80.4) | atAcatatgt | 9.387 (88.1) | -3.093 | No |
| rs9429942 | 204057025 | CF2-II | - | 204057018 | 204057027 | atGcatatgt | 5.775 (79.1) | atAcatatgt | 9.387 (88.1) | -3.612 | No |
| rs9429942 | 204057025 | CF2-II | + | 204057020 | 204057029 | atatgCatac | 4.797 (76.7) | gtatAcatat | 10.344 (90.4) | -5.547 | No |

**Table 2.** Contd.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs9429942 | 204057025 | CF2-II | - | 204057020 | 204057029 | gtatGcatat | 5.179 (77.6) | gtatAcatat | 10.344 (90.4) | -5.165 | No |
| rs9429942 | 204057025 | CF2-II | + | 204057022 | 204057031 | atgCatacat | 4.443 (75.8) | atgTatacat | 7.067 (82.3) | -2.624 | No |
| rs9429942 | 204057025 | CF2-II | + | 204057024 | 204057033 | gCatacatgt | 4.824 (76.8) | gTatacatgt | 9.031 (87.2) | -4.207 | No |
| rs9429942 | 204057025 | CF2-II | - | 204057024 | 204057033 | acatgtatGc | 2.889 (72.0) | gTatacatgt | 9.031 (87.2) | -6.142 | No |
| rs9429942 | 204057025 | FREAC-4 | - | 204057022 | 204057029 | gtatGcat | 5.095 (75.1) | gtatAcat | 8.269 (85.7) | -3.174 | No |
| rs9429942 | 204057025 | FREAC-4 | + | 204057024 | 204057031 | gCatacat | 4.030 (71.5) | gTatacat | 8.269 (85.7) | -4.239 | No |
| rs9429942 | 204057025 | HLF | - | 204057014 | 204057025 | Gcatatgtaaca | 3.516 (71.7) | tgttacatatgT | 7.823 (82.9) | -4.307 | No |
| rs9429942 | 204057025 | TBP | + | 204057024 | 204057038 | gCatacatgtgcgac | 3.296 (75.3) | gTatacatgtgcgac | 5.920 (80.5) | -2.624 | No |
| rs2296878 | 204058097 | AML-1 | + | 204058095 | 204058103 | GTCGTGGTG | 3.694 (74.7) | GTTGTGGTG | 6.980 (84.0) | -3.286 | No |
| rs4097396 | 204058347 | ARNT | + | 204058345 | 204058350 | CACGTG | 10.351 (100.0) | CAGGTG | 6.112 (83.1) | 4.239 | No |
| rs4097396 | 204058347 | ARNT | - | 204058345 | 204058350 | CACGTG | 10.351 (100.0) | CAGGTG | 6.112 (83.1) | 4.239 | No |
| rs4097396 | 204058347 | Max | + | 204058342 | 204058351 | TCCCACGTGC | 9.232 (88.2) | TCCCAGGTGC | 5.103 (75.6) | 4.129 | No |
| rs4097396 | 204058347 | SU_h | - | 204058334 | 204058349 | ACGTGGGAACTCCACG | 10.820 (83.3) | ACCTGGGAACTCCACG | 7.049 (74.3) | 3.771 | No |
| rs4097396 | 204058347 | Snail | + | 204058345 | 204058350 | CACGTG | 6.136 (83.5) | CAGGTG | 10.744 (100.0) | -4.608 | No |
| rs4097396 | 204058347 | Snail | - | 204058345 | 204058350 | CACGTG | 6.136 (83.5) | CAGGTG | 10.744 (100.0) | -4.608 | No |
| rs4097396 | 204058347 | USF | - | 204058344 | 204058350 | CACGTGG | 11.491 (100.0) | CACCTGG | 7.020 (84.6) | 4.471 | No |
| rs4097396 | 204058347 | USF | + | 204058345 | 204058351 | CACGTGC | 9.302 (92.5) | CAGGTGC | 4.879 (77.2) | 4.423 | No |
| rs4097396 | 204058347 | n-MYC | + | 204058345 | 204058350 | CACGTG | 10.533 (100.0) | CACCTG | 6.914 (86.5) | 3.619 | No |
| rs4097396 | 204058347 | n-MYC | - | 204058345 | 204058350 | CACGTG | 10.533 (100.0) | CACCTG | 6.914 (86.5) | 3.619 | No |
| **TNF-α (Tumor necrosis factor-alpha)** | | | | | | | | | | | |
| rs5875327 | 31648480 | NRF-2 | - | 31648476 | 31648485 | aGagaaagag | 6.514 (76.5) | aCagaaagag | 8.283 (82.2) | -1.769 | Yes |
| rs4647194 | 31648484 | NRF-2 | - | 31648476 | 31648485 | aGagaaagag | 6.514 (76.5) | aCagaaagag | 8.283 (82.2) | -1.769 | Yes |
| rs2857713 | 31648535 | Thing1-E47 | - | 31648528 | 31648537 | TGTGTGGCAC | 6.189 (80.6) | TGCGTGGCAC | 3.268 (72.0) | 2.921 | No |
| rs1041981 | 31648763 | AML-1 | - | 31648762 | 31648770 | TTTGAGGTT | 9.354 (90.7) | TTTGAGGGT | 5.508 (79.8) | 3.846 | No |
| rs4647195 | 31648936 | Chop-cEBP | + | 31648928 | 31648939 | CCCTGCCACCCC | 5.236 (75.2) | CCCTGCCATCCC | 7.428 (80.6) | -2.192 | No |
| rs4647195 | 31648936 | RREB-1 | + | 31648926 | 31648945 | CCCCCTGCCACCCCCCAGGA | 12.495 (81.3) | CCCCCTGCCATCCCCCAGGA | 9.256 (76.3) | 3.239 | No |
| rs4647195 | 31648936 | Spz1 | - | 31648929 | 31648939 | GGGGTGGCAGG | 7.123 (80.2) | GGGATGGCAGG | 3.347 (70.4) | 3.776 | No |
| rs3093544 | 31649758 | Spz1 | + | 31649750 | 31649760 | AGAGGAAGAGC | 8.473 (83.8) | AGAGGAAGGGC | 6.550 (78.8) | 1.923 | No |
| rs4645834 | 31649818 | Broad-complex_4 | + | 31649815 | 31649825 | AAGGCaaaaaa | 3.915 (71.0) | AAGTCaaaaaa | 7.111 (81.2) | -3.196 | Yes |
| rs4645834 | 31649818 | HMG-IY | + | 31649810 | 31649825 | TTATGAAGGCaaaaaa | 7.959 (81.0) | TTATGAAGTCaaaaaa | 5.752 (75.9) | 2.207 | Yes |
| rs4645834 | 31649818 | Hunchback | + | 31649815 | 31649824 | AAGGCaaaaa | 4.314 (73.8) | AAGTCaaaaa | 6.484 (80.5) | -2.17 | Yes |
| rs4645834 | 31649818 | Hunchback | + | 31649816 | 31649825 | AGGCaaaaaa | 6.899 (81.7) | AGTCaaaaaa | 4.577 (74.6) | 2.322 | Yes |
| rs4645834 | 31649818 | Hunchback | + | 31649817 | 31649826 | GGCaaaaaaa | 9.069 (88.4) | GTCaaaaaaa | 7.484 (83.5) | 1.585 | Yes |
| rs4645834 | 31649818 | Hunchback | + | 31649818 | 31649827 | GCaaaaaaat | 10.654 (93.3) | TCaaaaaaat | 9.069 (88.4) | 1.585 | Yes |
| rs3093547 | 31649827 | Hunchback | + | 31649818 | 31649827 | GCaaaaaaat | 10.654 (93.3) | TCaaaaaaat | 9.069 (88.4) | 1.585 | Yes |
| rs4645834 | 31649818 | bZIP910 | + | 31649812 | 31649818 | ATGAAGG | 2.446 (70.4) | ATGAAGT | 7.070 (85.2) | -4.624 | Yes |
| rs3093547 | 31649827 | Athb-1 | - | 31649822 | 31649829 | taTttttt | 2.890 (68.7) | taAttttt | 7.282 (82.6) | -4.392 | Yes |
| rs3093547 | 31649827 | Athb-1 | - | 31649826 | 31649833 | aatttaTt | 6.670 (80.7) | aatttaAt | 3.182 (69.7) | 3.488 | Yes |
| rs3093547 | 31649827 | Broad-complex_1 | + | 31649820 | 31649833 | aaaaaaaAtaaatt | 8.373 (81.2) | aaaaaaaTtaaatt | 6.226 (75.6) | 2.147 | Yes |
| rs3093547 | 31649827 | Broad-complex_4 | + | 31649821 | 31649831 | aaaaaaAtaaa | 7.688 (83.0) | aaaaaaTtaaa | 4.492 (72.8) | 3.196 | Yes |

