Open Journal of Blood Diseases, **2013, 3, 116-129**
Published Online December 2013 (http://www.scirp.org/journal/ojbd)
http://dx.doi.org/10.4236/ojbd.2013.34023

AKT3 rSNPs, Transcriptional Factor Binding Sites and Human Disease

Norman E. Buroker

Department of Pediatrics, University of Washington, Seattle, USA.
Email: nburoker@u.washington.edu

Received August 28th, 2013; revised September 28th, 2013; accepted October 5th, 2013

ABSTRACT

Seven rSNPs (rs10157763, rs10927067, rs12031994, rs2125230, rs2345994, rs4132509 and rs4590646) in intron one of thev-akt murine thymoma viral oncogene homolog 3 (*AKT*3) gene have been significantly associated with either Chronic Mountain Sickness, Renal Cell Carcinoma risk or Aggressive Prostate Cancer. These rSNP alleles alter the DNA landscape for potential transcriptional factors (TFs) to attach, resulting in changes in transcriptional factor binding sites (TFBS). The alleles of each rSNP were found to produce unique TFBS resulting in potential changes in TF *AKT*3 regulation. These regulatory changes are discussed with respect to the three diseases.

Keywords: *AKT*3; rSNPs; TFBS; Human Disease

1. Introduction

The phosphatidylinositol 3-kinase (PI3K)/AKT pathway plays a key role in numerous cellular functions including proliferation, adhesion, migration, invasion, metabolism and survival [1] as well as angiogenesis. The PI3K/AKT pathway modulates the expression of angiogenic factors such as nitric oxide and angiopoietins [2]. The v-akt murine thymoma viral oncogene homolog 3 (*AKT*3) is one of three isoforms of the AKTs which are major downstream targets of growth factor receptor tyrosine kinases that signal through PI3K [3]. The *AKT*3 single nucleotide polymorphism (SNP, rs4590656) from intron one was recently found to be significantly associated with three physiological parameters (hemoglobin, hematocrit and red blood cell count) in chronic mountain sickness patients indicating the strong association of this gene with angiogenesis [4]. Other intron one AKT3 SNPS (rs10157763, rs10927067 and rs2125230) have been significantly associated with aggressive prostate cancer (PCA) [5] in addition to the SNPs (rs4132509, rs12031994, rs2345994) which have been recently found to be significantly associated with renal cell carcinoma risk (RCC) [6]. This suggests that intron one maybe a regulatory region for the *AKT*3 gene. To follow up on this possibility these seven SNPs were examined for associations to potential transcription factor binding sites (TFBS).

Single nucleotide changes that affect gene expression by impacting gene regulatory sequences such as promoters, enhances, and silencers are known as regulatory SNPs (rSNPs) [7-10]. A rSNPs within a transcriptional factor binding site (TFBS) can change a transcriptional factor's (TF) ability to bind its TFBS [11-14] in which case the TF would be unable to effectively regulate its target gene [15-19]. This concept is examined for the above *AKT*3 rSNPs of intron one and their allelic association with TFBS. In this report, the rSNP associations with changes in potential TFBS and their possible relationship to the reported diseases or conditions have been examined.

2. Materials and Methods

2.1. Identifying TFBS

The JASPAR CORE database [20,21] and ConSite [22] 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 TFBS and rSNP loca-

tion within the binding sites have previously been discussed [4]. The Vector NTI Advance 11 computer program (Invitrogen, Life Technologies) was used to locate the TFBS in the *AKT*3 gene (NCBI Ref Seq NM_005465) from 11 kbp upstream of the transcriptional start site to 400 bp past the 3'UTR which represents a total of 35.5 Kbp.

3. Results

3.1. AKT3 rSNPs and TFBS

The *AKT*3 gene transcribes a serine/threonine kinase that plays a key role in regulating cell survival, insulin signaling, angiogenesis and tumor formation. The *AKT*3 SNP (rs4590656) has been found to be significantly associated with Hb and Hct in Tibetan Chinese with CMS. The *AKT*3-C allele creates two unique TFBS for the ARNT: AHR and HIF1α: ARNT TFs which are involved xenobiotic metabolism and cellular and systemic responses to hypoxia, respectively. The *AKT*3-T allele creates other unique TFBS for the HNF4α and PAX2 TFs which are involved in the expression of several hepatic genes and the conserved DNA-binding paired box, respectively which have previously been reported. In addition, the *AKT*3-T allele also creates two unique TFBS for the GFI TF which functions as a repressor and plays a role in hematopoiesis and oncogenesis and the SPIB TF which is a transcriptional activator that acts as a lymphoid-specific enhancer (**Table 1** and **Supplement**).

The *AKT*3 SNPs (rs4132509, rs12031994 and rs2345994) are significantly associated with the risk of RCC [6]. The rs4132509 *AKT*3-C allele creates one unique TFBS for the NFE2L1:MAFG TF, where NFE2L1 coordinates the up-regulation of cytoprotective genes via the antioxidant response element while MAFG is involved in cell differentiation of erythrocytes. The minor *AKT*3-A allele does not create any uniquely different TFBS (**Table 1** and **Supplement**). The rs12031994 *AKT*3-G allele creates seven unique TFBS for the FOXD1, FOXF2, FOXO3, NFYA, NKX2-5, SOX5 and SOX9 TFs those functions are described in the supplement. One TF standout is that FOXD1 may play a role in tumor formation which should have an effect on RCC. The minor *AKT*3-A allele creates nine unique TFBS for the EN1, GATA1, 2 & 3, HLTF, PDX1, PRRX2, TAL1: GALA1 and ZNF354C TFs those functions are also described in the supplement. Of these the TAL1: GALA1 TF is of interest since it is involved with the genesis of hemopoietic malignancies and may play a role in hemopoietic differentiation and a positive regulator of erythroid differentiation. The rs2345994 *AKT*3-G allele creates two unique TFBS for the SRY and SOX10 TFs which are part of a family of TFs involved in the regulation of embryonic development and in the determination of the cell fate (**Supplement**). The minor *AKT*3-A allele creates four unique TFBS for the ARID3A, FOXL1, MEF2A and NKX2-5 TFs of which the MEF2A TF is probably the most interesting since it activates many muscle-specific, growth factor-induced and stress-induced genes (**Table 1** & **Supplement**).

