Transcription Regulation and Candidate Diagnostic Markers of Esophageal Cancer

Magbubah Essack

Thesis presented in fulfillment of the requirements for the Degree of
Doctor Philosophiae at the South African National Bioinformatics
Institute, Faculty of Natural Sciences, University of the Western Cape

April 2009

Supervisor: Prof. Vladimir B. Bajic

DEDICATION

This thesis is dedicated to my mother,

Mariam Essack

and to the loving memory of my father,

Mogamat Gabieb Essack

who laid the foundation to make this accomplishment possible.

ABSTRACT

Transcription regulation and candidate diagnostic markers of esophageal cancer

PhD thesis, South African National Bioinformatics Institute, Faculty of Natural Sciences, University of the Western Cape

Esophageal cancer (EC) ranks among the ten most frequent cancers worldwide. Mortality rates associated with EC are very similar to the incidence rates due to the relatively late stage of diagnosis and the poor efficacy of treatment. The aim of this study was to enhance our insights of putative transcriptional circuitry of EC genes, thereby potentially positively impacting our knowledge of therapeutic targets, providing indications as to more appropriate lines of treatment, and additionally allowing for the determination of putative candidate diagnostic markers for the early stage detection of EC.

This thesis reports on the development of a novel comprehensive database (Dragon Database of Genes Implicated in Esophageal Cancer, DDEC) as an integrated knowledge database aimed at representing a gateway to esophageal cancer related data. More importantly, it illustrates how the biocurated genes in the database may represent a reliable starting point for divulging transcriptional regulation, diagnostic markers and the biology related to esophageal cancer.

DDEC contains known and novel information for 529 differentially expressed EC genes compiled using scientific publications from PubMed and is freely accessible for academic and non-profit users at http://apps.sanbi.ac.za/ddec/. The novel information provided to users of the DDEC is the lists of putative transcription factors that potentially control the 529 manually curated genes. The value of the information accessible through the database was further refined by providing precompiled text-mined and data-mined reports about each of these genes to allow for easy exploration of information about associations of EC-implicated genes with other human genes and proteins, metabolites and enzymes, toxins, chemicals with pharmacological effects, disease concepts and human anatomy. This feature has the capacity to display potential associations that are rarely reported and thus difficult to identify, and it enables the inspection of potentially new ‘association hypotheses’ generated based on the precompiled reports.

This study further illustrates how the biocurated esophageal squamous cell carcinoma (ESCC) genes in the database may represent a reliable starting point for exploring beyond current knowledge of the transcriptional circuitry of estrogen related hormone therapy. The genes were used to develop a method that identified 44 combinations of transcription factors (TFs) that characterize the promoter sequence of estrogen responsive genes implicated in ESCC. These significantly over-represented combinations of TFs were then used to increase confidence in the 47 novel putative estrogen response genes that may be related to ESCC too. Coincidently, two of the novel putative estrogen response genes were verified by current (2009), experimental publications.

The biocurated ESCC genes in the database were additionally used to expand on the current knowledge of the transcriptional circuitry of ESCC. The significantly over-represented transcription factors binding sites (TFBSs) were associated with the TFs that potentially bind at these sites. Since the TFs are known to regulate the gene expression, TFs were ranked and the top 10% of most regulating TFs; including, paired box 5 (PAX5), transcription factor AP-2 alpha (activating enhancer binding protein 2 alpha) (TFAP2A), Zic family member 2 (odd-paired homolog, Drosophila) (ZIC2), cAMP responsive element binding protein 1 (CREB1), E2F transcription factor 1 (E2F1), transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma) (TFAP2C), transcription factor Dp-1 (TFDP1), E2F transcription factor 3 (E2F3) and E2F transcription factor 4, p107/p130-binding (E2F4), were identified as the putative master regulators for ESCC. Additionally, the TFBSs that bind these master regulators were assessed for overlapping single nucleotide polymorphisms (SNPs) that may account for the deregulated expression of the genes under the control of these master regulators. Moreover, SNPs overlapping these TFBSs increase probability that such an altered TFBS and/or the associated TFs may be candidate biomarkers for ESCC.

This research establishes the value of the developed DDEC database to the EC research community and reports on combinations of TFs and estrogen responsive genes that possibly play a role in the underlying mechanisms of gene expression affected by estrogen treatment. Additionally, the method developed for determining the combination of TFs that characterize the hormone responsive genes could be applicable to other hormone responsive cancers. Also, the identification of the ESCC master regulators as candidate diagnostic markers, may enhance our knowledge of the transcriptional circuitry that may positively impact our knowledge of therapeutic targets and side-effects associated with the implemented therapy, and should provide indications as to the most appropriate line of treatment.

April 2009

KEYWORDS: Esophageal cancer

Esophageal squamous cell carcinoma

Database

Dragon Database of Genes Implicated in Esophageal Cancer

Differentially expressed genes

Estrogen

Estrogen Receptor

Transcription regulation

Transcription factor

Transcription factor binding sites

Estrogen response elements

Transcriptional networks

Diagnostic markers

Single nucleotide polymorphisms

DECLARATION

I declare that *"Transcription regulation and candidate diagnostic markers of esophageal cancer"* is my own work, that has not been submitted for any degree or examination in any other university, and that all the resources I have quoted have been indicated and acknowledged by full references.

Magbubah Essack

April 2009

Signed:…….. ……

ACKNOWLEDGEMENTS

I begin in the name of Allah, the Most Compassionate, the Most Merciful

My most fervent appreciation is extended to my Supervisor Prof. Vladimir Bajic for his guidance, intellectual insight and overwhelming support. You have been my pillar of strength, my inspiration and my exemplar; thus, nurturing my personal growth that has made my completion of this thesis possible.

My gratitude is extended to the National Research Foundation for the financial support of this project. I am also indebted to Sebastian Schmeier, Cameron MacPherson, Dr Adam Dawe, Aleksandar Radovanovic, Ulf Schafer, Dr. Mandeep Kaur, Feziwe Mpondo, Kavisha Ramadayal, Mark Wamalwa and Monique Maqungo for their never-ending assistance. My boundless appreciation goes to all fellow SANBItes and the administrative staff; who have carved their memory in my heart with a mere greeting or a smile.

To my mom, siblings and their kids (all thirteen of you) my life and accomplishments would be meaningless without you - love you always.

To my beloved husband, Akef Samih Mohammed Ali Almadadha, thank you for your unwavering support and constant motivation in driving the completion of this thesis.

PUBLICATIONS ARISING FROM THIS THESIS

Kaur, M., Radovanovic, A., Essack, M., Schaefer, U., Maqungo, M., Kibler, T., Schmeier, S., Christoffels, A., Narasimhan, K., Choolani, M., and Bajic, V. B. (2009) **Database for exploration of functional context of genes implicated in ovarian cancer**, *Nucleic Acids Res. 37*, D820-D823.

Essack, M., Radovanovic, A., Schaefer, U., Schmeier, S., Seshadri, S. V., Christoffels, A., Kaur, M., and Bajic, V. B. (2009) **DDEC: Dragon database of genes implicated in esophageal cancer**, *BMC Cancer 9*, 219.

Suzuki, H., Forrest, A. R., van, N. E., Daub, C. O., Balwierz, P. J., Irvine, K. M., Lassmann, T., Ravasi, T., Hasegawa, Y., de Hoon, M. J., Katayama, S., Schroder, K., Carninci, P., Tomaru, Y., Kanamori-Katayama, M., Kubosaki, A., Akalin, A., Ando, Y., Arner, E., Asada, M., Asahara, H., Bailey, T., Bajic, V. B., Bauer, D., Beckhouse, A. G., Bertin, N., Bjorkegren, J., Brombacher, F., Bulger, E., Chalk, A. M., Chiba, J., Cloonan, N., Dawe, A., Dostie, J., Engstrom, P. G., Essack, M., Faulkner, G. J., Fink, J. L., Fredman, D., Fujimori, K., Furuno, M., Gojobori, T., Gough, J., Grimmond, S. M., Gustafsson, M., Hashimoto, M., Hashimoto, T., Hatakeyama, M., Heinzel, S., Hide, W., Hofmann, O., Hornquist, M., Huminiecki, L., Ikeo, K., Imamoto, N., Inoue, S., Inoue, Y., Ishihara, R., Iwayanagi, T., Jacobsen, A., Kaur, M., Kawaji, H., Kerr, M. C., Kimura, R., Kimura, S., Kimura, Y., Kitano, H., Koga, H., Kojima, T., Kondo, S., Konno, T., Krogh, A., Kruger, A., Kumar, A., Lenhard, B., Lennartsson, A., Lindow, M., Lizio, M., Macpherson, C., Maeda, N., Maher, C. A., Maqungo, M., Mar, J.,

Matigian, N. A., Matsuda, H., Mattick, J. S., Meier, S., Miyamoto, S., Miyamoto-Sato, E., Nakabayashi, K., Nakachi, Y., Nakano, M., Nygaard, S., Okayama, T., Okazaki, Y., Okuda-Yabukami, H., Orlando, V., Otomo, J., Pachkov, M., Petrovsky, N., Plessy, C., Quackenbush, J., Radovanovic, A., Rehli, M., Saito, R., Sandelin, A., Schmeier, S., Schonbach, C., Schwartz, A. S., Semple, C. A., Sera, M., Severin, J., Shirahige, K., Simons, C., St, L. G., Suzuki, M., Suzuki, T., Sweet, M. J., Taft, R. J., Takeda, S., Takenaka, Y., Tan, K., Taylor, M. S., Teasdale, R. D., Tegner, J., Teichmann, S., Valen, E., Wahlestedt, C., Waki, K., Waterhouse, A., Wells, C. A., Winther, O., Wu, L., Yamaguchi, K., Yanagawa, H., Yasuda, J., Zavolan, M., Hume, D. A., Arakawa, T., Fukuda, S., Imamura, K., Kai, C., Kaiho, A., Kawashima, T., Kawazu, C., Kitazume, Y., Kojima, M., Miura, H., Murakami, K., Murata, M., Ninomiya, N., Nishiyori, H., Noma, S., Ogawa, C., Sano, T., Simon, C., Tagami, M., Takahashi, Y., Kawai, J., and Hayashizaki, Y. (2009) **The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line**, *Nat. Genet. 41*, 553-562.

Schmeier, S., MacPherson CR., Essack M., Kaur, M., Schaefer, U., Suzuki, H., Hayashizaki Y., and Bajic, V.B. (2009) **Deciphering the transcriptional circuitry of microRNA genes expressed during human monocytic differentiation**, BMC Genomics (accepted).

TABLE OF CONTENTS

CHAPTER 4: Pinpointing Master Regulators of Esophageal Cancer as Candidate Diagnostic Markers-- 59

LIST OF FIGURES AND TABLES

Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Tables

Chapter 2

Chapter 3

Chapter 4

ABBREVIATIONS

ADC	adenocarcinoma
ADH	aldehyde dehydrogenase
ADH1B	alcohol dehydrogenase 1B (class I), beta polypeptide
ALD	alcohol dehydrogenase
ALDH2	aldehyde dehydrogenase 2 family
APAF1	apoptotic protease-activating factor-1
APAF1	apoptotic peptidase activating factor 1
APC	adenomatous polyposis coli
AXIN1	axin 1
Bax	BCL-2-associated X protein
BAX	BCL2-associated X protein
Bcl-2	B cell leukaemia-2
Bcl-2	B cell leukaemia-2 family
BCL2	B-cell CLL/lymphoma 2
BCL2	B-cell CLL/lymphoma 2
BCL2L1	BCL2-like 1
BID	BH3 interacting domain death agonist
BIRC5	baculoviral IAP repeat-containing 5
BMP6	bone morphogenetic protein 6
BRD7	bromodomain containing 7
Caspase	**c**ysteine **asp**artic-specific prote**ases**
CCND1	cyclin D1
CDH1	cadherin 1, type 1, E-cadherin (epithelial)

CDH2	cadherin 2, type 1, N-cadherin (neuronal)
CDK4	cyclin-dependent kinase 4
CDK6	cyclin-dependent kinase 6
CDKN1A	cyclin-dependent kinase inhibitor 1A (p21, Cip1)
CDKN1B	cyclin-dependent kinase inhibitor 1B (p27, KIP1)
CDKN2A	cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)
CREB1	cAMP responsive element binding protein 1
CRYAB	crystallin, alpha B
CTNNB1	catenin (cadherin-associated protein), beta 1, 88kDa
CTSL1	cathepsin L1
CYP1A1	cytochrome P450, family 1, subfamily A, polypeptide 1
CYP2E1	cytochrome P450, family 2, subfamily E, polypeptide 1
DAXX	FAS death domain associated protein
DAXX	death-domain associated protein
DBD	DNA binding domains
DDEC	**D**ragon **D**atabase of Genes Implicated in **E**sophageal **C**ancer
DNA	deoxyribonucleic acid
DNAJA1	DnaJ (Hsp40) homolog, subfamily A, member 1
DNAJB1	DnaJ (Hsp40) homolog, subfamily B, member 1
E2F1	E2F transcription factor 1 (E2F1)
E2F3	E2F transcription factor 3 (E2F3)
E2F4	E2F transcription factor 4, p107/p130-binding
EBV	Epstein-Barr virus
EC	esophageal cancer

EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
EPCAM	epithelial cell adhesion molecule
ER	estrogen receptor
ERBB2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)
ESCC	esophageal squamous cell carcinoma
FADD	Fas (TNFRSF6)-associated via death domain
FAS	Fas (TNF receptor superfamily, member 6)
GSK3B	glycogen synthase kinase 3 beta (GSK3B)
GST	glutathione S-transferase
GSTM1	glutathione S-transferase mu 1
GSTP1	glutathione S-transferase pi 1
GSTT1	glutathione S-transferase theta 1
HDAC	histone deacetylase
HGF	hepatocyte growth factor (hepapoietin A; scatter factor)
HHV4	human herpesvirus 4
HPV	human papilomavirus
HSP90B1	heat shock protein 90kDa beta (Grp94), member 1
HSPA4	heat shock 70kDa protein 4
HSPB1	heat shock 27kDa protein 1
HSPs	heat shock proteins
IAPs	inhibitors of apoptosis
IL6	interleukin 6 (interferon, beta 2)
IL6R	interleukin 6 receptor

IRF7	interferon regulatory factor 7
IRF8	interferon regulatory factor 8
KRT1	keratin 1
KRT10	keratin 10
KRT13	keratin 13
KRT18	keratin 18
KRT19	keratin 19
LEF1	lymphoid enhancer-binding factor 1
LGALS1	lectin, galactoside-binding, soluble, 1
LIF	leukemia inhibitory factor (cholinergic differentiation factor)
LIN28B	lin-28 homolog B
MAP3K5	mitogen-activated protein kinase kinase kinase 5
MDM2	Mdm2 p53 binding protein homolog (mouse) (MDM2)
MTA1	metastasis associated 1
MYC	v-myc myelocytomatosis viral oncogene homolog (avian)
MYST3	MYST histone acetyltransferase (monocytic leukemia) 3
NME1	non-metastatic cells 1, protein (NM23A) expressed in
PAX5	paired box 5 (PAX5)
PDGFA	platelet-derived growth factor alpha
PDGFB	platelet-derived growth factor beta
PDGFRB	platelet-derived growth factor receptor, beta polypeptide
PTEN	phosphatase and tensin homolog
PTTG	pituitary tumor-transforming 1
RB1	Retinoblastoma 1
S100A4	S100 calcium binding protein A4

SMAD4	SMAD family member 4
SNAI1	snail homolog 1 (Drosophila)
SNAI2	snail homolog 2 (Drosophila)
SPIB	Spi-B transcription factor (Spi-1/PU.1 related
SPP1	secreted phosphoprotein 1
STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)
TACSTD2	tumor-associated calcium signal transducer 2
tBid	truncated BH3 Interacting Domain Death Agonist
TCF3	transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)
TCF4	transcription factor 4
TF	transcription factor
TFAP2A	transcription factor AP-2 alpha (activating enhancer binding protein 2 alpha)
TFAP2C	transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma)
TFBS	transcription factor binding site
TFDP1	transcription factor Dp-1 (TFDP1)
TGFA	transforming growth factor alpha
TGFβ1	transforming growth factor, beta 1
TNF	tumor necrosis factor
TP53	tumor protein p53
TWIST1	twist homolog 1 (Drosophila)
VEGF	vascular endothelial growth factor A

VHL	von Hippel-Lindau tumor suppressor
VIM	vimentin
WT1	Wilms tumor 1
ZEB1	zinc finger E-box binding homeobox 1
ZEB2	zinc finger E-box binding homeobox 2
ZIC2	Zic family member 2 (odd-paired homolog, Drosophila)

CHAPTER 1

GENERAL INTRODUCTION AND LITERATURE REVIEW

1.1 GENERAL INTRODUCTION

Reproduction, the source of new unique organisms, starts with the fusing of the ovum with the sperm cell to form a single cell called a zygote. This zygote undergoes a series of mitotic cell divisions. During the life cycle of these cells they adapt to their environment to perform a particular function, thereby developing into tissue, organs and biological systems that comprise the multi-cellular organisms. This life cycle that incorporates decision steps like replication and differentiation solicit the questions: Where does this instruction come from and how is it translated?

The control architecture of the cell is the hereditary information encoded in DNA. It is exerted through the selective production of RNA via a process called transcription. Different tissues and cells express distinct transcriptional profiles. Thus, the life of a cell is embodied in the transcriptional network. The transcriptional process necessitates the binding of specific proteins, called transcriptions factors (TFs), to the promoter sequence to drive gene expression. These TFs bind to short specific DNA sequences called transcription factor binding sites (TFBSs). The combination of TFs that bind the TFBSs and their co-factors determine whether or not gene expression will be activated or inhibited. These transcriptional regulatory processes must be rigidly controlled throughout development to maintain normal cellular processing. Many diseases arise as a consequence of errors occurring in the cellular regulatory system or errors being introduced into the genome as mutations causing cellular behavior to deviate from the norm. Identifying the transcriptional network should provide further insights into partially understood cellular and molecular functioning that characterizes a disease.

1.2 BACKGROUND OF ESOPHAGEAL CANCER

Esophageal cancer (EC) is a malignancy of the esophagus (see Figure 1.1). EC is comprised of heterogeneous groups of tumors that differ in pathogenesis and etiological and pathological features. The most recurrent histological subtype is esophageal squamous cell carcinoma (ESCC), followed by adenocarcinoma (ADC) (*1*). ESCC and ADC rank among the ten most frequent cancers worldwide with incidence of EC and histological subtype varying among regions (*2;3*). Mortality rates associated with EC is very similar to incidence rates due to the relatively late stage of diagnosis and the poor efficacy of treatment [3]. The poor prognosis of EC effect a five year survival rate of 5-20% (*4*).

The poor prognosis of EC patients could be an indirect consequence of our limited knowledge of the transcriptional circuitry related to EC. Enhancing our insights of the functioning of the transcriptional circuitry could potentially positively impact our knowledge of therapeutic targets and side-effects associated with the implemented therapy, and should provide indications as to the most appropriate line of treatment, and additionally allow for the determination of suitable master regulators as putative candidate diagnostic markers for the early stage detection of EC. This review and subsequent analyses focuses on the molecular biology of ESCC.

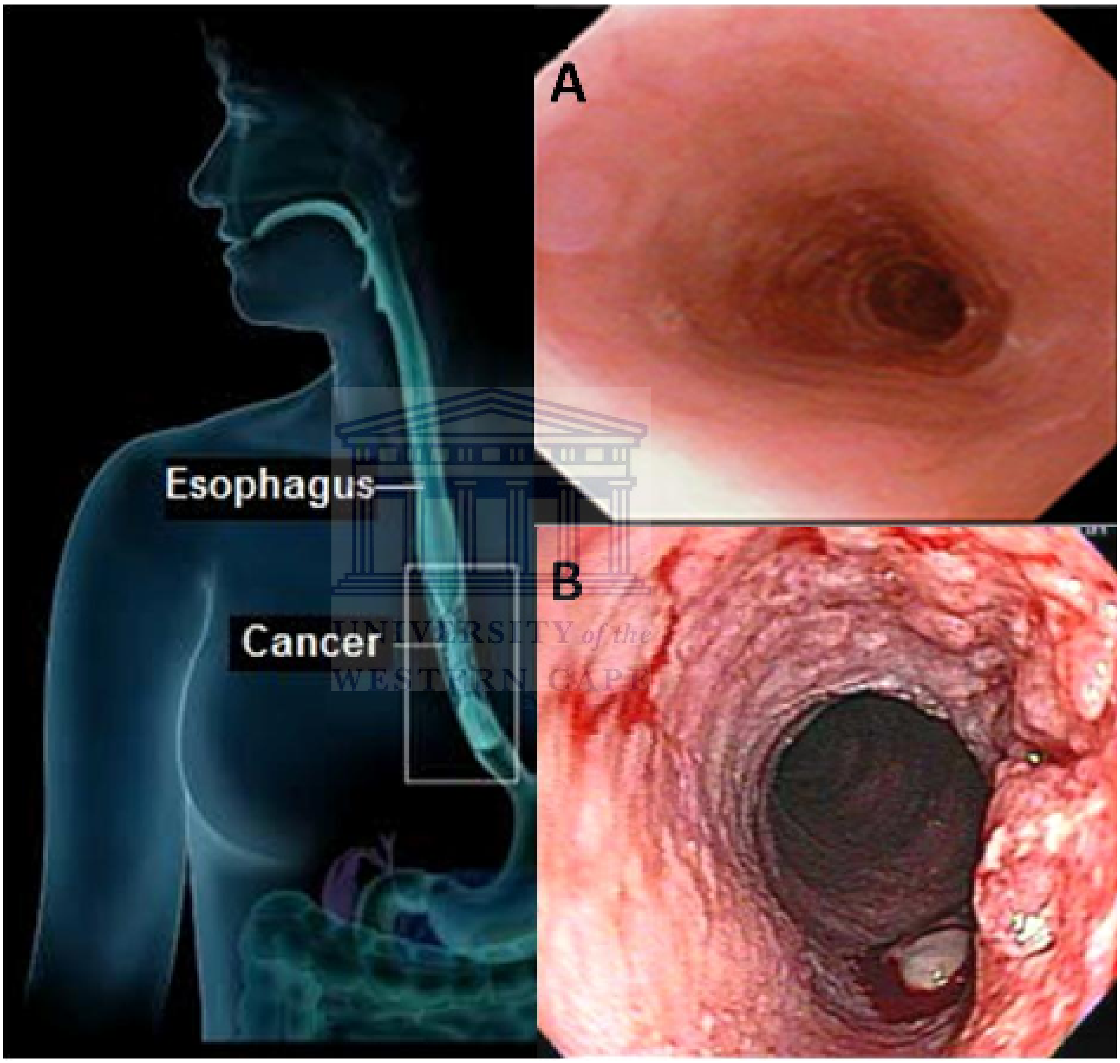


Figure 1.1: Diagram depicting the location of the esophagus and the physical changes associated with ESCC. A) Endoscopic view of proximal esophagus showing normal vascular pattern and distal narrowing (*5*). B) Endoscopic view of ESCC revealing a large tumor.

1.3 ETIOPATHOGENESIS

Epidemiological studies in ESCC have identified tobacco and alcohol as the major risk factors in Europe and North America, but not in the high-risk areas in China and Iran, where dietary habits has been reported to be most relevant. These factors render the esophageal epithelia more susceptible to injury by carcinogens and viruses (*6-9*).

1.3.1 GENES SUSCEPTIBLE TO CHEMICALS

Genetic polymorphisms have been detected in the genes of enzymes involved in the metabolism of carcinogens that have been shown to influence susceptibility to cancer. These include xenobiotic metabolism enzymes in the cytochrome P450 (CYP) family which, may in some cases, be responsible for the bioactivation of carcinogens rather than their detoxification (as in the case with the fungal toxin, Aflatoxin B1 – implicated in EC) , glutathione S-transferases (GSTs) that detoxify carcinogens, alcohol dehydrogenases (ALDs) and aldehyde dehydrogenases (ADHs) that are alcohol metabolizing enzymes.

A study in Japan found that heavy smokers that carry specific polymorphisms of the cytochrome P450, family 1, subfamily A, polypeptide 1 (*CYP1A1*) and glutathione S-transferase mu 1 (*GSTM1*) genes are associated with high risk for ESCC (*10*). Contrastingly, (*11*) and (*12*) have found that polymorphisms of the *CYP1A1*, cytochrome P450, family 2, subfamily E, polypeptide 1 (*CYP2E1*) and *GSTM1* genes are not associated with high risk for ESCC. However, they have found polymorphisms of the alcohol dehydrogenase 1B (class I), beta polypeptide

(*ADH1B*) and aldehyde dehydrogenase 2 family (*ALDH2*) genes are associated with high risk for ESCC.

A similar study in China found that the polymorphism of the *CYPE21* gene is associated with high risk for ESCC but no risk-associated polymorphisms of the *GSTM1*, glutathione S-transferase theta 1 (*GSTT1*) and glutathione S-transferase pi 1 (*GSTP1*) genes were found (*13*). Heavy alcohol consumption and polymorphisms of the *CYP2E1*, *ADH1B* and *ALDH2* genes were also identified as important risk factors for ESCC (*14*).

A recent study in Taiwan found that heavy smokers and alcohol consumers who carry polymorphisms of the *ADH1B* and *ALDH2* genes are associated with high risk for the development of ESCC (*15*).

1.3.2 HUMAN PAPILLOMAVIRUS (HPV) AND EPSTEIN-BARR VIRUS (EBV)

HPV is a papillomavirus that infects the skin and mucous membrane of humans. HPV has been detected in ESCC samples from high risk areas such as Portugal (*16*), Japan (*17*), China (*18*), India (*19*), Pakistan (*20*), South Africa (*21*), Alaska (*22*) and Australia (*23*); but has been detected at low frequency in areas such as France (*24*), Slovenia (*25*), Korea (*26*), North America (*27*), and Netherlands (*28*). A study by Lam concludes that the overall viral prevalence (22%) suggests that other risk factors may play a role and furthermore that HPV may have a synergistic effect with these in the pathogenesis of ESCC (*29*).

EBV, also called Human herpesvirus 4 (HHV4), is a virus of the herpes family (which include the herpes simplex virus, and is one of the most common viruses in humans, initially thought to play a role in ESCC but was later discounted. The EBV was detected at low prevalence in ESCC samples in the United States of America (*30*) and Pakistan (*20*) with an overall prevalence of 7-8%. A study in the high risk area of Japan reports that EBV was only detected in the undifferentiated carcinoma rather than ESCC (*31*). A similar study in Hong Kong detected EBV in lymphocytes in the stroma of ESCC but not in the tumor cells. Thus, EBV detected in the tissue from ESCC may likely have been from the stromal lymphocytes indicating that EBV is not important in the pathogenesis of ESCC (*32*).

1.4 GENES DIFFERENTIALLY EXPRESSED IN EC

1.4.1 ONCOGENES

1.4.1.1 GROWTH FACTORS AND GROWTH FACTOR RECEPTORS

Growth factor proteins are capable of stimulating cellular growth, proliferation and cellular differentiation. Growth factors such as cytokines and hormones typically act as signaling molecules between cells, as they bind to specific receptors on the surface of their target cells.

Other growth factors promote cell differentiation and maturation. There is a wide variety of these growth factors as their presence depends on the type of cell differentiation. For example, bone morphogenic proteins stimulate bone cell differentiation (*33*), while fibroblast growth factors and vascular endothelial growth factors stimulate blood vessel differentiation (*34*).

Numerous growth factors have been demonstrated to be differentially expressed in ESCC including; epidermal growth factor (EGF) (*35*), transforming growth factor alpha(TGFA) (*36*), epidermal growth factor receptor (EGFR) (*37*), platelet-derived growth factor (PDGF) A and B chain (*38*), platelet-derived growth factor receptor, beta polypeptide (PDGFRB) (*39*), transforming growth factor, beta 1 (TGFβ1) (*40*), v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian) (ERBB2) (*41*) and estrogen receptor (ER) (*42*). Yoshida *et al*., demonstrated an increase in accumulation of TGFA, PDGFA, PDGFB and ERBB2 mRNA after treatment with EGF; and an increase in accumulation of EGF, TGFA, EGFR, PDGF A and B chain and the ERBB2 mRNA after treatment with TGFA (*39*). ERBB2 expression is not commonly found in ESCC but has been identified focally in well-differentiated tumour cells, dysplastic epithelium and non-tumour inflamed epithelium. ERBB2 may thus participate at the early stages in the carcinogenesis of ESCC (*43*). Also, EGFR activation is crucial for the regulation of expression of bone morphogenetic protein 6 (BMP6). BMP6 in involved in the early steps of keratinogenesis (*44;45*). Over-expression of EGFR was associated with minimal response to chemoradiotherapy and thus may be considered a prognostic marker in ESCC (*46*). Furthermore, EGF and TGFA exhibits a regulatory role on the multi-growth-factor receptor expression and may thus be key regulators for tumor invasion and metastasis as autocrine modulators for human esophageal carcinoma (*39*).

1.4.1.2 SIGNAL TRANDUCERS

Signal transduction is the process by which signals external to the cell are transduced to the cell interior, generally through the use of membrane-bound receptors and cell signaling pathways (*47*). Proteins involved in this process are known as signal transducers. Some signal transducers reported to be differentially expressed in ESCC are; cyclin-dependent kinase inhibitor 1A (p21, Cip1) (CDKN1A) (*48*), signal transducer and activator of transcription 3 (acute-phase response factor) (STAT3) (*49*), epithelial cell adhesion molecule (EPCAM) (*50*), and tumor-associated calcium signal transducer 2 (TACSTD2) (*51*). *CDKN1A* is mutated in numerous cancer types but no mutation has been detected in ESCC (*52*), even though CDKN1A is over-expressed alongside gene products, secreted phosphoprotein 1 (SPP1) and cathepsin L1 (CTSL1), which it is documented to suppress. Increases in the SPP1 and CTSL1 gene products have been associated with tumor invasion and metastasis in ESCC. Casson and colleagues proved that the increases in these gene products are not a consequence of infiltrating macrophages (*53*).

STAT3 and its target gene products vascular endothelial growth factor A (VEGF) and B-cell CLL/lymphoma 2 (BCL2) were found to be over-expressed in ESCC (*49*) and is activated through phosphorylation in response to various cytokines and growth factors including EGF (*49*), interleukin 6 (interferon, beta 2) (IL6) (*49*), hepatocyte growth factor (hepapoietin A; scatter factor) (HGF) (*54*) and leukemia inhibitory factor (cholinergic differentiation factor) (LIF) (*55*). STAT3 is reported to regulate numerous genes as a transcription factor, but has also been reported to mediate epithelial-mesenchymal transition (EMT) in ovarian cancer cell lines. In this process STAT3 is directly activated by EGFR and interleukin 6 receptor (IL6R), and STAT3 induces cadherin 2, type 1, N-cadherin (neuronal) (CDH2) and vimentin (VIM) expression and IL6 production (*56*). These genes (CDH2, VIM and IL6) have been also demonstrated to be differentially expressed in ESCC.

EPCAM modulates cell adhesion with its over-expression correlating with tumor cell proliferation. This proliferative responses to EpCAM has been demonstrated to require regulated intramembrane proteolysis and a nucleocytoplasmic intracellular domain fragment (*57*).

1.4.1.3 TRANSCRIPTION FACTORS

A transcription factor (TF) is a protein that contains one or more DNA binding domains (DBDs) that binds to specific DNA sequences, thereby either activating or repressing the recruitment of RNA polymerase (the enzyme which performs the transcription of genetic information from DNA to RNA) to specific genes (*58-60*). Additional proteins such as co-activators, chromatin remodelers, histone acetylases, histone deacetylases, kinases and methylases, also play a crucial role in gene regulation, but most lack DNA binding domains, and therefore are not classified as TFs (*61*). Numerous TFs have been documented in the literature. The v-myc myelocytomatosis viral oncogene homolog (avian) (MYC) and transcription factor 4 (TCF4) have specifically been demonstrated to function as TFs in ESCC.

The MYC protein is a multifunctional, nuclear phosphoprotein that plays a role in cell cycle progression, apoptosis and cellular transformation and functions as a transcription factor that regulates transcription of specific target genes such as lin-28 homolog B (*LIN28B)* (*62*), cyclin-dependent kinase inhibitor 1B (p27, Kip1) (*CDKN1B*) (*63*) and bromodomain containing 7 (*BRD7*) (*64*). Pituitary tumor-transforming 1 (PTTG) was reported to be over-expressed in the ESCC cell line, EC9706, and increases *in vitro* cell migration and invasion and promoted *in vivo* lymph node metastasis. PTTG induced S100 calcium binding protein A4 (S100A4) and lectin, galactoside-binding, soluble, 1 (LGALS1) expression and MYC was confirmed to bind the *LGALS1* promoter in this process (*65*). However, over-expression of the *MYC* oncogene gene, as well as gene alteration of the *MYC* locus, appears to be limited, and analysis of the MYC gene yielded limited prognostic value in ESCC (*66*).

TCF4 has been found to participate in EMT. The most common biochemical change associated with EMT is the loss of cadherin 1, type 1, E-cadherin (epithelial) (CDH1) expression. CDH1 is the membrane substrate for catenin (cadherin-associated protein), beta 1, 88kDa (CTNNB1). Thus, *CDH1* transcriptional repressors such as snail homolog 1 (Drosophila) (SNAI1), snail homolog 2 (Drosophila) (SNAI2), zinc finger E-box binding homeobox 1 (ZEB1), zinc finger E-box binding homeobox 2 (ZEB2), transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47) (TCF3) and twist homolog 1 (Drosophila) (TWIST1) have been implicated in promoting EMT in tumor progression (*67;68*). TGFβ1 or TGFβ2 stimulates this signaling mechanism by increasing expression of SNAI1 and SNAI2. SNAI1 and SNAI2 promote formation of CTNNB1–TCF4 transcription complexes that bind to the promoter of the *TGFβ3* gene to increase its transcription. TGFβ3 signaling increases lymphoid enhancer-binding factor 1 (LEF1) expression causing formation of CTNNB1–LEF1 complexes that initiate EMT (*69*). TCF4 in human Pitt-Hopkins syndrome patients was associated with aberrant expression profile and impaired IFN response of the plasmacytoid dendritic cells (PDC). It has been shown to directly activate multiple PDC-enriched genes, including transcription factors involved in PDC development (Spi-B transcription factor (Spi-1/PU.1 related) (SPIB), interferon regulatory factor

8 (IRF8)) and function (interferon regulatory factor 7 (IRF7)) (*70*). TCF4 has also been reported to bind the promoter of the *STAT3* gene expressed in ESCC (*71*).

1.4.2 TUMOUR SUPRESSOR GENES

A tumor suppressor gene inhibits the cell from becoming cancerous. When a mutation in a tumor suppressor gene causes a loss or reduction in its function, the cell can progress to cancer, usually in combination with other genetic changes. Tumor suppressor genes generally follow the ‘two-hit hypothesis’ which implies that both alleles that code for a particular gene must be affected before an effect is manifested i.e. mutant tumor suppressor genes are usually recessive (*72*). The functions of proteins of tumor suppressor genes fall into several categories including the following: 1) Repression of genes essential for the continuing of the cell cycle, 2) Coupling the cell cycle to DNA damage so that the cell cycle can only continue if the DNA repair mechanism has been successful in the event of DNA damage, 3) If the DNA damage cannot be repaired, these tumor suppressor proteins initiate programmed cell death (apoptosis) to prevent aberrant cell growth, and 4) Other proteins involved in cell adhesion prevent tumor cells from dispersing, promote contact inhibition and inhibit metastasis (metastasis suppressors) (*73-75*).

Numerous tumor suppressor genes have been reported to be differentially expressed in ESCC including; adenomatous polyposis coli (*APC*) (*76*), cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4) (*CDKN2A*) (*77*), tumor protein p53 (*TP53*) (*78*), *CDKN1A* (*48*), *CDKN1B* (*79*), SMAD family member 4 (*SMAD4*) (*80*), phosphatase and tensin homolog (*PTEN*) (*81*), Retinoblastoma 1

(*RB1*) (*82*), von Hippel-Lindau tumor suppressor (*VHL*) (*83*) and Wilms tumor 1(*WT1*) (*84*).

The protein transcribed from the APC gene plays a critical role in several cellular processes that determine whether a cell may develop into a tumor by associating with other proteins involved in cell attachment and signaling. As an example, APC regulates CTNNB1 thereby preventing genes that stimulate cell division from being turned on too often and prevents cell overgrowth. The APC protein accomplishes this by complexing with glycogen synthase kinase 3 beta (GSK3B) and axin 1 (AXIN1) and then binds CTNNB1 (*85*). Keratin 1 (KRT1) then phosphorylates CTNNB1 with ensuing phosphorylation by GSK3B. This targets CTNNB1 for ubuiquitination and degradation by cellular proteosomes thereby preventing it from translocating into the nucleus, where it acts as a transcription factor for proliferation genes (*86*). Mutated *APC* has been associated with different cancer types (*87;88*). *APC* markers (D5S210, D5S346, D5S299, and D5S82) were tested in ESCC and showed an overall high frequency of allelic imbalance/loss of heterozygosity (62.48%) and microsatellite instability (41.27%), but did not show prognostic significance in the study cohort and were not correlated with the immunohistochemical data (*89*).

CDKN2A and *CDKN2B* genetic mutations have been detected in ESCC (*90*). CDKN2A and CDKN2B bind to cyclin-dependent kinase 4 (CDK4) and cyclin-dependent kinase 6 (CDK6) respectively to prevent entry into the S-phase of the cell cycle (*91;92*). *CDKN2B* mutations are more often associated with hematopoietic malignancies and gliomas while *CDKN2A* mutations are more often detected in epithelial derived tumors. The *CDKN2A* gene generates various transcripts that functions as inhibitors of CDK4 kinase and as a stabilizer of tumor suppressor proteins. CDKN2A functions as a stabilizer of the TP53 as it binds to this protein and sequester the Mdm2 p53 binding protein homolog (mouse) (MDM2) to degrade it (*93*). TP53 is a cell cycle regulator that functions via anti-cancer mechanisms including; 1) inducing growth arrest by holding the cell cycle at the G_1/S regulatory check point on DNA damage recognition, 2) activating DNA repair proteins, and 3) initiating apoptosis if DNA damage is irreparable. MYST histone acetyltransferase (monocytic leukemia) 3 (MYST3) complexes with TP53 and activates the expression of several genes that includes CDKN1A. The CDKN1A protein forms a complex with CDK2 to stall cell division at the check point. Thus, if a mutation in *TP53* prevents it from expressing the CDKN1A protein, this protein will not be available to act as the 'stop signal' for cell division. Thus the cell will divide uncontrollably, and form a tumor (*94*). Inherited *TP53* mutations have been documented as leading to the onset of Li-Fraumeni syndrome (*95*). Numerous cancers have been associated with a mutation or deletion of the *TP53* gene (*96-98*). *TP53* mutations have been reported in ESCC and frequency of mutations appears to be higher in high incidence areas (*99-102*). The mutations identified are primarily point mutations with G→T transversions that have been associated with chemical carcinogens such as cigarette smoke and aflatoxins (*100*). It has also been reported that patients with *TP53* mutations have a shorter survival (*103*).

RB1 prevents the cell from replicating damaged DNA by preventing its progression along the cell cycle through G1 into S phase (*91*). RB1 binds and inhibits TFs of the E2F family, which are composed of dimers of an E2F protein and a DP protein, as these complexes promote the cell cycle into S phase (*104-106*). The RB1-E2F/DP complex also attracts a histone deacetylase (HDAC) protein to the chromatin, further suppressing DNA synthesis. When it is time for a cell to enter S phase, complexes of RB1 are phosphorylated by cyclin D1 (CCND1)/CDK4/CDK6 and subsequently phosphorylated by CDKN1B/CDK2 thereby allowing E2F-DP to dissociate from RB1. RB1 remains phosphorylated throughout S, G2 and M phases (*91;104*). Mutations in *RB1* were found in 33% - 50% of cases with ESCC and CCND1 over-expression was associated with expression of RB1, and CDKN2A was associated with loss of RB1 (*107-110*).

1.4.3 METASTASIS GENES

Metastasis is the process by which cancer cells leave the original tumor site and migrate through the body via the bloodstream or the lymphatic system. When oral cancers metastasize, they commonly travel through the lymph system to the lymph nodes in the neck. The body resists the metastasis process by an array of mechanisms actioned by metastasis suppressor proteins (*74*).

Metastasis associated 1 (MTA1) protein was identified in a screen for genes expressed in metastatic adenocarcinoma cell lines (*111*). MTA1 plays a role in promoting cancer cell invasion, adhesion and movement and have been found to be up-regulated in numerous cancer types (*112-114*). MTA1 expression increased in the nucleus with the development of esophageal squamous cell carcinoma from normal epithelial cell, dysplasia, to invasive cancer (*115*).

The non-metastatic cells 1, protein (NM23A) expressed in (*NME1*) is a metastasis suppressor gene that is down-regulated in highly metastatic cells (*116*). Mutations in this *NME1* gene have been identified in aggressive neuroblastomas (*117*). Furthermore, loss of NME1 expression was detected in ESCC and corresponds to a reduction in the overall survival rate (*118*).

1.4.4 APOPTOSIS GENES

Apoptosis is a tightly regulated cell death program which requires the interplay of a multitude of genes, as opposed to chaotic, unstructured cell death coined necrosis (*119*). The components of the apoptotic signaling network are genetically encoded in a nucleated cell equipped to be activated by a death-inducing stimulus (*120;121*). In mammalian pathways it is known that the B cell leukaemia-2 (Bcl-2) family of proteins plays a central role in the decision step of apoptosis whereas the **c**ysteine **asp**artic-specific prote**ases** (caspase) family is essential in the execution of cell death (*122;123*). Furthermore, apoptosis is regulated by inhibitors of apoptosis (IAPs) (*124;125*) and heat shock proteins (HSPs) (*126;127*). Numerous apoptosis genes have been reported to be differentially expressed in ESCC including; B-cell CLL/lymphoma 2 (*BCL2)* (*128*), BCL2-like 1 (*BCL2L1)* (*129*), baculoviral IAP repeat-containing 5 (*BIRC5)* (*130*), heat shock protein 90kDa beta (Grp94), member 1 (*HSP90B1)* (*130*), heat shock 70kDa protein 4 (*HSPA4)* (*131*), heat shock 27kDa protein 1 (*HSPB1)* (*132*).

The BCL2 family direct mitochondrial outer membrane permeabilization and can be either pro-apoptotic (BCL2-associated X protein (BAX)) or anti-apoptotic (BCL2, BCL2L1) (*133*). As an example, if DNA damage is irreparable TP53 initiates the expression of the *BAX* gene (*134*). This excess of BAX gives rise to the BAX/BAX complexes (proapoptotic complex) that replace the BCL2/BAX complexes (anti-apoptotic complex) as the two complexes compete for the same receptor site on the mitochondrial membrane (*135*). The attachment of the BAX/BAX complex to the mitochondrial membrane brings forth the release of the apoptogenic factors (*134;136*).

HSPs including HSP90B1, HSPA4 and HSPB1, function as regulators of apoptosis. FAS (TNF receptor superfamily, member 6) (FAS) induced apoptosis typically involves ligation with the adaptor molecule FAS (TNFRSF6)-associated via death domain (FADD) or alternatively death-domain associated protein (DAXX). Activated DAXX leads to activation of the mitogen-activated protein kinase kinase kinase 5 (MAP3K5) to induce apoptosis via the MAPK/JNK pathway (*137;138*). HSPB1 appear to suppress FAS-inducd apoptosis by binding to DAXX (*139*). HSPB1 and HSPA4 suppresses caspase-dependent apoptotic signaling by binding to pro-caspase-3 and pro-caspase-7 and preventing their activation in U-937 cells. HSPA4 was however, unable to directly bind to activated caspase-3 and -7 (*140;141*). Interestingly, the small Hsp crystallin, alpha B (CRYAB) can suppress caspase-8 and cytochrome *c* mediated activation of caspase-3 via a direct interaction with caspase-3 to prevent apoptosis (*142*). HSPB1 is also reported to mediate BH3 interacting domain death agonist (Bid) translocation to the mitochondria in response to TNF-induced apoptosis (*143;144*). In addition, HSPA4 and its co-chaperones DnaJ (Hsp40) homolog, subfamily B, member 1 (DNAJB1) or DnaJ (Hsp40) homolog, subfamily A, member 1 (DNAJA1) inhibit BAX translocation to the mitochondria to prevent nitric-oxide induced apoptosis (*145*). Most apoptotic pathways converge at the point of cytochrome *c* release from the mitochondria. Cytochrome *c* disrupts the binding between BCL2 and apoptotic peptidase activating factor 1 (APAF1). APAF1/caspase-9 complexes from active caspase-9 that together with cytochrome *c* form the apoptosome (*146;147*). HSPA4 also inhibits apoptosome formation by directly associating with APAF1 to prevent activation of pro-caspase-9 (*148;149*). Furthermore, it has been demonstrated that

HSP90B1 and HSPB1 prevent APAF1 oligomerization by directly associating with APAF1 (*150*) and cytochrome *c*, respectively (*151*).

1.3.5 KERATINS

Keratins are a family of filament forming cytoskeletal proteins. Keratinization, an intracellular accumulation of cytokeratin, is a common feature of esophageal squamous-cell carcinoma (*152*). Their expression in ESCC has been demonstrated to be different when compared to normal esophageal epithelia and varies with differentiation. Lam and co-workers found that keratin 10 (KRT10), keratin 13 (KRT13), keratin 18 (KRT18) and keratin 19 (KRT19) expression levels are up-regulated in ESCC (*153*). The KRT19 appears to be a potential diagnostic marker as it correlates with tumor size, stage and responsiveness to treatment (*154*).

1.5 CURRENT TRENDS IN ESOPHAGEAL CANCER RESEARCH

For decades scientists have been investigating single genes, their functions and roles. However, modern technology enables the large scale study of genomics and proteomics, fundamentally changing our ability to study the molecular basis of cells and tissues in health and diseases through giving a new comprehensive view. The PubMed database was queried with keyword expression: "(ESCC OR esophageal squamous cell carcinoma) AND *array" restricted to the last five years on 31/01/2009 and 91 entries were retrieved. The 91 entries include 66 DNA microarray studies but no ChIP-on-chip studies. The lack of comprehensive transcription factor investigations such as ChIP-on-chip and the limited studies focused on TFs that regulate the expression of ESCC, emphasises the importance of the current study.

1.6 STUDY AIMS AND OBJECTIVES

The recent growth of information from new fields of molecular biology including genomics, proteomics, metabolics and high throughput microarray technologies, has given rise to an additional discipline called bioinformatics. Bioinformatics aims to classify, organise, integrate and analyze biological information on a large-scale (*155*). In this research project I will utilize existing bioinformatics tools to achieve the overall objective of my project. The aim of my PhD research project is to uncover the transcriptional regulation and identify putative diagnostic markers associated with EC. The findings of this research may provide the scientific community with directions where research should be focused with respect to therapeutic targets, appropriate line of treatment, and allow for the determination of suitable master regulators as putative candidate diagnostic markers for the early stage detection of EC.

The aim of the proposed study is increase insight into the transcriptional circuitry that characterizes ESCC by: 1) The development of an esophageal cancer database (DDEC) that allows for the easy exploration of EC-related information. 2) Using data housed in DDEC to identify TFBS combinations that characterise estrogen responsive ESCC genes. 3) To identify putative master regulators for ESCC and search for single nucleotide polymorphisms (SNPs) within the associated TFBSs that is presumed to underlie the differences in human traits and disease susceptibility. Presence of a SNP within an over-represented TFBS increases the likelihood of a good candidate biomarker for ESCC.

CHAPTER 2

DATA INTEGRATION FOR THE DEVELOPMENT OF A SPECIALIZED DATABASE ON ESOPHAGEAL CANCER

2.1 ABSTRACT

Efforts made by the scientific community to improve the survival rate of esophageal cancer have resulted in a wealth of scattered information that is difficult to find and not easily amendable to data-mining. To reduce this gap and to complement available cancer-related bioinformatic resources, a comprehensive database (Dragon Database of Genes Implicated in Esophageal Cancer) with EC related information was developed, as an integrated knowledge database aimed at representing a gateway to EC related data.

Manually curated 529 genes which are differentially expressed in EC are contained in the database. The promoter regions of these genes were extracted and analyzed, thereby complementing the gene-related information with transcription factors that potentially control them. Additionally, text-mined and data-mined reports about each of these genes were precompiled to allow for the easy exploration of information about associations of EC-implicated genes with other human genes and proteins, metabolites and enzymes, toxins, chemicals with pharmacological effects, disease concepts and human anatomy. The resulting database, DDEC, has a useful feature to display potential associations that are rarely reported and thus difficult to identify. Moreover, DDEC enables inspection of potentially new ‘association hypotheses’ generated based on the precompiled reports.

DDEC is freely accessible for academic and non-profit users at http://apps.sanbi.ac.za/ddec/. The semi-automated methodology used to populate DDEC genes and related data will be used to update the database twice a year. DDEC has further been recently published in BMC Cancer.

2.2 INTRODUCTION

The scientific research process is hampered by the wealth of scattered research data. Researchers need to sieve through this scattered research data to identify relevant research findings. However, this affects the research process negatively as the compiling of the relevant information is tedious and time consuming (*155*). In an attempt to enhance research endeavors related to EC, the DDEC has been developed as an integrated knowledge database that contains information about various genes differentially expressed in EC.

It should be noted that there are two initiatives aimed at coordinating activities in producing resources related to cancer research, such as the International Cancer Genome Consortium - ICGC (http://www.icgc.org/) and caBIG (cancer Biomedical Informatics Grid^{TM}, http://cabig.cancer.gov/). These two initiatives intend to promote specific data formats and other conditions that will enable easier integration of cancer-related resources. There are cancer-related databases that include information on EC, such as Cancer Gene Expression Database (CGED) (*156*), PDQ (*157*) and Oncomine (*158*). CGED houses a collection of gene expression and clinical data from a large number of patients with major cancers including EC. CGED expression data have been obtained by adaptor-tagged competitive PCR (ATAC-PCR) and allows researchers to explore the correlation between gene expression and clinical data for future diagnostic application (*156*). PDQ is the National Cancer Institute’s (NCI) cancer database that includes peer-reviewed summaries on cancer treatment, screening, prevention, genetics, and complementary and alternative medicine (*157*). The Oncomine initiative collects and analyzes all published cancer microarray data and currently houses EC-related microarray data (*158*).

However, none of the current public databases focuses on genes implicated in EC and their potential associations with other relevant biological, biochemical and medical entities. Moreover, DDEC provides a combination of features for exploration of information related to EC-implicated genes that cannot be found elsewhere, such as filtering for putative transcription factors shared amongst promoters of EC-implicated genes, inference of association networks and precompiled reports that provide insights into other human genes and proteins, metabolites and enzymes, toxins, chemicals with pharmacological effects, disease concepts and human anatomy associated with differentially expressed EC-implicated genes. It also enables finding rare information that will be likely missed in the common literature search. As a special feature, DDEC provides a module for generation of ‘association hypotheses’ between concepts related to EC-implicated genes. Batch queries and a database dump are also provided. DDEC represents a useful complement to the existing databases and will contribute to more efficient EC-related research. DDEC is freely accessible for academic and non-profit users at http://apps.sanbi.ac.za/ddec/. The semi-automated methodology used to populate DDEC genes and related data will be used to update the database twice a year.

2.3 CONSTRUCTION AND CONTENT

The DDEC is based on a three-tier (layer) (data, logic and presentation) architecture (see Figure 2.1). The presentation layer is web-based and implemented in DHTML and Javascript. The logic layer has been implemented as a number of server side PHP and Perl modules interfaced with the data layer. The data layer is MySQL, and for the text-mining purposes, file system based. The relational database design strictly distinguishes between tables that contain data entities and tables that establish logical connections between these data entities. Central data entity is the gene to which most other data entities are linked. Other important data entities are transcription-related such as transcription start sites (TSSs) and transcription factors (TFs). This is reflected in the entry points that a user can chose between on the top level of the web-interface. Information in the DDEC is structured into four distinct parts:

1. A platform that can be used to search the integrated gene information through standardized vocabularies.
2. Selection of genes of interest from the list. Search criteria provide users with gene details such as: general information, gene in other resources, experimental evidence, related proteins, associated pathways, associated diseases, orthologous genes, regulations and text-mined reports that can support building interactive association networks.
3. Transcription regulation information which includes all putative TFBSs for the EC-implicated genes in DDEC. This segment is useful for gene regulation studies since TFBSs of interest can be selected and the results will list each

TFBS and gene promoter with corresponding TFBSs. Genes sharing all the selected TFBSs are listed as well.

4. Batch queries and data download interface is provided to increase utility for users.

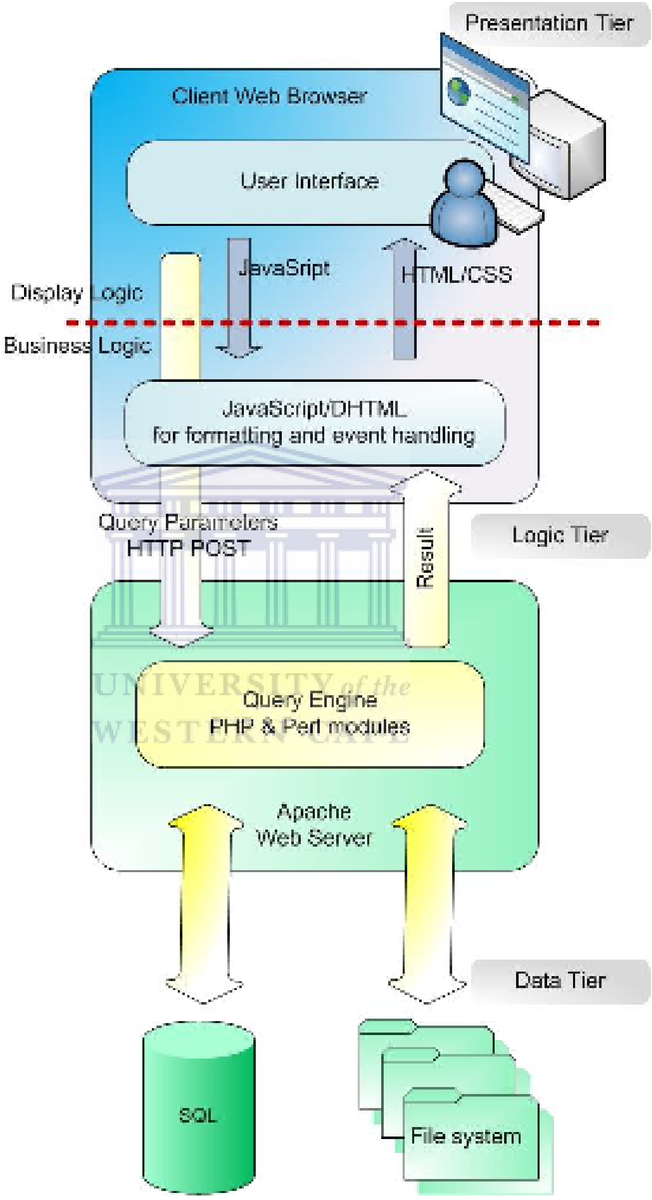


Figure 2.1: Schematic representation of the DDEC structure.

DDEC contains information on EC-implicated genes compiled based on scientific publications from PubMed. The PubMed database was queried with keyword expression: “esophageal (cancer OR cancers OR tumor OR tumors OR carcino*

OR adenocarc* OR malign* OR neoplasm*)" on 31/01/2008 and 35,892 PubMed abstracts were retrieved. The search for relevant publications was further refined using the licensed Dragon Exploration System (DES) from OrionCell (http://www.orioncell.org), that has an integrated Biomedical Text-Miner tool. DES retrieved a list of 1677 putative genes associated with EC from the extracted abstracts. Biologists then evaluated information about experimental conditions these genes have been subjected to using full-text articles whenever possible, and abstracts in other cases. When the available information was insufficient to deduce the correct experimental conditions, the gene was discarded. Taking into account that experimental conditions influence gene expression, DDEC provides details of the cell line, tissue or cell type, expression status, disease stage, tumor grade, esophageal cancer type and laboratory method reported in literature.

A final list of 529 genes was identified in this way and used to populate the database. The general information about the genes, which includes HGNC ID, approved symbol, approved name, entrez ID, previous symbol, previous name, aliases, OMIM-related information, and chromosome location, was extracted from sources such as HUGO (*159*) (http://www.genenames.org/) and GeneCards (*160*) (http://www.genecards.org/index.shtml). Included in the database are gene related identifiers such as EMBL (*161*) (http://www.ebi.ac.uk/embl/), Ensembl (*162*) (http://www.ensembl.org/index.html), Refseq (*163*), Genbank (*164*) (http://www.ncbi.nlm.nih.gov/), Unigene (*165*) http://www.ncbi.nlm.nih.gov/sites/entrez?db=unigene&orig_db), Uniprot (*166*) http://www.ebi.ac.uk/uniprot/), Swiss-Prot (*167*) (http://www.expasy.ch/sprot/) and PDB (*168*) (http://www.rcsb.org/pdb/home/home.do). ID conversion tools like

IDconvertor (*169*) (http://idconverter.bioinfo.cnio.es/) and Onto-tools (*170*) (http://vortex.cs.wayne.edu/ontoexpress/servlet/UserInfo) were used to convert between different types of identifiers. A summary of the statistics of the above mentioned features are listed in documentation. Links to the relevant sources of data such as gene ontologies (*171*) (http://www.geneontology.org/), Evoc (*172*) (http://www.evocontology.org/), and Reactome pathway data (*173*) (http://www.reactome.org/) are also provided.

As a useful feature, lists of putative TFBSs that map to the promoter regions of EC-implicated genes were generated to allow users to identify genes that share common TFBSs. For this purpose, promoter sequences were extracted using mainly Fantom3 CAGE tag data (*174*), as well as Toucan v. 3.0.2 (*175*). The TRANSFAC Professional database v.11.4 was used to map the TFBSs to the promoter sequences(*176*). All TRANSFAC mammalian matrix models of binding sites were mapped using the Match^{TM} program with *minFP* profiles for optimized thresholds of the matrix models (*177*). The complete list of 529 genes was used to extract promoter sequences for the identification of putative TFBSs. Promoter sequences of 409 genes (1200 bp upstream and 200 bp downstream from the transcription start site, TSS) were extracted from the Fantom3 CAGE tag data corresponding to 1582 transcription start sites (TSSs) where each had at least five tags in the tag cluster and a minimum of three tags in the representative tag (174). An additional 108 promoter sequences (1200 bp upstream and 200 bp downstream from the TSS) were extracted using Toucan v. 3.0.2 (*175*).

As an additional feature, all related PubMed documents for each of the 529 EC-implicated genes were extracted and analyzed them using DES. DES uses a dictionary-based text mining approach to extract information used for the precompiled reports by mapping the entities from the dictionaries to the submitted PubMed documents. Six manually curated DES dictionaries namely; human genes and proteins, metabolites and enzymes, toxins, chemicals with pharmacological effects, disease concepts and human anatomy, were applied. These dictionaries were compiled from literature and public databases. The accuracy of this integrated data has been evaluated in Sagar *et al.* in terms of precision, recall and F-measure. The analysis of the results displayed precision and recall ranging from 81%-100% and with an average F-measure of 92.9% for the *SCN1A* gene (*178*). The precompiled reports in this study are incorporated in the DDEC and provide the user with a possibility to inspect possible interactions associated with the genes of interest and associated networks of relevant biomedical entities. An additional feature in DES allows for hypotheses to be generated between two dictionary entries that are linked to a common dictionary entry. This tool enables the user to test the hypotheses generated by retrieving PubMed documents related to the two dictionary terms linked through the hypothesis. If no PubMed documents are retrieved the hypothesis may warrant further exploration. This functioning of the text-mining modules of DDEC is based on similar concepts as used in Pan *et al.* and Bajic *et al.* (*179*),(*180*). DES has also been employed in the creation of a module for the ovarian cancer database, DDOC (*181*). Batch queries and data download are provided to increase utility for users. Further, a database dump has been provided to support integration with other database resources.

The above outlined process of biocurated data collection and integration will be repeated twice yearly as an update process. Updates will incorporate extracting abstracts from the last update day to current day. This semi-automated process is more time consuming than current automated update systems but has the advantage of reducing redundant information.

The existence of this database is owing to me, Mandeep Kaur and Vladimir B. Bajic whom conceptualized the study. I further contributed the content in this database and a team comprising of Aleksandar Radovanovic, Ulf Schaefer, Sebastian Schmeier, Sundararajan V. Seshadr and Alan Christoffels developed the database. Alexsandar Radovanovic and Vladimir B. Bajic also developed the DES system

2.4 UTILITY AND DISCUSSION

DDEC provides a comprehensive compilation of information obtained from published EC research, complemented with the information from public databases and information derived from computational analysis. The information captured in DDEC is centered on genes differentially expressed in EC. The information used for selection of genes to be included in DDEC was curated. Only genes that satisfy all conditions listed below are included in DDEC:

(i) Genes that are differentially expressed in human EC with experimental proof.

(ii) Differential expression of EC-implicated genes has not been influenced by anti-cancer therapy.

(iii) Differentially expressed EC-implicated genes have not been artificially constructed.

Microarray data has been excluded at this stage as the results obtained using high throughput technologies are debatable in terms of deciding about a meaningful level of gene expression and statistical methods used for analysis and interpretation of data (*182;183*). However, as a future prospect the database will be expanded by adding a subset for raw expression data and analysis of the EC-related microarray data.

DDEC contains precompiled text-mined and data-mined reports that allow for easy exploration of information about associations of EC-implicated genes with other genes and proteins, metabolites and enzymes, toxins, chemicals with pharmacological effects, disease concepts, human anatomy, pathways and pathway reactions. Moreover, DDEC provides for potentially new 'association hypotheses' generated in the precompiled reports. It also provides frequency of associations that allows users to observe rare associations with the genes of interest that will usually be overlooked in a typical literature search taking into account the huge volume of data available. DDEC can be used to answer questions such as:

1. Is my gene of interest differentially expressed in EC, i.e. is it an EC-implicated gene as defined here?
2. Which putative transcription factors regulate the expression of an EC-implicated gene or sets of these genes?
3. Which of the other EC-implicated genes in DDEC are regulated by the same transcription factor (or factors) as the gene of interest?

4. My gene of interest has putative associations with other biomedical concepts. What are these concepts and what are the documents from which such associations are deduced so that I can explore them?

The potential uses and advantages of the database are described in the documentation section (http://apps.sanbi.ac.za/ddec/DDEC.pdf). An example of data analysis has been included in the documentation and should help users to understand and utilize different functions implemented in this database to maximize information exploration and extraction.

Kaur *et al*. recently published DDOC, an ovarian cancer (OC) database housing 379 OC-related genes using the same database model and query interface (*181*). To explore whether the EC and OC database content characterize functionally distinct groups of genes, the categories where probed for statistical over-representation of GO terms. EC and OC gene lists were compared for this analysis (see Table 2.1). Only 123 genes were found to be common to both cancer types while 406 genes were unique to EC and 256 genes were unique to OC. Generally, all categories were characterized by the majority of genes forming part of the broad terms, apoptosis and cell cycle. However, these categories were primarily over-represented for the genes common to both EC and OC. The gene list unique to EC was found to be enriched in functionally distinct groups such as ‘neuron differentiation and development’ and ‘epidermis development’ while the gene list unique to OC was found to be enriched in functionally distinct groups such as ‘sex differentiation and development’ and ‘embryonic development’.

KEGG pathways enriched for the genes unique to EC, genes unique to OC and the genes common to EC and OC were identified (*184*). The MAPK signaling pathway, ErbB signaling pathway and p53 signaling pathway were found to be the most pronounced pathways for genes common to EC and OC. The pathways most pronounced for the genes unique to EC were the MAPK signaling pathway, Wnt signaling pathway, with androgen and estrogen metabolism being unique to this group. The MAPK signaling pathway, ErbB signaling pathway and TGF-beta signaling pathways were most pronounced for the genes unique to OC.

Table 2.1: A comparison of Esophageal and Ovarian Cancer genes by characterizing functionally distinct groups based on Gene Ontology terms

Gene Ontology terms representing functionally distinct groups	Genes unique for Esophageal Cancer (EC)	Gene unique for Ovarian Cancer (OC)	Genes common to EC and OC
Neuron differentiation and development	3.03	0.77	0
Epidermis development	5.91	1.06	1.66
Sex differentiation and development	0.36	1.36	0
Embryonic development	0.76	1.85	0
Regulation of apoptosis	8.09	9.84	13.22
Regulation of cell cycle	11.99	11.31	14.73

* The values tabulated in Table 2.1 represent the overall enrichment score for the group based on EASE scores of each term members. The higher the score the more enriched (*185*).

The above analysis suggests that distinct categories of genes participating in specific pathways are involved in pathogenesis of different types of cancers. These cancer specific categories of genes can be investigated as potential biomarkers for prognosis and diagnosis of the disease.

In future, the effect of current therapeutic drugs will be incorporated. Additional features that may enhance search and retrieval of DDEC information will be added in due course, as well as incorporation of DDEC into ICGC, caBIG and LinkOut. DDEC will further be updated twice a year and will continue to grow in both content and functionality.

2.5 CONCLUSION

DDEC is an integrated knowledge database aimed at representing a gateway to EC-related data. DDEC houses information associated with 529 hand-curated human genes implicated in EC and allows the users to easily access the wealth of EC related data that is typically difficult to find and not easily amendable to data mining. Users are also provided with the DES interface that allows for the easy exploration of information, viewing of potential associations that are rarely reported and thus difficult to identify and inspection of potentially new 'association hypotheses' generated based on the precompiled reports.

DDEC is freely accessible for academic and non-profit users at http://apps.sanbi.ac.za/ddec/ and has recently been published in BMC Cancer (*186*). This resource will hopefully serve as a useful complement to the existing public resources and as a good starting point for researchers and physicians interested in EC genetics.

CHAPTER 3

CHARACTERIZING THE TRANSCRIPIONAL CIRCUITRY OF ESTROGEN RESPONSIVE ESOPHAGEAL CANCER GENES

3.1 ABSTRACT

Prognosis for EC and various other malignancies remains poor despite the use of multidisciplinary therapy. This may be a consequence of our limited knowledge of underlying mechanisms affected by these lines of treatment. This study was initiated in an attempt to expand on the current knowledge of these underlying mechanisms in estrogen-based EC therapy. This will be accomplished by *in silico* analysis of estrogen response elements (ERE) and TFs in the promoter regions of ESCC esophageal cancer related genes extracted from the integrated EC database, DDEC.

Estrogen is known to activate estrogen receptors (ERs) that in some cases bind to specific DNA motifs called EREs in target gene promoters' thereby altering gene expression with the involvement of specific combinations of TFs. In this study, estrogen response genes and ESCC genes that have been verified with *in vitro* biological expression data were used to computationally interpret the transcriptional circuitry of estrogen-related hormone therapy in ESCC. The applied research methodology allowed for the identification of 47 novel putative estrogen responsive ESCC genes. Recent 2009 publications by (*187*) and (*188*) verified two of the predicted estrogen responsive ESCC genes.

These findings provide the research community with combinations of TFBSs that may be used to identify estrogen responsive genes and to divulge more genes that possibly play a role in the underlying mechanisms affected by estrogen treatment.

3.2 INTRODUCTION

Bearing in mind that the transcriptional circuitry is a coordination of three constituents: TFs, the sequence-specific DNA binding motifs (TFBSs) that bind the TFs in the promoter region of the affected transcripts, and the affected transcripts, that are responsible for the regulation of gene expression (*189*). The TFBSs that mediate this gene expression generally occur within 2 kb upstream of the transcription start site (*190*); although some binding sites have been identified in other regions (*191;192*). The literature further indicates that TFBSs have 80-fold higher functionality if located in the proximal promoter region (*193*) and that TFs have the tendency to compete for overlapping TFBSs thereby inducing or inhibiting gene expression (*194-196*).

Published scientific data demonstrates that estrogen has an inhibitory effect on esophageal carcinoma, which is regulated by the functioning of the TF, estrogen receptor (ER) (*197-199*). ER, binds to a specific TFBS known as the estrogen response element (ERE) (*200;201*). There are two ER subtypes, ERα and ERβ, that are located on human chromosomes 6q25 (*202*) and chromosome 14q22-24, (*203*), respectively. Both ERα and ERβ bind to the same estrogen responsive elements (EREs) but ERα does so with an approximately twofold higher affinity (*204*). Additionally, ERβ is known to bind to ERα thereby suppressing ERα function

(*205;206*). This inverse biological function associated with the two ER subtypes has been confirmed to exist in ESCC (*197*).

Estrogens are implicated as regulators of physiological processes such as development, growth and maintenance of reproductive function (*207*). The canonical model for ER-mediated regulation of gene expression involves cytosolic estrogens traversing through the plasma membrane to bind to and activate the nuclear ER thereby causing the ER to dissociate from the chaperone complex (*208*). Upon dissociation, ER's are free to dimerize and directly bind to the EREs thereby controlling the expression of select genes (*199*; 200). ERs are also activated by hormone-independent pathways but still bind directly to EREs (*209-212*). Alternatively, ERs are involved in non-genomic cell membrane signalling pathways that result in fast tissue responses without involving gene expression (*213*) and ERs can indirectly regulate gene expression by forming a complex with specific transcription factors, thereby acting as co-activators or co-repressors (*214*). Genes regulated by ERα include progesterone receptor, cyclin D1, cathepsin D and c-Myc (*215-218*). Transcriptional activities of ERs in ESCC have been studied in brief and consequently they are not well understood within this disease context.

3.3 METHODOLOGY

3.3.1 OVERVIEW

This study attempts to increase insight into the transcriptional circuitry associated with ERs that bind directly to the EREs in ESCC genes. A list of 418 genes, differentially expressed in ESCC, was extracted from the Dragon Database of

Genes Implicated in Esophageal Cancer (DDEC). The promoter sequences of these genes were extracted from the Fantom3 CAGE tag data (*174*). These core promoter regions were searched for EREs using the Dragon ERE Finder version 6.0 (http://apps.sanbi.ac.za/ere/index.php) (*219*), hereby classifying the 418 genes into a gene set with predicted EREs (A) and a gene set lacking predicted EREs (B). Additionally, the ESCC genes in A and B with biological expression data that verify the gene to be estrogen responsive were identified using the list of estrogen response genes taken from KBERG (*220*) and ERtargetDB (*221*). Thus, the 418 differentially expressed ESCC genes were classified into the following 4 categories:

C1: predicted estrogen responsive ESCC genes with experimental proof verifying estrogen responsiveness.

