**Name of Journal: *World Journal of Hematology***

**ESPS Manuscript NO: 27052**

**Manuscript Type: ORIGINAL ARTICLE**

***Basic Study***

**Identifying changes in punitive transcriptional factor binding sites from regulatory single nucleotide polymorphisms that are significantly associated with disease or sickness**

Buroker NE. Identifying changes in punitive TFBS from rSNPs

**Norman E Buroker**

**Norman E Buroker**, Department of Pediatrics, University of Washington, Seattle, WA 98195, United States

**Author contributions:** Buroker NE analyzed all the data and wrote the manuscript.

**Institutional review board statement:** The manuscript did not use human or animal subjects.

**Informed consent statement:** The manuscript did not use human or animal subjects.

**Conflict-of-interest statement**: To the best of my knowledge, no conflict of interest exists.

**Open-Access:** This article is an open-access article which was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

**Manuscript source:** Invited manuscript

**Correspondence to**: **Norman E Buroker, PhD,** Department of Pediatrics, University of Washington, 1959 NE Pacific Street, Seattle, WA 98195, United States. nburoker@u.washington.edu

**Telephone:** +1-206-6160472

**Fax:** +1-206-6160471

**Received:** May 8, 2016

**Peer-review started:** May 9, 2016

**First decision:** June 13, 2016

**Revised:** June 24, 2016

**Accepted:** August 11, 2016

**Article in press:**

**Published online:**

**Abstract**

**AIM:** To identify punitive transcriptional factor binding sites (TFBS) from regulatory single nucleotide polymorphisms (rSNPs) that are significantly associated with disease.

**METHODS:** The genome-wide association studies (GWAS) have provided us with nearly 6500 disease or trait-predisposing SNPs where 93% are located within non-coding regions such as gene regulatory or intergenic areas of the genome. In the regulatory region of a gene, a SNP can change the DNA sequence of a transcriptional factor (TF) motif and in turn may affect the process of gene regulation. SNP changes that affect gene expression and impact gene regulatory sequences such as promoters, enhancers, and silencers are known as rSNPs. Computational tools can be used to identify unique punitive TFBS created by rSNPs that are associated with disease or sickness. Computational analysis was used to identify punitive TFBS generated by the alleles of these rSNPs.

**RESULTS:** rSNPs within nine genes that have been significantly associated with disease or sickness were used to illustrate the tremendous diversity of punitive unique TFBS that can be generated by their alleles. The genes studied are the adrenergic, beta, receptor kinase 1 (*ADRBK1*), the v-akt murine thymoma viral oncogene homolog 3 (*AKT3*), the activating transcription factor 3 (*ATF*3), the type 2 deiodeinase gene (DIO2), the endothetal Per-Arnt-Sim (PAS) domain protein 1 (*EPAS1*), the lysosomal acid lipase A (*LIPA*), the signal Transducer and Activator of Transcription 4 (*STAT4),* the thromboxane A2 receptor (TBXA2R) and the vascular endothelial growth factor A (*VEGFA*). From this sampling of SNPs among the nine genes, there are 73 potential unique TFBS generated by the common alleles compared to 124 generated by the minor alleles indicating the tremendous diversity of potential TFs that are capable of regulating these genes.

**CONCLUSION:** From the diversity of unique punitive binding sites for TFs, it was found that some TFs play a role in the disease or sickness being studied.

**Key words**: Regulatory single nucleotide polymorphisms; Alleles; Transcriptional factors; Transcriptional factor binding sites; Linkage disequilibrium;Disease or sickness

**© The Author(s) 2016.** Published by Baishideng Publishing Group Inc. All rights reserved.

**Core tip:** Disease or trait-predisposing single nucleotide polymorphisms (SNPs) in or near genes can alter the TFBS for the TFs regulating the gene; thereby affecting the health of an individual. In this report, the disease or sickness associated regulatory SNPs (rSNPs) within a sampling of nine human genes were studied with respect to the alterations in transcriptional factor binding sites (TFBS). From this sampling there were 73 punitive unique TFBS generated by the common rSNP alleles compared to 124 generated by the minor alleles indicating the tremendous diversity of potential TFs that are capable of affecting the health of person.

Buroker NE. Identifying changes in punitive transcriptional factor binding sites from regulatory single nucleotide polymorphisms that are significantly associated with disease or sickness. *World J Hematol* 2016; In press

**INTRODUCTION**

The genome-wide association studies (GWAS) have over the past decade provided us with nearly 6500 disease or trait-predisposing single nucleotide polymorphisms (SNPs). Only seven percent of these SNPs are located in protein-coding regions of the genome[1,2] while the remaining 93% are located within non-coding regions[3,4] such as gene regulatory or intergenic areas of the genome. Much attention has been drawn to SNPs that occur in the putative regulatory region of a gene where a single nucleotide change in the DNA sequence of a potential transcriptional factor (TF) motif may affect the process of gene regulation[5-7]. A nucleotide change in a transcriptional factor binding site (TFBS) can have multiple consequences. Since a TF can usually recognize a number of different binding motifs in a gene, the SNP may not change the TFBS interaction with the TF and consequently not alter the process of gene expression. In other cases the nucleotide change may increase or decrease the TF’s ability to bind DNA which would result in allele-specific gene expression. In some cases a nucleotide change may eliminate the natural binding motif or generate a new binding site (BS) as a result the gene is no longer regulated by the original TF[8,9]. Single nucleotide changes that affect gene expression by impacting gene regulatory sequences such as promoters, enhancers, and silencers are known as regulatory SNPs (rSNPs)[5,6,10,11]. Therefore, functional rSNPs in TFBS may result in differences in gene expression, phenotypes and susceptibility to environmental exposure[7]. Examples of rSNPs associated with disease susceptibility are numerous and several reviews have been published[7,12-16]. Advances in understanding the functional relevance of SNPs in non-coding regions of the human genome using epi-genomics and genome engineering have been recently reviewed[17]. Computational tools can be used to identify punitive TFBS created by rSNPs which are associated with disease or sickness[18]. To this end, computational analysis has been used to identify punitive or potentially unique TFBS generated by the alleles of rSNPs[19] where unique TFBS occur with only one of the two rSNP alleles.

In this report, rSNPs within a sample of nine human genes (Table 1) which have been significantly associated with disease or sickness were selected to illustrate the tremendous diversity of unique punitive TFBS that can be generated by SNP alleles (Table 2)[8,9,20-27]. The SNP alleles from these reports were found to share common TFBS between alleles but each SNP allele can also create unique TFBS only for that allele (Table 2). As an example in Table 2, the rs948988 ADRBK1-G allele creates two potential unique TFBS for the Kruppel-like factors 1 and 4 (KLF1,4) TFs that do not occur with the alternate ADRBK1-A allele while the ADRBK1-A allele creates ten other punitive unique TFBS not found with ADRBK1-G allele. Many of the rSNPs have been reported to be in linkage disequilibrium (LD) (Table 1), where LD is considered to be the non-random association of SNP alleles within a gene. LD between SNPs in the regulatory region of a gene can indicate strong associations of certain haplotypes and TFBS with sickness or disease[28].

**MATERIALS AND METHODS**

***Identifying TFBS***

The JASPAR CORE database[29,30] and ConSite[31] were used to identify the TFBS in this study. JASPAR is a collection of transcription factor DNA-binding preferences used for scanning genomic sequences where ConSite is a web-based tool for finding cis-regulatory elements in genomic sequences. The Vector NTI Advance 11 computer program (Invitrogen, Life Technologies) was used to locate SNPs andTFBS within all genes listed in Table 1.

**RESULTS**

The protein and gene symbol, chromosome position of the gene, SNP number and location within the gene and nucleotide (mutation) change are listed in Table 1. Also listed is whether or not linkage disequilibrium occurs between the SNPs within each gene (Table 1). Nine genes, ethnic groups, disease or sickness, SNPs and alleles as well as potential unique TFBS per allele that have been reported are found in Table 2. Not all of the SNPs for each gene are listed in the tables but can be found in the accompanying references (Table 1). From Table 2, it can be seen that there occur incidences when the SNP common allele does not have any unique punitive TFBS while the minor allele provide several (*e.g.,* rs12885300 in DIO2; rs2238632 and rs2238634 in TBXA2R). There are other incidences where the SNP common allele provides one or two unique punitive TFBS while the minor alleles again provide several (*e.g.,* rs948988 in ADRBK1; rs4590656 in AKT3; rs11119982 in ATF3; rs8179673 in STAT4 and rs2238631 in TBXA2R). A near balance between SNP alleles in unique punitive TFBS can also be found in the table (*e.g.,* rs4370946 in ADRBK1; rs10157763 and rs2125230 in AKT3; rs3125289 in ATF3; rs6756667 in EPAS1; rs34357231 and rs3025039 in VEGFA). The minor allele of the SNP usually generates more punitive unique TFBS than the common allele (*e.g.,* rs948988 in ADRBK1; rs1412444 in LIPA and rs8179673 in STAT4). In fact, from this sampling of SNPs among the nine genes, there are 73 potential unique TFBS generated by the common alleles compared to 124 by the minor alleles (Table 2).