**Table 2.** Contd.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs3093547 | 31649827 | Broad-complex_4 | + | 31649825 | 31649835 | aaAtaaattat | 8.552 (85.8) | aaTtaaattat | 6.459 (79.1) | 2.093 | Yes |
| rs3093547 | 31649827 | HFH-1 | - | 31649826 | 31649836 | aataatttaTt | 8.701 (82.5) | aataatttaAt | 4.972 (72.9) | 3.729 | Yes |
| rs3093547 | 31649827 | HFH-2 | - | 31649820 | 31649831 | tttaTttttttt | 5.397 (77.7) | tttaAttttttt | 7.734 (82.4) | -2.337 | Yes |
| rs3093547 | 31649827 | HFH-2 | - | 31649821 | 31649832 | atttaTtttttt | 12.289 (91.5) | atttaAtttttt | 7.460 (81.8) | 4.829 | Yes |
| rs3093547 | 31649827 | HFH-2 | + | 31649825 | 31649836 | aaAtaaattatt | 2.154 (71.3) | aataatttaAtt | 7.189 (81.3) | -5.035 | Yes |
| rs3093547 | 31649827 | HFH-2 | - | 31649825 | 31649836 | aataatttaTtt | 9.400 (85.7) | aataatttaAtt | 7.189 (81.3) | 2.211 | Yes |
| rs3093547 | 31649827 | HFH-3 | - | 31649821 | 31649832 | atttaTtttttt | 8.881 (85.0) | atttaAtttttt | 4.340 (74.1) | 4.541 | Yes |
| rs3093547 | 31649827 | HFH-3 | - | 31649825 | 31649836 | aataatttaTtt | 7.260 (81.1) | aataatttaAtt | 3.071 (71.0) | 4.189 | Yes |
| rs3093547 | 31649827 | HNF-3beta | - | 31649821 | 31649832 | atttaTtttttt | 8.477 (84.1) | atttaAtttttt | 4.430 (73.8) | 4.047 | Yes |
| rs3093547 | 31649827 | HNF-3beta | - | 31649825 | 31649836 | aataatttaTtt | 9.702 (87.3) | aataatttaAtt | 6.571 (79.2) | 3.131 | Yes |
| rs3093547 | 31649827 | Hunchback | + | 31649819 | 31649828 | CaaaaaaaAt | 9.706 (90.4) | CaaaaaaaTt | 5.706 (78.1) | 4 | Yes |
| rs3093547 | 31649827 | Sox-5 | + | 31649821 | 31649827 | aaaaaaA | 2.289 (67.9) | aaaaaaT | 6.624 (84.1) | -4.335 | Yes |
| rs3093547 | 31649827 | Sox-5 | + | 31649822 | 31649828 | aaaaaAt | 6.624 (84.1) | aaaaaTt | 2.350 (68.1) | 4.274 | Yes |
| rs4248157 | 31650122 | NF-kappaB | + | 31650113 | 31650122 | GAGAACTTCC | 6.878 (80.2) | GAGAACTTCT | 3.472 (72.0) | 3.406 | No |
| rs4248157 | 31650122 | NF-kappaB | - | 31650114 | 31650123 | GGGAAGTTCT | 8.991 (85.3) | GAGAAGTTCT | 4.309 (74.0) | 4.682 | No |
| rs9282875 | 31650245 | Thing1-E47 | - | 31650244 | 31650253 | AGTCTGGCGG | 7.996 (85.9) | AGTCTGGCAG | 9.992 (91.8) | -1.996 | Yes |
| rs4647198 | 31650314 | COUP-TF | - | 31650309 | 31650322 | TGACCCCCGCCCCT | 9.975 (82.7) | TGACCCCCACCCCT | 7.938 (78.4) | 2.037 | No |
| rs4647198 | 31650314 | RORalfa-1 | + | 31650313 | 31650322 | GCGGGGGTCA | 4.801 (73.6) | GTGGGGGTCA | 7.654 (81.5) | -2.853 | No |
| rs2507961 | 31650433 | TBP | - | 31650424 | 31650438 | CTATGGAAGTCGAGT | 1.619 (72.0) | CTATGAAAGTCGAGT | 7.410 (83.5) | -5.791 | No |
| rs1800630 | 31650455 | NF-kappaB | + | 31650446 | 31650455 | GGGACCCCCA | 3.021 (70.9) | GGGGGGTCCC | 8.366 (83.8) | -5.345 | No |
| rs1800630 | 31650455 | NF-kappaB | - | 31650446 | 31650455 | TGGGGGTCCC | 3.685 (72.5) | GGGGGGTCCC | 8.366 (83.8) | -4.681 | No |
| rs1800630 | 31650455 | p50 | + | 31650445 | 31650455 | GGGGACCCCCA | 9.183 (82.2) | GGGGGGTCCCC | 13.895 (93.4) | -4.712 | No |
| rs1800630 | 31650455 | p50 | - | 31650445 | 31650455 | TGGGGGTCCCC | 9.727 (83.5) | GGGGGGTCCCC | 13.895 (93.4) | -4.168 | No |
| rs1800630 | 31650455 | p50 | - | 31650446 | 31650456 | GTGGGGGTCCC | 6.535 (75.9) | GGGGGGGTCCC | 10.702 (85.8) | -4.167 | No |
| rs4645836 | 31650456 | p50 | - | 31650446 | 31650456 | GTGGGGGTCCC | 6.535 (75.9) | GGGGGGGTCCC | 10.702 (85.8) | -4.167 | No |
| rs1800630 | 31650455 | p50 | + | 31650446 | 31650456 | GGGACCCCCCA | 4.155 (70.2) | GGGGGGGTCCC | 10.702 (85.8) | -6.547 | No |
| rs4645836 | 31650456 | p50 | + | 31650446 | 31650456 | GGGACCCCCCA | 4.155 (70.2) | GGGGGGGTCCC | 10.702 (85.8) | -6.547 | No |
| rs1800630 | 31650455 | p50 | - | 31650446 | 31650456 | TGGGGGGTCCC | 6.535 (75.9) | GGGGGGGTCCC | 10.702 (85.8) | -4.167 | No |
| rs4645836 | 31650456 | p50 | - | 31650446 | 31650456 | TGGGGGGTCCC | 6.535 (75.9) | GGGGGGGTCCC | 10.702 (85.8) | -4.167 | No |
| rs4248158 | 31650512 | E74A | - | 31650511 | 31650517 | CTGGAGG | 2.967 (69.2) | CTGGAAG | 7.095 (85.5) | -4.128 | No |
| rs4248158 | 31650512 | Thing1-E47 | - | 31650511 | 31650520 | TACCTGGAGG | 4.622 (76.0) | TACCTGGAAG | 6.618 (81.9) | -1.996 | No |
| rs4987086 | 31650536 | bZIP910 | + | 31650531 | 31650537 | GGGACAT | 1.099 (66.1) | GGGACGT | 5.723 (80.9) | -4.624 | No |
| rs4248159 | 31650559 | SOX17 | + | 31650555 | 31650563 | CACAATGGG | 2.910 (71.7) | CCCAGTGTG | 6.074 (82.0) | -3.164 | No |
| rs4248159 | 31650559 | SOX17 | - | 31650555 | 31650563 | CCCATTGTG | 10.569 (96.5) | CCCAGTGTG | 6.074 (82.0) | 4.495 | No |
| rs4248159 | 31650559 | Sox-5 | + | 31650554 | 31650560 | ACACAAT | 6.687 (84.3) | ACACACT | 3.288 (71.6) | 3.399 | No |
| rs2857712 | 31650632 | HFH-2 | + | 31650623 | 31650634 | TAATGCTGGGTT | 4.158 (75.3) | TAATGCTGGTTT | 8.539 (84.0) | -4.381 | Yes |
| rs2736196 | 31650634 | HFH-2 | + | 31650623 | 31650634 | TAATGCTGGGTT | 4.158 (75.3) | TAATGCTGGTTT | 8.539 (84.0) | -4.381 | Yes |
| rs2857712 | 31650632 | HNF-3beta | + | 31650623 | 31650634 | TAATGCTGGGTT | 4.638 (74.3) | TAATGCTGGTTT | 7.769 (82.3) | -3.131 | Yes |
| rs2736196 | 31650634 | HNF-3beta | + | 31650623 | 31650634 | TAATGCTGGGTT | 4.638 (74.3) | TAATGCTGGTTT | 7.769 (82.3) | -3.131 | Yes |
| rs2857712 | 31650632 | Irf-1 | - | 31650630 | 31650641 | AAGACTGAACCC | 6.671 (76.4) | AAGACTGAAACC | 10.910 (85.8) | -4.239 | No |
| rs2736196 | 31650634 | Irf-1 | - | 31650630 | 31650641 | AAGACTGAACCC | 6.671 (76.4) | AAGACTGAAACC | 10.910 (85.8) | -4.239 | No |