C2: predicted estrogen responsive ESCC genes lacking experimental proof verifying estrogen responsiveness.

C3: ESCC genes with experimental proof verifying estrogen responsiveness that were not predicted to be estrogen responsive.

C4: ESCC genes lacking experimental proof verifying estrogen responsiveness that were not predicted to be estrogen responsive.

To explore whether the categories set apart functionally distinct groups of genes, categories were analysed for statistical over-representation of GO terms. The categories were further analysed for TFBSs that characterise the promoter region of the estrogen responsive genes. The hypothesis was that since EREs act as cis-regulatory elements within gene promoters, significant enrichment of specific sequence motifs in ESCC genes known to be estrogen responsive should be detectable. Consiquently, the TFBSs in the promoter region of all ESCC genes were predicted using position weight matrices (PWMs) from TRANSFAC Professional database v.11.4. Categories C1 and C3 were combined (C1&3) to represent a set of estrogen responsive genes with experimental proof verifying estrogen responsiveness. In this search for overrepresented TFBSs, the C1&3 TFBS mappings were used for comparison against the background TFBS mappings from C4. The one-sided Fisher's exact test was used to calculate the Bonferroni-corrected p-values for each of the TFBSs predicted. Similarly, the p-values for combinations of paired TFBSs within estrogen responsive gene promoters were calculated - believing that pairs of TFBSs represent a basic form of coordinated regulation by multiple TFs. Being particularly interested in the coordinated regulation of estrogen responsive genes, the search was extended to include combinations of TFBSs containing more than two elements. This combinatorial problem quickly became intractable. The method below was designed to reduce this combinatorial complexity.

A ranked list of TFBSs was computed by summing the p-values for each unique TFBS mapping as well as its paired mappings. The summed p-values represent a score intended to describe the disproportion of mappings for a TFBS in the target set against the background set, relative to other TFBSs. The top 10 ranked TFBSs with the lowest sum of p-values were used to calculate new combinations of TFBSs made up of at least 2 and at most 10 TFBSs. Within these criteria, all possible combinations of TFBSs were determined. Subsequently, the null hypothesis was tested, that each combination is not present in proportionally higher frequencies within the target versus background promoter sets. This step allowed for the analysis of multiple combinations of TFBSs potentially contributing to the regulation of ER genes, whilst significantly reducing computational requirements. The significant combinations comprised of these 10 TFBSs were then used to annotate the promoter regions of the putative estrogen response genes in C2 to support the novel putative estrogen responsive genes.

3.3.2 EXTRACTING OF THE PROMOTER SEQUENCES OF THE ESCC GENES

A list of 418 genes, differentially expressed in ESCC, was extracted from the Dragon Database of Genes Implicated in Esophageal Cancer (DDEC). Promoter sequences of 418 genes (1200 bp upstream and 200 bp downstream from the transcription start site, TSS) were extracted from the Fantom3 CAGE tag data that correspond to 1582 transcription start sites (TSSs) that each has at least five tags in the tag cluster and a minimum of three tags in the representative tag (*174*). As a background set, 39,000 randomly selected sequences of length 1200 bases were used from the human genome.

3.3.3 CATEGORIZING ESCC GENES BASED ON ERE PREDICTIONS AND EXPERIMENTAL EVIDENCE OF ESTROGEN RESPONSIVENESS

Dragon ERE Finder version 6.0 (http://apps.sanbi.ac.za/ere/index.php) was used to predict the EREs in the promoter sequences. A sensitivity of 0.83 was used as recommended in (*222*). The ESCC genes in the predicted estrogen responsive gene set and the ESCC gene set lacking predicted estrogen responsiveness were further categorized based on which genes have biological expression data verifying the gene to be estrogen responsive. The list of estrogen response genes with experimental proof was taken from KBERG (*223*) and ERtargetDB (*224*). The current (March 2009) version of ERTargetDB (http://bioinformatics.med.ohio-state.edu/ERTargetDB/) was used. It contained:- (a) 40 genes with 48 experimentally verified ERE direct binding sites and 11 experimentally verified ERE tethering sites; (b) 42 genes identified via ChIP-on-chip assay for estrogen binding (c) 355 genes from gene expression microarrays; (d) and 2659 computationally predicted genes. An additional 1516 experimentally confirmed estrogen-responsive genes extracted from the KBERG database (http://research.i2r.a-star.edu.sg/kberg/) were used as well, to ensure the final list used was comprehensive.

3.3.4 FUNCTIONAL ANALYSIS OF ESCC GENES BASED ON ABOVE CATEGORIZATION

To explore whether the above mentioned categories based on ERE predictions and experimental evidence characterize functionally distinct groups of genes, the categories where probed for statistical over-representation of GO terms (*171*). GO term analysis was performed for all the genes using DAVID (The **D**atabase for **A**nnotation, **V**isualization and **I**ntegrated **D**iscovery) version 2.0 (*225*) and transcripts were clustered based on their functional annotations at GO level 4.

3.3.5 MAPPING TFBS MATRICES TO THE PROMOTER SEQUENCE OF ALL EC GENES

The TRANSFAC Professional database v.11.4 was used to map the TFBSs matrices to the ESCC genes (*176*). The 522 TRANSFAC mammalian matrix profiles of TFBSs were mapped to the promoter sequences using the Match^{TM} program with *minFP* profiles for optimized thresholds of the matrix models (*177*).

3.3.6 DETERMINING THE COMBINATIONS OF TFBS SIGNIFICANTLY OVER-REPRESENTED IN C1&3 AS OPPOSED TO C4

To determine the TFBSs specific to the target set of ESCC promoters, the overrepresentation index (ORI) was calculated using a background set of 40,101 randomly chosen unique human promoters. This calculation was done by in-house computer scripts. Readers interested in applying the same filtering process should contact Vladimir Bajic (vlad@orioncell.org). An ORI value of one means no overrepresentation of the TFBS in the target promoter group compared to the background set. A greater ORI signifies over representation of the TFBS in the target promoter group. All TFBSs with an ORI below two were filtered out. The remaining TFBS were used to annotate the promoters.

In the search for overrepresented TFBSs, the C1&3 TFBS mappings were used for comparison against the background TFBS mappings from C4. The one-sided Fisher's exact test was used to calculate p-values for each of the TFBSs predicted and for all combinations of paired TFBSs. Bonferroni corrections were applied to all p-values. A correction factor of 522 and $(522^2-522)/2$ was applied to p-values for single TFBSs and paired TFBSs, respectively. Over-representation of paired TFBSs within estrogen responsive gene promoters were considered on the basis that pairs of TFBSs represent a basic form of coordinated regulation by multiple TFs.

The search was further extended to include combinations of TFBSs containing more than multiple TFBSs. This combinatorial problem quickly became intractable. The method below was designed to reduce this combinatorial complexity.

All TFBSs were ranked and scored by summing the p-values for each single TFBS mapping as well as its involved paired mappings. The summed p-values represent a score intended to describe the disproportion of mappings for a TFBS in the target set against the background set, relative to other TFBSs. The top 10 ranked TFBSs (i.e. those with the lowest sum of p-values) were selected. This selection was used to calculate new combinations of TFBSs made up of at least 2 and at most 10 TFBSs. Within these criteria, all possible number of TFBS combinations were determined. Subsequently, we tested the Null Hypothesis that each combination is not present in proportionally higher frequencies within the target versus background promoter sets. The one-sided Fisher's exact test was used to calculate the new p-values for all combinations of TFBSs. Bonferroni corrections were applied to all p-values. A correction factor of $(522^x-522)/x$ was applied to p-values for multiple TFBSs. The ranking step significantly reduced the complexity of the problem, whilst retaining those TFBSs most highly overrepresented, and therefore potentially likely to regulate transcription of ER genes. Once the Bonferroni corrected p-values for all combinations of TFBSss were obtained, only combinations that were significant was selected (p-value <= 0.05).

3.3.7 USING THE SIGNIFICANTLY OVER-REPRESENTED COMBINATIONS OF TFBSS TO INCREASE CONFIDENCE IN THE NOVEL PUTATIVE ESTROGEN RESPONSE GENES IN C2

Using the significant selection of top ranked combinations of TFBS (see above), the promoter regions of the predicted estrogen response genes in C2 were annotated for the presence of these combinations of TFBSs to support the novel putative estrogen response genes. These annotations were made viewable in the form of a heatmap using Cytoscape software (*226*).

3.4 RESULTS AND DISCUSSION

3.4.1 THE PREDICTION AND IDENTIFICATION OF ESTROGEN RESPONSIVE GENES DIFFERENTIALLY EXPRESSED IN ESCC

A sequential two-step process was used to predict and verify the ESCC genes that are estrogen responsive. This two-step process is: (a) the prediction of EREs in the promoter region of ESCC genes, and (b) the categorization of the genes in (a) based on published biological expression data that confirm which genes are estrogen responsive.

EREs in the promoter region of a gene suggests that the gene expression has a potential to be induced by activated ER occupying the ERE site, along with the necessary complement of transcriptional machinery primed. The 418 differentially expressed ESCC genes were extracted from the DDEC (http://apps.sanbi.ac.za/DDEC/index.php). The TSSs used to identify the core promoter sequences of the 418 ESCC genes (1200 bp upstream and 200 bp downstream from the TSS) were extracted from the Fantom3 CAGE tag data that correspond to 1645 TSSs. Thus the 1645 putative promoters were analysed for ERE via Dragon ERE Finder version 6.0 (http://apps.sanbi.ac.za/ere/index.php). This tool's detection algorithm was tested on several large datasets and makes one prediction in 13,300 nt while achieving a sensitivity of 83% (*227*). In the dataset, of the 1645 promoters, EREs were detected in 242 promoter regions corresponding to 128 ESCC genes (see appendix I).

The curated estrogen responsive gene lists stored in the KBERG and ERTargetDB were used to confirm which genes have published biological expression data demonstrating estrogen responsiveness (see appendix II). Of the 128 putative estrogen responsive genes, 56 genes had biological expression data confirming estrogen responsiveness. Thus, the additional 72 genes predicted can lay the foundation for increasing insights into the molecular events triggered by estrogen. In the dataset there are 291 genes that had no EREs predicted in their promoter regions. Of the 290 genes, 146 had biological expression data confirming estrogen responsiveness. The promoter regions of the 146 gene that did not contain an over-represented ERE motifs but have been documented to be estrogen responsive may exemplify ERs acting in a co-activator or co-repressor capacity (*228*). This analysis generated four categories of genes (Table 3.1) that are used in subsequent analyses.

Table 3.1: ESCC genes categorized based on ERE predictions and experimental evidence of estrogen responsiveness.

Category	Category Description	No. of genes
C1	genes with predicted EREs and have experimental evidence of estrogen responsiveness	56
C2	genes with predicted EREs and lack experimental evidence	72
C3	genes without ERE predictions but have experimental evidence of estrogen responsiveness	146
C4	genes without ERE predictions and also do not have experimental evidence of estrogen responsiveness	144

3.4.2 FUNCTIONAL ANALYSIS OF ESCC GENES BASED ON ERE CATEGORIZATION

To explore whether the gene categories, C1 to C4, characterize functionally distinct groups of genes, the categories where probed for statistical over-representation of GO terms (*171*).

For these analyses, categories C1 and C3 were combined (C1&3) as they represent a set of experimentally validated estrogen response genes. C1&3, C2 and C4 constitute 47.97%, 17.18% and 34.84% of genes under study, respectively. Generally, all categories were annotated by the predominant gene annotation from one of the broad terms, apoptosis and cell differentiation. These broad terms are defined, at least in part, by underlying processes (see Figure 3.1) such as cellular protein metabolic process, phosphate metabolic process, phosphotransferase activity, kinase activity, adenyl nucleotide binding, biopolymer modification, and purine ribonucleotide binding. Figure 3.1 illustrates the over-represented processes. The nucleus defined processes such as adenyl nucleotide binding, biopolymer modification and purine ribonucleotide binding are dominant in C1&3 and C2 compared to C4.

3.4.3 TFBS ANALYSIS OF ESTROGEN RESPONSIVE ESCC GENES

The analysis of the TFBSs in the promoter regions of the ESCC genes required the sequential execution of three processes: (a) mapping the TFBSs matrices to the promoter sequences of all ESCC genes, (b) determining the combinations of TFBSs significantly over-represented in C1&3 as opposed to C4 and (c) using the significantly over-represented combinations of TFBSs to increase confidence in the putative estrogen response genes in C2.

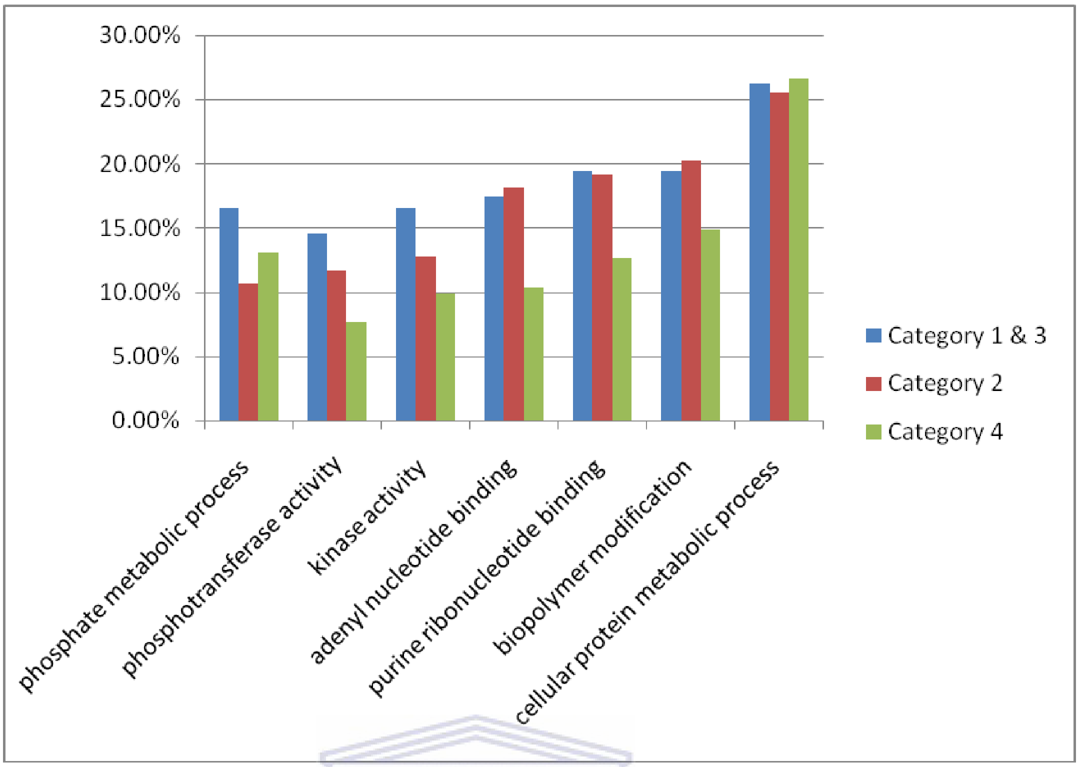
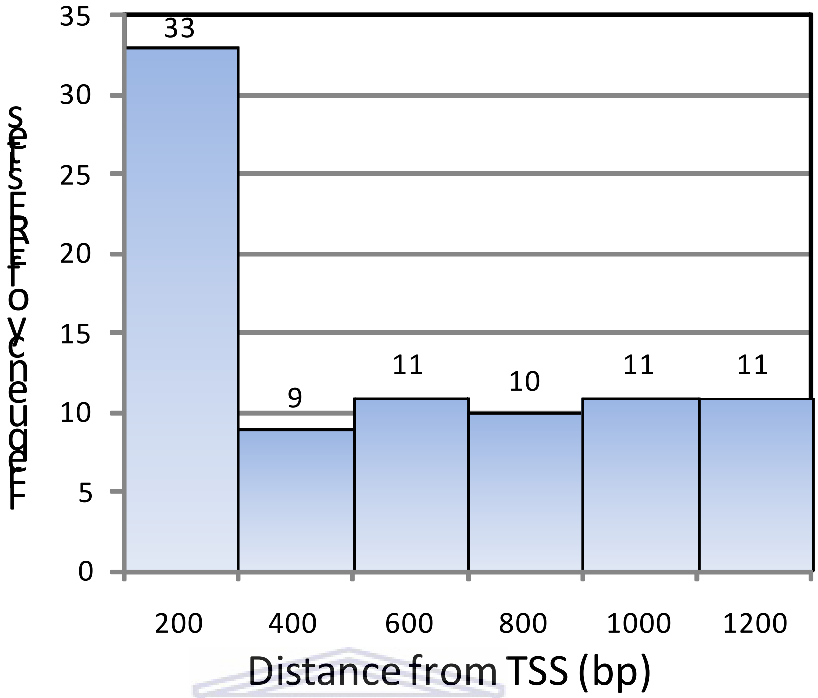


Figure 3.1: A graphical representation of the underlying processes over-represented in the GO analysis of C1&3, C2 and C4. This figure illustrates that processes such as adenyl nucleotide binding, biopolymer modification and purine ribonucleotide binding are dominant in C1&3 and C2 compared to C4.

3.4.3.1 TFBS MATRICES MAPPED TO THE PROMOTER SEQUENCE OF ALL ESCC GENES

The same promoter sequences extracted for the mapping of EREs in the promoter region of ESCC genes were also used in this analysis. The 522 TRANSFAC mammalian matrix profiles of TFBSs were mapped to the promoter sequences using Match^{TM} (176; 177). Of the 522 matrices mapped, 492 unique matrices mapped to the promoter sequences of all the ESCC genes at 165,788 positions. Of the 492 matrices mapped, 472 unique matrices were mapped to the promoter sequences of the putative estrogen responsive genes (C1 and C2) at 21,241 positions, and of the 21,241 TFBSs, 551 TFBSs overlapped with predicted EREs (see appendix III). Knowledge of overlapping TFBSs is essential to understanding underlying biological processes as these overlapping motifs have been shown to behave as regulatory mechanisms, in the expression of the gene of interest and they may occur via a multistage mechanism where the transcription factors bind the promoter region in a stepwise fashion (*193*; *194*).

The distance of functional promoter elements to the TSS of a gene is related to the transcriptional regulation of that gene (*199*). For example, Zou *et al.* demonstrated the 80-fold higher functionality of motifs in the proximal promoter region, defining a minimum proximal promoter region of the *NF1* gene (-270 to +230) and further suggesting that a repressor region exists within the 300 bases upstream of the *NF1* translation start site (+354 to +539) (*192*). The distribution of the 86 experimentally validated EREs (C1) and the 99 putative EREs (C2) from the TSSs are depicted in Figure 3.2. Figure 3.2A clearly depicts that the distribution of EREs in C1 (33/86) is concentrated in the proximal promoter region (-200bp to +200bp). This distribution pattern depicts higher functionality of motifs in the proximal promoter region. Figure 3.2B clearly depicts that the distribution of the putative EREs in C2 (22/99) is concentrated in the proximal promoter region. This suggests the probability that these putative EREs may be functional. Furthermore, the overall distribution pattern in Figure 3.2A and 3.2B are relatively similar.


A) Distribution of ERE sites with respect to TSS position in each promoter in C1

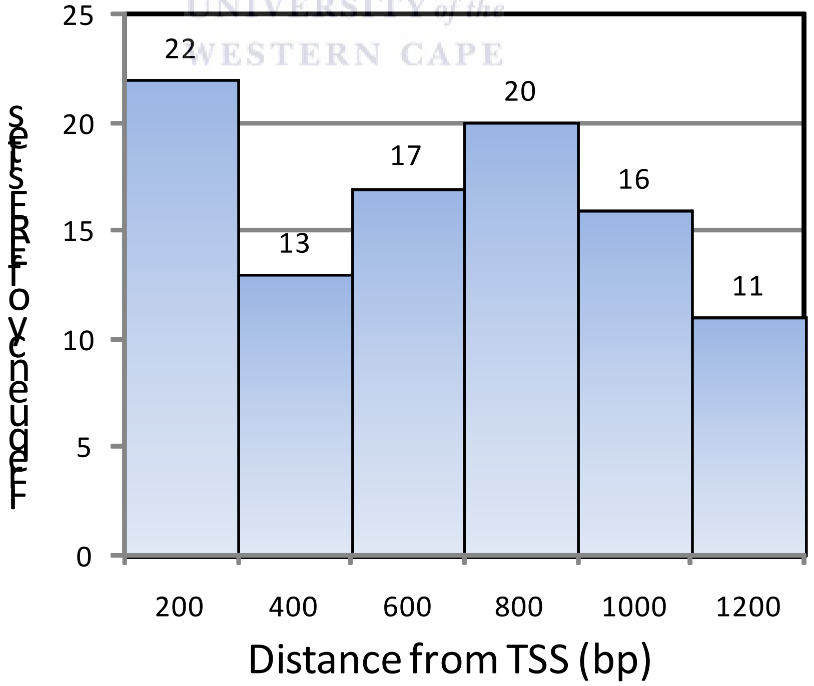

B) Distribution of ERE sites with respect to TSS position in each promoter in C2

Figure 3.2: Histograms representing the distribution of ERE matrices mapped to the promoter sequences of genes in C1 (A) and C2 (B) from the TSSs, respectively.

3.4.3.2 COMBINATIONS OF TFBS SIGNIFICANTLY OVER-REPRESENTED IN C1&3 AS OPPOSED TO C4

Gene expression is driven by the cohesive action of multiple TFs binding to specific TFBSs. Thus, comparable patterns of binding site motifs are able to define co-regulated genes (*229;230*). To identify the combinations of TFBS, significantly overrepresented in C1&3 as opposed to C4, the one-sided Fisher's exact test was computed to calculate the Bonferroni-corrected p-values for each of the TFBSs (correction factor = 522) predicted and the p-values for all combinations of paired TFBSs (correction factor = 522^2-522/2). To minimise the computational complexity, a ranked list of TFBSs was further computed by summing the above mentioned p-values for each unique TFBS. The 10 TFBSs with the lowest sum of p-values were used to calculate the p-values for every possible TFBS combination. The significant combinations with a p-value < 0.05 were selected (see Table 3.2). Significant combinations consisting of 12 combinations of two TFBS, 18 combinations of three TFBS, 10 combinations of four TFBS, 3 combinations of five TFBS and 1 combination of six TFBS were identified. These combinations, listed in Table 3.2, can be used to increase confidence that the putative estrogen responsive genes may be estrogen responsive, as these combinations of TFBSs were found to be significantly over-represented in a known estrogen responsive gene set. The 10 TFBSs that comprise these combinations are V$ELK1_01, V$CETS1P54_01, V$YY1_01, V$GATA3_01, V$TAXCREB_02, V$FREAC4_01, V$AREB6_01, V$CREB_Q3, V$E2A_Q6 and V$EBOX_Q6_01. Of the 44 significant combinations, eight combinations were not present in category C4. The most significant combination in C1&3 compared to C4 is V$AREB6_01,

Table 3.2: TBFS combinations significantly over-represented in ESCC genes

ERE Category	C1+C3		C1		C2		C3		C4	
Total Number of Genes	202		56		72		146		144	
	Counts	Percentage	Counts	Percentage	Counts	Percentage	Counts	Percentage	Counts	Percentage
M00412|M00059|M00032|M00973	17	8.42%	8	14.29%	9	12.50%	9	6.16%	0	0.00%
M00412|M00292|M01034|M00973	16	7.92%	11	19.64%	7	9.72%	5	3.42%	0	0.00%
M00115|M00412|M00032|M01034|M00801|M00973	16	7.92%	9	16.07%	7	9.72%	7	4.79%	0	0.00%
M00115|M00412|M01034|M00801|M00973	19	9.41%	9	16.07%	11	15.28%	10	6.85%	0	0.00%
M00115|M00292|M00973	18	8.91%	10	17.86%	8	11.11%	8	5.48%	0	0.00%
M00115|M00412|M00801|M00973	22	10.89%	10	17.86%	14	19.44%	12	8.22%	0	0.00%
M00115|M00412|M00007|M00801	18	8.91%	8	14.29%	8	11.11%	10	6.85%	0	0.00%
M00115|M00412|M00032|M00801|M00973	17	8.42%	9	16.07%	9	12.50%	8	5.48%	0	0.00%
M00115|M00007|M01034|M00801	19	9.41%	10	17.86%	8	11.11%	9	6.16%	1	0.69%
M00412|M00007|M00801	19	9.41%	8	14.29%	8	11.11%	11	7.53%	1	0.69%
M00115|M00412|M01034|M00801	21	10.40%	11	19.64%	11	15.28%	10	6.85%	1	0.69%
M00412|M00292|M00973	19	9.41%	13	23.21%	7	9.72%	6	4.11%	1	0.69%
M00412|M00032|M00801|M00973	19	9.41%	10	17.86%	9	12.50%	9	6.16%	1	0.69%
M00115|M00292|M00032	20	9.90%	14	25.00%	10	13.89%	6	4.11%	1	0.69%
M00115|M01034|M00801|M00973	24	11.88%	12	21.43%	12	16.67%	12	8.22%	1	0.69%
M00115|M00412|M00032|M01034|M00973	21	10.40%	12	21.43%	12	16.67%	9	6.16%	1	0.69%
M00292|M00032|M01034	18	8.91%	13	23.21%	10	13.89%	5	3.42%	1	0.69%
M00412|M01034|M00801|M00973	21	10.40%	10	17.86%	11	15.28%	11	7.53%	1	0.69%
M00292|M00032|M00973	18	8.91%	12	21.43%	7	9.72%	6	4.11%	1	0.69%
M00412|M01034|M00801	23	11.39%	12	21.43%	11	15.28%	11	7.53%	2	1.39%
M00115|M01034|M00801	27	13.37%	14	25.00%	13	18.06%	13	8.90%	2	1.39%
M00115|M00007|M00032|M01034	23	11.39%	14	25.00%	16	22.22%	9	6.16%	2	1.39%

Table 3.2: continued

ERE Category Total Number of Genes	C1+C3		C1		C2		C3		C4	
	202		56		72		146		144	
	Counts	Percentage	Counts	Percentage	Counts	Percentage	Counts	Percentage	Counts	Percentage
M00412|M00801|M00973	24	11.88%	11	19.64%	14	19.44%	13	8.90%	2	1.39%
M00292|M01034	22	10.89%	15	26.79%	13	18.06%	7	4.79%	3	2.08%
M00115|M00412|M00801	27	13.37%	12	21.43%	17	23.61%	15	10.27%	3	2.08%
M00412|M00032|M00801	24	11.88%	11	19.64%	9	12.50%	13	8.90%	3	2.08%
M00115|M00007|M00801	28	13.86%	12	21.43%	11	15.28%	16	10.96%	3	2.08%
M00412|M00292	26	12.87%	15	26.79%	11	15.28%	11	7.53%	4	2.78%
M00115|M00801|M00973	30	14.85%	14	25.00%	15	20.83%	16	10.96%	4	2.78%
M00292|M00973	25	12.38%	14	25.00%	9	12.50%	11	7.53%	4	2.78%
M01034|M00801|M00973	27	13.37%	13	23.21%	13	18.06%	14	9.59%	4	2.78%
M00007|M00801	32	15.84%	12	21.43%	11	15.28%	20	13.70%	5	3.47%
M00292|M00032	27	13.37%	17	30.36%	12	16.67%	10	6.85%	5	3.47%
M00115|M00032|M01034	34	16.83%	18	32.14%	21	29.17%	16	10.96%	5	3.47%
M00412|M00032|M00973	30	14.85%	16	28.57%	16	22.22%	14	9.59%	5	3.47%
M00412|M00801	31	15.35%	13	23.21%	17	23.61%	18	12.33%	6	4.17%
M00115|M00032|M00973	33	16.34%	15	26.79%	19	26.39%	18	12.33%	6	4.17%
M01034|M00801	32	15.84%	16	28.57%	15	20.83%	16	10.96%	7	4.86%
M00412|M01034|M00973	38	18.81%	16	28.57%	21	29.17%	22	15.07%	8	5.56%
M00007|M01034	36	17.82%	17	30.36%	18	25.00%	19	13.01%	9	6.25%
M00032|M00801	38	18.81%	15	26.79%	14	19.44%	23	15.75%	10	6.94%
M00032|M01034	41	20.30%	21	37.50%	24	33.33%	20	13.70%	11	7.64%
M00115|M00801	43	21.29%	18	32.14%	23	31.94%	25	17.12%	12	8.33%
M00412|M00973	50	24.75%	23	41.07%	25	34.72%	27	18.49%	16	11.11%

V$TAXCREB_02, V$TAXCREB_02, V$E2A_Q6. This combination was not present in C4, but it was found in 14.29% of C1&3 and 12.50% of C2. There were other combinations of TFBSs that were not found in C4 too (see Table 3.2)

V$AREB6_01 binds AREB6 (also known as ZEB1) (*231*). ZEB1 has been implicated in the transcriptional repression of genes such as IL2 (*232*), CDH13 (found in C3) (*233*), CDH1 (found in C2) (*234*), ARHGAP24 (*235*), PKP3 (*236*) and BZLF1 (*237*). ZEB1 and SIP1 are also implicated in the transcriptional repression of miR-200b (*238*) and conversely, the expression levels of ZEB1 is regulated by the microRNA-200 family (miR-200a, miR-200b, miR-200c, miR-141 and miR-429) and miR-205 (*239*) thereby producing a double negative feedback loop. Published biological data demonstrate that the miR-200 family plays a major role in specifying the epithelial phenotype by preventing expression of the transcription repressors, ZEB1/deltaEF1 and SIP1/ZEB2 (*240;241*). This epithelial to mesenchymal transition (EMT) occurs during embryologic development to allow tissue remodelling and is proposed to be a key step in the metastasis of epithelial-derived tumours. Thus, the double negative feedback loop regulates cellular phenotype and has direct relevance to the role of these factors in tumour progression. ZEB1 has also been found to mediate disintegration of intercellular adhesion and EMT via the transcriptional repression of genes such as CRB3, LLGL2 and MPP5 (*242*). Thus, it has been documented as a key biomarker of aggressive cancer at high risk of recurrence and is recognised as a therapeutic target (*243*).

V$TAXCREB_02 binds CREB, deltaCREB and Tax/CREB complex (*244;245*). Over-expression of CREB has been linked to numerous cancer types such as ovarian cancer (*246*), gastrointestinal cancer (*247*), breast cancer (*248*), duodenal cancer (*249*) and prostate cancer (*250*). CREB transcribes more than 5000 target genes. Expression of CREB is directly regulated by miR-34b and the miR-34b promoter was shown to be methylated in the leukemia cell lines used (*251*). CREB has been documented as a prognostic marker in acute myeloid leukemia (*252*). V$CREB_Q3 binds CREB1, CREMalpha, deltaCREB, ATF-1, ATF-2, ATF-3, ATF-4, ATF-a, and ATF-2-xbb4. Tax represses the transcription of TATA-less CCNA2, CCND3 and POLA1 promoters by attaching to the CREB/ATF complex bound to the CREB/ATF binding site (*253*). The interaction between Tax and CREB is highly specific, whereas Tax interacts with ATF only marginally despite the extensive sequence similarities between CREB and ATF. Thus, if CREB is not bound to the motif then transcriptional repression is absent (*254*). ATF-1 and ATF-2 is up-regulated and bind to Tax-responsive elements in infected bovine B lymphocytes (*255*). Additionally, BATF is up-regulated in human B lymphocytes infected by Epstein-Barr virus (EBV) and was shown to negatively impact the expression of a *BZLF1* reporter gene and to reduce the frequency of lytic replication in latently infected cells (*256*), suggesting that BATF impact viral and cellular gene expression by promote viral latency and control lytic-cycle entry. ZEB1, CREB1 and ATF1 are members of the bZIP family of leucine-zipper transactivators (*257;258*).

The PWMs designated V$CREB_Q3 and V$E2A_Q6 maps motifs known to bind multiple TFs, as this matrix was derived from a collection of TFBSs for many TFs (176). Thus, the TF associated with this significant combination cannot be defined for these matrices.

3.4.3.3 THE SIGNIFICANTLY OVER-REPRESENTED COMBINATIONS OF TFBSS UTILIZED TO INCREASE CONFIDENCE IN THE NOVEL PUTATIVE ESTROGEN RESPONSE GENES IN C2

Combinations that were significantly over-represented in C1&3 were further compared to the TFBSs predicted in the promoter regions of the genes in C2. A general comparison of these combinations to the TFBSs predicted in the promoter regions of the genes in C1, C2, C3 and C4 displayed matches of 574, 567, 561 and 153 combinations, respectively. It was determined that 71.43% (40/56) of the genes in C1, 66.67% (48/72) of the genes in C2, 46.57% (68/146) of the genes in C3 and 34.72% (50/144) of the genes in C4, had at least one of the significant combinations in the promoter sequence (see Table 3.2). These subsets of gene categories will be referred to with the postfix 'A', i.e. C1A, C2A, C3A, and C4A.

Even though the significant combinations were determined based on C1&3, the findings show that C1, representing genes both predicted and confirmed to be estrogen responsive has 24.86% more genes with a significant combination in the promoter sequence as compared to C3. This result indicates that C1 gene promoters contain a distinctive combination of TFBSs that are likely to define co-regulated genes. Thus the 66.67% of putative estrogen responsive genes in C2A that display at least one of these combinations may be an additional fraction of these co-regulated genes. These results strengthen support the putative estrogen responsive genes in C2A.

An overview of the regulatory effects of the combinations on the C2A genes is shown in Figure 3.3. The figure represents each association in C2A, in the form of a color dot in a heat map format using TMEV (*259;260*). The different shades of red

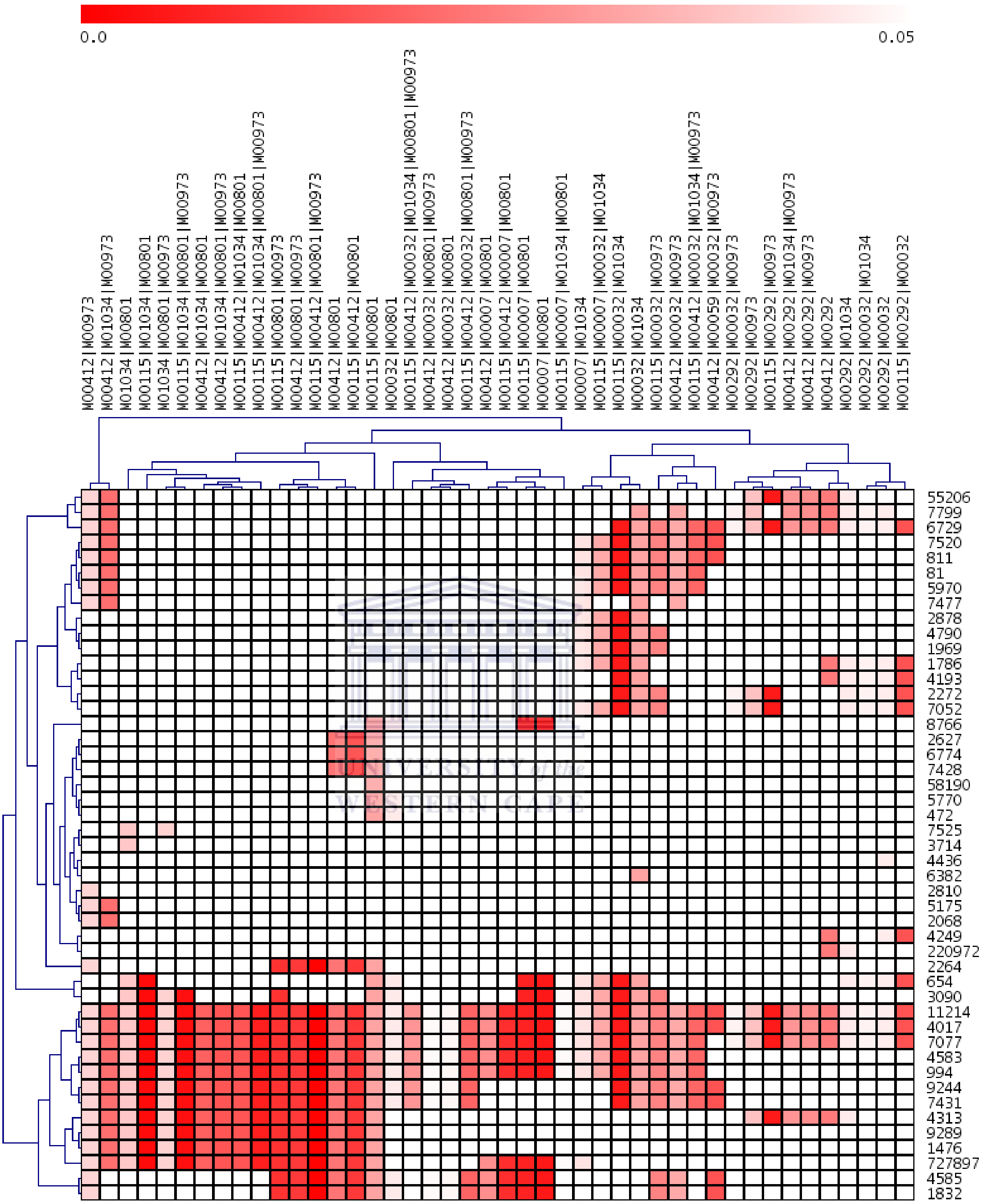


Figure 3.3: Heat map of combinations of TFBSs that are significantly found in the promoter sequence of estrogen response genes (columns) maps to the promoter sequence of putative estrogen response genes (rows). The more intense the shades of red depict the strength of association in terms of the calculated p-Value for the combination.

depict the strength of association in terms of the calculated p-value for the combination. If there is no colour dot, then no association has been made between the combination and gene. The heat-map has been clustered using hierarchical clustering with average linkage and Euclidian distance. The heat map depicts clusters of genes that may be regulated by the same combination/s. Of the 46 combinations identified 45 combinations are present in C2A and form clusters based on these combinations. Of the 48 genes in C2A all the genes have multiple combinations present in their promoter sequences except for SFN, JAG2, MSH2, SDC1 and CTDSP1. The heat map clearly depicts genes clusters based on multiple combinations common to the promoter sequences of numerous genes.

As an example, AKAP13 (Gene ID: 11214), LOXL2 (Gene ID: 4017), TIMP2 (Gene ID: 7077) cluster together and 33 combinations are common between their promoter sequences. Pathway analysis of these three genes using KEGG (via DAVID) produced no results. However, a literature survey draws attention to AKAP13 binding ER in estrogen responsive reproductive tissue. Also, polymorphisms in the AKAP13 gene have been linked to breast cancer (*261*). LOXL2 interacts with Snail to down-regulate CDH1 expression, and control the embryonic developmental step, EMT and consequently metastasis of epithelial derived tumours (*262*). TIMP2 is abundant in the ovarian perifollicular stroma. Its median level increased in the early ovulatory phase (*263*). TIMP2 inhibits MMP2 (*264*). MMP2 is also one of the predicted estrogen responsive gene. Prast *et al.*, demonstrated that human chorionic gonadotropin promotes trophoblast invasion and migration through activation of ERK and AKT signaling involving their downstream effector MMP2 (*265*). Gonadotropin-releasing hormone (GnRH), also known as Luteinizing-hormone releasing hormone (LHRH), activates gonadotropin-releasing hormone receptor (GNRHR), a seven transmembrane G-protein coupled receptor that sequentially lead to the release of follicle stimulating hormone (FSH) and Luteinizing-hormone (LH) from the anterior pituitary gland under the influence of the hypothalamus (*266*). FSH stimulates the growth and recruitment of immature ovarian follicles in the ovary. As the follicle matures, one becomes dominant and releases inhibin and estradiol, both of which decrease FSH production by inhibiting GnRH production in the hypothalamus (*267*). Overall, these genes display some link with estrogen and EMT.

Recent 2009 publication verifies that two of the predicted estrogen response genes are estrogen responsive, namely; Choi *et al*., demonstrated that estrogen induced increased MUC5B expression in human airway epithelial (*268*) and Grandas et al., demonstrated estrogen induced increased MMP2 expression in vascular smooth muscle cells (*188*).

3.5 CONCLUSION

These findings provide the research community with combinations of transcription factor binding sites that can be used to identify estrogen responsive genes of potential relevance for ESCC and divulge more genes that possibly play a role in the underlying mechanisms affected by estrogen treatment. Furthermore, this method for determining genes with hormone responsive elements could be applicable to other hormone responsive cancers.

CHAPTER 4

PINPOINTING MASTER REGULATORS OF ESOPHAGEAL CANCER AS CANDIDATE DIAGNOSTIC MARKERS

4.1 ABSTRACT

Mortality rates associated with EC are very similar to the incidence rates due to the relatively late stage of diagnosis. A lack of understanding of the transcriptional underpinning that characterizes EC prevents the effective identification of candidate diagnostic markers that may enable the detection of EC in the early stage. In this section, the transcriptional circuitry characterizing ESCC genes extracted from the DDEC is computationally analyzed to pinpoint putative master regulators of ESCC.

Key to pinpointing the master regulators is to comprehend that TFs bind to TFBSs in the gene promoter sequence, thereby driving the expression of the transcripts; thus, TFs are the regulators. This section, demonstrates how one can computationally infer the candidate diagnostic markers based on SNPs and ESCC genes verified with the *in vitro* biological expression data. Nine master regulators that include PAX5, TFAP2A, ZIC2, CREB1, E2F1, TFAP2C, E2F3, E2F4 and TFDP1 were predicted as candidate diagnostic markers.

These findings provide the research community with the master regulators characterizing ESCC and these can be explored as potential therapeutic targets. We further identify the presence of a SNP occurring within overrepresented TFBSs of the master regulators, as this finding increases confidence that such an altered

TFBS (with the SNP) is a candidate biomarker for ESCC. To my knowledge this is the first study identifying master regulators for ESCC.

4.2 INTRODUCTION

Gene expression is a multifunctional process influenced by chromatin remodelling and the interplay between transcription factors (TFs) and DNA sequences functioning as transcription factor binding sites (TFBSs) (*269;270*). This combination of TFs, TFBSs, and affected transcripts, defines the transcription regulatory networks (TRNs) that are responsible for the regulation of every transcript encoded in the genome. TFs bind to the TFBSs thereby mediating gene expression. Numerous factors such as the position of the TFBSs in the promoter sequence (*192*), the presence or absence of co-factors, overlapping TFBSs (*193*; *194*; *195*) and genetic alterations (*271;272*) may alter the TFs binding affinity or outright inhibit the TFs binding and consequently gene expression.

Genetic alterations (called polymorphisms) occur at >1% in the human population, while the impact of most polymorphisms are not known, an abundance of scientific research demonstrate that certain variants are associated with specific population groups, disease susceptibility, and individual response to therapeutic treatments and environmental exposure (*273*). Common polyphorphisms include tandemly repeated segments (minisatellite, 0.1-2.0 kb; microsatellite 2-100 nucleotides), large and small deletions/insertions/duplications and SNPs. SNPs account for approximately >90% of all known sequence variants occurring once per 100-300 bases (*273;274*). SNPs that affect gene expression occur in both the gene coding and regulatory regions of the genome. SNPs located in the coding region (cSNPs) of a gene cause errors in protein folding, polarity shift and improper phosphorylation (*275;276*). SNPs located in the gene regulatory region (rSNPs) such as promoters, enhancers, silencers and introns have been demonstrated to effect disease susceptibility including acute lung injury (*277*), breast cancer (*278*) and sudden infant death syndrome (*279*). There are several examples of rSNPs causing either complete elimination of TFBSs (*280*) or the formation of a novel TFBSs that diminishes TF binding capacity or/and introduce a scenario where TFs compete for the TFBS (*281;282*). Susceptability and prognosis of complex diseases such as cancer is determined by a combination of factors including; a spectrum of SNPs and environmental factors that cause DNA damage. The high frequency of stable SNPs makes them excellent biomarkers and drug response is thought to be dependent on an individual's SNP profile.

The onset of ESCC has been associated with genetic variations and unfavorable environmental factors. The PubMed database was queried with keyword expression: "(SNP AND esophageal squamous cell carcin*) on 31/01/2009 and 25 entries were retrieved. Literature indicates that SNP in the promoter of genes such as 53BP1 T885G (*283*), IL10 G1082A (*284*), MMP1 insertion/deletion polymorphism 2G/1G (*285*), CDH1 C-160A and G-347GA (*286*) and MMP-13 A-77G (*287*) are not associated with ESCC, whilst SNP such as MMP3 insertion/deletion polymorphism 6A/5A (*288*), TS 3RC (*289*), and ARL6IP5 G-76C (*290*) are strongly associated with ESCC. This result indicates that SNPs that characterize ESCC has not been extensively researched.

4.3 METHODOLOGY

4.3.1 OVERVIEW

This study specifically focuses on identifying the master regulators of ESCC and TFBSs with overlapping single nucleotide polymorphisms (SNPs) that are presumed to underlie the differences in human biological traits and disease susceptibility (*291;292*). A list of 418 differentially expressed ESCC genes and the associated promoter sequences were extracted as per section 3.3.2. In succession, the putative TFBSs that characterize these promoter sequences were determined using position weight matrices (PWMs) from TRANSFAC Professional database v.11.4 as per section 3.3.5. The hypothesis was that since TFs bind to TFBSs to drive gene expression, the TFs that bind to the statistically significant TFBSs mapped to the most promoter sequences are the master regulators or Hub TFs of the ESCC genes. The hub TFs were then screened for overlapping SNPs. SNPs was extracted from HapMap (*293*). The presence of a SNP overlapping the TFBSs will increase confidence that such an altered TFBS (with the SNP) is a candidate biomarker for ESCC (*294*).

4.3.2 EXTRACTING OF THE PROMOTER SEQUENCES OF THE ESCC GENES

As per section 3.3.2

4.3.3 MAPPING TFBS MATRICES TO THE PROMOTER SEQUENCE OF ALL ESCC GENES

As per section 3.3.5

4.3.4 DETERMINING THE MASTER REGULATORS OF ESCC

To determine the TFBSs specific to the target set of ESCC promoters, the overrepresentation index (ORI) was calculated using a background set of 40,101 randomly chosen unique human promoters. This calculation was done by in-house computer scripts. Readers interested in applying the same filtering process should contact Vladimir B. Bajic (vlad@orioncell.org). An ORI value of one means no overrepresentation of the TFBS in the target promoter group compared to the background set. A greater ORI signifies over representation of the TFBS in the target promoter group. The one-sided Fisher's exact test was used to calculate p-Values for each of the TFBSs predicted and Bonferroni corrections were applied to all p-values. All TFBSs with an ORI below two and p-values above 0.05 were filtered out. The remaining TFBS were used to annotate the promoters.

4.3.5 CONSTRUCTING A TRANSCRIPTIONAL REGULATORY NETWORK (TRN)

All network graphics have been produced with cytoscape (*226*). The networks comprise TFBS→gene edges. These associations were created by mapping TFBSs to the promoters of the genes of interest (gene→promoter→TFBS). The green rectangles in the networks represent TFBS nodes and the orange circles represent the genes. The TFBS nodes and genes are further linked to eachother by black edges.

4.3.6 DETERMINING THE TFBS'S WITH OVERLAPPING SNP'S

The SNP analysis was performed utilising 4,766,707 unique SNPs collected from the HapMap database. All the TFBS predictions were tested for a possible overlap with one or more SNPs. The absolute chromosomal positions of the TFBS predictions have been calculated with respect to the positions of the TSS that were used to define the promoter regions of the genes. An overlap is found if the chromosomal position of a SNP falls within the start and stop position of a predicted TFBS. The above calculations were done by in-house computer scripts. Readers interested in applying the same filtering process should contact Vladimir B. Bajic (vlad@orioncell.org).

4.4 RESULTS AND DISCUSSION

As a consequence of gene expression being driven by the cohesive action of multiple TFs binding to specific TFBSs; the TFs that bind to the statistically significant TFBSs mapped to the most promoter sequences, thereby driven the expression of the majority of transcripts involved in ESCC, are the master regulators or Hub TFs of the ESCC (*295;296*).

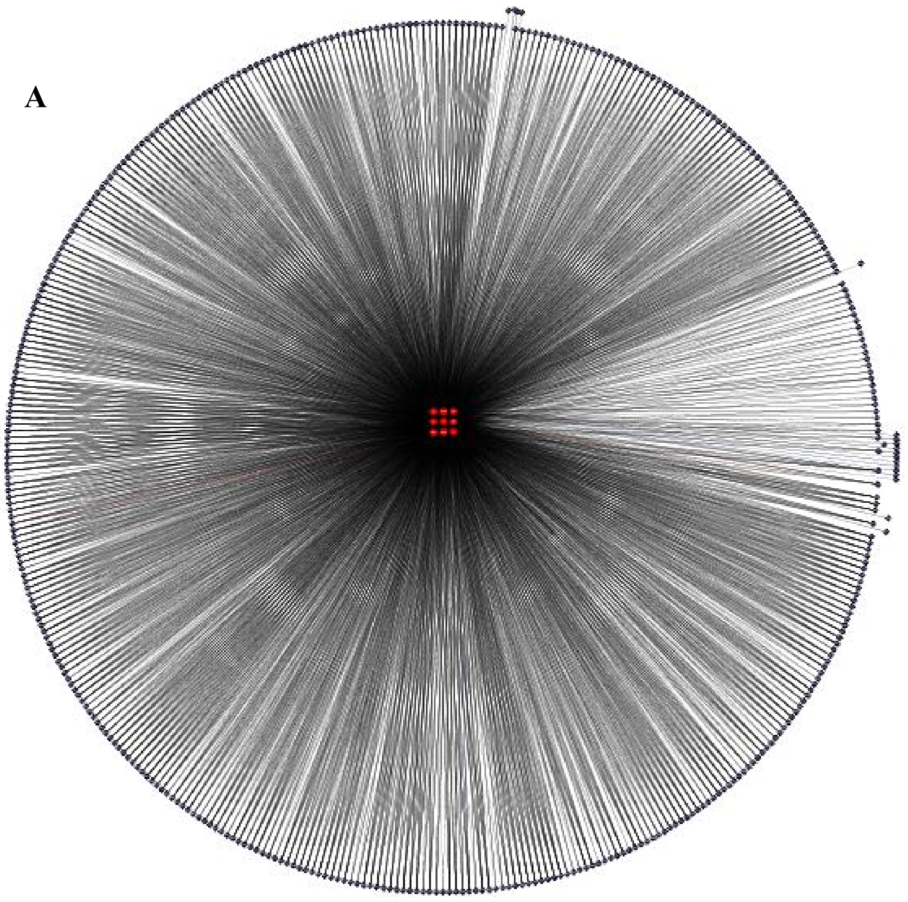
4.4.1 IDENTIFYING THE MASTER REGULATORS ASSOCIATED WITH ESCC

TFBSs matrices were mapped to the promoter sequences of the differentially expressed ESCC genes as per section 3.3.5. To determine the TFBSs specific to the target set of ESCC promoters, the ORI were calculated and the one-sided Fisher's exact test was used to calculate p-Values for each of the TFBSs predicted and Bonferroni corrections were applied to all p-values. All TFBSs with an ORI above 2 and p-values below 0.05 were used to annotate the promoter sequences used to drive the expression of ESCC genes. To infer the gene regulatory network from the TFBS predictions, this list of TFBSs were expanded to the TFs that bind the motifs, TFs could thus be connected to genes that have the predicted TFBSs for that TF (TF→TFBS→gene). Owing to the use of differentially expressed ESCC genes and the filtering procedure used, 53 385 unique edges that are predominantly active in ESCC were identified. The combinatorial complexity of a TRN produced from 53

385 edges makes it impractical to interpret. Nonetheless, it provides valuable information as different subnetworks of the global TRN define different cellular states. A subset of the TRN for ESCC is shown in Figure 4.1a

Master regulators may be defined as proteins whose ectopic expression affects a number of critical components that can mediate the entire developmental process. Master regulators have been demonstrated to act concurrently and their loss has been demonstrated to arrest the response in question (*297*). Since TFs mediate pathways through the expression of specific downstream genes, the TFs predominantly driving the expression of ESCC genes are the mater regulators. The master regulators were determined by expanding the TFBSs used to annotate the promoter sequences of the ESCC genes to the associated TFs known to bind these TFBSs (also called Hub TFs). These TFs were ranked by the number of transcripts each TF regulates (out-degree). The top 10% of most regulating TFs or master regulators are tabulated in Figure 4.1b.

The most regulating TF, PAX5 is predicted to regulate 375 transcripts. This TF has been demonstrated to play a vital role in B-lineage commitment in the midbrain development (*298*), progression of lymphoid development (*299*) and it has been demonstrated that PAX5 expression is down-regulated in plasma cell differentiation (*300*) and up-regulated in low- to intermediate-grade pulmonary neuroendocrine neoplasms (*301*). Chromosomal abnormalities in B-cell precursor acute lymphoblastic leukemia includes a breakpoint on 9p that results in the partial or complete deletion of PAX5; consequently it was demonstrated that PAX5 expression was down-regulated along with



B

Out-degree	Gene symbol
375	PAX5
344	TFAP2A
318	ZIC2
318	CREB1
307	E2F1
305	TFAP2C
300	TFDP1
300	E2F4
300	E2F3

Figure 4.1 Putative master regulators for ESCC. **A)** Subset of the TRN for ESCC for the nine putative master regulators (red) as central nodes. **B)** Tabulation of the nine putative master regulators based on the number of transcripts each TF regulates (out-degree).

PAX5 target genes such as EBF1, ALDH1A1, ATP9A, and FLT3 (*302*). Another study demonstrated that this chromosomal breakpoint results in unbalanced translocation of PAX5 that leads to the formation of fusion genes; the PAX5 gene was rearranged to a variety of partner genes including ETV6, FOXP1, AUTS2, and C20orf112 in pediatric acute lymphoblastic leukemia (ALL). The expression of the PAX5 fusion products were demonstrated to suppress the PAX5 transcriptional activity (*303*). Additionally, PAX5 expression was not detected in ESCC (*304*).

The second most regulating TF, TFAP2A is predicted to regulate 344 transcripts. This TF plays a vital role in vertebrate development (*305*) and transformation (*306*). Reduced expression of TFAP2A has been demonstrated in numerous cancer types including cervical cancer (*307*), breast cancer (*308*), colon cancer (*309*) and prostate cancer (*310*). It has been demonstrated that loss of TFAP2A expression in clear cell renal cell carcinoma is owing to hypermethylation of this tumor suppressor gene (*311*). It has also been demonstrated that TFAP2A reduces chemotherapy-induced apoptosis, migration, motility whilst increasing cell adhesion (*307*); and inhibits the growth of cells by inducing cell cycle arrest and apoptosis (*312*). Loss of TFAP2A has also been demonstrated as a marker of early stage human gliomas as normal brain tissue and grade 2 astrocytomas maintained TFAP2A expression in all tissue (100%), whilst TFAP2A expression is lost in most grade 3 anaplastic astrocytomas and grade 4 glioblastomas (98-99%) (*313*). To my knowledge the role of TFAP2A in ESCC was not investigated to date.

The third most regulating TF, ZIC2 is predicted to regulate 318 transcripts. This TF has not been extensively studied but is known to play a vital role in mammalian development. It has been demonstrated that loss of ZIC2 expression causes holoprosencephaly (HPE) in humans and mice (*314;315*). ZIC2 mutant mice showed additional abnormalities, including spina bifida, axial/limb skeletal malformation, retinal axon mis-projection, and reduced production of neural crest cells (*315*). Reduced ZIC2 expression has also been correlated with poor clinical outcome of pediatric medulloblastomas owing to the hypermethylation of this gene (*316*). ZIC2 was investigated as a possible ESCC tumor antigen, however, ZIC2 expression was not significantly differential in ESCC patients compared to healthy controls (*317*).

The fourth most regulating TF, CREB1 is predicted to regulate 318 transcripts. This TF plays a vital role in cellular growth, proliferation and oncogenesis and is responsible for more than 5000 target genes (*318*). It has been demonstrated that CREB1 is frequently overexpressed in ovarian cancer (*319*), breast cancer (*320*), duodenal cancer (*321*) and acute myeloid leukemia (*322*). Loss of heterozygosity (LOH) and internal tandem duplications (ITD) of the CREB1 gene have been shown to be frequent genetic events in human ESCC (*323*). The genetic alteration hot spot was identified in close proximity (within 40 kb) of a 13-kb chromosomal translocation breakpoint of *CBP* that has been reported earlier in leukemia. These genetic abnormalities have been associated with intrachromosomal recombination including *CBP-MLL*, *CBP-MOZ*, and *CBP-MORF* translocations found in leukemia patients (*324-326*).

The fifth most regulating TF, E2F1 is predicted to regulate 307 transcripts. This TF is a member of the E2F family plays a crucial role in cell proliferation (*327*), apoptosis (*328*) and differentiation (*329*). It has been demonstrated that E2F1 is frequently overexpressed in ovarian cancer (*330*), oral squamous cell carcinoma (*331*), non-small-cell lung cancer (NSCLC) (*332*), clear cell renal cell carcinoma (ccRCC) (*333*) and gastric cancer (*334*). Gastric cancer tissue were also screened for mutations in the E2F1 gene, mutations were not detected for E2F1 gene (*334*). E2F1 over-expression was also detected in ESCC (*335;336*) and E2F1 amplification was detected in ESCC and correlates with a significantly lower survival rate (*337*).

The sixth most regulating TF, TFAP2C is predicted to regulate 305 transcripts. This TF plays a vital role in development (*338*). It has been demonstrated that TFAP2C expression is frequently reduced in breast cancer (*339*). However, increased expression of TFAP2C was observed in breast cancer patients resistant to anti-hormone therapy as it contributes to failure of cells to growth arrest (*340*). TFAP2C expression was also demonstrated in neuroblastoma and increased expression was correlated with tumour grade (*341*). To my knowledge the role of TFAP2C in ESCC was not investigated to date.

The seventh most regulating TF, TFDP1 is predicted to regulate 300 transcripts. This TF plays a vital role in cell cycle progression and cell growth. It was demonstrated that SOCS3 and ARF binds TFDP1 thereby inhibiting E2F/TFDP-1 transcriptional activity and consequently cell cycle progression and cell growth under E2F/TFDP-1 control (*342;343*). TFDP1 was found to be over-expressed in hepatocellular carcinomas (HCC) (*344*). Comparative genomic hybridization (CGH) revealed frequent amplification at 13q34, in HCC and ESCC. Genes such as TFDP1, CUL4A and CDC16 showed distinct amplification in this region for HCC (*345*). Over-expression of TFDP1 may thus contribute to the progression of these cancer types by promoting growth of the tumor cells. It has also been demonstrated that TFDP1 and E2F1, E2F2, and E2F3 form a transcriptionally active heterodimer that is required for nuclear localization. Unlike E2F1, E2F2, and E2F3; E2F4 did not accumulate in the nucleus unless it heterodimerized with TFDP2, p107 and p130. RB binds these heterodimers thereby preventing cell cycle progression. (*346;347*).

The eighth and ninth most regulating TFs, E2F3, and E2F4 are predicted to regulate 300 transcripts. These TFs are also members of the E2F family that play distinctive roles in differentiation and proliferation (*329*). It has been demonstrated that E2F3 is significantly over-expressed in retinoblastoma (*348*) and small cell lung cancers (SCLC) (*349*). Also, Amplification and over-expression of the E2F3 gene at 6p22 in human bladder cancer is associated with increased tumor stage and grade, and in prostate cancer E2F3 over-expression is linked to tumor aggressiveness (*350;351*). It has also been demonstrated that E2F1, E2F2, and E2F3 transcribes the mir-17-92, suggesting an autoregulatory feedback loop between E2F factors and the mir-17-92 cluster. A member of the mir-17-92 cluster, miR-20a, modulates the translation of the E2F2 and E2F3 mRNAs. Over-expression of this miRNA has been demonstrated to decreases apoptosis in prostate cancer cell lines (*352*). E2F4 was demonstrated to drive the expression of most genes involved in cellular quiescence (*353*). Mutations of the E2F4 gene were detected in hereditary non-polyposis colorectal cancer (HNPCC) and sporadic colorectal cancer (CRC) and E2F4 is considered to play an important role in this carcinogenic process (*354*). E2F4 acts in cooperation with pRB to restrain cell proliferation. It was demonstrated that reduction in E2F4 expression delayed the development, and reduced the incidence, of Rb mutant tumors in the pituitary (*355*). To my knowledge the role of E2F3 and E2F4 in ESCC was not investigated to date.

The literature is summarized in Table 4.1. Literature provides evidence that PAX5, TFAP2A, ZIC2 and TFAP2C are actively involved in normal development and are generally down-regulated in numerous cancer types. In fulfilling this function they maintain a homeostatic environment by protecting the cell against foreign objects and aberrant cell growth, as a consequence their dysregulated expression is restricted to early stage cancers as demonstrated for PAX5 and TFAP2A by Sica *et al*. and Heimberger *et al*., respectively. Literature provides further evidence that CREB1, TFDP1, E2F1, E2F3, and E2F4 are actively involved in proliferation and the E2F family members are also involved in differentiation and have been demonstrated to be regulated by miRNAs in an auto-regulatory feedback loop. CREB1, TFDP1, E2F1, E2F3, and E2F4 are generally up-regulated driving aberrant cell growth in numerous cancer types and the increased expression of TFDP1, E2F1 and E2F3 is associated with genetic amplifications. Knowledge of which master regulators are up-regulated and down-regulators provided an idea of the transition that has taken place in the TRN for normal esophagus compared to the TRN for ESCC.

Additionally, these master regulator matrices mapped to the promoter sequences of the differentially expressed ESCC genes allows for the identification of co-expressed genes. An overview of the regulatory effects of the master regulators on the differentially expressed ESCC genes is shown in Figure 4.2. Figure 4.2 represents each association in Figure 4.1A, in the form of a colour dot in a heat map format using TMEV (*259;356*). The heat-map has been clustered using hierarchical clustering with average linkage and Euclidian distance.

Table 4.1: Summary of the literature review for the putative master regulators.

TF	Expression status cancer types	Mutations type	Expression status in ESCC	Literature
PAX5	down-regulated	deletion	down-regulated	Kawamata *et al.* 2008
TFAP2A	down-regulated	hypermethylation	–	–
CREB1	up-regulated	LOH and ITD	up-regulated	So *et al.* 2004
ZIC2	down-regulated	hypermethylation	normal	Shimada *et al.* 2005
E2F1	up-regulated	amplification	up-regulated	Yamazaki *et al.* 2005
TFAP2C	down-regulated	–	–	–
E2F3	up-regulated	amplification	–	–
E2F4	up-regulated	–	–	–
TFDP1	up-regulated	amplification	–	–

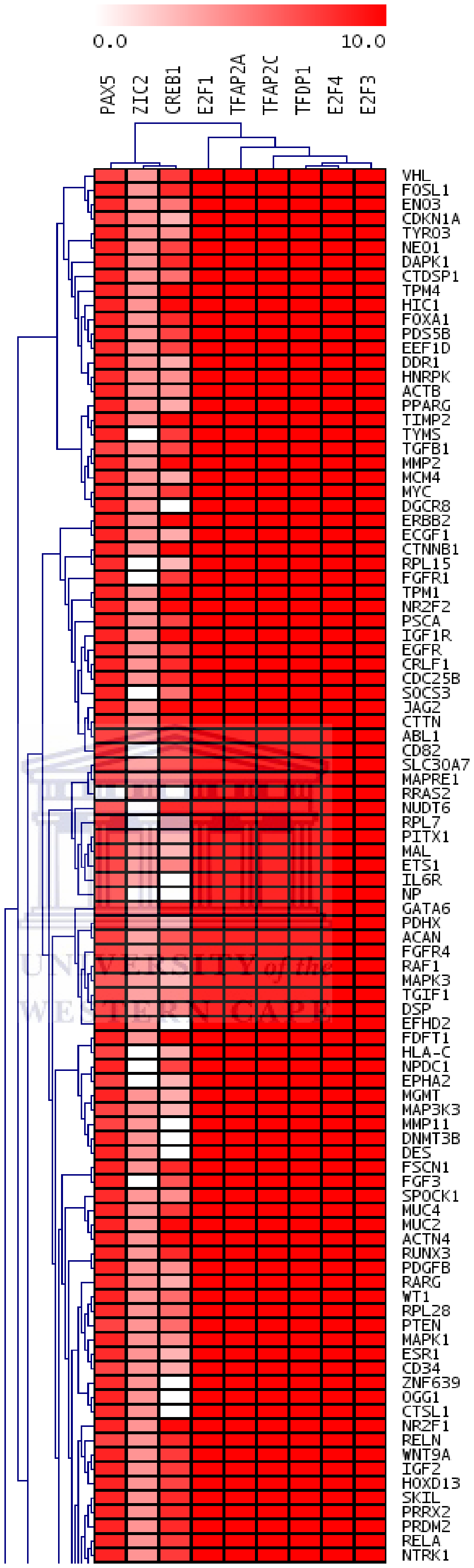


Figure 4.2 Heat map of the master regulators matrices (columns) maps to the promoter sequence of differentially expressed ESCC genes (rows). The different shades of red depict the strength of association in terms of the calculated ORI for the matrices. If there is no colour dot, then no association has been made between the matrix and gene.

Figure 4.2 continues…

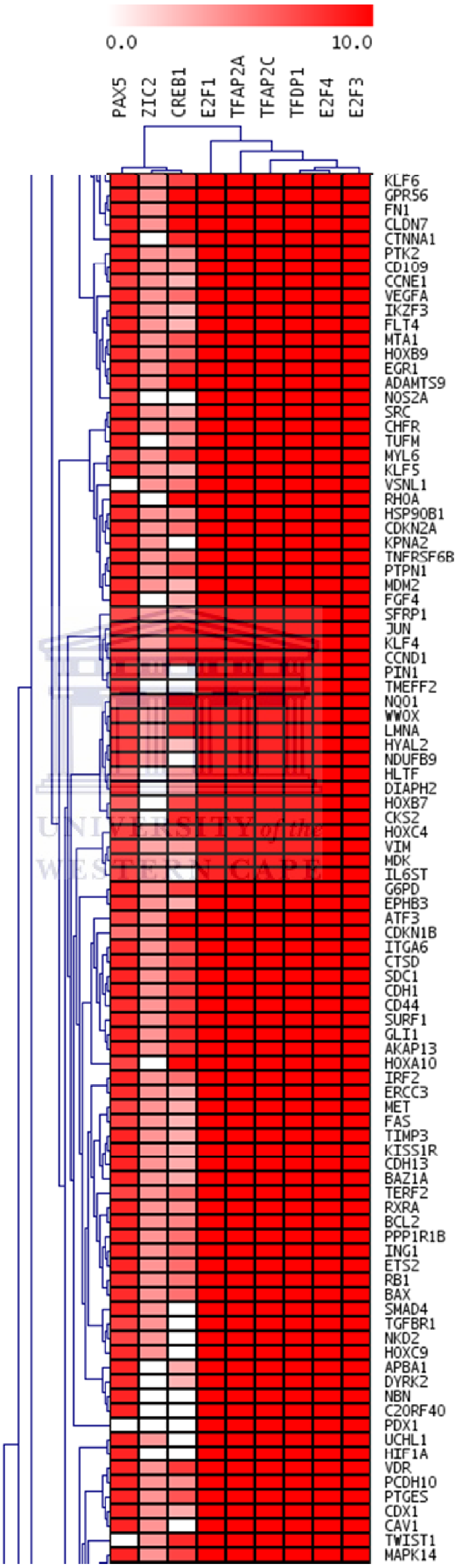


Figure 4.2 continues…

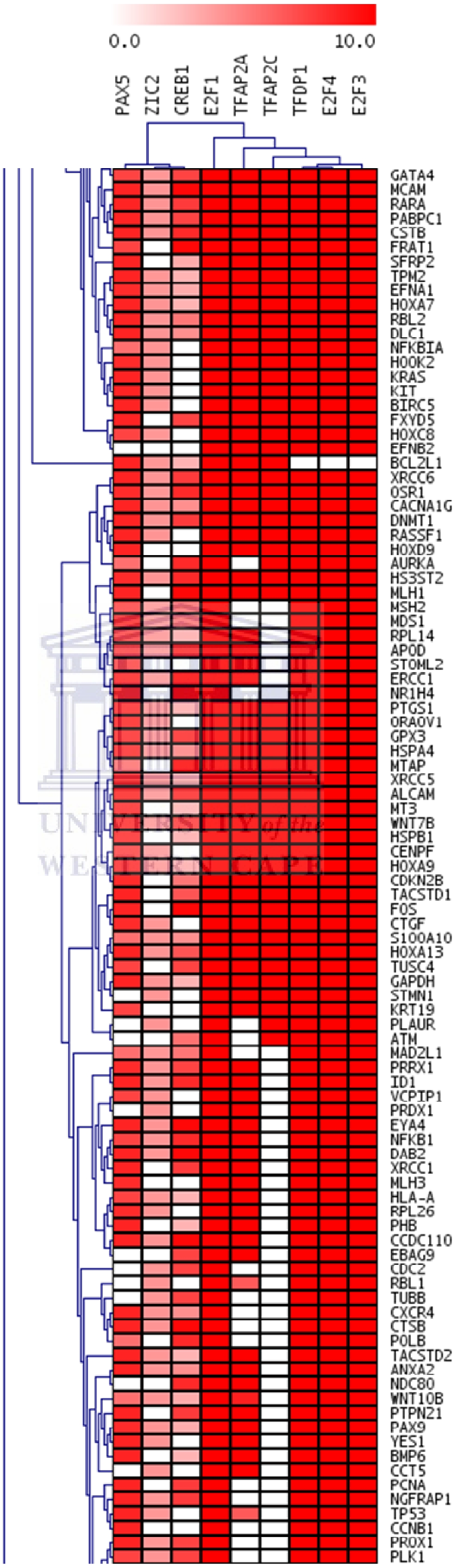
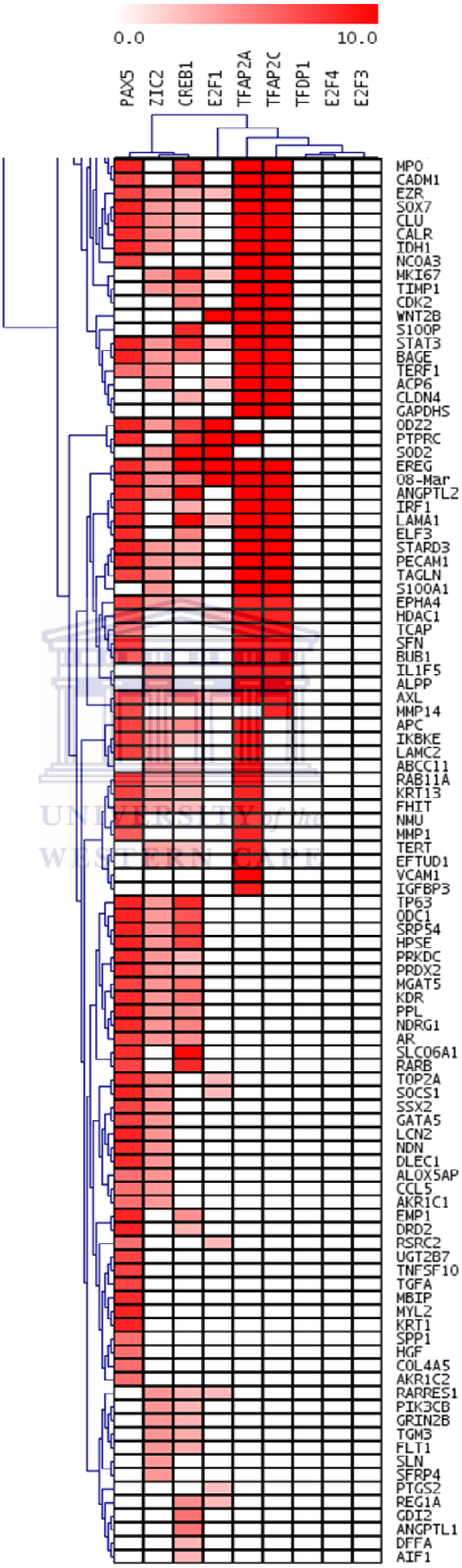


Figure 4.2 continues…



All of the differentially expressed ESCC genes had at least one master regulator matrix mapped to the promoter sequence. Figure 4.2 displays clusters of genes that are regulated by the same combinations of master regulators, 316 differentially expressed ESCC genes had 5 to 9 matrix combinations mapped to the promoter sequence. The genes that did not have 5 to 9 combinations of master regulator matrices mapped to the promoter sequences displayed mapping for the down-regulated master regulators that generally drive the developmental processes. However, the matrices for the up-regulated master regulators that are known to drive carcinogenesis display less pronounces mapping. Indicating that these genes may be up-regulated downstream in the activated pathways or the expression is driven by another combination of up-regulated TFs.