Table 1. Human diseases and AKT3 rSNPs found to be significantly associated in the referenced study. Common allele in bold. rSNPs alleles alter the transcriptional factor binding sites (TFBS) in intron one of the gene. The rSNP alleles are found only in these TFBS. CMS = Chronic Mountain Sickness; RCC = Renal Cell Carcinoma risk; PCA = Aggressive Prostate Cancer.

Disease	Ethnic Group	Sample size (N)	rSNPs	Allele	Disease	Normal	Reference	TFBS
CMS	Tibetan Chinese	48	rs4590656	**C**	0.59	0.6	4	ARNT:AHR, HIF1a:ARNT
				T	0.41	0.4		GFI, HNF4, PAX2, SPIB
RCC	Caucasians	577	rs4132509	**C**	0.83	n/a	6	NFE2L1:MAFG
				A	0.17	n/a		
			rs12031994	**G**	0.87	n/a	6	FOXD1, FOXF2, FOXO3, NFYA, NKX2-5, SOX5, SOX9
				A	0.13	n/a		EN1, GATA1, GATA2, GATA3, HLTF, PDX1, PRRX2, TAL1:GATA1, ZNF354C
			rs2345994	**G**	0.68	n/a	6	SRY, SOX10
				A	0.32	n/a		ARID3A, FOXL1, MEF2A, NKX2-5,
PCA	Caucasian males	688	rs10157763	**C**	0.67	0.71	5	ELF5, ELK1, Mycn, SPIB, SPI1, TFAP2A
				T	0.33	0.29		CTCF, NFATC2, SOX17, ZNF354C
			rs10927067	**G**	0.81	0.84	5	MAFB
				A	0.19	0.16		HNF1B, PDX1, PRRX2,
			rs2125230	**G**	0.8	0.83	5	ARNT:AHR, FEV, HIF1a:ARNT, SPI1,
				A	0.2	0.17		GATA1, HNF4a, HOXA5, IRF1, NR2F1, SOX17

The *AKT*3 SNPs (rs10157763, rs10927067 and rs-2125230) are significantly associated with PCA [5]. The rs10157763 *AKT*3-C allele creates six unique TFBS for the ELF5, ELK1, Mycn, SPIB, SPI1 and TFAP2A TFs whose function are described in the **Supplement**. Of these TFs the Mycn is probably the most interesting since it is a member of the MYC family and is associated with a variety of tumors. The minor *AKT*3-T allele creates four unique TFBS for the CTCF, NFATC2, SOX17 and ZNF354C TFs which are part of the transcription machinery for gene regulation (**Table 1** and **Supplement**). The rs10927067 *AKT*3-G allele creates one unique TFBS for the MAFB TF that is important in the regulation of lineage-specific hematopoiesis. The minor *AKT*3-A allele creates three unique TFBS for the HNF1B, PDX1 and PRRX2 TFs of which the HNF1B TF is of interest because it is a member of the homeodomain-containing superfamily of transcription factors and regulates development of the embryonic pancreas (**Table 1** and **Supplement**). The rs2125230 *AKT*3-G allele creates four unique TFBS for the ARNT: AHR, FEV, HIF1α: ARNT and SPI1 TFs. The ARNT: AHR and HIF1α: ARNT TFs are discussed above while the FEV and SPI1 TFs are involved as a transcriptional regulator and involved in myeloid and B-lymphoid cell development, respectively (**Table 1** and **Supplement**). The minor *AKT*3-A allele creates six unique TFBS for the GATA1, IRF1, HNF4a, NR2F1, SOX17 and HOXA5 TFs where the HOXA5 TF is of interest since it upregulates the tumor suppressor p53 and has an important role in tumorigenesis.

4. Discussion

Genome-wide association studies (GWAS) over the last decade have identified nearly 6500 disease or trait-predisposing SNPs where only 7% of these are located in protein-coding regions of the genome [23,24] and the remaining 93% are located within non-coding areas [25,26] such as regulatory or intergenic regions. SNPs which occur in the putative regulatory region of a gene where a single base change in the DNA sequence of a potential TFBS may affect the process of gene expression are drawing more attention [7,9,27]. A SNP in a TFBS can have multiple consequences. Often the SNP does not change the TFBS interaction nor does it alter gene expression since a transcriptional factor (TF) will usually recognize a number of different binding sites in the gene. In some cases the SNP may increase or decrease the TF binding which results in allele-specific gene expression. In rare cases, a SNP may eliminate the natural binding site or generate a new binding site. In which cases the gene is no longer regulated by the original TF. Therefore, functional rSNPs in TFBS may result in differences in gene expression, phenotypes and susceptibility to environmental exposure [27]. Examples of rSNPs associated with disease susceptibility are numerous and several reviews have been published [27-30].

The *AKT*3 SNP (rs4590656) in the potential TFBS and the corresponding nucleotide occurrence (%) in the BS (**Supplement**) have previously been discussed [4] as well as the number of other similar TFBS found in the gene. The rs4132509 *AKT*3-C allele located in the NFE2L1: MAFG TFBS has a 76% occurrence in humans; however, this BS occurs 140 other times in the gene (**Supplement**) and therefore the SNP would probably not have any impact on the regulation of the gene. The rs4132509 *AKT*3-A allele does not have any unique TFBS and therefore should not have any impact on gene regulation. The rs12031994 *AKT*3-G allele occurs in seven unique TFBS (**Table 1**), but perhaps the most interesting two are FOXF2 and NFYA which each occur only once in the gene. The rs12031994 *AKT*3-G allele [G (−strand) or C (+strand)] is a highly conserved nucleotide (97% and 93%) in these respective TFBS. The rs12031994 *AKT*3-A allele occurs in nine unique TFBS (**Table 1**), but perhaps the most interesting two are GATA1 and TAL1: GATA1 which again each only occurs once in the gene. The rs12031994 *AKT*3-A allele [A (−strand) or T (+strand)] is a highly conserved nucleotide (96% and 98%) in these respective TFBS. This SNP might be expected to have an influence on *AKT*3 regulation and could impact human disease. The rs2345994 *AKT*3-G allele occurs in two unique TFBS (SRY and SOX10) which both have multiple copies in the gene (**Table 1**). The rs2345994 *AKT*3-A allele occurs in four unique TFBS (**Table 1**), but perhaps the most interesting one is the MEF2A TFBS since it occurs only three times in the gene and the A nucleotide has a 98% occurrence (**Supplement**). The rs10157763 *AKT*3-C allele occurs in six unique TFBS (**Table 1**), but perhaps the most interesting TFBS is TFAP2A since it occurs four times in the gene and the allele is 100% conserved (**Supplement**). The rs10157763 *AKT*3-T allele occurs in four unique TFBS (**Table 1**) but perhaps the most important TFBS is the CTCF TF which occurs only once in the gene and is a transcriptional regulator protein with 11 highly conserved zinc finger domains. The rs10927067 *AKT*3-G allele occurs only in one unique TFBS which is MAFB. The MAFB TFBS occurs only two times in the gene and the G nucleotide has an 80% occurrence in the sequence for humans. The rs10927067 *AKT*3-A allele occurs in three unique TFBS, but perhaps only the HNF1B TFBS would be of interest since it occurs once in the gene and the A nucleotide has an 80% occurrence in humans. The rs2125230 *AKT*3-G allele occurs in four unique TFBS, but only the FEV and HIF1a:ARNT TFBS would be of interest since they both occur three times in the gene and G nucleotide occurrence is 100% and 99%, respectively. The rs2125230 *AKT*3-A allele has six unique