**Table 2.** Contd.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs2857712 | 31650632 | Athb-1 | - | 31650628 | 31650635 | CCAGCATT | 3.344 (70.2) | CAAGCATT | 6.650 (80.6) | -3.306 | Yes |
| rs2736196 | 31650634 | Athb-1 | - | 31650628 | 31650635 | CCAGCATT | 3.344 (70.2) | CAAGCATT | 6.650 (80.6) | -3.306 | Yes |
| rs2857712 | 31650632 | HFH-2 | + | 31650627 | 31650638 | TAATGCTGGTTT | 8.539 (84.0) | TAATGCTTGTTT | 10.064 (87.0) | -1.525 | No |
| rs2736196 | 31650634 | HFH-2 | + | 31650627 | 31650638 | TAATGCTGGTTT | 8.539 (84.0) | TAATGCTTGTTT | 10.064 (87.0) | -1.525 | No |
| rs2857712 | 31650632 | HFH-3 | + | 31650627 | 31650638 | TAATGCTGGTTT | 4.259 (73.9) | TAATGCTTGTTT | 8.800 (84.8) | -4.541 | No |
| rs2736196 | 31650634 | HFH-3 | + | 31650627 | 31650638 | TAATGCTGGTTT | 4.259 (73.9) | TAATGCTTGTTT | 8.800 (84.8) | -4.541 | No |
| rs2736196 | 31650634 | Irf-1 | - | 31650634 | 31650645 | AAGACTGAAACC | 10.910 (85.8) | AAGACTGAAACA | 8.174 (79.7) | 2.736 | No |
| rs2857712 | 31650632 | Sox-5 | - | 31650632 | 31650638 | AAACCAG | 2.228 (67.7) | AAACAAG | 6.563 (83.8) | -4.335 | No |
| rs2736196 | 31650634 | Sox-5 | - | 31650632 | 31650638 | AAACCAG | 2.228 (67.7) | AAACAAG | 6.563 (83.8) | -4.335 | No |
| rs2736195 | 31650670 | NF-kappaB | - | 31650661 | 31650670 | GGGCCTGCCC | 7.010 (80.5) | AGGCCTGCCC | 2.329 (69.2) | 4.681 | Yes |
| rs4248160 | 31650672 | Snail | - | 31650671 | 31650676 | CAGGTC | 6.099 (83.4) | CAGGCC | 1.382 (66.5) | 4.717 | Yes |
| rs1800195 | 31650743 | Spz1 | - | 31650740 | 31650750 | GGGGTAGTAGG | 6.969 (79.8) | GGGGTAGCAGG | 10.479 (89.0) | -3.51 | No |
| rs4248161 | 31650746 | Spz1 | - | 31650740 | 31650750 | GGGGTAGTAGG | 6.969 (79.8) | GGGGTAGCAGG | 10.479 (89.0) | -3.51 | No |
| rs4248162 | 31650748 | Spz1 | - | 31650740 | 31650750 | GGGGTAGTAGG | 6.969 (79.8) | GGGGTAGCAGG | 10.479 (89.0) | -3.51 | No |
| rs2857711 | 31650750 | Spz1 | - | 31650740 | 31650750 | GGGGTAGTAGG | 6.969 (79.8) | GGGGTAGCAGG | 10.479 (89.0) | -3.51 | No |
| rs1800195 | 31650743 | Spz1 | - | 31650740 | 31650750 | GGGGTAGCAGG | 10.479 (89.0) | GGGGGAGCAGG | 7.938 (82.4) | 2.541 | No |
| rs4248161 | 31650746 | Spz1 | - | 31650740 | 31650750 | GGGGTAGCAGG | 10.479 (89.0) | GGGGGAGCAGG | 7.938 (82.4) | 2.541 | No |
| rs4248162 | 31650748 | Spz1 | - | 31650740 | 31650750 | GGGGTAGCAGG | 10.479 (89.0) | GGGGGAGCAGG | 7.938 (82.4) | 2.541 | No |
| rs2857711 | 31650750 | Spz1 | - | 31650740 | 31650750 | GGGGTAGCAGG | 10.479 (89.0) | GGGGGAGCAGG | 7.938 (82.4) | 2.541 | No |
| rs1800195 | 31650743 | Spz1 | - | 31650740 | 31650750 | GGTGTAGCAGG | 6.830 (79.5) | GGGGTAGCAGG | 10.479 (89.0) | -3.649 | No |
| rs4248161 | 31650746 | Spz1 | - | 31650740 | 31650750 | GGTGTAGCAGG | 6.830 (79.5) | GGGGTAGCAGG | 10.479 (89.0) | -3.649 | No |
| rs4248162 | 31650748 | Spz1 | - | 31650740 | 31650750 | GGTGTAGCAGG | 6.830 (79.5) | GGGGTAGCAGG | 10.479 (89.0) | -3.649 | No |
| rs2857711 | 31650750 | Spz1 | - | 31650740 | 31650750 | GGTGTAGCAGG | 6.830 (79.5) | GGGGTAGCAGG | 10.479 (89.0) | -3.649 | No |
| rs2857711 | 31650750 | Spz1 | + | 31650750 | 31650760 | CGGGTAGCAGG | 8.321 (83.4) | GGGGTAGCAGG | 10.479 (89.0) | -2.158 | No |
| rs3093548 | 31650799 | Hen-1 | + | 31650790 | 31650801 | CCCCAGCTCCTT | 8.205 (80.7) | CCCCAGCTCTTT | 5.896 (75.7) | 2.309 | No |
| rs3093548 | 31650799 | Myf | + | 31650790 | 31650801 | CCCCAGCTCCTT | 2.679 (69.6) | AAAGAGCTGGGG | 7.632 (80.8) | -4.953 | No |
| rs4248163 | 31650807 | E74A | - | 31650802 | 31650808 | CGGGGAG | 2.967 (69.2) | CCGGGAG | 6.627 (83.7) | -3.66 | No |
| rs4248163 | 31650807 | E74A | + | 31650805 | 31650811 | CCCGCAG | 2.498 (67.4) | CCGGCAG | 6.627 (83.7) | -4.129 | No |
| rs4248164 | 31651327 | Hen-1 | + | 31651320 | 31651331 | CAGCAGACGCTC | 5.499 (74.9) | CAGCAGATGCTC | 9.745 (84.0) | -4.246 | Yes |
| rs4248164 | 31651327 | Hen-1 | - | 31651320 | 31651331 | GAGCGTCTGCTG | 4.037 (71.7) | CAGCAGATGCTC | 9.745 (84.0) | -5.708 | Yes |
| rs4248164 | 31651327 | Myf | + | 31651317 | 31651328 | AGCCAGCAGACG | 6.113 (77.4) | AGCCAGCAGATG | 9.573 (85.2) | -3.46 | Yes |
| rs4248164 | 31651327 | Myf | + | 31651320 | 31651331 | CAGCAGACGCTC | 4.504 (73.7) | GAGCATCTGCTG | 10.559 (87.5) | -6.055 | Yes |
| rs4248164 | 31651327 | Myf | - | 31651320 | 31651331 | GAGCGTCTGCTG | 6.471 (78.2) | GAGCATCTGCTG | 10.559 (87.5) | -4.088 | Yes |
| rs4248164 | 31651327 | Myf | - | 31651323 | 31651334 | AGGGAGCGTCTG | 5.012 (74.9) | AGGGAGCATCTG | 7.819 (81.3) | -2.807 | Yes |
| rs4248164 | 31651327 | Snail | + | 31651323 | 31651328 | CAGACG | 2.561 (70.8) | CAGATG | 7.278 (87.6) | -4.717 | Yes |
| rs4248164 | 31651327 | Tal1beta-E47S | - | 31651321 | 31651332 | GGAGCGTCTGCT | 4.566 (74.4) | GGAGCATCTGCT | 9.349 (84.9) | -4.783 | Yes |
| rs4645838 | 31651391 | Dorsal_1 | - | 31651389 | 31651400 | GTTGTTTTCAGG | 3.911 (70.8) | GGTTGTTTTCAG | 7.863 (80.7) | -3.952 | No |
| rs4645838 | 31651391 | E74A | + | 31651390 | 31651396 | CTGAAAA | 2.096 (65.8) | CCTGAAA | 5.756 (80.2) | -3.66 | No |
| rs2515924 | 31651470 | CREB | + | 31651463 | 31651474 | GAGGGTGAAGCC | 7.769 (83.5) | GAGGGTGGAGCC | 3.681 (72.5) | 4.088 | No |
| rs2228088 | 31651584 | Myf | - | 31651578 | 31651589 | AAGCACCGCCTG | 5.012 (74.9) | AAGCAACGCCTG | 8.012 (81.7) | -3 | Yes |
| rs3179060 | 31651651 | COUP-TF | + | 31651649 | 31651662 | TGAACTTTGGAGTG | 13.003 (89.2) | TGCACTTTGGAGTG | 9.164 (81.0) | 3.839 | Yes |

**Table 2.** Contd.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs3093661 | 31651737 | HMG-IY | + | 31651723 | 31651738 | GGGGAAATGGAGACAC | 10.540 (86.9) | GGGGAAATGGAGACGC | 8.939 (83.2) | 1.601 | No |
| rs4645839 | 31651804 | ARNT | + | 31651803 | 31651808 | aAcgtg | 8.609 (93.0) | aTcgtg | 4.440 (76.4) | 4.169 | No |
| rs1800610 | 31651806 | ARNT | + | 31651803 | 31651808 | aAcgtg | 8.609 (93.0) | aTcgtg | 4.440 (76.4) | 4.169 | No |
| rs4645839 | 31651804 | ARNT | - | 31651803 | 31651808 | cacgTt | 6.112 (83.1) | aTcgtg | 4.440 (76.4) | 1.672 | No |
| rs1800610 | 31651806 | ARNT | - | 31651803 | 31651808 | cacgTt | 6.112 (83.1) | aTcgtg | 4.440 (76.4) | 1.672 | No |
| rs4645839 | 31651804 | HFH-2 | - | 31651798 | 31651809 | ccacgTtttttt | 7.249 (81.4) | ccacgAtttttt | 2.421 (71.8) | 4.828 | No |
| rs1800610 | 31651806 | HFH-2 | - | 31651798 | 31651809 | ccacgTtttttt | 7.249 (81.4) | ccacgAtttttt | 2.421 (71.8) | 4.828 | No |
| rs4645839 | 31651804 | Hunchback | + | 31651796 | 31651805 | agaaaaaaAc | 6.484 (80.5) | agaaaaaaTc | 2.484 (68.1) | 4 | No |
| rs4645839 | 31651804 | Hunchback | + | 31651797 | 31651806 | gaaaaaaAcg | 6.876 (81.7) | gaaaaaaTcg | 2.969 (69.6) | 3.907 | No |
| rs1800610 | 31651806 | Hunchback | + | 31651797 | 31651806 | gaaaaaaAcg | 6.876 (81.7) | gaaaaaaTcg | 2.969 (69.6) | 3.907 | No |
| rs4645839 | 31651804 | Max | + | 31651800 | 31651809 | aaaaAcgtgg | 8.321 (85.4) | aaaaTcgtgg | 4.274 (73.1) | 4.047 | No |
| rs1800610 | 31651806 | Max | + | 31651800 | 31651809 | aaaaAcgtgg | 8.321 (85.4) | aaaaTcgtgg | 4.274 (73.1) | 4.047 | No |
| rs4645839 | 31651804 | Myc-Max | + | 31651800 | 31651810 | aaaaAcgtgga | 8.801 (82.1) | aaaaTcgtgga | 4.528 (71.3) | 4.273 | No |
| rs1800610 | 31651806 | Myc-Max | + | 31651800 | 31651810 | aaaaAcgtgga | 8.801 (82.1) | aaaaTcgtgga | 4.528 (71.3) | 4.273 | No |
| rs4645839 | 31651804 | Sox-5 | + | 31651798 | 31651804 | aaaaaaA | 2.289 (67.9) | aaaaaaT | 6.624 (84.1) | -4.335 | No |
| rs4645839 | 31651804 | USF | - | 31651802 | 31651808 | cacgTtt | 5.702 (80.1) | cacgAtt | 1.280 (64.9) | 4.422 | No |
| rs1800610 | 31651806 | USF | - | 31651802 | 31651808 | cacgTtt | 5.702 (80.1) | cacgAtt | 1.280 (64.9) | 4.422 | No |
| rs4645839 | 31651804 | USF | + | 31651803 | 31651809 | aAcgtgg | 6.973 (84.5) | aTcgtgg | 2.455 (68.9) | 4.518 | No |
| rs1800610 | 31651806 | USF | + | 31651803 | 31651809 | aAcgtgg | 6.973 (84.5) | aTcgtgg | 2.455 (68.9) | 4.518 | No |
| rs4645839 | 31651804 | n-MYC | + | 31651803 | 31651808 | aAcgtg | 5.992 (83.1) | aTcgtg | 2.829 (71.3) | 3.163 | No |
| rs1800610 | 31651806 | n-MYC | + | 31651803 | 31651808 | aAcgtg | 5.992 (83.1) | aTcgtg | 2.829 (71.3) | 3.163 | No |
| rs4645839 | 31651804 | n-MYC | - | 31651803 | 31651808 | cacgTt | 5.992 (83.1) | aTcgtg | 2.829 (71.3) | 3.163 | No |
| rs1800610 | 31651806 | n-MYC | - | 31651803 | 31651808 | cacgTt | 5.992 (83.1) | aTcgtg | 2.829 (71.3) | 3.163 | No |
| rs4645839 | 31651804 | ARNT | + | 31651804 | 31651809 | caCgtt | 6.112 (83.1) | aacAtg | 4.370 (76.1) | 1.742 | No |
| rs1800610 | 31651806 | ARNT | + | 31651804 | 31651809 | caCgtt | 6.112 (83.1) | aacAtg | 4.370 (76.1) | 1.742 | No |
| rs4645839 | 31651804 | ARNT | - | 31651804 | 31651809 | aacGtg | 8.609 (93.0) | aacAtg | 4.370 (76.1) | 4.239 | No |
| rs1800610 | 31651806 | ARNT | - | 31651804 | 31651809 | aacGtg | 8.609 (93.0) | aacAtg | 4.370 (76.1) | 4.239 | No |
| rs4645839 | 31651804 | HFH-2 | + | 31651803 | 31651814 | ccaCgttttttt | 7.249 (81.4) | ccaTgttttttt | 10.185 (87.3) | -2.936 | No |
| rs1800610 | 31651806 | HFH-2 | + | 31651803 | 31651814 | ccaCgttttttt | 7.249 (81.4) | ccaTgttttttt | 10.185 (87.3) | -2.936 | No |
| rs4645839 | 31651804 | HFH-3 | + | 31651803 | 31651814 | ccaCgttttttt | 3.705 (72.5) | ccaTgttttttt | 8.246 (83.5) | -4.541 | No |
| rs1800610 | 31651806 | HFH-3 | + | 31651803 | 31651814 | ccaCgttttttt | 3.705 (72.5) | ccaTgttttttt | 8.246 (83.5) | -4.541 | No |
| rs1800610 | 31651806 | HMG-IY | - | 31651805 | 31651820 | ggagagaaaaaaacGt | 7.221 (79.3) | ggagagaaaaaaacAt | 8.822 (82.9) | -1.601 | No |
| rs4645839 | 31651804 | HNF-3beta | + | 31651803 | 31651814 | ccaCgttttttt | 5.819 (77.3) | ccaTgttttttt | 8.887 (85.2) | -3.068 | No |
| rs1800610 | 31651806 | HNF-3beta | + | 31651803 | 31651814 | ccaCgttttttt | 5.819 (77.3) | ccaTgttttttt | 8.887 (85.2) | -3.068 | No |
| rs1800610 | 31651806 | Hunchback | - | 31651806 | 31651815 | gaaaaaaacG | 6.876 (81.7) | gaaaaaaacA | 8.613 (87.0) | -1.737 | No |
| rs4645839 | 31651804 | Max | - | 31651803 | 31651812 | aaaaacGtgg | 8.321 (85.4) | aaaaacAtgg | 5.253 (76.1) | 3.068 | No |
| rs1800610 | 31651806 | Max | - | 31651803 | 31651812 | aaaaacGtgg | 8.321 (85.4) | aaaaacAtgg | 5.253 (76.1) | 3.068 | No |
| rs4645839 | 31651804 | Myc-Max | - | 31651802 | 31651812 | aaaaacGtgga | 8.801 (82.1) | aaaaacAtgga | 5.500 (73.8) | 3.301 | No |
| rs1800610 | 31651806 | Myc-Max | - | 31651802 | 31651812 | aaaaacGtgga | 8.801 (82.1) | aaaaacAtgga | 5.500 (73.8) | 3.301 | No |
| rs4645839 | 31651804 | USF | - | 31651803 | 31651809 | aacGtgg | 6.973 (84.5) | aacAtgg | 3.293 (71.8) | 3.68 | No |
| rs1800610 | 31651806 | USF | - | 31651803 | 31651809 | aacGtgg | 6.973 (84.5) | aacAtgg | 3.293 (71.8) | 3.68 | No |