4.4.2 ALTERED TFBSS USED TO INCREASE CONFIDENCE IN THE NINE MASTER REGULATORS AS CANDIDATE DIAGNOSTIC MARKERS FOR ESCC

The presence of a SNP occurring within over-represented TFBSs for master regulators may increase confidence that such an altered TFBS (with the SNP) is a candidate biomarker for ESCC (*357*).

For this SNP analysis, all the TFBS predictions (determined in section 3.3.5) were tested for a possible overlap with one or more SNPs. A total of 4,766,707 unique SNPs were extracted from the HapMap database. The absolute chromosomal positions of each TFBS prediction had been calculated with respect to the positions of the TSSs that were used to define the promoter regions of the genes. An overlap was defined by the chromosomal position of a SNP mapping within the start and stop position of the predicted TFBSs.

Of the 4,766,707 unique SNPs, 259 SNPs mapped within the predicted TFBSs (see appendix IV). Of the 259 unique SNPs overlapping TFBSs, 70 were identified as binding domains for the putative master regulators (see Table 4.2). These results indicate that 27.03% (70/259) of the identified SNPs overlap TFBSs that are binding domains for the putative master regulators (see Figure 4.2). The 70 TFBSs with SNPs include, 35 TFBSs that bind PAX5, 9 TFBSs that bind either TFAP2A or TFAP2C (TFAP2A + 2C), 11 TFBSs that bind ZIC2, 11 TFBSs that bind CREB1 and 9 TFBSs that bind E2F family members and TFDP1 (E2F + DP) (see Figure 4.3). Thus, 71,43% of the TFBSs with SNPs are motifs for the master regulators (PAX5, TFAP2A, ZIC2 and TFAP2C) known to be down-regulated in numerous cancer types. Additionally, 50% of the TFBSs with SNPs are motifs that bind PAX5, indicating that the TRN of this master regulator of developmental processes may likely be transformed in ESCC. The up-regulated master regulators are characterized by 29% of the TFBSs with SNPs, indicating that certain underlying processes driven by these master regulators in the normal ESCC tissue are inhibited or adjusted.

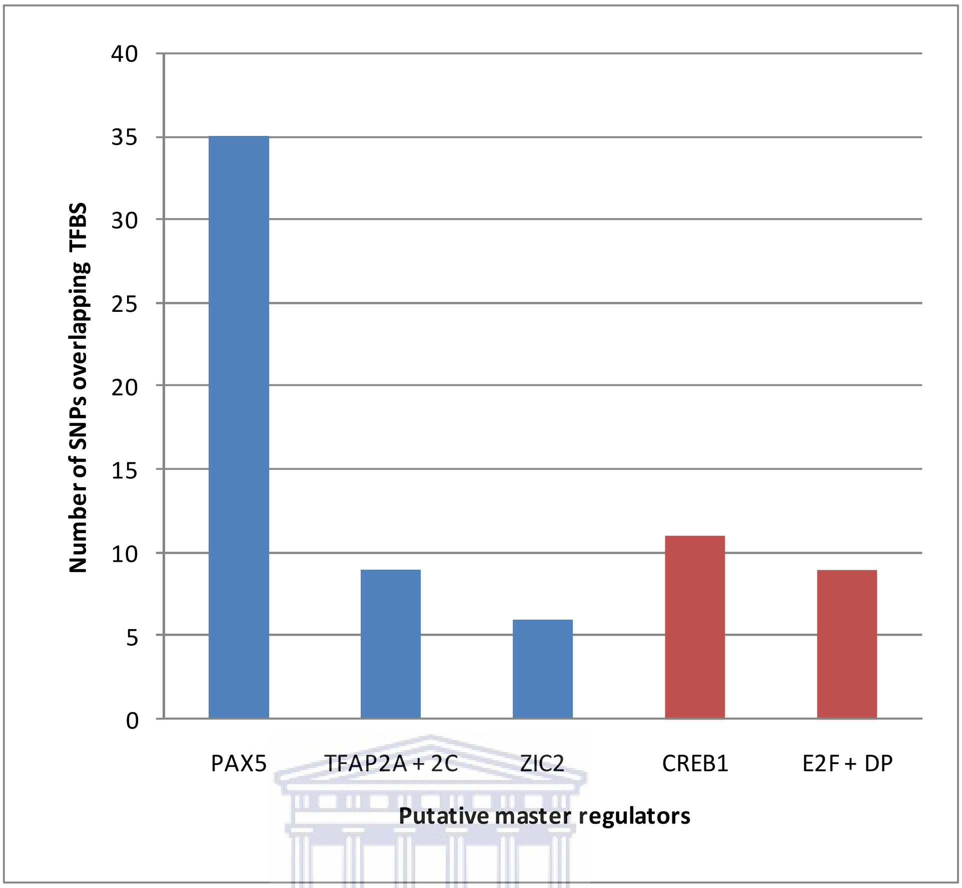


Figure 4.3: Histograms representing the distribution of TFBSs with SNPs for the master regulators. Master regulators down-regulated (blue) and up-regulated (red) in numerous cancer types.

Table 4.2: TFBSs with overlapping SNPs

TSS ID	Gene	TSSStart	Strand	SequenceStart	SequenceStop	RefSNP ID	SNP	SNP sition	Source
T18F00347AEA	7050	3439338	+	3439039	3439053	rs8090978	A/G	3439046	../snp/forward/non-redundant/genotypes_chr18_YRI_r27_nr.b36_fwd.txt
T20R033E16A6	6790	54400678	-	54401025	54401011	rs2104015	A/G	54401019	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T12F0323C141	3225	52674881	+	52674735	52674749	rs4759316	A/G	52674740	../snp/forward/non-redundant/genotypes_chr12_YRI_r27_nr.b36_fwd.txt
T08R00C94BFB	10395	13192187	-	13192559	13192545	rs7013667	A/G	13192552	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T11F0217D992	960	35117458	+	35117648	35117655	rs3751031	A/G	35117652	../snp/reverse/non-redundant/genotypes_chr11_YRI_r24_nr.b36.txt
T02R009FAB95	4953	10464149	-	10464612	10464601	rs3771134	C/T	10464611	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T22F0267A678	2547	40347256	+	40347196	40347210	rs132770	A/G	40347210	../snp/forward/non-redundant/genotypes_chr22_YRI_r27_nr.b36_fwd.txt
T15F050034C5	11214	83899589	+	83898783	83898797	rs12442369	C/T	83898796	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T11F00106482	4583	1074306	+	1073942	1073956	rs11245935	C/T	1073946	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T16F033936C4	4313	54081220	+	54081039	54081053	rs17242473	A/G	54081046	../snp/forward/non-redundant/genotypes_chr16_YRI_r27_nr.b36_fwd.txt
T11R00203EC8	3481	2113224	-	2113116	2113102	rs3213225	C/T	2113112	../snp/reverse/non-redundant/genotypes_chr11_JPT_r24_nr.b36.txt
T11R01ED6902	7490	32336130	-	32337118	32337091	rs7118294	C/T	32337097	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T03R00C05A1D	5894	12605981	-	12606883	12606856	rs7637824	A/C	12606865	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T09F0766CBF3	5742	124177395	+	124177455	124177482	rs10306133	C/T	124177465	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T01F0AD047C6	3918	181422022	+	181421916	181421943	rs2276543	A/G	181421928	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T11F00108922	4583	1083682	+	1083702	1083729	rs7944723	C/G	1083710	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T15F05C8D01D	3480	97046557	+	97045825	97045852	rs1317460	G/T	97045839	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T01R0D7ACD43	7483	226151747	-	226152520	226152493	rs10127974	A/G	226152497	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T15F03CFC853	8766	63948883	+	63948777	63948804	rs12909265	C/G	63948777	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T08R00A2115F	83595	10621279	-	10621636	10621609	rs10100209	C/T	10621622	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T18R006BB679	284217	7059065	-	7059373	7059346	rs7238600	A/G	7059358	../snp/forward/non-redundant/genotypes_chr18_YRI_r27_nr.b36_fwd.txt
T11F041F6031	595	69165105	+	69164266	69164293	rs3212860	A/T	69164288	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T07F0348F5A8	1956	55113128	+	55112813	55112840	rs9642564	A/G	55112828	../snp/forward/non-redundant/genotypes_chr7_YRI_r27_nr.b36_fwd.txt
T04F047F457E	2069	75449726	+	75449794	75449821	rs6836436	G/T	75449794	../snp/forward/non-redundant/genotypes_chr4_YRI_r27_nr.b36_fwd.txt
T20F01D672D4	1789	30831316	+	30830622	30830649	rs6058884	A/G	30830632	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt

T10R003A27CE	1316	3811278	-	3811565	3811538	rs17731	A/G	3811561	../snp/forward/non-redundant/genotypes_chr10_YRI_r27_nr.b36_fwd.txt
T09F055621E6	1514	89530854	+	89530659	89530686	rs3118869	A/C	89530683	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T19F029D1069	81	43847785	+	43847698	43847725	rs3786835	A/G	43847720	../snp/forward/non-redundant/genotypes_chr19_YRI_r27_nr.b36_fwd.txt
T20F003867FE	994	3696638	+	3695978	3696005	rs2295346	A/C	3695992	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T11R002029A6	3481	2107814	-	2108644	2108617	rs3802971	C/T	2108628	../snp/reverse/non-redundant/genotypes_chr11_YRI_r24_nr.b36.txt
T03R02F15DF3	387	49372659	-	49372747	49372720	rs11552757	C/T	49372746	../snp/reverse/non-redundant/genotypes_chr3_YRI_r24_nr.b36.txt
T09R021779C1	30968	35092929	-	35093143	35093116	rs2298310	A/C	35093117	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T20R02283F93	7052	36192147	-	36192298	36192271	rs2076378	C/T	36192272	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T17R01DC8DF5	6352	31231477	-	31232320	31232293	rs4239253	C/T	31232303	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt
T09F0554EF7C	1612	89452412	+	89452459	89452486	rs1927975	C/T	89452483	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T20R02284781	7052	36194177	-	36194454	36194427	rs2076386	A/G	36194428	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T16R004A5922	5493	4872482	-	4872684	4872657	rs1049223	A/G	4872660	../snp/forward/non-redundant/genotypes_chr16_YRI_r27_nr.b36_fwd.txt
T11R00203EC8	3481	2113224	-	2113740	2113713	rs3213219	C/T	2113726	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T08F08915B2E	8000	143743790	+	143743417	143743444	rs754957	A/G	143743418	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T14F04759734	2353	74815284	+	74815184	74815211	rs4645852	A/C	74815200	../snp/forward/non-redundant/genotypes_chr14_YRI_r27_nr.b36_fwd.txt
T0XR09252298	2539	153428632	-	153429307	153429280	rs5986992	A/C	153429301	../snp/forward/non-redundant/genotypes_chrX_YRI_r27_nr.b36_fwd.txt
T02R0CDFD2D4	2335	215995092	-	215995746	215995719	rs16854041	A/G	215995742	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T05R0ABA043F	2324	179962943	-	179963756	179963729	rs11739750	C/T	179963737	../snp/forward/non-redundant/genotypes_chr5_YRI_r27_nr.b36_fwd.txt
T03R03DA74A0	56999	64648352	-	64648857	64648830	rs6789298	C/T	64648855	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T03R0A5AD10A	8743	173723914	-	173724460	173724433	rs3136586	A/G	173724453	../snp/reverse/non-redundant/genotypes_chr3_YRI_r24_nr.b36.txt
T14R05401820	11099	88086560	-	88086772	88086745	rs891751	C/T	88086770	../snp/reverse/non-redundant/genotypes_chr14_YRI_r24_nr.b36.txt
T02R0900C475	390	151045237	-	151045093	151045082	rs1528429	C/T	151045088	../snp/reverse/non-redundant/genotypes_chr2_YRI_r24_nr.b36.txt
T03F01844198	5915	25444760	+	25443840	25443851	rs9827454	A/G	25443850	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T15F03A4E791	7168	61138833	+	61138535	61138546	rs4238372	A/G	61138541	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T05F09F51304	57451	167056132	+	167056301	167056308	rs898172	C/T	167056303	../snp/reverse/non-redundant/genotypes_chr5_YRI_r24_nr.b36.txt
T03F00950821	4968	9766945	+	9766950	9766958	rs1801126	G/T	9766953	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T08F07AD9FF8	4609	128819192	+	128818917	128818925	rs4645954	A/G	128818920	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T16F035A03F9	9289	56230905	+	56230961	56230969	rs3760061	A/G	56230961	../snp/forward/non-redundant/genotypes_chr16_YRI_r27_nr.b36_fwd.txt

T05F08E8D7A0	1044	149477280	+	149477363	149477371	rs246391	C/T	149477370	../snp/forward/non-redundant/genotypes_chr5_YRI_r27_nr.b36_fwd.txt
T06R0989C054	6648	160022612	-	160022705	160022697	rs5746141	A/G	160022700	../snp/reverse/non-redundant/genotypes_chr6_YRI_r24_nr.b36.txt
T03R00C0B214	5894	12628500	-	12628582	12628574	rs3730268	C/T	12628578	../snp/reverse/non-redundant/genotypes_chr3_YRI_r24_nr.b36.txt
T06F01D86475	780	30958709	+	30958556	30958561	rs1264327	C/T	30958561	../snp/reverse/non-redundant/genotypes_chr6_YRI_r24_nr.b36.txt
T05F009D37D1	22948	10303441	+	10302790	10302795	rs2548551	G/T	10302790	../snp/forward/non-redundant/genotypes_chr5_YRI_r27_nr.b36_fwd.txt
T09R051CF96F	3190	85784943	-	85785554	85785549	rs296890	G/T	85785551	../snp/reverse/non-redundant/genotypes_chr9_YRI_r24_nr.b36.txt
T08F0285ADA4	5423	42315172	+	42315119	42315132	rs2307155	A/C	42315132	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T09R051CF6BF	3190	85784255	-	85784901	85784886	rs296887	C/T	85784890	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T10F07D2423C	4255	131220028	+	131219103	131219114	rs7900620	G/T	131219110	../snp/forward/non-redundant/genotypes_chr10_JPT_r27_nr.b36_fwd.txt
T03F00950B65	4968	9767781	+	9767402	9767417	rs3219001	C/T	9767417	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T0XR0924A9F6	2539	153397750	-	153398734	153398719	rs5986989	A/G	153398725	../snp/forward/non-redundant/genotypes_chrX_YRI_r27_nr.b36_fwd.txt
T11R00203166	3481	2109798	-	2110776	2110761	rs3741214	A/G	2110767	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T02R0CE00838	2335	216008760	-	216008730	216008715	rs1250259	A/T	216008727	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T01R08ADD126	51205	145608998	-	145609327	145609312	rs2353522	A/G	145609315	../snp/reverse/non-redundant/genotypes_chr1_YRI_r24_nr.b36.txt
T11R01EE8567	7490	32408935	-	32409510	32409496	rs10742277	C/G	32409501	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T02F06522C1D	84417	106048541	+	106048030	106048044	rs755651	C/T	106048032	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T20R01C6508C	598	29773964	-	29774103	29774089	rs6087772	C/T	29774094	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt

These master regulators altered TFBSs may contribute to the TRN that characterizes ESCC. However, it must be acknowledged that eventhough the control regions of these master regulators were not in the scope of this study, the differential expression of the master regulators in numerous cancer types indicate that characterizing these control regions are also crucial to understanding the altered TRN of ESCC. To date the differential expression of these master regulators have been attributed to chromosomal translocation, hypermethylation and amplification (see Table 4.1), but none have been attributed to SNPs.

4.5 CONCLUSION

This study pinpoints key master regulators for ESCC and provide the possible stimulus of the observed dysregulated expression, as the TFBSs that binds the master regulators have overlapping SNPs. These putative TFBSs with SNPs provide a stepping stone to identify key changes in the TRN, as TFBSs with SNPs will not only affect the binding capacity of the TF but may even abolish its binding entirely and/or allow for the binding of a different TF thereby changing the TRN profile. Furthermore, the presence of a SNP occurring within overrepresented TFBSs further increase confidence that such an altered TFBS (with the SNP) may be a successful biomarker for ESCC.

CHAPTER 5

CONCLUSION, PITFALLS AND PROMISES

The overall objective of this project was to develop a novel comprehensive ESCC database as an integrated knowledge base aimed at representing a gateway to ESCC related data, in an attempt to enhance the scientific research process that is currently hampered by a wealth of scattered research data. More importantly, two illustrations are provided of how the biocurated genes in the database may represent a reliable starting point for divulging transcriptional regulation, diagnostic markers and the biology related to ESCC.

The Dragon Database of Genes Implicated in Esophageal Cancer (DDEC) was successfully developed and is freely accessible for academic and non-profit users at http://apps.sanbi.ac.za/ddec/. The ESCC genes were further used to develop a method that identifies combinations of TFBSs that characterize estrogen responsive genes. This method allowed for the prediction of 47 novel estrogen responsive genes, of which two were verified in current 2009 publications. The ESCC genes were additionally used to identify the master regulators (PAX5, TFAP2A, ZIC2, CREB1, E2F1, TFAP2C, TFDP1, E2F3 and E2F4) as candidate diagnostic markers for ESCC.

Although this study was adequate to develop a method to identify estrogen response genes and putative candidate diagnostic markers, shortcomings make the straight forward analysis difficult, namely: 1/ the limited number of differentially expressed

ESCC genes; 2/ undefined promoter regions (*358*); and 3/ non-exhaustive sets of TFBS and ERE models. Currently, promoter regions are identified by determination of the transcription start site (TSS). The core promoter region is defined as the region proximal to the TSS and is required for normal gene transcription regulation. The exact span of nucleotides making up the core promoter varies from study to study, and from gene to gene. There is no agreed upon definition regarding core promoter size. However, many authors report that trans-regulators tend to target cis-regulatory motifs within the first 2Mb upstream of the TSS (*359*). In this study, the promoter sets were defined to be -1200nt upstream and +200nt downstream of identifiable TSSs. The detection of EREs and TFBSs are dependent on the limited number of matrices derived from known experimental data. Thus, our ability to detect complete lists of EREs and TFBSs is restricted by a non-exhaustive set of models (*360;361*). Additionally, the analysis of the SNPs are limited as the overlapping SNPs could be identified but the SNP could not be verified as the promoter sequences extracted were not sequences from ESCC.

In conclusion this study has demonstrated the prediction of novel estrogen response gene and candidate diagnostic markers through the exploitation of the ESCC genes and provides a method for determining hormone responsive genes that could be applicable to other hormone responsive cancers. Furthermore, the circumstantial and imperical evidence presented here suggests that these findings may present a reliable starting point for the development of screens for novel therapeutic targets.

APPENDIX I: Esophageal cancer genes with ERE predictions

Gene Symbol	Gene ID	TSS	CTSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Tags	RepTags	Distance:		Sequence	Pattern
ABL1	25	T09F07E6FAFC	C09F07E6FAFC	chr9	+	132578044	132576844	132578244	13	3	1047	634	TT-GGTCA-GGC-TGGTC-TC	known
ABL1	25	T09F07E6FBCC	C09F07E6FBCC	chr9	+	132578252	132577052	132578452	37	9	865	426	TT-GGTCA-GGC-TGGTC-TC	known
ABL1	25	T09F07E6FBFD	C09F07E6FBFD	chr9	+	132578301	132577101	132578501	9	6	809	377	TT-GGTCA-GGC-TGGTC-TC	known
ACTN4	81	T19F029CFFE2	C19F029CFFE2	chr19	+	43843554	43842354	43843754	7	3	-13383	726	TT-GGTCA-GGC-TGGCC-TT	new
ACTN4	81	T19F029D9BDC	C19F029D9BDC	chr19	+	43883484	43882284	43883684	39	14	-53316	1058	GG-GGCCA-GAG-TGGCC-TG	new
ACTN4	81	T19F029DBE32	C19F029DBE32	chr19	+	43892274	43891074	43892474	10	8	-62100	1179	AA-GGGCA-AGG-TGGCT-GA	new
ACTN4	81	T19F029DD0C0	C19F029DD0C0	chr19	+	43897024	43895824	43897224	6	3	-66842	1268	GA-GGTCA-CTG-TGACT-CT	new
ANXA2	302	T15R037B8B75	C15R037B8B75	chr15	-	58428277	58428077	58429477	7	7	-1616	928	AG-TGTCA-GCC-TGCCC-TG	new
APC	324	T05F06B00439	C05F06B00439	chr5	+	112198713	112197513	112198913	5	3	-97226	394	CG-GCTCA-CTG-CGACC-TC	new
ATM	472	T11F0669D4A8	C11F0669D4A8	chr11	+	107599016	107597816	107599216	55	3	-182	562	AA-GGTCC-TTC-TGTCC-AG	new
BAX	581	T19F033A5DEB	C19F033A5DEB	chr19	+	54156779	54155579	54156979	9	7	-6850	57	TT-GGTCA-GGC-TGGTC-TT	new
BCL2	596	T18R0383753E	C18R0383753E	chr18	-	58946878	58946678	58948078	33	11	-5298	1240	GA-AGTCA-ACA-TGCCT-GC	new
BCL2L1	598	T20R01C5892A	C20R01C5892A	chr20	-	29722922	29722722	29724122	8	5	-6980	380	TT-GGTCA-GGC-TGGTC-TT	new
BCL2L1	598	T20R01C5892A	C20R01C5892A	chr20	-	29722922	29722722	29724122	8	5	-6980	458	TG-AGTCA-TTG-TGCCC-GG	new
BMP6	654	T06F007686F4	C06F007686F4	chr6	+	7767796	7766596	7767996	5	5	-95787	686	TT-GGTCT-TGT-TGCCC-AG	new
BMP6	654	T06F0076E486	C06F0076E486	chr6	+	7791750	7790550	7791950	7	3	-119739	40	AA-GGTGG-GGC-TGGCC-CC	new
CALR	811	T19F00C5082D	C19F00C5082D	chr19	+	12912685	12911485	12912885	41	6	-2210	1303	CA-GGTGA-GTC-TGACT-CA	new
RUNX3	864	T01R017FF283	C01R017FF283	chr1	-	25162371	25162171	25163571	6	6	-63707	132	GT-GGTCT-CTG-TGGTC-TC	new

Gene Symbol	Gene ID	TSS	CTSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Tags	RepTags	Distance:		Sequence	Pattern
RUNX3	864	T01R017FF283	C01R017FF283	chr1	-	25162371	25162171	25163571	6	6	-63707	1117	TG-GGTCT-GCC-TGCCT-GC	new
CDC2	983	T10F03B498BF	C10F03B498BF	chr10	+	62167231	62166031	62167431	7	6	38459	1048	CA-AATCA-CCA-TGACC-TA	new
CDC25B	994	T20F00383158	C20F00383158	chr20	+	3682648	3681448	3682848	6	4	41753	1134	AA-GGCCA-CTC-CACCC-CC	new
CDC25B	994	T20F0038F8EC	C20F0038F8EC	chr20	+	3733740	3732540	3733940	6	6	-9339	113	CT-GGTCC-ATA-TGACT-GC	new
CDH1	999	T16F04039AE1	C16F04039AE1	chr16	+	67345121	67343921	67345321	38	23	-16425	1324	CA-GGGCA-CGG-TGGCC-GA	new
CDKN1A	1026	T06F0230D425	C06F0230D425	chr6	+	36754469	36753269	36754669	774	483	3	605	AA-AATCA-TTC-TGGCC-TC	new
CDX1	1044	T05F08E92640	C05F08E92640	chr5	+	149497408	149496208	149497608	139	86	29145	442	CA-GGGCA-GTG-TGCCT-GG	new
CLU	1191	T08R01A3FB63	C08R01A3FB63	chr8	-	27523939	27523739	27525139	22	8	-13551	1357	TG-GGTCT-GGG-TGCCT-GT	new
CLDN4	1364	T07F04581C3F	C07F04581C3F	chr7	+	72883263	72882063	72883463	358	75	-132	903	TG-GGTCT-GGT-TGCCC-CC	new
CSTB	1476	T21R029FB389	C21R029FB389	chr21	-	44020617	44020417	44021817	154	11	-2323	914	CC-GGTCC-GTG-TGACT-CG	new
CTNNA1	1495	T05F083B7FCD	C05F083B7FCD	chr5	+	138117069	138115869	138117269	504	56	22	185	AA-GGTGG-CTC-TAGCC-TG	new
CTNNB1	1499	T03F0274EA18	C03F0274EA18	chr3	+	41216536	41215336	41216736	6	3	-532	1029	CC-GGTCG-CGC-TGCCT-CT	new
CTSD	1509	T11R001A5981	C11R001A5981	chr11	-	1726849	1726649	1728049	5	3	3730	252	CG-GGACA-CTT-TGCCC-TG	new
CTSD	1509	T11R001A59B6	C11R001A59B6	chr11	-	1726902	1726702	1728102	5	3	3682	305	CG-GGACA-CTT-TGCCC-TG	new
CTSD	1509	T11R001A6B3F	C11R001A6B3F	chr11	-	1731391	1731191	1732591	9	3	-810	11	GG-GGTCA-GAC-CAGCC-CT	new
DAB2	1601	T05R0259A87B	C05R0259A87B	chr5	-	39430267	39430067	39431467	7	3	-21747	1320	AT-GGTCA-GCC-CGACC-AA	new
NQO1	1728	T16R0412730C	C16R0412730C	chr16	-	68317964	68317764	68319164	769	272	-17134	736	GC-AGTCA-CAG-TGACT-CA	known
DNMT1	1786	T19R009AB2E0	C19R009AB2E0	chr19	-	10138336	10138136	10139536	13	4	-33267	445	AA-GATCA-GCC-TGGCC-AA	new
DSP	1832	T06F0072B44E	C06F0072B44E	chr6	+	7517262	7516062	7517462	7	4	-30393	1156	GA-GGGCA-GGA-TGACT-CT	new
EFNB2	1948	T13R06511B39	C13R06511B39	chr13	-	105978681	105978481	105979881	6	4	-38567	400	CA-GGACA-CAT-TGTCC-CT	new
EFNB2	1948	T13R06512E8E	C13R06512E8E	chr13	-	105983630	105983430	105984830	13	10	-43518	346	GC-GGTCT-CGA-TGGCC-GA	new

Gene Symbol	Gene ID	TSS	CTSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Tags	RepTags	Distance:		Sequence	Pattern
EGFR	1956	T07F0347ACD5	C07F0347ACD5	chr7	+	55028949	55027749	55029149	5	5	25270	776	GG-GGTCA-GCC-TGCCC-TA	new
EGFR	1956	T07F03480EB7	C07F03480EB7	chr7	+	55054007	55052807	55054207	21	5	215	853	GC-GGGCA-CGG-CGACC-TC	new
EGFR	1956	T07F03480EEA	C07F03480EEA	chr7	+	55054058	55052858	55054258	5	4	161	802	GC-GGGCA-CGG-CGACC-TC	new
EGFR	1956	T07F0347ACD5	C07F0347ACD5	chr7	+	55028949	55027749	55029149	5	5	25270	776	GG-GGTCA-GCC-TGCCC-TA	new
EGFR	1956	T07F03480EB7	C07F03480EB7	chr7	+	55054007	55052807	55054207	21	5	215	853	GC-GGGCA-CGG-CGACC-TC	new
EGFR	1956	T07F03480EEA	C07F03480EEA	chr7	+	55054058	55052858	55054258	5	4	161	802	GC-GGGCA-CGG-CGACC-TC	new
EGFR	1956	T07F03480F7D	C07F03480F7D	chr7	+	55054205	55053005	55054405	589	65	66	655	GC-GGGCA-CGG-CGACC-TC	new
EGFR	1956	T07F0348102B	C07F0348102B	chr7	+	55054379	55053179	55054579	7	4	-143	481	GC-GGGCA-CGG-CGACC-TC	new
EGFR	1956	T07F03481051	C07F03481051	chr7	+	55054417	55053217	55054617	58	10	-192	443	GC-GGGCA-CGG-CGACC-TC	new
EGFR	1956	T07F03481158	C07F03481158	chr7	+	55054680	55053480	55054880	15	3	-443	180	GC-GGGCA-CGG-CGACC-TC	new
EPHA2	1969	T01R00F98B55	C01R00F98B55	chr1	-	16354133	16353933	16355333	9	3	-30695	612	TT-GGTCG-CAC-GGGCC-TC	new
ENO3	2023	T01R00870718	C01R00870718	chr1	-	8849176	8848976	8850376	5	3	-5507	205	CA-GGTGA-CAA-TAGCC-TG	new
ERCC2	2068	T19R0303913C	C19R0303913C	chr19	-	50565436	50565236	50566636	11	5	-18733	396	TT-GGTCA-GGC-TGGTC-TT	new
ETS2	2114	T21F0254E0A0	C21F0254E0A0	chr21	+	39116960	39115760	39117160	11	5	-17212	923	GA-GGTCG-CCG-GGACC-AC	new
ETS2	2114	T21F0254E128	C21F0254E128	chr21	+	39117096	39115896	39117296	24	7	-17368	787	GA-GGTCG-CCG-GGACC-AC	new
ETS2	2114	T21F0254E24F	C21F0254E24F	chr21	+	39117391	39116191	39117591	20	8	-17655	492	GA-GGTCG-CCG-GGACC-AC	new
ETS2	2114	T21F0254E2D4	C21F0254E2D4	chr21	+	39117524	39116324	39117724	9	3	-17804	359	GA-GGTCG-CCG-GGACC-AC	new
MDS1	2122	T03R0A26D0B7	C03R0A26D0B7	chr3	-	170315959	170315759	170317159	6	4	-30695	182	TG-AGTCA-TGC-CAACC-CG	new
MDS1	2122	T03R0A26D50C	C03R0A26D50C	chr3	-	170317068	170316868	170318268	5	3	-31787	266	AA-GGTTA-GCA-CAACC-AC	new
MDS1	2122	T03R0A26D50C	C03R0A26D50C	chr3	-	170317068	170316868	170318268	5	3	-31787	1291	TG-AGTCA-TGC-CAACC-CG	new
FGF4	2249	T11R0421678A	C11R0421678A	chr11	-	69298058	69297858	69299258	7	3	-1063	265	GC-GGTCC-TGC-TGGCC-TT	new

Gene Symbol	Gene ID	TSS	CTSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Tags	RepTags	Distance:		Sequence	Pattern
FGFR4	2264	T05F0A843489	C05F0A843489	chr5	+	176436361	176435161	176436561	6	6	10166	32	TA-GGTCA-CTG-CAGCC-TC	new
FHIT	2272	T03R03A24F58	C03R03A24F58	chr3	-	60968792	60968592	60969992	66	39	-1258696	1270	CA-GGTGA-AAT-TAACC-TG	new
FLT1	2321	T13R01AA206D	C13R01AA206D	chr13	-	27926637	27926437	27927837	32	24	-152830	190	TA-GGTCT-CTA-TGATC-CA	new
FN1	2335	T02R0CDFC7B3	C02R0CDFC7B3	chr2	-	215992243	215992043	215993443	364	103	-58795	322	TG-GGTCA-GGC-TGGTC-TC	new
FN1	2335	T02R0CDFDEBD	C02R0CDFDEBD	chr2	-	215998141	215997941	215999341	23	5	-64646	357	AA-GGTGG-AGG-TGACC-TC	new
GAPDH	2597	T12F0063651F	C12F0063651F	chr12	+	6513951	6512751	6514151	93	20	66	456	CA-GGACA-TCG-TGACC-TT	new
GAPDH	2597	T12F00636637	C12F00636637	chr12	+	6514231	6513031	6514431	21	3	-297	176	CA-GGACA-TCG-TGACC-TT	new
GATA4	2626	T08F00B0A470	C08F00B0A470	chr8	+	11576432	11575232	11576632	5	3	22731	624	AT-GGTCA-CTG-TGTCC-CC	new
GATA4	2626	T08F00B0D09C	C08F00B0D09C	chr8	+	11587740	11586540	11587940	60	55	11423	148	CA-GGCCA-CAG-CAACC-TG	new
GATA6	2627	T18F0112BFD1	C18F0112BFD1	chr18	+	18005969	18004769	18006169	12	3	-2526	891	CC-GGTCT-ACG-TGCCC-AC	new
SFN	2810	T01F019CD661	C01F019CD661	chr1	+	27055713	27054513	27055913	10	6	6507	504	GA-GGCCA-TCC-TGTCC-CC	new
GPX3	2878	T05F08F6BD67	C05F08F6BD67	chr5	+	150388071	150386871	150388271	9	4	-7878	573	AG-GGTCT-CAT-TGGCC-CA	new
GRB7	2886	T17F02184FDB	C17F02184FDB	chr17	+	35147739	35146539	35147939	49	3	22	958	CA-GGTGG-TCG-GGACT-CT	new
GRB7	2886	T17F02185167	C17F02185167	chr17	+	35148135	35146935	35148335	118	21	-364	562	CA-GGTGG-TCG-GGACT-CT	new
HIC1	3090	T17F001D11FB	C17F001D11FB	chr17	+	1905147	1903947	1905347	215	93	-106	1235	CA-GGACA-GAC-CGACC-GG	new
HIC1	3090	T17F001D1223	C17F001D1223	chr17	+	1905187	1903987	1905387	8	4	-169	1195	CA-GGACA-GAC-CGACC-GG	new
HLA-C	3106	T06R01DF9F64	C06R01DF9F64	chr6	-	31432548	31432348	31433748	10	5	-2899	838	CT-GGTCA-TGG-CGCCC-CG	new
HOXC4	3221	T12F03246972	C12F03246972	chr12	+	52717938	52716738	52718138	63	11	-20980	282	CA-GGGCA-CAC-CAGCC-CC	new
HOXC4	3221	T12F03246972	C12F03246972	chr12	+	52717938	52716738	52718138	63	11	-20980	857	AA-GATCA-ACA-CGACC-TC	new
HOXC4	3221	T12F032469A6	C12F032469A6	chr12	+	52717990	52716790	52718190	23	9	-21081	230	CA-GGGCA-CAC-CAGCC-CC	new
HOXC4	3221	T12F032469A6	C12F032469A6	chr12	+	52717990	52716790	52718190	23	9	-21081	805	AA-GATCA-ACA-CGACC-TC	new

Gene Symbol	Gene ID	TSS	CTSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Tags	RepTags	Distance:		Sequence	Pattern
HOXC9	3225	T12F0323C141	C12F0323C141	chr12	+	52674881	52673681	52675081	66	5	5268	1352	TG-GGTCT-CTC-TGCCT-GA	new
HSPA4	3308	T05F07E4C83C	C05F07E4C83C	chr5	+	132433980	132432780	132434180	11	6	-18407	266	TT-GGTCA-CGC-TGGTC-TT	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	638	GG-GGTCT-TTC-TGCCC-AG	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	1110	GA-GGCCA-GCG-CAGCC-CC	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	325	CA-GGGCA-GGG-TGGCC-AG	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	406	AA-GGTCC-ACC-TGCCT-CC	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	1050	TT-GGTCA-ATG-CAGCC-CT	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	1037	TT-GGTCA-GGA-CAGCC-CA	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	744	TT-AGTCA-GGC-TGATC-TC	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	1235	TG-AGTCA-GGC-TGCCC-CG	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	897	GA-CGTCA-AAC-TGCCC-TC	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	520	TT-GGTCA-GGC-TGGTC-TC	known
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	37	AA-AGCCA-AAG-TGACC-AA	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	1291	CG-GGTCA-TGG-CGCCC-CG	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	1033	AG-AGTCA-CCC-TGGCC-CG	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	794	AA-GGTCA-CGG-TGGCC-AC	known
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	850	CA-CGTCA-CTG-TGACC-TA	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	1287	CT-GGTCC-TCC-TGACT-GG	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	58	AG-GGGCA-TGT-TGACC-CA	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	332	GA-GGACA-GAA-TGCCC-TC	new
ID1	3397	T20F01C48AB7	C20F01C48AB7	chr20	+	29657783	29656583	29657983	13	3	-1029	906	GC-AGTCA-CCG-TGGTC-CT	new

Gene Symbol	Gene ID	TSS	CTSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Tags	RepTags	Distance:		Sequence	Pattern
IGF1R	3480	T15F05CC041F	C15F05CC041F	chr15	+	97256479	97255279	97256679	5	3	-246191	233	CC-AGTCA-CAG-TGGCT-CA	new
IGF1R	3480	T15F05CC12C5	C15F05CC12C5	chr15	+	97260229	97259029	97260429	9	9	-249941	1284	TT-GGTCT-CCT-TGTCC-TT	new
IGF1R	3480	T15F05CC17A8	C15F05CC17A8	chr15	+	97261480	97260280	97261680	5	4	-251192	33	TT-GGTCT-CCT-TGTCC-TT	new
IGF2	3481	T11R00203063	C11R00203063	chr11	-	2109539	2109339	2110739	21	4	-2579	477	TG-AGTCA-AAT-TGGCC-TG	new
IGF2	3481	T11R002030C6	C11R002030C6	chr11	-	2109638	2109438	2110838	40	14	-2692	576	TG-AGTCA-AAT-TGGCC-TG	new
IGF2	3481	T11R0020310A	C11R0020310A	chr11	-	2109706	2109506	2110906	31	7	-2736	644	TG-AGTCA-AAT-TGGCC-TG	new
IGF2	3481	T11R0020313A	C11R0020313A	chr11	-	2109754	2109554	2110954	25	13	-2808	692	TG-AGTCA-AAT-TGGCC-TG	new
IGF2	3481	T11R00203166	C11R00203166	chr11	-	2109798	2109598	2110998	17	13	-2852	736	TG-AGTCA-AAT-TGGCC-TG	new
IGF2	3481	T11R00203190	C11R00203190	chr11	-	2109840	2109640	2111040	20	4	-2873	778	TG-AGTCA-AAT-TGGCC-TG	new
IGF2	3481	T11R00203254	C11R00203254	chr11	-	2110036	2109836	2111236	2060	468	-2918	974	TG-AGTCA-AAT-TGGCC-TG	new
IGF2	3481	T11R002032EE	C11R002032EE	chr11	-	2110190	2109990	2111390	338	76	-3183	1128	TG-AGTCA-AAT-TGGCC-TG	new
IL6R	3570	T01F09192B5F	C01F09192B5F	chr1	+	152644447	152643247	152644647	23	5	-102	654	CA-GGTCT-ATC-TGACT-GC	new
IL6R	3570	T01F09192BAD	C01F09192BAD	chr1	+	152644525	152643325	152644725	42	3	-176	576	CA-GGTCT-ATC-TGACT-GC	new
PDX1	3651	T13F01A208FB	C13F01A208FB	chr13	+	27396347	27395147	27396547	8	3	-4167	423	CT-GGCCA-ACG-TGACC-CC	new
ITGA6	3655	T02F0A50A70A	C02F0A50A70A	chr2	+	173057802	173056602	173058002	7	3	-57186	387	TT-GGTCA-GGC-TGGTC-TC	known
ITGA6	3655	T02F0A50A76A	C02F0A50A76A	chr2	+	173057898	173056698	173058098	32	8	-57258	291	TT-GGTCA-GGC-TGGTC-TC	known
ITGA6	3655	T02F0A50A87B	C02F0A50A87B	chr2	+	173058171	173056971	173058371	18	8	-57512	18	TT-GGTCA-GGC-TGGTC-TC	known
JAG2	3714	T14R063C89CC	C14R063C89CC	chr14	-	104630732	104630532	104631932	9	7	48407	230	CA-GCTCA-GCC-TGCCC-TC	new
JAG2	3714	T14R063C89CC	C14R063C89CC	chr14	-	104630732	104630532	104631932	9	7	48407	793	CC-AGTCA-AGC-CACCC-CA	new
CD82	3732	T11F02A6FE32	C11F02A6FE32	chr11	+	44498482	44497282	44498682	17	9	45235	116	GT-AGTCA-GCT-TGGCC-TG	new
KDR	3791	T04R0350F74B	C04R0350F74B	chr4	-	55637835	55637635	55639035	6	6	1590	1294	TG-GGTCA-GGC-TGGCC-TT	new

Gene Symbol	Gene ID	TSS	CTSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Tags	RepTags	Distance:		Sequence	Pattern
KIT	3815	T04F034A92C1	C04F034A92C1	chr4	+	55218881	55217681	55219081	378	97	112	547	GA-GGCCA-CGC-CGACC-AA	new
KRT13	3860	T17R02333152	C17R02333152	chr17	-	36909394	36909194	36910594	58	58	1382	813	GG-AGTCA-TCC-TGGCC-TC	new
LAMC2	3918	T01F0AD047C6	C01F0AD047C6	chr1	+	181422022	181420822	181422222	190	20	79	363	AA-GGTCA-GCC-TGCCC-AA	new
LCN2	3934	T09F07BEE741	C09F07BEE741	chr9	+	129951553	129950353	129951753	37	10	-13	380	AT-GGTCT-CAG-TGACC-TC	new
LOXL2	4017	T08R016224CB	C08R016224CB	chr8	-	23209163	23208963	23210363	6	5	1211	1128	CA-GGGGA-TGC-TGACC-CA	new
LOXL2	4017	T08R01628357	C08R01628357	chr8	-	23233367	23233167	23234567	15	8	-22988	1169	GA-GGCCA-TGG-TGGTC-TG	new
LOXL2	4017	T08R01628357	C08R01628357	chr8	-	23233367	23233167	23234567	15	8	-22988	1335	CA-GAGCA-AGT-TGCCC-TG	new
LOXL2	4017	T08R016283A7	C08R016283A7	chr8	-	23233447	23233247	23234647	5	3	-23069	1249	GA-GGCCA-TGG-TGGTC-TG	new
LOXL2	4017	T08R016283C8	C08R016283C8	chr8	-	23233480	23233280	23234680	6	4	-23106	1282	GA-GGCCA-TGG-TGGTC-TG	new
LOXL2	4017	T08R0162B82E	C08R0162B82E	chr8	-	23246894	23246694	23248094	15	4	-36520	969	GG-CGTCA-TGA-TGCCC-TG	new
LOXL2	4017	T08R0162B8AD	C08R0162B8AD	chr8	-	23247021	23246821	23248221	10	4	-36626	1096	GG-CGTCA-TGA-TGCCC-TG	new
TACSTD1	4072	T02F02D40719	C02F02D40719	chr2	+	47449881	47448681	47450081	64	7	118	546	TT-GGTCA-GGC-TGGTC-TC	known
TACSTD1	4072	T02F02D40762	C02F02D40762	chr2	+	47449954	47448754	47450154	516	162	21	473	TT-GGTCA-GGC-TGGTC-TC	known
MCM4	4173	T08F02EC40A7	C08F02EC40A7	chr8	+	49037479	49036279	49037679	12	5	-1431	263	TG-AGTCA-TAA-TGCCC-CA	new
MDM2	4193	T12F0405CD2F	C12F0405CD2F	chr12	+	67489071	67487871	67489271	15	5	-802	986	GA-GGTCC-GGA-TGATC-GC	new
MET	4233	T07F06EC9654	C07F06EC9654	chr7	+	116168276	116167076	116168476	6	3	-68580	781	CA-GGGCA-CAA-TGTCC-CT	new
MGAT5	4249	T02F0807A47B	C02F0807A47B	chr2	+	134718587	134717387	134718787	10	5	7051	78	CC-AGTCA-GGC-TGACC-CA	new
MGAT5	4249	T02F0807A47B	C02F0807A47B	chr2	+	134718587	134717387	134718787	10	5	7051	642	GG-GGCCA-TTT-TGGCC-TG	new
MGAT5	4249	T02F08090E08	C02F08090E08	chr2	+	134811144	134809944	134811344	6	6	-85506	1075	CA-GGCCA-TAT-TGGTC-TC	new
MKI67	4288	T10R07BC8C6D	C10R07BC8C6D	chr10	-	129797229	129797029	129798429	22	11	-12296	464	CA-GGGCA-TGG-CAACC-TT	new
MKI67	4288	T10R07BC8C6D	C10R07BC8C6D	chr10	-	129797229	129797029	129798429	22	11	-12296	1357	CC-AGTCA-TTA-CAACC-AG	new

Gene Symbol	Gene ID	TSS	CTSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Tags	RepTags	Distance:		Sequence	Pattern
MMP2	4313	T16F03390C37	C16F03390C37	chr16	+	54070327	54069127	54070527	28	6	286	676	GG-GGTCT-TTG-TGACC-TC	new
MMP2	4313	T16F03390C5D	C16F03390C5D	chr16	+	54070365	54069165	54070565	234	20	240	638	GG-GGTCT-TTG-TGACC-TC	new
MMP2	4313	T16F03390D53	C16F03390D53	chr16	+	54070611	54069411	54070811	7890	1816	77	392	GG-GGTCT-TTG-TGACC-TC	new
MMP2	4313	T16F03391C2C	C16F03391C2C	chr16	+	54074412	54073212	54074612	6	4	-3823	1214	CA-GGTGA-TCT-TGACC-AG	new
MSH2	4436	T02F02D48712	C02F02D48712	chr2	+	47482642	47481442	47482842	19	18	1125	656	TT-AGTCA-GGC-TGGTC-TC	new
MSH2	4436	T02F02D48B97	C02F02D48B97	chr2	+	47483799	47482599	47483999	543	43	72	1125	GT-GGTCG-CCG-TGGCC-GG	new
MT3	4504	T16F0349FE8D	C16F0349FE8D	chr16	+	55180941	55179741	55181141	56	29	-169	340	TT-GGTCA-GGC-TGGTC-TC	known
MUC2	4583	T11F000FF0A6	C11F000FF0A6	chr11	+	1044646	1043446	1044846	6	3	20229	70	CA-GGGCA-GGC-TGCCT-GG	new
MUC2	4583	T11F00105A6B	C11F00105A6B	chr11	+	1071723	1070523	1071923	71	18	-6780	596	CA-GGTGA-ACC-TGCCC-CA	new
MUC2	4583	T11F00105A9E	C11F00105A9E	chr11	+	1071774	1070574	1071974	10	4	-6888	545	CA-GGTGA-ACC-TGCCC-CA	new
MUC2	4583	T11F00105E29	C11F00105E29	chr11	+	1072681	1071481	1072881	32	7	-7759	899	AA-GGTGG-GTG-GGACC-CA	new
MUC2	4583	T11F00106014	C11F00106014	chr11	+	1073172	1071972	1073372	68	6	-8188	408	AA-GGTGG-GTG-GGACC-CA	new
MUC2	4583	T11F0010618D	C11F0010618D	chr11	+	1073549	1072349	1073749	20	3	-8674	31	AA-GGTGG-GTG-GGACC-CA	new
MUC2	4583	T11F00106A1D	C11F00106A1D	chr11	+	1075741	1074541	1075941	24	3	-10866	1161	GG-GGTCC-AGC-TGTCC-TG	new
MUC2	4583	T11F00106AE1	C11F00106AE1	chr11	+	1075937	1074737	1076137	13	3	-11060	965	GG-GGTCC-AGC-TGTCC-TG	new
MUC2	4583	T11F00106B33	C11F00106B33	chr11	+	1076019	1074819	1076219	50	9	-11105	883	GG-GGTCC-AGC-TGTCC-TG	new
MUC2	4583	T11F001096FD	C11F001096FD	chr11	+	1087229	1086029	1087429	41	12	-22353	1286	TC-CGTCA-GGC-TGCCC-TA	new
MUC4	4585	T03R0BBDC7A3	C03R0BBDC7A3	chr3	-	196986787	196986587	196987987	6	3	-27425	769	TT-GGTCA-GGC-TGGTC-TC	known
MUC6	4588	T11R000F6745	C11R000F6745	chr11	-	1009477	1009277	1010677	6	4	-6638	174	CT-GGTCC-CCA-TGGCC-CT	new
MUC6	4588	T11R000F6A1D	C11R000F6A1D	chr11	-	1010205	1010005	1011405	15	7	-7366	902	CT-GGTCC-CCA-TGGCC-CT	new
NFKB1	4790	T04F062D7027	C04F062D7027	chr4	+	103641127	103639927	103641327	5	4	391	163	TT-GGTCA-GGC-TGGTC-TC	known

Gene Symbol	Gene ID	TSS	CTSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Tags	RepTags	Distance:		Sequence	Pattern
NFKB1	4790	T04F062D70B7	C04F062D70B7	chr4	+	103641271	103640071	103641471	23	7	282	19	TT-GGTCA-GGC-TGGTC-TC	known
NFKB1	4790	T04F062E9997	C04F062E9997	chr4	+	103717271	103716071	103717471	5	3	-75750	1138	AA-GGTGG-AGG-GGACC-GG	new
NOS2A	4843	T17R016143A8	C17R016143A8	chr17	-	23151528	23151328	23152728	12	8	-43573	868	AA-GATCA-GGT-CACCC-AC	new
NP	4860	T14F013149F1	C14F013149F1	chr14	+	20007409	20006209	20007609	57	30	8	112	CA-GGTCA-CTG-CAACC-TC	new
OGG1	4968	T03F009506EC	C03F009506EC	chr3	+	9766636	9765436	9766836	431	44	-888	418	AA-GGTGA-GAA-TGGCC-CA	new
OGG1	4968	T03F00950821	C03F00950821	chr3	+	9766945	9765745	9767145	136	8	-1224	109	AA-GGTGA-GAA-TGGCC-CA	new
PBX1	5087	T01F09B4110C	C01F09B4110C	chr1	+	162795788	162794588	162795988	9	3	-226	1133	GG-GGTCG-GGA-TGGCC-GG	new
PECAM1	5175	T17R0390211C	C17R0390211C	chr17	-	59777308	59777108	59778508	10	3	-23147	648	AA-GGGCA-GGC-TGGCC-TG	new
PECAM1	5175	T17R0390BECF	C17R0390BECF	chr17	-	59817679	59817479	59818879	7	3	-63511	493	TT-GGTCA-GGC-TGGTC-TC	known
PPL	5493	T16R004A5922	C16R004A5922	chr16	-	4872482	4872282	4873682	5	3	49	622	AA-GGACA-AGC-CAACC-AC	new
MAPK3	5595	T16R01CA66BE	C16R01CA66BE	chr16	-	30041790	30041590	30042990	23	4	-8820	154	TT-GGTCA-GGC-TGGTC-TC	known
PTPN1	5770	T20F02E4F8B1	C20F02E4F8B1	chr20	+	48560305	48559105	48560505	698	84	7	161	TT-GGTCA-GGC-TGGTC-TC	known
RARA	5914	T17F0221073D	C17F0221073D	chr17	+	35718973	35717773	35719173	23	7	14	348	GA-GGTCC-CTC-TGCCC-CT	new
RARA	5914	T17F02213EF7	C17F02213EF7	chr17	+	35733239	35732039	35733439	5	5	-14267	230	AG-TGTCA-CCG-TGACC-CA	new
RARB	5915	T03F0184DAEC	C03F0184DAEC	chr3	+	25484012	25482812	25484212	38	8	-39246	83	TT-GGTCA-GGC-TGGTC-TC	known
RBL2	5934	T16F03198981	C16F03198981	chr16	+	52005249	52004049	52005449	6	4	20652	1353	TT-GGTCG-GGC-TGGTC-TC	new
RELA	5970	T11R03E25271	C11R03E25271	chr11	-	65163889	65163689	65165089	8	5	14522	909	CC-AGTCA-CCC-TGCCC-TG	new
RELA	5970	T11R03E29B2D	C11R03E29B2D	chr11	-	65182509	65182309	65183709	18	5	-4082	59	GG-GGTGA-GTG-TGGCC-TG	new
RELA	5970	T11R03E29C52	C11R03E29C52	chr11	-	65182802	65182602	65184002	17	6	-4388	352	GG-GGTGA-GTG-TGGCC-TG	new
RELA	5970	T11R03E29DFF	C11R03E29DFF	chr11	-	65183231	65183031	65184431	9	3	-4813	781	GG-GGTGA-GTG-TGGCC-TG	new
RXRA	6256	T09F08217724	C09F08217724	chr9	+	136410916	136409716	136411116	17	16	-52685	859	CA-AGTCA-GAT-TGTCC-TG	new

Gene Symbol	Gene ID	TSS	CTSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Tags	RepTags	Distance:		Sequence	Pattern
S100A10	6281	T01R08F45E53	C01R08F45E53	chr1	-	150232659	150232459	150233859	8	3	-10624	476	GA-GGTCG-ATA-GGACT-CC	new
S100A10	6281	T01R08F45F62	C01R08F45F62	chr1	-	150232930	150232730	150234130	184	30	-10859	747	GA-GGTCG-ATA-GGACT-CC	new
SDC1	6382	T02R01353CC3	C02R01353CC3	chr2	-	20266179	20265979	20267379	11	3	-2117	363	TT-GGTCC-TAA-TGCCT-GG	new
SDC1	6382	T02R01353DBC	C02R01353DBC	chr2	-	20266428	20266228	20267628	5	3	-2370	612	TT-GGTCC-TAA-TGCCT-GG	new
SKIL	6498	T03F0A396CBC	C03F0A396CBC	chr3	+	171535548	171534348	171535748	24	18	22660	1136	AA-GGTCT-TTA-TGACC-TG	new
FSCN1	6624	T07F00556EFD	C07F00556EFD	chr7	+	5598973	5597773	5599173	1564	819	21	574	GG-GGTCC-TGG-TGGTC-TT	new
FSCN1	6624	T07F00556EFD	C07F00556EFD	chr7	+	5598973	5597773	5599173	1564	819	21	917	GG-GGTCG-GGG-TGTCC-GA	new
SRC	6714	T20F021C47E4	C20F021C47E4	chr20	+	35407844	35406644	35408044	45	4	-1238	999	GA-GGGCA-TTT-TGGCC-TG	new
SRC	6714	T20F021C4838	C20F021C4838	chr20	+	35407928	35406728	35408128	176	29	-1416	915	GA-GGGCA-TTT-TGGCC-TG	new
SRP54	6729	T14F020EC227	C14F020EC227	chr14	+	34521639	34520439	34521839	35	6	398	799	CT-GGTCA-ATC-TACCC-CC	new
SRP54	6729	T14F020EC352	C14F020EC352	chr14	+	34521938	34520738	34522138	309	42	72	500	CT-GGTCA-ATC-TACCC-CC	new
STAT3	6774	T17R02400D8B	C17R02400D8B	chr17	-	37752203	37752003	37753403	10	4	-33295	486	TA-GGCCA-TCC-CAGCC-TG	new
AURKA	6790	T20R033E16A6	C20R033E16A6	chr20	-	54400678	54400478	54401878	35	4	-22750	1005	CA-GGTCT-GGC-TGGCC-GT	new
TERF2	7014	T16R040D3FB6	C16R040D3FB6	chr16	-	67977142	67976942	67978342	42	10	-30020	631	TT-GGTCA-GGC-TGGTC-TC	known
TERF2	7014	T16R040D4055	C16R040D4055	chr16	-	67977301	67977101	67978501	384	19	-30177	790	TT-GGTCA-GGC-TGGTC-TC	known
TERT	7015	T05R0013C2D7	C05R0013C2D7	chr5	-	1295063	1294863	1296263	11	4	11236	41	CG-GGTCA-GCC-CAACC-CC	new
TGM2	7052	T20R02283F93	C20R02283F93	chr20	-	36192147	36191947	36193347	6	4	-1851	1004	GA-GGGCA-GAT-TGTCC-AG	new
TGM3	7053	T20F00228137	C20F00228137	chr20	+	2261303	2260103	2261503	6	4	-36628	704	GA-AGTCA-ACC-TGGTC-CT	new
TIMP2	7077	T17R046EAB4C	C17R046EAB4C	chr17	-	74361676	74361476	74362876	16	4	-1001	690	CA-GGACA-CTA-TGGCC-TG	new
TIMP2	7077	T17R046EAB7F	C17R046EAB7F	chr17	-	74361727	74361527	74362927	20	9	-1053	741	CA-GGACA-CTA-TGGCC-TG	new
TIMP2	7077	T17R046EAC82	C17R046EAC82	chr17	-	74361986	74361786	74363186	56	20	-1301	1000	CA-GGACA-CTA-TGGCC-TG	new

Gene Symbol	Gene ID	TSS	CTSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Tags	RepTags	Distance:		Sequence	Pattern
TIMP2	7077	T17R046EACAC	C17R046EACAC	chr17	-	74362028	74361828	74363228	8	4	-1354	1042	CA-GGACA-CTA-TGGCC-TG	new
TPM1	7168	T15F03A4D4F8	C15F03A4D4F8	chr15	+	61134072	61132872	61134272	14	5	-12163	266	TT-GGTCA-GGC-TGGTC-TC	known
TPM1	7168	T15F03A4D4F8	C15F03A4D4F8	chr15	+	61134072	61132872	61134272	14	5	-12163	579	CT-GGTCT-CCA-TGCCC-AG	new
VEGFA	7422	T06F029D06FA	C06F029D06FA	chr6	+	43845370	43844170	43845570	6	3	561	920	CC-AGTCA-CTC-CAGCC-TG	new
VEGFA	7422	T06F029D086D	C06F029D086D	chr6	+	43845741	43844541	43845941	5	3	191	549	CC-AGTCA-CTC-CAGCC-TG	new
VEGFA	7422	T06F029D0926	C06F029D0926	chr6	+	43845926	43844726	43846126	14045	3244	37	364	CC-AGTCA-CTC-CAGCC-TG	new
VHL	7428	T03F009B0605	C03F009B0605	chr3	+	10159621	10158421	10159821	6	4	-1302	813	TT-GGTCA-GGC-TGGTC-TC	known
VHL	7428	T03F009B0605	C03F009B0605	chr3	+	10159621	10158421	10159821	6	4	-1302	1344	AA-GGTGA-TTT-TACCC-CT	new
EZR	7430	T06R097C10CE	C06R097C10CE	chr6	-	159125710	159125510	159126910	21	4	-18900	614	GC-AGTCA-TGC-CGTCC-TC	new
VIM	7431	T10F0108240B	C10F0108240B	chr10	+	17310731	17309531	17310931	10	5	573	190	AA-GGTTA-AGG-TGACC-GA	new
WNT7B	7477	T22R02A9FF20	C22R02A9FF20	chr22	-	44695328	44695128	44696528	11	7	1007	670	CT-GGTCC-TTT-TACCC-TG	new
WT1	7490	T11R01EE8567	C11R01EE8567	chr11	-	32408935	32408735	32410135	10	4	-43008	1266	GA-GGTCT-GCG-TGTCC-GG	new
XRCC5	7520	T02F0CEA6C29	C02F0CEA6C29	chr2	+	216689705	216688505	216689905	8	5	-7325	1030	GA-GGTCT-GGT-TGTCC-TG	new
YES1	7525	T18R000B9C0B	C18R000B9C0B	chr18	-	760843	760643	762043	5	4	-49234	862	TT-GGTCA-GGC-TGGTC-TT	new
PRDM2	7799	T01F00D55012	C01F00D55012	chr1	+	13979666	13978466	13979866	11	5	-75727	739	AA-AGTCA-CAG-CGACT-CA	new
RAB11A	8766	T15F03CFD712	C15F03CFD712	chr15	+	63952658	63951458	63952858	5	3	-3808	572	TT-GGTCA-GGC-TGGTC-TC	known
CRLF1	9244	T19R011B5610	C19R011B5610	chr19	-	18568720	18568520	18569920	7	4	-3653	161	GA-GGCCA-GCT-CACCC-TG	new
CRLF1	9244	T19R011B566F	C19R011B566F	chr19	-	18568815	18568615	18570015	38	6	-3718	256	GA-GGCCA-GCT-CACCC-TG	new
GPR56	9289	T16F0359B808	C16F0359B808	chr16	+	56211464	56210264	56211664	32	6	22	824	CA-GGTGA-GTC-TGGCC-CA	new
GPR56	9289	T16F0359B808	C16F0359B808	chr16	+	56211464	56210264	56211664	32	6	22	1059	AG-GGTCA-GGT-TGGCC-CC	new
DLEC1	9940	T03F0245814A	C03F0245814A	chr3	+	38109514	38108314	38109714	32	7	-53813	1253	CA-AGCCA-GTG-TGGCC-TC	new

Gene Symbol	Gene ID	TSS	CTSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Tags	RepTags	Distance:		Sequence	Pattern
HS3ST2	9956	T16F015AE0B8	C16F015AE0B8	chr16	+	22732984	22731784	22733184	248	45	378	854	CA-GGGCA-GGG-TGCCT-GC	new
HS3ST2	9956	T16F015AE0B8	C16F015AE0B8	chr16	+	22732984	22731784	22733184	248	45	378	880	TG-GGTCT-GTC-TGGTC-TG	new
NDRG1	10397	T08R0801A109	C08R0801A109	chr8	-	134324489	134324289	134325689	5	3	-5873	1123	GG-GGTCA-GCC-CAGCC-TC	new
AKAP13	11214	T15F0503078B	C15F0503078B	chr15	+	84084619	84083419	84084819	8	6	-359744	834	TA-GGTCA-CTG-TGGTC-TT	new
CADM1	23705	T11R06D3F6E4	C11R06D3F6E4	chr11	-	114554596	114554396	114555796	54	18	-4346	736	AT-AGTCA-GGT-CAGCC-CC	new
GAPDHS	26330	T19F026D32A1	C19F026D32A1	chr19	+	40710817	40709617	40711017	5	3	5387	28	CT-GGCCA-AAC-TGCCC-TG	new
PABPC1	26986	T08R061164DC	C08R061164DC	chr8	-	101803228	101803028	101804428	10	3	-18861	14	AA-AGCCA-CTT-TGACC-AA	new
PRRX2	51450	T09F07D6641E	C09F07D6641E	chr9	+	131490846	131489646	131491046	5	3	-23105	801	TG-AGTCA-CCG-TGCCC-GG	new
SBNO1	55206	T12R074B39D3	C12R074B39D3	chr12	-	122370515	122370315	122371715	10	8	-24088	58	GA-GATCA-GCC-TGACC-AA	new
CTDSP1	58190	T02F0D0D4708	C02F0D0D4708	chr2	+	218973960	218972760	218974160	15	9	-1238	1123	TG-GGTCT-GGC-TGCCC-CG	new
OSR1	130497	T02R01283138	C02R01283138	chr2	-	19411256	19411056	19412456	53	5	3515	172	CT-GGTCC-CAG-TGACT-GT	new
OSR1	130497	T02R01283138	C02R01283138	chr2	-	19411256	19411056	19412456	53	5	3515	1037	GT-GGTCT-CCC-TGCCC-GC	new
MARCH8	220972	T10R02B4E5B9	C10R02B4E5B9	chr10	-	45409721	45409521	45410921	7	4	-136877	945	TG-GGTCT-CTC-TGCCC-CG	new
MARCH8	220972	T10R02B4E65E	C10R02B4E65E	chr10	-	45409886	45409686	45411086	31	28	-137044	1110	TG-GGTCT-CTC-TGCCC-CG	new
MARCH8	220972	T10R02B4E6DD	C10R02B4E6DD	chr10	-	45410013	45409813	45411213	245	26	-137118	1237	TG-GGTCT-CTC-TGCCC-CG	new
MUC5B	727897	T11F001252E8	C11F001252E8	chr11	+	1200872	1199672	1201072	920	139	187	61	GA-GGCCA-CAC-CACCC-GA	new
MUC5B	727897	T11F001252E8	C11F001252E8	chr11	+	1200872	1199672	1201072	920	139	187	1117	AA-GAGCA-TGG-TGCCC-TG	new

APPENDIX II: ESCC genes categorized by predictions and experimental proof

Category C1: predicted estrogen responsive ESCC genes with experimental proof verifying estrogen responsiveness.