TFBS, but perhaps the IRF1 and NR2F1 would be of interest since they both occur only once in the gene and A nucleotide occurrence is 95% and 92%, respectively.

Kidney cancer accounted for an estimated 4% of new cancer cases in the United States in 2012 [31] and approximately 85% of the kidney cancers are renal cell carcinomas (RCCs) [32]. There has been compelling evidence for genetic susceptibility to RCC [33-35]. Recently, a RCC study involving SNP polymorphisms in the mTOR pathway which includes the *AKT*3 gene found that obesity, weight gain, physical activity and genetic variants may influence susceptibility to the disease [6]. After analyzing three AKT3 SNPs which were found to be statistically significant in the RCC study, it was determined that these SNPs do alter the potential TFBS for TFs regulating the gene (**Table 1**). Perhaps the most interesting TFBS change occurs with the SNP rs12031994; as an example, TAL1: GATA1 occurs only with the minor *AKT*3-A allele and is implicated in the genesis of hemopoietic malignancies.

Prostate cancer (PCA) is the second leading cause of cancer-related deaths among men in the United States [36] and the American Cancer Society estimates that 26% - 29% of all new cancer cases and cancer-related deaths are attributed to PCA cancer [36]. PCA risk factors include older age, ethnic race and family history while other potential contributors include lifestyle, genetic factors and imbalances in biological pathways [5]. GWAS has resulted in the detection of numerous PCA susceptibility loci [37] from which three *AKT*3 SNPs that have been previously associated with PCA [5] have been found to be also associated with changes in TFBS within intron one. Perhaps the most interesting TFBS change occurs with the rs10157763 SNP whose minor allele *AKT*3-T creates the chromatin insulator-binding factor (CTCF) TFBS that is found only once in the gene (**Table 1** and **Supplement**). CTCF is a highly conserved zinc finger TF implicated in diverse genomic regulatory functions, including transcriptional activation/repression, insulation, imprinting [38] which restricts enhancers from activating *AKT*3 and thereby restraining angiogenesis [39]. The CTCF protein has been shown to be involved with chromatin loop formation and co-associates with the zinc finger domains [38] such as the Zfx TFBS. Also worth mentioning is the change in the interferon (IRF1) TFBS which occurs with the SNP rs2125230 minor allele *AKT*3-A allele. Although other IRF1 TFBS are found in the gene this BS has a 95% occurrence of the A nucleotide which is regulated by the SNP (**Supplement**). Changes in the IRS1 TFBS for the gene are important since interferon is involved in fighting tumors.

Human diseases or conditions that have been significantly associated with rSNPs of the AKT3 gene are shown in the **Table 1** along with rSNP allele-specific TFBS. 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 regulatory landscape can alter gene regulation which in turn can result in human disease, a change in condition or illness. In this report, examples have been described to illustrate that a change in rSNP alleles can provide different TFBS which in turn are also significantly associated with human disease.

REFERENCES

[1] A. G. Bader, S. Kang, L. Zhao and P. K. Vogt, "Oncogenic PI3K Deregulates Transcription and Translation," *Nature Reviews. Cancer*, Vol. 5, No. 12, 2005, pp. 921-929.

[2] J. Karar and A. Maity, "PI3K/AKT/mTOR Pathway in Angiogenesis," *Frontiers in Molecular Neuroscience*, Vol. 4, 2011, p. 1. http://dx.doi.org/10.3389/fnmol.2011.00051

[3] J. R. Testa and A. Bellacosa, "AKT Plays a Central Role in Tumorigenesis," *Proceedings of the National Academy Sciences of the United States of America*, Vol. 98, No. 20, 2001, pp. 10983-10985. http://dx.doi.org/10.1073/pnas.211430998

[4] N. E. Buroker, X. H. Ning, Z. N. Zhou, K. Li, W.J. Cen, X. F. Wu, W. Z. Zhu, C. R. Scott and S. H. Chen, "*AKT*3, *ANGPTL*4, *eNOS*3, and *VEGFA* Associations with high Altitude Sickness in Han and Tibetan Chinese at the Qinghai-Tibetan Plateau," *International Journal of Hematology*, Vol. 96, No. 2, 2012, pp. 200-213. http://dx.doi.org/10.1007/s12185-012-1117-7

[5] N. A. Lavender, E. N. Rogers, S. Yeyeodu, J. Rudd, T. Hu, J. Zhang, G. N. Brock, K. S. Kimbro, J. H. Moore, D. W. Hein and L. C. Kidd, "Interaction among Apoptosis-Associated Sequence Variants and Joint Effects on Aggressive Prostate Cancer," *BMC Medical Genomics*, Vol. 5, No. 11, 2012 , p. 1-15. http://dx.doi.org/10.1186/1755-8794-5-11

[6] X. Shu, J. Lin, C. G. Wood, N. M. Tannir and X. Wu, "Energy Balance, Polymorphisms in the mTOR Pathway, and Renal Cell Carcinoma Risk," *Journal of the National Cancer Institute*, Vol. 105, No. 6, 2013, pp. 424-432. http://dx.doi.org/10.1093/jnci/djt005