**Table 2.** Contd.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs4645839 | 31651804 | USF | + | 31651804 | 31651810 | caCgttt | 5.702 (80.1) | caTgttt | 2.579 (69.3) | 3.123 | No |
| rs1800610 | 31651806 | USF | + | 31651804 | 31651810 | caCgttt | 5.702 (80.1) | caTgttt | 2.579 (69.3) | 3.123 | No |
| rs4645839 | 31651804 | n-MYC | + | 31651804 | 31651809 | caCgtt | 5.992 (83.1) | aacAtg | 2.877 (71.5) | 3.115 | No |
| rs1800610 | 31651806 | n-MYC | + | 31651804 | 31651809 | caCgtt | 5.992 (83.1) | aacAtg | 2.877 (71.5) | 3.115 | No |
| rs4645839 | 31651804 | n-MYC | - | 31651804 | 31651809 | aacGtg | 5.992 (83.1) | aacAtg | 2.877 (71.5) | 3.115 | No |
| rs1800610 | 31651806 | n-MYC | - | 31651804 | 31651809 | aacGtg | 5.992 (83.1) | aacAtg | 2.877 (71.5) | 3.115 | No |
| **IFR-1 (Interferon regulatory factor-1)** | | | | | | | | | | | |
| rs960757 | 131854209 | E2F | + | 131854207 | 131854214 | CTCCGCGC | 4.558 (72.3) | CTTCGCGC | 8.329 (85.3) | -3.771 | No |
| rs10900809 | 131854221 | ARNT | + | 131854217 | 131854222 | CACTCG | 1.872 (66.2) | CACTTG | 6.112 (83.1) | -4.24 | No |
| rs10900809 | 131854221 | ARNT | - | 131854217 | 131854222 | CGAGTG | 2.864 (70.1) | CACTTG | 6.112 (83.1) | -3.248 | No |
| rs10900809 | 131854221 | Snail | - | 131854217 | 131854222 | CGAGTG | 2.989 (72.3) | CAAGTG | 7.670 (89.0) | -4.681 | No |
| rs10900809 | 131854221 | USF | - | 131854216 | 131854222 | CGAGTGG | 2.550 (69.2) | CAAGTGG | 7.068 (84.8) | -4.518 | No |
| rs10900809 | 131854221 | USF | + | 131854217 | 131854223 | CACTCGG | 3.388 (72.1) | CACTTGG | 7.020 (84.6) | -3.632 | No |
| rs10900809 | 131854221 | n-MYC | + | 131854217 | 131854222 | CACTCG | 3.390 (73.4) | CACTTG | 6.133 (83.6) | -2.743 | No |
| rs10900809 | 131854221 | n-MYC | - | 131854217 | 131854222 | CGAGTG | 1.589 (66.7) | CACTTG | 6.133 (83.6) | -4.544 | No |
| rs2549009 | 131854664 | CFI-USP | + | 131854658 | 131854667 | GGGGCCAGGG | 8.380 (81.3) | GGGGCCGGGG | 5.371 (73.5) | 3.009 | No |
| rs2549009 | 131854664 | Snail | + | 131854663 | 131854668 | CAGGGG | 6.028 (83.1) | CGGGGG | 1.346 (66.4) | 4.682 | No |
| rs2549008 | 131854752 | c-REL | - | 131854744 | 131854753 | GGGGTATATC | 6.482 (80.1) | GAGGTATATC | 2.523 (68.0) | 3.959 | No |
| rs2549007 | 131854774 | ARNT | + | 131854773 | 131854778 | CACCTG | 6.112 (83.1) | CGCCTG | 2.864 (70.1) | 3.248 | No |
| rs2549007 | 131854774 | ARNT | - | 131854773 | 131854778 | CAGGTG | 6.112 (83.1) | CGCCTG | 2.864 (70.1) | 3.248 | No |
| rs2549007 | 131854774 | Hen-1 | + | 131854770 | 131854781 | CCTCACCTGCGT | 10.729 (86.1) | CCTCGCCTGCGT | 5.803 (75.5) | 4.926 | No |
| rs2706384 | 131854779 | Hen-1 | + | 131854770 | 131854781 | CCTCACCTGCGT | 10.729 (86.1) | CCTCGCCTGCGT | 5.803 (75.5) | 4.926 | No |
| rs2549007 | 131854774 | Hen-1 | - | 131854770 | 131854781 | ACGCAGGTGAGG | 9.731 (84.0) | CCTCGCCTGCGT | 5.803 (75.5) | 3.928 | No |
| rs2706384 | 131854779 | Hen-1 | - | 131854770 | 131854781 | ACGCAGGTGAGG | 9.731 (84.0) | CCTCGCCTGCGT | 5.803 (75.5) | 3.928 | No |
| rs2549007 | 131854774 | Snail | + | 131854773 | 131854778 | CACCTG | 1.490 (66.9) | CAGGCG | 6.028 (83.1) | -4.538 | No |
| rs2549007 | 131854774 | Snail | - | 131854773 | 131854778 | CAGGTG | 10.744 (100.0) | CAGGCG | 6.028 (83.1) | 4.716 | No |
| rs2549007 | 131854774 | USF | - | 131854772 | 131854778 | CAGGTGA | 6.501 (82.8) | CAGGCGA | 2.869 (70.3) | 3.632 | No |
| rs2549007 | 131854774 | n-MYC | + | 131854773 | 131854778 | CACCTG | 6.914 (86.5) | CAGGCG | 3.295 (73.0) | 3.619 | No |
| rs2549007 | 131854774 | n-MYC | - | 131854773 | 131854778 | CAGGTG | 6.038 (83.2) | CAGGCG | 3.295 (73.0) | 2.743 | No |
| rs2706384 | 131854779 | NF-kappaB | + | 131854777 | 131854786 | GGGGATCGCC | 7.158 (80.9) | GGCGATCACC | 3.714 (72.6) | 3.444 | No |
| rs2706384 | 131854779 | Snail | + | 131854775 | 131854780 | GAGGGG | 2.053 (68.9) | GAGGTG | 6.770 (85.8) | -4.717 | No |
| rs2549007 | 131854774 | bZIP910 | + | 131854773 | 131854779 | GTGAGGG | 1.099 (66.1) | GTGAGGT | 5.723 (80.9) | -4.624 | No |
| rs2706384 | 131854779 | bZIP910 | + | 131854773 | 131854779 | GTGAGGG | 1.099 (66.1) | GTGAGGT | 5.723 (80.9) | -4.624 | No |
| rs2706384 | 131854779 | c-REL | - | 131854779 | 131854788 | CGGGCGATCC | 6.649 (80.6) | CGGGCGATCA | 3.581 (71.2) | 3.068 | No |
| rs2706384 | 131854779 | p50 | - | 131854776 | 131854786 | GGCGATCCCCT | 8.523 (80.7) | GGCGATCACCT | 4.794 (71.8) | 3.729 | No |
| rs2706384 | 131854779 | p50 | + | 131854777 | 131854787 | GGGGATCGCCC | 11.924 (88.8) | GGTGATCGCCC | 7.756 (78.8) | 4.168 | No |
| rs2706384 | 131854779 | p50 | - | 131854777 | 131854787 | GGGCGATCCCC | 10.873 (86.2) | GGTGATCGCCC | 7.756 (78.8) | 3.117 | No |
| rs2706384 | 131854779 | Spz1 | + | 131854776 | 131854786 | AGGGGATCGCC | 7.357 (80.9) | AGGTGATCGCC | 4.688 (73.9) | 2.669 | No |
| rs2549006 | 131854964 | Thing1-E47 | + | 131854955 | 131854964 | TTTCTGCCTC | 5.811 (79.5) | TTTCTGCCTT | 7.509 (84.5) | -1.698 | No |
| rs2549005 | 131855090 | Athb-1 | + | 131855088 | 131855095 | ctAttatt | 10.419 (92.5) | ctGttatt | 6.027 (78.6) | 4.392 | No |