Gene Symbol	GeneID	TSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Promotersize	Total Tags	RepTags	Distance:	ERE Position	ERE Pattern		
ABL1	25	T09F07E6FAFC	chr9	+	132578044	132576844	132578244	1400	13	3	1047	634	TT-GGTCA-GGC-TGGTC-TC	known	pattern
ABL1	25	T09F07E6FBCC	chr9	+	132578252	132577052	132578452	1400	37	9	865	426	TT-GGTCA-GGC-TGGTC-TC	known	pattern
ABL1	25	T09F07E6FBFD	chr9	+	132578301	132577101	132578501	1400	9	6	809	377	TT-GGTCA-GGC-TGGTC-TC	known	pattern
ANXA2	302	T15R037B8B75	chr15	-	58428277	58428077	58429477	1400	7	7	-1616	928	AG-TGTCA-GCC-TGCCC-TG	new	pattern
APC	324	T05F06B00439	chr5	+	112198713	112197513	112198913	1400	5	3	-97226	394	CG-GCTCA-CTG-CGACC-TC	new	pattern
BCL2	596	T18R0383753E	chr18	-	58946878	58946678	58948078	1400	33	11	-5298	1240	GA-AGTCA-ACA-TGCCT-GC	new	pattern
BCL2L1	598	T20R01C5892A	chr20	-	29722922	29722722	29724122	1400	8	5	-6980	380	TT-GGTCA-GGC-TGGTC-TT	new	pattern
BCL2L1	598	T20R01C5892A	chr20	-	29722922	29722722	29724122	1400	8	5	-6980	458	TG-AGTCA-TTG-TGCCC-GG	new	pattern
CDC2	983	T10F03B498BF	chr10	+	62167231	62166031	62167431	1400	7	6	38459	1048	CA-AATCA-CCA-TGACC-TA	new	pattern
CDKN1A	1026	T06F0230D425	chr6	+	36754469	36753269	36754669	1400	774	483	3	605	AA-AATCA-TTC-TGGCC-TC	new	pattern
CDX1	1044	T05F08E92640	chr5	+	149497408	149496208	149497608	1400	139	86	29145	442	CA-GGGCA-GTG-TGCCT-GG	new	pattern
CLU	1191	T08R01A3FB63	chr8	-	27523939	27523739	27525139	1400	22	8	-13551	1357	TG-GGTCT-GGG-TGCCT-GT	new	pattern
CLDN4	1364	T07F04581C3F	chr7	+	72883263	72882063	72883463	1400	358	75	-132	903	TG-GGTCT-GGT-TGCCC-CC	new	pattern
CTNNA1	1495	T05F083B7FCD	chr5	+	138117069	138115869	138117269	1400	504	56	22	185	AA-GGTGG-CTC-TAGCC-TG	new	pattern
CTNNB1	1499	T03F0274EA18	chr3	+	41216536	41215336	41216736	1400	6	3	-532	1029	CC-GGTCG-CGC-TGCCT-CT	new	pattern
CTSD	1509	T11R001A5981	chr11	-	1726849	1726649	1728049	1400	5	3	3730	252	CG-GGACA-CTT-TGCCC-TG	new	pattern
CTSD	1509	T11R001A59B6	chr11	-	1726902	1726702	1728102	1400	5	3	3682	305	CG-GGACA-CTT-TGCCC-TG	new	pattern
CTSD	1509	T11R001A6B3F	chr11	-	1731391	1731191	1732591	1400	9	3	-810	11	GG-GGTCA-GAC-CAGCC-CT	new	pattern
DAB2	1601	T05R0259A87B	chr5	-	39430267	39430067	39431467	1400	7	3	-21747	1320	AT-GGTCA-GCC-CGACC-AA	new	pattern
NQO1	1728	T16R0412730C	chr16	-	68317964	68317764	68319164	1400	769	272	-17134	736	GC-AGTCA-CAG-TGACT-CA	known	pattern
EFNB2	1948	T13R06511B39	chr13	-	105978681	105978481	105979881	1400	6	4	-38567	400	CA-GGACA-CAT-TGTCC-CT	new	pattern
EFNB2	1948	T13R06512E8E	chr13	-	105983630	105983430	105984830	1400	13	10	-43518	346	GC-GGTCT-CGA-TGGCC-GA	new	pattern

EGFR	1956	T07F0347ACD5	chr7	+	55028949	55027749	55029149	1400	5	5	25270	776	GG-GGTCA-GCC-TGCCC-TA	new	pattern
EGFR	1956	T07F03480EB7	chr7	+	55054007	55052807	55054207	1400	21	5	215	853	GC-GGGCA-CGG-CGACC-TC	new	pattern
EGFR	1956	T07F03480EEA	chr7	+	55054058	55052858	55054258	1400	5	4	161	802	GC-GGGCA-CGG-CGACC-TC	new	pattern
EGFR	1956	T07F03480F7D	chr7	+	55054205	55053005	55054405	1400	589	65	66	655	GC-GGGCA-CGG-CGACC-TC	new	pattern
EGFR	1956	T07F0348102B	chr7	+	55054379	55053179	55054579	1400	7	4	-143	481	GC-GGGCA-CGG-CGACC-TC	new	pattern
EGFR	1956	T07F03481051	chr7	+	55054417	55053217	55054617	1400	58	10	-192	443	GC-GGGCA-CGG-CGACC-TC	new	pattern
EGFR	1956	T07F03481158	chr7	+	55054680	55053480	55054880	1400	15	3	-443	180	GC-GGGCA-CGG-CGACC-TC	new	pattern
ENO3	2023	T01R00870718	chr1	-	8849176	8848976	8850376	1400	5	3	-5507	205	CA-GGTGA-CAA-TAGCC-TG	new	pattern
MDS1	2122	T03R0A26D0B7	chr3	-	170315959	170315759	170317159	1400	6	4	-30695	182	TG-AGTCA-TGC-CAACC-CG	new	pattern
MDS1	2122	T03R0A26D50C	chr3	-	170317068	170316868	170318268	1400	5	3	-31787	266	AA-GGTTA-GCA-CAACC-AC	new	pattern
MDS1	2122	T03R0A26D50C	chr3	-	170317068	170316868	170318268	1400	5	3	-31787	1291	TG-AGTCA-TGC-CAACC-CG	new	pattern
FGF4	2249	T11R0421678A	chr11	-	69298058	69297858	69299258	1400	7	3	-1063	265	GC-GGTCC-TGC-TGGCC-TT	new	pattern
FLT1	2321	T13R01AA206D	chr13	-	27926637	27926437	27927837	1400	32	24	-152830	190	TA-GGTCT-CTA-TGATC-CA	new	pattern
FN1	2335	T02R0CDFC7B3	chr2	-	215992243	215992043	215993443	1400	364	103	-58795	322	TG-GGTCA-GGC-TGGTC-TC	new	pattern
FN1	2335	T02R0CDFDEBD	chr2	-	215998141	215997941	215999341	1400	23	5	-64646	357	AA-GGTGG-AGG-TGACC-TC	new	pattern
GAPDH	2597	T12F0063651F	chr12	+	6513951	6512751	6514151	1400	93	20	66	456	CA-GGACA-TCG-TGACC-TT	new	pattern
GAPDH	2597	T12F00636637	chr12	+	6514231	6513031	6514431	1400	21	3	-297	176	CA-GGACA-TCG-TGACC-TT	new	pattern
GATA4	2626	T08F00B0A470	chr8	+	11576432	11575232	11576632	1400	5	3	22731	624	AT-GGTCA-CTG-TGTCC-CC	new	pattern
GATA4	2626	T08F00B0D09C	chr8	+	11587740	11586540	11587940	1400	60	55	11423	148	CA-GGCCA-CAG-CAACC-TG	new	pattern
HOXC4	3221	T12F03246972	chr12	+	52717938	52716738	52718138	1400	63	11	-20980	282	CA-GGGCA-CAC-CAGCC-CC	new	pattern
HOXC4	3221	T12F03246972	chr12	+	52717938	52716738	52718138	1400	63	11	-20980	857	AA-GATCA-ACA-CGACC-TC	new	pattern
HOXC4	3221	T12F032469A6	chr12	+	52717990	52716790	52718190	1400	23	9	-21081	230	CA-GGGCA-CAC-CAGCC-CC	new	pattern
HOXC4	3221	T12F032469A6	chr12	+	52717990	52716790	52718190	1400	23	9	-21081	805	AA-GATCA-ACA-CGACC-TC	new	pattern
HOXC9	3225	T12F0323C141	chr12	+	52674881	52673681	52675081	1400	66	5	5268	1352	TG-GGTCT-CTC-TGCCT-GA	new	pattern
HSPA4	3308	T05F07E4C83C	chr5	+	132433980	132432780	132434180	1400	11	6	-18407	266	TT-GGTCA-CGC-TGGTC-TT	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	638	GG-GGTCT-TTC-TGCCC-AG	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	1110	GA-GGCCA-GCG-CAGCC-CC	new	pattern

ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	325	CA-GGGCA-GGG-TGGCC-AG	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	406	AA-GGTCC-ACC-TGCCT-CC	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	1050	TT-GGTCA-ATG-CAGCC-CT	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	1037	TT-GGTCA-GGA-CAGCC-CA	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	744	TT-AGTCA-GGC-TGATC-TC	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	1235	TG-AGTCA-GGC-TGCCC-CG	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	897	GA-CGTCA-AAC-TGCCC-TC	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	520	TT-GGTCA-GGC-TGGTC-TC	known	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	37	AA-AGCCA-AAG-TGACC-AA	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	1291	CG-GGTCA-TGG-CGCCC-CG	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	1033	AG-AGTCA-CCC-TGGCC-CG	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	794	AA-GGTCA-CGG-TGGCC-AC	known	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	850	CA-CGTCA-CTG-TGACC-TA	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	1287	CT-GGTCC-TCC-TGACT-GG	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	58	AG-GGGCA-TGT-TGACC-CA	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	332	GA-GGACA-GAA-TGCCC-TC	new	pattern
ID1	3397	T20F01C48AB7	chr20	+	29657783	29656583	29657983	1400	13	3	-1029	906	GC-AGTCA-CCG-TGGTC-CT	new	pattern
IGF1R	3480	T15F05CC041F	chr15	+	97256479	97255279	97256679	1400	5	3	-246191	233	CC-AGTCA-CAG-TGGCT-CA	new	pattern
IGF1R	3480	T15F05CC12C5	chr15	+	97260229	97259029	97260429	1400	9	9	-249941	1284	TT-GGTCT-CCT-TGTCC-TT	new	pattern
IGF1R	3480	T15F05CC17A8	chr15	+	97261480	97260280	97261680	1400	5	4	-251192	33	TT-GGTCT-CCT-TGTCC-TT	new	pattern
IGF2	3481	T11R00203063	chr11	-	2109539	2109339	2110739	1400	21	4	-2579	477	TG-AGTCA-AAT-TGGCC-TG	new	pattern
IGF2	3481	T11R002030C6	chr11	-	2109638	2109438	2110838	1400	40	14	-2692	576	TG-AGTCA-AAT-TGGCC-TG	new	pattern
IGF2	3481	T11R0020310A	chr11	-	2109706	2109506	2110906	1400	31	7	-2736	644	TG-AGTCA-AAT-TGGCC-TG	new	pattern
IGF2	3481	T11R0020313A	chr11	-	2109754	2109554	2110954	1400	25	13	-2808	692	TG-AGTCA-AAT-TGGCC-TG	new	pattern
IGF2	3481	T11R00203166	chr11	-	2109798	2109598	2110998	1400	17	13	-2852	736	TG-AGTCA-AAT-TGGCC-TG	new	pattern
IGF2	3481	T11R00203190	chr11	-	2109840	2109640	2111040	1400	20	4	-2873	778	TG-AGTCA-AAT-TGGCC-TG	new	pattern
IGF2	3481	T11R00203254	chr11	-	2110036	2109836	2111236	1400	2060	468	-2918	974	TG-AGTCA-AAT-TGGCC-TG	new	pattern

IGF2	3481	T11R002032EE	chr11	-	2110190	2109990	2111390	1400	338	76	-3183	1128	TG-AGTCA-AAT-TGGCC-TG	new	pattern
IL6R	3570	T01F09192B5F	chr1	+	152644447	152643247	152644647	1400	23	5	-102	654	CA-GGTCT-ATC-TGACT-GC	new	pattern
IL6R	3570	T01F09192BAD	chr1	+	152644525	152643325	152644725	1400	42	3	-176	576	CA-GGTCT-ATC-TGACT-GC	new	pattern
ITGA6	3655	T02F0A50A70A	chr2	+	173057802	173056602	173058002	1400	7	3	-57186	387	TT-GGTCA-GGC-TGGTC-TC	known	pattern
ITGA6	3655	T02F0A50A76A	chr2	+	173057898	173056698	173058098	1400	32	8	-57258	291	TT-GGTCA-GGC-TGGTC-TC	known	pattern
ITGA6	3655	T02F0A50A87B	chr2	+	173058171	173056971	173058371	1400	18	8	-57512	18	TT-GGTCA-GGC-TGGTC-TC	known	pattern
KDR	3791	T04R0350F74B	chr4	-	55637835	55637635	55639035	1400	6	6	1590	1294	TG-GGTCA-GGC-TGGCC-TT	new	pattern
KIT	3815	T04F034A92C1	chr4	+	55218881	55217681	55219081	1400	378	97	112	547	GA-GGCCA-CGC-CGACC-AA	new	pattern
LCN2	3934	T09F07BEE741	chr9	+	129951553	129950353	129951753	1400	37	10	-13	380	AT-GGTCT-CAG-TGACC-TC	new	pattern
MCM4	4173	T08F02EC40A7	chr8	+	49037479	49036279	49037679	1400	12	5	-1431	263	TG-AGTCA-TAA-TGCCC-CA	new	pattern
MET	4233	T07F06EC9654	chr7	+	116168276	116167076	116168476	1400	6	3	-68580	781	CA-GGGCA-CAA-TGTCC-CT	new	pattern
MUC6	4588	T11R000F6745	chr11	-	1009477	1009277	1010677	1400	6	4	-6638	174	CT-GGTCC-CCA-TGGCC-CT	new	pattern
MUC6	4588	T11R000F6A1D	chr11	-	1010205	1010005	1011405	1400	15	7	-7366	902	CT-GGTCC-CCA-TGGCC-CT	new	pattern
NP	4860	T14F013149F1	chr14	+	20007409	20006209	20007609	1400	57	30	8	112	CA-GGTCA-CTG-CAACC-TC	new	pattern
PBX1	5087	T01F09B4110C	chr1	+	162795788	162794588	162795988	1400	9	3	-226	1133	GG-GGTCG-GGA-TGGCC-GG	new	pattern
RARA	5914	T17F0221073D	chr17	+	35718973	35717773	35719173	1400	23	7	14	348	GA-GGTCC-CTC-TGCCC-CT	new	pattern
RARA	5914	T17F02213EF7	chr17	+	35733239	35732039	35733439	1400	5	5	-14267	230	AG-TGTCA-CCG-TGACC-CA	new	pattern
RARB	5915	T03F0184DAEC	chr3	+	25484012	25482812	25484212	1400	38	8	-39246	83	TT-GGTCA-GGC-TGGTC-TC	known	pattern
RBL2	5934	T16F03198981	chr16	+	52005249	52004049	52005449	1400	6	4	20652	1353	TT-GGTCG-GGC-TGGTC-TC	new	pattern
RXRA	6256	T09F08217724	chr9	+	136410916	136409716	136411116	1400	17	16	-52685	859	CA-AGTCA-GAT-TGTCC-TG	new	pattern
S100A10	6281	T01R08F45E53	chr1	-	150232659	150232459	150233859	1400	8	3	-10624	476	GA-GGTCG-ATA-GGACT-CC	new	pattern
S100A10	6281	T01R08F45F62	chr1	-	150232930	150232730	150234130	1400	184	30	-10859	747	GA-GGTCG-ATA-GGACT-CC	new	pattern
SKIL	6498	T03F0A396CBC	chr3	+	171535548	171534348	171535748	1400	24	18	22660	1136	AA-GGTCT-TTA-TGACC-TG	new	pattern
SRC	6714	T20F021C47E4	chr20	+	35407844	35406644	35408044	1400	45	4	-1238	999	GA-GGGCA-TTT-TGGCC-TG	new	pattern
SRC	6714	T20F021C4838	chr20	+	35407928	35406728	35408128	1400	176	29	-1416	915	GA-GGGCA-TTT-TGGCC-TG	new	pattern
AURKA	6790	T20R033E16A6	chr20	-	54400678	54400478	54401878	1400	35	4	-22750	1005	CA-GGTCT-GGC-TGGCC-GT	new	pattern
TERT	7015	T05R0013C2D7	chr5	-	1295063	1294863	1296263	1400	11	4	11236	41	CG-GGTCA-GCC-CAACC-CC	new	pattern

TPM1	7168	T15F03A4D4F8	chr15	+	61134072	61132872	61134272	1400	14	5	-12163	266	TT-GGTCA-GGC-TGGTC-TC	known	pattern
TPM1	7168	T15F03A4D4F8	chr15	+	61134072	61132872	61134272	1400	14	5	-12163	579	CT-GGTCT-CCA-TGCCC-AG	new	pattern
VEGFA	7422	T06F029D06FA	chr6	+	43845370	43844170	43845570	1400	6	3	561	920	CC-AGTCA-CTC-CAGCC-TG	new	pattern
VEGFA	7422	T06F029D086D	chr6	+	43845741	43844541	43845941	1400	5	3	191	549	CC-AGTCA-CTC-CAGCC-TG	new	pattern
VEGFA	7422	T06F029D0926	chr6	+	43845926	43844726	43846126	1400	14045	3244	37	364	CC-AGTCA-CTC-CAGCC-TG	new	pattern
WT1	7490	T11R01EE8567	chr11	-	32408935	32408735	32410135	1400	10	4	-43008	1266	GA-GGTCT-GCG-TGTCC-GG	new	pattern
NDRG1	10397	T08R0801A109	chr8	-	134324489	134324289	134325689	1400	5	3	-5873	1123	GG-GGTCA-GCC-CAGCC-TC	new	pattern
CADM1	23705	T11R06D3F6E4	chr11	-	114554596	114554396	114555796	1400	54	18	-4346	736	AT-AGTCA-GGT-CAGCC-CC	new	pattern
PABPC1	26986	T08R061164DC	chr8	-	101803228	101803028	101804428	1400	10	3	-18861	14	AA-AGCCA-CTT-TGACC-AA	new	pattern
OSR1	130497	T02R01283138	chr2	-	19411256	19411056	19412456	1400	53	5	3515	172	CT-GGTCC-CAG-TGACT-GT	new	pattern
OSR1	130497	T02R01283138	chr2	-	19411256	19411056	19412456	1400	53	5	3515	1037	GT-GGTCT-CCC-TGCCC-GC	new	pattern

Category C2: predicted estrogen responsive ESCC genes lacking experimental proof verifying estrogen responsiveness.

Gene Symbol	GeneID	TSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Promotersize	Total Tags	RepTags	Distance:	ERE Position	ERE Pattern		
ACTN4	81	T19F029CFFE2	chr19	+	43843554	43842354	43843754	1400	7	3	-13383	726	TT-GGTCA-GGC-TGGCC-TT	new	pattern
ACTN4	81	T19F029D9BDC	chr19	+	43883484	43882284	43883684	1400	39	14	-53316	1058	GG-GGCCA-GAG-TGGCC-TG	new	pattern
ACTN4	81	T19F029DBE32	chr19	+	43892274	43891074	43892474	1400	10	8	-62100	1179	AA-GGGCA-AGG-TGGCT-GA	new	pattern
ACTN4	81	T19F029DD0C0	chr19	+	43897024	43895824	43897224	1400	6	3	-66842	1268	GA-GGTCA-CTG-TGACT-CT	new	pattern
ATM	472	T11F0669D4A8	chr11	+	107599016	107597816	107599216	1400	55	3	-182	562	AA-GGTCC-TTC-TGTCC-AG	new	pattern
BAX	581	T19F033A5DEB	chr19	+	54156779	54155579	54156979	1400	9	7	-6850	57	TT-GGTCA-GGC-TGGTC-TT	new	pattern
BMP6	654	T06F007686F4	chr6	+	7767796	7766596	7767996	1400	5	5	-95787	686	TT-GGTCT-TGT-TGCCC-AG	new	pattern
BMP6	654	T06F0076E486	chr6	+	7791750	7790550	7791950	1400	7	3	-119739	40	AA-GGTGG-GGC-TGGCC-CC	new	pattern
CALR	811	T19F00C5082D	chr19	+	12912685	12911485	12912885	1400	41	6	-2210	1303	CA-GGTGA-GTC-TGACT-CA	new	pattern
RUNX3	864	T01R017FF283	chr1	-	25162371	25162171	25163571	1400	6	6	-63707	132	GT-GGTCT-CTG-TGGTC-TC	new	pattern
RUNX3	864	T01R017FF283	chr1	-	25162371	25162171	25163571	1400	6	6	-63707	1117	TG-GGTCT-GCC-TGCCT-GC	new	pattern
CDC25B	994	T20F00383158	chr20	+	3682648	3681448	3682848	1400	6	4	41753	1134	AA-GGCCA-CTC-CACCC-CC	new	pattern

CDC25B	994	T20F0038F8EC	chr20	+	3733740	3732540	3733940	1400	6	6	-9339	113	CT-GGTCC-ATA-TGACT-GC	new	pattern
CDH1	999	T16F04039AE1	chr16	+	67345121	67343921	67345321	1400	38	23	-16425	1324	CA-GGGCA-CGG-TGGCC-GA	new	pattern
CSTB	1476	T21R029FB389	chr21	-	44020617	44020417	44021817	1400	154	11	-2323	914	CC-GGTCC-GTG-TGACT-CG	new	pattern
DNMT1	1786	T19R009AB2E0	chr19	-	10138336	10138136	10139536	1400	13	4	-33267	445	AA-GATCA-GCC-TGGCC-AA	new	pattern
DSP	1832	T06F0072B44E	chr6	+	7517262	7516062	7517462	1400	7	4	-30393	1156	GA-GGGCA-GGA-TGACT-CT	new	pattern
EPHA2	1969	T01R00F98B55	chr1	-	16354133	16353933	16355333	1400	9	3	-30695	612	TT-GGTCG-CAC-GGGCC-TC	new	pattern
ERCC2	2068	T19R0303913C	chr19	-	50565436	50565236	50566636	1400	11	5	-18733	396	TT-GGTCA-GGC-TGGTC-TT	new	pattern
ETS2	2114	T21F0254E0A0	chr21	+	39116960	39115760	39117160	1400	11	5	-17212	923	GA-GGTCG-CCG-GGACC-AC	new	pattern
ETS2	2114	T21F0254E128	chr21	+	39117096	39115896	39117296	1400	24	7	-17368	787	GA-GGTCG-CCG-GGACC-AC	new	pattern
ETS2	2114	T21F0254E24F	chr21	+	39117391	39116191	39117591	1400	20	8	-17655	492	GA-GGTCG-CCG-GGACC-AC	new	pattern
ETS2	2114	T21F0254E2D4	chr21	+	39117524	39116324	39117724	1400	9	3	-17804	359	GA-GGTCG-CCG-GGACC-AC	new	pattern
FGFR4	2264	T05F0A843489	chr5	+	176436361	176435161	176436561	1400	6	6	10166	32	TA-GGTCA-CTG-CAGCC-TC	new	pattern
FHIT	2272	T03R03A24F58	chr3	-	60968792	60968592	60969992	1400	66	39	-1258696	1270	CA-GGTGA-AAT-TAACC-TG	new	pattern
GATA6	2627	T18F0112BFD1	chr18	+	18005969	18004769	18006169	1400	12	3	-2526	891	CC-GGTCT-ACG-TGCCC-AC	new	pattern
SFN	2810	T01F019CD661	chr1	+	27055713	27054513	27055913	1400	10	6	6507	504	GA-GGCCA-TCC-TGTCC-CC	new	pattern
GPX3	2878	T05F08F6BD67	chr5	+	150388071	150386871	150388271	1400	9	4	-7878	573	AG-GGTCT-CAT-TGGCC-CA	new	pattern
GRB7	2886	T17F02184FDB	chr17	+	35147739	35146539	35147939	1400	49	3	22	958	CA-GGTGG-TCG-GGACT-CT	new	pattern
GRB7	2886	T17F02185167	chr17	+	35148135	35146935	35148335	1400	118	21	-364	562	CA-GGTGG-TCG-GGACT-CT	new	pattern
HIC1	3090	T17F001D11FB	chr17	+	1905147	1903947	1905347	1400	215	93	-106	1235	CA-GGACA-GAC-CGACC-GG	new	pattern
HIC1	3090	T17F001D1223	chr17	+	1905187	1903987	1905387	1400	8	4	-169	1195	CA-GGACA-GAC-CGACC-GG	new	pattern
HLA-C	3106	T06R01DF9F64	chr6	-	31432548	31432348	31433748	1400	10	5	-2899	838	CT-GGTCA-TGG-CGCCC-CG	new	pattern
PDX1	3651	T13F01A208FB	chr13	+	27396347	27395147	27396547	1400	8	3	-4167	423	CT-GGCCA-ACG-TGACC-CC	new	pattern
JAG2	3714	T14R063C89CC	chr14	-	104630732	104630532	104631932	1400	9	7	48407	230	CA-GCTCA-GCC-TGCCC-TC	new	pattern
JAG2	3714	T14R063C89CC	chr14	-	104630732	104630532	104631932	1400	9	7	48407	793	CC-AGTCA-AGC-CACCC-CA	new	pattern
CD82	3732	T11F02A6FE32	chr11	+	44498482	44497282	44498682	1400	17	9	45235	116	GT-AGTCA-GCT-TGGCC-TG	new	pattern

KRT13	3860	T17R02333152	chr17	-	36909394	36909194	36910594	1400	58	58	1382	813	GG-AGTCA-TCC-TGGCC-TC	new	pattern
LAMC2	3918	T01F0AD047C6	chr1	+	181422022	181420822	181422222	1400	190	20	79	363	AA-GGTCA-GCC-TGCCC-AA	new	pattern
LOXL2	4017	T08R016224CB	chr8	-	23209163	23208963	23210363	1400	6	5	1211	1128	CA-GGGGA-TGC-TGACC-CA	new	pattern
LOXL2	4017	T08R01628357	chr8	-	23233367	23233167	23234567	1400	15	8	-22988	1169	GA-GGCCA-TGG-TGGTC-TG	new	pattern
LOXL2	4017	T08R01628357	chr8	-	23233367	23233167	23234567	1400	15	8	-22988	1335	CA-GAGCA-AGT-TGCCC-TG	new	pattern
LOXL2	4017	T08R016283A7	chr8	-	23233447	23233247	23234647	1400	5	3	-23069	1249	GA-GGCCA-TGG-TGGTC-TG	new	pattern
LOXL2	4017	T08R016283C8	chr8	-	23233480	23233280	23234680	1400	6	4	-23106	1282	GA-GGCCA-TGG-TGGTC-TG	new	pattern
LOXL2	4017	T08R0162B82E	chr8	-	23246894	23246694	23248094	1400	15	4	-36520	969	GG-CGTCA-TGA-TGCCC-TG	new	pattern
LOXL2	4017	T08R0162B8AD	chr8	-	23247021	23246821	23248221	1400	10	4	-36626	1096	GG-CGTCA-TGA-TGCCC-TG	new	pattern
TACSTD1	4072	T02F02D40719	chr2	+	47449881	47448681	47450081	1400	64	7	118	546	TT-GGTCA-GGC-TGGTC-TC	known	pattern
TACSTD1	4072	T02F02D40762	chr2	+	47449954	47448754	47450154	1400	516	162	21	473	TT-GGTCA-GGC-TGGTC-TC	known	pattern
MDM2	4193	T12F0405CD2F	chr12	+	67489071	67487871	67489271	1400	15	5	-802	986	GA-GGTCC-GGA-TGATC-GC	new	pattern
MGAT5	4249	T02F0807A47B	chr2	+	134718587	134717387	134718787	1400	10	5	7051	78	CC-AGTCA-GGC-TGACC-CA	new	pattern
MGAT5	4249	T02F0807A47B	chr2	+	134718587	134717387	134718787	1400	10	5	7051	642	GG-GGCCA-TTT-TGGCC-TG	new	pattern
MGAT5	4249	T02F08090E08	chr2	+	134811144	134809944	134811344	1400	6	6	-85506	1075	CA-GGCCA-TAT-TGGTC-TC	new	pattern
MKI67	4288	T10R07BC8C6D	chr10	-	129797229	129797029	129798429	1400	22	11	-12296	464	CA-GGGCA-TGG-CAACC-TT	new	pattern
MKI67	4288	T10R07BC8C6D	chr10	-	129797229	129797029	129798429	1400	22	11	-12296	1357	CC-AGTCA-TTA-CAACC-AG	new	pattern
MMP2	4313	T16F03390C37	chr16	+	54070327	54069127	54070527	1400	28	6	286	676	GG-GGTCT-TTG-TGACC-TC	new	pattern
MMP2	4313	T16F03390C5D	chr16	+	54070365	54069165	54070565	1400	234	20	240	638	GG-GGTCT-TTG-TGACC-TC	new	pattern
MMP2	4313	T16F03390D53	chr16	+	54070611	54069411	54070811	1400	7890	1816	77	392	GG-GGTCT-TTG-TGACC-TC	new	pattern
MMP2	4313	T16F03391C2C	chr16	+	54074412	54073212	54074612	1400	6	4	-3823	1214	CA-GGTGA-TCT-TGACC-AG	new	pattern
MSH2	4436	T02F02D48712	chr2	+	47482642	47481442	47482842	1400	19	18	1125	656	TT-AGTCA-GGC-TGGTC-TC	new	pattern
MSH2	4436	T02F02D48B97	chr2	+	47483799	47482599	47483999	1400	543	43	72	1125	GT-GGTCG-CCG-TGGCC-GG	new	pattern
MT3	4504	T16F0349FE8D	chr16	+	55180941	55179741	55181141	1400	56	29	-169	340	TT-GGTCA-GGC-TGGTC-TC	known	pattern
MUC2	4583	T11F000FF0A6	chr11	+	1044646	1043446	1044846	1400	6	3	20229	70	CA-GGGCA-GGC-TGCCT-GG	new	pattern
MUC2	4583	T11F00105A6B	chr11	+	1071723	1070523	1071923	1400	71	18	-6780	596	CA-GGTGA-ACC-TGCCC-CA	new	pattern
MUC2	4583	T11F00105A9E	chr11	+	1071774	1070574	1071974	1400	10	4	-6888	545	CA-GGTGA-ACC-TGCCC-CA	new	pattern

MUC2	4583	T11F00105E29	chr11	+	1072681	1071481	1072881	1400	32	7	-7759	899	AA-GGTGG-GTG-GGACC-CA	new	pattern
MUC2	4583	T11F00106014	chr11	+	1073172	1071972	1073372	1400	68	6	-8188	408	AA-GGTGG-GTG-GGACC-CA	new	pattern
MUC2	4583	T11F0010618D	chr11	+	1073549	1072349	1073749	1400	20	3	-8674	31	AA-GGTGG-GTG-GGACC-CA	new	pattern
MUC2	4583	T11F00106A1D	chr11	+	1075741	1074541	1075941	1400	24	3	-10866	1161	GG-GGTCC-AGC-TGTCC-TG	new	pattern
MUC2	4583	T11F00106AE1	chr11	+	1075937	1074737	1076137	1400	13	3	-11060	965	GG-GGTCC-AGC-TGTCC-TG	new	pattern
MUC2	4583	T11F00106B33	chr11	+	1076019	1074819	1076219	1400	50	9	-11105	883	GG-GGTCC-AGC-TGTCC-TG	new	pattern
MUC2	4583	T11F001096FD	chr11	+	1087229	1086029	1087429	1400	41	12	-22353	1286	TC-CGTCA-GGC-TGCCC-TA	new	pattern
MUC4	4585	T03R0BBDC7A3	chr3	-	196986787	196986587	196987987	1400	6	3	-27425	769	TT-GGTCA-GGC-TGGTC-TC	known	pattern
NFKB1	4790	T04F062D7027	chr4	+	103641127	103639927	103641327	1400	5	4	391	163	TT-GGTCA-GGC-TGGTC-TC	known	pattern
NFKB1	4790	T04F062D70B7	chr4	+	103641271	103640071	103641471	1400	23	7	282	19	TT-GGTCA-GGC-TGGTC-TC	known	pattern
NFKB1	4790	T04F062E9997	chr4	+	103717271	103716071	103717471	1400	5	3	-75750	1138	AA-GGTGG-AGG-GGACC-GG	new	pattern
NOS2A	4843	T17R016143A8	chr17	-	23151528	23151328	23152728	1400	12	8	-43573	868	AA-GATCA-GGT-CACCC-AC	new	pattern
OGG1	4968	T03F009506EC	chr3	+	9766636	9765436	9766836	1400	431	44	-888	418	AA-GGTGA-GAA-TGGCC-CA	new	pattern
OGG1	4968	T03F00950821	chr3	+	9766945	9765745	9767145	1400	136	8	-1224	109	AA-GGTGA-GAA-TGGCC-CA	new	pattern
PECAM1	5175	T17R0390211C	chr17	-	59777308	59777108	59778508	1400	10	3	-23147	648	AA-GGGCA-GGC-TGGCC-TG	new	pattern
PECAM1	5175	T17R0390BECF	chr17	-	59817679	59817479	59818879	1400	7	3	-63511	493	TT-GGTCA-GGC-TGGTC-TC	known	pattern
PPL	5493	T16R004A5922	chr16	-	4872482	4872282	4873682	1400	5	3	49	622	AA-GGACA-AGC-CAACC-AC	new	pattern
MAPK3	5595	T16R01CA66BE	chr16	-	30041790	30041590	30042990	1400	23	4	-8820	154	TT-GGTCA-GGC-TGGTC-TC	known	pattern
PTPN1	5770	T20F02E4F8B1	chr20	+	48560305	48559105	48560505	1400	698	84	7	161	TT-GGTCA-GGC-TGGTC-TC	known	pattern
RELA	5970	T11R03E25271	chr11	-	65163889	65163689	65165089	1400	8	5	14522	909	CC-AGTCA-CCC-TGCCC-TG	new	pattern
RELA	5970	T11R03E29B2D	chr11	-	65182509	65182309	65183709	1400	18	5	-4082	59	GG-GGTGA-GTG-TGGCC-TG	new	pattern
RELA	5970	T11R03E29C52	chr11	-	65182802	65182602	65184002	1400	17	6	-4388	352	GG-GGTGA-GTG-TGGCC-TG	new	pattern
RELA	5970	T11R03E29DFF	chr11	-	65183231	65183031	65184431	1400	9	3	-4813	781	GG-GGTGA-GTG-TGGCC-TG	new	pattern
SDC1	6382	T02R01353CC3	chr2	-	20266179	20265979	20267379	1400	11	3	-2117	363	TT-GGTCC-TAA-TGCCT-GG	new	pattern

SDC1	6382	T02R01353DBC	chr2	-	20266428	20266228	20267628	1400	5	3	-2370	612	TT-GGTCC-TAA-TGCCT-GG	new	pattern
FSCN1	6624	T07F00556EFD	chr7	+	5598973	5597773	5599173	1400	1564	819	21	574	GG-GGTCC-TGG-TGGTC-TT	new	pattern
FSCN1	6624	T07F00556EFD	chr7	+	5598973	5597773	5599173	1400	1564	819	21	917	GG-GGTCG-GGG-TGTCC-GA	new	pattern
SRP54	6729	T14F020EC227	chr14	+	34521639	34520439	34521839	1400	35	6	398	799	CT-GGTCA-ATC-TACCC-CC	new	pattern
SRP54	6729	T14F020EC352	chr14	+	34521938	34520738	34522138	1400	309	42	72	500	CT-GGTCA-ATC-TACCC-CC	new	pattern
STAT3	6774	T17R02400D8B	chr17	-	37752203	37752003	37753403	1400	10	4	-33295	486	TA-GGCCA-TCC-CAGCC-TG	new	pattern
TERF2	7014	T16R040D3FB6	chr16	-	67977142	67976942	67978342	1400	42	10	-30020	631	TT-GGTCA-GGC-TGGTC-TC	known	pattern
TERF2	7014	T16R040D4055	chr16	-	67977301	67977101	67978501	1400	384	19	-30177	790	TT-GGTCA-GGC-TGGTC-TC	known	pattern
TGM2	7052	T20R02283F93	chr20	-	36192147	36191947	36193347	1400	6	4	-1851	1004	GA-GGGCA-GAT-TGTCC-AG	new	pattern
TGM3	7053	T20F00228137	chr20	+	2261303	2260103	2261503	1400	6	4	-36628	704	GA-AGTCA-ACC-TGGTC-CT	new	pattern
TIMP2	7077	T17R046EAB4C	chr17	-	74361676	74361476	74362876	1400	16	4	-1001	690	CA-GGACA-CTA-TGGCC-TG	new	pattern
TIMP2	7077	T17R046EAB7F	chr17	-	74361727	74361527	74362927	1400	20	9	-1053	741	CA-GGACA-CTA-TGGCC-TG	new	pattern
TIMP2	7077	T17R046EAC82	chr17	-	74361986	74361786	74363186	1400	56	20	-1301	1000	CA-GGACA-CTA-TGGCC-TG	new	pattern
TIMP2	7077	T17R046EACAC	chr17	-	74362028	74361828	74363228	1400	8	4	-1354	1042	CA-GGACA-CTA-TGGCC-TG	new	pattern
VHL	7428	T03F009B0605	chr3	+	10159621	10158421	10159821	1400	6	4	-1302	813	TT-GGTCA-GGC-TGGTC-TC	known	pattern
VHL	7428	T03F009B0605	chr3	+	10159621	10158421	10159821	1400	6	4	-1302	1344	AA-GGTGA-TTT-TACCC-CT	new	pattern
EZR	7430	T06R097C10CE	chr6	-	159125710	159125510	159126910	1400	21	4	-18900	614	GC-AGTCA-TGC-CGTCC-TC	new	pattern
VIM	7431	T10F0108240B	chr10	+	17310731	17309531	17310931	1400	10	5	573	190	AA-GGTTA-AGG-TGACC-GA	new	pattern
WNT7B	7477	T22R02A9FF20	chr22	-	44695328	44695128	44696528	1400	11	7	1007	670	CT-GGTCC-TTT-TACCC-TG	new	pattern
XRCC5	7520	T02F0CEA6C29	chr2	+	216689705	216688505	216689905	1400	8	5	-7325	1030	GA-GGTCT-GGT-TGTCC-TG	new	pattern
YES1	7525	T18R000B9C0B	chr18	-	760843	760643	762043	1400	5	4	-49234	862	TT-GGTCA-GGC-TGGTC-TT	new	pattern
PRDM2	7799	T01F00D55012	chr1	+	13979666	13978466	13979866	1400	11	5	-75727	739	AA-AGTCA-CAG-CGACT-CA	new	pattern
RAB11A	8766	T15F03CFD712	chr15	+	63952658	63951458	63952858	1400	5	3	-3808	572	TT-GGTCA-GGC-TGGTC-TC	known	pattern
CRLF1	9244	T19R011B5610	chr19	-	18568720	18568520	18569920	1400	7	4	-3653	161	GA-GGCCA-GCT-CACCC-TG	new	pattern
CRLF1	9244	T19R011B566F	chr19	-	18568815	18568615	18570015	1400	38	6	-3718	256	GA-GGCCA-GCT-CACCC-TG	new	pattern
GPR56	9289	T16F0359B808	chr16	+	56211464	56210264	56211664	1400	32	6	22	824	CA-GGTGA-GTC-TGGCC-CA	new	pattern
GPR56	9289	T16F0359B808	chr16	+	56211464	56210264	56211664	1400	32	6	22	1059	AG-GGTCA-GGT-TGGCC-CC	new	pattern

DLEC1	9940	T03F0245814A	chr3	+	38109514	38108314	38109714	1400	32	7	-53813	1253	CA-AGCCA-GTG-TGGCC-TC	new	pattern
HS3ST2	9956	T16F015AE0B8	chr16	+	22732984	22731784	22733184	1400	248	45	378	854	CA-GGGCA-GGG-TGCCT-GC	new	pattern
HS3ST2	9956	T16F015AE0B8	chr16	+	22732984	22731784	22733184	1400	248	45	378	880	TG-GGTCT-GTC-TGGTC-TG	new	pattern
AKAP13	11214	T15F0503078B	chr15	+	84084619	84083419	84084819	1400	8	6	-359744	834	TA-GGTCA-CTG-TGGTC-TT	new	pattern
GAPDHS	26330	T19F026D32A1	chr19	+	40710817	40709617	40711017	1400	5	3	5387	28	CT-GGCCA-AAC-TGCCC-TG	new	pattern
PRRX2	51450	T09F07D6641E	chr9	+	131490846	131489646	131491046	1400	5	3	-23105	801	TG-AGTCA-CCG-TGCCC-GG	new	pattern
SBNO1	55206	T12R074B39D3	chr12	-	122370515	122370315	122371715	1400	10	8	-24088	58	GA-GATCA-GCC-TGACC-AA	new	pattern
CTDSP1	58190	T02F0D0D4708	chr2	+	218973960	218972760	218974160	1400	15	9	-1238	1123	TG-GGTCT-GGC-TGCCC-CG	new	pattern
08-Mar	220972	T10R02B4E5B9	chr10	-	45409721	45409521	45410921	1400	7	4	-136877	945	TG-GGTCT-CTC-TGCCC-CG	new	pattern
08-Mar	220972	T10R02B4E65E	chr10	-	45409886	45409686	45411086	1400	31	28	-137044	1110	TG-GGTCT-CTC-TGCCC-CG	new	pattern
08-Mar	220972	T10R02B4E6DD	chr10	-	45410013	45409813	45411213	1400	245	26	-137118	1237	TG-GGTCT-CTC-TGCCC-CG	new	pattern
MUC5B	727897	T11F001252E8	chr11	+	1200872	1199672	1201072	1400	920	139	187	61	GA-GGCCA-CAC-CACCC-GA	new	pattern
MUC5B	727897	T11F001252E8	chr11	+	1200872	1199672	1201072	1400	920	139	187	1117	AA-GAGCA-TGG-TGCCC-TG	new	pattern

Category C3: ESCC genes with experimental proof verifying estrogen responsiveness that were not predicted to be estrogen responsive.

Gene Symbol	GeneID	TSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Promotersize	Total Tags	RepTags	Distance:
ACTB	60	T07R005470CD	chr7	-	5533901	5533701	5535101	1400	5	3	-569
ACTB	60	T07R00547397	chr7	-	5534615	5534415	5535815	1400	9	3	-1274
ACTB	60	T07R00547411	chr7	-	5534737	5534537	5535937	1400	9	3	-1391
ACTB	60	T07R005479E9	chr7	-	5536233	5536033	5537433	1400	31	4	-2796
AIF1	199	T06F01E390EE	chr6	+	31690990	31689790	31691190	1400	18	3	39
ALCAM	214	T03F065A1A3D	chr3	+	106568253	106567053	106568453	1400	13	6	159
ALCAM	214	T03F065A1A66	chr3	+	106568294	106567094	106568494	1400	5	4	112
ALCAM	214	T03F065A1AB8	chr3	+	106568376	106567176	106568576	1400	6	4	39
ALCAM	214	T03F065A1AFB	chr3	+	106568443	106567243	106568643	1400	124	37	3
ALCAM	214	T03F065A2237	chr3	+	106570295	106569095	106570495	1400	14	8	-1890

ALCAM	214	T03F065C217E	chr3	+	106701182	106699982	106701382	1400	5	3	-132778
ALCAM	214	T03F065C5C6B	chr3	+	106716267	106715067	106716467	1400	7	5	-147864
ALCAM	214	T03F065C931B	chr3	+	106730267	106729067	106730467	1400	18	14	-161864
ALPP	250	T02F0DE28F99	chr2	+	232951705	232950505	232951905	1400	12	7	-113
ANXA1	301	T09F0477BF39	chr9	+	74956601	74955401	74956801	1400	123	54	12
APBA1	320	T09R0442A6E4	chr9	-	71476964	71476764	71478164	1400	5	3	-241917
APBA1	320	T09R0442A719	chr9	-	71477017	71476817	71478217	1400	136	18	-241958
BIRC5	332	T17F0464E849	chr17	+	73721929	73720729	73722129	1400	49	7	-29
APOD	347	T03R0BBA304E	chr3	-	196751438	196751238	196752638	1400	250	28	25490
APOD	347	T03R0BBABD6B	chr3	-	196787563	196787363	196788763	1400	25	8	-10659
APOD	347	T03R0BBACF24	chr3	-	196792100	196791900	196793300	1400	812	182	-15203
AR	367	T0XF03F976EE	chrX	+	66680558	66679358	66680758	1400	6	3	41
RHOA	387	T03R02F15DF3	chr3	-	49372659	49372459	49373859	1400	48	9	-1015
RHOA	387	T03R02F15E30	chr3	-	49372720	49372520	49373920	1400	21	4	-1097
RHOA	387	T03R02F15E63	chr3	-	49372771	49372571	49373971	1400	46	9	-1155
RHOA	387	T03R02F17E5C	chr3	-	49380956	49380756	49382156	1400	18	3	-9347
RHOA	387	T03R02F19389	chr3	-	49386377	49386177	49387577	1400	5	3	-14759
RHOA	387	T03R02F199FD	chr3	-	49388029	49387829	49389229	1400	8	5	-16399
RHOA	387	T03R02F226CA	chr3	-	49424074	49423874	49425274	1400	21	3	-52405
RND3	390	T02R0900C475	chr2	-	151045237	151045037	151046437	1400	12	6	-12252
RND3	390	T02R0900DCBD	chr2	-	151051453	151051253	151052653	1400	6	5	-18478
ATF3	467	T01F0C908991	chr1	+	210798993	210797793	210799193	1400	82	25	6384
ATF3	467	T01F0C914BD6	chr1	+	210848726	210847526	210848926	1400	228	122	-43347
ATF3	467	T01F0C915BBB	chr1	+	210852795	210851595	210852995	1400	70	69	-47420
ATF3	467	T01F0C916AE7	chr1	+	210856679	210855479	210856879	1400	5	4	-51305
BAD	572	T11R03CD2647	chr11	-	63776327	63776127	63777527	1400	7	4	17569
BAD	572	T11R03CD89BF	chr11	-	63801791	63801591	63802991	1400	13	3	-7894

CCND1	595	T11F041F6031	chr11	+	69165105	69163905	69165305	1400	521	131	43
CCND1	595	T11F041F6956	chr11	+	69167446	69166246	69167646	1400	5	3	-2387
CCND1	595	T11F041F8CE6	chr11	+	69176550	69175350	69176750	1400	12	5	-11475
BUB1	699	T02R06A00BF2	chr2	-	111152114	111151914	111153314	1400	55	10	-40175
DDR1	780	T06F01D85CE7	chr6	+	30956775	30955575	30956975	1400	13	3	3084
DDR1	780	T06F01D85D98	chr6	+	30956952	30955752	30957152	1400	26	13	2888
DDR1	780	T06F01D86475	chr6	+	30958709	30957509	30958909	1400	42	10	1166
DDR1	780	T06F01D8696B	chr6	+	30959979	30958779	30960179	1400	13	3	-80
DDR1	780	T06F01D86ACE	chr6	+	30960334	30959134	30960534	1400	47	8	-434
DDR1	780	T06F01D86B39	chr6	+	30960441	30959241	30960641	1400	20	4	-560
CAV1	857	T07F06E949DA	chr7	+	115952090	115950890	115952290	1400	5	4	0
CAV1	857	T07F06E94AC6	chr7	+	115952326	115951126	115952526	1400	64	29	-206
CAV1	857	T07F06E94B09	chr7	+	115952393	115951193	115952593	1400	6	3	-308
CAV1	857	T07F06E94B6A	chr7	+	115952490	115951290	115952690	1400	60	11	-375
CAV1	857	T07F06E94CA0	chr7	+	115952800	115951600	115953000	1400	108	28	-714
CAV1	857	T07F06E94CDA	chr7	+	115952858	115951658	115953058	1400	5	3	-772
CAV1	857	T07F06E94F56	chr7	+	115953494	115952294	115953694	1400	22	9	-1419
CAV1	857	T07F06E94FF4	chr7	+	115953652	115952452	115953852	1400	130	80	-1566
CCNB1	891	T05F0415352D	chr5	+	68498733	68497533	68498933	1400	243	78	-10
CCNE1	898	T19F02156C4D	chr19	+	34958413	34957213	34958613	1400	10	7	36329
CCNE1	898	T19F0215F99A	chr19	+	34994586	34993386	34994786	1400	5	4	155
CCNE1	898	T19F0215FA36	chr19	+	34994742	34993542	34994942	1400	24	3	60
CCNE1	898	T19F0215FBB3	chr19	+	34995123	34993923	34995323	1400	7	3	-382
CD34	947	T01R0C48F9BB	chr1	-	206109115	206108915	206110315	1400	16	4	17415
CD34	947	T01R0C493BBF	chr1	-	206126015	206125815	206127215	1400	7	3	512
CD34	947	T01R0C499D5A	chr1	-	206151002	206150802	206152202	1400	11	8	-24475
CD34	947	T01R0C499DB5	chr1	-	206151093	206150893	206152293	1400	176	15	-24505

CD44	960	T11F0217D8EE	chr11	+	35117294	35116094	35117494	1400	3679	931	-137
CD44	960	T11F0217D992	chr11	+	35117458	35116258	35117658	1400	23	12	-405
CD44	960	T11F02180E2A	chr11	+	35130922	35129722	35131122	1400	5	4	-13929
CD44	960	T11F02186B57	chr11	+	35154775	35153575	35154975	1400	8	4	-37782
CD44	960	T11F02189F5F	chr11	+	35168095	35166895	35168295	1400	12	3	-51062
CD44	960	T11F0218E813	chr11	+	35186707	35185507	35186907	1400	6	5	-69714
CDH13	1012	T16F04D74A1F	chr16	+	81218079	81216879	81218279	1400	27	11	12
CDH13	1012	T16F04D74A61	chr16	+	81218145	81216945	81218345	1400	25	11	-51
CDH13	1012	T16F04D8D131	chr16	+	81318193	81316993	81318393	1400	5	3	-100108
CDH13	1012	T16F04E0C86A	chr16	+	81840234	81839034	81840434	1400	30	21	-622155
CDH13	1012	T16F04E875E0	chr16	+	82343392	82342192	82343592	1400	25	13	-1125313
CDK2	1017	T12F0341D842	chr12	+	54646850	54645650	54647050	1400	47	7	11
CDKN1B	1027	T12F00C2B999	chr12	+	12761497	12760297	12761697	1400	15	7	79
CDKN1B	1027	T12F00C2B9E8	chr12	+	12761576	12760376	12761776	1400	154	23	38
CDKN1B	1027	T12F00C2C966	chr12	+	12765542	12764342	12765742	1400	6	3	-3966
CKS2	1164	T09F056E519A	chr9	+	91115930	91114730	91116130	1400	339	172	16
COL4A6	1288	T0XR06649710	chrX	-	107255568	107255368	107256768	1400	5	5	29951
KLF6	1316	T10R003A27CE	chr10	-	3811278	3811078	3812478	1400	12	6	-3066
KLF6	1316	T10R003A38F0	chr10	-	3815664	3815464	3816864	1400	8	4	-7456
KLF6	1316	T10R003A3B75	chr10	-	3816309	3816109	3817509	1400	8	4	-8100
KLF6	1316	T10R003A3D7E	chr10	-	3816830	3816630	3818030	1400	40	22	-8622
KLF6	1316	T10R003A3E6C	chr10	-	3817068	3816868	3818268	1400	6	3	-8860
KLF6	1316	T10R003A3E90	chr10	-	3817104	3816904	3818304	1400	7	3	-8896
KLF6	1316	T10R003A3F42	chr10	-	3817282	3817082	3818482	1400	8	3	-9073
CTGF	1490	T06R07E2E8D3	chr6	-	132311251	132311051	132312451	1400	9	3	-997
CTGF	1490	T06R07E2F28C	chr6	-	132313740	132313540	132314940	1400	5	4	-3523
EEF1D	1936	T08R08A0B77E	chr8	-	144750462	144750262	144751662	1400	7	3	-17378

EEF1D	1936	T08R08A0B80F	chr8	-	144750607	144750407	144751807	1400	72	31	-17545
EFNA1	1942	T01F0924302B	chr1	+	153366571	153365371	153366771	1400	16	4	409
EFNA1	1942	T01F092431C0	chr1	+	153366976	153365776	153367176	1400	120	34	26
EGR1	1958	T05F083719B0	chr5	+	137828784	137827584	137828984	1400	19	5	331
EGR1	1958	T05F08371ACB	chr5	+	137829067	137827867	137829267	1400	5861	1380	23
EGR1	1958	T05F08371B5A	chr5	+	137829210	137828010	137829410	1400	32	5	-92
EGR1	1958	T05F08371B9F	chr5	+	137829279	137828079	137829479	1400	7	7	-199
EGR1	1958	T05F08371CB5	chr5	+	137829557	137828357	137829757	1400	105	35	-473
EGR1	1958	T05F08371CFD	chr5	+	137829629	137828429	137829829	1400	66	15	-528
EGR1	1958	T05F08371EF7	chr5	+	137830135	137828935	137830335	1400	15	5	-1034
EGR1	1958	T05F08371F1C	chr5	+	137830172	137828972	137830372	1400	10	3	-1092
EGR1	1958	T05F08371FF7	chr5	+	137830391	137829191	137830591	1400	11	5	-1294
EGR1	1958	T05F08372059	chr5	+	137830489	137829289	137830689	1400	20	4	-1351
EGR1	1958	T05F083721D0	chr5	+	137830864	137829664	137831064	1400	10	3	-1783
EMP1	2012	T12F00C9482E	chr12	+	13191214	13190014	13191414	1400	15	8	49710
EMP1	2012	T12F00C9A571	chr12	+	13215089	13213889	13215289	1400	6	6	25835
EMP1	2012	T12F00CA0A5F	chr12	+	13240927	13239727	13241127	1400	12	4	0
EMP1	2012	T12F00CA0A9B	chr12	+	13240987	13239787	13241187	1400	81	24	-49
CTTN	2017	T11F042AEDFE	chr11	+	69922302	69921102	69922502	1400	299	32	27
CTTN	2017	T11F042B050E	chr11	+	69928206	69927006	69928406	1400	5	4	-5911
CTTN	2017	T11F042B1129	chr11	+	69931305	69930105	69931505	1400	7	6	-9013
CTTN	2017	T11F042B6F0C	chr11	+	69955340	69954140	69955540	1400	8	3	-33048
EPHA4	2043	T02R0D3BA331	chr2	-	222012209	222012009	222013409	1400	12	10	-21201
EPHA4	2043	T02R0D3BA80D	chr2	-	222013453	222013253	222014653	1400	12	8	-22445
EPHA4	2043	T02R0D3C1B54	chr2	-	222042964	222042764	222044164	1400	6	5	-51953
EPHA4	2043	T02R0D3C277C	chr2	-	222046076	222045876	222047276	1400	9	7	-55066
EPHB3	2049	T03F0B1281DB	chr3	+	185762267	185761067	185762467	1400	249	49	47

EPHB3	2049	T03F0B128604	chr3	+	185763332	185762132	185763532	1400	7	6	-1051
ERBB2	2064	T17F02176491	chr17	+	35087505	35086305	35087705	1400	9	4	10414
ERBB2	2064	T17F02178D11	chr17	+	35097873	35096673	35098073	1400	25	6	62
ERBB2	2064	T17F02178D3C	chr17	+	35097916	35096716	35098116	1400	20	5	13
ERBB2	2064	T17F0217BAF1	chr17	+	35109617	35108417	35109817	1400	8	3	-11686
ERBB2	2064	T17F0217BB7B	chr17	+	35109755	35108555	35109955	1400	452	41	-11768
ERBB2	2064	T17F02181C78	chr17	+	35134584	35133384	35134784	1400	517	14	-36660
EYA4	2070	T06F07F6A353	chr6	+	133604179	133602979	133604379	1400	22	5	27
EYA4	2070	T06F07F71E52	chr6	+	133635666	133634466	133635866	1400	21	8	-31456
EYA4	2070	T06F07F71E6A	chr6	+	133635690	133634490	133635890	1400	5	3	-31484
EYA4	2070	T06F07F71EB8	chr6	+	133635768	133634568	133635968	1400	1606	213	-31557
EYA4	2070	T06F07F87E5E	chr6	+	133725790	133724590	133725990	1400	66	31	-121584
ESR1	2099	T06F0911F00D	chr6	+	152170509	152169309	152170709	1400	9	3	-128
ESR1	2099	T06F0911F2B8	chr6	+	152171192	152169992	152171392	1400	5	4	-813
FDFT1	2222	T08F00B27E19	chr8	+	11697689	11696489	11697889	1400	472	42	6
FDFT1	2222	T08F00B2944F	chr8	+	11703375	11702175	11703575	1400	39	9	-5744
FGF3	2248	T11R042213BE	chr11	-	69342142	69341942	69343342	1400	148	73	-8208
FGF3	2248	T11R042214B5	chr11	-	69342389	69342189	69343589	1400	7	5	-8454
FLT4	2324	T05R0AB9D601	chr5	-	179951105	179950905	179952305	1400	5	3	-5275
FLT4	2324	T05R0ABA043F	chr5	-	179962943	179962743	179964143	1400	5	3	-17113
FLT4	2324	T05R0ABAB71B	chr5	-	180008731	180008531	180009931	1400	10	9	-62899
FOS	2353	T14F0475906E	chr14	+	74813550	74812350	74813750	1400	20	12	1735
FOS	2353	T14F04759734	chr14	+	74815284	74814084	74815484	1400	136	64	5
G6PD	2539	T0XR0924A9F6	chrX	-	153397750	153397550	153398950	1400	29	13	15076
G6PD	2539	T0XR09251E56	chrX	-	153427542	153427342	153428742	1400	8	5	-14723
G6PD	2539	T0XR0925211F	chrX	-	153428255	153428055	153429455	1400	2810	815	-15377
G6PD	2539	T0XR09252298	chrX	-	153428632	153428432	153429832	1400	31	6	-15798

XRCC6	2547	T22F026703EA	chr22	+	40305642	40304442	40305842	1400	6	4	41599
XRCC6	2547	T22F02674663	chr22	+	40322659	40321459	40322859	1400	20	15	24583
XRCC6	2547	T22F0267A678	chr22	+	40347256	40346056	40347456	1400	312	29	35
XRCC6	2547	T22F02680B29	chr22	+	40373033	40371833	40373233	1400	17	9	-25791
GLI1	2735	T12F0358A1E4	chr12	+	56140260	56139060	56140460	1400	7	6	-59
GLI1	2735	T12F0358A820	chr12	+	56141856	56140656	56142056	1400	5	3	-1655
GLI1	2735	T12F0358CD4C	chr12	+	56151372	56150172	56151572	1400	6	6	-11171
HGF	3082	T07R04D68E57	chr7	-	81170007	81169807	81171207	1400	11	7	-607
HIF1A	3091	T14F03A652F4	chr14	+	61231860	61230660	61232060	1400	9	4	161
HIF1A	3091	T14F03A65377	chr14	+	61231991	61230791	61232191	1400	495	78	73
FOXA1	3169	T14R02367B73	chr14	-	37124979	37124779	37126179	1400	5	4	3979
FOXA1	3169	T14R023692AC	chr14	-	37130924	37130724	37132124	1400	79	19	-1964
FOXA1	3169	T14R02369EA1	chr14	-	37133985	37133785	37135185	1400	86	4	-4957
FOXA1	3169	T14R02369F72	chr14	-	37134194	37133994	37135394	1400	225	73	-5217
HOXA7	3204	T07R019E59D8	chr7	-	27154904	27154704	27156104	1400	142	92	4977
HOXA7	3204	T07R019E6266	chr7	-	27157094	27156894	27158294	1400	19	9	2798
HOXA7	3204	T07R019E68C7	chr7	-	27158727	27158527	27159927	1400	26	8	1154
HOXA7	3204	T07R019E763A	chr7	-	27162170	27161970	27163370	1400	11	3	-2252
HOXA7	3204	T07R019E7704	chr7	-	27162372	27162172	27163572	1400	5	5	-2493
HOXA7	3204	T07R019E77C9	chr7	-	27162569	27162369	27163769	1400	10	3	-2690
HOXA7	3204	T07R019E7867	chr7	-	27162727	27162527	27163927	1400	6	4	-2846
HOXA10	3206	T07R019EAB59	chr7	-	27175769	27175569	27176969	1400	5	4	986
HOXA10	3206	T07R019ED342	chr7	-	27185986	27185786	27187186	1400	12	5	-9232
HOXA13	3209	T07R019EE8A6	chr7	-	27191462	27191262	27192662	1400	6	3	11581
HOXA13	3209	T07R019EF42B	chr7	-	27194411	27194211	27195611	1400	5	3	8630
HOXA13	3209	T07R019F1C57	chr7	-	27204695	27204495	27205895	1400	5	5	-1654
HOXA13	3209	T07R019F21D0	chr7	-	27206096	27205896	27207296	1400	7	3	-3053

HOXC8	3224	T12F0323F82C	chr12	+	52688940	52687740	52689140	1400	6	3	226
HOXC8	3224	T12F0323F8A8	chr12	+	52689064	52687864	52689264	1400	57	6	163
HOXD9	3235	T02F0A8825ED	chr2	+	176694765	176693565	176694965	1400	9	8	569
HOXD9	3235	T02F0A8826B6	chr2	+	176694966	176693766	176695166	1400	5	4	369
HOXD9	3235	T02F0A8826E1	chr2	+	176695009	176693809	176695209	1400	5	3	325
HOXD10	3236	T02F0A87FC88	chr2	+	176684168	176682968	176684368	1400	6	4	5572
HOXD10	3236	T02F0A87FECB	chr2	+	176684747	176683547	176684947	1400	7	3	4992
HOXD10	3236	T02F0A880D8C	chr2	+	176688524	176687324	176688724	1400	5	3	1214
HSD17B4	3295	T05F0714FD5B	chr5	+	118816091	118814891	118816291	1400	123	8	55
HSPB1	3315	T07F04842846	chr7	+	75769926	75768726	75770126	1400	7340	2898	82
HSPB1	3315	T07F048428C9	chr7	+	75770057	75768857	75770257	1400	6	4	-198
HSPB1	3315	T07F04842947	chr7	+	75770183	75768983	75770383	1400	13	4	-296
HSPB1	3315	T07F04842962	chr7	+	75770210	75769010	75770410	1400	9	6	-345
HSPB1	3315	T07F04842BB6	chr7	+	75770806	75769606	75771006	1400	44	6	-943
HSPB1	3315	T07F04842C4C	chr7	+	75770956	75769756	75771156	1400	66	19	-1070
IGFBP3	3486	T07R02BC79E0	chr7	-	45906400	45906200	45907600	1400	17	12	11986
IL6ST	3572	T05R034B60FB	chr5	-	55271675	55271475	55272875	1400	5	3	796
IL6ST	3572	T05R034C36C7	chr5	-	55326407	55326207	55327607	1400	10	3	-53918
IL6ST	3572	T05R034C3715	chr5	-	55326485	55326285	55327685	1400	26	6	-54000
IRF1	3659	T05R07DB6C78	chr5	-	131820664	131820464	131821864	1400	5	3	26032
JUN	3725	T01R0384964B	chr1	-	59020875	59020675	59022075	1400	60	48	-1805
JUN	3725	T01R03849996	chr1	-	59021718	59021518	59022918	1400	8	7	-2645
JUN	3725	T01R03849BDF	chr1	-	59022303	59022103	59023503	1400	788	187	-3191
KPNA2	3838	T17F03C7B6FE	chr17	+	63420158	63418958	63420358	1400	6	5	42159
KPNA2	3838	T17F03C7B83F	chr17	+	63420479	63419279	63420679	1400	8	4	41831
KPNA2	3838	T17F03C85BA9	chr17	+	63462313	63461113	63462513	1400	155	21	42
KRAS	3845	T12R0181F76E	chr12	-	25294702	25294502	25295902	1400	6	3	-45230

STMN1	3925	T01R018E5699	chr1	-	26105497	26105297	26106697	1400	241	110	-6282
TACSTD2	4070	T01R03816902	chr1	-	58812674	58812474	58813874	1400	8	3	574
TACSTD2	4070	T01R03817418	chr1	-	58815512	58815312	58816712	1400	140	34	-2244
TACSTD2	4070	T01R038174CD	chr1	-	58815693	58815493	58816893	1400	5	3	-2442
TACSTD2	4070	T01R03817519	chr1	-	58815769	58815569	58816969	1400	7	6	-2520
MAD2L1	4085	T04R07397A5D	chr4	-	121207389	121207189	121208589	1400	30	10	-7296
MCAM	4162	T11R071318F9	chr11	-	118692089	118691889	118693289	1400	6	4	-7625
MDK	4192	T11F02C3619A	chr11	+	46358938	46357738	46359138	1400	9	3	281
MDK	4192	T11F02C362BD	chr11	+	46359229	46358029	46359429	1400	18	3	29
MDK	4192	T11F02C36521	chr11	+	46359841	46358641	46360041	1400	777	114	-569
MLH1	4292	T03F0234BA6C	chr3	+	37010028	37008828	37010228	1400	22	3	1
MMP1	4312	T11R06170D96	chr11	-	102174102	102173902	102175302	1400	31	11	-8218
MUC1	4582	T01R0924F926	chr1	-	153418022	153417822	153419222	1400	8	3	6930
MYC	4609	T08F07ACFDA7	chr8	+	128777639	128776439	128777839	1400	8	8	39859
MYC	4609	T08F07AD9959	chr8	+	128817497	128816297	128817697	1400	68	23	2
MYC	4609	T08F07AD99FB	chr8	+	128817659	128816459	128817859	1400	76	18	-161
MYC	4609	T08F07AD9FF8	chr8	+	128819192	128817992	128819392	1400	12	3	-1675
MYL6	4637	T12F0344C48A	chr12	+	54838410	54837210	54838610	1400	406	186	-11
NDN	4692	T15R01473600	chr15	-	21444096	21443896	21445296	1400	199	71	37576
NDUFB9	4715	T08F077CD141	chr8	+	125620545	125619345	125620745	1400	142	11	1
NFKBIA	4792	T14R02153258	chr14	-	34943576	34943376	34944776	1400	6	6	-3090
NME1	4830	T17F02C6D871	chr17	+	46585969	46584769	46586169	1400	546	55	4
NTRK1	4914	T01F093DCE48	chr1	+	155045448	155044248	155045648	1400	7	6	6718
NTRK1	4914	T01F093E98BF	chr1	+	155097279	155096079	155097479	1400	10	3	-45113
ODC1	4953	T02R009FAB95	chr2	-	10464149	10463949	10465349	1400	5	4	33827
PCNA	5111	T20R004D091C	chr20	-	5048604	5048404	5049804	1400	168	46	-4954
PDGFB	5155	T22R02435B39	chr22	-	37968697	37968497	37969897	1400	25	23	-19009

PDGFB	5155	T22R02436009	chr22	-	37969929	37969729	37971129	1400	26	9	-20233
PDGFB	5155	T22R024363F5	chr22	-	37970933	37970733	37972133	1400	86	31	-21227
PIK3CA	5290	T03F0ABFE269	chr3	+	180347497	180346297	180347697	1400	7	4	1510
PIK3CA	5290	T03F0ABFE7A7	chr3	+	180348839	180347639	180349039	1400	17	9	175
PIK3CA	5290	T03F0ABFE7CB	chr3	+	180348875	180347675	180349075	1400	11	7	139
PITX1	5307	T05R0802BE53	chr5	-	134397523	134397323	134398723	1400	35	11	-6136
PITX1	5307	T05R0802BF5D	chr5	-	134397789	134397589	134398989	1400	108	14	-6397
PLK1	5347	T16F0168128F	chr16	+	23597711	23596511	23597911	1400	203	17	58
PRRX1	5396	T01F0A1130CC	chr1	+	168898764	168897564	168898964	1400	36	4	1279
PRRX1	5396	T01F0A11316F	chr1	+	168898927	168897727	168899127	1400	124	25	1062
PRRX1	5396	T01F0A1135A5	chr1	+	168900005	168898805	168900205	1400	18	3	-68
PPARG	5468	T03F00BBBF6E	chr3	+	12304238	12303038	12304438	1400	8	3	142
PPARG	5468	T03F00BBC027	chr3	+	12304423	12303223	12304623	1400	181	24	16
PPARG	5468	T03F00BBC090	chr3	+	12304528	12303328	12304728	1400	7	3	-160
PPARG	5468	T03F00BBC165	chr3	+	12304741	12303541	12304941	1400	12	8	-382
PPARG	5468	T03F00BBC2A3	chr3	+	12305059	12303859	12305259	1400	6	3	-700
PPARG	5468	T03F00BBC4A0	chr3	+	12305568	12304368	12305768	1400	127	49	-1158
PRKDC	5591	T08R02EAC363	chr8	-	48939875	48939675	48941075	1400	6	3	-91633
PTEN	5728	T10F05575F54	chr10	+	89612116	89610916	89612316	1400	7	4	1077
PTEN	5728	T10F0557600C	chr10	+	89612300	89611100	89612500	1400	104	19	936
PTEN	5728	T10F055762C7	chr10	+	89612999	89611799	89613199	1400	5	3	190
PTEN	5728	T10F05576323	chr10	+	89613091	89611891	89613291	1400	26	3	94
PTEN	5728	T10F05576431	chr10	+	89613361	89612161	89613561	1400	3421	275	7
PTEN	5728	T10F05576488	chr10	+	89613448	89612248	89613648	1400	21	3	-262
PTEN	5728	T10F05585AA0	chr10	+	89676448	89675248	89676648	1400	8	3	-63268
PTGS1	5742	T09F0766B959	chr9	+	124172633	124171433	124172833	1400	27	6	439
PTGS1	5742	T09F0766BB44	chr9	+	124173124	124171924	124173324	1400	454	107	-21

PTGS1	5742	T09F0766CBF3	chr9	+	124177395	124176195	124177595	1400	7	4	-4333
PTGS1	5742	T09F0766D679	chr9	+	124180089	124178889	124180289	1400	6	5	-7039
PTGS2	5743	T01R0B04FE28	chr1	-	184876584	184876384	184877784	1400	6	3	31027
PTK2	5747	T08R0873A9DA	chr8	-	141797850	141797650	141799050	1400	27	5	-60096
PTK2	5747	T08R08743F41	chr8	-	141836097	141835897	141837297	1400	7	5	-98396
PTK2	5747	T08R0874EA13	chr8	-	141879827	141879627	141881027	1400	12	3	-142092
PTK2	5747	T08R087582EE	chr8	-	141918958	141918758	141920158	1400	5	3	-181254
PTK2	5747	T08R08759CCC	chr8	-	141925580	141925380	141926780	1400	15	3	-187876
PTK2	5747	T08R0875F5EC	chr8	-	141948396	141948196	141949596	1400	10	9	-210693
PTK2	5747	T08R08761E85	chr8	-	141958789	141958589	141959989	1400	21	4	-221085
PTK2	5747	T08R08763BBD	chr8	-	141966269	141966069	141967469	1400	9	3	-228566
PTK2	5747	T08R0876C15D	chr8	-	142000477	142000277	142001677	1400	15	3	-262751
PTK2	5747	T08R0877F91D	chr8	-	142080285	142080085	142081485	1400	86	14	-342506
PTK2	5747	T08R0877F993	chr8	-	142080403	142080203	142081603	1400	8	3	-342683
RAF1	5894	T03R00C05A1D	chr3	-	12605981	12605781	12607181	1400	5	3	-5853
RAF1	5894	T03R00C06818	chr3	-	12609560	12609360	12610760	1400	12	12	-9432
RAF1	5894	T03R00C08240	chr3	-	12616256	12616056	12617456	1400	8	3	-16124
RAF1	5894	T03R00C090CD	chr3	-	12619981	12619781	12621181	1400	11	7	-19851
RAF1	5894	T03R00C090E9	chr3	-	12620009	12619809	12621209	1400	5	3	-19880
RAF1	5894	T03R00C0935A	chr3	-	12620634	12620434	12621834	1400	5	3	-20506
RAF1	5894	T03R00C0B214	chr3	-	12628500	12628300	12629700	1400	28	9	-28372
RAF1	5894	T03R00C0B241	chr3	-	12628545	12628345	12629745	1400	5	4	-28418
RAF1	5894	T03R00C17BC8	chr3	-	12680136	12679936	12681336	1400	9	4	-80006
RB1	5925	T13F02D90090	chr13	+	47775888	47774688	47776088	1400	736	87	74
RB1	5925	T13F02D99872	chr13	+	47814770	47813570	47814970	1400	10	5	-38856
RB1	5925	T13F02D9DC5B	chr13	+	47832155	47830955	47832355	1400	11	4	-56242
RB1	5925	T13F02D9E777	chr13	+	47834999	47833799	47835199	1400	10	3	-59083

RB1	5925	T13F02DA1EB7	chr13	+	47849143	47847943	47849343	1400	5	3	-73231
RB1	5925	T13F02DA4EE1	chr13	+	47861473	47860273	47861673	1400	8	4	-85561
RPL7	6129	T08R046EA732	chr8	-	74360626	74360426	74361826	1400	5	3	4819
S100A1	6271	T01F090D505B	chr1	+	151867483	151866283	151867683	1400	13	3	28
S100A1	6271	T01F090D5092	chr1	+	151867538	151866338	151867738	1400	48	20	-34
S100P	6286	T04F0066F188	chr4	+	6746504	6745304	6746704	1400	198	66	-36
CCL5	6352	T17R01DC8DF5	chr17	-	31231477	31231277	31232677	1400	158	23	-8845
SFRP1	6422	T08R02754CBD	chr8	-	41241789	41241589	41242989	1400	12	3	-3135
SFRP1	6422	T08R0275F977	chr8	-	41286007	41285807	41287207	1400	12	3	-47336
SFRP2	6423	T04R093C09C4	chr4	-	154929604	154929404	154930804	1400	9	6	-8387
SFRP4	6424	T07R0241F462	chr7	-	37876834	37876634	37878034	1400	7	7	35252
SLN	6588	T11R06613608	chr11	-	107034120	107033920	107035320	1400	6	6	49216
TAGLN	6876	T11F06F2CC43	chr11	+	116575299	116574099	116575499	1400	623	184	-43
TAGLN	6876	T11F06F2DF86	chr11	+	116580230	116579030	116580430	1400	8	3	-4980
NR2F2	7026	T15F05A47F2D	chr15	+	94666541	94665341	94666741	1400	5	4	8409
NR2F2	7026	T15F05A48D39	chr15	+	94670137	94668937	94670337	1400	6	3	4821
NR2F2	7026	T15F05A49F6B	chr15	+	94674795	94673595	94674995	1400	120	26	157
NR2F2	7026	T15F05A49FF4	chr15	+	94674932	94673732	94675132	1400	230	29	48
NR2F2	7026	T15F05A4A034	chr15	+	94674996	94673796	94675196	1400	66	36	-41
NR2F2	7026	T15F05A4A04E	chr15	+	94675022	94673822	94675222	1400	76	27	-68
NR2F2	7026	T15F05A4A09F	chr15	+	94675103	94673903	94675303	1400	6	3	-148
NR2F2	7026	T15F05A4A11A	chr15	+	94675226	94674026	94675426	1400	6	3	-276
NR2F2	7026	T15F05A4A136	chr15	+	94675254	94674054	94675454	1400	13	6	-304
NR2F2	7026	T15F05A4A9FA	chr15	+	94677498	94676298	94677698	1400	13	6	-2546
TFRC	7037	T03R0BC205B6	chr3	-	197264822	197264622	197266022	1400	12	3	-4025
TFRC	7037	T03R0BC234B9	chr3	-	197276857	197276657	197278057	1400	7	3	-16093
TFRC	7037	T03R0BC24B8B	chr3	-	197282699	197282499	197283899	1400	19	7	-21934