[7] J. C. Knight, "Functional Implications of Genetic Variation in Non-Coding DNA for Disease Susceptibility and Gene Regulation," *Clinical Science*, Vol. 104, No. 5, 2003, pp. 493-501. http://dx.doi.org/10.1042/CS20020304

[8] J. C. Knight, "Regulatory Polymorphisms Underlying Complex Disease Traits," *Journal of Molecular Medicine*, Vol. 83, No. 2, 2005, pp. 97-109. http://dx.doi.org/10.1007/s00109-004-0603-7

[9] X. Wang, D. J. Tomso, X. Liu and D. A. Bell, "Single Nucleotide Polymorphism in Transcriptional Regulatory Regions and Expression of Environmentally Responsive Genes," *Toxicology and Applied Pharmacology*, Vol. 207, No. 2, 2005, pp. 84-90. http://dx.doi.org/10.1016/j.taap.2004.09.024

[10] X. Wang, D. J. Tomso, B. N. Chorley, H. Y. Cho, V. G. Cheung, S. R. Kleeberger and D. A. Bell, "Identification of Polymorphic Antioxidant Response Elements in the Human Genome," *Human Molecular Genetics*, Vol. 16, No. 10, 2007, pp. 1188-1200. http://dx.doi.org/10.1093/hmg/ddm066

[11] F. Claessens, G. Verrijdt, E. Schoenmakers, A. Haelens, B. Peeters, G. Verhoeven and W. Rombauts, "Selective DNA Binding by the Androgen Receptor as a Mechanism for Hormone-Specific Gene Regulation," *The Journal of Steroid Biochemistry and Molecular Biology*, Vol. 76, No. 1-5, 2001, pp. 23-30. http://dx.doi.org/10.1016/S0960-0760(00)00154-0

[12] M. H. Hsu, U. Savas, K. J. Griffin and E. F. Johnson, "Regulation of Human Cytochrome P450 4F2 Expression by Sterol Regulatory Element-Binding Protein and Lovastatin," *Journal of Biological Chemistry*, Vol. 282, No. 8, 2007, pp. 5225-5236. http://dx.doi.org/10.1074/jbc.M608176200

[13] H. Takai, S. Araki, M. Mezawa, D. S. Kim, X. Li, L. Yang, Z. Li, Z. Wang, Y. Nakayama and Y. Ogata, "AP1 Binding Site Is Another Target of FGF2 Regulation of Bone Sialoprotein Gene Transcription," *Gene*, Vol. 410, No. 1, 2008, pp. 97-104. http://dx.doi.org/10.1016/j.gene.2007.11.017

[14] N. E. Buroker, J. Y. Huang, J. Barboza, D. R. Ledee, R. J. Eastman Jr., H. Reinecke, X. H. Ning, J. A. Bassuk and M. A. Portman, "The Adaptor-Related Protein Complex 2, Alpha 2 Subunit (AP2α2) Gene Is a Peroxisome Proliferator-Activated Receptor Cardiac Target Gene," *The Protein Journal*, Vol. 31, No. 1, 2012, pp. 75-83. http://dx.doi.org/10.1007/s10930-011-9379-0

[15] C. N. Huang, S. P. Huang, J. B. Pao, T. C. Hour, T. Y. Chang, Y. H. Lan, T. L. Lu, H. Z. Lee, S. H. Juang, P. P. Wu, C. Y. Huang, C. J. Hsieh and B. Y. Bao, "Genetic Polymorphisms in Oestrogen Receptor-Binding Sites Affect Clinical Outcomes in Patients with Prostate Cancer Receiving Androgen-Deprivation Therapy," *Journal of Internal Medicine*, Vol. 271, No. 5, 2012, pp. 499-509. http://dx.doi.org/10.1111/j.1365-2796.2011.02449.x

[16] C. N. Huang, S. P. Huang, J. B. Pao, T. Y. Chang, Y. H. Lan, T. L. Lu, H. Z. Lee, S. H. Juang, P. P. Wu, Y. S. Pu, C. J. Hsieh and B. Y. Bao, "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*, Vol. 23, No. 3, 2012, pp. 707-713. http://dx.doi.org/10.1093/annonc/mdr264

[17] B. Yu, H. Lin, L. Yang, K. Chen, H. Luo, J. Liu, X. Gao, X. Xia and Z. Huang, "Genetic Variation in the Nrf2 Promoter Associates with Defective Spermatogenesis in Humans," *Journal of Molecular Medicine*, Vol. 90, No. 11, 2012, pp. 1333-1342. http://dx.doi.org/10.1007/s00109-012-0914-z

[18] J. Wu, M. H. Richards, J. Huang, L. Al-Harthi, X. Xu, R. Lin, F. Xie, A. W. Gibson, J. C. Edberg and R. P. Kimberly, "Human *FasL* Gene Is a Target of β-Catenin/T-Cell Factor Pathway and Complex *FasL* Haplotypes Alter Promoter Functions," *PLoS ONE*, Vol. 6, No. 10, 2011, Article ID: e26143. http://dx.doi.org/10.1371/journal.pone.0026143

[19] M. Alam, V. Pravica, A. A. Fryer, C. P. Hawkins and I. V. Hutchinson, "Novel Polymorphism in the Promoter Region of the Human Nerve Growth-Factor Gene," *International Journal of Immunogenetics*, Vol. 32, No. 6, 2005, pp. 379-382. http://dx.doi.org/10.1111/j.1744-313X.2005.00541.x

[20] J. C. Bryne, E. Valen, M. H. Tang, T. Marstrand, O. Winther, I. da Piedade, A. Krogh, B. Lenhard and A. Sandelin, "JASPAR, the Open Access Database of Transcription Factor-Binding Profiles: New Content and Tools in the 2008 Update," *Nucleic Acids Research*, Vol. 36, Suppl. 1, 2008, pp. D102-D106. http://dx.doi.org/10.1093/nar/gkm955

[21] A. Sandelin, W. Alkema, P. Engstrom, W. W. Wasserman and B. Lenhard, "JASPAR: An Open-Access Database for Eukaryotic Transcription Factor Binding Profiles," *Nucleic Acids Research*, Vol. 32, Suppl. 1, 2004, pp. D91-D94. http://dx.doi.org/10.1093/nar/gkh012

[22] A. Sandelin, W. W. Wasserman and B. Lenhard, "Consite: web-Based Prediction of Regulatory Elements Using Cross-Species Comparison," *Nucleic Acids Research*, Vol. 32, Suppl. 2, 2004, pp. W249-W252. http://dx.doi.org/10.1093/nar/gkh372