**DISCUSSION**

The possible relationship of these punitive unique TFBS to disease and sickness has previously been discussed for each gene in the accompanying references (Table 1). The use of rSNPs that are in LD within a gene to identify punitive TFBS can be illustrated with a few SNPs from these nine genes. The *ADRBK1* gene, which transcribes the GRK2 kinase, is an important regulator of beta-adrenergic signaling and plays a central role in heart failure (HF) pathology[32-34]. Two rSNPs in LD within the *ADRBK1* gene are rs948988 and rs4370946 whose minor alleles create punitive unique TFBS for ESR2 that is a binding site for the beta estrogen receptor which is expressed in blood monocytes and pulmonary epithelial cells (Tables 1-3). The ESR2 TFBS is not found with the common (rs948988 and rs4370946) alleles of the gene and may be related to HF. The NR3C1 TFBS for the glucocorticoid receptor which regulates carbohydrate, protein and fat metabolism is also only found with the minor alleles of these rSNPs (Tables 2 and 3) and should have an impact on HF. Other TFBS generated by the rs948988 minor allele of interest in HF might be the MYB and NFE2L.1:MAF TFs which are involved with hematopoietic progenitor cells and cell differentiation of erythrocytes as well as the rs4370946 common allele for the NRF1 TF which is involved with heme biosynthesis and mitochondrial DNA transcription and replication (Tables 2 and 3).

The type 2 deiodeinase gene (*DIO2*) encodes a deiodinase that coverts the thyroid prohormone, thyroxine (T4), to the biologically active triiodothyronine (T3) where T3 plays an important role in the regulation of energy balance and glucose metabolism[35-38]. Two rSNPs in LD within the *DIO2* gene are rs225015 and rs225011 whose major alleles create unique punitive TFBS for TFs that are involved with energy balance and glucose metabolism (Tables 1-3). The ESRR an alpha estrogen-related receptor that is involved with regulating thyroid hormone receptor genes while PPAR:RXR and RXR are involved with the regulation of adipocyte differentiation and glucose homeostasis (Tables 2 and 3). The minor allele of the rs225015 rSNP creates a unique punitive TFBS for the TCF7L2 TF whose protein is implicated in blood glucose homeostasis (Tables 2 and 3). The minor allele of the rs225011 rSNP creates a unique punitive TFBS for the PDX1 TF whose protein activates insulin, somatostatin, glucokinase, islet amyloid polypeptide and glucose transporter type 2 gene transcription (Tables 2 and 3).

The thromboxane A2 receptor (*TBXA2R*) gene is a member of the seven-transmembrane G-protein-coupled receptor super family, which interacts with intracellular G proteins, regulates different downstream signaling cascades, and induces many cellular responses including the intracellular calcium influx, cell migration and proliferation as well as apoptosis[39]. Two rSNPs in LD within the *TBXA2R* gene are rs2238631 and rs2238634 whose minor alleles create unique punitive TFBS for TFs that are involved in signaling cascades and apoptosis (Table 1-3). The ELK1 and SPZ1 TFs are involved with the ras-raf-MAPK signaling cascade while the ETS1 TF is involved with cell death (Tables 2 and 3). NR2E3 is part of a large family of nuclear receptor TFs involved in signaling pathways (Tables 2 and 3).

The other six genes can be analyzed in the same manner to identify punitive TFBS created by the rSNP alleles of these genes (Tables 2 and 3). What a change in the rSNP alleles can do, is to alter the DNA landscape around the SNP for potential TFs to attach and regulate a gene. This change in the DNA landscape can alter gene regulation which in turn can result in a change of a biological process or signaling pathway resulting in disease or illness. The process laid out in this report is a convenient way of identifying potential TFBS created by rSNP alleles that have been found to be significantly associated with disease or sickness. Any potential alterations in TFBS obtained by computational analyses need to be verified by protein/DNA electrophoretic mobility gel shift assays and gene expression studies[40]. CHIP-seq[41] experiments have become the standard method of validating TFBS and studying gene regulation[42-44].

In conclusion, SNPs in the regulatory region of a gene can alter the DNA landscape for TFs resulting in TFBS changes. Consequently, alterations in TF binding can affect gene regulation. Examples of this for nine genes are presented in this report where SNP alleles will either have no effect on TF binding or each allele will create unique punitive TFBS and alter a TFs ability to bind the DNA and regulate the gene.

**COMMENTS**

***Background***

Transcriptional factors (TFs) bind the DNA near a gene at transcriptional factor binding sites (TFBS) in order to regulate the gene. Single nucleotide polymorphisms (SNPs) that occur in the TFBS can alter the TFs ability to bind the DNA and thereby affect gene regulation. Such regulatory (r)SNPs have been associated with human disease and sickness. In this report, the alteration of TFBS created by rSNP alleles associated with disease has been documented for nine human genes. Sometimes the rSNP alleles will have no effect on the TFBS and not change the TF ability to bind the DNA. Other times each allele will create unique punitive TFBS that alter the TFs ability to regulate the gene.

***Research frontiers***

This article addresses an emerging concept in understanding how rSNPs which are significantly associated with disease can alter the TFBS for TFs that regulate a gene.

***Innovations and breakthroughs***

TFBS alteration by rSNPs is a newly emerging field of research and provides a different direction in examining changes in gene regulation resulting in human disease and sickness.

***Applications***

Given the great diversity of punitive unique TFBS generated by each allele of a rSNP, the author suspects that alterations in TFBS affect how well a gene is expressed. The outcome may result in disease or sickness. The methods outlined in the article should be applied to all rSNPs that are associated with disease or sickness of a regulatory nature.

***Terminology***

rSNP: A regulatory single nucleotide polymorphism that affects gene expression; TF: Transcriptional factor that is involved with regulating a gene; TFBS: Transcription factor DNA binding site in the regulatory region of a gene; Unique TFBS: A TFBS created by one rSNP allele and not the alternate allele.

***Peer review***

This study is technically well performed and a very interesting result. The interpretation was also sound. The report applied a computational approach to predict functional rSNPs in TFBS, focusing on several genes published earlier. Computational modeling and analysis for functional prediction is one of the approaches recently developed, particularly to address GWAS findings.

**REFERENCES**

1 **Pennisi E**. The Biology of Genomes. Disease risk links to gene regulation. *Science* 2011; **332**: 1031 [PMID: 21617055 DOI: 10.1126/science.332.6033.1031]

2 **Kumar V**, Wijmenga C, Withoff S. From genome-wide association studies to disease mechanisms: celiac disease as a model for autoimmune diseases. *Semin Immunopathol* 2012; **34**: 567-580 [PMID: 22580835 DOI: 10.1007/s00281-012-0312-1]

3 **Hindorff LA**, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 2009; **106**: 9362-9367 [PMID: 19474294 DOI: 10.1073/pnas.0903103106]

4 **Kumar V**, Westra HJ, Karjalainen J, Zhernakova DV, Esko T, Hrdlickova B, Almeida R, Zhernakova A, Reinmaa E, Võsa U, Hofker MH, Fehrmann RS, Fu J, Withoff S, Metspalu A, Franke L, Wijmenga C. Human disease-associated genetic variation impacts large intergenic non-coding RNA expression. *PLoS Genet* 2013; **9**: e1003201 [PMID: 23341781 DOI: 10.1371/journal.pgen.1003201]

5 **Knight JC**. Functional implications of genetic variation in non-coding DNA for disease susceptibility and gene regulation. *Clin Sci (Lond)* 2003; **104**: 493-501 [PMID: 12513691 DOI: 10.1042/CS20020304]

6 **Wang X**, Tomso DJ, Liu X, Bell DA. Single nucleotide polymorphism in transcriptional regulatory regions and expression of environmentally responsive genes. *Toxicol Appl Pharmacol* 2005; **207**: 84-90 [PMID: 16002116 DOI: 10.1016/j.taap.2004.09.024]

7 **Chorley BN**, Wang X, Campbell MR, Pittman GS, Noureddine MA, Bell DA. Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: current and developing technologies. *Mutat Res* 2008; **659**: 147-157 [PMID: 18565787 DOI: 10.1016/j.mrrev.2008.05.001]

8 **Buroker NE**. AKT3 rSNPs, Transcritional Factor Binding Sites and Human Disease. *Open Journal of Blood Diseases* 2013; **3**: 116-129 [DOI: 10.4236/ojbd.2013.34023]

9 **Buroker NE**. ADRBK1 (GRK2) rSNPs, Transcriptional Factor Binding Sites and Cardiovascular Disease in the Black Population. *Journal of Cardiovascular Disease* 2014; **2**: 1

10 **Knight JC**. Regulatory polymorphisms underlying complex disease traits. *J Mol Med (Berl)* 2005; **83**: 97-109 [PMID: 15592805 DOI: 10.1007/s00109-004-0603-7]

11 **Wang X**, Tomso DJ, Chorley BN, Cho HY, Cheung VG, Kleeberger SR, Bell DA. Identification of polymorphic antioxidant response elements in the human genome. *Hum Mol Genet* 2007; **16**: 1188-1200 [PMID: 17409198 DOI: 10.1093/hmg/ddm066]

12 **Prokunina L**, Alarcón-Riquelme ME. Regulatory SNPs in complex diseases: their identification and functional validation. *Expert Rev Mol Med* 2004; **6**: 1-15 [PMID: 15122975 DOI: 10.1017/S1462399404007690]

13 **Buckland PR**. The importance and identification of regulatory polymorphisms and their mechanisms of action. *Biochim Biophys Acta* 2006; **1762**: 17-28 [PMID: 16297602 DOI: 10.1016/j.bbadis.2005.10.004]

14 **Sadee W**, Wang D, Papp AC, Pinsonneault JK, Smith RM, Moyer RA, Johnson AD. Pharmacogenomics of the RNA world: structural RNA polymorphisms in drug therapy. *Clin Pharmacol Ther* 2011; **89**: 355-365 [PMID: 21289622 DOI: 10.1038/clpt.2010.314]

15 **Li G**, Pan T, Guo D, Li LC. Regulatory Variants and Disease: The E-Cadherin -160C/A SNP as an Example. *Mol Biol Int* 2014; **2014**: 967565 [PMID: 25276428 DOI: 10.1155/2014/967565]

