British Journal of Medicine & Medical Research
9(4): 1-24, 2015, Article no.BJMMR.18535
ISSN: 2231-0614

SCIENCEDOMAIN *international*
www.sciencedomain.org

Computational *DIO2* rSNP Analysis, Transcriptional Factor Binding Sites and Disease

Norman E. Buroker[1*]

[1]Department of Pediatrics, University of Washington, Seattle, WA 98195, USA.

Author's contribution

The sole author designed, analyzed and interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/BJMMR/2015/18535
Editor(s):
(1) Alex Xiucheng Fan, Department of Biochemistry and Molecular Biology, University of Florida, USA.
Reviewers:
(1) Anonymous, Council of Scientific & Industrial Research, India.
(2) Safaa Ibrahim Tayel, Biochemistry Department, Faculty of Medicine, Menoufia Umiversity, Egypt.
(3) Leila Abid, Cardiology department, medicine university, Sfax, Tunisia.
Complete Peer review History: http://sciencedomain.org/review-history/9926

Short Research Article

Received 28th April 2015
Accepted 9th June 2015
Published 27th June 2015

ABSTRACT

Purpose: The *DIO2* gene transcribes the deiodinase type 2 enzyme that changes the thyroid prohormone, thyroxine (T4), to the biologically active triiodothyronine (T3) hormone. T3 plays a vital part in the regulation of energy balance and glucose metabolism. *DIO2* single-nucleotide polymorphisms (SNPs) were computationally examined with respect to changes in punitive transcriptional factor binding sites (TFBS) and these changes were discussed in relation to human disease.
Methods: The JASPAR CORE and ConSite databases were instrumental in identifying the TFBS. The Vector NTI Advance 11.5 computer program was employed in locating all the TFBS in the *DIO2* gene from 2.4 kb upstream of the transcriptional start site to 508 bp past the 3'UTR. The JASPAR CORE database was also involved in computing each nucleotide occurrence (%) within the TFBS.
Results: Regulatory SNPs (rSNPs) in the promoter region novel SNP (-2035bp), 5'UTR (rs12885300), intron one (rs225010, 225011 and rs225012), exon two [rs225014 (Thr92Ala)] and 3' UTR (rs6574549, rs225015 and rs225017) of the *DIO2* gene are in linkage disequilibrium. These rSNP alleles were found to alter the DNA landscape for potential transcriptional factors (TFs) to attach resulting in changes in TFBS.
Conclusion: The alleles of each rSNP were found to generate unique TFBS resulting in potential

**Corresponding author: Email: nburoker@u.washington.edu;*

changes in TF *DIO2* regulation. These regulatory changes were discussed with respect to changes in human health resulting in disease or sickness.

Keywords: DIO2; rSNPs; TFBS; disease.

1. INTRODUCTION

The type 2 deiodeinase gene (DIO2) encodes a deiodinase that coverts the thyroid prohormone, thyroxine (T4), to the biologically active triiodothyronine (T3) hormone where T3 is involved in the vital role of regulating energy balance and glucose metabolism [1-4]. DIO2 is found in the thyroid gland, cardiac and skeletal muscle, brown adipose tissue, placenta, pituitary, central nervous system (CNS) and at low levels in kidney and pancreases [5-7]. The DIO2 gene maps to human chromosome 14q24.3 and is about 15 kb in size. The coding region consists of two exons separated by a gap of approximately 7.4 kb [8]. Several single nucleotide polymorphisms (SNPs) have been found in the gene which have been studied is association with mental retardation (MR) [9], osteoarthritis [10], early-onset type 2 diabetes mellitus (T2DM) [11] and insulin resistance (IR) [12,13]. Three of the common SNPs in the gene (rs225014, rs225012 and rs225010) have been found to be in strong linkage disequilibrium (LD) with each other while the rs225012 and rs225010 SNPs have been shown to have a positive association with MR [9]. The haplotypes of two SNPs (rs225014 and rs12885300) have been shown to have a significant association with symptomatic osteoarthritis in Dutch women [10]. Three SNPs (rs225011, rs225014 and rs225015) which are in LD were found to be modestly associated with early-onset of T2DM in Pima Indians while these SNPs and rs6574549 were found to be nominally associated with hepatic glucose output [11]. Two of the SNPs (rs225014 and rs225017) which are in partial LD have been found to be associated with IR in Caucasian T2DM patients [13]. The rs6574549 SNP was also found to be associated with fasting insulin, insulin action and energy expenditure [11]. These studies suggest that some DIO2 SNPs may be affecting the regulatory network for the gene expression in humans. When LD is found between SNPs in a gene's regulatory region, it can result from strong associations of certain haplotypes with sickness or disease [14-16]. Consequently, a computational examination was made between *DIO2* SNPs in LD and the transcription factor binding site (TFBS) changes resulting from the SNPs. In this report LD is considered to be the non-random association of SNP alleles within the gene.

Nucleotide changes that influence gene expression by altering gene regulatory sequences such as in promoters, enhances, and silencers are known as regulatory SNPs (rSNPs) [17-20]. A rSNP within a transcriptional factor binding motif can alter a transcriptional factor's (TF) ability to bind the motif [21-24] in which case the TF would not effectively regulate the gene [25-29]. This concept is examined for the above *DIO2* rSNPs and their allelic association with TFBS, where computation analyses [30-33] was used to identify TFBS alterations created by the *DIO2* rSNPs. In this study, the rSNP associations with nucleotide substitutions in punitive TFBS are examined with their possible relationship to disease in humans.

2. METHODS

The JASPAR CORE database [34,35] and ConSite [36] were used to identify the potential *DIO2* TFBS in this study. JASPAR is a database 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 TFBS and rSNP location within the binding sites have previously been discussed [14,16,37]. The Vector NTI Advance 11.5 computer program (Invitrogen, Life Technologies) was used to locate the TFBS in the *DIO2* gene (NCBI Ref Seq NM_013989) from 2.4 kb upstream of the transcriptional start site to 508 bp past the 3'UTR which represents a total of 16.9 kb. The JASPAR CORE database was also used to calculate each nucleotide occurrence (%) within the TFBS, where upper case lettering indicate that the nucleotide occurs 90% or greater and lower case less than 90%. The occurrence of each SNP allele in the TFBS is also computed from the database (Table & Supplement).

3. RESULTS

3.1 *DIO2* rSNPs and TFBS

The *DIO2* gene transcribes the deiodinase type 2 enzyme that changes the thyroid prohormone, thyroxine (T4), to the biologically active

triiodothyronine (T3) hormone. The thyroid hormones play an important role in energy homeostasis and glucose metabolism. Due to the importance of this gene in energy homeostasis, *DIO2* SNPs associated with disease were computationally evaluated with regard to TFBS. The novel -2035bp SNP is located 5' upstream from the TSS, the rs12885300 SNP is located in the 5'UTR and the rs225010, rs225011 and rs225014 SNPs are found in intron one. The rs224014 (Thr92Ala) SNP is located in exon two while the rs674549, rs225015 and rs225017 SNPs are located in the 3' UTR. The novel -2035bp, rs225011, rs225014, rs674549 and rs225015 SNPs are all in LD with each other [11]. The rs225010 and rs225012 SNPs are also in LD but not with the other SNPs [9]. The novel -2035bp and rs6574549 SNPs have very rare alleles with a frequency of 0.004 and 0.007, respectively. Since the minor allele frequencies (MAF) of the other SNPs are rather large ranging from 0.229 to 0.421, the minor alleles that alter BS which give rise to different TFs would be expected to have an impact on *DIO2* regulation (Table).