**Table 2.** Contd.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs2549005 | 131855090 | SQUA | - | 131855084 | 131855097 | aaaataaTagtagc | 7.040 (80.1) | aaaataaCagtagc | 3.010 (71.8) | 4.03 | No |
| rs2549005 | 131855090 | Sox-5 | - | 131855087 | 131855093 | taaTagt | 3.225 (71.4) | taaCagt | 6.624 (84.1) | -3.399 | No |
| rs2549005 | 131855090 | Sox-5 | - | 131855090 | 131855096 | aaataaT | 7.499 (87.3) | aaataaC | 3.164 (71.2) | 4.335 | No |
| rs3840527 | 131855552 | Dorsal_2 | + | 131855547 | 131855556 | GGGGCTTACC | 9.321 (85.5) | GGGGCCTTAC | 7.217 (79.3) | 2.104 | No |
| rs3840527 | 131855552 | NF-kappaB | + | 131855547 | 131855556 | GGGGCTTACC | 11.146 (90.5) | GGGCCTTACC | 7.908 (82.7) | 3.238 | No |
| rs3840527 | 131855552 | NF-kappaB | - | 131855547 | 131855556 | GGTAAGCCCC | 4.319 (74.0) | GGGCCTTACC | 7.908 (82.7) | -3.589 | No |
| rs3840527 | 131855552 | c-REL | + | 131855546 | 131855555 | AGGGGCTTAC | 2.628 (68.3) | GGGCCTTACC | 8.033 (84.9) | -5.405 | No |
| rs3840527 | 131855552 | c-REL | + | 131855547 | 131855556 | GGGGCTTACC | 10.338 (92.0) | GGGCCTTACC | 8.033 (84.9) | 2.305 | No |
| rs3840527 | 131855552 | c-REL | - | 131855548 | 131855557 | TGGTAAGCCC | 1.515 (64.9) | GGGCCTTACC | 8.033 (84.9) | -6.518 | No |
| rs3840527 | 131855552 | p65 | + | 131855547 | 131855556 | GGGGCTTACC | 9.235 (84.2) | GGGCCTTACC | 6.309 (76.6) | 2.926 | No |
| rs2549004 | 131855724 | CFI-USP | + | 131855723 | 131855732 | GCGGTGAAGG | 5.096 (72.8) | GGGGTGAAGG | 9.703 (84.7) | -4.607 | Yes |
| rs12657668 | 131856363 | AGL3 | + | 131856361 | 131856370 | CCATATATGC | 9.097 (87.8) | CCGTATATGC | 5.199 (78.8) | 3.898 | No |
| rs12657668 | 131856363 | AGL3 | - | 131856361 | 131856370 | GCATATATGG | 8.347 (86.1) | CCGTATATGC | 5.199 (78.8) | 3.148 | No |
| rs12657668 | 131856363 | CF2-II | - | 131856359 | 131856368 | ATATATGGGT | 7.393 (83.1) | ATATACGGGT | 3.781 (74.2) | 3.612 | No |
| rs12657668 | 131856363 | CF2-II | + | 131856361 | 131856370 | CCATATATGC | 7.217 (82.7) | GCATATACGG | 5.186 (77.7) | 2.031 | No |
| rs12657668 | 131856363 | CF2-II | - | 131856361 | 131856370 | GCATATATGG | 7.830 (84.2) | GCATATACGG | 5.186 (77.7) | 2.644 | No |
| rs12657668 | 131856363 | CF2-II | - | 131856363 | 131856372 | ATGCATATAT | 7.088 (82.4) | GTATATGCAT | 11.134 (92.4) | -4.046 | No |
| rs12657668 | 131856363 | SQUA | + | 131856351 | 131856364 | ATAAATATACCCAT | 7.676 (81.4) | ATAAATATACCCGT | 5.320 (76.5) | 2.356 | No |
| rs12657668 | 131856363 | SQUA | + | 131856361 | 131856374 | CCATATATGCATAT | 7.962 (82.0) | CCGTATATGCATAT | 5.722 (77.4) | 2.24 | No |
| rs12657668 | 131856363 | SRF | + | 131856359 | 131856370 | ACCCATATATGC | 12.487 (86.0) | ACCCGTATATGC | 8.174 (78.0) | 4.313 | No |
| rs12657668 | 131856363 | TBP | - | 131856356 | 131856370 | GCATATATGGGTATA | 7.456 (83.5) | GCATATACGGGTATA | 3.273 (75.2) | 4.183 | No |
| rs12657668 | 131856363 | TBP | + | 131856361 | 131856375 | CCATATATGCATATG | 7.470 (83.6) | CCGTATATGCATATG | 1.784 (72.3) | 5.686 | No |
| rs2549003 | 131857196 | E74A | - | 131857194 | 131857200 | TAGGGAA | 1.962 (65.2) | TAGGAAA | 6.091 (81.6) | -4.129 | Yes |
| rs2549003 | 131857196 | Irf-2 | - | 131857180 | 131857197 | GGAAGATGAAACTAAAGG | 12.031 (81.5) | GAAAGATGAAACTAAAGG | 10.108 (78.1) | 1.923 | Yes |
| rs2549003 | 131857196 | Staf | + | 131857190 | 131857209 | CATCTTCCCTACTGCCAAGG | 12.148 (83.3) | CATCTTTCCTACTGCCAAGG | 8.377 (76.8) | 3.771 | No |
| **TGF-β-1 (Transforming growth factor-beta-1)** | | | | | | | | | | | |
| rs11466316 | 46551176 | E74A | - | 46551175 | 46551181 | CGGGAGG | 2.967 (69.2) | CGGGAAG | 7.095 (85.5) | -4.128 | Yes |
| rs11551226 | 46551242 | CFI-USP | + | 46551237 | 46551246 | GGGGGCAGGG | 8.380 (81.3) | GGGGGTAGGG | 4.495 (71.2) | 3.885 | No |
| rs1799753 | 46551551 | NF-kappaB | - | 46551551 | 46551560 | ggggtgcccc | 6.444 (79.2) | ggggcaCccc | 8.508 (84.2) | -2.064 | Yes |
| rs1800999 | 46551557 | NF-kappaB | - | 46551551 | 46551560 | ggggtgcccc | 6.444 (79.2) | ggggcaCccc | 8.508 (84.2) | -2.064 | Yes |
| rs1800999 | 46551557 | NF-kappaB | + | 46551552 | 46551561 | gggcaccccc | 4.019 (73.3) | ggggcaCccc | 8.508 (84.2) | -4.489 | Yes |
| rs1800999 | 46551557 | NF-kappaB | + | 46551553 | 46551562 | ggcacccccc | 2.983 (70.8) | ggggcaCccc | 8.508 (84.2) | -5.525 | Yes |
| rs1800999 | 46551557 | NF-kappaB | - | 46551553 | 46551562 | ggggggtgcc | 4.777 (75.1) | ggggcaCccc | 8.508 (84.2) | -3.731 | Yes |
| rs1799753 | 46551551 | p50 | + | 46551550 | 46551560 | cggggcacccc | 4.691 (71.5) | ggggGtgcccc | 13.895 (93.4) | -9.204 | Yes |
| rs1800999 | 46551557 | p50 | + | 46551550 | 46551560 | cggggcacccc | 4.691 (71.5) | ggggGtgcccc | 13.895 (93.4) | -9.204 | Yes |
| rs1799753 | 46551551 | p50 | - | 46551550 | 46551560 | ggggtgccccg | 7.205 (77.5) | ggggGtgcccc | 13.895 (93.4) | -6.69 | Yes |
| rs1800999 | 46551557 | p50 | - | 46551550 | 46551560 | ggggtgccccg | 7.205 (77.5) | ggggGtgcccc | 13.895 (93.4) | -6.69 | Yes |
| rs1799753 | 46551551 | p50 | + | 46551551 | 46551561 | ggggcaccccc | 12.111 (89.2) | ggggGtgcccc | 13.895 (93.4) | -1.784 | Yes |
| rs1800999 | 46551557 | p50 | + | 46551551 | 46551561 | ggggcaccccc | 12.111 (89.2) | ggggGtgcccc | 13.895 (93.4) | -1.784 | Yes |
| rs1800999 | 46551557 | p50 | + | 46551552 | 46551562 | gggcacccccc | 8.138 (79.7) | ggggGtgcccc | 13.895 (93.4) | -5.757 | Yes |