16 **Sadee W**, Hartmann K, Seweryn M, Pietrzak M, Handelman SK, Rempala GA. Missing heritability of common diseases and treatments outside the protein-coding exome. *Hum Genet* 2014; **133**: 1199-1215 [PMID: 25107510 DOI: 10.1007/s00439-014-1476-7]

17 **Tak YG**, Farnham PJ. Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics Chromatin* 2015; **8**: 57 [PMID: 26719772 DOI: 10.1186/s13072-015-0050-4]

18 **Buroker NE**, Ning XH, Zhou AN, Li K, Cen WJ, Wu XF, Zhu WH, Scott CR, Chen SH. SNPs amd TFBS Associated with High Altitude Sickness. *Open Journal of Blood Diseases* 2013; **3**: 85-93 [DOI: 10.4236/OJBD.2013.33018]

19 **Buroker NE**. Regulatory SNPs and transcriptional factor binding sites in ADRBK1, AKT3, ATF3, DIO2, TBXA2R and VEGFA. *Transcription* 2014; **5**: e964559 [PMID: 25483406 DOI: 10.4161/21541264.2014.964559]

20 **Buroker NE**. ATF3 rSNPs, transcriptional factor binding sites and human etiology. *Open Journal of Genetics* 2013; **3**: 253-261 [DOI: 10.4236/ojgen.2013.34028]

21 **Buroker NE**. DIO2 rSNPs, transcription factor binding sites and disease. *British Journal of Medicine & Medical Research* 2014; **9**: 1-24 [DOI: 9734/BLMMR/2014/18535]

22 **Buroker NE**. Computational EPAS1 rSNP analysis, transcriptional factor binding sites and high altitude sickness or adaptation. *Journal of Proteomics and Genomics Research* 2016; **1**: 31-59 [DOI: 10.14302/issn.2326-0793.jpgr-15-889]

23 **Buroker NE**. LIPA rSNPs (rs1412444 and rs2246833), Transcriptional Factor Binding Sites and Disease. *British Biomedical Bulletin* 2015; **3**: 281-294 [DOI: 10.9734/BLMMR/2015/18535]

24 **Buroker NE**. Computational STAT4 rSNP analysis, transcriptional factor binding sites and disease. *Bioinformatics and Diabetes* 2016; **1**: 1-36 [DOI: 10.14302/ISSN: 2374-9431.jbd-15-890]

25 **Buroker NE**. VEGFA rSNPs, transcriptional factor binding sites and human disease. *J Physiol Sci* 2014; **64**: 73-76 [PMID: 24097272 DOI: 10.1007/s12576-013-0293-4]

26 **Buroker NE**. VEGFA SNPs (rs34357231 & rs35569394), Transcriptional Factor Binding Sites and Human Disease. *British Journal of Medicine & Medical Research* 2015; **10**: 1-11 [DOI: 10.9734/BJMMR/2015/19777]

27 **Buroker NE**, Ning XH, Li K, Zhou ZN, Cen WJ, Wu XF, Zhu WZ, Scott CR, Chen SH. SNPs, Linkage Disequilibrium and Transcriptional Factor Binding Sites Associated with Acute Mountain Sickness among Han Chinese at the Qinghai-Tibetan Plateau. *International Journal of Genomic Medicine* 2015; **3**: 1 [DOI: 10.4172/2332-0672.1000120]

28 **Buroker NE**, Ning XH, Zhou ZN, Li K, Cen WJ, Wu XF, Zhu WZ, Scott CR, Chen SH. VEGFA SNPs and transcriptional factor binding sites associated with high altitude sickness in Han and Tibetan Chinese at the Qinghai-Tibetan Plateau. *J Physiol Sci* 2013; **63**: 183-193 [PMID: 23553563 DOI: 10.1007/s12576-013-0257-8]

29 **Bryne JC**, Valen E, Tang MH, Marstrand T, Winther O, da Piedade I, Krogh A, Lenhard B, Sandelin A. JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Res* 2008; **36**: D102-D106 [PMID: 18006571 DOI: 10.1093/nar/gkm955]

30 **Sandelin A**, Alkema W, Engström P, Wasserman WW, Lenhard B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res* 2004; **32**: D91-D94 [PMID: 14681366 DOI: 10.1093/nar/gkh012]

31 **Sandelin A**, Wasserman WW, Lenhard B. ConSite: web-based prediction of regulatory elements using cross-species comparison. *Nucleic Acids Res* 2004; **32**: W249-W252 [PMID: 15215389 DOI: 10.1093/nar/gkh372]

32 **Lobmeyer MT**, Wang L, Zineh I, Turner ST, Gums JG, Chapman AB, Cooper-DeHoff RM, Beitelshees AL, Bailey KR, Boerwinkle E, Pepine CJ, Johnson JA. Polymorphisms in genes coding for GRK2 and GRK5 and response differences in antihypertensive-treated patients. *Pharmacogenet Genomics* 2011; **21**: 42-49 [PMID: 21127457 DOI: 10.1097/FPC.0b013e328341e911]

33 **Cannavo A**, Liccardo D, Koch WJ. Targeting cardiac β-adrenergic signaling via GRK2 inhibition for heart failure therapy. *Front Physiol* 2013; **4**: 264 [PMID: 24133451 DOI: 10.3389/fphys.2013.00264]

34 **Lymperopoulos A**. Physiology and pharmacology of the cardiovascular adrenergic system. *Front Physiol* 2013; **4**: 240 [PMID: 24027534 DOI: 10.3389/fphys.2013.00240]

35 **Danforth E**. The role of thyroid hormones and insulin in the regulation of energy metabolism. *Am J Clin Nutr* 1983; **38**: 1006-1017 [PMID: 6359854]

36 **Krotkiewski M**. Thyroid hormones in the pathogenesis and treatment of obesity. *Eur J Pharmacol* 2002; **440**: 85-98 [PMID: 12007527 DOI: 10.1016/S0014-2999(02)01420-6]

37 **Silva JE**. Thyroid hormone control of thermogenesis and energy balance. *Thyroid* 1995; **5**: 481-492 [PMID: 8808101 DOI: 10.1089/thy.1995.5.481]

38 **Freake HC**, Oppenheimer JH. Thermogenesis and thyroid function. *Annu Rev Nutr* 1995; **15**: 263-291 [PMID: 8527221 DOI: 10.1146/annurev.nu.15.070195.001403]

39 **Huang JS**, Ramamurthy SK, Lin X, Le Breton GC. Cell signalling through thromboxane A2 receptors. *Cell Signal* 2004; **16**: 521-533 [PMID: 14751539 DOI: 10.1016/j.cellsig.2003.10.008]

40 **Tiwari P**, Tripathi LP, Nishikawa-Matsumura T, Ahmad S, Song SN, Isobe T, Mizuguchi K, Yoshizaki K. Prediction and experimental validation of a putative non-consensus binding site for transcription factor STAT3 in serum amyloid A gene promoter. *Biochim Biophys Acta* 2013; **1830**: 3650-3655 [PMID: 23391827 DOI: 10.1016/j.bbagen.2013.01.024]

41 **Robertson G**, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, Euskirchen G, Bernier B, Varhol R, Delaney A, Thiessen N, Griffith OL, He A, Marra M, Snyder M, Jones S. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* 2007; **4**: 651-657 [PMID: 17558387 DOI: 10.1038/nmeth1068]

42 **Mundade R**, Ozer HG, Wei H, Prabhu L, Lu T. Role of ChIP-seq in the discovery of transcription factor binding sites, differential gene regulation mechanism, epigenetic marks and beyond. *Cell Cycle* 2014; **13**: 2847-2852 [PMID: 25486472 DOI: 10.4161/15384101.2014.949201]

43 **Levitsky VG**, Kulakovskiy IV, Ershov NI, Oshchepkov DY, Makeev VJ, Hodgman TC, Merkulova TI. Application of experimentally verified transcription factor binding sites models for computational analysis of ChIP-Seq data. *BMC Genomics* 2014; **15**: 80 [PMID: 24472686 DOI: 10.1186/1471-2164-15-80]

44 **Tehranchi AK**, Myrthil M, Martin T, Hie BL, Golan D, Fraser HB. Pooled ChIP-Seq Links Variation in Transcription Factor Binding to Complex Disease Risk. *Cell* 2016; **165**: 730-741 [PMID: 27087447 DOI: 10.1016/j.cell.2016.03.041]