The *DIO2* SNPs (rs225010 and rs225012) which are in LD have been found to be significantly associated with MR in Chinese [9]. The common rs225010 SNP *DIO2*-C allele creates six unique TFBS for the ELF5, ELK1, GATA2, GATA4, JUN:FOS and SREBF1 TFs, which are involved with the ETS transcriptional factor family, the ras-raf-MAPK signaling cascade, the proliferation of hematopoietic and endocrine cell lineages, myocardial differentiation and function, steroidogenic gene expression, and lipid homeostasis, respectively (Table, supplement). The minor *TIO2*-T allele creates four unique TFBS for the HOXA9, JUND, NFATC2 and ZNF354C TFs which are involved with development, transcription enhancement, inducible expression of cytokine genes in T-cells and transcription repression, respectively (Table, supplement). There are also five conserved TBFS for the GATA1, KLF5, MAFB, NFE2L1:MAFG and PAX TFs which are involved with erythroid development, transcription activation, up-regulation of cytotoprotective genes, cell differentiation of erythrocytes and kidney cell differentiation, respectively (Table, supplement). The common rs225012 SNP *DIO2*-G allele creates seven unique TFBS for the E2F6, EGR1, ELF1, ERG, SP1, SPI1 and ZNF263 TFs which are involved with control of cell cycle and action of tumor suppressor proteins, mitogenesis and differentiation, transcription enhancement and repression, regulator of embryonic development, activation or repression of transcription, myeloid and B-lymphoid cell development, and transcription repression, respectively (Table, supplement). The minor rs225012 SNP *DIO2*-A allele creates eight unique TFBS for the EHF, ELF5, EN1, HLTF, HOXA5, NKX3-2, PDX1 and PRRX2 TFs which are involved with epithelial-specific expression, controlling development, altering chromatin structure, cell development, negative regulator of chondrocyte maturation, glucose-dependent regulation of insulin gene transcription, and proliferating fetal fibroblasts, respectively (Table, supplement). There are also two conserved TBFS for the FOXC1 and SPIB TFs which are involved with regulation of cell viability and resistance to oxidative stress as well as transcriptional activation, respectively (Table, supplement).

Two DIO2 SNPs (rs225014 and rs12885300) have been shown to have a significant association with symptomatic osteoarthritis in Dutch women [10] while a third SNP (rs225017) has been found to be significantly associated with IR [13]. The rs225014 SNP results in a non-synonymous amino acid substitution (Thr92Ala) in exon 2 and has also been associated with IR in obese Caucasian women [12]. The common rs225014 SNP *DIO2*-T allele creates five unique TFBS for the FOXC1, HOXA5, SPI1, STAT5A:STAT5B and THAP1 TFs which are involved with cell viability and resistance to oxidative stress, cell development and myeloid and B-lymphoid cell development, signal transduction and activation of transcription, and G1/8 cell-cycle progression respectively (Table, supplement). The minor rs225014 SNP *DIO2*-C allele creates six unique TFBS for the EBF1, NKX3-2, SP2, SPIB, TFAP2C and ZNF354C TFs which are involved with transcription activation, negative regulation of the chondrocyte maturation, activation of mRNA synthesis and transcription repression, lymphoid-specific enhancement, respectively (Table, supplement). There are also four conserved TFBS for the EGR1, ELF1, HLTF and RXR::RAR_DR5 TFs which are involved with mitogensis and differentiation, enhancement and repression, altering chromatin structure, regulation of development, respectively (Table, supplement). The common rs12885300 *DIO2*-C allele creates no unique TFBS for TFs. The minor rs12885300 *DIO2*-T alleles creates six unique TFBS for the ARID3A, BATF:JUN, IRF1, JUN:FOS, PAX2 and SOX6 TFs which are involved with B-cell

differentiation, negative regulation of AP-1/ATF transcription events, regulation of cellular responses, signal transduction, cell proliferation and differentiation, kidney cell differentiation and maintenance of cardiac and skeletal muscle cells, respectively (Table, supplement). There are also four conserved TFBS between the rs12885300 SNP alleles for the FOXP1, PRDM1, SOX3 and TCF7L2 TFs which are involved with lung epithelium, repression of beta-interferon gene expression, neuronal development and blood glucose homeostasis, respectively (Table, supplement). The common rs225017 *DIO2*-T allele creates six unique potential TFBS for the HOXA5, JUND(var.2), NFE2L1::MAFG, PDX1, SOX3 and STAT3 TFs which are involved with development regulation, enhancer binding, erythrocyte development, insulin activation, neuronal development and signal transduction (Table, supplement). The minor rs225017 *DIO2*-A allele creates five unique potential TFBS for the CEBPA & B, NKX2-5, PRRX2 and SRY TFs which are involved enhancer binding, inflammation, hemopoiesis, chondrocyte maturation, fetal fibroblasts and male development (Table, supplement). There are also three conserved TFBS between the rs225017 SNP alleles for the BATF::JUN, GATA4 and HAND1:TCFE2□ TFs which are involved with negative regulation of AP-1/ATF transcription events, myocardial differentiation and function as well as B lymphopoiesis (Table, supplement).

Three DIO2 SNPS (rs225011, rs225014 and rs225015) have been modestly associated with early-onset T2DM in Pima Indians [11] while the rs225014 and rs225017 SNPs have been found to be associated with IR in Caucasian morbidly obese subjects and T2DM patients [12,13]. The common rs225011 SNP *DIO2*-C allele creates two unique TFBS for the CRX and RXRA TFs which are involved with photoreceptor cells and retinoic acid-mediated gene activation, respectively (Table, supplement). The minor rs225011 SNP *DIO2*-T allele creates three unique TFBS for the FOXL1, MEF2A and PDX1 TFs which are involved with metabolism, cell proliferation and gene expression, skeletal and cardiac muscle development, glucose-dependent regulation of insulin gene transcription, respectively (Table, supplement). There are also eight conserved TFBS between the rs125011 SNPs alleles for the ESRRA, ESRRB, GATA4, NKX2-5, NR5A2, PRRX2, RORA_1 and RORA_2 TFs which are involved with site-specific transcription regulation, myocardial differentiation and function, negative regulation of chondrocyte maturation, regulation of cholesterol expression in liver, proliferating fetal fibroblasts, and nuclear hormone receptors, respectively (Table, supplement).

The common rs225015 SNP *DIO2*-G allele creates five unique TFBS for the EBF1, ESRRA, PPARG:RXRA, RFX5 and THAP1 TFs which are involved with transcription activation, site-specific transcription regulation, regulation of adipocyte differentiation and glucose homeostasis, and regulation of endothelial cell proliferation and G1/8 cell-cycle progression, respectively (Fig. 1, Table and supplement). The minor rs225015 SNP *DIO2*-A allele creates nine unique TFBS for the ELF1, ELK1, ERG, ETS1, FLI1, RUNX1, SOX9, SPI1 and TCF7L2 TFs which are involved with lymphoid cells, ras-raf-MAPK signaling cascade, regulation of embryonic development, TTRAP, UBE2I and Death associated proteins, transcription activation, normal hematopoiesis, skeletal development, myeloid and B-lymphoid cell development and blood glucose homeostasis, respectively (Table, supplement). There are also ten conserved TFBS between the SNPs alleles for the BRCA1, ELF5, HLTF, NFATC2, NFKB1, SOX2, SOX3, SOX6, SOX10 and SPIB TFs which are involved with genomic stability, epithelium cells, altering chromatin structure, cytokine genes in T-cells, signal transduction, regulation of embryonic development, neuronal development, central nervous system, and lymphoid-specific enhancement, respectively (Table, supplement).