**Table 2.** Contd.

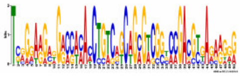
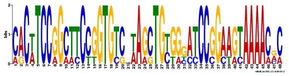
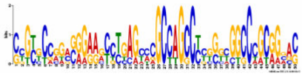
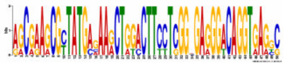
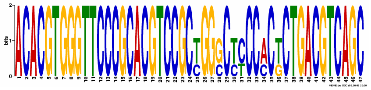
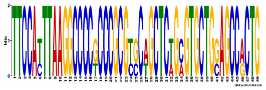
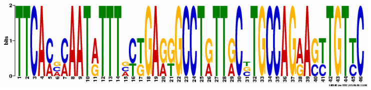
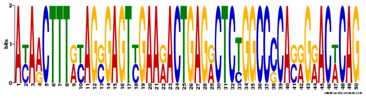
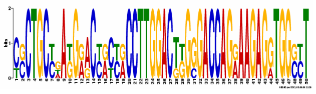
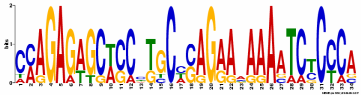
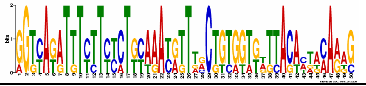
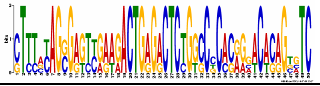
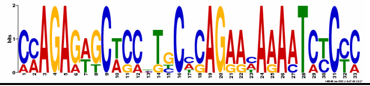
| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs1800999 | 46551557 | p50 | - | 46551552 | 46551562 | ggggggtgccc | 10.166 (84.6) | ggggGtgcccc | 13.895 (93.4) | -3.729 | Yes |
| rs12985978 | 46551621 | E74A | + | 46551617 | 46551623 | GCGGAAG | 8.651 (91.7) | GCGGAGG | 4.523 (75.4) | 4.128 | Yes |
| rs12977601 | 46551622 | E74A | + | 46551617 | 46551623 | GCGGAAG | 8.651 (91.7) | GCGGAGG | 4.523 (75.4) | 4.128 | Yes |
| rs12985978 | 46551621 | Myf | + | 46551618 | 46551629 | CGGAAGGAGGTG | 7.425 (80.4) | CGGAGGGAGGTG | 3.337 (71.0) | 4.088 | Yes |
| rs12977601 | 46551622 | Myf | + | 46551618 | 46551629 | CGGAAGGAGGTG | 7.425 (80.4) | CGGAGGGAGGTG | 3.337 (71.0) | 4.088 | Yes |
| rs12985978 | 46551621 | TEF-1 | - | 46551618 | 46551629 | CACCTCCTTCCG | 5.595 (73.3) | CACCTCCCTCCG | 8.264 (80.1) | -2.669 | Yes |
| rs12977601 | 46551622 | TEF-1 | - | 46551618 | 46551629 | CACCTCCTTCCG | 5.595 (73.3) | CACCTCCCTCCG | 8.264 (80.1) | -2.669 | Yes |
| rs11466314 | 46552076 | Thing1-E47 | - | 46552070 | 46552079 | GGTTTGTCAA | 4.192 (74.7) | GGTCTGTCAA | 7.836 (85.4) | -3.644 | No |
| rs1800469 | 46552136 | E74A | - | 46552135 | 46552141 | CCTGAGG | 2.498 (67.4) | CCTGAAG | 6.627 (83.7) | -4.129 | No |
| rs1800468 | 46552427 | CREB | - | 46552424 | 46552435 | GGTGGTGATGTT | 7.935 (83.9) | GGTGGTGACGTT | 12.023 (94.9) | -4.088 | No |
| rs1800468 | 46552427 | bZIP910 | - | 46552425 | 46552431 | GTGATGT | 5.723 (80.9) | GTGACGT | 10.348 (95.7) | -4.625 | No |
| rs1054797 | 46552599 | ARNT | + | 46552596 | 46552601 | CACAGG | 1.872 (66.2) | CCCGTG | 6.182 (83.4) | -4.31 | Yes |
| rs1054797 | 46552599 | ARNT | - | 46552596 | 46552601 | CCTGTG | 1.942 (66.4) | CCCGTG | 6.182 (83.4) | -4.24 | Yes |
| rs1054797 | 46552599 | Max | - | 46552595 | 46552604 | AAGCCTGTGC | 3.101 (69.5) | AAGCCCGTGC | 7.230 (82.1) | -4.129 | Yes |
| rs1054797 | 46552599 | c-REL | + | 46552598 | 46552607 | CAGGCTTTCG | 4.153 (73.0) | CGGGCTTTCG | 8.112 (85.1) | -3.959 | Yes |
| rs1054797 | 46552599 | n-MYC | + | 46552596 | 46552601 | CACAGG | 3.017 (72.0) | CACGGG | 6.133 (83.6) | -3.116 | Yes |
| rs1054797 | 46552599 | n-MYC | - | 46552596 | 46552601 | CCTGTG | 2.370 (69.6) | CACGGG | 6.133 (83.6) | -3.763 | Yes |
| rs1054797 | 46552599 | p65 | + | 46552598 | 46552607 | CAGGCTTTCG | 3.732 (69.8) | CGGGCTTTCG | 7.822 (80.5) | -4.09 | Yes |
| rs1800820 | 46552615 | Myf | + | 46552605 | 46552616 | GGGCAGTTGGAG | 8.255 (82.2) | CGCCAACTGCCC | 5.986 (77.1) | 2.269 | Yes |
| rs2317130 | 46553514 | Snail | + | 46553514 | 46553519 | Cagatg | 7.278 (87.6) | Tagatg | 2.596 (70.9) | 4.682 | Yes |
| rs2317130 | 46553514 | Tal1beta-E47S | + | 46553510 | 46553521 | agatCagatggt | 9.456 (85.1) | agatTagatggt | 4.704 (74.7) | 4.752 | Yes |
| rs2317130 | 46553514 | Tal1beta-E47S | - | 46553512 | 46553523 | ccaccatctGat | 11.614 (89.8) | ccaccatctAat | 6.863 (79.4) | 4.751 | Yes |
| rs11466311 | 46553588 | CFI-USP | - | 46553579 | 46553588 | Gggatcaagg | 8.906 (82.6) | Aggatcaagg | 4.262 (70.6) | 4.644 | No |
| rs11466310 | 46553698 | AML-1 | + | 46553692 | 46553700 | cttgctAtt | 1.650 (68.9) | cttgctGtt | 6.332 (82.2) | -4.682 | No |
| rs11466310 | 46553698 | Androgen | + | 46553685 | 46553706 | cttggcacttgctAttccctct | 7.046 (75.2) | cttggcacttgctGttccctct | 11.411 (82.2) | -4.365 | No |
| rs11466310 | 46553698 | Myf | - | 46553690 | 46553701 | gaaTagcaagtg | 3.352 (71.1) | gaaCagcaagtg | 7.352 (80.2) | -4 | No |
| rs11466310 | 46553698 | NRF-2 | - | 46553690 | 46553699 | aTagcaagtg | 4.725 (70.8) | aCagcaagtg | 7.823 (80.7) | -3.098 | No |
| rs11466310 | 46553698 | Spz1 | - | 46553695 | 46553705 | gagggaaTagc | 3.807 (71.6) | gagggaaCagc | 7.317 (80.8) | -3.51 | No |
| rs11666933 | 46554093 | CF2-II | + | 46554089 | 46554098 | ggatAtattt | 6.213 (80.2) | aaataCatcc | 1.968 (69.7) | 4.245 | No |
| rs13345981 | 46554303 | Chop-cEBP | + | 46554295 | 46554306 | aagtgaaaCctc | 7.396 (80.5) | aagtgaaaTctc | 9.588 (86.0) | -2.192 | No |
| rs13345981 | 46554303 | Irf-1 | + | 46554293 | 46554304 | aaaagtgaaaCc | 16.637 (98.6) | aaaagtgaaaTc | 12.978 (90.4) | 3.659 | No |
| rs13345981 | 46554303 | Irf-2 | + | 46554292 | 46554309 | taaaagtgaaaCctcagc | 12.732 (82.8) | taaaagtgaaaTctcagc | 9.745 (77.5) | 2.987 | No |
| rs13345981 | 46554303 | bZIP910 | - | 46554302 | 46554308 | ctgagGt | 7.070 (85.2) | ctgagAt | 2.446 (70.4) | 4.624 | No |
| **MBL-2 (Mannose binding lectin-2)** | | | | | | | | | | | |
| rs3737613 | 11041105 | bZIP910 | + | 11041100 | 11041106 | CTGATCT | 2.446 (70.4) | CTGATGT | 7.070 (85.2) | -4.624 | No |
| rs12751228 | 11042418 | AML-1 | - | 11042417 | 11042425 | tttgaggTg | 6.411 (82.4) | tttgaggGg | 2.565 (71.5) | 3.846 | No |
| rs12751228 | 11042418 | Hunchback | + | 11042418 | 11042427 | Acctcaaaaa | 6.524 (80.6) | Ccctcaaaaa | 8.109 (85.5) | -1.585 | No |
| rs12751228 | 11042418 | NF-kappaB | + | 11042410 | 11042419 | gagactccAc | 3.068 (71.0) | gGggagtctc | 7.535 (81.8) | -4.467 | No |
| rs12751228 | 11042418 | NF-kappaB | - | 11042410 | 11042419 | gTggagtctc | 2.854 (70.5) | gGggagtctc | 7.535 (81.8) | -4.681 | No |
| rs12751228 | 11042418 | Snail | - | 11042417 | 11042422 | gaggTg | 6.770 (85.8) | gaggGg | 2.053 (68.9) | 4.717 | No |

**Table 2.** Contd.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs12751228 | 11042418 | p50 | - | 11042409 | 11042419 | gTggagtctca | 4.405 (70.8) | gGggagtctca | 8.573 (80.8) | -4.168 | No |
| rs12737148 | 11043642 | NF-kappaB | + | 11043641 | 11043650 | gCgagactcc | 2.326 (69.2) | gGgagactcc | 7.007 (80.5) | -4.681 | No |
| rs12737148 | 11043642 | c-REL | + | 11043641 | 11043650 | gCgagactcc | 3.527 (71.1) | gGgagactcc | 6.508 (80.2) | -2.981 | No |
| rs12737148 | 11043642 | c-REL | - | 11043642 | 11043651 | tggagtctcG | 2.489 (67.9) | tggagtctcC | 6.535 (80.3) | -4.046 | No |
| rs12737148 | 11043642 | p65 | - | 11043642 | 11043651 | tggagtctcG | 4.236 (71.1) | tggagtctcC | 8.403 (82.0) | -4.167 | No |
| rs12757026 | 11043657 | Broad-complex_1 | + | 11043653 | 11043666 | ctcaAaaaaaaaaa | 9.103 (83.1) | ctcaCaaaaaaaaa | 6.781 (77.0) | 2.322 | No |
| rs12757026 | 11043657 | Broad-complex_1 | + | 11043654 | 11043667 | tcaAaaaaaaaaaa | 8.220 (80.8) | tcaCaaaaaaaaaa | 4.851 (71.9) | 3.369 | No |
| rs12757026 | 11043657 | Broad-complex_4 | + | 11043655 | 11043665 | caAaaaaaaaa | 7.688 (83.0) | caCaaaaaaaa | 5.595 (76.3) | 2.093 | No |
| rs12757026 | 11043657 | Broad-complex_4 | + | 11043656 | 11043666 | aAaaaaaaaaa | 7.688 (83.0) | aCaaaaaaaaa | 4.775 (73.7) | 2.913 | No |
| rs12757026 | 11043657 | HFH-2 | - | 11043656 | 11043667 | ttttttttttTt | 9.407 (85.7) | ttttttttttGt | 6.539 (80.0) | 2.868 | No |
| rs12757026 | 11043657 | HFH-2 | - | 11043657 | 11043668 | tttttttttttT | 9.407 (85.7) | tttttttttttG | 4.986 (76.9) | 4.421 | No |
| rs12757026 | 11043657 | HMG-IY | + | 11043653 | 11043668 | ctcaAaaaaaaaaaaa | 8.479 (82.2) | ctcaCaaaaaaaaaaa | 4.806 (73.7) | 3.673 | No |
| rs12757026 | 11043657 | HMG-IY | + | 11043655 | 11043670 | caAaaaaaaaaaaaaa | 11.524 (89.1) | caCaaaaaaaaaaaaa | 9.164 (83.7) | 2.36 | No |
| rs12757026 | 11043657 | HMG-IY | + | 11043656 | 11043671 | aAaaaaaaaaaaaaaa | 10.932 (87.8) | aCaaaaaaaaaaaaaa | 7.847 (80.7) | 3.085 | No |
| rs12757026 | 11043657 | Hunchback | + | 11043652 | 11043661 | tctcaAaaaa | 6.747 (81.3) | tctcaCaaaa | 2.659 (68.7) | 4.088 | No |
| rs12757026 | 11043657 | Hunchback | + | 11043653 | 11043662 | ctcaAaaaaa | 6.899 (81.7) | ctcaCaaaaa | 4.091 (73.1) | 2.808 | No |
| rs12757026 | 11043657 | Hunchback | + | 11043657 | 11043666 | Aaaaaaaaaa | 9.443 (89.6) | Caaaaaaaaa | 11.028 (94.5) | -1.585 | No |
| rs12757026 | 11043657 | SOX17 | - | 11043655 | 11043663 | ttttttTtg | 1.814 (68.2) | ttttttGtg | 6.355 (82.9) | -4.541 | No |
| rs12757026 | 11043657 | SQUA | + | 11043654 | 11043667 | tcaAaaaaaaaaaa | 9.768 (85.7) | tcaCaaaaaaaaaa | 7.402 (80.8) | 2.366 | No |
| rs12757026 | 11043657 | SQUA | + | 11043655 | 11043668 | caAaaaaaaaaaaa | 7.192 (80.4) | caCaaaaaaaaaaa | 2.981 (71.7) | 4.211 | No |
| rs12757026 | 11043657 | SQUA | + | 11043656 | 11043669 | aAaaaaaaaaaaaa | 6.879 (79.8) | aCaaaaaaaaaaaa | 10.972 (88.2) | -4.093 | No |
| rs12402525 | 11044252 | CFI-USP | - | 11044245 | 11044254 | ggGttcacgc | 10.323 (86.3) | ggCttcacgc | 5.642 (74.2) | 4.681 | No |
| rs12402525 | 11044252 | CREB | - | 11044244 | 11044255 | cggGttcacgcc | 7.593 (83.0) | cggCttcacgcc | 6.008 (78.7) | 1.585 | No |
| rs11577987 | 11044286 | ARNT | + | 11044282 | 11044287 | ctcgCg | 1.942 (66.4) | ctcgTg | 6.182 (83.4) | -4.24 | No |
| rs11577987 | 11044286 | ARNT | - | 11044282 | 11044287 | cGcgag | 2.864 (70.1) | ctcgTg | 6.182 (83.4) | -3.318 | No |
| rs11577987 | 11044286 | n-MYC | + | 11044282 | 11044287 | ctcgCg | 4.627 (78.0) | ctcgTg | 7.369 (88.2) | -2.742 | No |
| rs11577987 | 11044286 | n-MYC | - | 11044282 | 11044287 | cGcgag | 1.684 (67.0) | ctcgTg | 7.369 (88.2) | -5.685 | No |
| rs754311 | 11044480 | Snail | + | 11044477 | 11044482 | cagCtt | 2.632 (71.0) | cagGtt | 7.278 (87.6) | -4.646 | No |
| rs754311 | 11044480 | Snail | - | 11044477 | 11044482 | aaGctg | 1.417 (66.7) | cagGtt | 7.278 (87.6) | -5.861 | No |