[23] E. Pennisi, "The Biology of Genomes. Disease Risk Links to Gene Regulation," *Science*, Vol. 332, No. 6033, 2011, p. 1031. http://dx.doi.org/10.1126/science.332.6033.1031

[24] V. Kumar, C. Wijmenga and S. Withoff, "From Genome-Wide Association Studies to Disease Mechanisms: Celiac Disease as a Model for Autoimmune Diseases," *Seminars in Immunopathology*, Vol. 34, No. 4, 2012, pp. 567-580. http://dx.doi.org/10.1007/s00281-012-0312-1

[25] L. A. Hindorff, P. Sethupathy, H. A. Junkins, E. M. Ramos, J. P. Mehta, F. S. Collins and T. A. Manolio, "Potential Etiologic and Functional Implications of Genome-Wide Association Loci for Human Diseases and Traits," *Proceedings of the National Academy Sciences of the United States of America*, Vol. 106, No. 23, 2009, pp. 9362-9367. http://dx.doi.org/10.1073/pnas.0903103106

[26] V. Kumar, H. J. Westra, J. Karjalainen, D. V. Zhernakova, T. Esko, B. Hrdlickova, R. Almeida, A. Zhernakova, E. Reinmaa, U. Vosa, M. H. Hofker, R. S. Fehrmann, J. Fu, S. Withoff, A. Metspalu, L. Franke and C. Wijmenga, "Human Disease-Associated Genetic Variation Impacts Large Intergenic Non-Coding RNA Expression," *PLoS Genet*, Vol. 9, No. 1, 2013, Article ID: e1003201. http://dx.doi.org/10.1371/journal.pgen.1003201

[27] B. N. Chorley, X. Wang, M. R. Campbell, G. S. Pittman, M. A. Noureddine and D. A. Bell, "Discovery and Verification of Functional Single Nucleotide Polymorphisms in Regulatory Genomic Regions: Current and Developing Technologies," *Mutation Research*, Vol. 659, No. 1-2, 2008, pp. 147-157. http://dx.doi.org/10.1016/j.mrrev.2008.05.001

[28] L. Prokunina and M. E. Alarcon-Riquelme, "Regulatory SNPs in Complex Diseases: Their Identification and Functional Validation," *Expert Reviews in Molecular Medicine*, Vol. 6, No. 10, 2004, pp. 1-15.

http://dx.doi.org/10.1017/S1462399404007690

[29] P. R. Buckland, "The Importance and Identification of Regulatory Polymorphisms and Their Mechanisms of Action," *Biochimica et Biophysica Acta*, Vol. 1762, No. 1, 2006, pp. 17-28. http://dx.doi.org/10.1016/j.bbadis.2005.10.004

[30] W. Sadee, D. Wang, A. C. Papp, J. K. Pinsonneault, R. M. Smith, R. A. Moyer and A. D. Johnson, "Pharmacogenomics of the RNA World: Structural RNA Polymorphisms in Drug Therapy," *Clinical Pharmacology and Therapeutics*, Vol. 89, No. 3, 2011, pp. 355-365. http://dx.doi.org/10.1038/clpt.2010.314

[31] R. Siegel, D. Naishadham and A. Jemal, "Cancer Statistics, 2012," *CA*: *A Cancer Journal for Clinicians*, Vol. 62, No. 1, 2012, pp. 10-29. http://dx.doi.org/10.3322/caac.20138

[32] W. H. Chow, S. S. Devesa, J. L. Warren and J. F. Fraumeni Jr., "Rising Incidence of Renal Cell Cancer in the United States," *JAMA*, Vol. 281, No. 17, 1999, pp. 1628-1631. http://dx.doi.org/10.1001/jama.281.17.1628

[33] B. Schlehofer, W. Pommer, A. Mellemgaard, J. H. Stewart, M. McCredie, S. Niwa, P. Lindblad, J. S. Mandel, J. K. McLaughlin and J. Wahrendorf, "International Renal-Cell-Cancer Study. VI. The Role of Medical and Family History," *International Journal of Cancer*, Vol. 66, No. 6, 1996, pp. 723-726. http://dx.doi.org/10.1002/(SICI)1097-0215(19960611)66:6<723::AID-IJC2>3.0.CO;2-1

[34] J. Clague, J. Lin, A. Cassidy, S. Matin, N. M. Tannir, P. Tamboli, C. G. Wood and X. Wu, "Family History and Risk of Renal Cell Carcinoma: Results from a Case-Control Study and Systematic Meta-Analysis," *Cancer Epidemiology Biomarkers & Prevention*, Vol. 18, No. 3, 2009, pp. 801-807. http://dx.doi.org/10.1158/1055-9965.EPI-08-0601

[35] M. L. Nickerson, M. B. Warren, J. R. Toro, V. Matrosova, G. Glenn, M. L. Turner, P. Duray, M. Merino, P. Choyke, C. P. Pavlovich, N. Sharma, M. Walther, D. Munroe, R. Hill, E. Maher, C. Greenberg, M. I. Lerman, W. M. Linehan, B. Zbar and L. S. Schmidt, "Mutations in a Novel Gene Lead to Kidney Tumors, Lung Wall Defects, and Benign Tumors of the Hair Follicle in Patients with the Birt-Hogg-Dube Syndrome," *Cancer Cell*, Vol. 2, No. 2, 2002, pp. 157-164. http://dx.doi.org/10.1016/S1535-6108(02)00104-6

[36] American Cancer Society, "Cancer Facts and Figures," 2012.

[37] M. Yeager, N. Orr, R. B. Hayes, K. B. Jacobs, P. Kraft, S. Wacholder, M. J. Minichiello, P. Fearnhead, K. Yu, N. Chatterjee, Z. Wang, R. Welch, B. J. Staats, E. E. Calle, H. S. Feigelson, M. J. Thun, C. Rodriguez, D. Albanes, J. Virtamo, S. Weinstein, F. R. Schumacher, E. Giovannucci, W. C. Willett, G. Cancel-Tassin, O. Cussenot, A. Valeri, G. L. Andriole, E. P. Gelmann, M. Tucker, D. S. Gerhard, J. F. Fraumeni Jr., R. Hoover, D. J. Hunter, S. J. Chanock and G. Thomas, "Genome-Wide Association Study of Prostate Cancer Identifies a Second Risk Locus at 8q24," *Nature Genetics*, Vol. 39, No. 5, 2007, pp. 645-649.