Four *DIO2* SNPs (rs225011, rs225014, rs225015 and rs6574549) were nominally associated with hepatic glucose output while the rs6574549 SNP was also associated with fasting insulin, insulin action and energy expenditure in Pima Indians [11]. The common rs6574549 SNP *DIO2*-T allele creates five unique TFBS for the ARID3A, HNF1B, HOXA5, LHX3 and NKX3-2 TFs which are involved with cell cycle progression, embryonic pancreas development, specific positional identities of cells, pituitary development and chondrocyte maturation, respectively (Table, supplement). The minor rs6574549 SNP *DIO2*-G allele creates four unique TFBS for the FOXA1, FOXA2, NFIL3 and POU2F2 TFs which are involved with embryonic development, expression of interleukin-3, POU domain family, respectively (Table, supplement). There are also seven conserved TFBS between the SNP alleles for the FOXC1, FOXD3, FOXI1, FOXL1, HLTF, NKX3-1 and NKX2-5 TFs which

are involved with cell viability and resistance to oxidative stress, activation and repression, normal hearing, sense of balance and kidney function, ontogenesis, altering chromatin structure, epithelial cell growth, and chondrocyte maturation, respectively (Table, supplement). These four *DIO2* SNPS and a novel SNP in the 5'UTR flanking region were found to be in LD in the Pima Indian study [11]. The common novel SNP *DIO2*-C allele creates four unique TFBS for the BRCA1, NFYA, RUNX1 and RUNX2 TFs which are involved with genomic stability, stimulation of transcription of many genes, development of normal hematopoiesis and maturation of osteoblasts, respectively (Table, supplement). The rare novel SNP *DIO2*-T allele creates nine unique TFBS for the ARID3A, CDX2, GFI1, HOXA9, NKX2-5, NOBOX, PBX1, SOX17 and STAT3 TFs which are involved with cell cycle progression, cell growth and differentiation, hematopoiesis and oncogenesis, the developmental regulatory system, negative regulation of chondrocyte maturation, oogenesis, glucose-dependent regulation of insulin gene transcription, transcription repression, and cellular responses to interleukins, respectively (Fig. 2, Table, supplement). There are also seven conserved TFBS between the SNP alleles for the FOXD3, FOXI1, FOXQ1, MEIS1, NFYB, SRY and THAP1 TFs which are involved with transcriptional activation and repression, kidney function, follicle differentiation, normal development, binding CCAAT motifs in the promoter, male development and regulation of endothelial cell proliferation, respectively (Table, supplement).

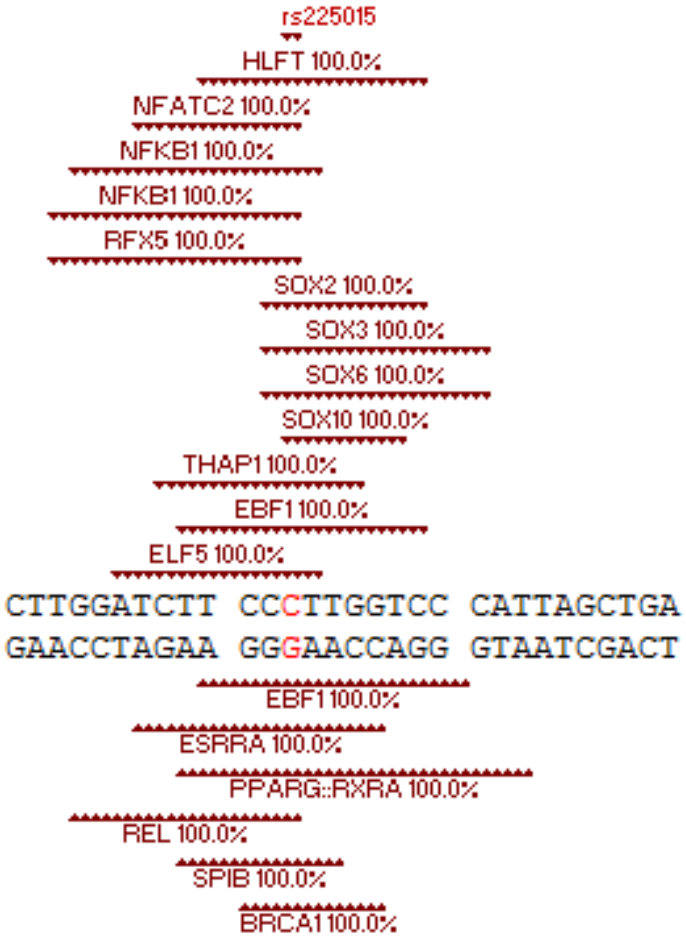

Fig. 1. Double stranded DNA from the DIO2 3'UTR showing the potential TFBS for eighteen different TFs which can bind their respective DNA sequence either above (+) or below (-) the duplex (cf. Table). The rs225015 common DIO2-G allele is found in each of these TFBS. As shown, this rSNP is located in the 3'UTR of the DIO2 gene. Also included with the potential TFBS is their % sequence homology to the duplex

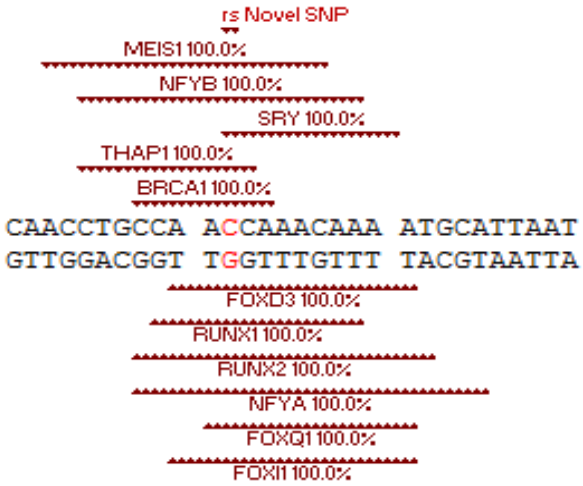

Fig. 2. Double stranded DNA from the DIO2 5'UTR showing the potential TFBS for eleven different TFs which can bind their respective DNA sequence either above (+) or below (-) the duplex (cf. Table). The novel -2035bp rSNP common DIO2-C allele is found in each of these TFBS. As shown, this rSNP is located in the 5'UTR of the DIO2 gene. Also included with the potential TFBS is their % sequence homology to the duplex

4. DISCUSSION

The genome-wide association studies (GWAS) has over the past decade provided us with nearly 6,500 disease or trait-predisposing SNPs. Only seven percent of these SNPs are located in protein-coding regions of the genome [38,39] while the remaining 93% are located within non-coding regions [40,41] such as gene regulatory or intergenic areas of the genome. Much attention has been drawn to SNPs that occur in the putative regulatory of a gene where a single nucleotide change in the DNA sequence of a potential TF motif may affect the process of gene regulation [17,19,42]. A nucleotide change in a 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 BS as a result the gene is no longer regulated by the original TF [14,16]. Therefore, functional rSNPs in TFBS may result in differences in gene expression, phenotypes and susceptibility to environmental exposure [42]. Examples of rSNPs associated with disease susceptibility are numerous and several reviews have been published [42-45].