example of this is the selection of sickle cell anemia, thalassemia, and glucose-6-phosphate dehydrogenase deficiency against the pressure of malaria. Falciparum malaria is an example of evolutionary selection. Certain traits are positively selected in malaria endemic region causing aminoacid change, modified protein in such a way as it provide resistance against malaria. Several host genes selects non-synonymous amino acid changes. What happens in regulatory region? It will be clear from this example. Sickle-cell trait, non-synonymous amino acid changes (HbAS) protects against severe outcomes of malaria, and it is highly prevalent in East African populations who live in malaria endemic areas (Aidoo et al., 2002). iNOS, –1173 C→T polymorphism is as prevalent as the sickle-cell trait in the two East African populations studied (Figure 1). The –1173 C→T polymorphism is significantly associated with protection from both cerebral malaria and severe malarial anemia, and with increased NO production *in vivo*. The C to T change predicts the creation of a new sequence recognition site for the GATA-1 or GATA-2 transcription factors and is responsible for, an increased degree of transcription from the -1173 C→T promoters (Heinemeyer et al., 1998). Some investigators

**Table 3.** Conservation pattern in the promoter region of genes involved in malaria pathogenesis and their comparison with other genes.

| Gene | Motif-1 | Motif-2 | Motif-3 |
|---|---|---|---|
| G6PD |  |  |  |
| Cdk-7 |  |  |  |
| Ldha |  |  |  |
| GSTP1 |  |  |  |
| iNOS |  |  |  |
| iNOS |  |  |  |

**CDK7-** *H. sapiens, P. troglodytes, C. lupus, B. taurus, M. musculus, R. norvegicus, D. rerio, D. melanogaster and C.elegans. LDHA, GSTP1,* **iNOSb-** *Human, rat, mouse.* **iNOSc-** *H. sapiens, P. troglodytes, C. lupus, B. taurus, M. musculus, R. norvegicus, D. rerio) with different length of sequences.* **G6PD-** *H. sapiens, C. lupus G6PD, B. taurus G6PD, M. musculus G6pdx, R. norvegicus G6pdx, D. rerio LOC100148915, D. rerio wu:fj78b06, D. melanogaster Zw gene, C. elegans G6PD, S. pombe zwf1.*

have reported significant protection by the -954 G→C or CCTTT long repeat polymorphisms genes from *NOS2* promoters with deletions in the proximal portion of the *NOS2* promoter. In the light of these facts and our finding in the supplementary data Table 2, our hypothesis is that in the regulatory region, those variations are selected, which alters the expression of genes in such a way as to reduce the infection of parasite. High level of iNOS is beneficial against the infection of *Plasmodium falciparum,* as it reduces the TNF-alpha toxicity and expression of PECAM1, ICAM1 (adhesion molecules). Changes in their structure or levels in individuals can influence the outcome of infection. Interestingly, a SNP of the *PECAM1* gene (rs668, exon 3, C/G) with low minor allele frequency in populations of the endemic region compared to the non-endemic region exhibited differential association with disease in these regions; the G allele was a risk factor for malaria in the endemic region, but exhibited significant association with protection from disease in the non-endemic region (Sinha et al., 2008, 2009). So variations occur according to disease prevalence and selected if it play beneficial role against the disease.

In mammalian genome, methylation of cytosine

**Table 4.** Comparison of percentage conservation in the promoters of different genes in different organisms. Polymorphisms in conserved region are more likely to influence the regulation of gene.

| Gene | % Conservation | | |
|---|---|---|---|
| | Logo1 | Logo2 | Logo3 |
| LDHA (Human, rat, mouse) | 87.23 | 81.63 | 65.21 |
| HSPCB* | 38 | 32 | 30 |
| GSTP1 (Human, rat, mouse) | 85 | 58.62 | 69.23 |
| iNOS** | 72 | 62 | 17.64 |

**H .sapiens*, *P. troglodytes*, *C .lupus*, *B. taurus*, *M. musculus*, *R. norvegicus*, *D. rerio*, *D. melanogaster* and *C. elegans.* ** *H. sapiens*, *P. troglodytes*, *C .lupus*, *B .taurus*, *M. musculus*, *R. norvegicus* and *D. rerio.*

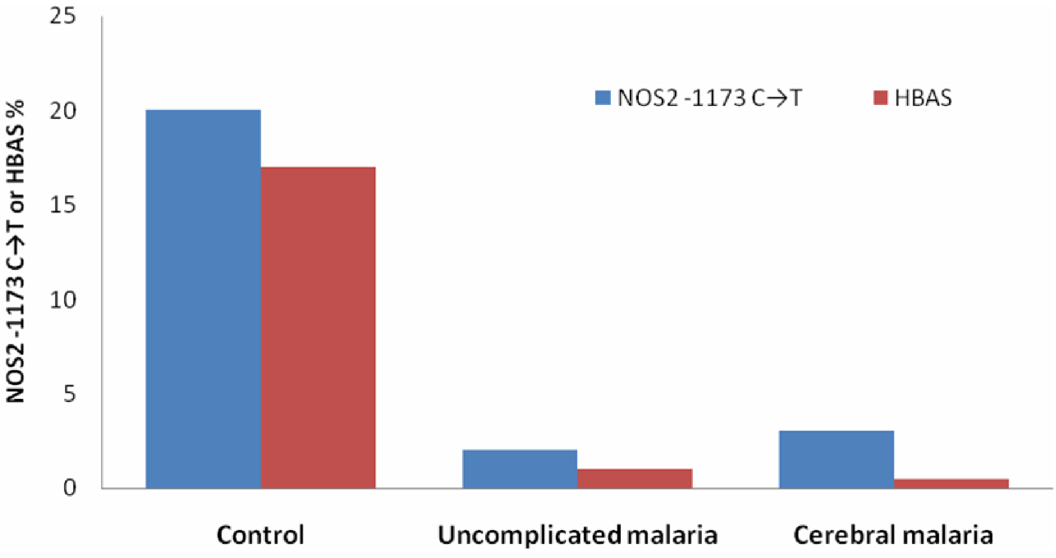


**Figure 1.** NOS2 –1173 C→T genotype or sickle cell trait (HBAS) relative to disease category in Tanzanian children. Source: Maurine et al. (2002), Lancet [13].

in CpG dinucleotide play role in outcome of disease. Half of the genes in mammalian genome in the proximal promoter regions (0.5 to 4 kb) have CpG Islands. In normal cells, these are unmethylated whereas DNA methylation in the promoter of certain genes is associated with transcriptional silencing. Methylation affects gene expression directly by interfering with transcription factor binding and/or indirectly by recruiting histone deacetylases through methyl-DNA binding proteins (Wei-Guo et al., 2003). The reports indicate that hypermethylation may interfere with Sp1/Sp3 binding. Methylation outside the consensus of Sp1 element induces a significant decrease in Sp1/Sp3 binding.

## Polymorphisms in transcription factor binding sites play role in expression divergence, fitness and evolution

In case of HIV, there is evidence of positive selection in early HIV-1 infection which appears to be driven in many cases by escape from early cytotoxic T lymphocyte (CTL) responses via mutations in the APOBEC sequence, suggesting a role for APOBEC in determining the pathway of immune escape (Wood et al., 2009). The study holds significance in understanding the regulation of gene expression at the level of TF and TF binding sites using the aforesaid tools. Further, the variations in the DNA region alter the interaction thereby modulating the host parasite interaction or the genetic disorders. Wittkopp (2010) reported that, natural selection has played some role in expression divergence, but the relative frequency of adaptive and neutral changes remains unclear. Bradley et al. (2010) observed differences in transcription factor binding between species that were similar in regions of the genome thought to act as functional *cis*-regulatory elements and those thought to have no such activity. The majority of changes in transcription factor binding may have little to no effect on gene expression. All changes in transcription factor binding are neutral. They also reported that changes in transcription factor binding alter gene expression and may have contributed to adaptive evolution (Sequence variation to transcription factor binding, transcription factor binding to gene expression, gene expression to phenotypic variation, and phenotypic variation to fitness in the wild).