TFRC	7037	T03R0BC2742B	chr3	-	197293099	197292899	197294299	1400	5	5	-32337
TGFA	7039	T02R04342F2A	chr2	-	70528810	70528610	70530010	1400	10	7	-869
TGFB1	7040	T19R02C65141	chr19	-	46551361	46551161	46552561	1400	400	44	-22746
TGFB1	7040	T19R02C651D4	chr19	-	46551508	46551308	46552708	1400	82	7	-22944
TGFBI	7045	T05F0811EC9C	chr5	+	135392412	135391212	135392612	1400	7	4	185
TGFBI	7045	T05F0811ED43	chr5	+	135392579	135391379	135392779	1400	3031	1092	94
TGFBI	7045	T05F0811FFEE	chr5	+	135397358	135396158	135397558	1400	7	3	-4759
TGFBI	7045	T05F0812316A	chr5	+	135410026	135408826	135410226	1400	9	3	-17398
TGFBR1	7046	T09F0603B884	chr9	+	100907140	100905940	100907340	1400	34	5	132
TGFBR1	7046	T09F0603B8D3	chr9	+	100907219	100906019	100907419	1400	1970	219	51
TGFBR1	7046	T09F0603BA32	chr9	+	100907570	100906370	100907770	1400	5	3	-337
TIMP1	7076	T0XF02D2255B	chrX	+	47326555	47325355	47326755	1400	17	5	79
TIMP1	7076	T0XF02D2263B	chrX	+	47326779	47325579	47326979	1400	930	109	3
TIMP3	7078	T22F01E1130B	chr22	+	31527691	31526491	31527891	1400	13867	7359	-798
TIMP3	7078	T22F01E11350	chr22	+	31527760	31526560	31527960	1400	7	4	-958
TIMP3	7078	T22F01E116B3	chr22	+	31528627	31527427	31528827	1400	63	46	-1825
TIMP3	7078	T22F01E11AD5	chr22	+	31529685	31528485	31529885	1400	53	27	-2880
TIMP3	7078	T22F01E1B404	chr22	+	31568900	31567700	31569100	1400	5	3	-42076
TOP2A	7153	T17R022294CD	chr17	-	35820749	35820549	35821949	1400	7	3	-22407
TP53	7157	T17R0072EC0A	chr17	-	7531530	7531330	7532730	1400	25	6	-19037
TP53	7157	T17R0072EC40	chr17	-	7531584	7531384	7532784	1400	14	4	-19098
HSP90B1	7184	T12F06215701	chr12	+	102848257	102847057	102848457	1400	14	3	89
HSP90B1	7184	T12F0621573E	chr12	+	102848318	102847118	102848518	1400	5759	909	42
HSP90B1	7184	T12F06215EA2	chr12	+	102850210	102849010	102850410	1400	6	4	-1876
HSP90B1	7184	T12F06216523	chr12	+	102851875	102850675	102852075	1400	8	4	-3554
HSP90B1	7184	T12F062165C9	chr12	+	102852041	102850841	102852241	1400	12	3	-3708
HSP90B1	7184	T12F06217641	chr12	+	102856257	102855057	102856457	1400	14	3	-7938

HSP90B1	7184	T12F0621825A	chr12	+	102859354	102858154	102859554	1400	5	3	-11016
HSP90B1	7184	T12F06218B83	chr12	+	102861699	102860499	102861899	1400	17	5	-13326
HSP90B1	7284	T16R01B692C4	chr16	-	28742340	28742140	28743540	1400	25	11	19301
HSP90B1	7284	T16R01B69315	chr16	-	28742421	28742221	28743621	1400	6	4	19214
HSP90B1	7284	T16R01B6945B	chr16	-	28742747	28742547	28743947	1400	19	15	18896
TWIST1	7291	T07R0123CE48	chr7	-	19123784	19123584	19124984	1400	44	21	-2149
TYMS	7298	T18F0009E1C0	chr18	+	647616	646416	647816	1400	1409	122	189
TYMS	7298	T18F0009E1F4	chr18	+	647668	646468	647868	1400	40	8	-16
TYMS	7298	T18F0009E21D	chr18	+	647709	646509	647909	1400	600	91	-44
TYMS	7298	T18F0009E316	chr18	+	647958	646758	648158	1400	5	3	-307
TYMS	7298	T18F0009E44F	chr18	+	648271	647071	648471	1400	5	4	-613
VCAM1	7412	T01F06047EBE	chr1	+	100957886	100956686	100958086	1400	70	17	4
VDR	7421	T12R02C55098	chr12	-	46485656	46485456	46486856	1400	8	7	35951
VDR	7421	T12R02C56D9B	chr12	-	46493083	46492883	46494283	1400	9	4	28522
VDR	7421	T12R02C58860	chr12	-	46499936	46499736	46501136	1400	255	22	21706
VDR	7421	T12R02C5891F	chr12	-	46500127	46499927	46501327	1400	21	13	21480
VDR	7421	T12R02C5A591	chr12	-	46507409	46507209	46508609	1400	9	4	14196
VDR	7421	T12R02C67E07	chr12	-	46562823	46562623	46564023	1400	9	4	-41204
VDR	7421	T12R02C6D37D	chr12	-	46584701	46584501	46585901	1400	15	5	-63080
VSNL1	7447	T02F010C552E	chr2	+	17585454	17584254	17585654	1400	244	28	-67
VSNL1	7447	T02F010DFE47	chr2	+	17694279	17693079	17694479	1400	12	3	-108989
WNT10B	7480	T12R02D6E495	chr12	-	47637653	47637453	47638853	1400	220	50	7780
WNT9A	7483	T01R0D7ACD43	chr1	-	226151747	226151547	226152947	1400	6	5	23590
WNT9A	7483	T01R0D7B2DAB	chr1	-	226176427	226176227	226177627	1400	7	4	-1089
WNT9A	7483	T01R0D7B3252	chr1	-	226177618	226177418	226178818	1400	37	13	-2281
WNT9A	7483	T01R0D7B8E5E	chr1	-	226201182	226200982	226202382	1400	15	6	-25844
XRCC1	7515	T19R02E78D8E	chr19	-	48729486	48729286	48730686	1400	6	3	9836

CXCR4	7852	T02R0823A48D	chr2	-	136553613	136553413	136554813	1400	6	4	34795
CXCR4	7852	T02R08243B3D	chr2	-	136592189	136591989	136593389	1400	32	18	-3776
PDHX	8050	T11F0213F6E7	chr11	+	34862823	34861623	34863023	1400	7	6	31918
PDHX	8050	T11F021472FC	chr11	+	34894588	34893388	34894788	1400	19	6	202
PDHX	8050	T11F02147395	chr11	+	34894741	34893541	34894941	1400	542	54	52
PDHX	8050	T11F02155D85	chr11	+	34954629	34953429	34954829	1400	6	4	-59888
FOSL1	8061	T11R03E62FF5	chr11	-	65417205	65417005	65418405	1400	15	4	-914
FOSL1	8061	T11R03E633A6	chr11	-	65418150	65417950	65419350	1400	8	3	-1844
FOSL1	8061	T11R03E64C33	chr11	-	65424435	65424235	65425635	1400	1586	261	-8090
NCOA3	8202	T20F02B740B7	chr20	+	45564087	45562887	45564287	1400	103	14	82
NCOA3	8202	T20F02B977A8	chr20	+	45709224	45708024	45709424	1400	7	3	-145160
SOCS1	8651	T16R00AB8E3C	chr16	-	11243068	11242868	11244268	1400	5	4	12724
SOCS1	8651	T16R00AB9FD7	chr16	-	11247575	11247375	11248775	1400	7	6	8217
HYAL2	8692	T03R030008D4	chr3	-	50333908	50333708	50335108	1400	153	47	-3587
TNFSF10	8743	T03R0A5AD10A	chr3	-	173723914	173723714	173725114	1400	16	13	-17737
TNFSF10	8743	T03R0A5AD12F	chr3	-	173723951	173723751	173725151	1400	14	4	-17761
SOCS3	9021	T17R04671CEB	chr17	-	73866475	73866275	73867675	1400	24	8	-1970
SOCS3	9021	T17R04671DF8	chr17	-	73866744	73866544	73867944	1400	10	5	-2247
SOCS3	9021	T17R04672104	chr17	-	73867524	73867324	73868724	1400	11	4	-3027
SOCS3	9021	T17R04672152	chr17	-	73867602	73867402	73868802	1400	7	3	-3119
RPL14	9045	T03F026994EA	chr3	+	40473834	40472634	40474034	1400	1062	423	29
EBAG9	9166	T08F0697F1B3	chr8	+	110621107	110619907	110621307	1400	21	5	4
EBAG9	9166	T08F0697F200	chr8	+	110621184	110619984	110621384	1400	10	4	-55
EBAG9	9166	T08F0697F247	chr8	+	110621255	110620055	110621455	1400	26	6	-150
PTTG1	9232	T05F0986124B	chr5	+	159781451	159780251	159781651	1400	24	3	3
KLF4	9314	T09R0683A318	chr9	-	109290264	109290064	109291464	1400	29	4	-3271
KLF4	9314	T09R0683A3D8	chr9	-	109290456	109290256	109291656	1400	6	3	-3481

KLF4	9314	T09R0683A436	chr9	-	109290550	109290350	109291750	1400	40	4	-3553
KLF4	9314	T09R0683A6D6	chr9	-	109291222	109291022	109292422	1400	6	4	-4246
IKBKE	9641	T01F0C33A22A	chr1	+	204710442	204709242	204710642	1400	174	35	23
IKBKE	9641	T01F0C3403C0	chr1	+	204735424	204734224	204735624	1400	6	6	-25005
TUSC4	10641	T03R03007B32	chr3	-	50363186	50362986	50364386	1400	195	11	-3209
NUDT6	11162	T04R07650DCD	chr4	-	124063181	124062981	124064381	1400	27	3	-29898
CCT5	22948	T05F009D37D1	chr5	+	10303441	10302241	10303641	1400	297	14	-73
ANGPTL2	23452	T09R07AE6608	chr9	-	128869896	128869696	128871096	1400	15	3	19573
ANGPTL2	23452	T09R07AF0377	chr9	-	128910199	128909999	128911399	1400	23	6	-20730
ANGPTL2	23452	T09R07AF0425	chr9	-	128910373	128910173	128911573	1400	6	3	-20903
ANGPTL2	23452	T09R07AF05D9	chr9	-	128910809	128910609	128912009	1400	6	5	-21340
TMEFF2	23671	T02R0B7D6687	chr2	-	192767623	192767423	192768823	1400	33	15	-244615
STOML2	30968	T09R021779C1	chr9	-	35092929	35092729	35094129	1400	19	4	-2977
ACP6	51205	T01R08ADD126	chr1	-	145608998	145608798	145610198	1400	5	3	-23177
LRP1B	53353	T02R08623F39	chr2	-	140656441	140656241	140657641	1400	5	4	49043
LRP1B	53353	T02R0862D8C3	chr2	-	140695747	140695547	140696947	1400	11	7	9739
LRP1B	53353	T02R0862DCDA	chr2	-	140696794	140696594	140697994	1400	10	9	8692
LRP1B	53353	T02R08636480	chr2	-	140731520	140731320	140732720	1400	6	6	-26036
LRP1B	53353	T02R0864549B	chr2	-	140792987	140792787	140794187	1400	8	6	-87501
LRP1B	53353	T02R087EF663	chr2	-	142538339	142538139	142539539	1400	8	5	-1832855
LRP1B	53353	T02R087FFAD1	chr2	-	142605009	142604809	142606209	1400	49	14	-1899507
FXYD5	53827	T19F02677F7E	chr19	+	40337278	40336078	40337478	1400	6	3	210
FXYD5	53827	T19F02678046	chr19	+	40337478	40336278	40337678	1400	395	45	12
FXYD5	53827	T19F0267822B	chr19	+	40337963	40336763	40338163	1400	21	3	-470
ADAMTS9	56999	T03R03D7EE55	chr3	-	64482901	64482701	64484101	1400	5	3	-6512
ADAMTS9	56999	T03R03D8566A	chr3	-	64509546	64509346	64510746	1400	15	4	-33154
ADAMTS9	56999	T03R03DA03BD	chr3	-	64619453	64619253	64620653	1400	29	10	-143056

ADAMTS9	56999	T03R03DA5C05	chr3	-	64642053	64641853	64643253	1400	44	21	-165661
ADAMTS9	56999	T03R03DA70E9	chr3	-	64647401	64647201	64648601	1400	7	4	-171012
ADAMTS9	56999	T03R03DA74A0	chr3	-	64648352	64648152	64649552	1400	31	4	-171920
VCPIP1	80124	T08R040929BC	chr8	-	67709372	67709172	67710572	1400	7	3	-4279
VCPIP1	80124	T08R0409A7CD	chr8	-	67741645	67741445	67742845	1400	6	4	-36571
TUBB	203068	T06F01D5E951	chr6	+	30796113	30794913	30796313	1400	2205	629	159

Category C4: ESCC genes lacking experimental proof verifying estrogen responsiveness that were not predicted to be estrogen responsive.

Gene Symbol	GeneID	TSS	Chromosome	Strand	TSSStart	SequenceStart	SequenceStop	Promotersize	Total Tags	RepTags	Distance:
ACAN	176	T15F0531C4C0	chr15	+	87147712	87146512	87147912	1400	66	18	304
ALOX5AP	241	T13F01CCEEC0	chr13	+	30207680	30206480	30207880	1400	18	3	6
ALOX5AP	241	T13F01CCF29A	chr13	+	30208666	30207466	30208866	1400	11	8	-995
FAS	355	T10F0568968C	chr10	+	90740364	90739164	90740564	1400	38	4	-85
FAS	355	T10F05689727	chr10	+	90740519	90739319	90740719	1400	60	4	-251
FAS	355	T10F056897C7	chr10	+	90740679	90739479	90740879	1400	24	3	-378
FAS	355	T10F056897F7	chr10	+	90740727	90739527	90740927	1400	8	6	-459
FAS	355	T10F056898D5	chr10	+	90740949	90739749	90741149	1400	11	7	-667
AXL	558	T19F02C44336	chr19	+	46416694	46415494	46416894	1400	125	13	0
AXL	558	T19F02C4444A	chr19	+	46416970	46415770	46417170	1400	1358	109	-200
AXL	558	T19F02C49618	chr19	+	46437912	46436712	46438112	1400	6	6	-21249
AXL	558	T19F02C4AC3C	chr19	+	46443580	46442380	46443780	1400	6	3	-26917
BAGE	574	T21R009A6D82	chr21	-	10120578	10120378	10121778	1400	18	5	-40893
KLF5	688	T13F0452BCC8	chr13	+	72531144	72529944	72531344	1400	810	262	37
KLF5	688	T13F0452BD65	chr13	+	72531301	72530101	72531501	1400	11	3	-158
KLF5	688	T13F0452BEF6	chr13	+	72531702	72530502	72531902	1400	15	4	-516
KLF5	688	T13F0452C98F	chr13	+	72534415	72533215	72534615	1400	18	4	-3266
KLF5	688	T13F0452C9FB	chr13	+	72534523	72533323	72534723	1400	15	4	-3342

KLF5	688	T13F0452CD07	chr13	+	72535303	72534103	72535503	1400	9	5	-4154
CDH5	1003	T16F03DF2E28	chr16	+	64957992	64956792	64958192	1400	8	3	72
CDH5	1003	T16F03DF2E87	chr16	+	64958087	64956887	64958287	1400	51	9	27
CDKN2A	1029	T09R014F27CD	chr9	-	21964749	21964549	21965949	1400	5	4	-6979
CDKN2A	1029	T09R014F2832	chr9	-	21964850	21964650	21966050	1400	42	3	-7059
CDKN2A	1029	T09R014F292E	chr9	-	21965102	21964902	21966302	1400	9	3	-7331
CDKN2A	1029	T09R014F74AA	chr9	-	21984426	21984226	21985626	1400	54	3	-26581
CDKN2B	1030	T09R014F782B	chr9	-	21985323	21985123	21986523	1400	6	5	7598
CENPF	1063	T01F0CAFBAA3	chr1	+	212843171	212841971	212843371	1400	250	27	36
CENPF	1063	T01F0CAFE824	chr1	+	212854820	212853620	212855020	1400	9	4	-11664
CENPF	1063	T01F0CAFF91E	chr1	+	212859166	212857966	212859366	1400	17	8	-16010
COL4A5	1287	T0XF0669621B	chrX	+	107569691	107568491	107569891	1400	48	9	151
COL4A5	1287	T0XF066AA941	chrX	+	107653441	107652241	107653641	1400	27	24	-83631
COL4A5	1287	T0XF066B8F4E	chrX	+	107712334	107711134	107712534	1400	22	6	-142500
COL4A5	1287	T0XF066C005C	chrX	+	107741276	107740076	107741476	1400	7	3	-171427
CLDN7	1366	T17R006C46D1	chr17	-	7096017	7095817	7097217	1400	96	12	8195
CLDN7	1366	T17R006C475E	chr17	-	7096158	7095958	7097358	1400	9	3	8045
CLDN7	1366	T17R006C6974	chr17	-	7104884	7104684	7106084	1400	5	4	-682
CTSB	1508	T08R00B2D17C	chr8	-	11719036	11718836	11720236	1400	8	4	18425
CTSB	1508	T08R00B2F717	chr8	-	11728663	11728463	11729863	1400	6	6	8796
CTSB	1508	T08R00B32137	chr8	-	11739447	11739247	11740647	1400	14	4	-1984
CTSB	1508	T08R00B321C8	chr8	-	11739592	11739392	11740792	1400	5	3	-2130
CTSB	1508	T08R00B3236B	chr8	-	11740011	11739811	11741211	1400	26	3	-2539
CTSB	1508	T08R00B343EA	chr8	-	11748330	11748130	11749530	1400	6	4	-10869
CTSL1	1514	T09F055621E6	chr9	+	89530854	89529654	89531054	1400	82	34	14
CTSL1	1514	T09F05562225	chr9	+	89530917	89529717	89531117	1400	41	3	-84
DAPK1	1612	T09F0552A64A	chr9	+	89302602	89301402	89302802	1400	3392	319	279

DAPK1	1612	T09F0552A6B2	chr9	+	89302706	89301506	89302906	1400	20	4	-119
DAPK1	1612	T09F0552A8D1	chr9	+	89303249	89302049	89303449	1400	48	4	-579
DAPK1	1612	T09F0552A995	chr9	+	89303445	89302245	89303645	1400	56	16	-851
DAPK1	1612	T09F0552EDCB	chr9	+	89320907	89319707	89321107	1400	13	12	-18330
DAPK1	1612	T09F05531156	chr9	+	89330006	89328806	89330206	1400	99	91	-27427
DAPK1	1612	T09F05533347	chr9	+	89338695	89337495	89338895	1400	13	6	-36119
DAPK1	1612	T09F055377BD	chr9	+	89356221	89355021	89356421	1400	9	4	-53645
DAPK1	1612	T09F0553C564	chr9	+	89376100	89374900	89376300	1400	10	7	-73524
DAPK1	1612	T09F0554CD0A	chr9	+	89443594	89442394	89443794	1400	5	3	-141013
DAPK1	1612	T09F0554EF7C	chr9	+	89452412	89451212	89452612	1400	222	73	-149834
DAPK1	1612	T09F055520E7	chr9	+	89465063	89463863	89465263	1400	8	8	-162487
DAPK1	1612	T09F055522C3	chr9	+	89465539	89464339	89465739	1400	5	4	-162963
AKR1C1	1645	T10F004BEEAB	chr10	+	4976299	4975099	4976499	1400	8	3	19158
AKR1C1	1645	T10F004C3A16	chr10	+	4995606	4994406	4995806	1400	6935	1730	-134
AKR1C1	1645	T10F004C45B6	chr10	+	4998582	4997382	4998782	1400	11	4	-3127
AKR1C1	1645	T10F004C726F	chr10	+	5010031	5008831	5010231	1400	5	3	-14577
AKR1C2	1646	T10R004CD811	chr10	-	5036049	5035849	5037249	1400	15826	12178	-14030
DES	1674	T02F0D1C66C1	chr2	+	219965121	219963921	219965321	1400	9	3	26222
DES	1674	T02F0D1CCD2F	chr2	+	219991343	219990143	219991543	1400	78	15	13
DFFA	1676	T01R009F5181	chr1	-	10441089	10440889	10442289	1400	7	3	2106
DIAPH2	1730	T0XF05B6324C	chrX	+	95826508	95825308	95826708	1400	160	7	-19
DIAPH2	1730	T0XF05B8F204	chrX	+	96006660	96005460	96006860	1400	6	4	-180295
DIAPH2	1730	T0XF05BF01DB	chrX	+	96403931	96402731	96404131	1400	25	4	-577564
DIAPH2	1730	T0XF05C09640	chrX	+	96507456	96506256	96507656	1400	6	3	-681091
DIAPH2	1730	T0XF05C18E4F	chrX	+	96570959	96569759	96571159	1400	6	6	-744594
DIAPH2	1730	T0XF05C1C729	chrX	+	96585513	96584313	96585713	1400	5	3	-759148
DIAPH2	1730	T0XF05C2E863	chrX	+	96659555	96658355	96659755	1400	9	5	-833187

DNMT3B	1789	T20F01D62E20	chr20	+	30813728	30812528	30813928	1400	13	3	125
DNMT3B	1789	T20F01D62E9C	chr20	+	30813852	30812652	30814052	1400	444	195	2
DNMT3B	1789	T20F01D62F6B	chr20	+	30814059	30812859	30814259	1400	64	5	-197
DNMT3B	1789	T20F01D66AD5	chr20	+	30829269	30828069	30829469	1400	6	6	-15417
DNMT3B	1789	T20F01D672D4	chr20	+	30831316	30830116	30831516	1400	24	4	-17438
DRD2	1813	T11R06B93E1D	chr11	-	112803357	112803157	112804557	1400	5	5	-17811
ECGF1	1890	T22R02F070BF	chr22	-	49311935	49311735	49313135	1400	6	3	-869
ECGF1	1890	T22R02F07217	chr22	-	49312279	49312079	49313479	1400	17	5	-1212
ECGF1	1890	T22R02F074AF	chr22	-	49312943	49312743	49314143	1400	11	3	-1865
ECGF1	1890	T22R02F07CFB	chr22	-	49315067	49314867	49316267	1400	15	5	-3967
ELF3	1999	T01F0BEF8442	chr1	+	200246338	200245138	200246538	1400	40	14	-2637
ELF3	1999	T01F0BEF8468	chr1	+	200246376	200245176	200246576	1400	57	14	-2676
ERCC1	2067	T19R03041DF2	chr19	-	50601458	50601258	50602658	1400	14	3	3290
ERCC1	2067	T19R030460F7	chr19	-	50618615	50618415	50619815	1400	169	30	-13795
EREG	2069	T04F047F457E	chr4	+	75449726	75448526	75449926	1400	10	3	1
EREG	2069	T04F047F5C4C	chr4	+	75455564	75454364	75455764	1400	6	3	-5840
EREG	2069	T04F047F5CF9	chr4	+	75455737	75454537	75455937	1400	11	4	-6008
ERCC3	2071	T02R079C73E5	chr2	-	127693797	127693597	127694997	1400	32	18	37557
ERCC3	2071	T02R079C75AB	chr2	-	127694251	127694051	127695451	1400	9	4	37110
ERCC3	2071	T02R079D967E	chr2	-	127768190	127767990	127769390	1400	125	14	-36799
ETS1	2113	T11R079EA343	chr11	-	127836995	127836795	127838195	1400	5	3	-3105
ETS1	2113	T11R079F5247	chr11	-	127881799	127881599	127882999	1400	10	6	-47908
ETS1	2113	T11R079F8EB4	chr11	-	127897268	127897068	127898468	1400	6	3	-63358
ETS1	2113	T11R079F8EDC	chr11	-	127897308	127897108	127898508	1400	11	3	-63408
FGFR1	2260	T08R024A4DF5	chr8	-	38424053	38423853	38425253	1400	7	6	-34583
FGFR1	2260	T08R024A7589	chr8	-	38434185	38433985	38435385	1400	6	3	-44717
FGFR1	2260	T08R024A9F21	chr8	-	38444833	38444633	38446033	1400	13	4	-55360

GDI2	2665	T10R0059D2A0	chr10	-	5886624	5886424	5887824	1400	10	4	-39413
GDI2	2665	T10R0059DB7C	chr10	-	5888892	5888692	5890092	1400	7	4	-41667
GRIN2B	2904	T12R00D160BC	chr12	-	13721788	13721588	13722988	1400	11	6	-116360
GRIN2B	2904	T12R00D47108	chr12	-	13922568	13922368	13923768	1400	6	3	-317126
HDAC1	3065	T01F01F04B4D	chr1	+	32525133	32523933	32525333	1400	10	9	5164
HDAC1	3065	T01F01F05EF4	chr1	+	32530164	32528964	32530364	1400	5	4	131
HDAC1	3065	T01F01F05F74	chr1	+	32530292	32529092	32530492	1400	992	102	38
HLA-A	3105	T06F01CA0AC1	chr6	+	30018241	30017041	30018441	1400	5	3	76
HLA-A	3105	T06F01CA0AFB	chr6	+	30018299	30017099	30018499	1400	20	4	14
HLA-A	3105	T06F01CA1741	chr6	+	30021441	30020241	30021641	1400	45	16	-3130
HNRPK	3190	T09R051CD4B0	chr9	-	85775536	85775336	85776736	1400	25	16	-2604
HNRPK	3190	T09R051CF6BF	chr9	-	85784255	85784055	85785455	1400	8	3	-11322
HNRPK	3190	T09R051CF6FC	chr9	-	85784316	85784116	85785516	1400	100	19	-11345
HNRPK	3190	T09R051CF766	chr9	-	85784422	85784222	85785622	1400	25	3	-11467
HNRPK	3190	T09R051CF96F	chr9	-	85784943	85784743	85786143	1400	544	45	-11941
HOXA9	3205	T07R019E7DB0	chr7	-	27164080	27163880	27165280	1400	47	4	4579
HOXA9	3205	T07R019E8942	chr7	-	27167042	27166842	27168242	1400	29	23	1558
HOXB7	3217	T17R02A00B5F	chr17	-	44043103	44042903	44044303	1400	9	7	-3488
HOXB9	3219	T17R02A044C2	chr17	-	44057794	44057594	44058994	1400	6	3	-4238
HOXB9	3219	T17R02A04670	chr17	-	44058224	44058024	44059424	1400	27	8	-4680
HOXD13	3239	T02F0A874EAD	chr2	+	176639661	176638461	176639861	1400	11	5	26120
HOXD13	3239	T02F0A87852A	chr2	+	176653610	176652410	176653810	1400	20	7	12169
HOXD13	3239	T02F0A87B45F	chr2	+	176665695	176664495	176665895	1400	23	6	84
IDH1	3417	T02R0C722DD4	chr2	-	208809428	208809228	208810628	1400	9	3	-210
IDH1	3417	T02R0C72734B	chr2	-	208827211	208827011	208828411	1400	15	5	-17993
IL8	3576	T04F04754026	chr4	+	74792998	74791798	74793198	1400	10	8	32141
IL8	3576	T04F0475467C	chr4	+	74794620	74793420	74794820	1400	7	5	30519

IL8	3576	T04F0475BDBE	chr4	+	74825150	74823950	74825350	1400	9	4	-11
ING1	3621	T13F0690F33A	chr13	+	110162746	110161546	110162946	1400	6	6	338
ING1	3621	T13F0690F3C1	chr13	+	110162881	110161681	110163081	1400	11	8	203
ING1	3621	T13F0690F850	chr13	+	110164048	110162848	110164248	1400	55	11	-893
ING1	3621	T13F069100AF	chr13	+	110166191	110164991	110166391	1400	13	3	-3106
IRF2	3660	T04R0B0F2DF0	chr4	-	185544176	185543976	185545376	1400	6	3	1752
IRF2	3660	T04R0B108190	chr4	-	185631120	185630920	185632320	1400	17	9	-85192
IRF2	3660	T04R0B1083F4	chr4	-	185631732	185631532	185632932	1400	7	3	-85782
IRF2	3660	T04R0B1084F8	chr4	-	185631992	185631792	185633192	1400	26	14	-86065
IRF2	3660	T04R0B1085C5	chr4	-	185632197	185631997	185633397	1400	25	3	-86262
KRT1	3848	T12R030F5D62	chr12	-	51338594	51338394	51339794	1400	9	9	16144
KRT1	3848	T12R030F9FA3	chr12	-	51355555	51355355	51356755	1400	9	3	-815
KRT1	3848	T12R030FA780	chr12	-	51357568	51357368	51358768	1400	12	5	-2825
LMNA	4000	T01F093316DA	chr1	+	154343130	154341930	154343330	1400	6	3	7992
LMNA	4000	T01F09333616	chr1	+	154351126	154349926	154351326	1400	2047	1249	39
SMAD4	4089	T18F02C9F1AC	chr18	+	46789036	46787836	46789236	1400	8	6	21575
SMAD4	4089	T18F02CA45DE	chr18	+	46810590	46809390	46810790	1400	278	72	41
SMAD4	4089	T18F02CA46D2	chr18	+	46810834	46809634	46811034	1400	5	3	-223
SMAD4	4089	T18F02CAB2F3	chr18	+	46838515	46837315	46838715	1400	14	9	-27904
MAL	4118	T02F05AA02C0	chr2	+	95027904	95026704	95028104	1400	12	4	27302
MAL	4118	T02F05AA6D4C	chr2	+	95055180	95053980	95055380	1400	106	35	58
MAL	4118	T02F05AAC59A	chr2	+	95077786	95076586	95077986	1400	10	6	-22580
MCC	4163	T05R06B5A282	chr5	-	112566914	112566714	112568114	1400	8	4	-177464
MAP3K3	4215	T17F038515CC	chr17	+	59053516	59052316	59053716	1400	127	7	64
MAP3K3	4215	T17F0385DC2F	chr17	+	59104303	59103103	59104503	1400	5	3	-50770
MAP3K3	4215	T17F0385DD17	chr17	+	59104535	59103335	59104735	1400	11	8	-51002
MGMT	4255	T10F07D0C483	chr10	+	131122307	131121107	131122507	1400	8	7	33149

MGMT	4255	T10F07D0DC8A	chr10	+	131128458	131127258	131128658	1400	11	6	26998
MGMT	4255	T10F07D1462A	chr10	+	131155498	131154298	131155698	1400	72	12	18
MGMT	4255	T10F07D2423C	chr10	+	131220028	131218828	131220228	1400	10	3	-64571
MGMT	4255	T10F07D439D8	chr10	+	131348952	131347752	131349152	1400	6	3	-193488
MMP9	4318	T20F029FE7FE	chr20	+	44034046	44032846	44034246	1400	10	3	36931
MMP9	4318	T20F029FE83C	chr20	+	44034108	44032908	44034308	1400	7	3	36848
MMP9	4318	T20F02A0782A	chr20	+	44070954	44069754	44071154	1400	418	155	35
MMP9	4318	T20F02A086C9	chr20	+	44074697	44073497	44074897	1400	5	5	-3743
MMP10	4319	T11R0615F78D	chr11	-	102102925	102102725	102104125	1400	16	9	43540
MMP11	4320	T22F01567BEC	chr22	+	22445036	22443836	22445236	1400	8	6	0
MMP14	4323	T14F01556CE9	chr14	+	22375657	22374457	22375857	1400	161	13	52
MPO	4353	T17R03336EA0	chr17	-	53702304	53702104	53703504	1400	7	4	-69
MPO	4353	T17R033392A7	chr17	-	53711527	53711327	53712727	1400	14	4	-9271
MPO	4353	T17R03339932	chr17	-	53713202	53713002	53714402	1400	8	6	-10967
MTAP	4507	T09F014C8795	chr9	+	21792661	21791461	21792861	1400	251	32	36
MTAP	4507	T09F014D530C	chr9	+	21844748	21843548	21844948	1400	8	3	-52111
MUC5AC	4586	T11F0011F6D7	chr11	+	1177303	1176103	1177503	1400	6	6	-7940
MYL2	4633	T12R068C0791	chr12	-	109840273	109840073	109841473	1400	7	6	-7245
NBN	4683	T08R056C4D7F	chr8	-	90983807	90983607	90985007	1400	20	9	30983
NEO1	4756	T15F043CC65C	chr15	+	71091804	71090604	71092004	1400	5	4	40124
NEO1	4756	T15F043D5FA1	chr15	+	71131041	71129841	71131241	1400	169	8	913
NEO1	4756	T15F043D60A0	chr15	+	71131296	71130096	71131496	1400	186	33	734
NEO1	4756	T15F043D6219	chr15	+	71131673	71130473	71131873	1400	12	10	255
NEO1	4756	T15F043D6261	chr15	+	71131745	71130545	71131945	1400	1713	184	219
NEO1	4756	T15F043D6302	chr15	+	71131906	71130706	71132106	1400	1906	143	103
NEO1	4756	T15F043D63F8	chr15	+	71132152	71130952	71132352	1400	30	6	-197
NEO1	4756	T15F043DBDD3	chr15	+	71155155	71153955	71155355	1400	9	4	-23227

NEO1	4756	T15F043E5E10	chr15	+	71196176	71194976	71196376	1400	7	3	-64247
NEO1	4756	T15F043E74EA	chr15	+	71202026	71200826	71202226	1400	7	5	-70098
NEO1	4756	T15F043E7549	chr15	+	71202121	71200921	71202321	1400	16	3	-70183
NEO1	4756	T15F043F476C	chr15	+	71255916	71254716	71256116	1400	7	3	-123962
NEO1	4756	T15F044088BC	chr15	+	71338172	71336972	71338372	1400	11	5	-206244
NEO1	4756	T15F0440A66E	chr15	+	71345774	71344574	71345974	1400	9	4	-213830
PRDX1	5052	T01R02BA3E8F	chr1	-	45760143	45759943	45761343	1400	570	261	-10827
PAX9	5083	T14F022861B5	chr14	+	36200885	36199685	36201085	1400	71	12	-160
PAX9	5083	T14F0228627E	chr14	+	36201086	36199886	36201286	1400	19	7	-430
PHB	5245	T17R02AC5068	chr17	-	44847208	44847008	44848408	1400	30	6	-10755
PIK3CB	5291	T03R08560A8E	chr3	-	139856526	139856326	139857726	1400	5	4	415
PIN1	5300	T19F0095A4A3	chr19	+	9807011	9805811	9807211	1400	192	38	30
PLAUR	5329	T19R02E9A370	chr19	-	48866160	48865960	48867360	1400	77	21	-24028
PLAUR	5329	T19R02E9A38D	chr19	-	48866189	48865989	48867389	1400	7	3	-24075
POLB	5423	T08F0285ADA4	chr8	+	42315172	42313972	42315372	1400	33	6	41
MAPK1	5594	T22R0138A7EF	chr22	-	20490223	20490023	20491423	1400	5	4	-46258
MAPK1	5594	T22R0138AF0C	chr22	-	20492044	20491844	20493244	1400	24	13	-48075
MAPK1	5594	T22R01396CAC	chr22	-	20540588	20540388	20541788	1400	41	18	-96620
MAPK1	5594	T22R013994D3	chr22	-	20550867	20550667	20552067	1400	5	3	-106902
MAPK1	5594	T22R013996DC	chr22	-	20551388	20551188	20552588	1400	5	3	-107423
MAPK1	5594	T22R013997CE	chr22	-	20551630	20551430	20552830	1400	27	4	-107666
MAPK1	5594	T22R01399855	chr22	-	20551765	20551565	20552965	1400	33	5	-107774
MAPK1	5594	T22R01399870	chr22	-	20551792	20551592	20552992	1400	19	3	-107827
PROX1	5629	T01F0CA65763	chr1	+	212227939	212226739	212228139	1400	54	9	589
PROX1	5629	T01F0CA679B9	chr1	+	212236729	212235529	212236929	1400	7	4	-8246
RELN	5649	T07R06228008	chr7	-	102924296	102924096	102925496	1400	12	3	-24801
RELN	5649	T07R0623276C	chr7	-	102967148	102966948	102968348	1400	17	6	-67638

RELN	5649	T07R06237004	chr7	-	102985732	102985532	102986932	1400	7	3	-86239
RELN	5649	T07R06237DD3	chr7	-	102989267	102989067	102990467	1400	10	3	-89763
RELN	5649	T07R06237F13	chr7	-	102989587	102989387	102990787	1400	5	3	-90080
RELN	5649	T07R0623FE11	chr7	-	103022097	103021897	103023297	1400	10	5	-122604
RELN	5649	T07R062A0432	chr7	-	103416882	103416682	103418082	1400	6	3	-517385
RELN	5649	T07R062A044B	chr7	-	103416907	103416707	103418107	1400	5	3	-517414
RELN	5649	T07R062A0499	chr7	-	103416985	103416785	103418185	1400	8	3	-517487
PTPRC	5788	T01F0BBC1229	chr1	+	196874793	196873593	196874993	1400	148	25	64
PTPRC	5788	T01F0BBC68E6	chr1	+	196896998	196895798	196897198	1400	8	7	-22150
PTPRC	5788	T01F0BBC8E3D	chr1	+	196906557	196905357	196906757	1400	14	11	-31709
PTPRC	5788	T01F0BBCB02D	chr1	+	196915245	196914045	196915445	1400	5	5	-40397
PTPRC	5788	T01F0BBCB27F	chr1	+	196915839	196914639	196916039	1400	7	4	-40989
RARG	5916	T12R0317E25C	chr12	-	51896924	51896724	51898124	1400	16	7	-6278
RARG	5916	T12R0317EFF3	chr12	-	51900403	51900203	51901603	1400	52	11	-9750
RARG	5916	T12R03181DBA	chr12	-	51912122	51911922	51913322	1400	27	4	-21454
RARRES1	5918	T03R09880DD8	chr3	-	159911384	159911184	159912584	1400	6	3	-13773
REG1A	5967	T02F04B88361	chr2	+	79201121	79199921	79201321	1400	83	26	-27
RNF6	6049	T13R0187F4A3	chr13	-	25687203	25687003	25688403	1400	5	4	-82931
RNF6	6049	T13R0187F501	chr13	-	25687297	25687097	25688497	1400	8	6	-83025
RNF6	6049	T13R01880EF5	chr13	-	25693941	25693741	25695141	1400	409	60	-89654
RNF6	6049	T13R01880FD2	chr13	-	25694162	25693962	25695362	1400	16	3	-89887
RNF6	6049	T13R018810C5	chr13	-	25694405	25694205	25695605	1400	840	77	-90113
RPL28	6158	T19F039C7BE5	chr19	+	60586981	60585781	60587181	1400	7	5	2131
RPL28	6158	T19F039C8438	chr19	+	60589112	60587912	60589312	1400	1659	1014	11
HLTF	6596	T03R08F3C046	chr3	-	150192198	150191998	150193398	1400	6	3	38426
HLTF	6596	T03R08F3C1DB	chr3	-	150192603	150192403	150193803	1400	5	3	38021
HLTF	6596	T03R08F4F720	chr3	-	150271776	150271576	150272976	1400	9	8	-41155

SOD2	6648	T06R0989C054	chr6	-	160022612	160022412	160023812	1400	15	4	-2449
SOD2	6648	T06R0989C1A3	chr6	-	160022947	160022747	160024147	1400	7	3	-2789
SOD2	6648	T06R0989C459	chr6	-	160023641	160023441	160024841	1400	8	4	-3475
SPOCK1	6695	T05R082859C6	chr5	-	136862150	136861950	136863350	1400	91	11	-523186
SPOCK1	6695	T05R08285A93	chr5	-	136862355	136862155	136863555	1400	21	3	-523365
SPP1	6696	T04F054FCCF2	chr4	+	89115890	89114690	89116090	1400	257	95	2
SSX2	6757	T0XR0324383E	chrX	-	52705342	52705142	52706542	1400	18	15	37347
SURF1	6834	T09R080F1F63	chr9	-	135208803	135208603	135210003	1400	26	6	-297
SURF1	6834	T09R080F3044	chr9	-	135213124	135212924	135214324	1400	11	5	-4618
PRDX2	7001	T19R00C28DA7	chr19	-	12750247	12750047	12751447	1400	10	5	18415
TERF1	7013	T08F046A12DE	chr8	+	74060510	74059310	74060710	1400	323	239	23153
TERF1	7013	T08F046A6D73	chr8	+	74083699	74082499	74083899	1400	166	25	37
TERF1	7013	T08F046AFF69	chr8	+	74121065	74119865	74121265	1400	6	3	-37403
NR2F1	7025	T05F058A38EE	chr5	+	92944622	92943422	92944822	1400	5	3	177
NR2F1	7025	T05F058A392C	chr5	+	92944684	92943484	92944884	1400	84	14	151
NR2F1	7025	T05F058A3AB6	chr5	+	92945078	92943878	92945278	1400	9	3	-279
NR2F1	7025	T05F058A3B1F	chr5	+	92945183	92943983	92945383	1400	19	4	-337
NR2F1	7025	T05F058A3B45	chr5	+	92945221	92944021	92945421	1400	8	3	-411
NR2F1	7025	T05F058A3E17	chr5	+	92945943	92944743	92946143	1400	11	3	-1142
NR2F1	7025	T05F058A3E63	chr5	+	92946019	92944819	92946219	1400	13	5	-1205
NR2F1	7025	T05F058A3EFA	chr5	+	92946170	92944970	92946370	1400	26	6	-1321
NR2F1	7025	T05F058A3F3D	chr5	+	92946237	92945037	92946437	1400	38	12	-1409
NR2F1	7025	T05F058A4089	chr5	+	92946569	92945369	92946769	1400	9	3	-1762
NR2F1	7025	T05F058A487F	chr5	+	92948607	92947407	92948807	1400	6	3	-3788
NR2F1	7025	T05F058A4E50	chr5	+	92950096	92948896	92950296	1400	10	9	-5297
TGIF1	7050	T18F0033E8D0	chr18	+	3401936	3400736	3402136	1400	5	3	149
TGIF1	7050	T18F00347AEA	chr18	+	3439338	3438138	3439538	1400	7	5	-37265

TGIF1	7050	T18F00347E2F	chr18	+	3440175	3438975	3440375	1400	31	12	-38089
TGIF1	7050	T18F003483C0	chr18	+	3441600	3440400	3441800	1400	10	5	-39519
TGIF1	7050	T18F00348403	chr18	+	3441667	3440467	3441867	1400	5	3	-39595
TPM2	7169	T09R02206E8B	chr9	-	35679883	35679683	35681083	1400	6	3	-7872
TPM2	7169	T09R02206EB2	chr9	-	35679922	35679722	35681122	1400	124	61	-7911
TPM4	7171	T19F00F4BDB4	chr19	+	16039348	16038148	16039548	1400	3970	1764	9153
TPM4	7171	T19F00F4BE5B	chr19	+	16039515	16038315	16039715	1400	25	3	8865
TPM4	7171	T19F00F4DF29	chr19	+	16047913	16046713	16048113	1400	10	3	412
TPM4	7171	T19F00F4DFB1	chr19	+	16048049	16046849	16048249	1400	8	3	278
TPM4	7171	T19F00F4E005	chr19	+	16048133	16046933	16048333	1400	439	106	243
TPM4	7171	T19F00F4E0B3	chr19	+	16048307	16047107	16048507	1400	6232	1231	51
TPM4	7171	T19F00F4E1BD	chr19	+	16048573	16047373	16048773	1400	22	3	-235
TPM4	7171	T19F00F4E291	chr19	+	16048785	16047585	16048985	1400	63	3	-438
TPM4	7171	T19F00F4EA98	chr19	+	16050840	16049640	16051040	1400	9	3	-2515
TPM4	7171	T19F00F4ED2B	chr19	+	16051499	16050299	16051699	1400	9	5	-3174
TPM4	7171	T19F00F50DE7	chr19	+	16059879	16058679	16060079	1400	192	62	-11550
UCHL1	7345	T04F0270493D	chr4	+	40913213	40912013	40913413	1400	9	4	40473
UCHL1	7345	T04F0270E766	chr4	+	40953702	40952502	40953902	1400	624	316	40
UCHL1	7345	T04F0270E971	chr4	+	40954225	40953025	40954425	1400	7	4	-536
UGT2B7	7364	T04F042C10ED	chr4	+	69996781	69995581	69996981	1400	14	5	33
WNT2B	7482	T01F06BA20CB	chr1	+	112861387	112860187	112861587	1400	5	3	-49824
PSCA	8000	T08F08915B2E	chr8	+	143743790	143742590	143743990	1400	7	5	15087
PSCA	8000	T08F08916C9B	chr8	+	143748251	143747051	143748451	1400	6	3	10626
PSCA	8000	T08F08916D0B	chr8	+	143748363	143747163	143748563	1400	23	18	10514
PSCA	8000	T08F08919643	chr8	+	143758915	143757715	143759115	1400	21	12	-28
DYRK2	8445	T12F03F418D8	chr12	+	66328792	66327592	66328992	1400	128	7	296
TCAP	8557	T17F02173438	chr17	+	35075128	35073928	35075328	1400	23	12	-3

CASK	8573	T0XR027B25E4	chrX	-	41625060	41624860	41626260	1400	21	19	-360754
TP73L	8626	T03F0B608ED2	chr3	+	190877394	190876194	190877594	1400	6	6	-45484
TNFRSF6B	8771	T20F03AE4FFE	chr20	+	61755390	61754190	61755590	1400	7	4	41348
TNFRSF6B	8771	T20F03AE6270	chr20	+	61760112	61758912	61760312	1400	106	9	36644
TNFRSF6B	8771	T20F03AE63C5	chr20	+	61760453	61759253	61760653	1400	68	4	36321
TNFRSF6B	8771	T20F03AEF831	chr20	+	61798449	61797249	61798649	1400	229	52	-1710
TNFRSF6B	8771	T20F03AEF87B	chr20	+	61798523	61797323	61798723	1400	5	3	-1778
CACNA1G	8913	T17F02BDCC73	chr17	+	45993075	45991875	45993275	1400	18	4	382
CACNA1G	8913	T17F02BDCD7B	chr17	+	45993339	45992139	45993539	1400	14	9	109
CACNA1G	8913	T17F02BE3EF0	chr17	+	46022384	46021184	46022584	1400	6	5	-28936
ANGPTL1	9068	T01R0A8DD58B	chr1	-	177067403	177067203	177068603	1400	8	5	17917
MTA1	9112	T14F06418378	chr14	+	104956792	104955592	104956992	1400	6	4	619
MTA1	9112	T14F0641853A	chr14	+	104957242	104956042	104957442	1400	918	76	298
MTA1	9112	T14F06418599	chr14	+	104957337	104956137	104957537	1400	114	12	85
MTA1	9112	T14F0641FBFC	chr14	+	104987644	104986444	104987844	1400	15	5	-30233
NR1H4	9971	T12F05EC319C	chr12	+	99365276	99364076	99365476	1400	6	5	26534
NR1H4	9971	T12F05EC991F	chr12	+	99391775	99390575	99391975	1400	266	88	96
NR1H4	9971	T12F05EC9981	chr12	+	99391873	99390673	99392073	1400	22	8	-46
NR1H4	9971	T12F05EC99AA	chr12	+	99391914	99390714	99392114	1400	11	7	-104
NR1H4	9971	T12F05ED0C55	chr12	+	99421269	99420069	99421469	1400	8	3	-29454
FRAT1	10023	T10F05E7AC5B	chr10	+	99069019	99067819	99069219	1400	8	3	1
NDC80	10403	T18F0027162C	chr18	+	2561580	2560380	2561780	1400	25	3	68
NUDC	10726	T01F019DB4E3	chr1	+	27112675	27111475	27112875	1400	14	7	8137
NUDC	10726	T01F019DD4B7	chr1	+	27120823	27119623	27121023	1400	428	30	13
HPSE	10855	T04R0508FC8E	chr4	-	84475022	84474822	84476222	1400	121	12	-39468
NMU	10874	T04R0359787E	chr4	-	56195198	56194998	56196398	1400	5	4	-39024
STARD3	10948	T17F0216C613	chr17	+	35046931	35045731	35047131	1400	131	14	79

STARD3	10948	T17F02171D4C	chr17	+	35069260	35068060	35069460	1400	6	4	-22322
PTPN21	11099	T14R05401820	chr14	-	88086560	88086360	88087760	1400	7	5	-82673
PTPN21	11099	T14R05401939	chr14	-	88086841	88086641	88088041	1400	6	3	-82953
PTPN21	11099	T14R05401C09	chr14	-	88087561	88087361	88088761	1400	61	18	-83649
BAZ1A	11177	T14R020C03D2	chr14	-	34341842	34341642	34343042	1400	12	8	-50134
BAZ1A	11177	T14R020D1D1B	chr14	-	34413851	34413651	34415051	1400	18	4	-122112
BAZ1A	11177	T14R020D1ECD	chr14	-	34414285	34414085	34415485	1400	102	14	-122491
RRAS2	22800	T11R00DA6EE1	chr11	-	14315233	14315033	14316433	1400	8	4	-58322
RRAS2	22800	T11R00DABDBF	chr11	-	14335423	14335223	14336623	1400	7	5	-78531
IKZF3	22806	T17R021891FD	chr17	-	35164669	35164469	35165869	1400	70	53	10076
MAPRE1	22919	T20F01D70F7C	chr20	+	30871420	30870220	30871620	1400	265	35	75
PDS5B	23047	T13F01E92CDD	chr13	+	32058589	32057389	32058789	1400	1189	160	143
PDS5B	23047	T13F01E92E2E	chr13	+	32058926	32057726	32059126	1400	6	3	-293
PDS5B	23047	T13F01E92F7F	chr13	+	32059263	32058063	32059463	1400	9	3	-630
PDS5B	23047	T13F01EA456B	chr13	+	32130411	32129211	32130611	1400	12	4	-71786
PDS5B	23047	T13F01EA8A46	chr13	+	32148038	32146838	32148238	1400	20	9	-89395
PDS5B	23047	T13F01EA9C86	chr13	+	32152710	32151510	32152910	1400	9	6	-94079
PDS5B	23047	T13F01EAB643	chr13	+	32159299	32158099	32159499	1400	8	6	-100673
JMJD2C	23081	T09F00665549	chr9	+	6706505	6705305	6706705	1400	8	3	41578
JMJD2C	23081	T09F00697823	chr9	+	6912035	6910835	6912235	1400	5	5	-163952
NAT6	24142	T03R02FFB028	chr3	-	50311208	50311008	50312408	1400	5	3	-2351
NAT6	24142	T03R02FFB07F	chr3	-	50311295	50311095	50312495	1400	97	16	-2409
IL1F5	26525	T02F06C46D95	chr2	+	113536405	113535205	113536605	1400	6	6	-3719
NGFRAP1	27018	T0XF061C4D6A	chrX	+	102518122	102516922	102518322	1400	111	19	-161
NGFRAP1	27018	T0XF061C4FB2	chrX	+	102518706	102517506	102518906	1400	11	9	-782
MLH3	27030	T14R04721F13	chr14	-	74587923	74587723	74589123	1400	40	8	-37658
HOOK2	29911	T19R00C25E1D	chr19	-	12738077	12737877	12739277	1400	12	7	-3240

HOOK2	29911	T19R00C27B97	chr19	-	12745623	12745423	12746823	1400	7	3	-10786
ZNF639	51193	T03F0AC2916F	chr3	+	180523375	180522175	180523575	1400	33	3	874
ZNF639	51193	T03F0AC291F9	chr3	+	180523513	180522313	180523713	1400	97	21	762
RHCG	51458	T15R053B408C	chr15	-	87769228	87769028	87770428	1400	8	8	46431
RHCG	51458	T15R053B693F	chr15	-	87779647	87779447	87780847	1400	8	6	36013
MBIP	51562	T14R0222D2C2	chr14	-	35836610	35836410	35837810	1400	5	3	928
WWOX	51741	T16F0492358B	chr16	+	76690827	76689627	76691027	1400	89	9	248
WWOX	51741	T16F0492367F	chr16	+	76691071	76689871	76691271	1400	1488	66	61
WWOX	51741	T16F0493F4DA	chr16	+	76805338	76804138	76805538	1400	9	4	-114285
WWOX	51741	T16F04947517	chr16	+	76838167	76836967	76838367	1400	9	9	-147115
DGCR8	54487	T22F01197D98	chr22	+	18447768	18446568	18447968	1400	78	4	111
UGT1A10	54575	T02F0DF64CA3	chr2	+	234245283	234244083	234245483	1400	10	3	-35411
UGT1A10	54575	T02F0DF7A5D7	chr2	+	234333655	234332455	234333855	1400	10	6	-123793
UGT1A8	54576	T02F0DF64CA3	chr2	+	234245283	234244083	234245483	1400	10	3	-54243
UGT1A8	54576	T02F0DF7A5D7	chr2	+	234333655	234332455	234333855	1400	10	6	-142625
CHFR	55743	T12R07DCDF01	chr12	-	131915521	131915321	131916721	1400	1621	50	11604
CHFR	55743	T12R07DCE002	chr12	-	131915778	131915578	131916978	1400	294	12	11343
CHFR	55743	T12R07DCFE81	chr12	-	131923585	131923385	131924785	1400	13	5	3465
CHFR	55743	T12R07DD0323	chr12	-	131924771	131924571	131925971	1400	24	11	2259
NPDC1	56654	T09R0849D1AE	chr9	-	139055534	139055334	139056734	1400	6	4	-1773
NPDC1	56654	T09R0849D8F0	chr9	-	139057392	139057192	139058592	1400	5	3	-3625
NPDC1	56654	T09R0849E3CE	chr9	-	139060174	139059974	139061374	1400	12	3	-6414
PCDH10	57575	T04F080119EB	chr4	+	134289899	134288699	134290099	1400	60	43	45
PCDH10	57575	T04F080124E9	chr4	+	134292713	134291513	134292913	1400	27	4	-2777
PCDH10	57575	T04F08012578	chr4	+	134292856	134291656	134293056	1400	19	3	-2910
PCDH10	57575	T04F080126B3	chr4	+	134293171	134291971	134293371	1400	11	6	-3250
RSRC2	65117	T12R073F0FD3	chr12	-	121573331	121573131	121574531	1400	6	3	-18085

EFHD2	79180	T01F00EE2D44	chr1	+	15609156	15607956	15609356	1400	15	7	-128
EFTUD1	79631	T15R04C7A2AE	chr15	-	80192174	80191974	80193374	1400	7	7	17469
EFTUD1	79631	T15R04C80380	chr15	-	80216960	80216760	80218160	1400	8	4	-7315
SOX7	83595	T08R00A2115F	chr8	-	10621279	10621079	10622479	1400	16	14	-2571
PPP1R1B	84152	T17F021632C7	chr17	+	35009223	35008023	35009423	1400	6	3	27482
PPP1R1B	84152	T17F021643B2	chr17	+	35013554	35012354	35013754	1400	6	3	23151
PPP1R1B	84152	T17F021653AA	chr17	+	35017642	35016442	35017842	1400	8	3	19063
PPP1R1B	84152	T17F02167B39	chr17	+	35027769	35026569	35027969	1400	30	14	8959
PPP1R1B	84152	T17F021686E0	chr17	+	35030752	35029552	35030952	1400	14	10	5953
PPP1R1B	84152	T17F02169DFE	chr17	+	35036670	35035470	35036870	1400	8	3	42
PPP1R1B	84152	T17F02169E3B	chr17	+	35036731	35035531	35036931	1400	70	16	8
C2orf40	84417	T02F06522C1D	chr2	+	106048541	106047341	106048741	1400	31	3	4
KISS1R	84634	T19F000CA74C	chr19	+	829260	828060	829460	1400	5	5	39098
KISS1R	84634	T19F000D3070	chr19	+	864368	863168	864568	1400	8	6	3991
KISS1R	84634	T19F000D4628	chr19	+	869928	868728	870128	1400	8	4	-1570
KISS1R	84634	T19F000D4C92	chr19	+	871570	870370	871770	1400	6	3	-3212
ABCC11	85320	T16R02C99B0F	chr16	-	46766863	46766663	46768063	1400	6	5	-8523
NKD2	85409	T05F0010341D	chr5	+	1061917	1060717	1062117	1400	40	21	251
NKD2	85409	T05F00103505	chr5	+	1062149	1060949	1062349	1400	6	3	22
NKD2	85409	T05F0010358A	chr5	+	1062282	1061082	1062482	1400	8	5	-114
SLCO6A1	133482	T05R06124CC3	chr5	-	101862595	101862395	101863795	1400	11	4	-127009
CD109	135228	T06F046F912C	chr6	+	74420524	74419324	74420724	1400	10	4	42028
CD109	135228	T06F046F9158	chr6	+	74420568	74419368	74420768	1400	86	12	42002
CD109	135228	T06F0470348E	chr6	+	74462350	74461150	74462550	1400	106	10	358
CD109	135228	T06F04703543	chr6	+	74462531	74461331	74462731	1400	99	13	126
CD109	135228	T06F047035B7	chr6	+	74462647	74461447	74462847	1400	272	21	-25
CD109	135228	T06F04703622	chr6	+	74462754	74461554	74462954	1400	79	47	-197

CD109	135228	T06F04714E6F	chr6	+	74534511	74533311	74534711	1400	5	3	-71963
CD109	135228	T06F04720B0E	chr6	+	74582798	74581598	74582998	1400	8	5	-120250
GATA5	140628	T20R039A3B3B	chr20	-	60439355	60439155	60440555	1400	26	12	32611
SLC30A7	148867	T01F06072B02	chr1	+	101133058	101131858	101133258	1400	25	10	1208
SLC30A7	148867	T01F06072EBF	chr1	+	101134015	101132815	101134215	1400	5	3	263
SLC30A7	148867	T01F06072FA4	chr1	+	101134244	101133044	101134444	1400	1390	168	71
SLC30A7	148867	T01F06073027	chr1	+	101134375	101133175	101134575	1400	6	4	-109
ORAOV1	220064	T11R041FE5C0	chr11	-	69199296	69199096	69200496	1400	58	6	-9731
CCDC110	256309	T04R0B1F0C3B	chr4	-	186584123	186583923	186585323	1400	62	3	19271
CCDC110	256309	T04R0B1F2FAC	chr4	-	186593196	186592996	186594396	1400	16	10	10163
LAMA1	284217	T18R0069967C	chr18	-	6919804	6919604	6921004	1400	7	3	12100
LAMA1	284217	T18R006A80E4	chr18	-	6979812	6979612	6981012	1400	7	3	-47908
LAMA1	284217	T18R006BB679	chr18	-	7059065	7058865	7060265	1400	5	4	-127163

APPENDIX III: TFBSs with overlapping putative EREs

ClassOfOverlap	TargetGeneID	TargetHeaderID	MatrixID	TF_GeneID	Overlap	TFBSlength	TFstart	TFstop	EREstart	EREstop
2	1832	T06F0072B44E	M00222	6929	13	16	154	169	157	178
3	2114	T21F0254E128	M00240	None	7	7	-197	-191	-212	-191
2	2114	T21F0254E128	M00640	3201	2	8	-218	-211	-212	-191
1	2114	T21F0254E128	M00747	3659	5	7	-195	-189	-212	-191
2	7428	T03F009B0605	M01020	6910	3	10	-193	-184	-186	-165
2	7428	T03F009B0605	M00749	None	3	7	-190	-184	-186	-165
1	7428	T03F009B0605	M00514	468	6	12	-170	-159	-186	-165
2	7428	T03F009B0605	M00993	6886	9	10	-187	-178	-186	-165
1	7428	T03F009B0605	M00373	None	17	21	-181	-161	-186	-165
2	7428	T03F009B0605	M00066	6929	12	16	-190	-175	-186	-165
2	7428	T03F009B0605	M00066	6886	12	16	-190	-175	-186	-165
2	7428	T03F009B0605	M00070	6925	12	16	-190	-175	-186	-165
2	7428	T03F009B0605	M00070	6886	12	16	-190	-175	-186	-165
2	7428	T03F009B0605	M00065	6886	12	16	-190	-175	-186	-165
2	7428	T03F009B0605	M00065	6929	12	16	-190	-175	-186	-165
3	220972	T10R02B4E65E	M00217	7391	8	8	-931	-924	-931	-910
2	220972	T10R02B4E65E	M00539	405	14	20	-937	-918	-931	-910
3	220972	T10R02B4E65E	M00726	7392	6	6	-930	-925	-931	-910
2	220972	T10R02B4E65E	M00236	405	12	16	-935	-920	-931	-910
3	220972	T10R02B4E65E	M00799	4149	7	7	-930	-924	-931	-910
3	220972	T10R02B4E65E	M00799	4609	7	7	-930	-924	-931	-910
3	4585	T03R0BBDC7A3	M00803	1871	6	6	-577	-572	-590	-569

3	4585	T03R0BBDC7A3	M00803	7027	6	6	-577	-572	-590	-569
3	4585	T03R0BBDC7A3	M00803	1874	6	6	-577	-572	-590	-569
3	4585	T03R0BBDC7A3	M00803	1869	6	6	-577	-572	-590	-569
2	51450	T09F07D6641E	M00327	5077	2	21	-217	-197	-198	-177
1	51450	T09F07D6641E	M00179	1386	10	12	-186	-175	-198	-177
1	51450	T09F07D6641E	M00178	1385	10	12	-186	-175	-198	-177
4	51450	T09F07D6641E	M00460	6776	22	24	-200	-177	-198	-177
1	51450	T09F07D6641E	M00538	None	16	24	-192	-169	-198	-177
1	51450	T09F07D6641E	M00981	1385	8	9	-184	-176	-198	-177
1	51450	T09F07D6641E	M00981	22809	8	9	-184	-176	-198	-177
1	51450	T09F07D6641E	M00981	11016	8	9	-184	-176	-198	-177
1	51450	T09F07D6641E	M00981	468	8	9	-184	-176	-198	-177
1	51450	T09F07D6641E	M00981	1386	8	9	-184	-176	-198	-177
1	51450	T09F07D6641E	M00981	466	8	9	-184	-176	-198	-177
1	51450	T09F07D6641E	M00981	1390	8	9	-184	-176	-198	-177
1	51450	T09F07D6641E	M00981	22926	8	9	-184	-176	-198	-177
1	51450	T09F07D6641E	M00981	467	8	9	-184	-176	-198	-177
3	51450	T09F07D6641E	M00039	1385	8	8	-184	-177	-198	-177
3	51450	T09F07D6641E	M00483	22926	8	8	-184	-177	-198	-177
1	51450	T09F07D6641E	M00017	None	11	14	-187	-174	-198	-177
2	4588	T11R000F6A1D	M01109	51385	10	15	-728	-714	-723	-702
2	4588	T11R000F6A1D	M00518	6256	15	19	-727	-709	-723	-702
2	4588	T11R000F6A1D	M00518	5465	15	19	-727	-709	-723	-702
2	1364	T07F04581C3F	M00415	6935	6	9	-99	-91	-96	-75
1	3651	T13F01A208FB	M00729	None	11	14	-565	-552	-576	-555
3	4017	T08R016224CB	M00751	861	6	6	-938	-933	-949	-928
3	4017	T08R016224CB	M00271	861	6	6	-938	-933	-949	-928

3	4017	T08R016224CB	M00722	None	8	8	-940	-933	-949	-928
4	4017	T08R016224CB	M00999	326	22	26	-952	-927	-949	-928
3	4017	T08R01628357	M00975	5989	9	9	-982	-974	-990	-969
3	4017	T08R01628357	M00975	5992	9	9	-982	-974	-990	-969
3	4017	T08R01628357	M00975	8625	9	9	-982	-974	-990	-969
3	4017	T08R01628357	M00975	5990	9	9	-982	-974	-990	-969
3	4017	T08R01628357	M00975	5993	9	9	-982	-974	-990	-969
3	4017	T08R01628357	M00975	5991	9	9	-982	-974	-990	-969
3	4017	T08R01628357	M00975	5994	9	9	-982	-974	-990	-969
3	4017	T08R01628357	M00117	1051	14	14	-982	-969	-990	-969
3	4017	T08R01628357	M00116	1050	14	14	-982	-969	-990	-969
3	4017	T08R01628357	M00770	1053	12	12	-980	-969	-990	-969
3	4017	T08R01628357	M00770	1050	12	12	-980	-969	-990	-969
3	4017	T08R01628357	M00770	1054	12	12	-980	-969	-990	-969
3	4017	T08R01628357	M00770	1051	12	12	-980	-969	-990	-969
3	4017	T08R01628357	M00770	1052	12	12	-980	-969	-990	-969
3	4017	T08R01628357	M00912	1051	12	12	-980	-969	-990	-969
3	4017	T08R01628357	M00912	1050	12	12	-980	-969	-990	-969
3	4017	T08R01628357	M00912	1052	12	12	-980	-969	-990	-969
1	581	T19F033A5DEB	M00033	2033	1	14	-921	-908	-942	-921
3	581	T19F033A5DEB	M00638	3172	13	13	-941	-929	-942	-921
3	581	T19F033A5DEB	M00764	3172	13	13	-942	-930	-942	-921
3	581	T19F033A5DEB	M00764	3174	13	13	-942	-930	-942	-921
3	581	T19F033A5DEB	M00762	7026	13	13	-942	-930	-942	-921
3	581	T19F033A5DEB	M00762	3172	13	13	-942	-930	-942	-921
3	581	T19F033A5DEB	M00762	7025	13	13	-942	-930	-942	-921
3	581	T19F033A5DEB	M00762	3174	13	13	-942	-930	-942	-921

3	581	T19F033A5DEB	M00765	7025	13	13	-942	-930	-942	-921
3	581	T19F033A5DEB	M00765	7026	13	13	-942	-930	-942	-921
1	581	T19F033A5DEB	M00245	1960	2	12	-922	-911	-942	-921
1	581	T19F033A5DEB	M00246	1959	2	12	-922	-911	-942	-921
2	581	T19F033A5DEB	M00158	3172	13	14	-943	-930	-942	-921
2	581	T19F033A5DEB	M00158	7025	13	14	-943	-930	-942	-921
2	581	T19F033A5DEB	M00468	11278	2	7	-947	-941	-942	-921
2	581	T19F033A5DEB	M00083	7593	4	8	-946	-939	-942	-921
3	4233	T07F06EC9654	M00192	2908	19	19	-217	-199	-218	-197
1	4233	T07F06EC9654	M00206	6927	1	17	-197	-181	-218	-197
1	3480	T15F05CC17A8	M00373	None	7	21	-951	-931	-966	-945
1	3480	T15F05CC17A8	M00821	4780	10	13	-954	-942	-966	-945
1	3480	T15F05CC17A8	M00514	468	9	12	-953	-942	-966	-945
1	3480	T15F05CC17A8	M00378	None	11	12	-955	-944	-966	-945
1	3480	T15F05CC17A8	M00983	4094	9	11	-953	-943	-966	-945
1	3480	T15F05CC17A8	M00983	4779	9	11	-953	-943	-966	-945
1	3480	T15F05CC17A8	M00983	4778	9	11	-953	-943	-966	-945
1	3480	T15F05CC17A8	M00983	4097	9	11	-953	-943	-966	-945
1	3480	T15F05CC17A8	M00983	4780	9	11	-953	-943	-966	-945
1	3480	T15F05CC17A8	M00983	571	9	11	-953	-943	-966	-945
1	3480	T15F05CC17A8	M00983	60468	9	11	-953	-943	-966	-945
1	3480	T15F05CC17A8	M00983	23764	9	11	-953	-943	-966	-945
1	3480	T15F05CC17A8	M00983	9935	9	11	-953	-943	-966	-945
1	3480	T15F05CC17A8	M00983	7975	9	11	-953	-943	-966	-945
2	598	T20R01C5892A	M00192	2908	9	19	-211	-193	-201	-180
2	598	T20R01C5892A	M01058	None	8	12	-283	-272	-279	-258
2	598	T20R01C5892A	M00955	2908	13	27	-215	-189	-201	-180

2	598	T20R01C5892A	M00956	None	13	27	-215	-189	-201	-180
1	598	T20R01C5892A	M00097	5080	8	21	-187	-167	-201	-180
2	598	T20R01C5892A	M00250	None	14	24	-289	-266	-279	-258
1	598	T20R01C5892A	M00143	5079	10	28	-189	-162	-201	-180
1	1509	T11R001A6B3F	M00468	11278	5	7	185	191	168	189
2	1509	T11R001A6B3F	M00246	1959	10	12	166	177	168	189
3	3481	T11R00203190	M00415	6935	9	9	-592	-584	-599	-578
2	3481	T11R00203190	M00687	None	5	11	-605	-595	-599	-578
3	3481	T11R00203190	M00749	None	7	7	-586	-580	-599	-578
2	3481	T11R00203190	M00099	None	7	16	-608	-593	-599	-578
2	3481	T11R00203190	M00380	None	19	30	-610	-581	-599	-578
2	3481	T11R00203190	M00377	None	3	11	-607	-597	-599	-578
3	3481	T11R00203190	M00221	6720	11	11	-589	-579	-599	-578
3	3397	T20F01C48AB7	M00340	2114	14	14	-353	-340	-361	-340
3	3397	T20F01C48AB7	M00415	6935	9	9	-956	-948	-962	-941
2	3397	T20F01C48AB7	M00233	4205	1	22	17	38	38	59
2	3397	T20F01C48AB7	M00233	4205	5	22	17	38	34	55
1	3397	T20F01C48AB7	M00955	2908	3	27	-130	-104	-149	-128
1	3397	T20F01C48AB7	M00160	6736	6	12	127	138	111	132
1	3397	T20F01C48AB7	M00069	7528	16	20	-199	-180	-205	-184
1	3397	T20F01C48AB7	M01047	7020	3	15	-83	-69	-102	-81
1	3397	T20F01C48AB7	M01047	7020	12	15	-83	-69	-93	-72
1	3397	T20F01C48AB7	M00042	None	5	10	128	137	111	132
3	3397	T20F01C48AB7	M00716	None	8	8	-101	-94	-102	-81
3	3397	T20F01C48AB7	M00450	None	9	9	-585	-577	-593	-572
3	3397	T20F01C48AB7	M00803	1871	6	6	-146	-141	-149	-128
3	3397	T20F01C48AB7	M00803	7027	6	6	-146	-141	-149	-128