The rs225012 rSNP *DIO2*-G allele [G (- strand) or C (+ strand)] located in the E2F6 and ELF1 TFBSs have a 100% occurrence in humans while the EGR1 and SPI1 TFBS have a 94% and 92% occurrence, respectively (Table). Since these binding sites (BS) occurs only once in the gene, this rSNP would probably have a major impact on these TFs regulating the gene. The ERG and SP1 TFBSs also have a 100% occurrence in humans but these BS occur more than once in the gene and should be the rSNP would not have much of an impact gene regulation (Table). The alternate rs225012 rSNP *DIO2*-A allele [A (- strand) or T (+ strand)] located in the PRRX2 TFBS has a 100% occurrence in humans but occurs 55 times in the gene therefore the rSNP would not be expected to have an impact on the TFs regulating the gene (Table). On the other hand, the rs225012 rSNP *DIO2*-A allele located in HOXA5 and NKX3-2 TFBSs also have a 100% occurrence in humans and occur only once in the gene and therefore, should have a major impact on gene regulation since these BS only occur with the minor allele (Table). The E2F6 TFBS provided by the rs225012 rSNP common G allele and not present with the minor A allele is a BS for a TF which is involved with the control of the cell cycle and the action of tumor suppressor proteins. Consequently individuals carrying the rs225012 rSNP *DIO2*-A allele maybe at risk for sickness or disease. In fact, the rs225012 rSNP *DIO2* AA genotype [TT genotype (- strand)] frequency has been significantly associated with MR [9] in Chinese patients.

Table 1. The DIO2 SNPs that were examined in this study where the minor allele is in red. Also listed are the transcriptional factors (TF), their potential binding sites (TFBS) containing these SNPs and DNA strand orientation. TFs in red differ between the SNP alleles. Where upper case nucleotide designates the 90% conserved BS region and red is the SNP location of the alleles in the TFBS. Below the TFBS is the nucleotide occurrence (%) obtained from the Jaspar Core database. Also listed are the number (#) of binding sites in the gene for the given TF. Note: TFs can bind to more than one nucleotide sequence.

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
Novel Rare (-)2035bp TSS	C	BRCA1	breast cancer 1, early onset	3	ccAacca c=67%	minus
		FOXD3	Forkhead box D3	1	tttTgTTtggtt g=1%	plus
		FOXI1	Forkhead box I1	1	tttTgTTTggtt g=23%	plus
		FOXQ1	Forkhead box Q1	1	ttttGTTTggt g=22%	plus
		MEIS1	Meis homeobox 1	1	accTGcCAaccaaac c=21%	minus
		NFYA	Nuclear transcription factor Y, alpha	1	tgcattttgtTTGGttgg G=99%	plus
		NFYB	nuclear transcription factor Y, beta	1	ctgccaaCCAAacaa C=100%	minus
		RUNX2	Runt-related transcription factor 12	1	attttgTtTGGttgg G=100%	plus
		RUNX1	Runt-related transcription factor 1	1	ttgTttGGttg G=92%	plus
		SRY	Sex determining region Y	1	ccaaACAAa c=18%	minus
		THAP1	THAP domain containing, apoptosis associated protein 1	1	ctgCCaacc c=36%	minus
	T 0.004	ARID3A	AT rich interactive domain 3A (BRIGHT-like)	12	ATcAAa T=100%	minus
		CDX2	Caudal type homeobox 2	2	tgccaATcAAa T=100%	minus
		FOXI1	Forkhead box I1	1	tttTgTTTgatt a=1%	plus

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
		FOXD3	Forkhead box D3	1	tttTgTTtgatt a=13%	plus
		GFI1	Growth factor independent 1 transcription repressor	1	ccAATCaaac T=98%	minus
		HOXA9	Homeobox A9	1	ccaATcAAaCa T=100%	minus
		HOXC9	Homeobox C9	1	gccaaTcAAacaa T=99%	minus
		MEIS1	Meis homeobox 1	1	accTGcCAatcaaac t=57%	minus
		NFYB	nuclear transcription factor Y, beta	1	caacctgCCAATcaa T=100%	minus
		NKX2-5	Natural killer 3 homeobox 2	2	ttgAttg A=100%	plus
		NKX2-5 (var.2)	Natural killer 2 homeobox 5	2	tgcCaaTCaaa T=100%	minus
		Nobox	NOBOX oogenesis homeobox	1	TgATTggc A=100%	plus
		PBX1	Pre-B-cell leukemia homeobox 1	2	ctgcCAATCAaa T=94%	minus
		PBX1	Pre-B-cell leukemia homeobox 1	1	caaTCAAaCaaa T=90%	minus
		SOX17	SRY (sex determining region Y)-box 17	4	ttgATTggc A=97%	plus
		SRY	Sex determining region Y	1	tcaaACAAa t=29%	minus
		STAT3	Signal transducer and activator of transcription 3 (acute-phase response factor)	2	tTgattGGcAg a=5%	plus
		THAP1	THAP domain containing, apoptosis associated protein 1	1	ctgCCaatc t=29%	minus
rs12885300 (C/T) 5' UTR	C	FOXP1	Forkhead box P1	1	aaaggctAAAgAaaa g=28%	plus
		PRDM1	PR domain containing 1, with ZNF domain	1	agacAatGAAAGgct	plus

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
					g=11%	
		SOX3	SRY (sex determining region Y)-box 3	1	cctTTcattg c=72%	minus
		TCF7L2	Transcription factor 7-like 2 (T-cell specific, HMG-box)	1	gacaaTgAAAGgct g=48%	plus
	T 0.229	ARID3A	AT rich interactive domain 3A (BRIGHT-like)	15	ActAAa A=100%	plus
		BATF::JUN	Basic leucine zipper transcription factor, ATF-like Jun proto-oncogene	3	tgaaaGAcTaA A=100%	plus
		FOXP1	Forkhead box P1	1	aaagactAAAgAaaa a=44%	plus
		IRF1	Interferon regulatory factor 1	1	ttagtcTTTCatTgtctctat t=12%	minus
		JUN:FOS	*Jun* proto-oncogene FBJ murine osteosarcoma viral oncogene homolog	10	TgAaagA A=90%	plus
		PAX2	Paired box gene 2	1	agtCtttc t=84%	minus
		PRDM1	PR domain containing 1, with ZNF domain	1	agacAatGAAAGact a=3%	plus
		SOX3	SRY (sex determining region Y)-box 3	3	cttTaGTctt T=100%	minus
		SOX6	SRY (sex determining region Y)-box 6	3	cttTaGTctt T=100%	minus
		TCF7L2	Transcription factor 7-like 2 (T-cell specific, HMG-box)	1	gacaaTgAAAGact a=25%	plus
rs225010 (C/T)	C	ELF5	E74-like factor 5	1	cttaTCCgt g=25%	plus
		ELK1	ELK1, member of ETS oncogene family	1	gtgacgGAta c=86%	minus
		GATA1	GATA binding protein 1	12	agcTTATCCcgt g=14%	plus
		GATA2	GATA binding protein 2	1	gtgagcTTATCcgt g=45%	plus
		GATA4	GATA binding protein 4	1	gcTTATCcgtc	plus

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
					g=34%	
		JUN:FOS	*Jun* proto-oncogene FBJ murine osteosarcoma viral oncogene homolog	1	TgAcggA c=67%	minus
		KLF5	Kruppel-like factor 5 (intestinal)	1	ccgtCaCCCa g=1%	plus
		MAFB	v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	1	Ggtgacgg c=53%	minus
		NFE2L1:MAFG	Nuclear factor erythroid 2-related factor 1 Transcription factor MafG	7	ggTGAc c=76%	minus
		PAX2	Paired box gene 2	1	ggtgacgg c=55%	minus
		PAX2	Paired box gene 2	1	cgtCaccc g=68%	plus
		SREBF1	Sterol regulatory element binding transcription factor 1	1	gTCAcccaaa g=28%	plus
	T 0.415	GATA1	GATA binding protein 1	3	agctTATCcat a=14%	plus
		HOXA9	Hoxa9	4	cttATccATCa A=99%	plus
		JUND (var.2)	Jun D proto-oncogene	1	atggaTaAgctcAct t=25%	minus
		KLF5	Kruppel-like factor 5 (intestinal)	1	ccatCaCCCa a=1%	plus
		MAFB	v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	12	Ggtgatgg t=20%	minus
		NFATC2	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2	1	ttaTCCa a=69%	plus
		NFE2L1:MAFG	Nuclear factor erythroid 2-related factor 1 Transcription factor MafG	6	caTcAc a=85%	plus
		PAX2	Paired box gene 2	12	ggtgatgg t=35%	minus
		ZNF354C	Zinc finger protein 354C	9	atCCAt A=100%	plus
rs225011	C	CRX	Cone-rod homeobox	1	cctagGaTTAt	plus