**P-Reviewer:** Jeong BH, Xuei XL **S-Editor:** Qiu S **L-Editor: E-Editor:**

|  |  |
| --- | --- |
| **Table 1 Genes and their single nucleotide polymorphisms that have been found to be associated with disease or sickness. Also listed is the gene chromosome location, single nucleotide polymorphism location in the gene and the resulting genetic mutation as well as the occurrence of linkage disequilibrium between single nucleotide polymorphisms found within each gene** |  |
| **Protein and gene symbol** | **Chromosome** | **SNP** | **SNP location** | **Mutation** | **LD**  | **Ref.** |
| Adrenergic, beta, receptor kinase 1 (*ADRBK1*) | 11q13.1 | rs948988  | intron 2 | c.190+653G>A | Yes | [9,19] |
|  |  | rs4370946 | 3'UTR | c.\*217C>T | Yes |  |
| v-akt murine thymoma viral oncogene homolog 3 (*AKT3*) | 1q44 | rs4590656  | intron 1 | c.46+3654C>T | Yes | [8,19] |
|  |  | rs10157763 | intron 1 | c.46+11386C>T | Yes |  |
|  |  | rs2125230 | intron 1 | c.47-26830G>A | Yes |  |
| Activating transcription factor 3 (*ATF*3)  | 1q32.3 | rs3125289  | promoter | c.-5+9322T>C | unknown | [19,20] |
|  |  | rs11119982  | promoter | c.-4-23516C>T | unknown |  |
| Type 2 deiodeinase gene (DIO2)  | 14q24.3  | rs225015  | 3'UTR | c.\*1453G>A | Yes | [19,21] |
|  |  | rs225011 | intron 1 | c. 330+366C>T | Yes |  |
|  |  | rs12885300  | 5'UTR | c.-451C>T | Yes |  |
| Endothetal Per-Arnt-Sim (PAS) domain protein 1 (*EPAS1*)  | 2p21 | rs6756667  | intron 2 | c.218-3881A>G | No | [22] |
|  |  | rs1868092 | 3'UTR | c.\*2403G>A | No |  |
| Lysosomal acid lipase A (*LIPA*)  | 10q23.31  | rs1412444  | intron 2 | c.229+2506C>T | n/a | [23] |
| Signal Transducer and Activator of Transcription 4 (*STAT4)*  | 2q32.3  | rs8179673  | intron 2 | c.274-28290T>C | Yes | [24] |
|  |  | rs10181656  | intron 2 | c.274-28828C>G | Yes |  |
| Thromboxane A2 receptor (TBXA2R) | 19p13.3  | rs2238631  | intron 1 | c.-84+2229G>A | Yes | [19,25] |
|  |  | rs2238632 | intron 1 | c.-84+2030C>T | Yes |  |
|  |  | rs2238634  | intron 1 | c.-84+1799G>T | Yes |  |
| Vascular endothelial growth factor A (*VEGFA*) | 6p21.1 | rs34357231 | promoter | c. -2550-2568D>I | Yes | [19,26,28] |
|  |  | rs1570360 | promoter | c.-614A>G | Yes |  |
|  |  | rs3025039 | 3'UTR | c.\*237C>T | Yes |  |

|  |
| --- |
| **Table 2 Genes whose single nucleotide polymorphisms are significantly associated with human disease or sickness** |
|  |
| **Gene Symbol** |  |  |  |  |  |  |  |
| **ADRBK1**  | Ethnic Group | B |  | B |  |  |  |
|  | Disease or Sickness | 12 |  | 12 |  |  |  |
|  | SNP | rs948988 (G/A) |  | rs4370946 (C/T) |  |  |  |
|  | Alleles (MAF) | G | A (0.29) | C | T (0.2) |  |  |
|  | Potential unique TFBS | KLF1, 4 | BATF:JUN | E2F1,3,4,6 | ARNT:AHR |  |  |
|  |  |  | ESR2 | EGR1 | ATOH1 |  |  |
|  |  |  | FOS | INSM1 | ELF1 |  |  |
|  |  |  | FOSL2 | KLF4 | ESR2 |  |  |
|  |  |  | JUND | NFKB1 | NR3C1 |  |  |
|  |  |  | JUN:FOS | NRF1 |  |  |  |
|  |  |  | MYB | SP1, 2 |  |  |  |
|  |  |  | NFE2L1:MAF |  |  |  |  |
|  |  |  | NR3C1 |  |  |  |  |
|  |  |  | SOX17 |  |  |  |  |
|  |  |  |  |  |  |  |  |
| **AKT3** | Ethnic Group | G |  | C  |  | C |  |
|  | Disease or Sickness | 1 |  | 14 |  | 14 |  |
|  | SNP | rs4590656 (C/T) |  | rs10157763 (C/T) |  | rs2125230 (G/A) |  |
|  | Alleles (MAF) | C | T (0.41) | C  | T (0.33) | G | A (0.2) |
|  | Potential unique TFBS | ARNT:AHR | GFI | ELF5 | CTCF | ARNT:AHR | GATA1 |
|  |  | HIF1a:ARNT | HNF4A | ELK1 | NFATC2 | FEV |  HNF4a |
|  |  |  | PAX2 |  MYCN  | SOX17 | HIF1a:ARNT | HOXA5 |
|  |  |  | SPIB | SPIB  | ZNF354C | SPI1 | IRF1 |
|  |  |  |  | SPI1 |  |  | NR2F1 |
|  |  |  |  | TFAP2A |  |  | SOX17 |
| **ATF3** | Ethnic Group | C  |  | C |  |  |  |
|  | Disease or Sickness | 15 |  | 15 |  |  |  |
|  | SNP | rs3125289 (C/T) |  | rs11119982 (C/T) |  |  |  |
|  | Alleles (MAF) | C  | T (0.10) | C | T (0.36) |  |  |
|  | Potential unique TFBS | ARNT | FOXA1, 2 | HLTF | ARID3A |  |  |
|  |  | ARNT:AHR | FOXL1 |  | MAX |  |  |
|  |  | GABPa | FOXO3 |  | MYB |  |  |
|  |  | MYC | HLTF |  | USF1 |  |  |
|  |  | MYCN | SOX10 |  | ZEB1 |  |  |
|  |  | MZF1 | SOX17 |  |  |  |  |
|  |  | SPIB | SRY |  |  |  |  |
|  |  | USF1 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| **DIO2** | Ethnic Group | F |  | F |  | C |  |
|  | Disease or Sickness | 2 |  | 2 |  | 17 |  |
|  | SNP | rs225015 (G/A) |  | rs225011 (C/T) |  | rs12885300 (C/T) |  |
|  | Alleles (MAF) | G | A (0.4) | C | T (0.42) | C | T (0.23) |
|  | Potential unique TFBS | EBF1 | ELF1 | CRX | FOXL1 |  | ARID3A |
|  |  | ESRRA | ELK1 | RXRa | MEF2A |  | BATF:JUN |
|  |  | PPARg:RXRa | ERG |  | PDX1 |  | IRF1 |
|  |  | RFX5 | ETS1 |  |  |  | JUN:FOS |
|  |  | THAP1 | FLI1 |  |  |  | PAX2 |
|  |  |  | RUNX1 |  |  |  | SOX6 |
|  |  |  | SOX9 |  |  |  |  |
|  |  |  | SPI1 |  |  |  |  |
|  |  |  | TCF7L2 |  |  |  |  |
|  |  |  |  |  |  |  |  |
| **EPAS1** | Ethnic Group | G |  | G |  |  |  |
|  | Disease or Sickness | 1 |  | 1 |  |  |  |
|  | SNP | rs6756667 (A/G) |  | rs1868093 (A/G) |  |  |  |
|  | Alleles (MAF) | A | G (0.20) | A | G (0.25) |  |  |
|  | Potential unique TFBS | CEBPa | ATF7 | NR2C2 | HIC2 |  |  |
|  |  | NFIA | GMEB2 | NFIA | KLF5 |  |  |
|  |  | NRL | JDP2 | YY1 | MGA |  |  |
|  |  |  |  |  | TEAD1 |  |  |
|  |  |  |  |  | USF1 |  |  |
|  |  |  |  |  |  |  |  |
| **LIPA** | Ethnic Group | C, D, E |  |  |  |  |  |
|  | Disease or Sickness | 20, 21 |  |  |  |  |  |
|  | SNP | rs1412444 (C/T) |  |  |  |  |  |
|  | Alleles (MAF) | C | T (0.32) |  |  |  |  |
|  | Potential unique TFBS | ELF1 | FOXA1 |  |  |  |  |
|  |  | ETS1 | FOXL1 |  |  |  |  |
|  |  | GABPa | FOXO3 |  |  |  |  |
|  |  | HOXA5 | HNF1B |  |  |  |  |
|  |  | SPI1 | MEF2A |  |  |  |  |
|  |  |  | NFKB1 |  |  |  |  |
|  |  |  | NFIC |  |  |  |  |
|  |  |  | PAX2 |  |  |  |  |
|  |  |  | SOX6 |  |  |  |  |
|  |  |  | SOX9 |  |  |  |  |
|  |  |  | SRY |  |  |  |  |
|  |  |  | THAP1 |  |  |  |  |
|  |  |  |  |  |  |  |  |
| **STAT4** | Ethnic Group | A, C |  | A, C |  |  |  |
|  | Disease or Sickness | 2,6,10 |  | 2,10 |  |  |  |
|  | SNP | rs8179673 (T/C) |  | rs10181656 (C/G) |  |  |  |
|  | Alleles (MAF) | T | C (0.26) | C | G (0.26) |  |  |
|  | Potential unique TFBS | EN1 | FOXA2 | AR | HNF4g |  |  |
|  |  | NFIL3 | FOXH1 | E2F6 | STAT3 |  |  |
|  |  |  | FOXO1 | NR1H3:RXRa |  |  |  |
|  |  |  | FOXP1 | ZNF263 |  |  |  |
|  |  |  | FOXQ1 |  |  |  |  |
|  |  |  | HNF1a |  |  |  |  |
|  |  |  | HNF4g |  |  |  |  |
|  |  |  |  |  |  |  |  |
| **TBXA2R** | Ethnic Group | A |  | A |  | A |  |
|  | Disease or Sickness | 22 |  | 22 |  | 22 |  |
|  | SNP | rs2238631 (G/A) |  | rs2238632 (C/T) |  | rs2238634 (G/T) |  |
|  | Alleles (MAF) | G | A (0.2) | C | T (0.21) | G | T (0.22) |
|  | Potential unique TFBS | FOXC1 | ELK1 |  | ARNT |  | HLTF |
|  |  | TFAP2a | ELK4 |  | CREB1 |  | HNF4a |
|  |  |  | ETS1 |  | HIF1a:ARNT |  | NR2F1 |
|  |  |  | GATA2 |  | MAX |  | NR2E3 |
|  |  |  | HAND1:TCFE2a |  | USF1 |  | NR4A2 |
|  |  |  | SPZ1 |  |  |  |  |
| **VEGFA** | Ethnic Group | G |  | G |  | G |  |
|  | Disease or Sickness | 1 |  | 1 |  | 1 |  |
|  | SNP | rs34357231 (I/D) |  | rs1570360 (G/A) |  | rs3025039 (C/T) |  |
|  | Alleles (MAF) | D | I (0.28) | G | A (0.13) | C | T (0.09) |
|  | Potential unique TFBS | HNF4a | AR | EGR1 | EGR2 | BRCA1 | NFE2::MAF |
|  |  | HNF4g | EGR1,2 | MZF1 | EHF | ESR2 | RFX5 |
|  |  | JUN | KLF5 | SP2 | FOXH1 | HIF1A::ARNT | YY1 |
|  |  | MYB | MZF1\_1-4 |  | MAFK | NFE2L1::MAFG |  |
|  |  | NFIC | NFYB |  | SPIB |  |  |
|  |  | NR2C2 | NFATC2 |  | THAP1 |  |  |
|  |  | NR4A2 | NKX2-5 |  |  |  |  |
|  |  | PAX2 | NKX3-2 |  |  |  |  |
|  |  | RFX5 | SP1, 2 |  |  |  |  |
|  |  |  | STAT5A:STAT5B |  |  |  |  |
| Ethnic Group | Disease |  | Disease |  |  | Disease |  |
| A. Asian | 1. Chronic Mountain Sickness | 8. Juvenile Idiopathic Arthritis |  | 15. Hypospadias |  |
| B. Black | 2. Diabetes |  | 9. Primary biliary cirrhosis and Crohn's disease | 16. Mental Retardation |
| C. Caucasian | 3. Hepatitis B virus- related hepatocellular | 10. Lupus |  |  | 17. Osteoarthritis |  |
| D. Chinese | 4. Hepatitis B virus infection | 11. Ulcerative colitis |  | 18. Insulin Resistance |
| E. Hispanic | 5. HBV viral clearance |  | 12. Cardiovascular Disease |  | 19. Hepatic Glucose Output |
| F. Pima Indians | 6. Hepatocellular carcinoma | 13. Renal Cell Carcinoma risk |  | 20. Coronary Artery Disease |
| G. Tibetan | 7. Inflammatory bowel disease | 14. Aggressive Prostate Cancer |  | 21. Myocardial Infarction |
|  |  |  |  |  |  | 22. Asthma |  |
|  |  |  |  |  |  |  |  |