3	3397	T20F01C48AB7	M00803	1874	6	6	-146	-141	-149	-128
3	3397	T20F01C48AB7	M00803	1869	6	6	-146	-141	-149	-128
1	3397	T20F01C48AB7	M00797	3091	3	14	-648	-635	-667	-646
1	3397	T20F01C48AB7	M00531	1998	17	18	-356	-339	-361	-340
2	3397	T20F01C48AB7	M00327	5077	18	21	48	68	51	72
1	3397	T20F01C48AB7	M00327	5077	12	21	48	68	38	59
1	3397	T20F01C48AB7	M00327	5077	8	21	48	68	34	55
1	3397	T20F01C48AB7	M00415	6935	6	9	-463	-455	-479	-458
2	3397	T20F01C48AB7	M00245	1960	8	12	-153	-142	-149	-128
3	3397	T20F01C48AB7	M00518	6256	19	19	-671	-653	-674	-653
3	3397	T20F01C48AB7	M00518	5465	19	19	-671	-653	-674	-653
2	3397	T20F01C48AB7	M00518	6256	15	19	-671	-653	-667	-646
2	3397	T20F01C48AB7	M00518	5465	15	19	-671	-653	-667	-646
3	3397	T20F01C48AB7	M00394	None	9	9	-95	-87	-102	-81
2	3397	T20F01C48AB7	M00394	None	7	9	-95	-87	-93	-72
2	3397	T20F01C48AB7	M01079	None	10	16	-680	-665	-674	-653
2	3397	T20F01C48AB7	M01079	None	3	16	-680	-665	-667	-646
1	3397	T20F01C48AB7	M00538	None	7	24	-652	-629	-667	-646
4	3397	T20F01C48AB7	M00144	5079	22	28	49	76	51	72
1	3397	T20F01C48AB7	M00144	5079	11	28	49	76	38	59
1	3397	T20F01C48AB7	M00144	5079	7	28	49	76	34	55
2	3397	T20F01C48AB7	M00453	3665	10	18	-601	-584	-593	-572
1	3397	T20F01C48AB7	M00800	7020	6	16	-86	-71	-102	-81
1	3397	T20F01C48AB7	M00800	7021	6	16	-86	-71	-102	-81
1	3397	T20F01C48AB7	M00800	7020	15	16	-86	-71	-93	-72
1	3397	T20F01C48AB7	M00800	7021	15	16	-86	-71	-93	-72
1	3397	T20F01C48AB7	M00423	55810	9	14	124	137	111	132

1	3397	T20F01C48AB7	M00144	5079	7	28	-240	-213	-255	-234
1	3397	T20F01C48AB7	M01109	51385	6	15	-189	-175	-205	-184
2	3397	T20F01C48AB7	M00059	7528	10	17	44	60	51	72
1	3397	T20F01C48AB7	M00059	7528	16	17	44	60	38	59
1	3397	T20F01C48AB7	M00059	7528	12	17	44	60	34	55
1	3397	T20F01C48AB7	M00017	None	2	14	-647	-634	-667	-646
3	3397	T20F01C48AB7	M00098	5076	19	19	51	69	51	72
1	3397	T20F01C48AB7	M00098	5076	9	19	51	69	38	59
1	3397	T20F01C48AB7	M00098	5076	5	19	51	69	34	55
3	4968	T03F00950821	M00704	7003	6	6	-889	-884	-890	-869
1	4968	T03F00950821	M00993	6886	3	10	-871	-862	-890	-869
3	3308	T05F07E4C83C	M00241	None	8	8	-726	-719	-733	-712
3	3308	T05F07E4C83C	M00395	None	9	9	-726	-718	-733	-712
3	3308	T05F07E4C83C	M00640	3201	8	8	-725	-718	-733	-712
4	3308	T05F07E4C83C	M00023	None	22	30	-737	-708	-733	-712
2	3481	T11R00203166	M00991	1044	11	18	-564	-547	-557	-536
2	3481	T11R00203166	M00991	1045	11	18	-564	-547	-557	-536
2	10397	T08R0801A109	M00241	None	4	8	-948	-941	-944	-923
1	1044	T05F08E92640	M01034	4609	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	7391	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	84159	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	4613	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	8553	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	6886	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	6929	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	6887	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	4149	1	10	-536	-527	-557	-536

1	1044	T05F08E92640	M01034	4654	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	4084	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	79365	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	7392	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	3637	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	7942	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	6925	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	4617	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	6938	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	7030	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	9421	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	4656	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	4618	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	9464	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	4601	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	4807	1	10	-536	-527	-557	-536
1	1044	T05F08E92640	M01034	429	1	10	-536	-527	-557	-536
3	1044	T05F08E92640	M00076	2624	10	10	-546	-537	-557	-536
1	2886	T17F02185167	M00280	5989	4	17	-419	-403	-437	-416
3	2886	T17F02185167	M00241	None	8	8	-427	-420	-437	-416
3	5770	T20F02E4F8B1	M00644	None	7	7	-824	-818	-838	-817
1	5770	T20F02E4F8B1	M00993	6886	9	10	-825	-816	-838	-817
1	4968	T03F009506EC	M00993	6886	3	10	-562	-553	-581	-560
3	4968	T03F009506EC	M00704	7003	6	6	-580	-575	-581	-560
2	2122	T03R0A26D50C	M00622	1054	10	13	-90	-78	-87	-66
1	2122	T03R0A26D50C	M00432	7080	9	10	-74	-65	-87	-66
2	2122	T03R0A26D50C	M00201	1050	13	18	-92	-75	-87	-66

3	2122	T03R0A26D50C	M00240	None	7	7	-72	-66	-87	-66
2	2122	T03R0A26D50C	M00249	1649	9	13	-91	-79	-87	-66
2	2122	T03R0A26D50C	M00249	1050	9	13	-91	-79	-87	-66
3	2122	T03R0A26D50C	M00912	1051	12	12	-87	-76	-87	-66
3	2122	T03R0A26D50C	M00912	1050	12	12	-87	-76	-87	-66
3	2122	T03R0A26D50C	M00912	1052	12	12	-87	-76	-87	-66
2	2122	T03R0A26D50C	M00116	1050	12	14	-89	-76	-87	-66
2	2122	T03R0A26D50C	M00190	1050	12	14	-89	-76	-87	-66
1	4249	T02F0807A47B	M01107	None	3	10	-902	-893	-921	-900
3	4249	T02F0807A47B	M00638	3172	13	13	-356	-344	-357	-336
1	4249	T02F0807A47B	M00999	326	6	26	-905	-880	-921	-900
4	4249	T02F0807A47B	M00380	None	22	30	-359	-330	-357	-336
1	4249	T02F0807A47B	M00066	6929	14	16	-349	-334	-357	-336
1	4249	T02F0807A47B	M00066	6886	14	16	-349	-334	-357	-336
1	4249	T02F0807A47B	M00065	6886	14	16	-349	-334	-357	-336
1	4249	T02F0807A47B	M00065	6929	14	16	-349	-334	-357	-336
1	4249	T02F0807A47B	M01000	326	8	25	-907	-883	-921	-900
2	4249	T02F0807A47B	M00717	7849	7	15	-365	-351	-357	-336
1	4249	T02F0807A47B	M00097	5080	17	21	-352	-332	-357	-336
2	4588	T11R000F6745	M00717	7849	13	15	3	17	5	26
2	4588	T11R000F6745	M00378	None	7	12	0	11	5	26
2	4588	T11R000F6745	M00221	6720	9	11	3	13	5	26
3	4588	T11R000F6745	M00749	None	7	7	6	12	5	26
2	4588	T11R000F6745	M00776	6721	8	12	1	12	5	26
2	4588	T11R000F6745	M00776	6720	8	12	1	12	5	26
2	4588	T11R000F6745	M00114	1385	9	15	-1	13	5	26
1	81	T19F029D9BDC	M00097	5080	3	21	78	98	59	80

2	81	T19F029D9BDC	M01014	6666	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	8403	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	6663	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	11166	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	6658	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	6659	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	6664	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	6736	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	30812	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	6657	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	54345	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	6665	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	6660	7	13	53	65	59	80
2	81	T19F029D9BDC	M01014	6662	7	13	53	65	59	80
2	81	T19F029D9BDC	M00477	2309	6	14	51	64	59	80
2	81	T19F029D9BDC	M00477	2310	6	14	51	64	59	80
2	983	T10F03B498BF	M00160	6736	4	12	41	52	49	70
2	983	T10F03B498BF	M00290	2295	9	16	42	57	49	70
2	983	T10F03B498BF	M00291	2296	9	16	42	57	49	70
2	983	T10F03B498BF	M01137	2309	5	12	42	53	49	70
2	983	T10F03B498BF	M00377	None	1	11	39	49	49	70
2	983	T10F03B498BF	M00474	2308	7	14	42	55	49	70
2	983	T10F03B498BF	M00476	4303	7	14	42	55	49	70
2	983	T10F03B498BF	M00477	2309	7	14	42	55	49	70
2	983	T10F03B498BF	M00477	2310	7	14	42	55	49	70
2	7077	T17R046EAC82	M00373	None	14	21	-828	-808	-821	-800
2	7077	T17R046EAC82	M00450	None	1	9	-829	-821	-821	-800

3	7077	T17R046EAC82	M00444	7421	15	15	-820	-806	-821	-800
3	7077	T17R046EAC82	M00221	6720	11	11	-813	-803	-821	-800
1	2114	T21F0254E24F	M00747	3659	5	7	-490	-484	-507	-486
3	2114	T21F0254E24F	M00240	None	7	7	-492	-486	-507	-486
2	2114	T21F0254E24F	M00640	3201	2	8	-513	-506	-507	-486
3	2626	T08F00B0D09C	M00241	None	8	8	-837	-830	-851	-830
1	2626	T08F00B0D09C	M00416	8092	9	18	-838	-821	-851	-830
1	2626	T08F00B0D09C	M00133	5453	12	15	-841	-827	-851	-830
1	2335	T02R0CDFDEBD	M00380	None	17	30	-173	-144	-178	-157
2	654	T06F0076E486	M00463	5454	2	14	-971	-958	-959	-938
1	654	T06F0076E486	M00128	2623	3	13	-940	-928	-959	-938
1	654	T06F0076E486	M00346	2623	2	10	-939	-930	-959	-938
1	4072	T02F02D40719	M00746	1997	8	12	-439	-428	-453	-432
1	9244	T19R011B5610	M00821	4780	6	13	34	46	18	39
1	9244	T19R011B5610	M00983	4094	5	11	35	45	18	39
1	9244	T19R011B5610	M00983	4779	5	11	35	45	18	39
1	9244	T19R011B5610	M00983	4778	5	11	35	45	18	39
1	9244	T19R011B5610	M00983	4097	5	11	35	45	18	39
1	9244	T19R011B5610	M00983	4780	5	11	35	45	18	39
1	9244	T19R011B5610	M00983	571	5	11	35	45	18	39
1	9244	T19R011B5610	M00983	60468	5	11	35	45	18	39
1	9244	T19R011B5610	M00983	23764	5	11	35	45	18	39
1	9244	T19R011B5610	M00983	9935	5	11	35	45	18	39
1	9244	T19R011B5610	M00983	7975	5	11	35	45	18	39
3	9244	T19R011B5610	M00649	4150	8	8	23	30	18	39
2	3570	T01F09192B5F	M00975	5989	5	9	-349	-341	-345	-324
2	3570	T01F09192B5F	M00975	5992	5	9	-349	-341	-345	-324

2	3570	T01F09192B5F	M00975	8625	5	9	-349	-341	-345	-324
2	3570	T01F09192B5F	M00975	5990	5	9	-349	-341	-345	-324
2	3570	T01F09192B5F	M00975	5993	5	9	-349	-341	-345	-324
2	3570	T01F09192B5F	M00975	5991	5	9	-349	-341	-345	-324
2	3570	T01F09192B5F	M00975	5994	5	9	-349	-341	-345	-324
2	3570	T01F09192B5F	M01072	3090	7	15	-353	-339	-345	-324
2	4249	T02F08090E08	M00284	4779	7	22	61	82	76	97
1	4249	T02F08090E08	M00072	None	6	11	92	102	76	97
3	4249	T02F08090E08	M00415	6935	9	9	86	94	76	97
1	4249	T02F08090E08	M00380	None	5	30	93	122	76	97
3	4249	T02F08090E08	M00699	3394	12	12	83	94	76	97
2	4249	T02F08090E08	M00821	4780	3	13	66	78	76	97
3	4249	T02F08090E08	M00772	3664	15	15	82	96	76	97
3	4249	T02F08090E08	M00772	3661	15	15	82	96	76	97
3	4249	T02F08090E08	M00772	3663	15	15	82	96	76	97
3	4249	T02F08090E08	M00772	3662	15	15	82	96	76	97
3	4249	T02F08090E08	M00772	3659	15	15	82	96	76	97
3	4249	T02F08090E08	M00772	3665	15	15	82	96	76	97
3	4249	T02F08090E08	M00772	3394	15	15	82	96	76	97
3	4249	T02F08090E08	M00772	3660	15	15	82	96	76	97
2	4249	T02F08090E08	M00432	7080	6	10	72	81	76	97
1	4249	T02F08090E08	M00999	326	13	26	85	110	76	97
2	4288	T10R07BC8C6D	M00394	None	5	9	-289	-281	-285	-264
1	7168	T15F03A4D4F8	M00133	5453	11	15	-722	-708	-733	-712
2	7168	T15F03A4D4F8	M00378	None	1	12	-431	-420	-420	-399
2	7168	T15F03A4D4F8	M00377	None	1	11	-743	-733	-733	-712
1	7168	T15F03A4D4F8	M00999	326	18	26	-729	-704	-733	-712

3	7168	T15F03A4D4F8	M00395	None	9	9	-730	-722	-733	-712
1	5970	T11R03E29DFF	M00143	5079	1	28	-581	-554	-602	-581
1	4313	T16F03390D53	M00626	5989	9	14	-594	-581	-607	-586
2	6774	T17R02400D8B	M00220	6720	2	11	-316	-306	-307	-286
1	2627	T18F0112BFD1	M00360	None	2	13	-88	-76	-108	-87
2	2627	T18F0112BFD1	M00947	7024	11	15	-112	-98	-108	-87
1	2627	T18F0112BFD1	M00716	None	1	8	-87	-80	-108	-87
2	7422	T06F029D086D	M00378	None	2	12	-460	-449	-450	-429
2	7422	T06F029D086D	M00444	7421	3	15	-462	-448	-450	-429
2	7422	T06F029D086D	M01113	None	3	8	-455	-448	-450	-429
3	7422	T06F029D086D	M00159	1050	13	13	-450	-438	-450	-429
3	7422	T06F029D086D	M00704	7003	6	6	-436	-431	-450	-429
3	1728	T16R0412730C	M00794	7080	12	12	-556	-545	-557	-536
2	3934	T09F07BEE741	M00373	None	20	21	-620	-600	-619	-598
2	2810	T01F019CD661	M00222	6929	11	16	-500	-485	-495	-474
2	2810	T01F019CD661	M00947	7024	4	15	-506	-492	-495	-474
2	2810	T01F019CD661	M00733	4089	9	15	-501	-487	-495	-474
3	7477	T22R02A9FF20	M00272	7157	10	10	-481	-472	-491	-470
2	7422	T06F029D0926	M01113	None	3	8	-640	-633	-635	-614
2	7422	T06F029D0926	M00444	7421	3	15	-647	-633	-635	-614
3	7422	T06F029D0926	M00159	1050	13	13	-635	-623	-635	-614
3	7422	T06F029D0926	M00704	7003	6	6	-621	-616	-635	-614
2	7422	T06F029D0926	M00378	None	2	12	-645	-634	-635	-614
2	994	T20F00383158	M00245	1960	9	12	132	143	135	156
2	6256	T09F08217724	M00915	7022	3	13	-150	-138	-140	-119
2	6256	T09F08217724	M00915	7020	3	13	-150	-138	-140	-119
2	6256	T09F08217724	M00915	7021	3	13	-150	-138	-140	-119

2	6256	T09F08217724	M00800	7020	1	16	-155	-140	-140	-119
2	6256	T09F08217724	M00800	7021	1	16	-155	-140	-140	-119
2	6256	T09F08217724	M00448	None	8	9	-141	-133	-140	-119
2	6256	T09F08217724	M00450	None	8	9	-141	-133	-140	-119
2	6256	T09F08217724	M00378	None	8	12	-144	-133	-140	-119
2	7422	T06F029D06FA	M00444	7421	3	15	-91	-77	-79	-58
3	7422	T06F029D06FA	M00704	7003	6	6	-65	-60	-79	-58
3	7422	T06F029D06FA	M00159	1050	13	13	-79	-67	-79	-58
2	7422	T06F029D06FA	M00378	None	2	12	-89	-78	-79	-58
2	7422	T06F029D06FA	M01113	None	3	8	-84	-77	-79	-58
1	7053	T20F00228137	M01109	51385	10	15	-283	-269	-295	-274
3	7053	T20F00228137	M00444	7421	15	15	-289	-275	-295	-274
1	7053	T20F00228137	M00792	4089	8	9	-281	-273	-295	-274
1	7053	T20F00228137	M00792	4087	8	9	-281	-273	-295	-274
1	7053	T20F00228137	M00792	4086	8	9	-281	-273	-295	-274
1	7053	T20F00228137	M00792	4088	8	9	-281	-273	-295	-274
1	1956	T07F0347ACD5	M00340	2114	1	14	-202	-189	-223	-202
2	1956	T07F0347ACD5	M01075	7704	9	29	-243	-215	-223	-202
1	1956	T07F0347ACD5	M00463	5454	11	14	-212	-199	-223	-202
2	1956	T07F0347ACD5	M00135	5451	2	19	-240	-222	-223	-202
1	1476	T21R029FB389	M00144	5079	17	28	-730	-703	-735	-714
1	6382	T02R01353DBC	M00794	7080	9	12	-420	-409	-433	-412
1	6382	T02R01353DBC	M00717	7849	12	15	-423	-409	-433	-412
1	6382	T02R01353DBC	M00328	7849	14	18	-425	-408	-433	-412
3	6382	T02R01353DBC	M00776	6721	12	12	-425	-414	-433	-412
3	6382	T02R01353DBC	M00776	6720	12	12	-425	-414	-433	-412
3	6382	T02R01353DBC	M00450	None	9	9	-423	-415	-433	-412

1	7525	T18R000B9C0B	M00991	1044	5	18	-666	-649	-683	-662
1	7525	T18R000B9C0B	M00991	1045	5	18	-666	-649	-683	-662
1	7525	T18R000B9C0B	M00464	5454	7	10	-668	-659	-683	-662
1	7525	T18R000B9C0B	M00729	None	5	14	-666	-653	-683	-662
3	5970	T11R03E29C52	M00432	7080	10	10	-167	-158	-173	-152
2	4583	T11F00106AE1	M00448	None	8	9	-35	-27	-34	-13
1	4583	T11F00106AE1	M00800	7020	1	16	-13	2	-34	-13
1	4583	T11F00106AE1	M00800	7021	1	16	-13	2	-34	-13
2	3221	T12F03246972	M00695	None	3	7	-721	-715	-717	-696
3	3221	T12F03246972	M01047	7020	15	15	-710	-696	-717	-696
3	3221	T12F03246972	M00293	2300	16	16	-138	-123	-142	-121
3	3221	T12F03246972	M00464	5454	10	10	-131	-122	-142	-121
2	3481	T11R002032EE	M00436	3651	6	12	-955	-944	-949	-928
1	3481	T11R002032EE	M00284	4779	15	22	-942	-921	-949	-928
1	3481	T11R002032EE	M00821	4780	10	13	-937	-925	-949	-928
1	3481	T11R002032EE	M00098	5076	18	19	-945	-927	-949	-928
2	2122	T03R0A26D0B7	M00539	405	4	20	-19	0	-3	18
2	2122	T03R0A26D0B7	M00121	7391	1	14	-16	-3	-3	18
2	2122	T03R0A26D0B7	M00119	4149	1	14	-16	-3	-3	18
2	2122	T03R0A26D0B7	M00122	7391	1	14	-16	-3	-3	18
2	2122	T03R0A26D0B7	M00118	4149	1	14	-16	-3	-3	18
2	2122	T03R0A26D0B7	M00118	4609	1	14	-16	-3	-3	18
3	727897	T11F001252E8	M01100	51341	9	9	131	139	118	139
3	727897	T11F001252E8	M00412	6935	13	13	122	134	118	139
3	727897	T11F001252E8	M00984	860	15	15	-936	-922	-938	-917
3	727897	T11F001252E8	M00984	861	15	15	-936	-922	-938	-917
3	727897	T11F001252E8	M00984	865	15	15	-936	-922	-938	-917

3	727897	T11F001252E8	M00984	864	15	15	-936	-922	-938	-917
3	727897	T11F001252E8	M00159	1050	13	13	-937	-925	-938	-917
2	727897	T11F001252E8	M01047	7020	11	15	114	128	118	139
1	727897	T11F001252E8	M00450	None	8	9	132	140	118	139
3	727897	T11F001252E8	M00751	861	6	6	-932	-927	-938	-917
3	727897	T11F001252E8	M00271	861	6	6	-932	-927	-938	-917
1	727897	T11F001252E8	M00515	5468	18	23	122	144	118	139
2	23705	T11R06D3F6E4	M00136	5451	3	15	-569	-555	-557	-536
1	23705	T11R06D3F6E4	M00729	None	10	14	-545	-532	-557	-536
2	23705	T11R06D3F6E4	M00999	326	20	26	-563	-538	-557	-536
2	23705	T11R06D3F6E4	M00127	2623	12	14	-559	-546	-557	-536
2	23705	T11R06D3F6E4	M01003	None	6	11	-562	-552	-557	-536
3	23705	T11R06D3F6E4	M00128	2623	13	13	-549	-537	-557	-536
1	8766	T15F03CFD712	M00097	5080	16	21	-421	-401	-427	-406
1	8766	T15F03CFD712	M00463	5454	4	14	-409	-396	-427	-406
3	8766	T15F03CFD712	M00747	3659	7	7	-418	-412	-427	-406
2	3732	T11F02A6FE32	M00415	6935	7	9	-885	-877	-883	-862
3	7077	T17R046EAB7F	M00177	1385	12	12	-558	-547	-562	-541
3	7077	T17R046EAB7F	M00514	468	12	12	-558	-547	-562	-541
3	7077	T17R046EAB7F	M00179	1386	12	12	-558	-547	-562	-541
3	7077	T17R046EAB7F	M00178	1385	12	12	-558	-547	-562	-541
3	7077	T17R046EAB7F	M00039	1385	8	8	-556	-549	-562	-541
3	7077	T17R046EAB7F	M00981	1385	9	9	-557	-549	-562	-541
3	7077	T17R046EAB7F	M00981	22809	9	9	-557	-549	-562	-541
3	7077	T17R046EAB7F	M00981	11016	9	9	-557	-549	-562	-541
3	7077	T17R046EAB7F	M00981	468	9	9	-557	-549	-562	-541
3	7077	T17R046EAB7F	M00981	1386	9	9	-557	-549	-562	-541

3	7077	T17R046EAB7F	M00981	466	9	9	-557	-549	-562	-541
3	7077	T17R046EAB7F	M00981	1390	9	9	-557	-549	-562	-541
3	7077	T17R046EAB7F	M00981	22926	9	9	-557	-549	-562	-541
3	7077	T17R046EAB7F	M00981	467	9	9	-557	-549	-562	-541
1	9289	T16F0359B808	M00037	4778	4	11	-157	-147	-175	-154
3	9289	T16F0359B808	M01044	6910	10	10	-174	-165	-175	-154
1	9289	T16F0359B808	M00221	6720	9	11	-162	-152	-175	-154
3	9289	T16F0359B808	M00378	None	12	12	-165	-154	-175	-154
1	9289	T16F0359B808	M00373	None	20	21	-173	-153	-175	-154
1	6382	T02R01353CC3	M00340	2114	9	14	-171	-158	-184	-163
3	3714	T14R063C89CC	M00246	1959	12	12	-604	-593	-614	-593
1	3714	T14R063C89CC	M01044	6910	9	10	-601	-592	-614	-593
3	3714	T14R063C89CC	M00800	7020	16	16	-46	-31	-51	-30
3	3714	T14R063C89CC	M00800	7021	16	16	-46	-31	-51	-30
2	3714	T14R063C89CC	M00056	4782	10	29	-70	-42	-51	-30
2	3714	T14R063C89CC	M00144	5079	4	28	-75	-48	-51	-30
2	3714	T14R063C89CC	M00328	7849	1	18	-68	-51	-51	-30
1	3714	T14R063C89CC	M00373	None	10	21	-39	-19	-51	-30
2	3714	T14R063C89CC	M00143	5079	9	28	-70	-43	-51	-30
2	2626	T08F00B0A470	M00088	None	7	13	-381	-369	-375	-354
2	2626	T08F00B0A470	M01109	51385	10	15	-380	-366	-375	-354
2	2626	T08F00B0A470	M00208	None	9	12	-378	-367	-375	-354
3	6281	T01R08F45F62	M00915	7022	13	13	-562	-550	-568	-547
3	6281	T01R08F45F62	M00915	7020	13	13	-562	-550	-568	-547
3	6281	T01R08F45F62	M00915	7021	13	13	-562	-550	-568	-547
2	6498	T03F0A396CBC	M01036	7025	6	23	120	142	137	158
2	6498	T03F0A396CBC	M01036	7026	6	23	120	142	137	158

2	6498	T03F0A396CBC	M00727	2516	1	8	130	137	137	158
3	7799	T01F00D55012	M00632	2626	12	12	-251	-240	-260	-239
2	7799	T01F00D55012	M00468	11278	5	7	-262	-256	-260	-239
3	6714	T20F021C47E4	M00716	None	8	8	12	19	0	21
1	6714	T20F021C47E4	M00333	None	10	13	12	24	0	21
3	6714	T20F021C47E4	M00333	None	13	13	8	20	0	21
1	2068	T19R0303913C	M01043	1482	6	10	-201	-192	-217	-196
1	9244	T19R011B566F	M00143	5079	7	28	-62	-35	-77	-56
2	9244	T19R011B566F	M00808	5076	4	11	-84	-74	-77	-56
2	9244	T19R011B566F	M00808	7849	4	11	-84	-74	-77	-56
2	9244	T19R011B566F	M00808	5079	4	11	-84	-74	-77	-56
2	3481	T11R0020310A	M00075	None	6	10	-469	-460	-465	-444
2	3481	T11R0020310A	M00069	7528	12	20	-473	-454	-465	-444
1	3481	T11R0020310A	M00088	None	12	13	-455	-443	-465	-444
2	81	T19F029DBE32	M00765	7025	1	13	168	180	180	201
2	81	T19F029DBE32	M00765	7026	1	13	168	180	180	201
1	5087	T01F09B4110C	M00281	5989	9	18	147	164	134	155
1	5087	T01F09B4110C	M00410	6662	9	14	147	160	134	155
2	5087	T01F09B4110C	M00746	1997	10	12	132	143	134	155
2	5087	T01F09B4110C	M00658	6688	7	8	133	140	134	155
1	5087	T01F09B4110C	M00160	6736	8	12	148	159	134	155
3	5087	T01F09B4110C	M00415	6935	9	9	147	155	134	155
3	4436	T02F02D48712	M00396	None	7	7	-330	-324	-343	-322
2	4436	T02F02D48712	M00144	5079	8	28	-363	-336	-343	-322
1	4436	T02F02D48712	M00803	1871	4	6	-325	-320	-343	-322
1	4436	T02F02D48712	M00803	7027	4	6	-325	-320	-343	-322
1	4436	T02F02D48712	M00803	1874	4	6	-325	-320	-343	-322

1	4436	T02F02D48712	M00803	1869	4	6	-325	-320	-343	-322
3	4583	T11F00105A9E	M01118	7490	9	9	-446	-438	-454	-433
3	4583	T11F00105A9E	M00446	None	15	15	-452	-438	-454	-433
1	25	T09F07E6FAFC	M01010	3159	3	15	-346	-332	-365	-344
1	25	T09F07E6FAFC	M01010	8091	3	15	-346	-332	-365	-344
1	25	T09F07E6FAFC	M00341	2553	10	12	-353	-342	-365	-344
1	25	T09F07E6FAFC	M00341	2551	10	12	-353	-342	-365	-344
1	25	T09F07E6FAFC	M00460	6776	5	24	-348	-325	-365	-344
1	25	T09F07E6FAFC	M00459	6777	8	15	-351	-337	-365	-344
1	25	T09F07E6FAFC	M00224	6772	11	21	-354	-334	-365	-344
1	25	T09F07E6FAFC	M00225	6774	11	21	-354	-334	-365	-344
1	25	T09F07E6FAFC	M00223	6775	5	9	-348	-340	-365	-344
1	25	T09F07E6FAFC	M00223	6778	5	9	-348	-340	-365	-344
1	25	T09F07E6FAFC	M00223	6772	5	9	-348	-340	-365	-344
1	25	T09F07E6FAFC	M00223	6774	5	9	-348	-340	-365	-344
1	25	T09F07E6FAFC	M00223	6773	5	9	-348	-340	-365	-344
1	25	T09F07E6FAFC	M01004	None	4	11	-347	-337	-365	-344
2	1499	T03F0274EA18	M00378	None	10	12	28	39	30	51
3	1499	T03F0274EA18	M01100	51341	9	9	40	48	30	51
2	1499	T03F0274EA18	M00196	6667	4	13	21	33	30	51
2	1499	T03F0274EA18	M00932	6670	4	13	21	33	30	51
2	1499	T03F0274EA18	M00932	6668	4	13	21	33	30	51
2	1499	T03F0274EA18	M00932	6667	4	13	21	33	30	51
2	1499	T03F0274EA18	M00932	6671	4	13	21	33	30	51
3	1499	T03F0274EA18	M00450	None	9	9	39	47	30	51
2	1499	T03F0274EA18	M00008	6667	3	10	23	32	30	51
2	1499	T03F0274EA18	M00931	6670	3	10	23	32	30	51

2	1499	T03F0274EA18	M00931	6667	3	10	23	32	30	51
2	1499	T03F0274EA18	M00931	6671	3	10	23	32	30	51
2	1499	T03F0274EA18	M00933	6668	2	10	22	31	30	51
2	1499	T03F0274EA18	M00933	6671	2	10	22	31	30	51
2	1499	T03F0274EA18	M00933	6667	2	10	22	31	30	51
2	1499	T03F0274EA18	M00933	6670	2	10	22	31	30	51
2	1499	T03F0274EA18	M00982	1960	1	14	17	30	30	51
2	1499	T03F0274EA18	M00982	1959	1	14	17	30	30	51
2	1499	T03F0274EA18	M00982	1958	1	14	17	30	30	51
2	1499	T03F0274EA18	M00982	1961	1	14	17	30	30	51
1	1948	T13R06512E8E	M00102	1523	13	15	-158	-144	-167	-146
1	1948	T13R06512E8E	M00103	None	13	15	-158	-144	-167	-146
1	1948	T13R06512E8E	M01125	None	1	15	-146	-132	-167	-146
1	1948	T13R06512E8E	M00138	5451	2	23	-147	-125	-167	-146
1	1948	T13R06512E8E	M01075	7704	3	29	-148	-120	-167	-146
1	1948	T13R06512E8E	M00145	5454	5	16	-150	-135	-167	-146
3	4583	T11F00106014	M00285	4779	13	13	-582	-570	-591	-570
1	4583	T11F00106014	M00539	405	8	20	-577	-558	-591	-570
1	4583	T11F00106014	M00726	7392	1	6	-570	-565	-591	-570
3	4583	T11F00106014	M00395	None	9	9	-580	-572	-591	-570
1	1509	T11R001A5981	M00974	4091	7	11	-58	-48	-73	-52
1	1509	T11R001A5981	M00974	4089	7	11	-58	-48	-73	-52
1	1509	T11R001A5981	M00974	4092	7	11	-58	-48	-73	-52
1	1509	T11R001A5981	M00974	4087	7	11	-58	-48	-73	-52
1	1509	T11R001A5981	M00974	4086	7	11	-58	-48	-73	-52
1	1509	T11R001A5981	M00974	4090	7	11	-58	-48	-73	-52
1	1509	T11R001A5981	M00974	4088	7	11	-58	-48	-73	-52

2	4583	T11F00106B33	M00448	None	8	9	-117	-109	-116	-95
1	4583	T11F00106B33	M00800	7020	1	16	-95	-80	-116	-95
1	4583	T11F00106B33	M00800	7021	1	16	-95	-80	-116	-95
2	4790	T04F062D70B7	M00004	None	5	18	-993	-976	-980	-959
2	654	T06F007686F4	M00373	None	17	21	-317	-297	-313	-292
3	654	T06F007686F4	M00220	6720	11	11	-313	-303	-313	-292
2	654	T06F007686F4	M00514	468	6	12	-319	-308	-313	-292
2	654	T06F007686F4	M00513	467	7	14	-320	-307	-313	-292
2	654	T06F007686F4	M00373	None	4	21	-330	-310	-313	-292
3	654	T06F007686F4	M01029	4286	8	8	-312	-305	-313	-292
3	654	T06F007686F4	M01029	7030	8	8	-312	-305	-313	-292
3	654	T06F007686F4	M01029	7942	8	8	-312	-305	-313	-292
2	4436	T02F02D48B97	M00449	7546	8	9	125	133	126	147
3	4436	T02F02D48B97	M00333	None	13	13	133	145	126	147
3	4436	T02F02D48B97	M00751	861	6	6	129	134	126	147
3	4436	T02F02D48B97	M00271	861	6	6	129	134	126	147
2	3815	T04F034A92C1	M00733	4089	12	15	-455	-441	-452	-431
2	3815	T04F034A92C1	M00143	5079	13	28	-467	-440	-452	-431
2	2114	T21F0254E0A0	M00640	3201	2	8	-82	-75	-76	-55
1	2114	T21F0254E0A0	M00747	3659	5	7	-59	-53	-76	-55
3	2114	T21F0254E0A0	M00240	None	7	7	-61	-55	-76	-55
2	3570	T01F09192BAD	M00975	5989	5	9	-427	-419	-423	-402
2	3570	T01F09192BAD	M00975	5992	5	9	-427	-419	-423	-402
2	3570	T01F09192BAD	M00975	8625	5	9	-427	-419	-423	-402
2	3570	T01F09192BAD	M00975	5990	5	9	-427	-419	-423	-402
2	3570	T01F09192BAD	M00975	5993	5	9	-427	-419	-423	-402
2	3570	T01F09192BAD	M00975	5991	5	9	-427	-419	-423	-402

2	3570	T01F09192BAD	M00975	5994	5	9	-427	-419	-423	-402
2	3570	T01F09192BAD	M01072	3090	7	15	-431	-417	-423	-402
2	11214	T15F0503078B	M00281	5989	6	18	-177	-160	-165	-144
2	6624	T07F00556EFD	M00915	7022	8	13	-87	-75	-82	-61
2	6624	T07F00556EFD	M00915	7020	8	13	-87	-75	-82	-61
2	6624	T07F00556EFD	M00915	7021	8	13	-87	-75	-82	-61
1	6624	T07F00556EFD	M00143	5079	21	28	-81	-54	-82	-61
3	6624	T07F00556EFD	M00192	2908	19	19	-425	-407	-425	-404
1	6624	T07F00556EFD	M00800	7020	7	16	-67	-52	-82	-61
1	6624	T07F00556EFD	M00800	7021	7	16	-67	-52	-82	-61
3	6624	T07F00556EFD	M01113	None	8	8	-68	-61	-82	-61
1	6624	T07F00556EFD	M00626	5989	4	14	-64	-51	-82	-61
1	6624	T07F00556EFD	M00144	5079	2	28	-62	-35	-82	-61
2	6624	T07F00556EFD	M00144	5079	16	28	-94	-67	-82	-61
1	302	T15R037B8B75	M00195	5451	7	15	-734	-720	-749	-728
3	302	T15R037B8B75	M00770	1053	12	12	-743	-732	-749	-728
3	302	T15R037B8B75	M00770	1050	12	12	-743	-732	-749	-728
3	302	T15R037B8B75	M00770	1054	12	12	-743	-732	-749	-728
3	302	T15R037B8B75	M00770	1051	12	12	-743	-732	-749	-728
3	302	T15R037B8B75	M00770	1052	12	12	-743	-732	-749	-728
1	302	T15R037B8B75	M00795	5451	6	11	-733	-723	-749	-728
1	302	T15R037B8B75	M00795	5452	6	11	-733	-723	-749	-728
1	302	T15R037B8B75	M00795	5450	6	11	-733	-723	-749	-728
1	302	T15R037B8B75	M00795	5460	6	11	-733	-723	-749	-728
1	302	T15R037B8B75	M00795	5462	6	11	-733	-723	-749	-728
1	302	T15R037B8B75	M00795	5457	6	11	-733	-723	-749	-728
1	302	T15R037B8B75	M00795	5453	6	11	-733	-723	-749	-728

1	302	T15R037B8B75	M00795	5454	6	11	-733	-723	-749	-728
1	302	T15R037B8B75	M00135	5451	9	19	-736	-718	-749	-728
2	302	T15R037B8B75	M00991	1044	16	18	-751	-734	-749	-728
2	302	T15R037B8B75	M00991	1045	16	18	-751	-734	-749	-728
1	302	T15R037B8B75	M00210	5452	7	13	-734	-722	-749	-728
1	302	T15R037B8B75	M00210	5451	7	13	-734	-722	-749	-728
2	302	T15R037B8B75	M00380	None	8	30	-771	-742	-749	-728
1	302	T15R037B8B75	M00138	5451	11	23	-738	-716	-749	-728
1	302	T15R037B8B75	M00464	5454	4	10	-731	-722	-749	-728
1	302	T15R037B8B75	M00342	5451	5	10	-732	-723	-749	-728
1	302	T15R037B8B75	M00131	None	10	15	-737	-723	-749	-728
3	4173	T08F02EC40A7	M00466	3091	12	12	-727	-716	-736	-715
3	4843	T17R016143A8	M00240	None	7	7	-688	-682	-689	-668
2	4843	T17R016143A8	M00794	7080	10	12	-691	-680	-689	-668
1	4843	T17R016143A8	M00222	6929	4	16	-671	-656	-689	-668
3	4583	T11F00105E29	M00285	4779	13	13	-91	-79	-100	-79
1	4583	T11F00105E29	M00726	7392	1	6	-79	-74	-100	-79
1	4583	T11F00105E29	M00539	405	8	20	-86	-67	-100	-79
3	4583	T11F00105E29	M00395	None	9	9	-89	-81	-100	-79
1	5914	T17F02213EF7	M00340	2114	5	14	-752	-739	-769	-748
1	5914	T17F02213EF7	M00531	1998	6	18	-753	-736	-769	-748
1	5914	T17F02213EF7	M00341	2553	4	12	-751	-740	-769	-748
1	5914	T17F02213EF7	M00341	2551	4	12	-751	-740	-769	-748
1	5914	T17F02213EF7	M00678	51513	3	10	-750	-741	-769	-748
1	5914	T17F02213EF7	M00327	5077	17	21	-764	-744	-769	-748
1	5914	T17F02213EF7	M00339	2113	4	15	-751	-737	-769	-748
1	5914	T17F02213EF7	M00032	None	4	10	-751	-742	-769	-748

1	5914	T17F02213EF7	M01043	1482	1	10	-748	-739	-769	-748
1	5914	T17F02213EF7	M00655	2118	2	7	-749	-743	-769	-748
1	5914	T17F02213EF7	M00971	2005	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	2113	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	51513	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	2004	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	2002	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	1998	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	2078	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	2551	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	2114	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	6688	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	6689	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	2313	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	2553	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	2000	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	1997	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	2118	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00971	2120	3	8	-750	-743	-769	-748
1	5914	T17F02213EF7	M00771	1998	3	12	-750	-739	-769	-748
1	5914	T17F02213EF7	M00771	2077	3	12	-750	-739	-769	-748
1	5914	T17F02213EF7	M00771	51513	3	12	-750	-739	-769	-748
1	5914	T17F02213EF7	M00771	2113	3	12	-750	-739	-769	-748
1	5914	T17F02213EF7	M00771	2313	3	12	-750	-739	-769	-748
1	5914	T17F02213EF7	M00771	2078	3	12	-750	-739	-769	-748
1	5914	T17F02213EF7	M00771	2005	3	12	-750	-739	-769	-748
1	5914	T17F02213EF7	M00771	2114	3	12	-750	-739	-769	-748

2	4583	T11F00106A1D	M00448	None	8	9	161	169	162	183
1	4583	T11F00106A1D	M00800	7020	1	16	183	198	162	183
1	4583	T11F00106A1D	M00800	7021	1	16	183	198	162	183
2	3481	T11R00203254	M00245	1960	4	12	-803	-792	-795	-774
2	3481	T11R00203254	M00196	6667	1	13	-807	-795	-795	-774
2	3481	T11R00203254	M00932	6670	1	13	-807	-795	-795	-774
2	3481	T11R00203254	M00932	6668	1	13	-807	-795	-795	-774
2	3481	T11R00203254	M00932	6667	1	13	-807	-795	-795	-774
2	3481	T11R00203254	M00932	6671	1	13	-807	-795	-795	-774
3	3481	T11R00203254	M01100	51341	9	9	-792	-784	-795	-774
3	3481	T11R00203254	M01100	51341	9	9	-795	-787	-795	-774
2	994	T20F0038F8EC	M00143	5079	14	28	-900	-873	-886	-865
2	994	T20F0038F8EC	M00221	6720	1	11	-896	-886	-886	-865
3	6714	T20F021C4838	M00333	None	13	13	-76	-64	-84	-63
3	6714	T20F021C4838	M00716	None	8	8	-72	-65	-84	-63
1	6714	T20F021C4838	M00333	None	10	13	-72	-60	-84	-63
3	2886	T17F02184FDB	M00241	None	8	8	-31	-24	-41	-20
1	2886	T17F02184FDB	M00280	5989	4	17	-23	-7	-41	-20
1	2321	T13R01AA206D	M00999	326	14	26	-3	22	-11	10
3	2321	T13R01AA206D	M00640	3201	8	8	-7	0	-11	10
1	2264	T05F0A843489	M00023	None	9	30	-954	-925	-967	-946
1	4860	T14F013149F1	M00803	1871	1	6	-866	-861	-887	-866
1	4860	T14F013149F1	M00803	7027	1	6	-866	-861	-887	-866
1	4860	T14F013149F1	M00803	1874	1	6	-866	-861	-887	-866
1	4860	T14F013149F1	M00803	1869	1	6	-866	-861	-887	-866
2	4860	T14F013149F1	M00808	5076	9	11	-889	-879	-887	-866
2	4860	T14F013149F1	M00808	7849	9	11	-889	-879	-887	-866

2	4860	T14F013149F1	M00808	5079	9	11	-889	-879	-887	-866
2	7014	T16R040D4055	M00380	None	1	30	-640	-611	-611	-590
1	7014	T16R040D4055	M00144	5079	5	28	-594	-567	-611	-590
1	7015	T05R0013C2D7	M00201	1050	3	18	157	174	138	159
1	3481	T11R00203063	M00192	2908	7	19	-283	-265	-298	-277
3	3481	T11R00203063	M00192	2908	19	19	-295	-277	-298	-277
2	3481	T11R00203063	M00192	2908	13	19	-304	-286	-298	-277
3	3221	T12F032469A6	M00464	5454	10	10	-183	-174	-194	-173
2	3221	T12F032469A6	M00695	None	3	7	-773	-767	-769	-748
3	3221	T12F032469A6	M01047	7020	15	15	-762	-748	-769	-748
3	3221	T12F032469A6	M00293	2300	16	16	-190	-175	-194	-173
2	9956	T16F015AE0B8	M00649	4150	3	8	-150	-143	-145	-124
3	9956	T16F015AE0B8	M00716	None	8	8	-105	-98	-119	-98
1	9956	T16F015AE0B8	M00777	6778	10	13	-133	-121	-145	-124
1	9956	T16F015AE0B8	M00777	6776	10	13	-133	-121	-145	-124
1	9956	T16F015AE0B8	M00777	6777	10	13	-133	-121	-145	-124
1	9956	T16F015AE0B8	M00777	6775	10	13	-133	-121	-145	-124
1	9956	T16F015AE0B8	M00777	6772	10	13	-133	-121	-145	-124
1	9956	T16F015AE0B8	M00777	6773	10	13	-133	-121	-145	-124
1	9956	T16F015AE0B8	M00777	6774	10	13	-133	-121	-145	-124
1	9956	T16F015AE0B8	M00084	7593	3	13	-126	-114	-145	-124
2	9956	T16F015AE0B8	M00084	7593	6	13	-126	-114	-119	-98
3	9956	T16F015AE0B8	M00415	6935	9	9	-132	-124	-145	-124
2	4790	T04F062D7027	M00004	None	5	18	-849	-832	-836	-815
1	4583	T11F0010618D	M00539	405	8	20	-954	-935	-968	-947
1	4583	T11F0010618D	M00726	7392	1	6	-947	-942	-968	-947
3	4583	T11F0010618D	M00285	4779	13	13	-959	-947	-968	-947

3	4583	T11F0010618D	M00395	None	9	9	-957	-949	-968	-947
1	3480	T15F05CC041F	M00373	None	14	21	-758	-738	-766	-745
3	2335	T02R0CDFC7B3	M01004	None	11	11	-132	-122	-143	-122
1	2335	T02R0CDFC7B3	M00056	4782	5	29	-126	-98	-143	-122
3	2335	T02R0CDFC7B3	M00293	2300	16	16	-143	-128	-143	-122
2	2335	T02R0CDFC7B3	M00373	None	13	21	-151	-131	-143	-122
1	864	T01R017FF283	M00761	8626	9	10	60	69	47	68
1	864	T01R017FF283	M00761	7157	9	10	60	69	47	68
1	864	T01R017FF283	M00761	7161	9	10	60	69	47	68
2	864	T01R017FF283	M00373	None	15	21	41	61	47	68
3	864	T01R017FF283	M00088	None	13	13	52	64	47	68
3	864	T01R017FF283	M00223	6775	9	9	52	60	47	68
3	864	T01R017FF283	M00223	6778	9	9	52	60	47	68
3	864	T01R017FF283	M00223	6772	9	9	52	60	47	68
3	864	T01R017FF283	M00223	6774	9	9	52	60	47	68
3	864	T01R017FF283	M00223	6773	9	9	52	60	47	68
3	864	T01R017FF283	M00777	6778	13	13	47	59	47	68
3	864	T01R017FF283	M00777	6776	13	13	47	59	47	68
3	864	T01R017FF283	M00777	6777	13	13	47	59	47	68
3	864	T01R017FF283	M00777	6775	13	13	47	59	47	68
3	864	T01R017FF283	M00777	6772	13	13	47	59	47	68
3	864	T01R017FF283	M00777	6773	13	13	47	59	47	68
3	864	T01R017FF283	M00777	6774	13	13	47	59	47	68
1	864	T01R017FF283	M00955	2908	2	27	67	93	47	68
2	5970	T11R03E29B2D	M00695	None	1	7	114	120	120	141
1	130497	T02R01283138	M01044	6910	9	10	-845	-836	-858	-837
3	130497	T02R01283138	M00032	None	10	10	10	19	7	28

3	130497	T02R01283138	M00025	2002	14	14	7	20	7	28
1	130497	T02R01283138	M01109	51385	5	15	-841	-827	-858	-837
1	130497	T02R01283138	M00328	7849	8	18	21	38	7	28
1	25	T09F07E6FBCC	M00460	6776	5	24	-556	-533	-573	-552
1	25	T09F07E6FBCC	M00341	2553	10	12	-561	-550	-573	-552
1	25	T09F07E6FBCC	M00341	2551	10	12	-561	-550	-573	-552
1	25	T09F07E6FBCC	M00224	6772	11	21	-562	-542	-573	-552
1	25	T09F07E6FBCC	M00225	6774	11	21	-562	-542	-573	-552
1	25	T09F07E6FBCC	M00459	6777	8	15	-559	-545	-573	-552
1	25	T09F07E6FBCC	M01004	None	4	11	-555	-545	-573	-552
1	25	T09F07E6FBCC	M01010	3159	3	15	-554	-540	-573	-552
1	25	T09F07E6FBCC	M01010	8091	3	15	-554	-540	-573	-552
1	25	T09F07E6FBCC	M00223	6775	5	9	-556	-548	-573	-552
1	25	T09F07E6FBCC	M00223	6778	5	9	-556	-548	-573	-552
1	25	T09F07E6FBCC	M00223	6772	5	9	-556	-548	-573	-552
1	25	T09F07E6FBCC	M00223	6774	5	9	-556	-548	-573	-552
1	25	T09F07E6FBCC	M00223	6773	5	9	-556	-548	-573	-552
2	5970	T11R03E25271	M00808	5076	1	11	-740	-730	-730	-709
2	5970	T11R03E25271	M00808	7849	1	11	-740	-730	-730	-709
2	5970	T11R03E25271	M00808	5079	1	11	-740	-730	-730	-709
1	5970	T11R03E25271	M00360	None	9	13	-717	-705	-730	-709
1	5970	T11R03E25271	M00244	None	6	12	-714	-703	-730	-709
1	5970	T11R03E25271	M00246	1959	6	12	-714	-703	-730	-709
1	5970	T11R03E25271	M00378	None	11	12	-719	-708	-730	-709
2	5970	T11R03E25271	M00808	5076	3	11	-738	-728	-730	-709
2	5970	T11R03E25271	M00808	7849	3	11	-738	-728	-730	-709
2	5970	T11R03E25271	M00808	5079	3	11	-738	-728	-730	-709

1	5970	T11R03E25271	M01044	6910	1	10	-709	-700	-730	-709
1	26330	T19F026D32A1	M00515	5468	20	23	-969	-947	-971	-950
1	4193	T12F0405CD2F	M00446	None	2	15	7	21	-13	8
1	4193	T12F0405CD2F	M00143	5079	18	28	-9	18	-13	8
2	4017	T08R0162B8AD	M00133	5453	6	15	-926	-912	-917	-896
2	4017	T08R0162B8AD	M00821	4780	8	13	-922	-910	-917	-896
1	4017	T08R0162B8AD	M00991	1044	5	18	-900	-883	-917	-896
1	4017	T08R0162B8AD	M00991	1045	5	18	-900	-883	-917	-896
2	4017	T08R0162B8AD	M00284	4779	13	22	-926	-905	-917	-896
2	4017	T08R0162B8AD	M00037	4778	7	11	-921	-911	-917	-896
3	4017	T08R0162B8AD	M00630	None	9	9	-911	-903	-917	-896
2	4017	T08R0162B8AD	M00517	3727	7	13	-923	-911	-917	-896
2	4017	T08R0162B8AD	M00517	3725	7	13	-923	-911	-917	-896
2	4017	T08R0162B8AD	M00517	8061	7	13	-923	-911	-917	-896
2	4017	T08R0162B8AD	M00517	3726	7	13	-923	-911	-917	-896
2	4017	T08R0162B8AD	M00517	2353	7	13	-923	-911	-917	-896
3	4017	T08R0162B8AD	M00993	6886	10	10	-914	-905	-917	-896
2	2249	T11R0421678A	M01045	7020	4	15	-97	-83	-86	-65
3	4583	T11F000FF0A6	M01003	None	11	11	-927	-917	-929	-908
2	4583	T11F000FF0A6	M00395	None	8	9	-930	-922	-929	-908
3	4583	T11F000FF0A6	M00394	None	9	9	-929	-921	-929	-908
2	4790	T04F062E9997	M00380	None	18	30	127	156	139	160
1	5175	T17R0390211C	M00456	2186	9	14	-456	-443	-469	-448
2	5175	T17R0390211C	M00955	2908	21	27	-475	-449	-469	-448
2	5175	T17R0390211C	M00145	5454	1	16	-484	-469	-469	-448
3	5175	T17R0390211C	M00148	6736	7	7	-454	-448	-469	-448
3	5175	T17R0390211C	M00998	5089	12	12	-469	-458	-469	-448

3	5175	T17R0390211C	M00998	5087	12	12	-469	-458	-469	-448
3	5175	T17R0390211C	M00998	5090	12	12	-469	-458	-469	-448
2	220972	T10R02B4E5B9	M00378	None	10	12	-768	-757	-766	-745
1	220972	T10R02B4E5B9	M00192	2908	18	19	-762	-744	-766	-745
2	324	T05F06B00439	M00373	None	20	21	-606	-586	-605	-584
2	3860	T17R02333152	M01004	None	7	11	-638	-628	-634	-613
1	3860	T17R02333152	M00444	7421	3	15	-615	-601	-634	-613
1	3860	T17R02333152	M00966	7421	7	21	-619	-599	-634	-613
1	3860	T17R02333152	M00966	6256	7	21	-619	-599	-634	-613
1	3860	T17R02333152	M00966	8856	7	21	-619	-599	-634	-613
1	3860	T17R02333152	M00966	9970	7	21	-619	-599	-634	-613
2	6281	T01R08F45E53	M00446	None	7	15	-305	-291	-297	-276
3	6281	T01R08F45E53	M00695	None	7	7	-294	-288	-297	-276
3	6281	T01R08F45E53	M00466	3091	12	12	-292	-281	-297	-276
2	6281	T01R08F45E53	M00378	None	6	12	-303	-292	-297	-276
1	55206	T12R074B39D3	M00955	2908	2	27	141	167	121	142
2	3918	T01F0AD047C6	M00130	27022	5	12	-643	-632	-636	-615
3	3918	T01F0AD047C6	M00148	6736	7	7	-635	-629	-636	-615
2	3918	T01F0AD047C6	M00809	2302	9	13	-640	-628	-636	-615
2	3918	T01F0AD047C6	M00809	2294	9	13	-640	-628	-636	-615
2	3918	T01F0AD047C6	M00809	27022	9	13	-640	-628	-636	-615
2	3918	T01F0AD047C6	M00809	2295	9	13	-640	-628	-636	-615
2	3918	T01F0AD047C6	M00289	2299	9	13	-640	-628	-636	-615
2	3918	T01F0AD047C6	M01012	3169	11	18	-643	-626	-636	-615
2	3918	T01F0AD047C6	M01012	3170	11	18	-643	-626	-636	-615
2	3918	T01F0AD047C6	M01012	3171	11	18	-643	-626	-636	-615
2	3918	T01F0AD047C6	M00130	27022	9	12	-639	-628	-636	-615

2	3918	T01F0AD047C6	M00724	3169	6	11	-641	-631	-636	-615
2	3918	T01F0AD047C6	M00160	6736	11	12	-637	-626	-636	-615
2	3918	T01F0AD047C6	M00148	6736	4	7	-639	-633	-636	-615
2	3918	T01F0AD047C6	M00791	3170	8	13	-641	-629	-636	-615
2	3918	T01F0AD047C6	M00791	3169	8	13	-641	-629	-636	-615
2	3918	T01F0AD047C6	M00791	3171	8	13	-641	-629	-636	-615
2	7431	T10F0108240B	M00919	1869	9	11	-811	-801	-809	-788
2	7431	T10F0108240B	M00919	1871	9	11	-811	-801	-809	-788
2	7431	T10F0108240B	M00919	1874	9	11	-811	-801	-809	-788
2	7431	T10F0108240B	M00919	7027	9	11	-811	-801	-809	-788
2	7431	T10F0108240B	M00918	1871	8	9	-810	-802	-809	-788
2	7431	T10F0108240B	M00918	1869	8	9	-810	-802	-809	-788
2	7431	T10F0108240B	M00918	7027	8	9	-810	-802	-809	-788
2	7431	T10F0108240B	M00918	1874	8	9	-810	-802	-809	-788
3	7431	T10F0108240B	M00426	None	8	8	-809	-802	-809	-788
3	7431	T10F0108240B	M00427	None	8	8	-809	-802	-809	-788
3	7431	T10F0108240B	M00430	None	8	8	-809	-802	-809	-788
3	4583	T11F00105A6B	M00446	None	15	15	-401	-387	-403	-382
3	4583	T11F00105A6B	M01118	7490	9	9	-395	-387	-403	-382
1	5915	T03F0184DAEC	M00006	4205	12	16	-906	-891	-916	-895
1	4313	T16F03390C37	M00626	5989	9	14	-310	-297	-323	-302
1	7430	T06R097C10CE	M00056	4782	2	29	-415	-387	-435	-414
1	7430	T06R097C10CE	M00144	5079	20	28	-433	-406	-435	-414
1	7430	T06R097C10CE	M00097	5080	15	21	-428	-408	-435	-414
1	7430	T06R097C10CE	M00245	1960	7	12	-420	-409	-435	-414
2	7430	T06R097C10CE	M00056	4782	12	29	-452	-424	-435	-414
1	7430	T06R097C10CE	M00328	7849	13	18	-426	-409	-435	-414

2	7430	T06R097C10CE	M00704	7003	3	6	-438	-433	-435	-414

APPENDIX IV: TFBSs with overlapping SNPs

TSS	Gene ID	Accession No.	TSSStart	Strand	Chromosome	SequenceStart	SequenceStop	RefSNP ID	SNP	SNP position	Source
T05F08F6BD67	2878	M00539	150388071	+	chr5	150388123	150388142	rs8177451	C/T	150388140	../snp/forward/non-redundant/genotypes_chr5_YRI_r27_nr.b36_fwd.txt
T08R016224CB	4017	M00237	23209163	-	chr8	23210084	23210066	rs4348504	C/T	23210076	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T12R02D6E495	7480	M00075	47637653	-	chr12	47637594	47637585	rs863855	G/T	47637590	../snp/forward/non-redundant/genotypes_chr12_YRI_r27_nr.b36_fwd.txt
T11R01EE8567	7490	M01045	32408935	-	chr11	32409510	32409496	rs10742277	C/G	32409501	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T11R01ED6902	7490	M00143	32336130	-	chr11	32337118	32337091	rs7118294	C/T	32337097	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T19F02C44336	558	M00033	46416694	+	chr19	46416658	46416671	rs2301236	A/G	46416660	../snp/forward/non-redundant/genotypes_chr19_YRI_r27_nr.b36_fwd.txt
T18R03865B74	596	M00243	59136884	-	chr18	59136755	59136744	rs1800477	C/T	59136753	../snp/forward/non-redundant/genotypes_chr18_YRI_r27_nr.b36_fwd.txt
T09R021779C1	30968	M00144	35092929	-	chr9	35093143	35093116	rs2298310	A/C	35093117	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T02F06522C1D	84417	M01045	106048541	+	chr2	106048030	106048044	rs755651	C/T	106048032	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T03R00C05A1D	5894	M00143	12605981	-	chr3	12606883	12606856	rs7637824	A/C	12606865	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T20R02284781	7052	M00518	36194177	-	chr20	36194086	36194068	rs6023261	A/G	36194071	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T21F0254E2D4	2114	M00378	39117524	+	chr21	39117698	39117709	rs1051476	C/G	39117706	../snp/forward/non-redundant/genotypes_chr21_YRI_r27_nr.b36_fwd.txt
T15F05031C8F	11214	M00378	84089999	+	chr15	84089511	84089522	rs1053992	C/G	84089522	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T09R051CF96F	3190	M00333	85784943	-	chr9	85785555	85785543	rs296890	A/C	85785551	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T16F03DF2E28	1003	M00056	64957992	+	chr16	64957967	64957995	rs11860610	A/G	64957977	../snp/forward/non-redundant/genotypes_chr16_YRI_r27_nr.b36_fwd.txt
T20R01C63BD2	598	M00378	29768658	-	chr20	29769528	29769517	rs13339897	G/T	29769521	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T18R006BB679	284217	M00644	7059065	-	chr18	7059745	7059739	rs8095328	A/C	7059745	../snp/forward/non-redundant/genotypes_chr18_YRI_r27_nr.b36_fwd.txt

T18F00347AEA	7050	M00024	3439338	+	chr18	3439039	3439053	rs8090978	A/G	3439046	../snp/forward/non-redundant/genotypes_chr18_YRI_r27_nr.b36_fwd.txt
T20R02283F93	7052	M00144	36192147	-	chr20	36192298	36192271	rs2076378	C/T	36192272	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T08R00B2F717	1508	M01109	11728663	-	chr8	11728747	11728733	rs1293319	C/T	11728739	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T18R0383753E	596	M00360	58946878	-	chr18	58947433	58947421	rs10503077	C/T	58947425	../snp/forward/non-redundant/genotypes_chr18_YRI_r27_nr.b36_fwd.txt
T06F01D86475	780	M01109	30958709	+	chr6	30958549	30958563	rs1264327	A/G	30958561	../snp/forward/non-redundant/genotypes_chr6_YRI_r27_nr.b36_fwd.txt
T09F0766CBF3	5742	M00143	124177395	+	chr9	124177455	124177482	rs10306133	C/T	124177465	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T17R01DC8DF5	6352	M00144	31231477	-	chr17	31232320	31232293	rs4239253	C/T	31232303	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt
T09F0554EF7C	1612	M00144	89452412	+	chr9	89452459	89452486	rs1927975	C/T	89452483	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T01F0AD047C6	3918	M00143	181422022	+	chr1	181421916	181421943	rs2276543	A/G	181421928	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T01F01F04B4D	3065	M00695	32525133	+	chr1	32524659	32524665	rs11807976	A/G	32524661	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T09F0554EF7C	1612	M00072	89452412	+	chr9	89451827	89451837	rs12343313	C/T	89451828	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T03F00950B65	4968	M00800	9767781	+	chr3	9767402	9767417	rs3219001	C/T	9767417	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T15R053B693F	51458	M00749	87779647	-	chr15	87780302	87780296	rs8034740	C/T	87780299	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T15F050034C5	11214	M00115	83899589	+	chr15	83898783	83898797	rs12442369	C/T	83898796	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T11F00108922	4583	M00143	1083682	+	chr11	1083702	1083729	rs7944723	C/G	1083710	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T14F0641FBFC	9112	M00025	104987644	+	chr14	104987462	104987475	rs4983409	C/T	104987465	../snp/forward/non-redundant/genotypes_chr14_YRI_r27_nr.b36_fwd.txt
T02F0A87B45F	3239	M00033	176665695	+	chr2	176664753	176664766	rs3754983	A/G	176664756	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T01R02BA3E8F	5052	M00716	45760143	-	chr1	45760165	45760158	rs17522918	A/C	45760161	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T07R019E59D8	3204	M00280	27154904	-	chr7	27155722	27155706	rs3807597	C/T	27155720	../snp/forward/non-redundant/genotypes_chr7_YRI_r27_nr.b36_fwd.txt
T13F01CCEEC0	241	M00378	30207680	+	chr13	30206713	30206724	rs3809376	A/C	30206722	../snp/forward/non-redundant/genotypes_chr13_YRI_r27_nr.b36_fwd.txt
T07R005470CD	60	M00733	5533901	-	chr7	5534518	5534504	rs11546899	A/G	5534512	../snp/forward/non-redundant/genotypes_chr7_YRI_r27_nr.b36_fwd.txt
T01F06BA20CB	7482	M00518	112861387	+	chr1	112860734	112860752	rs351372	A/T	112860743	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T11F02180E2A	960	M00539	35130922	+	chr11	35130706	35130725	rs353620	A/G	35130712	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T11R0615F78D	4319	M00005	102102925	-	chr11	102103409	102103392	rs12278632	G/T	102103399	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T03R039429AA	2272	M00446	60041642	-	chr3	60042043	60042029	rs2736780	C/T	60042031	../snp/forward/non-redundant/genotypes_chr3_JPT_r27_nr.b36_fwd.txt
T11R01ED6902	7490	M00221	32336130	-	chr11	32336809	32336799	rs6484572	A/G	32336799	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T17R02AC5068	5245	M00737	44847208	-	chr17	44847180	44847173	rs2277636	A/G	44847176	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt

T15F05C8D01D	3480	M00143	97046557	+	chr15	97045825	97045852	rs1317460	G/T	97045839	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T06F01D86475	780	M00444	30958709	+	chr6	30958113	30958127	rs1264328	A/G	30958121	../snp/forward/non-redundant/genotypes_chr6_YRI_r27_nr.b36_fwd.txt
T19R011B39AB	9244	M00378	18561451	-	chr19	18561790	18561779	rs2074176	C/T	18561788	../snp/forward/non-redundant/genotypes_chr19_YRI_r27_nr.b36_fwd.txt
T17F02173438	8557	M00733	35075128	+	chr17	35074960	35074974	rs931992	G/T	35074961	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt
T20R033E16A6	6790	M00024	54400678	-	chr20	54401025	54401011	rs2104015	A/G	54401019	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T09R02206E8B	7169	M00237	35679883	-	chr9	35680320	35680302	rs2756892	C/T	35680310	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T20R02284781	7052	M00144	36194177	-	chr20	36194454	36194427	rs2076386	A/G	36194428	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T02F0D0D4708	58190	M00327	218973960	+	chr2	218973629	218973649	rs3795985	C/T	218973640	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T06F01D85D98	780	M00004	30956952	+	chr6	30957102	30957119	rs885910	A/T	30957103	../snp/forward/non-redundant/genotypes_chr6_YRI_r27_nr.b36_fwd.txt
T19F0215F99A	898	M00720	34994586	+	chr19	34993656	34993664	rs7257330	A/G	34993663	../snp/forward/non-redundant/genotypes_chr19_YRI_r27_nr.b36_fwd.txt
T17R0390211C	5175	M00947	59777308	-	chr17	59778062	59778048	rs8069746	A/G	59778061	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt
T03F00950821	4968	M00449	9766945	+	chr3	9766950	9766958	rs1801126	G/T	9766953	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T06R01DF9F64	3106	M00327	31432548	-	chr6	31433051	31433031	rs2596487	C/T	31433035	../snp/forward/non-redundant/genotypes_chr6_YRI_r27_nr.b36_fwd.txt
T08R01A3FB63	1191	M00378	27523939	-	chr8	27523874	27523863	rs9331894	C/T	27523863	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T16R004A5922	5493	M00144	4872482	-	chr16	4872684	4872657	rs1049223	A/G	4872660	../snp/forward/non-redundant/genotypes_chr16_YRI_r27_nr.b36_fwd.txt
T11R03E285CB	5970	M00448	65177035	-	chr11	65177341	65177333	rs10896027	C/G	65177336	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T01R0D7ACD43	7483	M00143	226151747	-	chr1	226152520	226152493	rs10127974	A/G	226152497	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T15F03A4BB08	7168	M00056	61127432	+	chr15	61127270	61127298	rs4075583	A/G	61127280	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T14F020EC352	6729	M00025	34521938	+	chr14	34521882	34521895	rs12890307	A/G	34521883	../snp/forward/non-redundant/genotypes_chr14_YRI_r27_nr.b36_fwd.txt
T18F003483C0	7050	M00056	3441600	+	chr18	3441759	3441787	rs238533	C/T	3441762	../snp/forward/non-redundant/genotypes_chr18_YRI_r27_nr.b36_fwd.txt
T08R0162E0DB	4017	M00539	23257307	-	chr8	23257268	23257249	rs9785120	C/T	23257258	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T15F03CFC853	8766	M00143	63948883	+	chr15	63948777	63948804	rs12909265	C/G	63948777	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T16F0359B808	9289	M01109	56211464	+	chr16	56210824	56210838	rs11862210	A/G	56210838	../snp/forward/non-redundant/genotypes_chr16_YRI_r27_nr.b36_fwd.txt
T14R02369EA1	3169	M00649	37133985	-	chr14	37134496	37134489	rs10400689	G/T	37134489	../snp/forward/non-redundant/genotypes_chr14_YRI_r27_nr.b36_fwd.txt
T08R00A2115F	83595	M00143	10621279	-	chr8	10621636	10621609	rs10100209	C/T	10621622	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T13F02D90090	5925	M00032	47775888	+	chr13	47775700	47775709	rs4151418	C/T	47775708	../snp/forward/non-redundant/genotypes_chr13_TSI_r27_nr.b36_fwd.txt
T10F07D2423C	4255	M00189	131220028	+	chr10	131219103	131219114	rs7900620	G/T	131219110	../snp/forward/non-redundant/genotypes_chr10_JPT_r27_nr.b36_fwd.txt