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
(C/T)					G=100%	
		ESRRA	Estrogen-related receptor alpha	1	cctAGGtCAgt c=57%	minus
		ESRRB	Estrogen-related receptor beta	1	aatcctAGGTCA c=34%	minus
		GATA4	GATA binding protein 4	1	tgTTATaatcc c=37%	minus
		NKX2-5	Natural killer 3 homeobox 2	55	atAAtcc c=12%	minus
		NR5A2	Nuclear receptor subfamily 5, group A, member 2	1	taatcCtAGGtCagt c=21%	minus
		PRRX2	Paired related homeobox 2	34	gATTA g=7%	plus
		RORA_1	RAR-related orphan receptor A	2	tcctaGGTCA c=4%	minus
		RORA_2	RAR-related orphan receptor A	1	taatccTaGGTCAg c=1%	minus
		RXRA	Retinoid X receptor, alpha	1	cctAGgtCAgt c=85%	minus
	T 0.415	ESRRA	Estrogen-related receptor alpha	1	tctAGGtCAgt t=30%	minus
		ESRRB	Estrogen-related receptor beta	1	aattctAGGTCA t=47%	minus
		FOXL1	Forkhead box L1	2	gaattATA a=43%	plus
		GATA4	GATA binding protein 4	1	tgTTATaattc t=36%	minus
		MEF2A	Myocyte enhancer factor 2A	1	acctagAAtTAtaac A=92%	plus
		NKX2-5	Natural killer 3 homeobox 2	3	atAAttc t=65%	minus
		NR5A2	Nuclear receptor subfamily 5, group A, member 2	1	taattCtAGGTCAgt t=74%	minus
		PDX1	Pancreatic and duodenal homeobox 1	23	aTAATt	minus

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
					t=61%	
		PRRX2	Paired related homeobox 2	68	aATTA	plus
					a=88%	
		RORA_1	RAR-related orphan receptor A	1	ttctaGGTCA	minus
					t=60%	
		RORA_2	RAR-related orphan receptor A	1	taattcTaGGTCAg	minus
					t=36%	
rs225012 (A/G)	A 0.419	EHF	Ets homologous factor	2	aCTTCtTc	plus
					T=100%	
		ELF5	E74-like factor 5	2	tactTCttc	plus
					T=98%	
		EN1	Engrailed homeobox 1	1	aagtagagata	minus
					a=50%	
		FOXC1	Forkhead box C1	4	aagaaGTA	minus
					a=44%	
		HLTF	Helicase-like transcription factor	1	ctaCtTcttc	plus
					t=58%	
		HOXA5	Hoxa5	1	ctctacTt	plus
					T=100%	
		NKX3-2	NK3 homeobox *2*	1	agaAGTaga	minus
					A=100%	
		PDX1	Pancreatic and duodenal homeobox 1	12	cTAcTt	plus
					T=97%	
		PRRX2	Paired related homeobox 2	55	aAgTA	minus
					A=100%	
		SPIB	Spi-B transcription factor (Spi-1/PU.1 related)	15	agaaGAA	minus
					A=96%	
	G	E2F6	E2F transcription factor 6	1	aaGaGGtAgag	minus
					G=100%	
		EGR1	Early growth response 1	1	ctacCtCttctgcc	plus
					C=94%	
		ELF1	E74-like factor 1 (ets domain transcription factor)	1	agaagaGGtAGag	minus
					G=100%	
		ERG	v-ets avian erythroblastosis virus E26	3	agAGGtAgaga	minus

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
			oncogene homolog		G=100%	
		FOXC1	Forkhead box C1	1	aagagGTA g=25%	minus
		SP1	Specificity Protein 1	3	tctCtaCCtct C=100%	plus
		SPI1	Spleen focus forming virus (SFFV) proviral integration oncogene spi1	1	cagaagaGGtAgaga G=92%	minus
		SPIB	Spi-B transcription factor (Spi-1/PU.1 related)	5	agaGGtA G=96%	minus
		ZNF263	Zinc finger protein 263	1	tggGgcagaagaggtagagat g=58%	minus
		ZNF263	Zinc finger protein 263	1	gggGcAgaagaggtagagata g=44%	minus
rs225013 (G/T)	G	FOXA1	Forkhead box A1	1	gtatTcTTtaCtatc g=19%	minus
		FOXC1	Forkhead box C1	10	ggctgGTA G=100%	minus
		HAND1:TCFE2α	Heart- and neural crest derivatives-expressed protein 1: transcription factor E2A	1	aggCTGgtat g=86%	minus
	T 0.364	BATF::JUN	Basic leucine zipper transcription factor, ATF-like Jun proto-oncogene	1	gtaaaGAaTaA A=98%	plus
		BRCA1	breast cancer 1, early onset	4	atAacag a=84%	plus
		FOXA1	Forkhead box A1	1	ttatTcTTtaCtatc t=40%	minus
		GFI1B	Growth factor independent 1B transcription repressor	1	gAAtaacaGcc a=59%	plus
		HNF1A	HNF1 homeobox A	1	tGTTAtTctTtact T=95%	minus
		HOXA5	Hoxa5	2	ctgttaTt t=56%	minus
		MYB	v-*myb* myeloblastosis viral oncogene homolog	1	atAACaGcCt A=93%	plus

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
		NFE2L1:MAFG	Nuclear factor erythroid 2-related factor 1 Transcription factor MafG	12	aaTaAc A=100%	plus
		PAX2	Paired box gene 2	2	tgttattc t=84%	minus
		SPZ1	Spermatogenic leucine zipper 1	1	agaataacagc a=50%	plus
rs225014 (C/T)	T	EGR1	Early growth response 1	1	tcacCtCCttCtgt t=14%	minus
		ELF1	E74-like factor 1 (ets domain transcription factor)	1	acagaaGGAgGtg a=31%	plus
		FOXC1	Forkhead box C1	1	cttctGTA T=100%	minus
		FOXC1	Forkhead box C1	1	ctccaGTA A=100%	plus
		HLTF	Helicase-like transcription factor	1	ctcCtTctgt t=25%	minus
		HOXA5	Hoxa5	1	ctgtacTg t=56%	minus
		RXR::RAR_DR 5	Retinoid X receptor: Retinoic acid receptor	1	aGtacAgaaggaggtgA a=1%	plus
		SPI1	Spleen focus forming virus (SFFV) proviral integration oncogene spi1	1	tacagaaGGAggtga a=42%	plus
		STAT5A::STAT5B	Signal transducer and activator of transcription 5A and transcription 5B	1	tgTctccaGtA A=90%	plus
		THAP1	THAP domain containing, apoptosis associated protein 1	1	tctCCagta a=58%	plus
	C 0.421	EBF1	Early B-cell factor 1	2	gtCtCcaGtGc G=100%	plus
		EGR1	Early growth response 1	1	tcacCtCCttCtgc c=46%	minus
		ELF1	E74-like factor 1 (ets domain transcription factor)	1	gcagaaGGAgGtg g=41%	plus