[38] J. E. Phillips and V. G. Corces, "CTCF: Master Weaver of the Genome," *Cell*, Vol. 137, No. 7, 2009, pp. 1194-1211. http://dx.doi.org/10.1016/j.cell.2009.06.001

[39] M. Tang, B. Chen, T. Lin, Z. Li, C. Pardo, C. Pampo, J. Chen, C. L. Lien, L. Wu, L. Ai, H. Wang, K. Yao, S. P. Oh, E. Seto, L. E. Smith, D. W. Siemann, M. P. Kladde, C. L. Cepko and J. Lu, "Restraint of Angiogenesis by Zinc Finger Transcription Factor CTCF-Dependent Chromatin Insulation," *Proceedings of the National Academy Sciences of the United States of America*, Vol. 108, No. 37, 2011, pp. 15231-15236. http://dx.doi.org/10.1073/pnas.1104662108

Appendices

Supplemental material is available for this article.

Supplement. The AKT3 SNPs from intron 1 (147,408 bp) that were examined in this study. Also listed are the transcriptional factors (TF), their potential binding sites (TFBS) containing these SNPs and DNA strand orientation. 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	TF discription
rs4590656 (**C**/T)	**C**	ARNT:AHR	Aryl hydrocarbon receptor nuclear translocator aryl hydrocarbon receptor	18	tCCGTG C = 96%	minus	The dimer alters transcription of target genes. Involved in the induction of several enzymes that participate in xenobiotic metabolism.
		HIF1a:ARNT	Hypoxia-inducible factor 1: Aryl hydrocarbon receptor nuclear translocator	3	ctcCGTGa C = 99%	minus	HIF1 is a homodimeric basic helix-loop-helix structure composed of HIF1a, 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.
		FOXA1	Forkhead box A1	1	TgTTTgCtccg c = 1%	minus	Transcription factor that is involved in embryonic development, establishment of tissue-specific gene expression and regulation of gene expression in differentiated tissues.
		FOXA2	Forkhead box A2	1	TgTTtgCtccgt c = 1%	minus	Transcription factor that is involved in embryonic development, establishment of tissue-specific gene expression and regulation of gene expression in differentiated tissues.
	T	GFI	Growth factor independent 1 transcription repressor	4	aacATCacag a = 36%	plus	This gene encodes a nuclear zinc finger protein that functions as a transcriptional repressor. This protein plays a 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.
		FOXA1	Forkhead box A1	11	TgTTTgCtctg t = 54%	minus	Transcription factor that is involved in embryonic development, establishment of tissue-specific gene expression and regulation of gene expression in differentiated tissues.
		FOXA2	Forkhead box A2	1	TgTTtgCtctgt t = 45%	minus	Transcription factor that is involved in embryonic development, establishment of tissue-specific gene expression and regulation of gene expression in differentiated tissues.
		HNF4a	Hepatocyte nuclear factor 4, alpha	3	aGagCAAAcatca a = 42%	plus	A nuclear transcription factor which binds DNA as a homodimer and controls the expression of several genes including several hepatic genes. Transcriptionally controlled transcription factor

Continued

		PAX2	Paired box gene 2	11	caTCAcag a = 6%	plus	Probable transcription factor that may have a role in kidney cell differentiation.
		PAX2	Paired box gene 2	23	tgtgatgt t = 29%	minus	Probable transcription factor that may have a role in kidney cell differentiation.
		SPIB	Spi-B transcription factor (Spi-1/PU.1 related)	76	agaGcAA a = 63%	plus	The protein encoded by this gene is a transcriptional activator that binds to the PU-box (5'-GAGGAA-3') and acts as a lymphoid-specific enhancer.
rs4132509 (**C**/A)	**C**	FOXL1	Forkhead box L1	2	tcgtcATA g = 26%	plus	This gene encodes a member of the forkhead/winged helix-box (FOX) family of transcription factors.
		FOXL1	Forkhead box L1	18	gtcatATA g = 17%	plus	See above.
		NFE2L1:MAFG	Nuclear factor erythroid 2-related factor 1 Transcription factor MafG	140	taTGAc c = 76%	minus	Nuclear factor erythroid 2-related factor (Nrf2) coordinates the up-regulation of cytoprotective genes via 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.
		PAX2	Paired box gene 2	2	cgtCatat g = 68%	plus	See above.
		PAX2	Paired box gene 2	2	agtCtcgt g = 61%	plus	See above.
		TBP	TATA box binding protein	1	gtATAtgAcgagact c = 11%	minus	General transcription factor that functions at the core of the DNA-binding multiprotein factor TFIID.
	A	FOXL1	Forkhead box L1	31	tcttcATA t = 26%	plus	See above.
		FOXL1	Forkhead box L1	38	ttcatATA t = 48%	plus	See above.
		PAX2	Paired box gene 2	23	agtCtctt t = 35%	plus	See above.
		TBP	TATA box binding protein		gtATAtgAagagact a = 40%	minus	See above.
rs12031994 (**G**/A)	**G**	FOXA1	Forkhead box A1	4	TgTcTaCccag g = 86%	minus	See above.
		FOXA2	Forkhead box A2	1	TgTcTACccaga g = 72%	minus	See above.
		FOXD1	Forkhead box D1	7	gTAgACAa C = 90%	plus	This gene belongs to the forkhead family of transcription factors which is characterized by a distinct forkhead domain. The specific function of this gene has not yet been determined; however, it may play a role in tumor formation.