T02F0DF7A5D7	54575	M00947	234333655	+	chr2	234332864	234332878	rs9646717	A/G	234332877	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T02F0DF7A5D7	54576	M00947	234333655	+	chr2	234332864	234332878	rs9646717	A/G	234332877	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T22F0267A678	2547	M00378	40347256	+	chr22	40346643	40346654	rs2267437	C/G	40346645	../snp/forward/non-redundant/genotypes_chr22_YRI_r27_nr.b36_fwd.txt
T02F0D0D42A3	58190	M00426	218972835	+	chr2	218972347	218972354	rs2739049	A/G	218972351	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T03R03007B32	10641	M00327	50363186	-	chr3	50363496	50363476	rs2233475	A/G	50363491	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T08F02EC3869	4173	M00378	49035369	+	chr8	49035049	49035060	rs8178000	C/T	49035058	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T12F0323C141	3225	M00008	52674881	+	chr12	52674394	52674403	rs12314066	A/G	52674397	../snp/forward/non-redundant/genotypes_chr12_YRI_r27_nr.b36_fwd.txt
T11F00106482	4583	M00115	1074306	+	chr11	1073942	1073956	rs11245935	C/T	1073946	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T11R00DA6EE1	22800	M01104	14315233	-	chr11	14315855	14315849	rs2349297	C/G	14315850	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T07R019EF42B	3209	M00695	27194411	-	chr7	27195054	27195048	rs17501146	C/T	27195049	../snp/forward/non-redundant/genotypes_chr7_YRI_r27_nr.b36_fwd.txt
T02F0DF64CA3	54575	M00646	234245283	+	chr2	234245231	234245238	rs17868322	A/G	234245233	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T02F0DF64CA3	54576	M00646	234245283	+	chr2	234245231	234245238	rs17868322	A/G	234245233	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T11F0011F6D7	4586	M00378	1177303	+	chr11	1177325	1177336	rs28514396	C/T	1177336	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T17F021643B2	84152	M00539	35013554	+	chr17	35013454	35013473	rs9910883	C/T	35013460	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt
T20F0038F8EC	994	M00341	3733740	+	chr20	3733808	3733819	rs11570020	C/T	3733819	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T11R0615F78D	4319	M01109	102102925	-	chr11	102103819	102103805	rs3758861	A/G	102103810	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T18R006BB679	284217	M00143	7059065	-	chr18	7059373	7059346	rs7238600	A/G	7059358	../snp/forward/non-redundant/genotypes_chr18_YRI_r27_nr.b36_fwd.txt
T17F001D11FB	3090	M00947	1905147	+	chr17	1904634	1904648	rs8065820	A/G	1904643	../snp/forward/non-redundant/genotypes_chr17_TSI_r27_nr.b36_fwd.txt
T18F0009E316	7298	M00243	647958	+	chr18	648060	648071	rs2853533	C/G	648064	../snp/forward/non-redundant/genotypes_chr18_YRI_r27_nr.b36_fwd.txt
T11R00203EC8	3481	M00144	2113224	-	chr11	2113740	2113713	rs3213219	C/T	2113726	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T08F07AD9FF8	4609	M00449	128819192	+	chr8	128818917	128818925	rs4645954	A/G	128818920	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T01F0C914BD6	467	M00327	210848726	+	chr1	210847807	210847827	rs11571516	G/T	210847808	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T22F0267A678	2547	M00114	40347256	+	chr22	40347196	40347210	rs132770	A/G	40347210	../snp/forward/non-redundant/genotypes_chr22_YRI_r27_nr.b36_fwd.txt
T13R01880EF5	6049	M00695	25693941	-	chr13	25694410	25694404	rs2277427	A/C	25694408	../snp/forward/non-redundant/genotypes_chr13_YRI_r27_nr.b36_fwd.txt
T12R074AE462	55206	M00716	122348642	-	chr12	122348823	122348816	rs12582153	C/T	122348819	../snp/forward/non-redundant/genotypes_chr12_YRI_r27_nr.b36_fwd.txt
T17F02184FDB	2886	M00716	35147739	+	chr17	35147256	35147263	rs4794818	C/G	35147259	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt
T17F021632C7	84152	M00056	35009223	+	chr17	35009235	35009263	rs998480	A/T	35009251	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt

T01F093DCE48	4914	M00665	155045448	+	chr1	155044952	155044965	rs6678281	A/G	155044963	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T01F0C3403C0	9641	M00539	204735424	+	chr1	204735079	204735098	rs17025213	C/T	204735080	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T06F01D86475	780	M00378	30958709	+	chr6	30957790	30957801	rs9295930	A/G	30957801	../snp/forward/non-redundant/genotypes_chr6_YRI_r27_nr.b36_fwd.txt
T01F0C33A22A	9641	M00056	204710442	+	chr1	204710155	204710183	rs3813977	C/T	204710157	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T11F041F6031	595	M00143	69165105	+	chr11	69164266	69164293	rs3212860	A/T	69164288	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T16F04E875E0	1012	M00033	82343392	+	chr16	82343051	82343064	rs448802	C/G	82343055	../snp/forward/non-redundant/genotypes_chr16_YRI_r27_nr.b36_fwd.txt
T15F03A515A7	7168	M00539	61150631	+	chr15	61150071	61150090	rs11632190	A/C	61150076	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T11R03E1F583	5970	M00518	65140099	-	chr11	65141073	65141055	rs1787031	G/T	65141068	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T19R03041DF2	2067	M00280	50601458	-	chr19	50601485	50601469	rs34014819	G/T	50601474	../snp/forward/non-redundant/genotypes_chr19_TSI_r27_nr.b36_fwd.txt
T05F0714FD5B	3295	M00075	118816091	+	chr5	118815408	118815417	rs35695978	A/G	118815408	../snp/forward/non-redundant/genotypes_chr5_TSI_r27_nr.b36_fwd.txt
T05F009D37D1	22948	M00803	10303441	+	chr5	10302790	10302795	rs2548551	G/T	10302790	../snp/forward/non-redundant/genotypes_chr5_YRI_r27_nr.b36_fwd.txt
T07R019EF42B	3209	M00025	27194411	-	chr7	27194819	27194806	rs17472294	C/T	27194811	../snp/forward/non-redundant/genotypes_chr7_YRI_r27_nr.b36_fwd.txt
T20R01C5863F	598	M00426	29722175	-	chr20	29722203	29722196	rs10439607	A/G	29722202	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T01R08F45E53	6281	M00539	150232659	-	chr1	150233375	150233356	rs1038745	G/T	150233369	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T19F029D82D2	81	M00450	43877074	+	chr19	43876655	43876663	rs7252630	G/T	43876662	../snp/forward/non-redundant/genotypes_chr19_YRI_r27_nr.b36_fwd.txt
T05F0415352D	891	M00644	68498733	+	chr5	68498504	68498510	rs2069429	A/G	68498509	../snp/forward/non-redundant/genotypes_chr5_YRI_r27_nr.b36_fwd.txt
T08R0162B82E	4017	M00025	23246894	-	chr8	23246886	23246873	rs1010156	C/T	23246886	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T09F0552EDCB	1612	M00221	89320907	+	chr9	89320713	89320723	rs12348866	A/G	89320720	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T19F00F4EA98	7171	M00446	16050840	+	chr19	16050115	16050129	rs10402928	A/G	16050123	../snp/forward/non-redundant/genotypes_chr19_YRI_r27_nr.b36_fwd.txt
T13F01EA456B	23047	M00426	32130411	+	chr13	32130330	32130337	rs7335546	A/G	32130334	../snp/forward/non-redundant/genotypes_chr13_YRI_r27_nr.b36_fwd.txt
T09F07D609E5	51450	M00378	131467749	+	chr9	131467116	131467127	rs28562184	C/T	131467122	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T12F0324A56A	3221	M00056	52733290	+	chr12	52733276	52733304	rs2277372	G/T	52733280	../snp/forward/non-redundant/genotypes_chr12_JPT_r27_nr.b36_fwd.txt
T18R006A80E4	284217	M01113	6979812	-	chr18	6979814	6979807	rs492449	A/G	6979809	../snp/forward/non-redundant/genotypes_chr18_YRI_r27_nr.b36_fwd.txt
T06F01D86475	780	M00444	30958709	+	chr6	30958102	30958116	rs13197561	G/T	30958111	../snp/forward/non-redundant/genotypes_chr6_YRI_r27_nr.b36_fwd.txt
T17R023FE9A7	6774	M00733	37743015	-	chr17	37743277	37743263	rs12721576	C/G	37743267	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt
T20R033E16A6	6790	M00448	54400678	-	chr20	54401572	54401564	rs6064387	A/G	54401566	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T03F0184FD6D	5915	M00749	25492845	+	chr3	25493034	25493040	rs17016482	C/G	25493037	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt

T0XR0924A9F6	2539	M00800	153397750	-	chrX	153398734	153398719	rs5986989	A/G	153398725	../snp/forward/non-redundant/genotypes_chrX_YRI_r27_nr.b36_fwd.txt
T08F08915B2E	8000	M00144	143743790	+	chr8	143743417	143743444	rs754957	A/G	143743418	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T01R0384964B	3725	M01079	59020875	-	chr1	59021634	59021619	rs2811894	G/T	59021619	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T11R06D3F6E4	23705	M00446	114554596	-	chr11	114554937	114554923	rs17118046	A/G	114554937	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T11F001252E8	727897	M00378	1200872	+	chr11	1199956	1199967	rs885454	A/G	1199967	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T14F04759734	2353	M00144	74815284	+	chr14	74815184	74815211	rs4645852	A/C	74815200	../snp/forward/non-redundant/genotypes_chr14_YRI_r27_nr.b36_fwd.txt
T12F0323C141	3225	M00024	52674881	+	chr12	52674735	52674749	rs4759316	A/G	52674740	../snp/forward/non-redundant/genotypes_chr12_YRI_r27_nr.b36_fwd.txt
T02F0DE28F99	250	M00056	232951705	+	chr2	232951806	232951834	rs1130335	C/T	232951830	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T20F02A086C9	4318	M00378	44074697	+	chr20	44073795	44073806	rs3918254	C/T	44073798	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T08R00C94BFB	10395	M00024	13192187	-	chr8	13192559	13192545	rs7013667	A/G	13192552	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T09F0766CBF3	5742	M00243	124177395	+	chr9	124177098	124177109	rs10306128	A/G	124177102	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T16F0168128F	5347	M00539	23597711	+	chr16	23597506	23597525	rs16972787	A/G	23597522	../snp/forward/non-redundant/genotypes_chr16_YRI_r27_nr.b36_fwd.txt
T20F0038D69D	994	M00947	3724957	+	chr20	3724966	3724980	rs6116042	A/T	3724968	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T03R08F3C046	6596	M00531	150192198	-	chr3	150192555	150192538	rs13100699	A/G	150192544	../snp/forward/non-redundant/genotypes_chr3_LWK_r27_nr.b36_fwd.txt
T16F035A03F9	9289	M00449	56230905	+	chr16	56230961	56230969	rs3760061	A/G	56230961	../snp/forward/non-redundant/genotypes_chr16_YRI_r27_nr.b36_fwd.txt
T11R00203166	3481	M00800	2109798	-	chr11	2110776	2110761	rs3741214	A/G	2110767	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T07F0348F5A8	1956	M00143	55113128	+	chr7	55112813	55112840	rs9642564	A/G	55112828	../snp/forward/non-redundant/genotypes_chr7_YRI_r27_nr.b36_fwd.txt
T03R039429AA	2272	M01079	60041642	-	chr3	60042447	60042432	rs17062113	A/C	60042436	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T07R019E6266	3204	M00531	27157094	-	chr7	27157195	27157178	rs886339	A/C	27157195	../snp/forward/non-redundant/genotypes_chr7_YRI_r27_nr.b36_fwd.txt
T0XR09252298	2539	M00144	153428632	-	chrX	153429307	153429280	rs5986992	A/C	153429301	../snp/forward/non-redundant/genotypes_chrX_YRI_r27_nr.b36_fwd.txt
T16R0412730C	1728	M00378	68317964	-	chr16	68318564	68318553	rs689458	C/G	68318561	../snp/forward/non-redundant/genotypes_chr16_YRI_r27_nr.b36_fwd.txt
T08F00B27E19	2222	M00378	11697689	+	chr8	11697769	11697780	rs1047643	C/T	11697771	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T03F00952607	4968	M00929	9774599	+	chr3	9774109	9774126	rs293795	A/G	9774113	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T05F08F6BD67	2878	M00056	150388071	+	chr5	150388106	150388134	rs2070593	A/G	150388133	../snp/forward/non-redundant/genotypes_chr5_YRI_r27_nr.b36_fwd.txt
T19F039C8438	6158	M00665	60589112	+	chr19	60588518	60588531	rs3810168	C/T	60588521	../snp/forward/non-redundant/genotypes_chr19_YRI_r27_nr.b36_fwd.txt
T19R011AE29B	9244	M00033	18539163	-	chr19	18539194	18539181	rs2286328	C/T	18539183	../snp/forward/non-redundant/genotypes_chr19_MEX_r27_nr.b36_fwd.txt
T05F08F6BD67	2878	M00056	150388071	+	chr5	150388106	150388134	rs8177448	A/G	150388128	../snp/forward/non-redundant/genotypes_chr5_YRI_r27_nr.b36_fwd.txt

T19F039C8438	6158	M00147	60589112	+	chr19	60588976	60588985	rs4806668	G/T	60588982	../snp/forward/non-redundant/genotypes_chr19_YRI_r27_nr.b36_fwd.txt
T08F0285ADA4	5423	M00916	42315172	+	chr8	42315119	42315132	rs2307155	A/C	42315132	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T15F05024278	11214	M00033	84034168	+	chr15	84034012	84034025	rs7181796	C/G	84034013	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T19F029D82D2	81	M00426	43877074	+	chr19	43876645	43876652	rs7251903	A/G	43876651	../snp/forward/non-redundant/genotypes_chr19_YRI_r27_nr.b36_fwd.txt
T11R000F6A1D	4588	M00378	1010205	-	chr11	1010785	1010774	rs11602663	C/T	1010778	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T02R009FAB95	4953	M00113	10464149	-	chr2	10464612	10464601	rs3771134	C/T	10464611	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T01F0C916AE7	467	M00539	210856679	+	chr1	210856017	210856036	rs9429888	A/G	210856022	../snp/forward/non-redundant/genotypes_chr1_YRI_r27_nr.b36_fwd.txt
T02R0CE00838	2335	M00800	216008760	-	chr2	216008730	216008715	rs1250259	A/T	216008727	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T20R01C6508C	598	M01045	29773964	-	chr20	29774103	29774089	rs6087772	C/T	29774094	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T17R02AC5068	5245	M00084	44847208	-	chr17	44847597	44847585	rs2233659	G/T	44847586	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt
T06F01E390EE	199	M00446	31690990	+	chr6	31690000	31690014	rs3132451	C/G	31690004	../snp/forward/non-redundant/genotypes_chr6_YRI_r27_nr.b36_fwd.txt
T11F00108CAD	4583	M00237	1084589	+	chr11	1084348	1084366	rs10902089	A/G	1084357	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T06R097C10CE	7430	M00716	159125710	-	chr6	159125626	159125619	rs9364995	C/T	159125620	../snp/forward/non-redundant/genotypes_chr6_YRI_r27_nr.b36_fwd.txt
T10R0059DB7C	2665	M01034	5888892	-	chr10	5889593	5889584	rs12264073	C/T	5889588	../snp/forward/non-redundant/genotypes_chr10_YRI_r27_nr.b36_fwd.txt
T04F047F457E	2069	M00143	75449726	+	chr4	75449794	75449821	rs6836436	G/T	75449794	../snp/forward/non-redundant/genotypes_chr4_YRI_r27_nr.b36_fwd.txt
T03F01844198	5915	M00017	25444760	+	chr3	25443839	25443852	rs9827454	A/G	25443850	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T20F01D672D4	1789	M00143	30831316	+	chr20	30830622	30830649	rs6058884	A/G	30830632	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T05F06B0193F	324	M00084	112204095	+	chr5	112204216	112204228	rs42427	A/G	112204224	../snp/forward/non-redundant/genotypes_chr5_YRI_r27_nr.b36_fwd.txt
T08R08018A7D	10397	M00378	134318717	-	chr8	134319357	134319346	rs1049697	A/G	134319347	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T17F0464E849	332	M00947	73721929	+	chr17	73721788	73721802	rs17885280	A/G	73721802	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt
T14R05401820	11099	M00947	88086560	-	chr14	88086770	88086756	rs891751	A/G	88086770	../snp/forward/non-redundant/genotypes_chr14_YRI_r27_nr.b36_fwd.txt
T02R087FFAD1	53353	M00072	142605009	-	chr2	142605332	142605322	rs13430169	C/G	142605323	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T05F08E8D7A0	1044	M00449	149477280	+	chr5	149477363	149477371	rs246391	C/T	149477370	../snp/forward/non-redundant/genotypes_chr5_YRI_r27_nr.b36_fwd.txt
T11F021472FC	8050	M00716	34894588	+	chr11	34893986	34893993	rs3763931	A/G	34893992	../snp/forward/non-redundant/genotypes_chr11_YRI_r27_nr.b36_fwd.txt
T16F033936C4	4313	M00115	54081220	+	chr16	54081039	54081053	rs17242473	A/G	54081046	../snp/forward/non-redundant/genotypes_chr16_YRI_r27_nr.b36_fwd.txt
T19F033A4354	581	M00004	54149972	+	chr19	54149746	54149763	rs4645878	A/G	54149750	../snp/forward/non-redundant/genotypes_chr19_YRI_r27_nr.b36_fwd.txt
T08R01632081	4017	M00733	23273601	-	chr8	23274052	23274038	rs7010362	A/G	23274046	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt

T16F03DF2E28	1003	M00448	64957992	+	chr16	64957650	64957658	rs11859765	G/T	64957656	../snp/forward/non-redundant/genotypes_chr16_YRI_r27_nr.b36_fwd.txt
T09R051CF6BF	3190	M00378	85784255	-	chr9	85784169	85784158	rs12002509	C/T	85784166	../snp/forward/non-redundant/genotypes_chr9_JPT_r27_nr.b36_fwd.txt
T06F01E390EE	199	M00539	31690990	+	chr6	31690255	31690274	rs2857600	C/T	31690266	../snp/forward/non-redundant/genotypes_chr6_TSI_r27_nr.b36_fwd.txt
T20F01C486B1	3397	M00378	29656753	+	chr20	29656558	29656569	rs6060260	C/T	29656568	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T15F03A4E791	7168	M00017	61138833	+	chr15	61138534	61138547	rs4238372	A/G	61138541	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T04F0475BDBE	3576	M00025	74825150	+	chr4	74825206	74825219	rs2227537	A/G	74825212	../snp/forward/non-redundant/genotypes_chr4_YRI_r27_nr.b36_fwd.txt
T02R0CDFD2D4	2335	M00144	215995092	-	chr2	215995746	215995719	rs16854041	A/G	215995742	../snp/forward/non-redundant/genotypes_chr2_YRI_r27_nr.b36_fwd.txt
T09F0766CBF3	5742	M00004	124177395	+	chr9	124177490	124177507	rs10306134	A/G	124177493	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T08F08919643	8000	M00947	143758915	+	chr8	143758134	143758148	rs2920280	C/G	143758146	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T09R051CF6BF	3190	M00938	85784255	-	chr9	85784901	85784886	rs296887	C/T	85784890	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T05R07DB6C78	3659	M00221	131820664	-	chr5	131821195	131821185	rs4705950	C/T	131821185	../snp/forward/non-redundant/genotypes_chr5_YRI_r27_nr.b36_fwd.txt
T12F06215EA2	7184	M00327	102850210	+	chr12	102849577	102849597	rs17034931	C/G	102849589	../snp/forward/non-redundant/genotypes_chr12_YRI_r27_nr.b36_fwd.txt
T05R0ABA043F	2324	M00144	179962943	-	chr5	179963756	179963729	rs11739750	C/T	179963737	../snp/forward/non-redundant/genotypes_chr5_YRI_r27_nr.b36_fwd.txt
T10R003A27CE	1316	M00143	3811278	-	chr10	3811565	3811538	rs17731	A/G	3811561	../snp/forward/non-redundant/genotypes_chr10_YRI_r27_nr.b36_fwd.txt
T09F055621E6	1514	M00143	89530854	+	chr9	89530659	89530686	rs3118869	A/C	89530683	../snp/forward/non-redundant/genotypes_chr9_YRI_r27_nr.b36_fwd.txt
T22R013994D3	5594	M01113	20550867	-	chr22	20551388	20551381	rs5999869	A/C	20551383	../snp/forward/non-redundant/genotypes_chr22_YRI_r27_nr.b36_fwd.txt
T08R08018A7D	10397	M00518	134318717	-	chr8	134319185	134319167	rs10675	C/T	134319168	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T18R0069967C	284217	M00378	6919804	-	chr18	6920064	6920053	rs9960153	A/G	6920055	../snp/forward/non-redundant/genotypes_chr18_JPT_r27_nr.b36_fwd.txt
T03R00C090CD	5894	M01113	12619981	-	chr3	12620011	12620004	rs2596828	C/T	12620007	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T19F029D1069	81	M00143	43847785	+	chr19	43847698	43847725	rs3786835	A/G	43847720	../snp/forward/non-redundant/genotypes_chr19_YRI_r27_nr.b36_fwd.txt
T02F0D1CCD2F	1674	M00196	219991343	+	chr2	219990558	219990570	rs1039896	A/G	219990565	../snp/forward/non-redundant/genotypes_chr2_TSI_r27_nr.b36_fwd.txt
T08R0162B82E	4017	M00450	23246894	-	chr8	23246787	23246779	rs2294129	A/C	23246787	../snp/forward/non-redundant/genotypes_chr8_YRI_r27_nr.b36_fwd.txt
T13F02D99872	5925	M00810	47814770	+	chr13	47814882	47814899	rs520342	C/T	47814896	../snp/forward/non-redundant/genotypes_chr13_YRI_r27_nr.b36_fwd.txt
T06R0989C054	6648	M00378	160022612	-	chr6	160022705	160022694	rs5746141	C/T	160022700	../snp/forward/non-redundant/genotypes_chr6_YRI_r27_nr.b36_fwd.txt
T03R0BBDB347	4585	M00139	196981575	-	chr3	196982338	196982321	rs2641781	C/T	196982329	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T17F02169DFE	84152	M00539	35036670	+	chr17	35035919	35035938	rs10852934	C/T	35035936	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt
T17F02BE3EF0	8913	M00665	46022384	+	chr17	46022017	46022030	rs2240222	C/G	46022030	../snp/forward/non-redundant/genotypes_chr17_YRI_r27_nr.b36_fwd.txt

T20F003867FE	994	M00143	3696638	+	chr20	3695978	3696005	rs2295346	A/C	3695992	../snp/forward/non-redundant/genotypes_chr20_YRI_r27_nr.b36_fwd.txt
T03R00C090CD	5894	M00280	12619981	-	chr3	12619892	12619876	rs5746212	A/T	12619881	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T11F041F6031	595	M00056	69165105	+	chr11	69164617	69164645	rs3212861	A/G	69164632	../snp/forward/non-redundant/genotypes_chr11_TSI_r27_nr.b36_fwd.txt
T15F05C9DC4E	3480	M00056	97115214	+	chr15	97115023	97115051	rs8027767	A/C	97115026	../snp/forward/non-redundant/genotypes_chr15_YRI_r27_nr.b36_fwd.txt
T03R03DA74A0	56999	M00144	64648352	-	chr3	64648857	64648830	rs6789298	C/T	64648855	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T03R09880DD8	5918	M00448	159911384	-	chr3	159911343	159911335	rs2307064	A/G	159911336	../snp/forward/non-redundant/genotypes_chr3_YRI_r27_nr.b36_fwd.txt
T06F01D86475	780	M00801	30958709	+	chr6	30958556	30958561	rs1264327	C/T	30958561	../snp/reverse/non-redundant/genotypes_chr6_YRI_r24_nr.b36.txt
T06R0989C054	6648	M00449	160022612	-	chr6	160022705	160022697	rs5746141	A/G	160022700	../snp/reverse/non-redundant/genotypes_chr6_YRI_r24_nr.b36.txt
T11R00203254	3481	M00378	2110036	-	chr11	2110976	2110965	rs14367	G/T	2110976	../snp/reverse/non-redundant/genotypes_chr11_YRI_r24_nr.b36.txt
T18R03865B74	596	M00982	59136884	-	chr18	59136756	59136743	rs1800477	A/G	59136753	../snp/reverse/non-redundant/genotypes_chr18_YRI_r24_nr.b36.txt
T11F06F2DF86	6876	M00378	116580230	+	chr11	116579499	116579510	rs2269397	A/G	116579509	../snp/reverse/non-redundant/genotypes_chr11_YRI_r24_nr.b36.txt
T08F07AD9959	4609	M00333	128817497	+	chr8	128817123	128817135	rs3824120	A/C	128817135	../snp/reverse/non-redundant/genotypes_chr8_YRI_r24_nr.b36.txt
T01R08ADD126	51205	M00800	145608998	-	chr1	145609327	145609312	rs2353522	A/G	145609315	../snp/reverse/non-redundant/genotypes_chr1_YRI_r24_nr.b36.txt
T03R0BBDB347	4585	M00327	196981575	-	chr3	196982344	196982324	rs2641781	A/G	196982329	../snp/reverse/non-redundant/genotypes_chr3_YRI_r24_nr.b36.txt
T08R0162B82E	4017	M00025	23246894	-	chr8	23246884	23246871	rs2294128	A/G	23246871	../snp/reverse/non-redundant/genotypes_chr8_YRI_r24_nr.b36.txt
T03R0A5AD10A	8743	M00144	173723914	-	chr3	173724460	173724433	rs3136586	A/G	173724453	../snp/reverse/non-redundant/genotypes_chr3_YRI_r24_nr.b36.txt
T01F0C3403C0	9641	M00327	204735424	+	chr1	204735210	204735230	rs944775	A/G	204735216	../snp/reverse/non-redundant/genotypes_chr1_YRI_r24_nr.b36.txt
T03R00C0B214	5894	M00449	12628500	-	chr3	12628582	12628574	rs3730268	C/T	12628578	../snp/reverse/non-redundant/genotypes_chr3_YRI_r24_nr.b36.txt
T08R00B2D17C	1508	M00327	11719036	-	chr8	11719987	11719967	rs2645427	G/T	11719975	../snp/reverse/non-redundant/genotypes_chr8_YRI_r24_nr.b36.txt
T11F0217D992	960	M00050	35117458	+	chr11	35117648	35117655	rs3751031	A/G	35117652	../snp/reverse/non-redundant/genotypes_chr11_YRI_r24_nr.b36.txt
T11R002029A6	3481	M00143	2107814	-	chr11	2108644	2108617	rs3802971	C/T	2108628	../snp/reverse/non-redundant/genotypes_chr11_YRI_r24_nr.b36.txt
T09R051CF96F	3190	M00803	85784943	-	chr9	85785554	85785549	rs296890	G/T	85785551	../snp/reverse/non-redundant/genotypes_chr9_YRI_r24_nr.b36.txt
T11R00203EC8	3481	M00327	2113224	-	chr11	2113741	2113721	rs3213219	A/G	2113726	../snp/reverse/non-redundant/genotypes_chr11_JPT_r24_nr.b36.txt
T11R00203EC8	3481	M00115	2113224	-	chr11	2113116	2113102	rs3213225	C/T	2113112	../snp/reverse/non-redundant/genotypes_chr11_JPT_r24_nr.b36.txt
T05R0259A87B	1601	M00056	39430267	-	chr5	39431164	39431136	rs2855510	C/T	39431154	../snp/reverse/non-redundant/genotypes_chr5_YRI_r24_nr.b36.txt
T08F02EC3869	4173	M00446	49035369	+	chr8	49035047	49035061	rs8178000	A/G	49035058	../snp/reverse/non-redundant/genotypes_chr8_YRI_r24_nr.b36.txt
T01R08F45E53	6281	M00539	150232659	-	chr1	150233375	150233356	rs1038745	A/C	150233369	../snp/reverse/non-redundant/genotypes_chr1_YRI_r24_nr.b36.txt

T12R00D160BC	2904	M00327	13721788	-	chr12	13722618	13722598	rs2193150	G/T	13722606	../snp/reverse/non-redundant/genotypes_chr12_YRI_r24_nr.b36.txt
T11F0669D4A8	472	M00446	107599016	+	chr11	107598406	107598420	rs228589	A/T	107598418	../snp/reverse/non-redundant/genotypes_chr11_YRI_r24_nr.b36.txt
T11F001252E8	727897	M00450	1200872	+	chr11	1199959	1199967	rs885454	C/T	1199967	../snp/reverse/non-redundant/genotypes_chr11_YRI_r24_nr.b36.txt
T20R039A3B3B	140628	M01079	60439355	-	chr20	60439807	60439792	rs1109400	C/T	60439800	../snp/reverse/non-redundant/genotypes_chr20_YRI_r24_nr.b36.txt
T11R03E29DFF	5970	M01113	65183231	-	chr11	65184151	65184144	rs732072	C/T	65184144	../snp/reverse/non-redundant/genotypes_chr11_YRI_r24_nr.b36.txt
T08R08018A7D	10397	M00327	134318717	-	chr8	134319174	134319154	rs10675	A/G	134319168	../snp/reverse/non-redundant/genotypes_chr8_YRI_r24_nr.b36.txt
T10R07BC8C6D	4288	M00221	129797229	-	chr10	129797883	129797873	rs2857027	A/G	129797874	../snp/reverse/non-redundant/genotypes_chr10_YRI_r24_nr.b36.txt
T03R0BBDB347	4585	M00929	196981575	-	chr3	196981772	196981755	rs2688492	A/G	196981763	../snp/reverse/non-redundant/genotypes_chr3_YRI_r24_nr.b36.txt
T12R07DCDF01	55743	M00025	131915521	-	chr12	131915600	131915587	rs3825109	C/T	131915595	../snp/reverse/non-redundant/genotypes_chr12_JPT_r24_nr.b36.txt
T19R0303913C	2068	M00448	50565436	-	chr19	50565595	50565587	rs3826908	A/G	50565589	../snp/reverse/non-redundant/genotypes_chr19_YRI_r24_nr.b36.txt
T11R00203190	3481	M00025	2109840	-	chr11	2110806	2110793	rs12993	G/T	2110796	../snp/reverse/non-redundant/genotypes_chr11_YRI_r24_nr.b36.txt
T17F02173438	8557	M01073	35075128	+	chr17	35074956	35074973	rs931992	A/C	35074961	../snp/reverse/non-redundant/genotypes_chr17_YRI_r24_nr.b36.txt
T03R02F15DF3	387	M00143	49372659	-	chr3	49372747	49372720	rs11552757	C/T	49372746	../snp/reverse/non-redundant/genotypes_chr3_YRI_r24_nr.b36.txt
T05F09F51304	57451	M00425	167056132	+	chr5	167056301	167056308	rs898172	C/T	167056303	../snp/reverse/non-redundant/genotypes_chr5_YRI_r24_nr.b36.txt
T05F009D37D1	22948	M00025	10303441	+	chr5	10303433	10303446	rs2578618	A/G	10303443	../snp/reverse/non-redundant/genotypes_chr5_YRI_r24_nr.b36.txt
T06F07F87E5E	2070	M00444	133725790	+	chr6	133725653	133725667	rs212831	A/G	133725656	../snp/reverse/non-redundant/genotypes_chr6_YRI_r24_nr.b36.txt
T02R0900C475	390	M00177	151045237	-	chr2	151045093	151045082	rs1528429	C/T	151045088	../snp/reverse/non-redundant/genotypes_chr2_YRI_r24_nr.b36.txt
T08R0162B82E	4017	M00056	23246894	-	chr8	23246958	23246930	rs2294127	A/G	23246940	../snp/reverse/non-redundant/genotypes_chr8_YRI_r24_nr.b36.txt
T09F0766D679	5742	M00025	124180089	+	chr9	124180139	124180152	rs876567	C/T	124180147	../snp/reverse/non-redundant/genotypes_chr9_YRI_r24_nr.b36.txt
T08F046AFF69	7013	M00360	74121065	+	chr8	74120577	74120589	rs3863242	A/G	74120587	../snp/reverse/non-redundant/genotypes_chr8_YRI_r24_nr.b36.txt
T20R02286A7B	7052	M01109	36203131	-	chr20	36203906	36203892	rs2229471	C/T	36203897	../snp/reverse/non-redundant/genotypes_chr20_YRI_r24_nr.b36.txt
T14R05401820	11099	M00144	88086560	-	chr14	88086772	88086745	rs891751	C/T	88086770	../snp/reverse/non-redundant/genotypes_chr14_YRI_r24_nr.b36.txt
T02R0CDF273B	2335	M00004	215951163	-	chr2	215952094	215952077	rs1250204	A/G	215952077	../snp/reverse/non-redundant/genotypes_chr2_YRI_r24_nr.b36.txt
T03R00C08240	5894	M00448	12616256	-	chr3	12616732	12616724	rs5746219	A/C	12616732	../snp/reverse/non-redundant/genotypes_chr3_YRI_r24_nr.b36.txt
T11F02180E2A	960	M00539	35130922	+	chr11	35130706	35130725	rs353620	C/T	35130712	../snp/reverse/non-redundant/genotypes_chr11_YRI_r24_nr.b36.txt
T11F02189F5F	960	M00280	35168095	+	chr11	35167372	35167388	rs996076	C/T	35167374	../snp/reverse/non-redundant/genotypes_chr11_YRI_r24_nr.b36.txt

REFERENCES

1. Stoner, G. D. and Rustgi, A. K. (1995) Biology of the esophageal squamous cell carcinoma, *Gastrointest. Cancers Biol. Diagn. 8*, 141-146.

2. Parkin, D. M., Pisani, P., and Ferlay, J. (1993) Estimates of the worldwide incidence of eighteen major cancers in 1985, *Int. J. Cancer 54*, 594-606.

3. Pisani, P., Parkin, D. M., and Ferlay, J. (1993) Estimates of the worldwide mortality from eighteen major cancers in 1985. Implications for prevention and projections of future burden, *Int. J. Cancer 55*, 891-903.

4. Reed, C. E. (1999) Surgical management of esophageal carcinoma, *Oncologist. 4*, 95-105.

5. Franciosi, J. P. and Liacouras, C. A. (2009) Eosinophilic esophagitis, *Immunol. Allergy Clin. North Am. 29*, 19-27, viii.

6. Cheng, K. K., Duffy, S. W., Day, N. E., and Lam, T. H. (1995) Oesophageal cancer in never-smokers and never-drinkers, *Int. J. Cancer 60*, 820-822.

7. Garidou, A., Tzonou, A., Lipworth, L., Signorello, L. B., Kalapothaki, V., and Trichopoulos, D. (1996) Life-style factors and medical conditions in relation to esophageal cancer by histologic type in a low-risk population, *Int. J. Cancer 68*, 295-299.

8. Chang-Claude, J. C., Wahrendorf, J., Liang, Q. S., Rei, Y. G., Munoz, N., Crespi, M., Raedsch, R., Thurnham, D. I., and Correa, P. (1990) An epidemiological study of precursor lesions of esophageal cancer among young persons in a high-risk population in Huixian, China, *Cancer Res. 50*, 2268-2274.

9. Hossein, S. M., Mirinezhad, K., Farhang, S., Jazayeri, E., Sani, A., Seif-Farshadi, M., Golzari, M., Kashef, S., and Sadegy, M. (2006) Gastrointestinal cancer occurrence in East Azarbaijan: a five year study from North Western Iran, *Asian Pac. J. Cancer Prev. 7*, 309-312.

10. Nimura, Y., Yokoyama, S., Fujimori, M., Aoki, T., Adachi, W., Nasu, T., He, M., Ping, Y. M., and Iida, F. (1997) Genotyping of the CYP1A1 and GSTM1 genes in esophageal carcinoma patients with special reference to smoking, *Cancer 80*, 852-857.

11. Morita, S., Yano, M., Shiozaki, H., Tsujinaka, T., Ebisui, C., Morimoto, T., Kishibuti, M., Fujita, J., Ogawa, A., Taniguchi, M., Inoue, M., Tamura, S., Yamazaki, K., Kikkawa, N., Mizunoya, S., and Monden, M. (1997) CYP1A1, CYP2E1 and GSTM1 polymorphisms are not associated with susceptibility to squamous-cell carcinoma of the esophagus, *Int. J. Cancer 71*, 192-195.

12. Hori, H., Kawano, T., Endo, M., and Yuasa, Y. (1997) Genetic polymorphisms of tobacco- and alcohol-related metabolizing enzymes and human esophageal squamous cell carcinoma susceptibility, *J. Clin. Gastroenterol. 25*, 568-575.

13. Lin, D. X., Tang, Y. M., Peng, Q., Lu, S. X., Ambrosone, C. B., and Kadlubar, F. F. (1998) Susceptibility to esophageal cancer and genetic polymorphisms in glutathione S-transferases T1, P1, and M1 and cytochrome P450 2E1, *Cancer Epidemiol. Biomarkers Prev. 7*, 1013-1018.

14. Guo, Y. M., Wang, Q., Liu, Y. Z., Chen, H. M., Qi, Z., and Guo, Q. H. (2008) Genetic polymorphisms in cytochrome P4502E1, alcohol and aldehyde dehydrogenases and the risk of esophageal squamous cell carcinoma in Gansu Chinese males, *World J. Gastroenterol. 14*, 1444-1449.

15. Lee, C. H., Lee, J. M., Wu, D. C., Goan, Y. G., Chou, S. H., Wu, I. C., Kao, E. L., Chan, T. F., Huang, M. C., Chen, P. S., Lee, C. Y., Huang, C. T., Huang, H. L., Hu, C. Y., Hung, Y. H., and Wu, M. T. (2008) Carcinogenetic impact of ADH1B and ALDH2 genes on squamous cell carcinoma risk of the esophagus with regard to the consumption of alcohol, tobacco and betel quid, *Int. J. Cancer 122*, 1347-1356.

16. Fidalgo, P. O., Cravo, M. L., Chaves, P. P., Leitao, C. N., and Mira, F. C. (1995) High prevalence of human papillomavirus in squamous cell carcinoma and matched normal esophageal mucosa: assessment by polymerase chain reaction, *Cancer 76*, 1522-1528.

17. Takahashi, A., Ogoshi, S., Ono, H., Ishikawa, T., Toki, T., Ohmori, N., Iwasa, M., Iwasa, Y., Furihata, M., and Ohtsuki, Y. (1998) High-risk human papillomavirus infection and overexpression of p53 protein in squamous cell carcinoma of the esophagus from Japan, *Dis. Esophagus. 11*, 162-167.

18. de Villiers, E. M., Lavergne, D., Chang, F., Syrjanen, K., Tosi, P., Cintorino, M., Santopietro, R., and Syrjanen, S. (1999) An interlaboratory study to determine the presence of human papillomavirus DNA in esophageal carcinoma from China, *Int. J. Cancer 81*, 225-228.

19. Agarwal, S. K., Chatterji, A., Bhambhani, S., and Sharma, B. K. (1998) Immunohistochemical co-expression of human papillomavirus type 16/18 transforming (E6) oncoprotein and p53 tumour suppressor gene proteins in oesophageal cancer, *Indian J. Exp. Biol. 36*, 559-563.

20. Khurshid, A., Kazuya, N., Hanae, I., and Manabu, I. (1998) Infection of human papillomavirus (HPV) and Epstein-Barr virus (EBV) and p53 expression in human esophageal carcinoma, *J. Pak. Med. Assoc. 48*, 138-142.

21. Lavergne, D. and de Villiers, E. M. (1999) Papillomavirus in esophageal papillomas and carcinomas, *Int. J. Cancer 80*, 681-684.

22. Miller, B. A., Davidson, M., Myerson, D., Icenogle, J., Lanier, A. P., Tan, J., and Beckmann, A. M. (1997) Human papillomavirus type 16 DNA in esophageal carcinomas from Alaska Natives, *Int. J. Cancer 71*, 218-222.

23. Kulski, J. K., Demeter, T., Mutavdzic, S., Sterrett, G. F., Mitchell, K. M., and Pixley, E. C. (1990) Survey of histologic specimens of human cancer for human papillomavirus types 6/11/16/18 by filter in situ hybridization, *Am. J. Clin. Pathol. 94*, 566-570.

24. Benamouzig, R., Jullian, E., Chang, F., Robaskiewicz, M., Flejou, J. F., Raoul, J. L., Coste, T., Couturier, D., Pompidou, A., and Rautureau, J. (1995) Absence of human papillomavirus DNA detected by polymerase chain reaction in French patients with esophageal carcinoma, *Gastroenterology 109*, 1876-1881.

25. Poljak, M., Cerar, A., and Seme, K. (1998) Human papillomavirus infection in esophageal carcinomas: a study of 121 lesions using multiple broad-spectrum polymerase chain reactions and literature review, *Hum. Pathol. 29*, 266-271.

26. Koh, J. S., Lee, S. S., Baek, H. J., and Kim, Y. I. (2008) No association of high-risk human papillomavirus with esophageal squamous cell carcinomas among Koreans, as determined by polymerase chain reaction, *Dis. Esophagus. 21*, 114-117.

27. Turner, J. R., Shen, L. H., Crum, C. P., Dean, P. J., and Odze, R. D. (1997) Low prevalence of human papillomavirus infection in esophageal squamous cell carcinomas from North America: analysis by a highly sensitive and specific polymerase chain reaction-based approach, *Hum. Pathol. 28*, 174-178.

28. Kok, T. C., Nooter, K., Tjong, A. H. S., Smits, H. L., and Ter Schegget, J. T. (1997) No evidence of known types of human papillomavirus in squamous cell cancer of the oesophagus in a low-risk area. Rotterdam Oesophageal Tumour Study Group, *Eur. J. Cancer 33*, 1865-1868.

29. Lam, A. K. (2000) Molecular biology of esophageal squamous cell carcinoma, *Crit Rev. Oncol. Hematol. 33*, 71-90.

30. Jenkins, T. D., Nakagawa, H., and Rustgi, A. K. (1996) The association of Epstein-Barr virus DNA with esophageal squamous cell carcinoma, *Oncogene 13*, 1809-1813.

31. Mori, M., Watanabe, M., Tanaka, S., Mimori, K., Kuwano, H., and Sugimachi, K. (1994) Epstein-Barr virus-associated carcinomas of the esophagus and stomach, *Arch. Pathol. Lab Med. 118*, 998-1001.

32. Lam, K. Y., Chan, A. C., Chan, K. W., Leung, M. L., and Srivastava, G. (1995) Absence of Epstein-Barr virus in penile carcinoma. A study of 42 cases using in situ hybridization, *Cancer 76*, 658-660.

33. Celeste, A. J., Iannazzi, J. A., Taylor, R. C., Hewick, R. M., Rosen, V., Wang, E. A., and Wozney, J. M. (1990) Identification of transforming growth factor beta family members present in bone-inductive protein purified from bovine bone, *Proc. Natl. Acad. Sci. U. S. A 87*, 9843-9847.

34. Rizvanov, A. A., Kiyasov, A. P., Gaziziov, I. M., Yilmaz, T. S., Kaligin, M. S., Andreeva, D. I., Shafigullina, A. K., Guseva, D. S., Kiselev, S. L., Matin, K., Palotas, A., and Islamov, R. R. (2008) Human umbilical cord blood cells transfected with VEGF and L(1)CAM do not differentiate into neurons but transform into vascular endothelial cells and secrete neuro-trophic factors to support neuro-genesis-a novel approach in stem cell therapy, *Neurochem. Int. 53*, 389-394.

35. Jaskiewicz, K. and De Groot, K. M. (1994) p53 gene mutants expression, cellular proliferation and differentiation in oesophageal carcinoma and non-cancerous epithelium, *Anticancer Res. 14*, 137-140.

36. Li, Z., Shimada, Y., Uchida, S., Maeda, M., Kawabe, A., Mori, A., Itami, A., Kano, M., Watanabe, G., and Imamura, M. (2000) TGF-alpha as well as VEGF, PD-ECGF and bFGF contribute to angiogenesis of esophageal squamous cell carcinoma, *Int. J. Oncol. 17*, 453-460.

37. Fukai, Y., Masuda, N., Kato, H., Fukuchi, M., Miyazaki, T., Nakajima, M., Sohda, M., Kuwano, H., and Nakajima, T. (2005) Correlation between laminin-5 gamma2 chain and epidermal growth factor receptor expression in esophageal squamous cell carcinomas, *Oncology 69*, 71-80.

38. Wong, F. H., Hu, C. P., Chiu, J. H., Huang, B. S., Chang, J. P., Lin, P. J., Chien, K. Y., and Chang, C. (1994) Expression of multiple oncogenes in human esophageal carcinomas, *Cancer Invest 12*, 121-131.

39. Yoshida, K., Kuniyasu, H., Yasui, W., Kitadai, Y., Toge, T., and Tahara, E. (1993) Expression of growth factors and their receptors in human esophageal carcinomas: regulation of expression by epidermal growth factor and transforming growth factor alpha, *J. Cancer Res. Clin. Oncol. 119*, 401-407.

40. Koliopanos, A., Friess, H., di Mola, F. F., Tang, W. H., Kubulus, D., Brigstock, D., Zimmermann, A., and Buchler, M. W. (2002) Connective tissue growth factor gene expression alters tumor progression in esophageal cancer, *World J. Surg. 26*, 420-427.

41. Bahnassy, A. A., Zekri, A. R., Abdallah, S., El-Shehaby, A. M., and Sherif, G. M. (2005) Human papillomavirus infection in Egyptian esophageal carcinoma: correlation with p53, p21, mdm2, C-erbB2 and impact on survival, *Pathol. Int. 55*, 53-62.

42. Eads, C. A., Lord, R. V., Wickramasinghe, K., Long, T. I., Kurumboor, S. K., Bernstein, L., Peters, J. H., DeMeester, S. R., DeMeester, T. R., Skinner, K. A., and Laird, P. W. (2001) Epigenetic patterns in the progression of esophageal adenocarcinoma, *Cancer Res. 61*, 3410-3418.

43. Lam, K. Y., Tin, L., and Ma, L. (1998) C-erbB-2 protein expression in oesophageal squamous epithelium from oesophageal squamous cell carcinomas, with special reference to histological grade of carcinoma and pre-invasive lesions, *Eur. J. Surg. Oncol. 24*, 431-435.

44. Drozdoff, V., Wall, N. A., and Pledger, W. J. (1994) Expression and growth inhibitory effect of decapentaplegic Vg-related protein 6: evidence for a regulatory role in keratinocyte differentiation, *Proc. Natl. Acad. Sci. U. S. A 91*, 5528-5532.

45. Blessing, M., Schirmacher, P., and Kaiser, S. (1996) Overexpression of bone morphogenetic protein-6 (BMP-6) in the epidermis of transgenic mice: inhibition or stimulation of proliferation depending on the pattern of transgene expression and formation of psoriatic lesions, *J. Cell Biol. 135*, 227-239.

46. Hickey, K., Grehan, D., Reid, I. M., O'Briain, S., Walsh, T. N., and Hennessy, T. P. (1994) Expression of epidermal growth factor receptor and proliferating cell nuclear antigen predicts response of esophageal squamous cell carcinoma to chemoradiotherapy, *Cancer 74*, 1693-1698.

47. Wong, B. R., Josien, R., Lee, S. Y., Vologodskaia, M., Steinman, R. M., and Choi, Y. (1998) The TRAF family of signal transducers mediates NF-kappaB activation by the TRANCE receptor, *J. Biol. Chem. 273*, 28355-28359.

48. Goan, Y. G., Hsu, H. K., Chang, H. C., Chou, Y. P., Chiang, K. H., and Cheng, J. T. (2005) Deregulated p21(WAF1) overexpression impacts survival of surgically resected esophageal squamous cell carcinoma patients, *Ann. Thorac. Surg. 80*, 1007-1016.

49. Wang, X. H., Li, S. S., Yan, A. H., Lu, C. X., and Guo, Y. P. (2006) [Constitutive activation of signal transducers and activators of transcription 3 and expression of its target gene products in human ESCC cell line], *Nan. Fang Yi. Ke. Da. Xue. Xue. Bao. 26*, 441-444.

50. Kimura, H., Kato, H., Faried, A., Sohda, M., Nakajima, M., Fukai, Y., Miyazaki, T., Masuda, N., Fukuchi, M., and Kuwano, H. (2007) Prognostic significance of EpCAM expression in human esophageal cancer, *Int. J. Oncol. 30*, 171-179.

51. Nakashima, K., Shimada, H., Ochiai, T., Kuboshima, M., Kuroiwa, N., Okazumi, S., Matsubara, H., Nomura, F., Takiguchi, M., and Hiwasa, T. (2004) Serological identification of TROP2 by recombinant cDNA expression cloning using sera of patients with esophageal squamous cell carcinoma, *Int. J. Cancer 112*, 1029-1035.

52. Hollstein, M. C., Peri, L., Mandard, A. M., Welsh, J. A., Montesano, R., Metcalf, R. A., Bak, M., and Harris, C. C. (1991) Genetic analysis of human esophageal tumors from two high incidence geographic areas: frequent p53 base substitutions and absence of ras mutations, *Cancer Res. 51*, 4102-4106.

53. Casson, A. G., Wilson, S. M., McCart, J. A., O'Malley, F. P., Ozcelik, H., Tsao, M. S., and Chambers, A. F. (1997) ras mutation and expression of the ras-regulated genes osteopontin and cathepsin L in human esophageal cancer, *Int. J. Cancer 72*, 739-745.

54. Trovato, M., Grosso, M., Vitarelli, E., Benvenga, S., Trimarchi, F., and Barresi, G. (2002) Immunoexpression of the hepatocyte growth factor (HGF), HGF-receptor (c-met) and STAT3 on placental tissues from malformed fetuses, *Histol. Histopathol. 17*, 691-698.

55. Poehlmann, T. G., Fitzgerald, J. S., Meissner, A., Wengenmayer, T., Schleussner, E., Friedrich, K., and Markert, U. R. (2005) Trophoblast invasion: tuning through LIF, signalling via Stat3, *Placenta 26 Suppl A*, S37-S41.

56. Colomiere, M., Ward, A. C., Riley, C., Trenerry, M. K., Cameron-Smith, D., Findlay, J., Ackland, L., and Ahmed, N. (2009) Cross talk of signals between EGFR and IL-6R through JAK2/STAT3 mediate epithelial-mesenchymal transition in ovarian carcinomas, *Br. J. Cancer 100*, 134-144.

57. Maetzel, D., Denzel, S., Mack, B., Canis, M., Went, P., Benk, M., Kieu, C., Papior, P., Baeuerle, P. A., Munz, M., and Gires, O. (2009) Nuclear signalling by tumour-associated antigen EpCAM, *Nat. Cell Biol. 11*, 162-171.

58. Latchman, D. S. (1997) Transcription factors: an overview, *Int. J. Biochem. Cell Biol. 29*, 1305-1312.

59. Lee, T. I. and Young, R. A. (2000) Transcription of eukaryotic protein-coding genes, *Annu. Rev. Genet. 34*, 77-137.

60. Mitchell, P. J. and Tjian, R. (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins, *Science 245*, 371-378.

61. Brivanlou, A. H. and Darnell, J. E., Jr. (2002) Signal transduction and the control of gene expression, *Science 295*, 813-818.

62. Chang, T. C., Zeitels, L. R., Hwang, H. W., Chivukula, R. R., Wentzel, E. A., Dews, M., Jung, J., Gao, P., Dang, C. V., Beer, M. A., Thomas-Tikhonenko, A., and Mendell, J. T. (2009) Lin-28B transactivation is necessary for Myc-mediated let-7 repression and proliferation, *Proc. Natl. Acad. Sci. U. S. A 106*, 3384-3389.

63. Bagui, T. K., Cui, D., Roy, S., Mohapatra, S., Shor, A. C., Ma, L., and Pledger, W. J. (2009) Inhibition of p27Kip1 gene transcription by mitogens, *Cell Cycle 8*, 115-124.

64. Liu, H., Zhou, M., Luo, X., Zhang, L., Niu, Z., Peng, C., Ma, J., Peng, S., Zhou, H., Xiang, B., Li, X., Li, S., He, J., Li, X., and Li, G. (2008) Transcriptional regulation of BRD7 expression by Sp1 and c-Myc, *BMC. Mol. Biol. 9*, 111.

65. Yan, S., Zhou, C., Lou, X., Xiao, Z., Zhu, H., Wang, Q., Wang, Y., Lu, N., He, S., Zhan, Q., Liu, S., and Xu, N. (2009) PTTG Overexpression Promotes Lymph Node Metastasis in Human Esophageal Squamous Cell Carcinoma, *Cancer Res.*

66. Miyazaki, S., Sasno, H., Shiga, K., Sawai, T., Nishihira, T., Okamoto, H., and Mori, S. (1992) Analysis of c-myc oncogene in human esophageal carcinoma: immunohistochemistry, in situ hybridization and northern and Southern blot studies, *Anticancer Res. 12*, 1747-1755.

67. Peinado, H., Portillo, F., and Cano, A. (2004) Transcriptional regulation of cadherins during development and carcinogenesis, *Int. J. Dev. Biol. 48*, 365-375.

68. Yang, J., Mani, S. A., Donaher, J. L., Ramaswamy, S., Itzykson, R. A., Come, C., Savagner, P., Gitelman, I., Richardson, A., and Weinberg, R. A. (2004) Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis, *Cell 117*, 927-939.

69. Medici, D., Hay, E. D., and Olsen, B. R. (2008) Snail and Slug promote epithelial-mesenchymal transition through beta-catenin-T-cell factor-4-dependent expression of transforming growth factor-beta3, *Mol. Biol. Cell 19*, 4875-4887.

70. Cisse, B., Caton, M. L., Lehner, M., Maeda, T., Scheu, S., Locksley, R., Holmberg, D., Zweier, C., den Hollander, N. S., Kant, S. G., Holter, W., Rauch, A., Zhuang, Y.,

and Reizis, B. (2008) Transcription factor E2-2 is an essential and specific regulator of plasmacytoid dendritic cell development, *Cell 135*, 37-48.

71. Yan, S., Zhou, C., Zhang, W., Zhang, G., Zhao, X., Yang, S., Wang, Y., Lu, N., Zhu, H., and Xu, N. (2008) beta-Catenin/TCF pathway upregulates STAT3 expression in human esophageal squamous cell carcinoma, *Cancer Lett. 271*, 85-97.

72. Knudson, A. G., Jr. (1971) Mutation and cancer: statistical study of retinoblastoma, *Proc. Natl. Acad. Sci. U. S. A 68*, 820-823.

73. Hirohashi, S. and Kanai, Y. (2003) Cell adhesion system and human cancer morphogenesis, *Cancer Sci. 94*, 575-581.

74. Yoshida, B. A., Sokoloff, M. M., Welch, D. R., and Rinker-Schaeffer, C. W. (2000) Metastasis-suppressor genes: a review and perspective on an emerging field, *J. Natl. Cancer Inst. 92*, 1717-1730.

75. Sherr, C. J. (2004) Principles of tumor suppression, *Cell 116*, 235-246.

76. Bektas, N., Donner, A., Wirtz, C., Heep, H., Gabbert, H. E., and Sarbia, M. (2000) Allelic loss involving the tumor suppressor genes APC and MCC and expression of the APC protein in the development of dysplasia and carcinoma in Barrett esophagus, *Am. J. Clin. Pathol. 114*, 890-895.

77. Zheng, J., Hou, P., and Zhou, C. (1996) [Loss of CDKN2 gene and amplification of cyclin D1 gene in human esophageal cancer], *Zhonghua Zhong. Liu Za Zhi. 18*, 408-411.

78. Michel, P., Magois, K., Robert, V., Chiron, A., Lepessot, F., Bodenant, C., Roque, I., Seng, S. K., Frebourg, T., and Paillot, B. (2002) Prognostic value of TP53 transcriptional activity on p21 and bax in patients with esophageal squamous cell carcinomas treated by definitive chemoradiotherapy, *Int. J. Radiat. Oncol. Biol. Phys. 54*, 379-385.

79. Shibata, H., Matsubara, O., Wakiyama, H., and Tanaka, S. (2001) The role of cyclin-dependent kinase inhibitor p27 in squamous cell carcinoma of the esophagus, *Pathol. Res. Pract. 197*, 157-164.

80. Fukuchi, M., Masuda, N., Miyazaki, T., Nakajima, M., Osawa, H., Kato, H., and Kuwano, H. (2002) Decreased Smad4 expression in the transforming growth factor-beta signaling pathway during progression of esophageal squamous cell carcinoma, *Cancer 95*, 737-743.

81. Tachibana, M., Shibakita, M., Ohno, S., Kinugasa, S., Yoshimura, H., Ueda, S., Fujii, T., Rahman, M. A., Dhar, D. K., and Nagasue, N. (2002) Expression and prognostic significance of PTEN product protein in patients with esophageal squamous cell carcinoma, *Cancer 94*, 1955-1960.

82. Nakajima, Y., Nagai, K., Miyake, S., Ohashi, K., Kawano, T., and Iwai, T. (2002) Evaluation of an indicator for lymph node metastasis of esophageal squamous cell carcinoma invading the submucosal layer, *Jpn. J. Cancer Res. 93*, 305-312.

83. Wijnhoven, B. P., Lindstedt, E. W., Abbou, M., Ijzendoorn, Y., de Krijger, R. R., Tilanus, H. W., and Dinjens, W. N. (2001) Molecular genetic analysis of the von Hippel-Lindau and human peroxisome proliferator-activated receptor gamma tumor-suppressor genes in adenocarcinomas of the gastroesophageal junction, *Int. J. Cancer 94*, 891-895.

84. Oji, Y., Yano, M., Nakano, Y., Abeno, S., Nakatsuka, S., Ikeba, A., Yasuda, T., Fujiwara, Y., Takiguchi, S., Yamamoto, H., Fujita, S., Kanato, K., Ito, K., Jomgeow, T., Kawakami, M., Tsuboi, A., Shirakata, T., Nishida, S., Hosen, N., Oka, Y., Aozasa, K., Monden, M., and Sugiyama, H. (2004) Overexpression of the Wilms' tumor gene WT1 in esophageal cancer, *Anticancer Res. 24*, 3103-3108.

85. Hulsken, J., Birchmeier, W., and Behrens, J. (1994) E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton, *J. Cell Biol. 127*, 2061-2069.

86. Caspi, M., Zilberberg, A., Eldar-Finkelman, H., and Rosin-Arbesfeld, R. (2008) Nuclear GSK-3beta inhibits the canonical Wnt signalling pathway in a beta-catenin phosphorylation-independent manner, *Oncogene 27*, 3546-3555.

87. Plawski, A., Lubinski, J., Banasiewicz, T., Paszkowski, J., Lipinski, D., Strembalska, A., Kurzawski, G., Byrski, T., Zajaczek, S., Hodorowicz-Zaniewska, D., Gach, T., Brozek, I., Nowakowska, D., Czkwaniec, E., Krokowicz, P., Drews, M., Zeyland, J., Juzwa, W., and Slomski, R. (2004) Novel germline mutations in the adenomatous polyposis coli gene in Polish families with familial adenomatous polyposis, *J. Med. Genet. 41*, e11.

88. Dundar, M., Caglayan, A. O., Saatci, C., Karaca, H., Baskol, M., Tahiri, S., and Ozkul, Y. (2007) How the I1307K adenomatous polyposis coli gene variant contributes in the assessment of risk of colorectal cancer, but not stomach cancer, in a Turkish population, *Cancer Genet. Cytogenet. 177*, 95-97.

89. Nair, K. S., Naidoo, R., and Chetty, R. (2006) Microsatellite analysis of the APC gene and immunoexpression of E-cadherin, catenin, and tubulin in esophageal squamous cell carcinoma, *Hum. Pathol. 37*, 125-134.

90. Tanaka, H., Shimada, Y., Imamura, M., Shibagaki, I., and Ishizaki, K. (1997) Multiple types of aberrations in the p16 (INK4a) and the p15(INK4b) genes in 30 esophageal squamous-cell-carcinoma cell lines, *Int. J. Cancer 70*, 437-442.

91. Hall, M., Bates, S., and Peters, G. (1995) Evidence for different modes of action of cyclin-dependent kinase inhibitors: p15 and p16 bind to kinases, p21 and p27 bind to cyclins, *Oncogene 11*, 1581-1588.

92. Fahraeus, R., Paramio, J. M., Ball, K. L., Lain, S., and Lane, D. P. (1996) Inhibition of pRb phosphorylation and cell-cycle progression by a 20-residue peptide derived from p16CDKN2/INK4A, *Curr. Biol. 6*, 84-91.

93. Rodway, H., Llanos, S., Rowe, J., and Peters, G. (2004) Stability of nucleolar versus non-nucleolar forms of human p14(ARF), *Oncogene 23*, 6186-6192.

94. Rokudai, S., Aikawa, Y., Tagata, Y., Tsuchida, N., Taya, Y., and Kitabayashi, I. (2009) Monocytic leukemia zinc finger (MOZ) interacts with p53 to induce p21 expression and cell-cycle arrest, *J. Biol. Chem. 284*, 237-244.

95. Bougeard, G., Sesboue, R., Baert-Desurmont, S., Vasseur, S., Martin, C., Tinat, J., Brugieres, L., Chompret, A., de Paillerets, B. B., Stoppa-Lyonnet, D., Bonaiti-Pellie, C., and Frebourg, T. (2008) Molecular basis of the Li-Fraumeni syndrome: an update from the French LFS families, *J. Med. Genet. 45*, 535-538.

96. Schlomm, T., Iwers, L., Kirstein, P., Jessen, B., Kollermann, J., Minner, S., Passow-Drolet, A., Mirlacher, M., Milde-Langosch, K., Graefen, M., Haese, A., Steuber, T., Simon, R., Huland, H., Sauter, G., and Erbersdobler, A. (2008) Clinical significance of p53 alterations in surgically treated prostate cancers, *Mod. Pathol. 21*, 1371-1378.

97. Dicker, F., Herholz, H., Schnittger, S., Nakao, A., Patten, N., Wu, L., Kern, W., Haferlach, T., and Haferlach, C. (2009) The detection of TP53 mutations in chronic lymphocytic leukemia independently predicts rapid disease progression and is highly correlated with a complex aberrant karyotype, *Leukemia 23*, 117-124.

98. Brankovic-Magic, M. V., Jankovic, R. N., Dobricic, J. D., Borojevic, N. D., Magic, Z. M., and Radulovic, S. S. (2008) TP53 mutations in breast cancer: association with ductal histology and early relapse of disease, *Int. J. Biol. Markers 23*, 147-153.

99. Wang, L. S., Chow, K. C., Liu, C. C., and Chiu, J. H. (1998) p53 gene alternation in squamous cell carcinoma of the esophagus detected by PCR-cold SSCP analysis, *Proc. Natl. Sci. Counc. Repub. China B 22*, 114-121.

100. Montesano, R., Hollstein, M., and Hainaut, P. (1996) Genetic alterations in esophageal cancer and their relevance to etiology and pathogenesis: a review, *Int. J. Cancer 69*, 225-235.

101. Wagata, T., Shibagaki, I., Imamura, M., Shimada, Y., Toguchida, J., Yandell, D. W., Ikenaga, M., Tobe, T., and Ishizaki, K. (1993) Loss of 17p, mutation of the p53 gene, and overexpression of p53 protein in esophageal squamous cell carcinomas, *Cancer Res. 53*, 846-850.

102. Casson, A. G., Tammemagi, M., Eskandarian, S., Redston, M., McLaughlin, J., and Ozcelik, H. (1998) p53 alterations in oesophageal cancer: association with clinicopathological features, risk factors, and survival, *Mol. Pathol. 51*, 71-79.

103. Lam, K. Y., Tsao, S. W., Zhang, D., Law, S., He, D., Ma, L., and Wong, J. (1997) Prevalence and predictive value of p53 mutation in patients with oesophageal squamous cell carcinomas: a prospective clinico-pathological study and survival analysis of 70 patients, *Int. J. Cancer 74*, 212-219.

104. De, V. L., Joubes, J., and Inze, D. (2003) Plant cell cycle transitions, *Curr. Opin. Plant Biol. 6*, 536-543.

105. Wu, C. L., Zukerberg, L. R., Ngwu, C., Harlow, E., and Lees, J. A. (1995) In vivo association of E2F and DP family proteins, *Mol. Cell Biol. 15*, 2536-2546.

106. Funk, J. O., Waga, S., Harry, J. B., Espling, E., Stillman, B., and Galloway, D. A. (1997) Inhibition of CDK activity and PCNA-dependent DNA replication by p21 is blocked by interaction with the HPV-16 E7 oncoprotein, *Genes Dev. 11*, 2090-2100.

107. Busatto, G., Shiao, Y. H., Parenti, A. R., Baffa, R., Ruol, A., Plebani, M., and Rugge, M. (1998) p16/CDKN2 alterations and pRb expression in oesophageal squamous carcinoma, *Mol. Pathol. 51*, 80-84.

108. Yang, G., Zhang, Z., Liao, J., Seril, D., Wang, L., Goldstein, S., and Yang, C. S. (1997) Immunohistochemical studies on Waf1p21, p16, pRb and p53 in human esophageal carcinomas and neighboring epithelia from a high-risk area in northern China, *Int. J. Cancer 72*, 746-751.

109. Roncalli, M., Bosari, S., Marchetti, A., Buttitta, F., Bossi, P., Graziani, D., Peracchia, A., Bonavina, L., Viale, G., and Coggi, G. (1998) Cell cycle-related gene abnormalities and product expression in esophageal carcinoma, *Lab Invest 78*, 1049-1057.

110. Chetty, R. and Chetty, S. (1997) Cyclin D1 and retinoblastoma protein expression in oesophageal squamous cell carcinoma, *Mol. Pathol. 50*, 257-260.

111. Hofer, M. D., Menke, A., Genze, F., Gierschik, P., and Giehl, K. (2004) Expression of MTA1 promotes motility and invasiveness of PANC-1 pancreatic carcinoma cells, *Br. J. Cancer 90*, 455-462.

112. Dannenmann, C., Shabani, N., Friese, K., Jeschke, U., Mylonas, I., and Bruning, A. (2008) The metastasis-associated gene MTA1 is upregulated in advanced ovarian cancer, represses ERbeta, and enhances expression of oncogenic cytokine GRO, *Cancer Biol. Ther. 7*, 1460-1467.

113. Martin, M. D., Hilsenbeck, S. G., Mohsin, S. K., Hopp, T. A., Clark, G. M., Osborne, C. K., Allred, D. C., and O'Connell, P. (2006) Breast tumors that overexpress nuclear metastasis-associated 1 (MTA1) protein have high recurrence risks but enhanced responses to systemic therapies, *Breast Cancer Res. Treat. 95*, 7-12.

114. Hofer, M. D., Kuefer, R., Varambally, S., Li, H., Ma, J., Shapiro, G. I., Gschwend, J. E., Hautmann, R. E., Sanda, M. G., Giehl, K., Menke, A., Chinnaiyan, A. M., and Rubin, M. A. (2004) The role of metastasis-associated protein 1 in prostate cancer progression, *Cancer Res. 64*, 825-829.

115. Qian, H., Lu, N., Xue, L., Liang, X., Zhang, X., Fu, M., Xie, Y., Zhan, Q., Liu, Z., and Lin, C. (2005) Reduced MTA1 expression by RNAi inhibits in vitro invasion and migration of esophageal squamous cell carcinoma cell line, *Clin. Exp. Metastasis 22*, 653-662.

116. Boissan, M., Poupon, M. F., and Lacombe, M. L. (2007) [NM23 and metastasis suppressor genes: update], *Med. Sci. (Paris) 23*, 1115-1123.

117. Georgescauld, F., Mocan, I., Lacombe, M. L., and Lascu, I. (2009) Rescue of the neuroblastoma mutant of the human nucleoside diphosphate kinase A/nm23-H1 by the natural osmolyte trimethylamine-N-oxide, *FEBS Lett. 583*, 820-824.

118. Patel, D. D., Bhatavdekar, J. M., Chikhlikar, P. R., Patel, Y. V., Shah, N. G., Ghosh, N., Suthar, T. P., and Balar, D. B. (1997) Clinical significance of p53, nm23, and bcl-2 in T3-4N1M0 oesophageal carcinoma: an immunohistochemical approach, *J. Surg. Oncol. 65*, 111-116.

119. Kerr, J. F., Wyllie, A. H., and Currie, A. R. (1972) Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics, *Br. J. Cancer 26*, 239-257.

120. Weil, M., Jacobson, M. D., Coles, H. S., Davies, T. J., Gardner, R. L., Raff, K. D., and Raff, M. C. (1996) Constitutive expression of the machinery for programmed cell death, *J. Cell Biol. 133*, 1053-1059.

121. Ishizaki, Y., Cheng, L., Mudge, A. W., and Raff, M. C. (1995) Programmed cell death by default in embryonic cells, fibroblasts, and cancer cells, *Mol. Biol. Cell 6*, 1443-1458.

122. Chao, D. T. and Korsmeyer, S. J. (1998) BCL-2 family: regulators of cell death, *Annu. Rev. Immunol. 16*, 395-419.

123. Alnemri, E. S., Livingston, D. J., Nicholson, D. W., Salvesen, G., Thornberry, N. A., Wong, W. W., and Yuan, J. (1996) Human ICE/CED-3 protease nomenclature, *Cell 87*, 171.

124. Takahashi, R., Deveraux, Q., Tamm, I., Welsh, K., ssa-Munt, N., Salvesen, G. S., and Reed, J. C. (1998) A single BIR domain of XIAP sufficient for inhibiting caspases, *J. Biol. Chem. 273*, 7787-7790.

125. Salvesen, G. S. and Duckett, C. S. (2002) IAP proteins: blocking the road to death's door, *Nat. Rev. Mol. Cell Biol. 3*, 401-410.

126. Haslbeck, M., Ignatiou, A., Saibil, H., Helmich, S., Frenzl, E., Stromer, T., and Buchner, J. (2004) A domain in the N-terminal part of Hsp26 is essential for chaperone function and oligomerization, *J. Mol. Biol. 343*, 445-455.

127. Nollen, E. A. and Morimoto, R. I. (2002) Chaperoning signaling pathways: molecular chaperones as stress-sensing 'heat shock' proteins, *J. Cell Sci. 115*, 2809-2816.

128. Sarbia, M., Bittinger, F., Porschen, R., Verreet, P., Dutkowski, P., Willers, R., and Gabbert, H. E. (1996) bcl-2 expression and prognosis in squamous-cell carcinomas of the esophagus, *Int. J. Cancer 69*, 324-328.

129. Takayama, T., Nagao, M., Sawada, H., Yamada, Y., Emoto, K., Fujimoto, H., Ueno, M., Hirao, S., and Nakajima, Y. (2001) Bcl-X expression in esophageal squamous cell carcinoma: association with tumor progression and prognosis, *J. Surg. Oncol. 78*, 116-123.

130. Ikeguchi, M. and Kaibara, N. (2002) survivin messenger RNA expression is a good prognostic biomarker for oesophageal carcinoma, *Br. J. Cancer 87*, 883-887.

131. Noguchi, T., Takeno, S., Shibata, T., Uchida, Y., Yokoyama, S., and Muller, W. (2002) Expression of heat shock protein 70 in grossly resected esophageal squamous cell carcinoma, *Ann. Thorac. Surg. 74*, 222-226.

132. Nakajima, M., Kuwano, H., Miyazaki, T., Masuda, N., and Kato, H. (2002) Significant correlation between expression of heat shock proteins 27, 70 and lymphocyte infiltration in esophageal squamous cell carcinoma, *Cancer Lett. 178*, 99-106.