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
		HLTF	Helicase-like transcription factor	3	ctcCtTctgc c=20%	minus
		NKX3-2	NK3 homeobox *2*	6	tccAGTgca g=75%	plus
		RXR::RAR_DR5	Retinoid X receptor: Retinoic acid receptor	1	aGtgcagaaggaggtgA g=1%	plus
		SP2	*Sp2* transcription factor	1	tcaCCtCCttctgca c=47%	minus
		SPIB	Spi-B transcription factor (Spi-1/PU.1 related)	7	tgcaGAA g=59%	plus
		TFAP2C	Transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma)	1	catgtCtccAGtgca g=55%	plus
		ZNF354C	Zinc finger protein 354C	9	ctgCAC C=100%	minus
rs6574549 (G/T)	G 0.007	FOXA1	Forkhead box A1	1	aaaaTacTtaCatat C=93%	plus
		FOXA2	Forkhead box A2	4	TacTACatatt C=90%	plus
		FOXC1	Forkhead box C1	38	aatatGTA G=100%	minus
		FOXC1	Forkhead box C1	25	tgtaaGTA g=38%	minus
		FOXD3	Forkhead box D3	1	gtaagTatttttt g=55%	minus
		FOXI1	Forkhead box I1	1	acaTaTTTgtaa c=26%	plus
		FOXL1	Forkhead box L1	38	aatatgTA g=9%	minus
		FOXL1	Forkhead box L1	23	cttacATA c=43%	plus
		HLTF	Helicase-like transcription factor	4	ttaCaTattt C=100%	plus
		HLTF	Helicase-like transcription factor	4	ataCtTacat	plus

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
					c=1%	
		NFIL3	Nuclear factor, interleukin 3 regulated	1	aTAtGTAAgta	minus
					G=91%	
		NKX3-1	NK3 homeobox *1*	2	tTACaTA	plus
					C=95%	
		NKX2-5	Natural killer 3 homeobox 2	3	gtAAgta	minus
					g=6%	
		POU2F2	POU class 2 homeobox 2	1	aatacTTaCATat	plus
					C=99%	
	T	ARID3A	AT rich interactive domain 3A (BRIGHT-like)	28	ATttAa	minus
					t=63%	
		ARID3A	AT rich interactive domain 3A (BRIGHT-like)	17	cTtAAa	plus
					A=100%	
		FOXC1	Forkhead box C1	2	tttaaGTA	minus
					t=19%	
		FOXD3	Forkhead box D3	1	aaaTaTTtaagt	minus
					T=98%	
		FOXI1	Forkhead box I1	1	aaaTaTTTgtaa	plus
					a=26%	
		FOXL1	Forkhead box L1	3	cttaaATA	plus
					a=17%	
		FOXL1	Forkhead box L1	2	tttaagTA	minus
					T=30%	
		HLTF	Helicase-like transcription factor	1	ataCtTaaat	plus
					a=1%	
		HNF1B	HNF1 homeobox B	1	caAaTatTTAAg	minus
					T=90%	
		HOXA5	Hoxa5	3	cttaaaTa	plus
					a=88%	
		LHX3	LIM homeobox 3	1	tacTTAAaTattt	plus
					A=100%	
		LHX3	LIM homeobox 3	1	tatTTAAgTattt	minus
					T=100%	
		NKX2-5	Natural killer 3 homeobox 2	4	ttAAgta	minus

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
					t=53%	
		NKX3-1	NK3 homeobox *1*	9	aTAtTTA	minus
					T=95%	
		NKX3-2	NK3 homeobox *2*	1	ttaAGTatt	minus
					t=38%	
rs225015 (A/G)	G	BRCA1	breast cancer 1, early onset	1	ccAaggg	minus
					g=14%	
		EBF1	Early B-cell factor 1	1	ggacCaaGgGa	minus
					G=97%	
		EBF1	Early B-cell factor 1	1	ttCcCttGgtc	plus
					C=100%	
		ELF5	E74-like factor 5	1	atctTCCct	plus
					c=7%	
		ESRRA	Estrogen-related receptor alpha	1	ccAAGGgaAga	minus
					G=98%	
		HLTF	Helicase-like transcription factor	1	tccCtTggtc	plus
					C=100%	
		NFATC2	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2	4	tcTTCCc	plus
					c=12%	
		NFKB1	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	1	gGatcTTcCCt	plus
					C=98%	
		NFKB1	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	1	tGGatcTtCCc	plus
					c=47%	
		PPARγ::RXRα	Peroxisome proliferator-activated receptor γ Retinoid X receptor, alpha	1	atgggaccaagggaa	minus
					g=16%	
		REL	v-*rel* avian reticuloendotheliosis viral oncogene homolog	1	gggaagatcC	minus
					g=47%	
		RFX5	Regulatory factor X, 5 (influences HLA class II expression)	1	tccCttgGtccCatt	plus
					C=95%	
		SOX2	SRY (sex determining region Y)-box 2	1	CCtTgGTc	plus
					C=100%	
		SOX3	SRY (sex determining region Y)-box 3	1	cctTgGTccc	plus
					c=75%	
		SOX6	SRY (sex determining region Y)-box 6	1	cCtTgGTccc	plus

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
					C=90%	
		SOX10	SRY (sex determining region Y)-box 10	11	cttggT	plus
					c=86%	
		SPIB	Spi-B transcription factor (Spi-1/PU.1 related)	8	aagGGAA	minus
					g=14%	
		THAP1	THAP domain containing, apoptosis associated protein 1	1	cttCCcttg	plus
					c=68%	
	A	BRCA1	breast cancer 1, early onset	2	ccAaagg	minus
	0.403				a=16%	
		ELF1	E74-like factor 1 (ets domain transcription factor)	1	accaaaGGAAGat	minus
					a=66%	
		ELF5	E74-like factor 5	3	atctTCCtt	plus
					t=36%	
		ELK1	ELK1, member of ETS oncogene family	1	ccaaagGAag	minus
					a=7%	
		ERG	v-ets avian erythroblastosis virus E26 oncogene homolog	3	aaAGGAAGatc	minus
					A=96%	
		ETS1	Protein C-ets-1	1	ggatctTCCTttggt	plus
					T=92%	
		FLI1	Fli-1 proto-oncogene, ETS transcription factor	3	aaaGGAAgatc	minus
					a=74%	
		HLTF	Helicase-like transcription factor	2	ttcCtTtggt	plus
					t=58%	
		NFATC2	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2	8	tcTTCCt	plus
					t=15%	
		NFKB1	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	1	tGGatcTtCCt	plus
					t=98%	
		RUNX1	Runt-related transcription factor 1	1	tccTttGGtcc	plus
					T=97%	
		SOX2	SRY (sex determining region Y)-box 2	12	CCtTTGgt	plus
					t=55%	
		SOX3	SRY (sex determining region Y)-box 3	2	cctTTGgtcc	plus
					t=87%	
		SOX3	SRY (sex determining region Y)-box 3	2	cttTgGTccc	plus