Also listed are the SNP alleles and frequencies within the ethnic group as well as the potential unique transcriptional factor binding site created with each SNP allele. For a complete list of significant gene SNPs see references (Table 1). SNP: Single nucleotide polymorphisms.

**Table 3 Transcriptional factors, protein name and their description or function**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **TF** | **Protein name** |  | **TF description/function** |  |  |  |  |  |  |  |
| AR | Androgen receptor |  | The protein functions as a steroid-hormone activated transcription factor |  |  |  |  |
|  |  |  |  | Upon binding the hormone ligand, the receptor dissociates from accessory proteins, translocates into the nucleus, |  |
|  |  |  |  | dimerizes, and then stimulates transcription of androgen responsive genes. They are expressed in bone marrow, |  |
|  |  |  |  | mammary gland, prostate, testicular and muscle tissues where they exist as dimers coupled to Hsp90 and HMGB proteins |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ARID3A | AT rich interactive domain 3A (BRIGHT-like)  | This gene encodes a member of the ARID (AT-rich interaction domain) family of DNA binding proteins |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ARNT | Aryl hydrocarbon receptor nuclear translocator | Involved in the induction of several enzymes that participate in xenobiotic metabolism |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ARNT:AHR | Hypoxia-inducible factor 1:Aryl hydrocarbon  | The dimer alters transcription of target genes. Involved in the induction |  |  |  |  |
|  | receptor nuclear translocator  |  | of several enzymes that participate in xenobiotic metabolism |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ATF7 | Activating Transcription Factor 7 |  | Plays important functions in early cell signaling |  |  |  |  |  |  |
|  |  |  |  | Has no intrinsic transcriptional activity, but activates transcription on formation of JUN or FOS heterodimers |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ATOH1 | Atonal homolog 1 |  | Transcriptional regulator. Activates E box-dependent transcription in collaboration with TCF3/E47 |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| BATF::JUN | Basic leucine zipper transcription factor, | The protein encoded by this gene is a nuclear basic leucine zipper protein that belongs to the AP-1/ATF superfamily of |  |
|  |  ATF-like Jun proto-oncogene  |  | transcription factors. The leucine zipper of this protein mediates dimerization with members of the Jun family of |  |
|  |  |  |  | proteins. This protein is thought to be a negative regulator of AP-1/ATF transcriptional events |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| BRCA1 | Breast cancer 1, early onset |  | This gene encodes a nuclear phosphoprotein that plays a role in maintaining genomic stability, and it also acts as a |  |
|  |  |  |  | tumor suppressor |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| CEBPA | CCAAT/enhancer binding protein (C/EBP), alpha | C/EBP is a DNA-binding protein that recognizes two different motifs: the CCAAT homology common to many |  |
|  |  |  |  | promoters and the enhanced core homology common to many enhancers |  |  |  |  |
| CREB1 | cAMP responsive element binding protein 1 | Phosphorylation-dependent transcription factor that stimulates transcription upon binding to the DNA |  |  |
|  |  |  |  | cAMP response element (CRE), a sequence present in many viral and cellular promoters |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| CRX | Cone-rod homeobox  |  | The protein encoded by this gene is a photoreceptor-specific transcription factor which plays a role in the |  |  |
|  |  |  |  | differentiation of photoreceptor cells. This homeodomain protein is necessary for the maintenance of normal cone |  |
|  |  |  |  | and rod function |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| CTCF | CCCTC-binding factor  |  | This gene is a member of the BORIS + **CTCF** gene family and encodes a transcriptional regulator protein |  |  |
|  | (zinc finger protein)  |  | with 11 highly conserved zinc finger (ZF) domains. This nuclear protein is able to use different combinations |  |  |
|  |  |  |  |  of the ZF domains to bind different DNA target sequences and proteins |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| E2F1-6 | E2F transcription factors 1-6 |  | The protein encoded by this gene is a member of the E2F family of transcription factors. The E2F family plays a |  |
|  |  |  |  | crucial role in the control of cell cycle and action of tumor suppressor proteins and is also a target of the |  |  |
|  |  |  |  | transforming proteins of small DNA tumor viruses. TheE2F proteins contain several evolutionally conserved |  |  |
|  |  |  |  | domains found in most members of the family. These domains include a DNA binding domain, a dimerization domain |  |
|  |  |  |  | which determines interaction with the differentiation regulated transcription factor proteins (DP), a |  |  |
|  |  |  |  | transactivation domain enriched in acidic amino acids, and a tumor suppressor protein association domain which is |  |
|  |  |  |  | embedded within the transactivation domain |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| EBF1  | Transcription factor COE1 |  | EBF1 has been shown to interact with ZNF423 and CREB binding proteins |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| EGR1 | Early growth response 1 |  | The protein encoded by this gene belongs to the EGR family of C2H2-type zinc-finger proteins. It is a nuclear |  |
|  |  |  |  | protein and functions as a transcriptional regulator. The products of target genes it activates are required for |  |  |
|  |  |  |  | differentiation and mitogenesis |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| EGR2 | Early growth response 2 |  | The protein encoded by this gene is a transcription factor with three tandem C2H2-type zinc fingers |  |  |
|  |  |  |  | Sequence-specific DNA-binding transcription factor |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| EHF | Ets homologous factor |  | This gene encodes a protein that belongs to an ETS transcription factor subfamily characterized by |  |  |
|  |  |  |  | epithelial-specific expression (ESEs). The encoded protein acts as a transcriptional repressor and may be |  |  |
|  |  |  |  | involved in epithelial differentiation and carcinogenesis |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ELF1 | E74-like factor 1 (ets domain transcription factor)  | The encoded protein is primarily expressed in lymphoid cells and acts as both an enhancer and a |  |  |  |
|  |  |  |  |  repressor to regulate transcription of various genes |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ELF5  | E74-like factor 5 |  | [A member of an epithelium-specific subclass of the Ets transcritpion factor family](http://en.wikipedia.org/wiki/Ets_transcritpion_factor_family) |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ELK1 | ELK1, member of ETS oncogene family  | This gene is a member of the Ets family of transcription factors and of the ternary complex factor (TCF) |  |  |
|  |  |  |  |  subfamily. The protein encoded by this gene is a nuclear target for the ras-raf-MAPK signaling cascade |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ELK4 | ELK4, ETS-domain protein (SRF accessory protein 1) | This gene is a member of the Ets family of transcription factors and of the ternary complex factor (TCF) |  |  |
|  |  |  |  | subfamily. Proteins of the TCF subfamily form a ternary complex by binding to the serum response factor and |  |
|  |  |  |  | the serum response element in the promoter of the c-fos proto-oncogene |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| EN1 | Engrailed homeobox 1  |  | Homeobox-containing genes are thought to have a role in controlling development |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ERG | v-ets avian erythroblastosis virus E26 oncogene homolog  | This gene encodes a member of the erythroblast transformation-specific (ETS) family of transcriptions factors. All |  |
|  |  |  |  | members of this family are key regulators of embryonic development, cell proliferation, differentiation, angiogenesis, |  |
|  |  |  |  | inflammation, and apoptosis |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ESR2  | Estrogen receptor beta |  | Estrogen receptor β is a member of the family of estrogen receptors and the superfamily of nuclear  |  |  |
|  |  |  |  | receptor transcription factors and is expressed by many tissues including blood monocytes and tissue  |  |  |
|  |  |  |  | macrophages, colonic and pulmonary epithelial cells |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ESRRA | Estrogen-related receptor alpha |  | This nuclear receptor acts as a site-specific transcription regulator and has been also shown to interact with estrogen |  |
|  |  |  |  | and the transcripton factor TFIIB by direct protein-protein contact. The binding and regulatory activities of this |  |
|  |  |  |  | protein have been demonstrated in the regulation of a variety of genes including lactoferrin, osteopontin, |  |  |
|  |  |  |  | medium-chain acyl coenzyme A dehydrogenase (MCAD) and thyroid hormone receptor genes |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ETS1  | Protein C-ets-1 |  | [The protein encoded by this gene belongs to the ETS family of transcription factors](http://en.wikipedia.org/wiki/Transcription_factor) |  |  |  |
|  |  |  |  | and has been shown to interact with TTRAP, UBE2I and Death associated protein |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FEV  | ETS oncogene family |  | It functions as a transcriptional repressor |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FLI1 | Fli-1 proto-oncogene, ETS transcription factor  | Sequence-specific transcriptional activator |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FOS | FBJ murine osteosarcoma viral oncogene homolog | The **Fos** gene family consists of 4 members: **FOS**, FOSB, FOSL1, and FOSL2. These genes encode leucine zipper proteins |  |
|  |  |  |  | that can dimerize with proteins of the JUN family, thereby forming the transcription factor complex AP-1. As |  |
|  |  |  |  | such, the **FOS** proteins have been implicated as regulators of cell proliferation, differentiation, and |  |  |
|  |  |  |  | transformation. In some cases, expression of the **FOS** gene has also been associated with apoptotic cell death |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FOSL1 and 2 | FOS-like antigen 1 and 2 |  | GO annotations related to this gene include *RNA polymerase II regulatory region sequence-specific DNA binding* |  |