133. Ranger, A. M., Malynn, B. A., and Korsmeyer, S. J. (2001) Mouse models of cell death, *Nat. Genet. 28*, 113-118.

134. Vogelstein, B., Lane, D., and Levine, A. J. (2000) Surfing the p53 network, *Nature 408*, 307-310.

135. Cheng, E. H., Kirsch, D. G., Clem, R. J., Ravi, R., Kastan, M. B., Bedi, A., Ueno, K., and Hardwick, J. M. (1997) Conversion of Bcl-2 to a Bax-like death effector by caspases, *Science 278*, 1966-1968.

136. Belizario, J. E., Alves, J., Occhiucci, J. M., Garay-Malpartida, M., and Sesso, A. (2007) A mechanistic view of mitochondrial death decision pores, *Braz. J. Med. Biol. Res. 40*, 1011-1024.

137. Chang, H. Y., Nishitoh, H., Yang, X., Ichijo, H., and Baltimore, D. (1998) Activation of apoptosis signal-regulating kinase 1 (ASK1) by the adapter protein Daxx, *Science 281*, 1860-1863.

138. Yang, X., Khosravi-Far, R., Chang, H. Y., and Baltimore, D. (1997) Daxx, a novel Fas-binding protein that activates JNK and apoptosis, *Cell 89*, 1067-1076.

139. Charette, S. J. and Landry, J. (2000) The interaction of HSP27 with Daxx identifies a potential regulatory role of HSP27 in Fas-induced apoptosis, *Ann. N. Y. Acad. Sci. 926*, 126-131.

140. Pandey, P., Farber, R., Nakazawa, A., Kumar, S., Bharti, A., Nalin, C., Weichselbaum, R., Kufe, D., and Kharbanda, S. (2000) Hsp27 functions as a negative regulator of cytochrome c-dependent activation of procaspase-3, *Oncogene 19*, 1975-1981.

141. Komarova, E. Y., Afanasyeva, E. A., Bulatova, M. M., Cheetham, M. E., Margulis, B. A., and Guzhova, I. V. (2004) Downstream caspases are novel targets for the antiapoptotic activity of the molecular chaperone hsp70, *Cell Stress. Chaperones. 9*, 265-275.

142. Kamradt, M. C., Chen, F., and Cryns, V. L. (2001) The small heat shock protein alpha B-crystallin negatively regulates cytochrome c- and caspase-8-dependent activation of caspase-3 by inhibiting its autoproteolytic maturation, *J. Biol. Chem. 276*, 16059-16063.

143. Paul, C., Manero, F., Gonin, S., Kretz-Remy, C., Virot, S., and Arrigo, A. P. (2002) Hsp27 as a negative regulator of cytochrome C release, *Mol. Cell Biol. 22*, 816-834.

144. Gabai, V. L. and Sherman, M. Y. (2002) Invited review: Interplay between molecular chaperones and signaling pathways in survival of heat shock, *J. Appl. Physiol 92*, 1743-1748.

145. Gotoh, T., Terada, K., Oyadomari, S., and Mori, M. (2004) hsp70-DnaJ chaperone pair prevents nitric oxide- and CHOP-induced apoptosis by inhibiting translocation of Bax to mitochondria, *Cell Death. Differ. 11*, 390-402.

146. Adams, J. M. and Cory, S. (2002) Apoptosomes: engines for caspase activation, *Curr. Opin. Cell Biol. 14*, 715-720.

147. Acehan, D., Jiang, X., Morgan, D. G., Heuser, J. E., Wang, X., and Akey, C. W. (2002) Three-dimensional structure of the apoptosome: implications for assembly, procaspase-9 binding, and activation, *Mol. Cell 9*, 423-432.

148. Saleh, A., Srinivasula, S. M., Balkir, L., Robbins, P. D., and Alnemri, E. S. (2000) Negative regulation of the Apaf-1 apoptosome by Hsp70, *Nat. Cell Biol. 2*, 476-483.

149. Beere, H. M., Wolf, B. B., Cain, K., Mosser, D. D., Mahboubi, A., Kuwana, T., Tailor, P., Morimoto, R. I., Cohen, G. M., and Green, D. R. (2000) Heat-shock protein 70 inhibits apoptosis by preventing recruitment of procaspase-9 to the Apaf-1 apoptosome, *Nat. Cell Biol. 2*, 469-475.

150. Pandey, P., Saleh, A., Nakazawa, A., Kumar, S., Srinivasula, S. M., Kumar, V., Weichselbaum, R., Nalin, C., Alnemri, E. S., Kufe, D., and Kharbanda, S. (2000) Negative regulation of cytochrome c-mediated oligomerization of Apaf-1 and activation of procaspase-9 by heat shock protein 90, *EMBO J. 19*, 4310-4322.

151. Bruey, J. M., Ducasse, C., Bonniaud, P., Ravagnan, L., Susin, S. A., az-Latoud, C., Gurbuxani, S., Arrigo, A. P., Kroemer, G., Solary, E., and Garrido, C. (2000) Hsp27 negatively regulates cell death by interacting with cytochrome c, *Nat. Cell Biol. 2*, 645-652.

152. Takahashi, H., Shikata, N., Senzaki, H., Shintaku, M., and Tsubura, A. (1995) Immunohistochemical staining patterns of keratins in normal oesophageal epithelium and carcinoma of the oesophagus, *Histopathology 26*, 45-50.

153. Lam, K. Y., Loke, S. L., Shen, X. C., and Ma, L. T. (1995) Cytokeratin expression in non-neoplastic oesophageal epithelium and squamous cell carcinoma of the oesophagus, *Virchows Arch. 426*, 345-349.

154. Yamamoto, K., Oka, M., Hayashi, H., Tangoku, A., Gondo, T., and Suzuki, T. (1997) CYFRA 21-1 is a useful marker for esophageal squamous cell carcinoma, *Cancer 79*, 1647-1655.

155. Hocquette, J. F. (2005) Where are we in genomics?, *J. Physiol Pharmacol. 56 Suppl 3*, 37-70.

156. Kato, K., Yamashita, R., Matoba, R., Monden, M., Noguchi, S., Takagi, T., and Nakai, K. (2005) Cancer gene expression database (CGED): a database for gene expression profiling with accompanying clinical information of human cancer tissues, *Nucleic Acids Res. 33*, D533-D536.

157. Thiemann, K. M., Frost, M. H., and Thompson, R. A. (1999) A multifaceted educational approach to increasing awareness and use of physician data query (PDQ), *J. Cancer Educ. 14*, 78-82.

158. Rhodes, D. R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., Barrette, T., Pandey, A., and Chinnaiyan, A. M. (2004) ONCOMINE: a cancer microarray database and integrated data-mining platform, *Neoplasia. 6*, 1-6.

159. Eyre, T. A., Ducluzeau, F., Sneddon, T. P., Povey, S., Bruford, E. A., and Lush, M. J. (2006) The HUGO Gene Nomenclature Database, 2006 updates, *Nucleic Acids Res. 34*, D319-D321.

160. Safran, M., Solomon, I., Shmueli, O., Lapidot, M., Shen-Orr, S., Adato, A., Ben-Dor, U., Esterman, N., Rosen, N., Peter, I., Olender, T., Chalifa-Caspi, V., and Lancet, D. (2002) GeneCards 2002: towards a complete, object-oriented, human gene compendium, *Bioinformatics. 18*, 1542-1543.

161. Kulikova, T., Akhtar, R., Aldebert, P., Althorpe, N., Andersson, M., Baldwin, A., Bates, K., Bhattacharyya, S., Bower, L., Browne, P., Castro, M., Cochrane, G., Duggan, K., Eberhardt, R., Faruque, N., Hoad, G., Kanz, C., Lee, C., Leinonen, R., Lin, Q., Lombard, V., Lopez, R., Lorenc, D., McWilliam, H., Mukherjee, G., Nardone, F., Pastor, M. P., Plaister, S., Sobhany, S., Stoehr, P., Vaughan, R., Wu, D., Zhu, W., and Apweiler, R. (2007) EMBL Nucleotide Sequence Database in 2006, *Nucleic Acids Res. 35*, D16-D20.

162. Flicek, P., Aken, B. L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T., Down, T., Dyer, S. C., Eyre, T., Fitzgerald, S., Fernandez-Banet, J., Graf, S., Haider, S., Hammond, M., Holland, R., Howe, K. L., Howe, K., Johnson, N., Jenkinson, A., Kahari, A., Keefe, D., Kokocinski, F., Kulesha, E., Lawson, D., Longden, I., Megy, K., Meidl, P., Overduin, B., Parker, A., Pritchard, B., Prlic, A., Rice, S., Rios, D., Schuster, M., Sealy, I., Slater, G., Smedley, D., Spudich, G., Trevanion, S., Vilella, A. J., Vogel, J., White, S., Wood, M., Birney, E., Cox, T., Curwen, V., Durbin, R., Fernandez-Suarez, X. M., Herrero, J., Hubbard, T. J., Kasprzyk, A., Proctor, G., Smith, J., Ureta-Vidal, A., and Searle, S. (2008) Ensembl 2008, *Nucleic Acids Res. 36*, D707-D714.

163. Pruitt, K. D., Tatusova, T., and Maglott, D. R. (2007) NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins, *Nucleic Acids Res. 35*, D61-D65.

164. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Sayers, E. W. (2009) GenBank, *Nucleic Acids Res. 37*, D26-D31.

165. Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., Dicuccio, M., Edgar, R., Federhen, S., Feolo, M., Geer, L. Y.,

Helmberg, W., Kapustin, Y., Khovayko, O., Landsman, D., Lipman, D. J., Madden, T. L., Maglott, D. R., Miller, V., Ostell, J., Pruitt, K. D., Schuler, G. D., Shumway, M., Sequeira, E., Sherry, S. T., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusov, R. L., Tatusova, T. A., Wagner, L., and Yaschenko, E. (2008) Database resources of the National Center for Biotechnology Information, *Nucleic Acids Res. 36*, D13-D21.

166. UniProt (2008) The universal protein resource (UniProt), *Nucleic Acids Res. 36*, D190-D195.

167. Gasteiger, E., Jung, E., and Bairoch, A. (2001) SWISS-PROT: connecting biomolecular knowledge via a protein database, *Curr. Issues Mol. Biol. 3*, 47-55.

168. Berman, H., Henrick, K., Nakamura, H., and Markley, J. L. (2007) The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data, *Nucleic Acids Res. 35*, D301-D303.

169. Alibes, A., Yankilevich, P., Canada, A., and az-Uriarte, R. (2007) IDconverter and IDClight: conversion and annotation of gene and protein IDs, *BMC. Bioinformatics. 8*, 9.

170. Khatri, P., Desai, V., Tarca, A. L., Sellamuthu, S., Wildman, D. E., Romero, R., and Draghici, S. (2006) New Onto-Tools: Promoter-Express, nsSNPCounter and Onto-Translate, *Nucleic Acids Res. 34*, W626-W631.

171. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium, *Nat. Genet. 25*, 25-29.

172. Kelso, J., Visagie, J., Theiler, G., Christoffels, A., Bardien, S., Smedley, D., Otgaar, D., Greyling, G., Jongeneel, C. V., McCarthy, M. I., Hide, T., and Hide, W. (2003) eVOC: a controlled vocabulary for unifying gene expression data, *Genome Res. 13*, 1222-1230.

173. Vastrik, I., D'Eustachio, P., Schmidt, E., Gopinath, G., Croft, D., de, B. B., Gillespie, M., Jassal, B., Lewis, S., Matthews, L., Wu, G., Birney, E., and Stein, L. (2007) Reactome: a knowledge base of biologic pathways and processes, *Genome Biol. 8*, R39.

174. Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C. A., Taylor, M. S., Engstrom, P. G., Frith, M. C., Forrest, A. R., Alkema, W. B., Tan, S. L., Plessy, C., Kodzius, R., Ravasi, T., Kasukawa, T., Fukuda, S., Kanamori-Katayama, M., Kitazume, Y., Kawaji, H., Kai, C., Nakamura, M., Konno, H., Nakano, K., Mottagui-Tabar, S., Arner, P., Chesi, A., Gustincich, S., Persichetti, F., Suzuki, H., Grimmond, S. M., Wells, C. A., Orlando, V., Wahlestedt, C., Liu, E. T., Harbers, M., Kawai, J., Bajic, V. B., Hume, D. A., and Hayashizaki, Y. (2006) Genome-wide analysis of mammalian promoter architecture and evolution, *Nat. Genet. 38*, 626-635.

175. Aerts, S., Van, L. P., Thijs, G., Mayer, H., de, M. R., Moreau, Y., and De, M. B. (2005) TOUCAN 2: the all-inclusive open source workbench for regulatory sequence analysis, *Nucleic Acids Res. 33*, W393-W396.

176. Matys, V., Kel-Margoulis, O. V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K., Voss, N., Stegmaier, P., Lewicki-Potapov, B., Saxel, H., Kel, A. E., and Wingender, E. (2006) TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes, *Nucleic Acids Res. 34*, D108-D110.

177. Kel, A. E., Gossling, E., Reuter, I., Cheremushkin, E., Kel-Margoulis, O. V., and Wingender, E. (2003) MATCH: A tool for searching transcription factor binding sites in DNA sequences, *Nucleic Acids Res. 31*, 3576-3579.

178. Sagar, S., Kaur, M., Dawe, A., Seshadri, S. V., Christoffels, A., Schaefer, U., Radovanovic, A., and Bajic, V. B. (2008) DDESC: Dragon database for exploration of sodium channels in human, *BMC. Genomics 9*, 622.

179. Pan, H., Zuo, L., Choudhary, V., Zhang, Z., Leow, S. H., Chong, F. T., Huang, Y., Ong, V. W., Mohanty, B., Tan, S. L., Krishnan, S. P., and Bajic, V. B. (2004) Dragon TF Association Miner: a system for exploring transcription factor associations through text-mining, *Nucleic Acids Res. 32*, W230-W234.

180. Bajic, V. B., Veronika, M., Veladandi, P. S., Meka, A., Heng, M. W., Rajaraman, K., Pan, H., and Swarup, S. (2005) Dragon Plant Biology Explorer. A text-mining tool for integrating associations between genetic and biochemical entities with genome annotation and biochemical terms lists, *Plant Physiol 138*, 1914-1925.

181. Kaur, M., Radovanovic, A., Essack, M., Schaefer, U., Maqungo, M., Kibler, T., Schmeier, S., Christoffels, A., Narasimhan, K., Choolani, M., and Bajic, V. B. (2009) Database for exploration of functional context of genes implicated in ovarian cancer, *Nucleic Acids Res. 37*, D820-D823.

182. Smyth, G. K., Yang, Y. H., and Speed, T. (2003) Statistical issues in cDNA microarray data analysis, *Methods Mol. Biol. 224*, 111-136.

183. Pritchard, C. C., Hsu, L., Delrow, J., and Nelson, P. S. (2001) Project normal: defining normal variance in mouse gene expression, *Proc. Natl. Acad. Sci. U. S. A 98*, 13266-13271.

184. Goto, S., Bono, H., Ogata, H., Fujibuchi, W., Nishioka, T., Sato, K., and Kanehisa, M. (1997) Organizing and computing metabolic pathway data in terms of binary relations, *Pac. Symp. Biocomput.* 175-186.

185. Huang, d. W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, *Nat. Protoc. 4*, 44-57.

186. Essack, M., Radovanovic, A., Schaefer, U., Schmeier, S., Seshadri, S. V., Christoffels, A., Kaur, M., and Bajic, V. B. (2009) DDEC: Dragon database of genes implicated in esophageal cancer, *BMC. Cancer 9*, 219.

187. Choi, H. J., Chung, Y. S., Kim, H. J., Moon, U. Y., Choi, Y. H., Van, S., I, Baek, S. J., Yoon, H. G., and Yoon, J. H. (2009) Signal pathway of 17beta-estradiol-induced MUC5B expression in human airway epithelial cells, *Am. J. Respir. Cell Mol. Biol. 40*, 168-178.

188. Grandas, O. H., Mountain, D. H., Kirkpatrick, S. S., Cassada, D. C., Stevens, S. L., Freeman, M. B., and Goldman, M. H. (2009) Regulation of vascular smooth muscle cell expression and function of matrix metalloproteinases is mediated by estrogen and progesterone exposure, *J. Vasc. Surg. 49*, 185-191.

189. Tan, K., Tegner, J., and Ravasi, T. (2008) Integrated approaches to uncovering transcription regulatory networks in mammalian cells, *Genomics 91*, 219-231.

190. Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R., and Young, R. A. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells, *Cell 122*, 947-956.

191. Rada-Iglesias, A., Wallerman, O., Koch, C., Ameur, A., Enroth, S., Clelland, G., Wester, K., Wilcox, S., Dovey, O. M., Ellis, P. D., Wraight, V. L., James, K., Andrews, R., Langford, C., Dhami, P., Carter, N., Vetrie, D., Ponten, F., Komorowski, J., Dunham, I., and Wadelius, C. (2005) Binding sites for metabolic disease related transcription factors inferred at base pair resolution by chromatin immunoprecipitation and genomic microarrays, *Hum. Mol. Genet. 14*, 3435-3447.

192. Tronche, F., Ringeisen, F., Blumenfeld, M., Yaniv, M., and Pontoglio, M. (1997) Analysis of the distribution of binding sites for a tissue-specific transcription factor in the vertebrate genome, *J. Mol. Biol. 266*, 231-245.

193. Zou, M. X., Butcher, D. T., Sadikovic, B., Groves, T. C., Yee, S. P., and Rodenhiser, D. I. (2004) Characterization of functional elements in the neurofibromatosis (NF1) proximal promoter region, *Oncogene 23*, 330-339.

194. Rim, J. S. and Kozak, L. P. (2002) Regulatory motifs for CREB-binding protein and Nfe2l2 transcription factors in the upstream enhancer of the mitochondrial uncoupling protein 1 gene, *J. Biol. Chem. 277*, 34589-34600.

195. Discenza, M. T., Dehbi, M., and Pelletier, J. (1997) Overlapping DNA recognition motifs between Sp1 and a novel trans-acting factor within the wt1 tumour suppressor gene promoter, *Nucleic Acids Res. 25*, 4314-4322.

196. Eisermann, K., Tandon, S., Bazarov, A., Brett, A., Fraizer, G., and Piontkivska, H. (2008) Evolutionary conservation of zinc finger transcription factor binding sites in promoters of genes co-expressed with WT1 in prostate cancer, *BMC. Genomics 9*, 337.

197. Ueo, H., Matsuoka, H., Sugimachi, K., Kuwano, H., Mori, M., and Akiyoshi, T. (1990) Inhibitory effects of estrogen on the growth of a human esophageal carcinoma cell line, *Cancer Res. 50*, 7212-7215.

198. Nozoe, T., Oyama, T., Takenoyama, M., Hanagiri, T., Sugio, K., and Yasumoto, K. (2007) Significance of immunohistochemical expression of estrogen receptors alpha and beta in squamous cell carcinoma of the esophagus, *Clin. Cancer Res. 13*, 4046-4050.

199. Utsumi, Y., Nakamura, T., Nagasue, N., Kubota, H., and Morikawa, S. (1989) Role of estrogen receptors in the growth of human esophageal carcinoma, *Cancer 64*, 88-93.

200. Klinge, C. M. (2001) Estrogen receptor interaction with estrogen response elements, *Nucleic Acids Res. 29*, 2905-2919.

201. Pettersson, K., Grandien, K., Kuiper, G. G., and Gustafsson, J. A. (1997) Mouse estrogen receptor beta forms estrogen response element-binding heterodimers with estrogen receptor alpha, *Mol. Endocrinol. 11*, 1486-1496.

202. Woo, I. S., Park, M. J., Choi, S. W., Kim, S. J., Lee, M. A., Kang, J. H., Hong, Y. S., and Lee, K. S. (2004) Loss of estrogen receptor-alpha expression is associated with hypermethylation near its ATG start codon in gastric cancer cell lines, *Oncol. Rep. 11*, 617-622.

203. Peng, B., Lu, B., Leygue, E., and Murphy, L. C. (2003) Putative functional characteristics of human estrogen receptor-beta isoforms, *J. Mol. Endocrinol. 30*, 13-29.

204. Hyder, S. M., Chiappetta, C., and Stancel, G. M. (1999) Interaction of human estrogen receptors alpha and beta with the same naturally occurring estrogen response elements, *Biochem. Pharmacol. 57*, 597-601.

205. Liu, M. M., Albanese, C., Anderson, C. M., Hilty, K., Webb, P., Uht, R. M., Price, R. H., Jr., Pestell, R. G., and Kushner, P. J. (2002) Opposing action of estrogen receptors alpha and beta on cyclin D1 gene expression, *J. Biol. Chem. 277*, 24353-24360.

206. Hayashi, S. I., Eguchi, H., Tanimoto, K., Yoshida, T., Omoto, Y., Inoue, A., Yoshida, N., and Yamaguchi, Y. (2003) The expression and function of estrogen receptor alpha and beta in human breast cancer and its clinical application, *Endocr. Relat Cancer 10*, 193-202.

207. Gruber, C. J., Tschugguel, W., Schneeberger, C., and Huber, J. C. (2002) Production and actions of estrogens, *N. Engl. J. Med. 346*, 340-352.

208. Fliss, A. E., Benzeno, S., Rao, J., and Caplan, A. J. (2000) Control of estrogen receptor ligand binding by Hsp90, *J. Steroid Biochem. Mol. Biol. 72*, 223-230.

209. Le, G. P., Montano, M. M., Schodin, D. J., and Katzenellenbogen, B. S. (1994) Phosphorylation of the human estrogen receptor. Identification of hormone-regulated sites and examination of their influence on transcriptional activity, *J. Biol. Chem. 269*, 4458-4466.

210. Ignar-Trowbridge, D. M., Pimentel, M., Parker, M. G., McLachlan, J. A., and Korach, K. S. (1996) Peptide growth factor cross-talk with the estrogen receptor requires the A/B domain and occurs independently of protein kinase C or estradiol, *Endocrinology 137*, 1735-1744.

211. Chalbos, D., Philips, A., and Rochefort, H. (1994) Genomic cross-talk between the estrogen receptor and growth factor regulatory pathways in estrogen target tissues, *Semin. Cancer Biol. 5*, 361-368.

212. Bunone, G., Briand, P. A., Miksicek, R. J., and Picard, D. (1996) Activation of the unliganded estrogen receptor by EGF involves the MAP kinase pathway and direct phosphorylation, *EMBO J. 15*, 2174-2183.

213. Diel, P. (2002) Tissue-specific estrogenic response and molecular mechanisms, *Toxicol. Lett. 127*, 217-224.

214. McKenna, N. J. and O'Malley, B. W. (2002) Combinatorial control of gene expression by nuclear receptors and coregulators, *Cell 108*, 465-474.

215. Altucci, L., Addeo, R., Cicatiello, L., Dauvois, S., Parker, M. G., Truss, M., Beato, M., Sica, V., Bresciani, F., and Weisz, A. (1996) 17beta-Estradiol induces cyclin D1 gene transcription, p36D1-p34cdk4 complex activation and p105Rb phosphorylation during mitogenic stimulation of G(1)-arrested human breast cancer cells, *Oncogene 12*, 2315-2324.

216. Foster, J. S. and Wimalasena, J. (1996) Estrogen regulates activity of cyclin-dependent kinases and retinoblastoma protein phosphorylation in breast cancer cells, *Mol. Endocrinol. 10*, 488-498.

217. Dubik, D., Dembinski, T. C., and Shiu, R. P. (1987) Stimulation of c-myc oncogene expression associated with estrogen-induced proliferation of human breast cancer cells, *Cancer Res. 47*, 6517-6521.

218. Yu, W. C., Leung, B. S., and Gao, Y. L. (1981) Effects of 17 beta-estradiol on progesterone receptors and the uptake of thymidine in human breast cancer cell line CAMA-1, *Cancer Res. 41*, 5004-5009.

219. Bajic, V. B., Tan, S. L., Chong, A., Tang, S., Strom, A., Gustafsson, J. A., Lin, C. Y., and Liu, E. T. (2003) Dragon ERE Finder version 2: A tool for accurate detection and analysis of estrogen response elements in vertebrate genomes, *Nucleic Acids Res. 31*, 3605-3607.

220. Tang, S., Zhang, Z., Tan, S. L., Tang, M. H., Kumar, A. P., Ramadoss, S. K., and Bajic, V. B. (2007) KBERG: KnowledgeBase for Estrogen Responsive Genes, *Nucleic Acids Res. 35*, D732-D736.

221. Jin, V. X., Sun, H., Pohar, T. T., Liyanarachchi, S., Palaniswamy, S. K., Huang, T. H., and Davuluri, R. V. (2005) ERTargetDB: an integral information resource of transcription regulation of estrogen receptor target genes, *J. Mol. Endocrinol. 35*, 225-230.

222. Bajic, V. B., Tan, S. L., Chong, A., Tang, S., Strom, A., Gustafsson, J. A., Lin, C. Y., and Liu, E. T. (2003) Dragon ERE Finder version 2: A tool for accurate detection and analysis of estrogen response elements in vertebrate genomes, *Nucleic Acids Res. 31*, 3605-3607.

223. Tang, S., Zhang, Z., Tan, S. L., Tang, M. H., Kumar, A. P., Ramadoss, S. K., and Bajic, V. B. (2007) KBERG: KnowledgeBase for Estrogen Responsive Genes, *Nucleic Acids Res. 35*, D732-D736.

224. Jin, V. X., Sun, H., Pohar, T. T., Liyanarachchi, S., Palaniswamy, S. K., Huang, T. H., and Davuluri, R. V. (2005) ERTargetDB: an integral information resource of transcription regulation of estrogen receptor target genes, *J. Mol. Endocrinol. 35*, 225-230.

225. Dennis, G., Jr., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., and Lempicki, R. A. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery, *Genome Biol. 4*, 3.

226. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Res. 13*, 2498-2504.

227. Bajic, V. B., Tan, S. L., Chong, A., Tang, S., Strom, A., Gustafsson, J. A., Lin, C. Y., and Liu, E. T. (2003) Dragon ERE Finder version 2: A tool for accurate detection and analysis of estrogen response elements in vertebrate genomes, *Nucleic Acids Res. 31*, 3605-3607.

228. McKenna, N. J. and O'Malley, B. W. (2002) Combinatorial control of gene expression by nuclear receptors and coregulators, *Cell 108*, 465-474.

229. Hyder, S. M., Chiappetta, C., and Stancel, G. M. (1999) Interaction of human estrogen receptors alpha and beta with the same naturally occurring estrogen response elements, *Biochem. Pharmacol. 57*, 597-601.

230. Liu, M. M., Albanese, C., Anderson, C. M., Hilty, K., Webb, P., Uht, R. M., Price, R. H., Jr., Pestell, R. G., and Kushner, P. J. (2002) Opposing action of estrogen receptors alpha and beta on cyclin D1 gene expression, *J. Biol. Chem. 277*, 24353-24360.

231. Ikeda, K. and Kawakami, K. (1995) DNA binding through distinct domains of zinc-finger-homeodomain protein AREB6 has different effects on gene transcription, *Eur. J. Biochem. 233*, 73-82.

232. Wang, J., Lee, S., Teh, C. E., Bunting, K., Ma, L., and Shannon, M. F. (2009) The transcription repressor, ZEB1, cooperates with CtBP2 and HDAC1 to suppress IL-2 gene activation in T cells, *Int. Immunol. 21*, 227-235.

233. Adachi, Y., Takeuchi, T., Nagayama, T., Ohtsuki, Y., and Furihata, M. (2009) Zeb1-mediated T-cadherin repression increases the invasive potential of gallbladder cancer, *FEBS Lett. 583*, 430-436.

234. Huang, W., Zhang, Y., Varambally, S., Chinnaiyan, A. M., Banerjee, M., Merajver, S. D., and Kleer, C. G. (2008) Inhibition of CCN6 (Wnt-1-induced signaling protein 3) down-regulates E-cadherin in the breast epithelium through induction of snail and ZEB1, *Am. J. Pathol. 172*, 893-904.

235. Dominguez, G., Pena, C., Silva, J., Garcia, J. M., Garcia, V., Rodriguez, R., Cantos, B., Citores, M. J., Espana, P., and Bonilla, F. (2006) The presence of an intronic deletion in p73 and high levels of ZEB1 alter the TAp73/DeltaTAp73 ratio in colorectal carcinomas, *J. Pathol. 210*, 390-397.

236. Aigner, K., Descovich, L., Mikula, M., Sultan, A., Dampier, B., Bonne, S., van, R. F., Mikulits, W., Schreiber, M., Brabletz, T., Sommergruber, W., Schweifer, N., Wernitznig, A., Beug, H., Foisner, R., and Eger, A. (2007) The transcription factor ZEB1 (deltaEF1) represses Plakophilin 3 during human cancer progression, *FEBS Lett. 581*, 1617-1624.

237. Kraus, R. J., Perrigoue, J. G., and Mertz, J. E. (2003) ZEB negatively regulates the lytic-switch BZLF1 gene promoter of Epstein-Barr virus, *J. Virol. 77*, 199-207.

238. Bracken, C. P., Gregory, P. A., Kolesnikoff, N., Bert, A. G., Wang, J., Shannon, M. F., and Goodall, G. J. (2008) A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition, *Cancer Res. 68*, 7846-7854.

239. Gregory, P. A., Bert, A. G., Paterson, E. L., Barry, S. C., Tsykin, A., Farshid, G., Vadas, M. A., Khew-Goodall, Y., and Goodall, G. J. (2008) The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1, *Nat. Cell Biol. 10*, 593-601.

240. Bracken, C. P., Gregory, P. A., Kolesnikoff, N., Bert, A. G., Wang, J., Shannon, M. F., and Goodall, G. J. (2008) A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition, *Cancer Res. 68*, 7846-7854.

241. Gregory, P. A., Bracken, C. P., Bert, A. G., and Goodall, G. J. (2008) MicroRNAs as regulators of epithelial-mesenchymal transition, *Cell Cycle 7*, 3112-3118.

242. Aigner, K., Dampier, B., Descovich, L., Mikula, M., Sultan, A., Schreiber, M., Mikulits, W., Brabletz, T., Strand, D., Obrist, P., Sommergruber, W., Schweifer, N., Wernitznig, A., Beug, H., Foisner, R., and Eger, A. (2007) The transcription factor ZEB1 (deltaEF1) promotes tumour cell dedifferentiation by repressing master regulators of epithelial polarity, *Oncogene 26*, 6979-6988.

243. Singh, M., Spoelstra, N. S., Jean, A., Howe, E., Torkko, K. C., Clark, H. R., Darling, D. S., Shroyer, K. R., Horwitz, K. B., Broaddus, R. R., and Richer, J. K. (2008) ZEB1 expression in type I vs type II endometrial cancers: a marker of aggressive disease, *Mod. Pathol. 21*, 912-923.

244. Paca-Uccaralertkun, S., Zhao, L. J., Adya, N., Cross, J. V., Cullen, B. R., Boros, I. M., and Giam, C. Z. (1994) In vitro selection of DNA elements highly responsive to

the human T-cell lymphotropic virus type I transcriptional activator, Tax, *Mol. Cell Biol. 14*, 456-462.

245. Cheng, J. C., Esparza, S., Sandoval, S., Shankar, D., Fu, C., and Sakamoto, K. M. (2007) Potential role of CREB as a prognostic marker in acute myeloid leukemia, *Future. Oncol. 3*, 475-480.

246. Linnerth, N. M., Greenaway, J. B., Petrik, J. J., and Moorehead, R. A. (2008) cAMP response element-binding protein is expressed at high levels in human ovarian adenocarcinoma and regulates ovarian tumor cell proliferation, *Int. J. Gynecol. Cancer 18*, 1248-1257.

247. Chinnappan, D., Qu, X., Xiao, D., Ratnasari, A., and Weber, H. C. (2008) Human gastrin-releasing peptide receptor gene regulation requires transcription factor binding at two distinct CRE sites, *Am. J. Physiol Gastrointest. Liver Physiol 295*, G153-G162.

248. Chhabra, A., Fernando, H., Watkins, G., Mansel, R. E., and Jiang, W. G. (2007) Expression of transcription factor CREB1 in human breast cancer and its correlation with prognosis, *Oncol. Rep. 18*, 953-958.

249. Qu, X., Xiao, D., and Weber, H. C. (2002) Human gastrin-releasing peptide receptor mediates sustained CREB phosphorylation and transactivation in HuTu 80 duodenal cancer cells, *FEBS Lett. 527*, 109-113.

250. Garcia, G. E., Nicole, A., Bhaskaran, S., Gupta, A., Kyprianou, N., and Kumar, A. P. (2006) Akt-and CREB-mediated prostate cancer cell proliferation inhibition by Nexrutine, a Phellodendron amurense extract, *Neoplasia. 8*, 523-533.

251. Pigazzi, M., Manara, E., Baron, E., and Basso, G. (2009) miR-34b targets cyclic AMP-responsive element binding protein in acute myeloid leukemia, *Cancer Res. 69*, 2471-2478.

252. Cheng, J. C., Esparza, S., Sandoval, S., Shankar, D., Fu, C., and Sakamoto, K. M. (2007) Potential role of CREB as a prognostic marker in acute myeloid leukemia, *Future. Oncol. 3*, 475-480.

253. Kibler, K. V. and Jeang, K. T. (2001) CREB/ATF-dependent repression of cyclin a by human T-cell leukemia virus type 1 Tax protein, *J. Virol. 75*, 2161-2173.

254. Adya, N., Zhao, L. J., Huang, W., Boros, I., and Giam, C. Z. (1994) Expansion of CREB's DNA recognition specificity by Tax results from interaction with Ala-Ala-Arg at positions 282-284 near the conserved DNA-binding domain of CREB, *Proc. Natl. Acad. Sci. U. S. A 91*, 5642-5646.

255. Adam, E., Kerkhofs, P., Mammerickx, M., Burny, A., Kettmann, R., and Willems, L. (1996) The CREB, ATF-1, and ATF-2 transcription factors from bovine leukemia virus-infected B lymphocytes activate viral expression, *J. Virol. 70*, 1990-1999.

256. Johansen, L. M., Deppmann, C. D., Erickson, K. D., Coffin, W. F., III, Thornton, T. M., Humphrey, S. E., Martin, J. M., and Taparowsky, E. J. (2003) EBNA2 and activated Notch induce expression of BATF, *J. Virol. 77*, 6029-6040.

257. Shaywitz, A. J. and Greenberg, M. E. (1999) CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals, *Annu. Rev. Biochem. 68*, 821-861.

258. Kraus, R. J., Perrigoue, J. G., and Mertz, J. E. (2003) ZEB negatively regulates the lytic-switch BZLF1 gene promoter of Epstein-Barr virus, *J. Virol. 77*, 199-207.

259. Saeed, A. I., Bhagabati, N. K., Braisted, J. C., Liang, W., Sharov, V., Howe, E. A., Li, J., Thiagarajan, M., White, J. A., and Quackenbush, J. (2006) TM4 microarray software suite, *Methods Enzymol. 411*, 134-193.

260. Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., and Quackenbush, J. (2003) TM4: a free, open-source system for microarray data management and analysis, *Biotechniques 34*, 374-378.

261. Wirtenberger, M., Tchatchou, S., Hemminki, K., Klaes, R., Schmutzler, R. K., Bermejo, J. L., Chen, B., Wappenschmidt, B., Meindl, A., Bartram, C. R., and Burwinkel, B. (2006) Association of genetic variants in the Rho guanine nucleotide exchange factor AKAP13 with familial breast cancer, *Carcinogenesis 27*, 593-598.

262. Peinado, H., Del Carmen Iglesias-de la Cruz, Olmeda, D., Csiszar, K., Fong, K. S., Vega, S., Nieto, M. A., Cano, A., and Portillo, F. (2005) A molecular role for lysyl oxidase-like 2 enzyme in snail regulation and tumor progression, *EMBO J. 24*, 3446-3458.

263. Lind, A. K., hm-Kahler, P., Weijdegard, B., Sundfeldt, A. K., and Brannstrom, M. (2006) Gelatinases and their tissue inhibitors during human ovulation: increased expression of tissue inhibitor of matrix metalloproteinase-1, *Mol. Hum. Reprod. 12*, 725-736.

264. Troeberg, L., Tanaka, M., Wait, R., Shi, Y. E., Brew, K., and Nagase, H. (2002) E. coli expression of TIMP-4 and comparative kinetic studies with TIMP-1 and TIMP-2: insights into the interactions of TIMPs and matrix metalloproteinase 2 (gelatinase A), *Biochemistry 41*, 15025-15035.

265. Prast, J., Saleh, L., Husslein, H., Sonderegger, S., Helmer, H., and Knofler, M. (2008) Human chorionic gonadotropin stimulates trophoblast invasion through extracellularly regulated kinase and AKT signaling, *Endocrinology 149*, 979-987.

266. Guzman, J. M., Rubio, M., Ortiz-Delgado, J. B., Klenke, U., Kight, K., Cross, I., Sanchez-Ramos, I., Riaza, A., Rebordinos, L., Sarasquete, C., Zohar, Y., and Mananos, E. L. (2009) Comparative gene expression of gonadotropins (FSH and LH) and peptide levels of gonadotropin-releasing hormones (GnRHs) in the pituitary of wild and cultured Senegalese sole (Solea senegalensis) broodstocks, *Comp Biochem. Physiol A Mol. Integr. Physiol.*

267. Miao, M. F. and Huang, H. F. (2009) Dynamic measurements of serum inhibin B and estradiol: a predictive evaluation of ovarian response to gonadotrophin stimulation in the early stage of IVF treatment, *J. Zhejiang. Univ Sci. B 10*, 35-45.

268. Choi, H. J., Chung, Y. S., Kim, H. J., Moon, U. Y., Choi, Y. H., Van, S., I, Baek, S. J., Yoon, H. G., and Yoon, J. H. (2009) Signal pathway of 17beta-estradiol-induced MUC5B expression in human airway epithelial cells, *Am. J. Respir. Cell Mol. Biol. 40*, 168-178.

269. Gilbert, N., Gilchrist, S., and Bickmore, W. A. (2005) Chromatin organization in the mammalian nucleus, *Int. Rev. Cytol. 242*, 283-336.

270. Cremer, T. and Cremer, C. (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells, *Nat. Rev. Genet. 2*, 292-301.

271. Zabetian, C. P., Anderson, G. M., Buxbaum, S. G., Elston, R. C., Ichinose, H., Nagatsu, T., Kim, K. S., Kim, C. H., Malison, R. T., Gelernter, J., and Cubells, J. F. (2001) A quantitative-trait analysis of human plasma-dopamine beta-hydroxylase activity: evidence for a major functional polymorphism at the DBH locus, *Am. J. Hum. Genet. 68*, 515-522.

272. Kuriki, C., Tanaka, T., Fukui, Y., Sato, O., and Motojima, K. (2002) Structural and functional analysis of a new upstream promoter of the human FAT/CD36 gene, *Biol. Pharm. Bull. 25*, 1476-1478.

273. Collins, F. S., Brooks, L. D., and Chakravarti, A. (1998) A DNA polymorphism discovery resource for research on human genetic variation, *Genome Res. 8*, 1229-1231.

274. Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., and Sirotkin, K. (2001) dbSNP: the NCBI database of genetic variation, *Nucleic Acids Res. 29*, 308-311.

275. Marcus, P. M., Vineis, P., and Rothman, N. (2000) NAT2 slow acetylation and bladder cancer risk: a meta-analysis of 22 case-control studies conducted in the general population, *Pharmacogenetics 10*, 115-122.

276. Bell, D. A., Taylor, J. A., Butler, M. A., Stephens, E. A., Wiest, J., Brubaker, L. H., Kadlubar, F. F., and Lucier, G. W. (1993) Genotype/phenotype discordance for human arylamine N-acetyltransferase (NAT2) reveals a new slow-acetylator allele common in African-Americans, *Carcinogenesis 14*, 1689-1692.

277. Reddy, A. J., Christie, J. D., Aplenc, R., Fuchs, B., Lanken, P. N., and Kleeberger, S. R. (2008) ASSOCIATION OF HUMAN NAD(P)H:QUINONE OXIDOREDUCTASE 1 (NQO1) POLYMORPHISM WITH DEVELOPMENT OF ACUTE LUNG INJURY, *J. Cell Mol. Med.*

278. Willems, P., Magri, V., Cretnik, M., Fasano, M., Jakubowska, A., Levanat, S., Lubinski, J., Marras, E., Musani, V., Thierens, H., Vandersickel, V., Perletti, G., and Vral, A. (2009) Characterization of the c.190T>C missense mutation in BRCA1 codon 64 (Cys64Arg), *Int. J. Oncol. 34*, 1005-1015.

279. Osawa, M., Kimura, R., Hasegawa, I., Mukasa, N., and Satoh, F. (2009) SNP association and sequence analysis of the NOS1AP gene in SIDS, *Leg. Med. (Tokyo)*.

280. Kidd, L. C., Paltoo, D. N., Wang, S., Chen, W., Akereyeni, F., Isaacs, W., Ahaghotu, C., and Kittles, R. (2005) Sequence variation within the 5' regulatory regions of the vitamin D binding protein and receptor genes and prostate cancer risk, *Prostate 64*, 272-282.

281. Hohjoh, H. and Tokunaga, K. (2001) Allele-specific binding of the ubiquitous transcription factor OCT-1 to the functional single nucleotide polymorphism (SNP) sites in the tumor necrosis factor-alpha gene (TNFA) promoter, *Genes Immun. 2*, 105-109.

282. Shah, R., Hurley, C. K., and Posch, P. E. (2006) A molecular mechanism for the differential regulation of TGF-beta1 expression due to the common SNP -509C-T (c. -1347C > T), *Hum. Genet. 120*, 461-469.

283. Cao, Y. Y., Ge, H., Chen, L. Q., Chen, Z. F., Wen, D. G., Li, Y., and Zhang, J. H. (2007) [Correlation of 53BP1 and p53 polymorphisms to susceptibility to esophageal squamous cell carcinoma and gastric cardiac adenocarcinoma], *Ai. Zheng. 26*, 1052-1057.

284. Guo, W., Wang, N., Wang, Y. M., Li, Y., Wen, D. G., Chen, Z. F., He, Y. T., and Zhang, J. H. (2005) Interleukin-10 -1082 promoter polymorphism is not associated with susceptibility to esophageal squamous cell carcinoma and gastric cardiac adenocarcinoma in a population of high-incidence region of north China, *World J. Gastroenterol. 11*, 858-862.

285. Jin, X., Kuang, G., Wei, L. Z., Li, Y., Wang, R., Guo, W., Wang, N., Fang, S. M., Wen, D. G., Chen, Z. F., and Zhang, J. H. (2005) No association of the matrix metalloproteinase 1 promoter polymorphism with susceptibility to esophageal squamous cell carcinoma and gastric cardiac adenocarcinoma in northern China, *World J. Gastroenterol. 11*, 2385-2389.

286. Zhang, X. F., Wang, Y. M., Wang, R., Wei, L. Z., Li, Y., Guo, W., Wang, N., and Zhang, J. H. (2005) [Correlation of E-cadherin polymorphisms to esophageal squamous cell carcinoma and gastric cardiac adenocarcinoma], *Ai. Zheng. 24*, 513-519.

287. Zhang, X. J., Guo, W., Wang, N., Zhou, R. M., Dong, X. J., and Li, Y. (2006) [The association of MMP-13 polymorphism with susceptibility to esophageal squamous cell carcinoma and gastric cardiac adenocarcinoma], *Yi. Chuan 28*, 1500-1504.

288. Zhang, J., Jin, X., Fang, S., Li, Y., Wang, R., Guo, W., Wang, N., Wang, Y., Wen, D., Wei, L., Kuang, G., and Dong, Z. (2004) The functional SNP in the matrix metalloproteinase-3 promoter modifies susceptibility and lymphatic metastasis in esophageal squamous cell carcinoma but not in gastric cardiac adenocarcinoma, *Carcinogenesis 25*, 2519-2524.

289. Tan, W., Miao, X., Wang, L., Yu, C., Xiong, P., Liang, G., Sun, T., Zhou, Y., Zhang, X., Li, H., and Lin, D. (2005) Significant increase in risk of gastroesophageal

cancer is associated with interaction between promoter polymorphisms in thymidylate synthase and serum folate status, *Carcinogenesis 26*, 1430-1435.

290. Tang, W. Y., Wang, L., Li, C., Hu, Z. B., Chen, R., Zhu, Y. J., Shen, H. B., Wei, Q. Y., and Zhou, J. W. (2007) Identification and functional characterization of JWA polymorphisms and their association with risk of gastric cancer and esophageal squamous cell carcinoma in a Chinese population, *J. Toxicol. Environ. Health A 70*, 885-894.

291. Wang, X., Tomso, D. J., Liu, X., and Bell, D. A. (2005) Single nucleotide polymorphism in transcriptional regulatory regions and expression of environmentally responsive genes, *Toxicol. Appl. Pharmacol. 207*, 84-90.

292. Ponomarenko, J. V., Orlova, G. V., Merkulova, T. I., Gorshkova, E. V., Fokin, O. N., Vasiliev, G. V., Frolov, A. S., and Ponomarenko, M. P. (2002) rSNP_Guide: an integrated database-tools system for studying SNPs and site-directed mutations in transcription factor binding sites, *Hum. Mutat. 20*, 239-248.

293. Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Gibbs, R. A., Belmont, J. W., Boudreau, A., Hardenbol, P., Leal, S. M., Pasternak, S., Wheeler, D. A., Willis, T. D., Yu, F., Yang, H., Zeng, C., Gao, Y., Hu, H., Hu, W., Li, C., Lin, W., Liu, S., Pan, H., Tang, X., Wang, J., Wang, W., Yu, J., Zhang, B., Zhang, Q., Zhao, H., Zhao, H., Zhou, J., Gabriel, S. B., Barry, R., Blumenstiel, B., Camargo, A., Defelice, M., Faggart, M., Goyette, M., Gupta, S., Moore, J., Nguyen, H., Onofrio, R. C., Parkin, M., Roy, J., Stahl, E., Winchester, E., Ziaugra, L., Altshuler, D., Shen, Y., Yao, Z., Huang, W., Chu, X., He, Y., Jin, L., Liu, Y., Shen, Y., Sun, W., Wang, H., Wang, Y., Wang, Y., Xiong, X., Xu, L., Waye, M. M., Tsui, S. K., Xue, H., Wong, J. T., Galver, L. M., Fan, J. B., Gunderson, K., Murray, S. S., Oliphant, A. R., Chee, M. S., Montpetit, A., Chagnon, F., Ferretti, V., Leboeuf, M., Olivier, J. F., Phillips, M. S., Roumy, S., Sallee, C., Verner, A., Hudson, T. J., Kwok, P. Y., Cai, D., Koboldt, D. C., Miller, R. D., Pawlikowska, L., Taillon-Miller, P., Xiao, M., Tsui, L. C., Mak, W., Song, Y. Q., Tam, P. K., Nakamura, Y., Kawaguchi, T., Kitamoto, T., Morizono, T., Nagashima, A., Ohnishi, Y., Sekine, A., Tanaka, T., Tsunoda, T., Deloukas, P., Bird, C. P., Delgado, M., Dermitzakis, E. T., Gwilliam, R., Hunt, S., Morrison, J., Powell, D., Stranger, B. E., Whittaker, P., Bentley, D. R., Daly, M. J., de Bakker, P. I., Barrett, J., Chretien, Y. R., Maller, J., McCarroll, S., Patterson, N., Pe'er, I., Price, A., Purcell, S., Richter, D. J., Sabeti, P., Saxena, R., Schaffner, S. F., Sham, P. C., Varilly, P., Altshuler, D., Stein, L. D., Krishnan, L., Smith, A. V., Tello-Ruiz, M. K., Thorisson, G. A., Chakravarti, A., Chen, P. E., Cutler, D. J., Kashuk, C. S., Lin, S., Abecasis, G. R., Guan, W., Li, Y., Munro, H. M., Qin, Z. S., Thomas, D. J., McVean, G., Auton, A., Bottolo, L., Cardin, N., Eyheramendy, S., Freeman, C., Marchini, J., Myers, S., Spencer, C., Stephens, M., Donnelly, P., Cardon, L. R., Clarke, G., Evans, D. M., Morris, A. P., Weir, B. S., Tsunoda, T., Mullikin, J. C., Sherry, S. T., Feolo, M., Skol, A., Zhang, H., Zeng, C., Zhao, H., Matsuda, I., Fukushima, Y., Macer, D. R., Suda, E., Rotimi, C. N., Adebamowo, C. A., Ajayi, I., Aniagwu, T., Marshall, P. A., Nkwodimmah, C., Royal, C. D., Leppert, M. F., Dixon, M., Peiffer, A., Qiu, R., Kent, A., Kato, K., Niikawa, N., Adewole, I. F., Knoppers, B. M., Foster, M. W., Clayton, E. W., Watkin, J., Gibbs, R. A., Belmont, J. W., Muzny, D., Nazareth, L., Sodergren, E., Weinstock, G. M., Wheeler, D. A., Yakub, I., Gabriel, S. B., Onofrio, R. C., Richter,

D. J., Ziaugra, L., Birren, B. W., Daly, M. J., Altshuler, D., Wilson, R. K., Fulton, L. L., Rogers, J., Burton, J., Carter, N. P., Clee, C. M., Griffiths, M., Jones, M. C., McLay, K., Plumb, R. W., Ross, M. T., Sims, S. K., Willey, D. L., Chen, Z., Han, H., Kang, L., Godbout, M., Wallenburg, J. C., L'Archeveque, P., Bellemare, G., Saeki, K., Wang, H., An, D., Fu, H., Li, Q., Wang, Z., Wang, R., Holden, A. L., Brooks, L. D., McEwen, J. E., Guyer, M. S., Wang, V. O., Peterson, J. L., Shi, M., Spiegel, J., Sung, L. M., Zacharia, L. F., Collins, F. S., Kennedy, K., Jamieson, R., and Stewart, J. (2007) A second generation human haplotype map of over 3.1 million SNPs, *Nature 449*, 851-861.

294. Wang, X., Tomso, D. J., Liu, X., and Bell, D. A. (2005) Single nucleotide polymorphism in transcriptional regulatory regions and expression of environmentally responsive genes, *Toxicol. Appl. Pharmacol. 207*, 84-90.

295. Lim, W. K., Lyashenko, E., and Califano, A. (2009) Master regulators used as breast cancer metastasis classifier, *Pac. Symp. Biocomput.* 504-515.

296. Nilsson, R., Bajic, V. B., Suzuki, H., di, B. D., Bjorkegren, J., Katayama, S., Reid, J. F., Sweet, M. J., Gariboldi, M., Carninci, P., Hayashizaki, Y., Hume, D. A., Tegner, J., and Ravasi, T. (2006) Transcriptional network dynamics in macrophage activation, *Genomics 88*, 133-142.

297. Borneman, A. R., Leigh-Bell, J. A., Yu, H., Bertone, P., Gerstein, M., and Snyder, M. (2006) Target hub proteins serve as master regulators of development in yeast, *Genes Dev. 20*, 435-448.

298. Urbanek, P., Wang, Z. Q., Fetka, I., Wagner, E. F., and Busslinger, M. (1994) Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP, *Cell 79*, 901-912.

299. Horcher, M., Souabni, A., and Busslinger, M. (2001) Pax5/BSAP maintains the identity of B cells in late B lymphopoiesis, *Immunity. 14*, 779-790.

300. Morrison, A. M., Nutt, S. L., Thevenin, C., Rolink, A., and Busslinger, M. (1998) Loss- and gain-of-function mutations reveal an important role of BSAP (Pax-5) at the start and end of B cell differentiation, *Semin. Immunol. 10*, 133-142.

301. Sica, G., Vazquez, M. F., Altorki, N., Port, J., Lee, P. C., Liu, Y., Hyjek, E., and Saqi, A. (2008) PAX-5 expression in pulmonary neuroendocrine neoplasms: its usefulness in surgical and fine-needle aspiration biopsy specimens, *Am. J. Clin. Pathol. 129*, 556-562.

302. An, Q., Wright, S. L., Konn, Z. J., Matheson, E., Minto, L., Moorman, A. V., Parker, H., Griffiths, M., Ross, F. M., Davies, T., Hall, A. G., Harrison, C. J., Irving, J. A., and Strefford, J. C. (2008) Variable breakpoints target PAX5 in patients with dicentric chromosomes: a model for the basis of unbalanced translocations in cancer, *Proc. Natl. Acad. Sci. U. S. A 105*, 17050-17054.

303. Kawamata, N., Ogawa, S., Zimmermann, M., Niebuhr, B., Stocking, C., Sanada, M., Hemminki, K., Yamatomo, G., Nannya, Y., Koehler, R., Flohr, T., Miller, C. W., Harbott, J., Ludwig, W. D., Stanulla, M., Schrappe, M., Bartram, C. R., and

Koeffler, H. P. (2008) Cloning of genes involved in chromosomal translocations by high-resolution single nucleotide polymorphism genomic microarray, *Proc. Natl. Acad. Sci. U. S. A 105*, 11921-11926.

304. Kanteti, R., Nallasura, V., Loganathan, S., Tretiakova, M., Kroll, T., Krishnaswamy, S., Faoro, L., Cagle, P., Husain, A. N., Vokes, E. E., Lang, D., and Salgia, R. (2009) PAX5 is expressed in small-cell lung cancer and positively regulates c-Met transcription, *Lab Invest 89*, 301-314.

305. Zhang, J., Hagopian-Donaldson, S., Serbedzija, G., Elsemore, J., Plehn-Dujowich, D., McMahon, A. P., Flavell, R. A., and Williams, T. (1996) Neural tube, skeletal and body wall defects in mice lacking transcription factor AP-2, *Nature 381*, 238-241.

306. Kannan, P., Buettner, R., Chiao, P. J., Yim, S. O., Sarkiss, M., and Tainsky, M. A. (1994) N-ras oncogene causes AP-2 transcriptional self-interference, which leads to transformation, *Genes Dev. 8*, 1258-1269.

307. Orso, F., Fassetta, M., Penna, E., Solero, A., De, F. K., Sismondi, P., De, B. M., and Taverna, D. (2007) The AP-2alpha transcription factor regulates tumor cell migration and apoptosis, *Adv. Exp. Med. Biol. 604*, 87-95.

308. Gee, J. M., Robertson, J. F., Ellis, I. O., Nicholson, R. I., and Hurst, H. C. (1999) Immunohistochemical analysis reveals a tumour suppressor-like role for the transcription factor AP-2 in invasive breast cancer, *J. Pathol. 189*, 514-520.

309. Schwartz, B., Melnikova, V. O., Tellez, C., Mourad-Zeidan, A., Blehm, K., Zhao, Y. J., McCarty, M., Adam, L., and Bar-Eli, M. (2007) Loss of AP-2alpha results in deregulation of E-cadherin and MMP-9 and an increase in tumorigenicity of colon cancer cells in vivo, *Oncogene 26*, 4049-4058.

310. Ruiz, M., Pettaway, C., Song, R., Stoeltzing, O., Ellis, L., and Bar-Eli, M. (2004) Activator protein 2alpha inhibits tumorigenicity and represses vascular endothelial growth factor transcription in prostate cancer cells, *Cancer Res. 64*, 631-638.

311. Dalgin, G. S., Drever, M., Williams, T., King, T., DeLisi, C., and Liou, L. S. (2008) Identification of novel epigenetic markers for clear cell renal cell carcinoma, *J. Urol. 180*, 1126-1130.

312. Wajapeyee, N. and Somasundaram, K. (2003) Cell cycle arrest and apoptosis induction by activator protein 2alpha (AP-2alpha) and the role of p53 and p21WAF1/CIP1 in AP-2alpha-mediated growth inhibition, *J. Biol. Chem. 278*, 52093-52101.

313. Heimberger, A. B., McGary, E. C., Suki, D., Ruiz, M., Wang, H., Fuller, G. N., and Bar-Eli, M. (2005) Loss of the AP-2alpha transcription factor is associated with the grade of human gliomas, *Clin. Cancer Res. 11*, 267-272.

314. Brown, L., Paraso, M., Arkell, R., and Brown, S. (2005) In vitro analysis of partial loss-of-function ZIC2 mutations in holoprosencephaly: alanine tract expansion modulates DNA binding and transactivation, *Hum. Mol. Genet. 14*, 411-420.

315. Nagai, T., Aruga, J., Minowa, O., Sugimoto, T., Ohno, Y., Noda, T., and Mikoshiba, K. (2000) Zic2 regulates the kinetics of neurulation, *Proc. Natl. Acad. Sci. U. S. A 97*, 1618-1623.

316. Pfister, S., Schlaeger, C., Mendrzyk, F., Wittmann, A., Benner, A., Kulozik, A., Scheurlen, W., Radlwimmer, B., and Lichter, P. (2007) Array-based profiling of reference-independent methylation status (aPRIMES) identifies frequent promoter methylation and consecutive downregulation of ZIC2 in pediatric medulloblastoma, *Nucleic Acids Res. 35*, e51.

317. Shimada, H., Nakashima, K., Ochiai, T., Nabeya, Y., Takiguchi, M., Nomura, F., and Hiwasa, T. (2005) Serological identification of tumor antigens of esophageal squamous cell carcinoma, *Int. J. Oncol. 26*, 77-86.

318. Siu, Y. T. and Jin, D. Y. (2007) CREB--a real culprit in oncogenesis, *FEBS J. 274*, 3224-3232.

319. Linnerth, N. M., Greenaway, J. B., Petrik, J. J., and Moorehead, R. A. (2008) cAMP response element-binding protein is expressed at high levels in human ovarian adenocarcinoma and regulates ovarian tumor cell proliferation, *Int. J. Gynecol. Cancer 18*, 1248-1257.

320. Chhabra, A., Fernando, H., Watkins, G., Mansel, R. E., and Jiang, W. G. (2007) Expression of transcription factor CREB1 in human breast cancer and its correlation with prognosis, *Oncol. Rep. 18*, 953-958.

321. Qu, X., Xiao, D., and Weber, H. C. (2002) Human gastrin-releasing peptide receptor mediates sustained CREB phosphorylation and transactivation in HuTu 80 duodenal cancer cells, *FEBS Lett. 527*, 109-113.

322. Shankar, D. B., Cheng, J. C., Kinjo, K., Federman, N., Moore, T. B., Gill, A., Rao, N. P., Landaw, E. M., and Sakamoto, K. M. (2005) The role of CREB as a proto-oncogene in hematopoiesis and in acute myeloid leukemia, *Cancer Cell 7*, 351-362.

323. So, C. K., Nie, Y., Song, Y., Yang, G. Y., Chen, S., Wei, C., Wang, L. D., Doggett, N. A., and Yang, C. S. (2004) Loss of heterozygosity and internal tandem duplication mutations of the CBP gene are frequent events in human esophageal squamous cell carcinoma, *Clin. Cancer Res. 10*, 19-27.

324. Borrow, J., Stanton, V. P., Jr., Andresen, J. M., Becher, R., Behm, F. G., Chaganti, R. S., Civin, C. I., Disteche, C., Dube, I., Frischauf, A. M., Horsman, D., Mitelman, F., Volinia, S., Watmore, A. E., and Housman, D. E. (1996) The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein, *Nat. Genet. 14*, 33-41.

325. Rowley, J. D., Reshmi, S., Sobulo, O., Musvee, T., Anastasi, J., Raimondi, S., Schneider, N. R., Barredo, J. C., Cantu, E. S., Schlegelberger, B., Behm, F., Doggett, N. A., Borrow, J., and Zeleznik-Le, N. (1997) All patients with the T(11;16)(q23;p13.3) that involves MLL and CBP have treatment-related hematologic disorders, *Blood 90*, 535-541.

326. Sobulo, O. M., Borrow, J., Tomek, R., Reshmi, S., Harden, A., Schlegelberger, B., Housman, D., Doggett, N. A., Rowley, J. D., and Zeleznik-Le, N. J. (1997) MLL is fused to CBP, a histone acetyltransferase, in therapy-related acute myeloid leukemia with a t(11;16)(q23;p13.3), *Proc. Natl. Acad. Sci. U. S. A 94*, 8732-8737.

327. Hayashi, R., Goto, Y., Ikeda, R., Yokoyama, K. K., and Yoshida, K. (2006) CDCA4 is an E2F transcription factor family-induced nuclear factor that regulates E2F-dependent transcriptional activation and cell proliferation, *J. Biol. Chem. 281*, 35633-35648.

328. Salon, C., Eymin, B., Micheau, O., Chaperot, L., Plumas, J., Brambilla, C., Brambilla, E., and Gazzeri, S. (2006) E2F1 induces apoptosis and sensitizes human lung adenocarcinoma cells to death-receptor-mediated apoptosis through specific downregulation of c-FLIP(short), *Cell Death. Differ. 13*, 260-272.

329. Sylvestre, Y., De, G., V, Querido, E., Mukhopadhyay, U. K., Bourdeau, V., Major, F., Ferbeyre, G., and Chartrand, P. (2007) An E2F/miR-20a autoregulatory feedback loop, *J. Biol. Chem. 282*, 2135-2143.

330. Suh, D. S., Yoon, M. S., Choi, K. U., and Kim, J. Y. (2008) Significance of E2F-1 overexpression in epithelial ovarian cancer, *Int. J. Gynecol. Cancer 18*, 492-498.

331. Du, Y., Zhang, S., Wang, Z., Zhou, W., Luan, M., Yang, X., and Chen, N. (2008) Induction of apoptosis and cell cycle arrest by NS398 in oral squamous cell carcinoma cells via downregulation of E2 promoter-binding factor-1, *Oncol. Rep. 20*, 605-611.

332. Huang, C. L., Liu, D., Nakano, J., Yokomise, H., Ueno, M., Kadota, K., and Wada, H. (2007) E2F1 overexpression correlates with thymidylate synthase and survivin gene expressions and tumor proliferation in non small-cell lung cancer, *Clin. Cancer Res. 13*, 6938-6946.

333. Turowska, O., Nauman, A., Pietrzak, M., Poplawski, P., Master, A., Nygard, M., Bondesson, M., Tanski, Z., and Puzianowska-Kuznicka, M. (2007) Overexpression of E2F1 in clear cell renal cell carcinoma: a potential impact of erroneous regulation by thyroid hormone nuclear receptors, *Thyroid 17*, 1039-1048.

334. Xiao, Q., Li, L., Xie, Y., Tan, N., Wang, C., Xu, J., Xia, K., Gardner, K., and Li, Q. Q. (2007) Transcription factor E2F-1 is upregulated in human gastric cancer tissues and its overexpression suppresses gastric tumor cell proliferation, *Cell Oncol. 29*, 335-349.

335. Ebihara, Y., Miyamoto, M., Shichinohe, T., Kawarada, Y., Cho, Y., Fukunaga, A., Murakami, S., Uehara, H., Kaneko, H., Hashimoto, H., Murakami, Y., Itoh, T., Okushiba, S., Kondo, S., and Katoh, H. (2004) Over-expression of E2F-1 in esophageal squamous cell carcinoma correlates with tumor progression, *Dis. Esophagus. 17*, 150-154.

336. Yamazaki, K., Hasegawa, M., Ohoka, I., Hanami, K., Asoh, A., Nagao, T., Sugano, I., and Ishida, Y. (2005) Increased E2F-1 expression via tumour cell proliferation

and decreased apoptosis are correlated with adverse prognosis in patients with squamous cell carcinoma of the oesophagus, *J. Clin. Pathol. 58*, 904-910.

337. Fujita, Y., Sakakura, C., Shimomura, K., Nakanishi, M., Yasuoka, R., Aragane, H., Hagiwara, A., Abe, T., Inazawa, J., and Yamagishi, H. (2003) Chromosome arm 20q gains and other genomic alterations in esophageal squamous cell carcinoma, as analyzed by comparative genomic hybridization and fluorescence in situ hybridization, *Hepatogastroenterology 50*, 1857-1863.

338. Friedrichs, N., Steiner, S., Buettner, R., and Knoepfle, G. (2007) Immunohistochemical expression patterns of AP2alpha and AP2gamma in the developing fetal human breast, *Histopathology 51*, 814-823.

339. Guler, G., Huebner, K., Himmetoglu, C., Jimenez, R. E., Costinean, S., Volinia, S., Pilarski, R. T., Hayran, M., and Shapiro, C. L. (2009) Fragile histidine triad protein, WW domain-containing oxidoreductase protein Wwox, and activator protein 2gamma expression levels correlate with basal phenotype in breast cancer, *Cancer 115*, 899-908.

340. Gee, J. M., Eloranta, J. J., Ibbitt, J. C., Robertson, J. F., Ellis, I. O., Williams, T., Nicholson, R. I., and Hurst, H. C. (2009) Overexpression of TFAP2C in invasive breast cancer correlates with a poorer response to anti-hormone therapy and reduced patient survival, *J. Pathol. 217*, 32-41.

341. Schulte, J. H., Kirfel, J., Lim, S., Schramm, A., Friedrichs, N., Deubzer, H. E., Witt, O., Eggert, A., and Buettner, R. (2008) Transcription factor AP2alpha (TFAP2a) regulates differentiation and proliferation of neuroblastoma cells, *Cancer Lett. 271*, 56-63.

342. Masuhiro, Y., Kayama, K., Fukushima, A., Baba, K., Soutsu, M., Kamiya, Y., Gotoh, M., Yamaguchi, N., and Hanazawa, S. (2008) SOCS-3 inhibits E2F/DP-1 transcriptional activity and cell cycle progression via interaction with DP-1, *J. Biol. Chem. 283*, 31575-31583.

343. Datta, A., Sen, J., Hagen, J., Korgaonkar, C. K., Caffrey, M., Quelle, D. E., Hughes, D. E., Ackerson, T. J., Costa, R. H., and Raychaudhuri, P. (2005) ARF directly binds DP1: interaction with DP1 coincides with the G1 arrest function of ARF, *Mol. Cell Biol. 25*, 8024-8036.

344. Yasui, K., Okamoto, H., Arii, S., and Inazawa, J. (2003) Association of over-expressed TFDP1 with progression of hepatocellular carcinomas, *J. Hum. Genet. 48*, 609-613.

345. Yasui, K., Arii, S., Zhao, C., Imoto, I., Ueda, M., Nagai, H., Emi, M., and Inazawa, J. (2002) TFDP1, CUL4A, and CDC16 identified as targets for amplification at 13q34 in hepatocellular carcinomas, *Hepatology 35*, 1476-1484.

346. Magae, J., Wu, C. L., Illenye, S., Harlow, E., and Heintz, N. H. (1996) Nuclear localization of DP and E2F transcription factors by heterodimeric partners and retinoblastoma protein family members, *J. Cell Sci. 109 (Pt 7)*, 1717-1726.

347. Rubin, S. M., Gall, A. L., Zheng, N., and Pavletich, N. P. (2005) Structure of the Rb C-terminal domain bound to E2F1-DP1: a mechanism for phosphorylation-induced E2F release, *Cell 123*, 1093-1106.

348. Orlic, M., Spencer, C. E., Wang, L., and Gallie, B. L. (2006) Expression analysis of 6p22 genomic gain in retinoblastoma, *Genes Chromosomes. Cancer 45*, 72-82.

349. Cooper, C. S., Nicholson, A. G., Foster, C., Dodson, A., Edwards, S., Fletcher, A., Roe, T., Clark, J., Joshi, A., Norman, A., Feber, A., Lin, D., Gao, Y., Shipley, J., and Cheng, S. J. (2006) Nuclear overexpression of the E2F3 transcription factor in human lung cancer, *Lung Cancer 54*, 155-162.

350. Olsson, A. Y., Feber, A., Edwards, S., Te, P. R., Giddings, I., Merson, S., and Cooper, C. S. (2007) Role of E2F3 expression in modulating cellular proliferation rate in human bladder and prostate cancer cells, *Oncogene 26*, 1028-1037.

351. Oeggerli, M., Tomovska, S., Schraml, P., Calvano-Forte, D., Schafroth, S., Simon, R., Gasser, T., Mihatsch, M. J., and Sauter, G. (2004) E2F3 amplification and overexpression is associated with invasive tumor growth and rapid tumor cell proliferation in urinary bladder cancer, *Oncogene 23*, 5616-5623.

352. Sylvestre, Y., De, G., V, Querido, E., Mukhopadhyay, U. K., Bourdeau, V., Major, F., Ferbeyre, G., and Chartrand, P. (2007) An E2F/miR-20a autoregulatory feedback loop, *J. Biol. Chem. 282*, 2135-2143.

353. Conboy, C. M., Spyrou, C., Thorne, N. P., Wade, E. J., Barbosa-Morais, N. L., Wilson, M. D., Bhattacharjee, A., Young, R. A., Tavare, S., Lees, J. A., and Odom, D. T. (2007) Cell cycle genes are the evolutionarily conserved targets of the E2F4 transcription factor, *PLoS. ONE. 2*, e1061.

354. Moriyama, H., Sasamoto, H., Kambara, T., Matsubara, N., Ikeda, M., Baba, S., Meltzer, S. J., Lynch, H. T., Shimizu, K., and Tanaka, N. (2002) E2F-4 mutation in hereditary non-polyposis colorectal cancer, *J. Exp. Clin. Cancer Res. 21*, 185-189.

355. Parisi, T., Bronson, R. T., and Lees, J. A. (2009) Inhibition of pituitary tumors in Rb mutant chimeras through E2f4 loss reveals a key suppressive role for the pRB/E2F pathway in urothelium and ganglionic carcinogenesis, *Oncogene 28*, 500-508.

356. Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., and Quackenbush, J. (2003) TM4: a free, open-source system for microarray data management and analysis, *Biotechniques 34*, 374-378.

357. Wang, X., Tomso, D. J., Liu, X., and Bell, D. A. (2005) Single nucleotide polymorphism in transcriptional regulatory regions and expression of environmentally responsive genes, *Toxicol. Appl. Pharmacol. 207*, 84-90.

358. Fickett, J. W. and Hatzigeorgiou, A. G. (1997) Eukaryotic promoter recognition, *Genome Res. 7*, 861-878.

359. Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R., and Young, R. A. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells, *Cell 122*, 947-956.

360. Ohler, U. (2006) Identification of core promoter modules in Drosophila and their application in accurate transcription start site prediction, *Nucleic Acids Res. 34*, 5943-5950.

361. Li, X., Zeng, J., and Yan, H. (2008) PCA-HPR: A principle component analysis model for human promoter recognition, *Bioinformation. 2*, 373-378.