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
					t=25%	
		SOX6	SRY (sex determining region Y)-box 6	2	cCtTTGgtcc	plus
					t=35%	
		SOX9	SRY (sex determining region Y)-box 9	3	cctTtGgtc	plus
					t=5%	
		SOX10	SRY (sex determining region Y)-box 10	14	cttTgg	plus
					t=45%	
		SPIB	Spi-B transcription factor (Spi-1/PU.1 related)	10	aaaGGAA	minus
					a=49%	
		SPI1	Spleen focus forming virus (SFFV) proviral integration oncogene spi1	1	gaccaaaGGAAgatc	minus
					a=71%	
		TCF7L2	Transcription factor 7-like 2 (T-cell specific, HMG-box)	1	tgggaccAAAGgaa	minus
					A=100%	
rs225017 (A/T)	T	BATF::JUN	Basic leucine zipper transcription factor, ATF-like Jun proto-oncogene	1	gaaaTGAgaaA	minus
					a=78%	
		GATA4	GATA binding protein 4	7	tcTTATCtcat	plus
					t=27%	
		GATA4	GATA binding protein 4	4	ttTTcTCAttt	plus
					t=21%	
		HAND1:TCFE2α	Heart- and neural crest derivatives-expressed protein 1: transcription factor E2A	3	tgtCTGaaat	minus
					a=55%	
		HOXA5	Hoxa5	19	ctgaaaTg	minus
					a=88%	
		JUND (var.2)	Jun D proto-oncogene	1	tgaaaTGAgaaaAga	minus
					a=75%	
		NFE2L1:MAFG	Nuclear factor erythroid 2-related factor 1 Transcription factor MafG	11	aaTGAg	minus
					a=85%	
		PDX1	Pancreatic and duodenal homeobox 1	11	cTcATt	plus
					T=97%	
		SOX3	SRY (sex determining region Y)-box 3	4	cTtTTcTcat	plus
					t=48%	
		STAT3	Signal transducer and activator of transcription 3 (acute-phase response factor)	1	tTgtctGaAAt	minus
					A=100%	

SNP	Allele	TFs	Protein name	# of Sites	TFBS	Strand
	A 0.309	BATF::JUN	Basic leucine zipper transcription factor, ATF-like Jun proto-oncogene	1	gaatTGAgaaA t=7%	minus
		CEBPα	CCAAT/enhancer binding protein (C/EBP), alpha	1	tTTtCtcaAtt A=100%	plus
		CEBPβ	CCAAT/enhancer binding protein (C/EBP), beta	1	ctTTtCtcaAt A=100%	plus
		GATA4	GATA binding protein 4	1	tcTTtTCtcaa a=14%	plus
		HAND1:TCFE2α	Heart- and neural crest derivatives-expressed protein 1: transcription factor E2A	1	tgtCTGaatt t=34%	minus
		NKX2-5	Natural killer 3 homeobox 2	7	tcAAttc A=100%	plus
		NKX2-5	Natural killer 3 homeobox 2	7	tgAAttg t=59%	minus
		PRRX2	Paired related homeobox 2	68	aAtTA t=98%	minus
		SRY	Sex determining region Y	2	tttctCAAt A=96%	plus

The rs225017 rSNP *DIO2*-T allele [A (- strand) or T (+ strand)] located in the JUND (var.2) and STAT3 TFBS have in humans a 75% and 100% occurrence, respectively (Table). Since these BS occurs only once in the gene, this rSNP would probably have a major impact on these TFs regulating the gene. The HOXA5, NFE2L1: TCFE2α and PDX1 TFBSs have an 88%, 85% and 97% occurrence, respectively, in humans but these BS occur more than once in the gene and consequently, the rs225017 rSNP might not have much of an impact on gene regulation by these TFs (Table). The minor rs225017 rSNP *DIO2*-A allele [T (- strand) or A (+ strand)] located in the CEBPα & β have a 100% occurrence in humans and are found only once in the gene. Since these BS only occur once in the gene, the SNP would probably have a major impact on these enhancer and inflammation TFs regulating the gene. The NKX2-5, PRRX2 and SRY TFBS have in humans a 100%, 98% and 96% occurrence, respectively; however, these BS occur more than once in the gene and consequently this rSNP might not have much of an impact on DIO2 regulation by these TFs (Table).

Similar logic can be used to evaluate the potential TFBS within the other *DIO2* rSNPs found in the Table. It should be noted that the minor -2035bp novel rSNP T allele creates ten unique potential TFBS compared to the common C allele which creates only four BS while the rs225012 rSNP *DIO2* alleles each generate eight unique potential TFBS. In fact, 57 potential TFBS are created by the minor alleles of the nine SNPs compared to 39 TFBS created by the common alleles with 51 TFBS being shared by both alleles. Since the MAF of the nine SNPs ranges from 0.004 to 0.421, the potential TFBS generated by the minor alleles should have a tremendous impact on thyroid related illnesses and other sickness in humans. As an example, the POU2F2 (POU class 2 homeobox 2) TFBS is only created by the minor rare allele of rs6574549 and occurs only once in the gene which is important because it's a TF that binds in immunoglobulin gene promoters (supplement). This rSNP has been associated with fasting insulin, insulin action and energy expenditure in Pima Indians [11].

Human diseases or conditions can be associated with rSNPs of the *DIO2* gene as illustrated above. 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. As an example, the potential TFBS associated with the novel -2035bp common rSNP *DIO2*-C allele from Table are illustrated in Fig. 2 as well as the rs225015 rSNP *DIO2*-G allele illustrated in Fig. 1. As can be seen in Table, these potential TFBS change when an individual carries the minor allele. The importance of this can be illustrated with the BRCA1 TFBS where the common allele has this function and the minor allele does not. The BRCA1 TF plays a role in maintaining genomic stability and also acts as a tumor suppressor. Another example would be the PPARG::RXRA TFBS where the common allele has this function while the minor allele does not. This TF has been implicated in the pathology of numerous diseases including obesity, diabetes, atherosclerosis and cancer.

5. CONCLUSION

SNPs that alter the TFBS are not only found in the promoter regions but in the introns, exons and the UTRs of a gene (Table). The nucleus of the cell is where epigenetic alterations occur and TFs operate to convert chromosomes into single stranded DNA for mRNA transcription while it is the cytoplasm where mRNA is processed by separating exons and introns for protein translation. Consequently, it doesn't matter where TFs bind the DNA in the nucleus because it is only there that TFs function. The SNPs outlined in this report should be considered as rSNPs since they change the DNA landscape for TF binding and have been associated with disease. In this report, examples have been described to illustrate that a change in rSNP alleles in the *DIO2* gene can provide different TFBS which in turn are also associated with disease in humans. The potential alterations in TFBS obtained by computational analyses need to be verified by future protein/DNA electrophoretic mobility gel shift assays and gene expression studies.

CONSENT

It is not applicable.

ETHICAL APPROVAL

It is not applicable.

COMPETING INTERESTS

Author has declared that no competing interests exist.