|  |  |  |  |  and sequence-specific DNA binding transcription factor activity |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FOXA1 | Forkhead box A1 |  | Transcription factor that is involved in embryonic development, establishment of tissue-specific gene |  |  |
|  |  |  |  | expression and regulation of gene expression in differentiated tissues. Is thought to act as a 'pioneer' factor |  |  |
|  |  |  |  | opening the compacted chromatin for other proteins through interactions with nucleosomal core histones and |  |
|  |  |  |  | thereby replacing linker histones at target enhancer and/or promoter sites |  |  |  |  |
|  |  |  |  | Involved in the development of multiple endoderm-derived organ systems such as liver, pancreas, lung and |  |  |
|  |  |  |  | prostate. Modulates the transcriptional activity of nuclear hormone receptors |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FOXA2 | Forkhead box A2 |  | Involved in embryonic development, establishment of tissue-specific gene |  |  |  |  |
|  |  |  |  | expression and regulation of gene expression in differentiated tissues |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FOXC1 | Forkhead box C1 |  | An important regulator of cell viability and resistance to oxidative stress in the eye |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FOXL1 | Forkhead box L1  |  | Transcription factor required for proper proliferation and differentiation in the gastrointestinal |  |  |  |
|  |  |  |  | epithelium. Target gene of the hedgehog (Hh) signaling pathway |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FOXO1 | Forkhead Box O1 |  | Transcription factor that is the main target of insulin signaling and regulates metabolic  |  |  |  |
|  |  |  |  | homeostasis in response to oxidative stress |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FOXO3 | Forkhead Box O3 |  | This gene belongs to the forkhead family of transcription factors which |  |  |  |  |
|  |  |  |  | are characterized by a distinct forkhead domain. This gene likely |  |  |  |  |  |
|  |  |  |  | functions as a trigger for apoptosis through expression of genes necessary for cell death |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FOXP1 | Forkhead box P1 |  | This gene belongs to subfamily P of the forkhead box (FOX) transcription factor family. Forkhead box transcription |  |
|  |  |  |  | factors play important roles in the regulation of tissue- and cell type-specific gene transcription during both |  |  |
|  |  |  |  | development and adulthood. Transcriptional repressor. It plays an important role in the |  |  |  |
|  |  |  |  | specification and differentiation of lung epithelium |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FOXQ1 | Forkhead box Q1 |  | This gene belongs to the forkhead family of transcription factors which is characterized by a |  |  |  |
|  |  |  |  |  distinct DNA-binding forkhead domain. Plays a role in hair follicle differentiation |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| GABPA  | GA-binding protein alpha chain |  | One of three GA-binding protein transcription factor subunits which functions as a DNA-binding subunit |  |  |
|  |  |  |  | which shares identity with a subunit encoding the nuclear respiratory factor 2 gene |  |  |  |  |
|  |  |  |  | and is likely involved in activation of cytochrome oxidase expression and nuclear control of mitochondrial  |  |  |
|  |  |  |  | function |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| GATA1 | GATA binding protein 1 |  | The protein plays an important role in erythroid development by regulating |  |  |  |  |
|  |  |  |  | the switch of fetal hemoglobin to adult hemoglobin |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| GATA2  | GATA binding protein 2 |  | A member of the GATA family of zinc-finger transcription factors that are named for the consensus |  |  |
|  |  |  |  | nucleotide sequence they bind in the promoter regions of target genes and play an essential role in |  |  |
|  |  |  |  | regulating transcription of genes involved in the development and proliferation of hematopoietic  |  |  |
|  |  |  |  | and endocrine cell lineages |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| GATA3  | GATA binding protein 3 |  | Plays an important role in endothelial cell biology |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| GFI | Growth factor independent 1  |  | This gene encodes a nuclear zinc finger protein that functions as a transcriptional repressor. This protein plays a |  |
|  | transcription repressor  |  | role in diverse developmental contexts, including hematopoiesis and oncogenesis. It functions as part of a complex |  |
|  |  |  |  | along with other cofactors to control histone modifications that lead to silencing of the target gene promoters |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| GMEB2 | Glucocorticoid Modulatory Element | This gene is a member of KDWK gene family. The product of this gene associates with GMEB1 protein,  |  |  |
|  |  Binding Protein 1 |  | and the complex is essential for parvovirus DNA replication |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HAND1:TCFE2  | Heart- and neural crest derivatives-expressed | Hand1 belongs to the basic helix-loop-helix family of transcription factors |  |  |  |  |
|  |  protein 1: transcription factor E2A | The Tcfe2a gene encodes the transcription factor E2A, a member of the “class I”  |  |  |  |
|  |  |  |  | a family of basic helix-loop-helix (bHLH) transcription factors (also known simply as "E-proteins"). |  |  |
|  |  |  |  | The transcription factor E2A controls the initiation of B lymphopoiesis |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HIC1 | Hypermethylated In Cancer 1 |  | This gene functions as a growth regulatory and tumor repressor gene |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HIF1A:ARNT  | Hypoxia-inducible factor 1:Aryl hydrocarbon  | HIF1 is a homodimeric basic helix-loop-helix structure composed of HIF1a,  |  |  |  |  |
|  | receptor nuclear translocator  |  | the alpha subunit, and the aryl hydrocarbon receptor nuclear translocator  |  |  |  |  |
|  |  |  |  | (Arnt), the beta subunit. The protein encoded by HIF1 is a Per-Arnt-Sim  |  |  |  |  |
|  |  |  |  | (PAS) transcription factor found in mammalian cells growing at low oxygen  |  |  |  |  |
|  |  |  |  | concentrations. It plays an essential role in cellular and systemic responses  |  |  |  |  |
|  |  |  |  | to hypoxia |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HLTF | Helicase-like transcription factor |  | Member of the SWI/SNF (SWItch/Sucrose NonFermentable) family which have helicase and  |  |  |  |
|  |  |  |  | ATPase activities and are thought to regulate transcription of certain genes by altering the chromatin |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HNF1A | Hepatocyte Nuclear Factor 1 homeobox A | Transcriptional activator that regulates the tissue specific expression of multiple genes, especially in |  |  |
|  |  |  |  | pancreatic islet cells and in liver |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HNF1B | HNF1 homeobox B |  | This gene encodes a member of the homeodomain-containing superfamily of transcription factors. The protein binds |  |
|  |  |  |  | to DNA as either a homodimer, or a heterodimer with the related protein hepatocyte nuclear factor 1-alpha. The |  |
|  |  |  |  | gene has been shown to function in nephron development, and regulates development of the embryonic pancreas |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HNF4 | Hepatocyte nuclear factor 4, alpha  | The protein encoded by this gene is a nuclear transcription factor which binds DNA as a homodimer. The encoded |  |
|  |  |  |  | protein controls the expression of several genes, including hepatocyte nuclear factor 1 alpha, a transcription |  |  |
|  |  |  |  | factor which regulates the expression of several hepatic genes. This gene may play a role in development of the |  |
|  |  |  |  | liver, kidney, and intestines |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HNF4 | Hepatocyte nuclear factor 4, gamma  | Steroid hormone receptor activity and sequence-specific DNA binding transcription factor activity. |  |  |
|  |  |  |  | An important paralog of this gene is RXRA |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HOXA5  | Homeobox protein Hox-A5 |  | DNA-binding transcription factor which may regulate gene expression, morphogenesis, and differentiation |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ISNM1 | Insulinoma-associated 1 |  | Insulinoma-associated 1 gene is intronless and encodes a protein containing both a zinc finger DNA-binding |  |  |
|  |  |  |  | domain and a putative prohormone domain. This gene is a sensitive marker for neuroendocrine differentiation of |  |
|  |  |  |  | human lung tumors |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| IRF1,2 | Interferon regulatory factor |  | Members of the interferon regulatory transcription factor (IRF) family that contain a conserved  |  |  |
|  |  |  |  | N-terminal region of about 120 amino acids, which folds into a structure that binds specifically to  |  |  |
|  |  |  |  | the interferon consensus sequence (ICS) |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| JDP2 | Jun Dimerization Protein 2 |  | Component of the AP-1 transcription factor that represses transactivation mediated by the Jun family of proteins.  |  |
|  |  |  |  | Involved in a variety of transcriptional responses associated with AP-1 such as UV-induced apoptosis, cell |  |  |
|  |  |  |  |  differentiation, tumorigenesis and antitumogeneris |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| JUN | Jun Proto-Oncogene |  | This gene is the putative transforming gene of avian sarcoma virus 17. It encodes a protein which is highly |  |  |
|  |  |  |  | similar to the viral protein, and which interacts directly with specific target DNA sequences to regulate gene |  |  |
|  |  |  |  | expression |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| JUND | Jun D proto-oncogene  |  | The protein encoded by this intronless gene is a member of the JUN family, and a functional component of the AP1 |  |
|  |  |  |  | transcription factor complex. This protein has been proposed to protect cells from p53-dependent senescence and |  |
|  |  |  |  | apoptosis |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| JUN::FOS | *Jun* proto-oncogene FBJ murine |  | Promotes activity of NR5A1 when phosphorylated by HIPK3 leading to increased |  |  |  |
|  |  osteosarcoma viral oncogene homolog  | steroidogenic gene expression upon cAMP signaling pathway stimulation |  |  |  |  |