Interactions of TF to TFBS are mediated by the DNA binding domain of the transcription factor and the nucleotides within the *cis*-regulatory DNA to which it binds. Most *cis*-regulatory sequences are bound by more than one type of transcription factor, and most transcription factors bind to the *cis*-regulatory sequences from several genes. The recruitment of different combinations of transcription factors to different genes allows expression of each gene to be regulated independently. Mutations that alter the activity or availability of transcription factors, as well as mutations that alter the *cis*-regulatory sequences to which they bind, can change gene expression. Both types of changes contribute to evolution; however, studies from a variety of organisms suggest that mutations affecting *cis*-regulatory activity are the predominant source of expression divergence between species (Tirosh et al., 2009, Wittkopp et al., 2008, Wilson et al., 2008, Graze et al., 2009). Changes in gene expression often alter phenotypes; mutations that affect gene expression can affect fitness and contribute to adaptive evolution. Bradley et al. (2010) consistently detected binding of the same transcription factors to regions of DNA in *D. melanogaster* and *D. yakuba* that have a common evolutionary origin; however, the relative affinity of these binding sites often differed between species. This suggests that evolutionary changes in the DNA sequence of *cis*-regulatory regions have occurred that alter the strength of the interaction between transcription factors and their binding sites without eliminating binding.

We have checked the conservation pattern in the promoter region and found that most of the region in the promoter is conserved. It is also suggestive of the fact that if a polymorphism or any variation in the DNA sequence occurs in the conserved region then it effects the interaction of TF with TFBS largely. Table 4 shows comparison of conservation patterns in the promoter region of some genes viz., *G6PD, CDK7, LDHA, GSTP1 and iNOS*. Understanding the conservation and change of regulatory sequences is critical to our knowledge of the unity as well as diversity of animal development and phenotypes. It can be deciphered from this data that as the number of organisms increases, the percent conservation decreases although certain position in the sequence remains constant throughout. These conserved sequences are thought to be the essential sites that are controlling the regulatory activity for the normal expression of the gene. Genome tries to preserve these conserved sequences irrespective of whether the gene is a housekeeping gene or a highly regulatory gene. Kim et al. (2009) reported that the evolution of sequences involved in the regulation of body patterning in *Drosophila* embryo. The mutations of nucleotides within a binding site are constrained by evolutionary forces to preserve the site's binding affinity to the cognate transcription factor. Functional binding sites are frequently destroyed during evolution and the rate of loss across evolutionary spans is roughly constant.

In this report, we focus those polymorphisms of TFBS in the regulatory region of host genes and hypothesize that these variation increases the susceptibility/resistance to a particular disease by alteration of gene product in the cell. Therefore, we have concluded that 124 promoter polymorphisms in the 9 genes involved in malaria pathogenesis play important role in susceptibility to falciparum malaria. The individual SNP has no contribution in risk of disease but SNP-SNP interaction is associated with risk of disease. SNP-SNP interaction analysis of the predicted SNPs will be important in case/control study. This is the first report in which an attempt has been made to associate SNP, variations in promoter sequences in relation with gene deregulation and susceptibility/resistance to various genetic and infectious diseases. A given transcription factor can have different effects on different promoters and multiple transcription factor affects the activity of a single gene. The regulatory sequence of DNA (*cis* elements) tries to preserve itself. A large number of diseases have been reported in which, variations in regulatory region is responsible for outcome of disease. Polymorphisms and methyl group resulting from methylation of promoter, affects the interaction of TF to TFBS in the studied gene (involved in malaria pathogenesis). DNA sequence variation in the regulatory region and aberrant promoter hypermethylation interfere with regulation of gene expression at transcription level, altering the level of gene product and thus responsible for susceptibility/ resistance to genetic and infectious disease. Further studies are required to validate the predicted SNPs in Indian and other populations. Haplotype and individual SNP analysis does not give the clear picture of case-control study. Therefore, SNP-SNP interaction (gene-gene interaction) and SNP-environment interactions (gene environment interaction) are required for the analysis.

## ACKNOWLEDGMENTS

The authors express their gratitude to Ministry of Science and Technology, Government of India and University Grants Commission, New Delhi for providing financial support to the department of Biochemistry, Dr. R.M.L. Avadh University, Faizabad, U. P., India, under the DST-FIST and UGC-SAP schemes, for developing infrastructure to carry out work on such aspects.

## REFERENCES

Aidoo M, Terlouw DJ, Kolczak MS, McElroy PD, Kuile FO, Kariuki S, Nahlen BL, Lal AA, Udhayakumar V (2002). Protective effects of the sickle cell gene against malaria morbidity and mortality. Lancet, 359: 1311–1312.

Bradley RK, Li XY, Trapnell C, Davidson S, Pachter L, Chu HC, Tonkin LA, Biggin MD, Eisen MB (2010). Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related *Drosophila* species. PLoS Biol., 8: e1000343.doi:10.1371/journal.pbio.1000343.

Frisch M, Frech K, Klingenhoff A, Cartharius K, Liebich I, Werner T (2002). *In silico* prediction of scaffold/matrix attachment regions in large genomic sequences. Genome Res., 12: 349–354.

Graze RM, McIntyre LM, Main BJ, Wayne ML, Nuzhdin SV (2009). Regulatory Divergence in *Drosophila melanogaster* and *D. simulans*, a Genome-wide Analysis of Allele-specific Expression. Genetics, 183: 547–551.

Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, Ignatieva EV, Ananko EA, Podkolodnaya OA, Kolpakov FA, Podkolodny NL, Kolchanov NA (1998). Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. Nucleic Acids Res., 26: 362–367.

Higuichi T, Seki N, Kamizono S, Yamada A, Kimura A, Kato H, Itoh K (1998). Polymorphism of the 5'-flanking region of the human tumor necrosis factor TNF-A gene in Japanese. Tissue Antigen, 51: 605-612.

Hobbs MR, Udhayakumar V, Levesque MC, Booth J, Roberts JM, Tkachuk AN (2002). A new *NOS2* promoter polymorphism associated with increased nitric oxide production and protection from severe malaria in Tanzanian and Kenyan children. Lancet, 360: 1468–1475.

Jones PA, Stephen BB (2002). The fundamental role of epigenetic events in cancer. Nat. Genet., 3: 415-428.

Kim J, Xin H, Sinha S (2009). Evolution of Regulatory Sequences in 12 Drosophila species PLoS Genet., 5(1): e1000330. doi:10.1371/journal.pgen.1000330.

Klingenhoff A, Frech K, Quandt K, Werner T (1999). Functional promoter modules can be detected by formal models independent of overall nucleotide sequence similarity. Bioinformatics, 15: 180–186.

Klingenhoff A, Frech K, Werner T (2002). Regulatory modules shared within gene classes as well as across gene classes can be detected by the same *in silico* approach. *In Silico* Biol., 2: S17–S26.

Krivan W, Wasserman WW (2001). A predictive model for regulatory sequences directing liver-specific transcription. Genome Res., 11: 1559–1566.

Murakami K, Kojima T, Sakaki Y (2004). Assessment of clusters of transcription factor binding sites in relationship to human promoter, CpG islands and gene expression. BMC Genom., 5: 16.

Phairote T, Ohashi J, Patarapotikul J, Kimura R, Nuchnoi P (2008). A Functional single nucleotide polymorphism in the CR1 promoter region contributes to protection against cerebral malaria. J. Infect. Dis., 198: 1880-1891.

Pilpel Y, Sudarsanam P, Church GM (2001). Identifying regulatory networks by combinatorial analysis of promoter elements. Nat. Genet., 29: 153–159.

Robertson KD (2002). DNA methylation and chromatin — Unraveling the tangled web. Oncogene, 21: 5361–5369.

Sinha S, Jha GN, Anand P, Qidwai T, Pati SS, Mishra SK, Tyagi PK, Sharma SK., IGVC, Venkatesh V, Habib S (2009). CR1 levels and gene polymorphisms exhibit differential association with falciparum malaria in regions of varying disease endemicity. Hum. Immunol., 70: 244-250.

Sinha S, Mishra SK, Sharma S, Patibandla PK, Mallick PK, Sharma SK, Mohanty S, Pati SS, Mishra SK, Ramteke BK, Bhatt RM, Joshi H, Dash AP, Ahuja RC, Awasthi S, IGVC, Venkatesh V, Habib S (2008). Polymorphisms of *TNF*-enhancer and gene for FcγRIIa correlate with the severity falciparum malaria in the ethnically diverse Indian population. Malaria J., 7: 12.

Stormo GD (2000). DNA binding sites: Representation and discovery. Bioinformatics, 16: 16–23.

Tirosh I, Reikhav S, Levy AA, Barkai N (2009). A yeast hybrid provides insight into the evolution of gene expression regulation. Science, 324: 659–662.

Wasserman WW, Fickett JW (1998). Identification of regulatory regions which confer muscle-specific gene expression. J. Mol. Biol., 278: 167–181.

Wei-Guo Z, Srinivasan K, Dai Z, Duan W, Druhan LJ, Haiming D, Lisa Y, Villalona-Calero MA, Christoph P, Gregory A (2003). Otterson. Methylation of adjacent CpG Sites affects Sp1/Sp3 binding and activity in the *p21*$^{Cip1}$ promoter. Mol. Cell. Biol., 27: 4056-4065.

Werner T (1999). Models for prediction and recognition of eukaryotic promoters. Mammalian Genom., 10: 168–175.

Wilson MD, Barbosa-Morais NL, Schmidt D, Conboy CM, Vanes L, Tybulewicz VLJ, Fisher EMC, Tavaré S, Odom DT (2008). Species-specific transcription in mice carrying human chromosome 21. Science, 322: 434–438.

Wittkopp PJ (2010). Variable Transcription Factor Binding: A Mechanism of Evolutionary Change. PLoS Biol., 8(3): e1000342. doi:10.1371/journal.pbio. 1000342.

Wood N, Bhattacharya T, Keele BF, Giorgi E, Liu M, *Gaschen B,* Daniels M, Ferrari G, Haynes BF, McMichael A, Shaw GM, Hahn BH, Korber B, Seoighe C (2009). HIV Evolution in Early Infection: Selection Pressures, Patterns of Insertion and Deletion, and the impact of APOBEC. PLoS Pathog., 5(5): e1000414. doi:10.1371/journal.ppat.

Xueping Yu, Jimmy L, Zack DJ, Qian J (2006). Computational analysis of tissue-specific combinatorial gene regulation: Predicting interaction between transcription factors in human tissues. Nucleic Acids Res., 34: 4925–4936.

Zhu Z, Pilpel Y, Church GM (2002). Computational identification of transcription factor binding sites via a transcription-factor-centric clustering (TFCC) algorithm. J. Mol. Biol., 318: 71–81.