REFERENCES

1. Danforth E Jr. The role of thyroid hormones and insulin in the regulation of energy metabolism. The American Journal of Clinical Nutrition. 1983;38:1006-17.
2. Krotkiewski M. Thyroid hormones in the pathogenesis and treatment of obesity. European Journal of Pharmacology 2002; 440:85-98.
3. Silva JE. Thyroid hormone control of thermogenesis and energy balance. Thyroid. 1995;5:481-92.
4. Freake HC, Oppenheimer JH. Thermogenesis and thyroid function. Annu Rev Nutr. 1995;15:263-91.
5. Bianco AC, Salvatore D, Gereben B, Berry MJ, Larsen PR. Biochemistry, cellular and molecular biology, and physiological roles of the iodothyronine selenodeiodinases. Endocr Rev. 2002;23:38-89.
6. Bianco AC, Larsen PR. Cellular and structural biology of the deiodinases. Thyroid. 2005;15:777-86.
7. Bartha T, Kim SW, Salvatore D, Gereben B, Tu HM, Harney JW, Rudas P, Larsen PR. Characterization of the 5'-flanking and 5'-untranslated regions of the cyclic adenosine 3',5'-monophosphate-responsive human type 2 iodothyronine deiodinase gene. Endocrinology. 2000; 141:229-37.
8. Celi FS, Canettieri G, Yarnall DP, Burns DK, Andreoli M, Shuldiner AR, Centanni M, Genomic characterization of the coding region of the human type II 5'-deiodinase gene. Mol Cell Endocrinol 1998;141: 49-52.
9. Guo TW, Zhang FC, Yang MS, Gao XC, Bian L, Duan SW, Zheng ZJ, Gao JJ, Wang H, Li RL, Feng GY, St Clair D, He L, Positive association of the DIO2 (deiodinase type 2) gene with mental retardation in the iodine-deficient areas of China. J Med Genet. 2004;41:585-90.
10. Meulenbelt I, Min JL, Bos S, Riyazi N, Houwing-Duistermaat JJ, van der Wijk HJ, Kroon HM, Nakajima M, Ikegawa S, Uitterlinden AG, van Meurs JB, van der Deure WM, Visser TJ, Seymour AB, Lakenberg N, van der Breggen R, Kremer D, van Duijn CM, Kloppenburg M, Loughlin J, Slagboom PE. Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. Hum Mol Genet. 2008;17: 1867-75.
11. Nair S, Muller YL, Ortega E, Kobes S, Bogardus C, Baier LJ. Association analyses of variants in the DIO2 gene with early-onset type 2 diabetes mellitus in Pima Indians. Thyroid. 2012;22:80-7.
12. Mentuccia D, Proietti-Pannunzi L, Tanner K, Bacci V, Pollin TI, Poehlman ET, Shuldiner AR, Celi FS. Association between a novel variant of the human type 2 deiodinase gene Thr92Ala and insulin resistance: Evidence of interaction with the Trp64Arg variant of the beta-3-adrenergic receptor. Diabetes. 2002;51:880-3.
13. Leiria LB, Dora JM, Wajner SM, Estivalet AA, Crispim D, Maia AL. The rs225017 polymorphism in the 3'UTR of the human DIO2 gene is associated with increased insulin resistance. PLoS One. 2014;9: e103960.
14. Buroker NE. ADRBK1 (GRK2) rSNPs, transcriptional factor binding sites and cardiovascular disease in the black population. Journal of Cardiovascular Disease. 2014;2.
15. 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.
16. Buroker NE. AKT3 rSNPs, transcritional factor binding sites and human disease. Open Journal of Blood Diseases. 2013;3: 116-29.
17. Knight JC. Functional implications of genetic variation in non-coding DNA for disease susceptibility and gene regulation. Clin Sci (Lond). 2003;104:493-501.
18. Knight JC. Regulatory polymorphisms underlying complex disease traits. Journal of Molecular Medicine. 2005;83:97-109.
19. 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.
20. 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;16:1188-200.
21. Claessens F, Verrijdt G, Schoenmakers E, Haelens A, Peeters B, Verhoeven G, Rombauts W. Selective DNA binding by the androgen receptor as a mechanism for hormone-specific gene regulation. The

Journal of Steroid Biochemistry and Molecular biology. 2001;76:23-30.

22. Hsu MH, Savas U, Griffin KJ, Johnson EF, Regulation of human cytochrome P450 4F2 expression by sterol regulatory element-binding protein and lovastatin. J Biol Chem. 2007;282:5225-36.
23. Takai H, Araki S, Mezawa M, Kim DS, Li X, Yang L, Li Z, Wang Z, Nakayama Y, Ogata Y. AP1 binding site is another target of FGF2 regulation of bone sialoprotein gene transcription. Gene. 2008;410:97-104.
24. Buroker NE, Huang JY, Barboza J, Ledee DR, Eastman RJ, Jr., Reinecke H, Ning XH, Bassuk JA, Portman MA. The adaptor-related protein complex 2, alpha 2 subunit (AP2alpha2) gene is a peroxisome proliferator-activated receptor cardiac target gene. The Protein Journal. 2012;31: 75-83.
25. Huang CN, Huang SP, Pao JB, Hour TC, Chang TY, Lan YH, Lu TL, Lee HZ, Juang SH, Wu PP, Huang CY, Hsieh CJ, Bao BY. Genetic polymorphisms in oestrogen receptor-binding sites affect clinical outcomes in patients with prostate cancer receiving androgen-deprivation therapy. Journal of Internal Medicine. 2012;271: 499-509.
26. Huang CN, Huang SP, Pao JB, Chang TY, Lan YH, Lu TL, Lee HZ, Juang SH, Wu PP, Pu YS, Hsieh CJ, Bao BY. Genetic polymorphisms in androgen receptor-binding sites predict survival in prostate cancer patients receiving androgen-deprivation therapy. Annals of oncology: Official Journal of the European Society for Medical Oncology / ESMO. 2012;23:707-13.
27. Yu B, Lin H, Yang L, Chen K, Luo H, Liu J, Gao X, Xia X, Huang Z. Genetic variation in the Nrf2 promoter associates with defective spermatogenesis in humans. Journal of Molecular Medicine; 2012.
28. Wu J, Richards MH, Huang J, Al-Harthi L, Xu X, Lin R, Xie F, Gibson AW, Edberg JC, Kimberly RP. Human FasL gene is a target of beta-catenin/T-cell factor pathway and complex FasL haplotypes alter promoter functions. PLoS One. 2011;6:e26143.
29. Alam M, Pravica V, Fryer AA, Hawkins CP, Hutchinson IV. Novel polymorphism in the promoter region of the human nerve growth-factor gene. International Journal of Immunogenetics. 2005;32:379-82.
30. Kumar A, Purohit R. Computational investigation of pathogenic nsSNPs in CEP63 protein. Gene. 2012;503:75-82.
31. Kamaraj B, Purohit R. Computational screening of disease-associated mutations in OCA2 gene. Cell Biochem Biophys. 2014;68:97-109.
32. Kumar A, Rajendran V, Sethumadhavan R, Shukla P, Tiwari S, Purohit R. Computational SNP analysis: Current approaches and future prospects. Cell Biochem Biophys. 2014;68:233-9.
33. Kumar A, Purohit R, Use of long term molecular dynamics simulation in predicting cancer associated SNPs. PLoS Comput Biol. 2014;10: e1003318.
34. 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:D102-6.
35. Sandelin A, Alkema W, Engstrom P, Wasserman WW, Lenhard B, JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res. 2004;32:D91-4.
36. Sandelin A, Wasserman WW, Lenhard B. ConSite: Web-based prediction of regulatory elements using cross-species comparison. Nucleic Acids Res. 2004;32: W249-52.
37. Buroker NE, Ning XH, Zhou ZN, Li K, Cen WJ, Wu XF, Zhu WZ, Scott CR, Chen SH. AKT3, ANGPTL4, eNOS3, and VEGFA associations with high altitude sickness in Han and Tibetan Chinese at the Qinghai-Tibetan Plateau. International Journal of Hematology. 2012;96:200-13.
38. Pennisi E. The biology of genomes. Disease risk links to gene regulation. Science. 2011;332:1031.
39. 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;34:567-80.
40. 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;106:9362-7.

41. Kumar V, Westra HJ, Karjalainen J, Zhernakova DV, Esko T, Hrdlickova B, Almeida R, Zhernakova A, Reinmaa E, Vosa 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.
42. 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; 659:147-57.
43. Prokunina L, Alarcon-Riquelme ME. Regulatory SNPs in complex diseases: Their identification and functional validation. Expert Rev Mol Med. 2004;6:1-15.
44. Buckland PR. The importance and identification of regulatory polymorphisms and their mechanisms of action. Biochim Biophys Acta. 2006;1762:17-28.
45. 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;89:355-65.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/9926