|  |  |  |  | Has a critical function in regulating the development of cells destined to form and |  |  |  |  |
|  |  |  |  | maintain the skeleton. It is thought to have an important role in signal transduction, |  |  |  |
|  |  |  |  | cell proliferation and differentiation |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| KLF1 | Kruppel-like factor 1 (erythroid)  |  | Transcription regulator of erythrocyte development that probably serves as a general switch factor during |  |  |
|  |  |  |  | erythropoiesis. Is a dual regulator of fetal-to-adult globin switching |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| KLF4 | Krueppel-like factor 4 |  | Transcription factor that can act both as activator and as repressor. |  |  |  |  |  |
|  |  |  |  | Regulates the expression of key transcription factors during |  |  |  |  |  |
|  |  |  |  | embryonic development |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| KLF5 | Krueppel-like factor 5 |  | This gene encodes a member of the Kruppel-like factor subfamily of zinc finger proteins. The encoded protein is a |  |
|  |  |  |  | transcriptional activator that binds directly to a specific recognition motif in the promoters of target genes |  |  |
|  |  |  |  | This protein acts downstream of multiple different signaling pathways and is regulated by post-translational |  |  |
|  |  |  |  | modification. It may participate in both promoting and suppressing cell proliferation. Expression of this gene |  |  |
|  |  |  |  | may be changed in a variety of different cancers and in cardiovascular disease. Alternative splicing results in |  |  |
|  |  |  |  | multiple transcript variants |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| MAX | MYC associated factor X |  | The protein encoded by this gene is a member of the basic helix-loop-helix leucine zipper (bHLHZ) family of |  |  |
|  |  |  |  | transcription factors |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| MAFK | v-maf avian musculoaponeurotic |  | Since they lack a putative transactivation domain, the small Mafs behave as transcriptional repressors |  |  |
|  |  fibrosarcoma oncogene homolog K  | when they dimerize among themselves. However, they seem to serve as transcriptional activators by dimerizing with |  |
|  |  |  |  | other (usually larger) basic-zipper proteins and recruiting them to specific DNA-binding sites |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| MEF2A | Myocyte enhancer factor 2A |  | The protein encoded by this gene is a DNA-binding transcription factor that activates many muscle-specific, growth |  |
|  |  |  |  | factor-induced, and stress-induced genes. Mediates cellular functions not only in skeletal and cardiac muscle development, |  |
|  |  |  |  | but also in neuronal differentiation and survival |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| MGA | MGA, MAX Dimerization Protein | Functions as a dual-specificity transcription factor, regulating the expression of both MAX-network and T-box family |  |
|  |  |  |  |  target genes. Functions as a repressor or an activator |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| MYB | Myb proto-oncogene protein |  | This gene encodes a transcription factor that is a member of the MYB family of transcription factor genes. |  |  |
|  |  |  |  | Transcriptional activator and plays an important role in the control of proliferation and differentiation of  |  |  |
|  |  |  |  | hematopoietic progenitor cells |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| MYC | v-*myc* myelocytomatosis viral oncogene homolog  | The protein encoded by this gene is a multifunctional, nuclear phosphoprotein that plays a role in cell cycle |  |  |
|  |  |  |  | progression, apoptosis and cellular transformation |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| MYCN | v-myc myelocytomatosis viral related | This gene is a member of the MYC family and encodes a protein with a basic helix-loop-helix (bHLH) domain |  |
|  | oncogene, neuroblastoma derived (avian)  | Amplification of this gene is associated with a variety of tumors, most notably neuroblastomas |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| MZF1\_1-4  | Myeloid zinc finger 1 |  | Binds to target promoter DNA and functions as transcription regulator. |  |  |  |  |
|  |  |  |  | May be one regulator of transcriptional events during hemopoietic development. |  |  |  |  |
|  |  |  |  | Isoforms of this protein have been shown to exist at protein level |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NFATC2  | Nuclear factor of activated T-cells, cytoplasmic 2 | This protein is present in the cytosol and only translocates to the nucleus upon T cell receptor (TCR)  |  |  |
|  |  |  |  | stimulation, where it becomes a member of the nuclear factors of activated T cells transcription complex. |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NFIA | Nuclear Factor I/A |  | Recognizes and binds the palindromic sequence 5-TTGGCNNNNNGCCAA-3 present in viral and cellular promoters |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  transcription and replication and in the origin of replication of adenovirus type 2. These proteins are  |  |  |
|  |  |  |  | individually capable of activating transcription and replication |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NFIC  | Nuclear factor 1 C-type |  | Recognizes and binds the palindromic sequence 5'-TTGGCNNNNNGCCAA-3' present in viral and cellular  |  |  |
|  |  |  |  | promoters and in the origin of replication of adenovirus type 2. These proteins are individually capable  |  |  |
|  |  |  |  | of activating transcription and replication |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NFE2::MAF | Nuclear factor, erythroid 2  |  | Regulates erythroid and megakaryocytic maturation and differentiation. Plays a role in all aspects  |  |  |
|  | V-*maf* avian musculoaponeurotic |  | of hemoglobin production from globin and heme synthesis to procurement of iron |  |  |  |
|  |  fibrosarcoma oncogene homolog |  | When overexpressed, represses anti-oxidant response element (ARE)-mediated transcription |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NFE2L1:MAFG | Nuclear factor erythroid 2-related factor 1 | Nuclear factor erythroid 2-related factor (Nrf2) coordinates the up-regulation of cytoprotective genes *via*  |  |  |
|  | Transcription factor MafG |  | the antioxidant response element (ARE). MafG is a ubiquitously expressed small maf protein that is  |  |  |
|  |  |  |  | involved in cell differentiation of erythrocytes. It dimerizes with P45 NF-E2 protein and activates  |  |  |
|  |  |  |  | expression of a and b-globin. |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NFIL3 | Nuclear factor, interleukin 3 regulated  | Expression of interleukin-3 (IL3; MIM 147740) is restricted to activated T cells, natural killer (NK) cells,  |  |
|  |  |  |  | and mast cell lines |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NFKB1 | Nuclear factor of kappa light polypeptide | NF-kappa-B is a pleiotropic transcription factor present in almost all cell types and is the endpoint of |  |  |
|  |  gene enhancer in B-cells 1  |  | a series of signal transduction events that are initiated by a vast array of stimuli related to many biological |  |  |
|  |  |  |  | processes such as inflammation, immunity, differentiation, cell growth, tumorigenesis and apoptosis |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NFYB | Nuclear transcription factor Y, beta | The protein encoded by this gene is one subunit of a trimeric complex, forming a highly conserved transcription |  |
|  |  |  |  | factor that binds with high specificity to CCAAT motifs in the promoter regions in a variety of genes. This gene |  |
|  |  |  |  | product, subunit B, forms a tight dimer with the C subunit, a prerequisite for subunit A association. The |  |  |
|  |  |  |  | resulting trimer binds to DNA with high specificity and affinity. Subunits B and C each contain a histone-like motif |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NHLH1 | Nescient helix loop helix 1 |  | The helix-loop-helix (HLH) proteins are a family of putative transcription factors, some of which have been shown |  |
|  |  |  |  | to play an important role in growth and development of a wide variety of tissues and species |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NKX2-5 | Natural killer 3 homeobox 2 |  | This gene encodes a member of the NK family of homeobox-containing proteins |  |  |  |  |
|  |  |  |  | Transcriptional repressor that acts as a negative regulator of chondrocyte maturation |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NKX3-2 | Natural killer 3 homeobox 2 |  | This gene encodes a member of the NK family of homeobox-containing proteins |  |  |  |  |
|  |  |  |  | Transcriptional repressor that acts as a negative regulator of chondrocyte maturation |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NR1H3:RXRa | Nuclear Receptor Subfamily 1, Group H, Member 3 | The protein encoded by this gene belongs to the NR1 subfamily of the nuclear receptor superfamily |  |  |
|  | Retinoid X receptor, alpha  |  | The NR1 family members are key regulators of macrophage function, controlling |  |  |  |  |
|  |  |  |  | transcriptional programs involved in lipid homeostasis and inflammation. This protein is highly |  |  |  |
|  |  |  |  | expressed in visceral organs, including liver, kidney and intestine. It forms a heterodimer with |  |  |  |
|  |  |  |  | retinoid X receptor (RXR), and regulates expression of target genes containing retinoid response elements |  |  |
|  |  |  |  | Studies in mice lacking this gene suggest that it may play an important role in the regulation of cholesterol homeostasis |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NR2C2 | Nuclear receptor subfamily 2, group C, member 2 | Orphan nuclear receptor that can act as a repressor or activator of transcription. An important |  |  |  |
|  |  |  |  | repressor of nuclear receptor signaling pathways such as retinoic acid receptor, retinoid X, vitamin D3 receptor, |  |
|  |  |  |  | thyroid hormone receptor and estrogen receptor pathways |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NR2E3 | Nuclear receptor subfamily 2, group E, member 3  | This protein is part of a large family of nuclear receptor transcription factors involved in signaling pathways |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NR2F1 | Nuclear receptor subfamily 2, group F, member 1 | Binds to the ovalbumin promoter and, in conjunction with another protein (S300-II) |  |  |  |
| (COUP) |  |  |  | stimulates initiation of transcription. Binds to both direct repeats and palindromes of the  |  |  |  |