Continued

	FOXF2	Forkhead box F2	1	tctgggTAgACAat C = 93%	plus	FOXF2 is expressed in lung and placenta, and has been shown to transcriptionally activate several lung-specific genes.
	FOXO3	Forkhead box O3	2	ggtAgACA C = 92%	plus	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.
	IRF1	Interferon regulatory factor 1	2	gAcAatGAAAgc c = 1%	plus	IRF1 encodes interferon regulatory factor 1, a member of the interferon regulatory transcription factor (**IRF**) family. Interferons share several common effects; they are antiviral agents and can fight tumors.
	NFYA	Nuclear transcription factor Y, alpha	1	ggtagaCAATgaaagc C = 97%	plus	The protein encoded by this gene is one subunit of a trimeric complex, forming a highly conserved transcription factor that binds to CCAAT motifs in the promoter regions in a variety of genes.
	NKX2-5	NK2 homeobox *5*	148	ttcAttg g = 65%	minus	This gene encodes a member of the NK family of homeobox-containing proteins. This transcription factor functions in heart formation and development.
	PAX2	Paired box gene 2	33	tttCattg g = 35%	minus	See above.
	SOX5	SRY (sex determining region Y)-box 5	67	agACAAT C = 96%	plus	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.
	SOX9	SRY (sex determining region Y)-box 9	9	agaCaATga C = 95%	plus	The protein encoded by this gene recognizes the sequence CCTTGAG along with other members of the HMG-box class DNA-binding proteins.
	SOX10	SRY (sex determining region Y)-box 10	267	catTgT g = 86%	minus	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	267	ttcATTGTc G = 100%	minus	Acts as transcription regulator that binds target promoter DNA and bends the DNA.
	SRY	Sex determining region Y	80	gtagACAAt C = 93%	plus	Transcriptional regulator that controls a genetic switch in male development.
	EN1	Engrailed homeobox 1	1	taatgaaagcc t = 30%	plus	Homeobox-containing genes are thought to have a role in controlling development.
A	FOXA1	Forkhead box A1	1	TaTcTaCccag a = 12%	minus	Transcription factor that is involved in embryonic development, establishment of tissue-specific gene expression and regulation of gene expression in differentiated tissues.

Continued

	FOXA2	Forkhead box A2	1	TaTcTACccaga a = 27%	minus	Transcription factor that is involved in embryonic development, establishment of tissue-specific gene expression and regulation of gene expression in differentiated tissues.
	GATA1	GATA binding protein 1 (globin transcription factor 1)	1	gtaGATAAtga T = 96%	plus	Transcriptional activator which probably serves as a general switch factor for erythroid development.
	GATA2	GATA binding protein 2	1011	aGATa T = 100%	plus	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	342	aGATaa T = 92%	plus	Plays an important role in endothelial cell biology.
	HLTF	Helicase-like transcription factor	2	ttcatTatct a = 41%	minus	This gene encodes a member of the SWI/SNF family. Members of this family have helicase and ATPase activities and are thought to regulate transcription of certain genes by altering the chromatin structure around those genes.
	IRF1	Interferon regulatory factor 1	1	gAtAatGAAAgc t = 5%	plus	See above.
	PRRX2	Paired related homeobox 2	1039	cATTA A = 98%	minus	The DNA-associated protein encoded by this gene is a member of the paired family of homeobox proteins.
	SOX10	SRY (sex determining region Y)-box 10	347	catTaT a = 1%	minus	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	2	ttcATTaTc a = 1%	minus	Acts as transcription regulator that binds target promoter DNA and bends the DNA.
	SRY	Sex determining region Y	4	gtagAtAAt t = 7%	plus	See above.
	TAL1:GATA1	T-cell acute lymphocytic leukemia 1: GATA binding protein 1	1	attctctgggtaGATAat T = 98%	plus	Implicated in the genesis of hemopoietic malignancies. It may play an important role in hemopoietic differentiation. Serves as a positive regulator of erythroid differentiation.
	ZNF354C	Zinc finger protein 354C	164	atCtAC a = 44%	minus	May function as a transcription repressor.
rs2345994 **G** (**G**/A)	FOXI1	Forkhead box I1	4	ttgTaTTTgggg g = 26%	plus	See above.
	SOX10	SRY (sex determining region Y)-box 10	267	catTgT g = 86%	plus	See above.
	SRY	Sex determining region Y	70	aaatACAAa C = 7%	minus	Transcriptional regulator that controls a genetic switch in male development.

Continued

	A	ARID3A	AT rich interactive domain 3A (BRIGHT-like)	727	ATaAAt T = 100%	minus	This gene encodes a member of the ARID (AT-rich interaction domain) family of DNA binding proteins.
		FOXI1	Forkhead box I1	1	ttaTaTTTgggg a = 35%	plus	This gene belongs to the forkhead family of transcription factors which is characterized by a distinct forkhead domain.
		FOXL1	Forkhead box L1	24	gatttATA A = 91%	plus	See above.
		FOXL1	Forkhead box L1	37	caaatATA T = 91%	minus	See above.
		MEF2A	Myocyte enhancer factor 2	3	tTATaTtTgg A = 98%	plus	The protein encoded by this gene is a DNA-binding transcription factor that activates many muscle-specific, growth factor-induced, and stress-induced genes.
		NKX2-5	Natural killer 3 homeobox 2	148	atAtttg a = 41%	plus	See above.
rs10157763 (**C**/T)	**C**	ELF5	E74-like factor 5 (ets domain transcription factor)	20	atctTCCgc g = 25%	plus	The protein encoded by this gene is a member of an epithelium-specific subclass of the Ets transcription factor family. In addition to its role in regulating the later stages of terminal differentiation of keratinocytes, it appears to regulate a number of epithelium-specific genes found in tissues containing glandular epithelium such as salivary gland and prostate.
		ELK1	ELK1, member of ETS oncogene family	12	aaggcgGAag c = 86%	minus	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.
		ETS1	Protein C-ets-1	18	ctTCCg g = 43%	plus	The protein encoded by this gene belongs to the ETS family of transcription factors and has been shown to interact with TTRAP, UBE2I and Death associated protein.
		HLTF	Helicase-like transcription factor	7	cgcCtTgtgg g = 8%	plus	See above.
		Mycn	v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)	17	tcCAcagGgc c = 61%	minus	This gene is a member of the MYC family and encodes a protein with a basic helix-loop-helix (bHLH) domain. Amplification of this gene is associated with a variety of tumors, most notably neuroblastomas.
		MZF1_5-13	Myeloid zinc finger 1	12	caAgGcggaa c = 12%	minus	Binds to target promoter DNA and functions as trancription regulator. May be one regulator of transcriptional events during hemopoietic development. Isoforms of this protein have been shown to exist at protein level.
		SPI1	Spleen focus forming virus (SFFV) proviral integration oncogene spi1	8	cgGAAga c = 7%	minus	This gene encodes an ETS-domain transcription factor that activates gene expression during myeloid and B-lymphoid cell development.

Continued

		SPIB	Spi-B transcription factor (Spi-1/PU.1 related)	3	ggcGGAA c = 14%	minus	See above.
		TFAP2A	Transcription factor AP-2 alpha (activating enhancer binding protein 2 alpha)	4	GCCttgtgg G = 100%	plus	The protein encoded by this gene is a transcription factor that binds the consensus sequence 5'-GCCNNNGGC-3' and activates the transcription of some genes while inhibiting the transcription of others.
		ZEB1	Zinc finger E-box binding homeobox 1	19	cGCCTt G = 93%	plus	This gene encodes a zinc finger transcription factor. Acts as a transcriptional repressor.
	T	CTCF	CCCTC-binding factor (zinc finger protein)	1	attcCacaAGgtGgaagat t = 37%	minus	This gene is a member of the BORIS + **CTCF** gene family and encodes a transcriptional regulator 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.
		ETS1	Protein C-ets-1	277	ctTCCa a = 13%	plus	The protein encoded by this gene belongs to the ETS family of transcription factors and has been shown to interact with TTRAP, UBE2I and Death associated protein.
		HLTF	Helicase-like transcription factor	1	cacCtTgtgg a = 38%	plus	See above.
		MZF1_5-13	Myeloid zinc finger 1	1	caAgGtggaa t = 12%	minus	See above.
		NFATC2	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2	107	tcTTCCa a = 69%	plus	Member of the nuclear factor of activated T cells (NFAT) family. This protein is present in the cytosol and only translocates to the nucleus upon T cell receptor (TCR) stimulation,
		SOX17	SRY (sex determining region Y)-box 17	2	cacCTTGTc a = 26%	plus	See above.
		ZEB1	Zinc finger E-box binding homeobox 1	158	cACCTt a = 93%	plus	See above.
		ZNF354C	Zinc finger protein 354C	187	ttCCAC A = 100%	plus	May function as a transcription repressor.
rs10927067 (**G**/A)	**G**	GATA2	GATA binding protein 2	91	cGATc c = 13%	plus	See above.
		GATA2	GATA binding protein 2	648	tGATc G = 91%	minus	See above.
		GATA3	GATA binding protein 3	23	tGATcg G = 98%	minus	See above.
		MAFB	v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	2	Gctgatcg g = 80%	minus	The protein encoded by this gene is a basic leucine zipper (bZIP) transcription factor that plays an important role in the regulation of lineage-specific hematopoiesis.
	A	GATA2	GATA binding protein 2	94	cGATt t = 11%	plus	See above.

Continued

		GATA3	GATA binding protein 3	23	cGATta t = 31%	plus	See above.
		HNF1B	HNF1 homeobox B	1	ctAaTcgTTAAt a = 78%	minus	This gene encodes a member of the homeodomain-containing superfamily of transcription factors and has been shown to function in nephron development, and regulates development of the embryonic pancreas.
		PDX1	Pancreatic and duodenal homeobox 1	131	cTAATc A = 100%	minus	The protein encoded by this gene is a transcriptional activator of several genes, including insulin, somatostatin, glucokinase, islet amyloid polypeptide, and glucose transporter type 2.
		PRRX2	Paired related homeobox 2	818	gATTA T = 100%	plus	The DNA-associated protein encoded by this gene is a member of the paired family of homeobox proteins.
rs2125230 (**G**/A)	**G**	ARNT:AHR	aryl hydrocarbon receptor nuclear translocator aryl hydrocarbon receptor	27	ttCGTG C = 96%	plus	See above.
		FEV	FEV (ETS oncogene family)	3	cacGAAgT G = 100%	minus	This gene belongs to the ETS transcription factor family. Functions as a transcriptional regulator. This gene is exclusively expressed in neurons of the central serotonin (5-HT) system, a system implicated in the pathogeny of such psychiatric diseases as depression, anxiety, and eating disorders.
		HIF1a:ARNT	Hypoxia-inducible factor 1:Aryl hydrocarbon receptor nuclear translocator	3	cttCGTGc C = 99%	plus	See above.
		HLTF	Helicase-like transcription factor	1	taaCtTcgtg c = 8%	plus	See above.
		SOX10	SRY (sex determining region Y)-box 10	32	cttcgT c = 1%	plus	See above.
		SPI1	Spleen focus forming virus (SFFV) proviral integration oncogene spi1	13	acGAAgt G = 90%	minus	This gene encodes an ETS-domain transcription factor that activates gene expression during myeloid and B-lymphoid cell development.
	A	FOXA1	Forkhead box A1	22	TtTgTgCtttg T = 99%	plus	See above.
		GATA1	GATA binding protein 1 (globin transcription factor 1)	1	aaaGtTAAaga a = 36%	minus	See above.
		GFI	Growth factor independent 1 transcription repressor	1	caAAgCacaa a = 36%	minus	See above.
		HLTF	Helicase-like transcription factor	2	taaCtTtgtg t = 31%	plus	See above.
		HNF4a	Hepatocyte nuclear factor 4, alpha	1	aGcaCAAAcatca A = 76%	minus	See above.
		HNF4a	Hepatocyte nuclear factor 4, alpha	1	aGacCAAAgcaca a = 63%	minus	See above.

Continued

	HOXA5	Homeobox A5	21	cacaAAgt a = 38%	minus	This gene is part of the A cluster of A, B, C & D on chromosome 7 and encodes a DNA-binding transcription factor which may regulate gene expression, morphogenesis, and differentiation. Methylation of this gene may result in the loss of its expression and, since the encoded protein upregulates the tumor suppressor p53, this protein may play an important role in tumorigenesis.
	IRF1	Interferon regulatory factor 1	1	cAAAgttAAAga A = 95%	minus	See above.
	NR2F1	Nuclear receptor subfamily 2, group F, member 1	1	ttAaCtTTgtgctt T = 92%	plus	Coup transcription factor binds to the ovalbumin promoter and, in conjunction with another protein (S300-II) stimulates initiation of transcription. Binds to both direct repeats and palindromes of the 5'-AGGTCA-3' motif
	RUNX1	Runt-related transcription factor 1	35	cttTGtGcttt T = 97%	plus	Core binding factor (CBF) is a 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 CBF and is thought to be involved in the development of normal hematopoiesis.
	SOX10	SRY (sex determining region Y)-box 10	349	cttTgT T = 100%	plus	See above.
	SOX17	SRY (sex determining region Y)-box 17	7	aactTTGTg T = 100%	plus	See above.