|  |  |  |  | 5'-AGGTCA-3' motif. An important paralog of this gene is RXRA |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NR3C1 | Nuclear receptor subfamily 3, group C, member 1  | Glucocorticoids regulate carbohydrate, protein and fat metabolism, |  |  |  |  |  |
|  | (glucocorticoid receptor)  |  | modulate immune responses through suppression of chemokine and cytokine production and have critical roles in |  |
|  |  |  |  | constitutive activity of the CNS, digestive, hematopoietic, renal and reproductive systems |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NR4A2 | Nuclear receptor subfamily 4, group A, member 2 | Transcriptional regulator which is important for the differentiation and maintenance of meso-diencephalic |  |  |
|  |  |  |  | dopaminergic (mdDA) neurons during development |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NRF1 | Nuclear respiratory factor 1 |  | This gene encodes a protein that homodimerizes and functions as a transcription factor which activates the |  |  |
|  |  |  |  | expression of some key metabolic genes regulating cellular growth and nuclear genes required for respiration, |  |  |
|  |  |  |  | heme biosynthesis, and mitochondrial DNA transcription and replication |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NRL | Neural Retina Leucine Zipper |  | This gene encodes a basic motif-leucine zipper transcription factor of the Maf subfamily.  |  |  |  |
|  |  |  |  | The encoded protein is conserved among vertebrates and is a critical intrinsic regulator of photoceptor  |  |  |
|  |  |  |  | development and function |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| PAX2 | Paired box gene 2 |  | Probable transcription factor that may have a role in kidney cell  |  |  |  |  |  |
|  |  |  |  | differentiation |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| PDX1 | Pancreatic and duodenal homeobox 1  | Activates insulin, somatostatin, glucokinase, islet amyloid polypeptide and glucose transporter type 2 gene |  |  |
|  |  |  |  | transcription. Particularly involved in glucose-dependent regulation of insulin gene transcription |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| PPAR:RXR | Peroxisome proliferator-activated receptor gamma | Peroxisome proliferator-activated receptor gamma (PPARgamma) is a member of the nuclear receptor family of |  |
|  | Retinoid X receptor, alpha |  | ligand-activated transcription factors that heterodimerize with the retinoic X receptor (RXR) to regulate |  |  |
|  |  |  |  | gene expression. PPARgamma is located primarily in the adipose tissue, lymphoid tissue, colon, liver and |  |  |
|  |  |  |  | heart and is thought to regulate adipocyte differentiation and glucose homeostasis |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| RXRa | Retinoid X receptor, alpha |  | Retinoid X receptors (RXRs) and retinoic acid receptors (RARs), are nuclear receptors that mediate the biological |  |
|  |  |  |  | effects of retinoids by their involvement in retinoic acid-mediated gene activation |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| RFX5 | Regulatory factor X, 5 |  | Activates transcription from class II MHC promoters. Recognizes X-boxes. Mediates cooperative binding |  |  |
|  |  |  |  | between RFX and NF-Y. RFX binds the X1 box of MHC-II promoters |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| RUNX1 | Runt-related transcription factor 1 | Heterodimeric transcription factor that binds to the core element of  |  |  |  |  |
|  |  |  |  | many enhancers and promoters. The protein encoded by this gene represents  |  |  |  |  |
|  |  |  |  | the alpha subunit of core binding factor and is thought to be involved in the development of normal hematopoiesis |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SOX6 | SRY (sex determining region Y)-box 6 | The encoded protein is a transcriptional activator that is required for normal |  |  |  |  |
|  |  |  |  | development of the central nervous system, chondrogenesis and maintenance of cardiac and skeletal muscle cells |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SOX9 | SRY (sex determining region Y)-box 9 | The protein encoded by this gene recognizes the sequence CCTTGAG along with other members of the  |  |  |
|  |  |  |  | involved in chondrogenesis by acting as a transcription factor for these genes |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SOX10 | SRY (sex determining region Y)-box 10  | This gene encodes a member of the SOX (SRY-related HMG-box) family of transcription factors involved in the |  |
|  |  |  |  | regulation of embryonic development and in the determination of the cell fate |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SOX17 | SRY (sex determining region Y)-box 17  | Acts as transcription regulator that binds target promoter DNA and bends the DNA |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SP1 | Specificity Protein 1 |  | Can activate or repress transcription in response to physiological and  |  |  |  |  |
|  |  |  |  | pathological stimuli. Regulates the expression of a large number of  |  |  |  |  |  |
|  |  |  |  | genes involved in a variety of processes such as cell growth,  |  |  |  |  |  |
|  |  |  |  | apoptosis, differentiation and immune responses |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SP2 | Specificity Protein 2 |  | This gene encodes a member of the Sp subfamily of Sp/XKLF transcription factors. Sp family proteins are |  |  |
|  |  |  |  | sequence-specific DNA-binding proteins characterized by an amino-terminal trans-activation domain and three |  |
|  |  |  |  | carboxy-terminal zinc finger motifs. This protein contains the least conserved DNA-binding domain within the Sp |  |
|  |  |  |  | subfamily of proteins, and its DNA sequence specificity differs from the other Sp proteins. It localizes |  |  |
|  |  |  |  | primarily within subnuclear foci associated with the nuclear matrix, and can activate or in some cases repress |  |  |
|  |  |  |  | expression from different promoters |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SPIB  | Transcription factor Spi-B |  | SPI1 and SPIB are members of a subfamily of ETS transcription factors |  |  |  |  |
|  |  |  |  | ETS proteins share a conserved ETS domain that mediates specific DNA binding |  |  |  |  |
|  |  |  |  | SPIB and SPI1 bind to a purine-rich sequence, the PU box (5-prime-GAGGAA-3-) |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SPI1 | Spleen focus forming virus (SFFV) |  | This gene encodes an ETS-domain transcription factor that activates gene expression during myeloid  |  |  |
|  |  proviral integration oncogene spi1 | and B-lymphoid cell development |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SPZ1 |  |  |  | This gene encodes a bHLH-zip transcription factor which functions in the mitogen-activate protein kinase (MAPK) |  |
|  |  |  |  | signaling pathway |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SRY | Sex determining region Y  |  | Transcriptional regulator that controls a genetic switch in male development |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| STAT3 | Signal transducer and activator of |  | Signal transducer and transcription activator that mediates cellular responses to interleukins, |  |  |  |
|  |  transcription 3 (acute-phase response factor)  | KITLG/SCF and other growth factors |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| STAT5A:STAT5B | Signal transducer and activator of  |  | Carries out a dual function: signal transduction and activation of transcription |  |  |  |  |
|  | transcription 5A and transcription 5B  | Regulates the expression of milk proteins during lactation |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| TCF7L2 | Transcription factor 7-like 2 |  | This gene encodes a high mobility group (HMG) box-containing transcription factor that plays a key role in the Wnt |  |
|  |  (T-cell specific, HMG-box)  |  | signaling pathway. The protein has been implicated in blood glucose homeostasis |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| TEAD1 | TEA Domain Family Member 1 |  | This gene encodes a ubiquitous transcriptional enhancer factor that is a member of the TEA/ATTS domain family |  |
|  |  |  |  | This protein directs the transactivation of a wide variety of genes and, in placental cells, also acts as a |  |  |
|  |  |  |  |  transcriptional repressor |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| TFAP2a | Activator protein 2 |  | [The AP2a protein acts as a sequence specific DNA-binding](http://en.wikipedia.org/wiki/Transcription_factor)  |  |  |  |  |  |
|  |  |  |  | transcription factor recognizing and binding to the specific DNA  |  |  |  |  |  |
|  |  |  |  | sequence and recruiting transcription machinery |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| THAP1 | THAP domain containing, apoptosis | DNA-binding transcription regulator that regulates endothelial cell proliferation and G1/S cell-cycle |  |  |
|  |  associated protein 1 |  | progression |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| USF1  | Upstream transcription factor 1  |  | This gene encodes a member of the basic helix-loop-helix leucine zipper family, and can function as a cellular |  |  |
|  |  |  |  | transcription factor. The encoded protein can activate transcription through pyrimidine-rich initiator (Inr) |  |  |
|  |  |  |  | elements and E-box motifs |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| YY1 | YY1 transcription factor |  | YY1 is a ubiquitously distributed transcription factor belonging to the GLI-Kruppel class of zinc finger proteins |  |
|  |  |  |  | The protein is involved in repressing and activating a diverse number of promoters. YY1 may direct histone |  |  |
|  |  |  |  | deacetylases and histone acetyltransferases to a promoter in order to activate or repress the promoter, thus |  |  |
|  |  |  |  | implicating histone modification in the function of YY1 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ZEB1  | Zinc finger E-box-binding homeobox 1 | A member of the delta-EF1 (TCF8)/Zfh1 family of 2-handed zinc finger/homeodomain proteins and  |  |  |
|  |  |  |  | interacts SMADs with receptor-mediated, activated full-length activated full-length SMADs |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ZNF263 | Zinc finger protein 263 |  | Might play an important role in basic cellular processes as a transcriptional repressor. An important paralog to ZNF496 |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ZNF354C | Zinc finger protein 354C |  | May function as a transcription repressor  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |