2012-05-01

Genetic Approaches to Study Transcriptional Activation and Tumor Suppression: A Dissertation

Ling Lin

*University of Massachusetts Medical School*

Let us know how access to this document benefits you.

Follow this and additional works at: [https://escholarship.umassmed.edu/gsbs_diss](https://escholarship.umassmed.edu/gsbs_diss)

Part of the [Cells Commons](https://escholarship.umassmed.edu/gsbs_diss), [Genetic Phenomena Commons](https://escholarship.umassmed.edu/gsbs_diss), [Genetics and Genomics Commons](https://escholarship.umassmed.edu/gsbs_diss), [Neoplasms Commons](https://escholarship.umassmed.edu/gsbs_diss), [Nucleic Acids, Nucleotides, and Nucleosides Commons](https://escholarship.umassmed.edu/gsbs_diss), and the [Respiratory Tract Diseases Commons](https://escholarship.umassmed.edu/gsbs_diss)

**Repository Citation**


This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in GSBS Dissertations and Theses by an authorized administrator of eScholarship@UMMS. For more information, please contact [Lisa.Palmer@umassmed.edu](mailto:Lisa.Palmer@umassmed.edu).
GENETIC APPROACHES TO STUDY
TRANSCRIPTIONAL ACTIVATION AND TUMOR SUPPRESSION

A Dissertation Presented

By

LING LIN

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
In partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

MAY 1, 2012
INTERDISCIPLINARY GRADUATE PROGRAM
GENETIC APPROACHES TO STUDY

TRANSCRIPTIONAL ACTIVATION AND TUMOR SUPPRESSION

A Dissertation Presented By

LING LIN

The signatures of the Dissertation Defense Committee signify completion and approval as to style and content of the Dissertation

Michael Green, Thesis Advisor

Paul Kaufman, Member of Committee

Kirsten Hagstrom, Member of Committee

Sharon Cantor, Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

Scot Wolfe, Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school

Anthony Carruthers, Ph.D.,
Dean of the Graduate School of Biomedical Sciences Program
May 1, 2012
DEDICATION

I would like to dedicate this work to my parents Jingfang Yao and Feng Lin. Their support and encouragement of my pursuits in science helped me conquer all the difficulties in the past eight years. I wish to share my joy of success with them.
ACKNOWLEDGEMENTS

First of all, I would like to thank Dr. Michael Green for providing me the opportunity to study the cutting edge of current science. His intelligence and expertise always inspire me and guide me to improve myself, scientific and personal. He forgave the mistakes that I made and has been patient with me even when experiments are not working. My graduate studies would not be completed without his guidance and support.

I would also like to thank Lynn Chamberlain. Lynn helped me with many laborious experiments and helped me get organized. Her companion for the last eight years is a great support.

I would also thank Green lab members, who provided great suggestions and discussions in many experiments.

Lastly, I would thank my Advisory Committee for their support and advice during my graduate studies.
ABSTRACT

The development of methods and techniques is the driving force of scientific research. In this work, we described two large-scale screens in studying transcriptional activation and tumor suppression.

In Part I, we studied transcriptional activation mechanisms by deriving and characterizing activation defective mutants. Promoter-specific transcriptional activators stimulate transcription through direct interactions with one or more components of the transcription machinery, termed the “target”. The identification of direct in vivo targets of activators has been a major challenge. We perform a large-scale genetic screen to derive and characterize tra1 alleles that are selectively defective for interaction with Gal4 in vivo. Utilizing these mutants, we demonstrated that Tra is an essential target for Gal4 activation, Gal4 and Tra1 bind cooperatively at the promoter and the Gal4–Tra1 interaction occurs predominantly on the promoter. In addition, we demonstrated that the Gal4-interaction site on Tra1 is highly selective.

In Part II, we described a functional genomics approach to discover new tumor suppressor genes. A goal of contemporary cancer research is to identify the genes responsible for neoplastic transformation. Cells that are immortalized but non-tumorigenic were stably transduced with pools of short hairpin RNAs
(shRNAs) and tested for their ability to form tumors in mice. ShRNAs in any resulting tumors were identified by sequencing to reveal candidate TSGs, which were then validated both experimentally and clinically by analysis of human tumor samples. Using this approach, we identified and validated 33 candidate TSGs. We found that most candidate TSGs were down-regulated in >70% of human lung squamous cell carcinoma (hLSCC) samples, and 17 candidate TSGs negatively regulate FGFR signalling pathway, and their ectopic expression inhibited growth of hLSCC xenografts. Furthermore, we suggest that by examining at the expression level of TSGs in lung cancer patients, we can predict their drug responsiveness to FGFR inhibitors. In conclusion, we have identified many new lung squamous cell cancer TSGs, using an experimental strategy that can be broadly applied to find TSGs in other tumor types.
# TABLE OF CONTENTS

## PART I

Studying Transcriptional Activation Mechanisms by Deriving and Characterizing Activation Defective Mutants

### CHAPTER I

1.1 General Introduction

1.1.1 Transcription

1.1.2 Pol II Transcription

   - General transcription factors (GTF)
   - PIC formation
   - PIC activation
   - Recycle of transcription factors

1.1.3 Transcription regulation

   - Trans-regulating factors
     - Repressors and mechanisms
     - Activators and mechanisms
   - Cis-regulating factors

   - Chromatin structure and functions

   - Chromatin modification
     - Classes and functions
     - Acetylation and transcription

1.1.4 Tra1 in transcriptional regulation

   - Tra1 functions
   - Tra1 structure

1.1.5 Gal4-directed transcriptional regulation

   - Gal4 structure
   - Gal4 functions
   - Gal4 regulation
### CHAPTER II

1.2 Analysis of Gal4-directed transcription activation using Tra1 mutants selectively defective for interaction with Gal4

- **Abstract**
- **Introduction**
- **Results**
- **Materials and Methods**
- **Discussion**
- **Chapter II Figures**
- **Chapter II Tables**

### PART II

Identification of New Tumor Suppressor Genes Through a Large-Scale shRNA Screen for Mouse Tumorigenesis

### CHAPTER I

2.1 General Introduction

- **2.1.1 Cancer**
- **2.1.2 Oncogenes**
  - History of oncogene discovery
  - The role of oncogene in cancer development
  - Activation of oncogenes
  - Classification of oncogenes
  - Receptor tyrosine kinases
- **2.1.3 FGFR1**
  - Structure
LIST OF TABLES

Table 1.1.1  Histone acetylation enzymes and complexes
Table 1.2.1  Microarray data comparing gene expression in a tra1-2ts strain relative to a wild-type TRA1 strain
Table 1.2.2  List of the top 20 genes most affected by Tra1 inactivation
Table 1.2.3  Microarray data comparing gene expression in a tra1-mut1 strain relative to a wild-type TRA1 strain
Table 1.2.4  Microarray data comparing gene expression in a tra1-mut8 strain relative to a wild-type TRA1 strain
Table 1.2.5  List of yeast strains used in this study
Table 1.2.6  Primer sequences for qRT-PCR, ChIP and yeast strain construction
Table 2.1.1  Genetic alterations in FGFR1 and related cancer
Table 2.1.2  Altered expression of FGFR1 in cancer
Table 2.1.3  Current status of FGFR-targeting therapies
Table 2.2.1  List of 41 shRNAs recovered from tumors
Table 2.2.2  List of 32 candidate TSGs that validate in colony formation and mouse tumorigenesis assays
Table 2.2.3  Summary of NCBI SKY/MFISH & CGH data for candidate TSGs
Table 2.2.4  Summary of Oncomine data analysis querying whether candidate TSGs are down-regulated in cancer versus normal tissue

Table 2.2.5  List of catalog numbers for shRNAs obtained from Open Biosystems

Table 2.2.6  List of primers used for quantitative real-time RT-PCR and for cloning TSGs

Table 2.2.7  Source and catalog numbers for human cDNAs used to clone TSGs
LIST OF FIGURES

Figure 1.1.1  Formation of the PIC on the promoter occurs in a series of ordered steps

Figure 1.1.2  SAGA is a modular complex with five distinct domains

Figure 1.1.3  Structure of Tra1 and Gal4

Figure 1.1.4  Gal4 regulation

Figure 1.2.1  Isolation of tra1 mutants that cannot support growth on galactose

Figure 1.2.2  Analysis of tra1 mutants that cannot support growth on galactose

Figure 1.2.3  Development of a bimolecular fluorescence complementation assay for detecting interactions between activators and Tra1 in vivo

Figure 1.2.4  Identification of Tra1 mutants that are unable to interact with Gal4

Figure 1.2.5  Identification of Tra1 mutants that are unable to interact with HATs complexes

Figure 1.2.6  Gal4 and Tra1 bind cooperatively to the GAL1 promoter

Figure 1.2.7  Schematic diagram of the on/off DNA strategy

Figure 1.2.8  Gal4 and Tra1 interaction occurs on the promoter

Figure 1.2.9  Gcn4 functionally interacts with the Tra1 GID mutants

Figure 1.2.10  The Gal4-interaction site on Tra1 is highly selective

Figure 1.2.11  Tra1 GID mutants can interact with various activators

Figure 1.2.12  Microarray data shows that Tra1 GID mutants are Gal4 specific
Figure 2.1.1  Transformation is a multistep process
Figure 2.1.2  Schematic structure of FGFR1
Figure 2.1.3  Strategy of Pooled RNAi Library Screen
Figure 2.2.1  large-scale shRNA screen for the identification of candidate TSGs
Figure 2.2.2  shRNA Knocking Down Efficiency
Figure 2.2.3  Down-regulation of candidate TSGs in hLSCC samples
Figure 2.2.4  Down-regulation of candidate TSGs in human lung adenocarcinoma samples
Figure 2.2.5  Knocking down TSGs in SA cell line activates FGFR signaling pathway
Figure 2.2.6  Model of TSGs negatively regulating FGFR1 signaling
Figure 2.2.7  Gene repressed by FGFR1 signaling at transcriptional level
Figure 2.2.8  Knocking Down TSGs Induces transformation of lung epithelia cells
Figure 2.2.9  SA cells with TSGs knocking down are sensitive to FGFR inhibitor
Figure 2.2.10 Candidate TSG promoter hypermethylation in hLSCC samples.
Figure 2.2.11 Over-expression of candidate TSGs inhibits growth of hLSCC cells
Figure 2.2.12 Mechanisms of candidate TSGs inhibiting growth of hLSCC cells
PREFACE

The work contained within this thesis is represented in the following publications:


**Lin L, Wajapeyee N, Chamberlain L, Zhu LJ, Green MR**

Identification of Squamous Cell Lung Cancer Tumor Suppressor Genes that Negatively Regulate FGFR1 Signaling. (manuscript in preparation).
PART I

STUDYING TRANSCRIPTIONAL ACTIVATION MECHANISMS BY DERIVING AND CHARACTERIZING ACTIVATION DEFECTIVE MUTANTS
CHAPTER I

GENERAL INTRODUCTION
1.1.1 Transcription

Transcription is a process of copying DNA into a sequence complementary RNA molecule. Messenger RNA (mRNA) is then tailored (splicing) and translated into a protein, which has cellular functions. Non-coding RNAs function in translation or gene regulation. Transcription is the first step of gene expression.

The regulation of transcription directly influences cell viability, fitness, response to environment, etc. During the transcription process, DNA sequence is read by RNA polymerases. Eukaryotes have several types of RNA polymerases. RNA polymerase I synthesizes ribosomal RNA (rRNAs), which are the RNA components of ribosome. RNA polymerase II (Pol II) synthesizes mRNAs and most small nuclear RNAs (snRNAs) and microRNAs. RNA polymerase III synthesizes transfer RNAs (tRNAs), rRNA 5S and other small RNAs. RNA polymerase IV and V are only found in plants. RNA Pol II is the most studied type of RNA polymerase, because it transcribes all the coding genes.

1.1.2 Pol II Transcription

In eukaryotes, transcriptional regulation mechanisms are conserved from single cell to multiple cell organisms. The discovery of RNA polymerase in 1959 by Weiss and Gladstone launched the field of transcription studies (Weiss and Gladstone, 1959). Ten years later, the purification of three eukaryotic RNA
polymerases by Roeder and Rutter provided the foundation for the study of gene expression regulation (Roeder and Rutter, 1969). These purified RNA polymerases can initiate transcription at selected promoters \textit{in vitro} with the help of basal or general transcription factors (GTFs). \textit{In vivo}, successful promoter specific transcription requires additional apparatuses such as activators, chromatin-modification factors, elongation factors, \textit{etc}. By sequentially recruiting these factors, eukaryotic transcription works in multi-steps: pre-initiation complex (PIC) assembly, PIC activation, initiation, promoter clearance, elongation and termination.

**General transcription factors (GTF)**

“Basal” transcription, which only occurs \textit{in vitro}, requires PIC, which consists of RNA Pol II and GTFs. Core RNA Pol II has 12 subunits (Rpb1-Rpb12). They are structurally and functionally conserved from yeast to human. Pol II alone cannot recognize specific promoters without additional factors. “Activated” transcription, which occurs \textit{in vivo}, requires additional regulatory molecules. GTFs include TFIIA, TFIIIB, TFIIID, TFIIE, TFIIIF, and TFIIH, named in the order of discovery by chromatographic elution (Matsui et al 1980).

**PIC formation**

Formation of PIC on promoter occurs in a series of ordered steps (Figure 1.1.1): (1) TFIIID recognizes of a promoter and forms the nucleus of transcription initiation complex; (2) TFIIIB recognizes the TFIIID-promoter complex and forms
DNA binding (DB) complex; (3) TFIIF/pol II complex is recruited to the promoter; (4) recruitment of TFIIE and TFIIH completes the PIC formation.

**PIC activation**

Immediately following the PIC formation is PIC activation (DNA melting), which form an “open” initiation complex. Unlike other steps, DNA melting requires energy. The helicase activity of ERCC3 (excision repair cross-complementing rodent repair deficiency, complementation group 3) in TFIIH facilitates DNA melting of a 10bp region upstream of the start site. TFIIE assists melting by stabilizing the single-stranded region near the start site. During this process, carboxy-terminal domain (CTD) of polymerase II is phosphorylated, which allows promoter clearance and elongation for synthesizing longer transcripts.

**Recycle of transcription factors**

PIC components will be released from Pol II in the order of TFIIB, TFIIE and TFIIH, and recycled for the next round of transcription initiation. After transcription termination, Pol II CTD will be dephosphorylated, and recycled for the PIC assembly.

**1.1.3 Transcriptional regulation**

Transcriptional regulation is conducted at different levels: cis-regulators, which are DNA elements and trans-regulators, such as transcription factors, chromatin structures, etc.
Trans-regulating factors

Repressors and mechanisms. The isolation and mechanism study of $lac$ repressor by Gilbert and Ptashne proved the visionary hypothesis of repressor proposed by Jacob and Monod in 1961 and evidenced the existence of trans-acting factors that could regulate transcription by binding to the upstream DNA sequences (Jacob 1961, Gilbert 1966, Ptashne 1967). Transcription repressors can be divided into two classes: general repressors and gene specific repressors. General repressors function through interacting with the TATA-binding protein (TBP) or preventing PIC assembly. Gene specific repressors function via: (1) binding to activators, for example, Gal80 represses Gal4 function by binding to its activation domain, (2) competing for activator binding sites, (3) interacting with transcription apparatuses, and (4) modifying chromatin by recruiting histone deacetylases (HDAC) to specific locations.

Activators and mechanisms. The existence of activators was not accepted until 1970 when Beckwith and Ptashne discovered and characterized an activator of the $lac$ operon (Zubay et al, 1970; Emmer et al, 1970). A single activator usually can activate multiple genes and a single gene can be regulated by multiple activators as well. Therefore, activators regulate gene expression in a coordinated and combinatorial fashion. Activators stimulate transcription by several mechanisms: (1) recruiting chromatin modifying complexes to the promoters, (2) recruiting transcription initiation apparatus to the promoters, (3) influencing the activity of the transcription apparatus.
Transcription activators have several broad classes based on their activation domains: proline-rich, glutamine-rich, and acidic activators. Mammalian proline-rich or glutamine-rich activator domains have no function in yeast (Ponticelli et al, 1995; Kunzler et al, 1994). Acidic activators function in all eukaryotes from yeast to human. Yeast acidic activators are modular proteins with two functionally and physically distinct domains. One is specific DNA-binding domain that reads promoter sequence. One domain is a short acidic region that can recognize transcription co-factors and the interaction with co-factor is independent on the structural complementarity (Sadowski et al, 1988).

Cis-regulating factors

The concept of promoter was first proposed by Ippen in 1968, and in 1971 Eron and Block confirmed it by *in vitro* transcription experiments (Ippen 1968, Eron 1971). The first detailed transcription factor binding sequence was described by Tjian in 1978 using SV40 DNA. For most protein-coding genes, their promoters consist of a core promoter and transcriptional regulator binding sequences. The core promoter includes the transcription start site and TATA box. Transcriptional regulator binding sequences include upstream activating sequences (UASs), enhancers, upstream repressing sequences (URSs), and silencers. UASs usually refer to the activator binding elements that are close to the transcription start site. Enhancers refer to the clusters of DNA binding sequences for transcriptional regulators distant from the start site, and it is independent of orientation.
Chromatin structure and function

Throughout the 1960s, DNA was thought to be “naked” in a linear form, until 1974, Roger Kornberg presented the “nucleosome hypothesis”, in which eukaryotic DNA is packaged around histones to form nucleosome arrays (Kornberg et al, 1974). By packaging DNA into chromatin, DNA can fit into the limiting confines of nucleus and be strengthened for mitosis and meiosis. In addition, the chromatin structure can prevent DNA damage, control gene expression and DNA replication. Chromatin consists of repeating nucleoprotein units named nucleosomes and linker DNA. A nucleosome contains three parts: (1) nucleosome core, which is 146bp of DNA wrapped in 1.67 left-handed superhelical turns around the histone octamer, (2) the linker DNA connects adjacent histone octamers, (3) linker histone, H1, that binds the linker DNA and nucleosome core, and forms the exit/entry of the DNA strand on the nucleosome. Transcription inhibition caused by packaging promoters in nucleosomes was found in vitro by Kornberg and colleagues in 1987, and was not accepted until it is proven in vivo by Han and Grunstein in 1988 (Knezetic and Iuse, 1986; Lorch et al., 1987; Han and Grunstein, 1988). Histone-DNA contact and higher order of chromatin structures can restrict access of transcription factors to the promoter elements. Disruption of nucleosome structure facilitates activator binding and transcription apparatus recruitment. However, in some cases, activators bind to
specific promoters in the distinct physical context provided by nucleosome (Chavez 1997, Cirillo 1999, Schild 1993).

**Chromatin modification**

**Classes and functions.** Higher order of chromatin structures, usually mediated by modification of N-terminal tails of histones, are correlated with transcription regulation (Durrin et al, 1991). Histones can be modified by acetylation, phosphorylation, methylation, ubiquitination and sumoylation. Histone acetylation is the best understood histone modification. Chromatin with histone hyperacetylation is more accessible to transcription apparatus and therefore is associated with active transcription. Histone acetylation activates transcription through several mechanisms: (1) providing greater access to DNA sequence for the transcription apparatus and its regulators by disrupting higher order chromatin structures, (2) decreasing their affinity for DNA or neighboring nucleosomes by disrupting nucleosome structure, (3) promoting or suppressing interactions with specific transcription factors. Histone phosphorylation and ubiquitination are correlated with increased transcriptional activity. Whether Histone methylation activates or represses transcription depends on the types of methylation of H3, H4 and H2.

**Acetylation and transcription.** Although RNA synthesis can be affected by histone acetylation was discovered early in 1964 by Allfrey et al, the link between transcription and acetylation wasn’t established until Brownell and Schreiber found that a histone acetyltransferase (HAT) and a histone de-
acetyltransferase (HDAC) they identified are homologs to yeast transcription factors Gcn5 and Rpd3 respectively (Allfrey et al, 1963; Brownell et al, 1996; Schreiber et al 1996). Since then, a large number of HATs have been discovered and are found to be previously identified transcriptional coactivators (Table 1.1.1). HATs can be divided into two classes based on their cellular locations: type A HATs are localized in nuclei and acetylate nuclear factors; type B HATs are localized in cytoplasm and acetylate newly synthesized histones in the process of histone assembly. In vivo, HATs function in association with multisubunit complexes. Each HAT may show different substrate specificity in different complexes. For example, Gcn5 is the catalytic component in SAGA, ADA and SLIK (SAGA like)/SALSA (SAGA altered) complexes (Grant et al 1997, Sterner et al 2002). SAGA (Spt-Ada-Gcn5 Acetyltransferase complex) is a 1.8-MDa complex, comprises of four classes of proteins: (1) Ada proteins (Ada1, Ada2, Ada3, Gcn5 and Ada5); (2) TBP-related Spt proteins (Spt3, Spt7, Spt8 and Spt20); (3) TBP-associated factors (TAF5, TAF6, TAF9, TAF10 and TAF12); and (4) Tra1. SAGA complex has a variant, SALSA (or SLIK), which functions in retrograde response pathway with a truncated form of Spt7 and lacking Spt8 (Grant 2002). The structure of SAGA complex was revealed by Schultz lab using immuno-electron microscopy method (Wu et al, 2004). They showed that SAGA is a modular complex with five distinct domains (Figure 1.1.2). Domain II&IV contain TAFs that may provide histone folding functions. Domain III is a central architectural domain with HAT activity. Domain V may represent TBP interacting
surface. Domain I contains a 400kDa protein, Tra1, representing the activator interacting surface.

1.1.4 Tra1 in transcriptional regulation

Tra1 functions

Tra1 is the yeast homolog of human TRRAP (Transformation/Transcription domain-Associated Protein). TRRAP was originally discovered by McMahon and Cole in 1998 (McMahon et al, 1998), who found TRRAP bound to c-Myc and E2F and mediated transformation. Yeast homolog Tra1 is a 400kDa protein and belongs to the phosphatidylinositol 3-kinase (PI3K)-related protein kinase (PIKK) protein family, which regulates signaling pathways by serine/threonine phosphorylation (Figure 1.1.3). Tra1 is an essential gene for yeast viability and a scaffold protein in SAGA, ADA and NuA4 acetyltransferase complexes (Grant et al, 1998; Saleh et al, 1998). Tra1 regulates many cellular functions, such as cell cycle progression, cell dividing, DNA repair, etc. Given its gigantic size and essential role in cell viability, it is challenging to study the structure and the functions of Tra1.

Tra1 structure

Although Tra1 retains the structure of the PI3K domain, it does not have the catalytic function, because it lacks several critical residues essential for protein phosphorylation (Knutson and Hahn, 2010). Mutiu and Brandl using targeted mutagenesis identified functional residues that are critical to cell
viability, temperature sensitive growth and growth on 6% ethanol (Mutiu et al 2007). Knutson and Hahn made 42 tra1 mutants with ~100 aa deletions and identified the regions that are critical for cell viability, association with SAGA and NuA4 complexes (Δ24), HAT specificity and activity (Δ32), activator interaction (e.g. Δ2, Δ5) (Knutson, 2011).

1.1.5 Gal4-directed transcriptional regulation

Gal4 Structure

Gal4 is one of the best investigated activators of yeast transcription. Gal4 is an 881aa protein member of the zn(II)2Cys6 (zinc cluster) family proteins (Figure 1.1.3). It has a Zn-Cys DNA binding domain, a linker domain, a dimerization domain and two acidic activation domains (Lohr, 1995, FASEB). The consensus DNA binding site for Gal4 is a 17mer sequence 5'-CGG-N11-CGG-3'. Using X-ray crystallography, Marmorstein and Harrison showed that N'terminal fragment (DNA binding domain) binds to the consensus UAS as a dimer (Marmorstein, 1992). Reece and Ptashne further identified that a 19aa region at the C-terminal to the Zn-Cys cluster directs the specificity of Gal4 binding. The gene structure of Gal4 is illustrated in Figure 1.1.3.

Gal4 functions

Gal4 regulates expression of most genes in galactose utilization pathway when glucose or other carbon sources are not available. These genes function in transportation of galactose into the cell and its metabolism through the
glycolytic pathway. There are two classes of GAL genes that are required for the
growth of yeast on galactose: structural genes (GAL1, GAL10, GAL2 and GAL7)
and regulatory genes (GAL4, GAL80 and GAL3). Transcription induction of the
GAL structural genes by galactose is dependent on Gal4 binding to the UAS in
their promoters. The number of Gal4 binding sites and the space between the
triplets in the consensus sequence determine their affinity for Gal4 activator and
lead to differential activation (Lohr, 1995). Besides GAL genes, some genes that
are required for global adaptation to growth on galactose, such as MTH1, PCL10
and FUR4, are also regulated by Gal4 (Ren, 2000).

**Gal4 regulation**

Gal4 activating transcription is the earliest and most well studied model
system for investigating transcription regulation in yeast (Figure 1.1.4). In the
absence of galactose, Gal4 is inactivated by binding of Gal80 at its activation
domain, rendering incapable to interact with transcriptional co-factors (Wu 1996;
Carrozza, 2002). In the presence of galactose, the interaction between Gal3 and
Gal80 causes conformational change in the Gal80-Gal4 complex, exposes the
Gal4 activation domain to the transcription co-factors and leads to the activation
of transcription (Zenke, 1996; Leuther, 1992). This is confirmed by *in vivo*
protein-protein interaction assay, fluorescence resonance energy transfer
(FRET), showing that Gal4 interacted with Gal80 in the presence of galactose
(Bhaumik, 2004). Peng and Hopper believed that Gal3 interacted with Gal80 and
dissociated it from Gal4 by “shuttling” it to the cytoplasm (Peng and Hopper,
2000, 2002). However, our work showed that Gal80 was still associated with Gal4 in nucleus after galactose induction.

**Gal4 targets**

Gal4 activates transcription by recruiting transcription co-activators and general transcription factors to the UAS through its activation domain. The pivotal questions are how does it recruit other factors and what are the targets. There are many proteins that have been shown to interact with Gal4-activation domain: TBP (Melcher, 1995; Wu, 1996), TFIIB (Wu, 1996), Gal11 (Jeong, 2001), Srb10 (Ansari, 2002) and SWI/SNF (Yudkovsky, 1999) were identified interacting with Gal4 AD using *in vitro* binding assay with purified proteins; Srb4 was identified using affinity chromatography, label transfer affinity-photo-cross-linking and surface plasmon resonance (Koh, 1998); Sug1 (Gonzalez, 2002) and Sug2 (Chang et al, 2001) were identified to interact with Gal4 using *in vitro* binding assay and *in vivo* cytotrap assay; SAGA was identified as Gal4 AD target using label transfer affinity-photo-cross-linking assay (Brown, 2001) and FRET (Bhaumik, 2004).

Gal4 alone is sufficient to recruit SAGA to the UAS, suggesting that it’s a direct target of Gal4 activation domain (Bhaumik et al 2001, Bryant et al 2003, Larschan et al 2001, 2005). In addition, Bhaumik and Green showed that SAGA works as a scaffold for PIC assembly at the promoter instead of (or other than) HAT, because its catalytic subunit, Gcn5, is not required for SAGA recruitment and PIC formation (Bhaumik et al 2001). Furthermore, they showed that one of
the 22 subunits of SAGA complex, Tra1, is the direct target of Gal4 activation domain, and the interaction happens on the promoter (Bhaumik et al 2004).

In many of these target finding experiments, mutations in Gal4 activation domain were used to compromise its interaction with proposed targets, therefore confirm the specificity of the interaction. However, we think it’s not sufficient to prove the direct interaction(s) \textit{in vivo}. What has been lacking is a reciprocal mutation in a target that prevents interaction with Gal4 and does not disrupt interactions between the Gal4 AD and any of its other putative targets. In the next chapter, we described Tra1 mutants that are selectively defective for interaction with Gal4, and use them to study how Gal4 stimulates transcription \textit{in vivo} and the basis by which Tra1 is recognized by Gal4 and other activators.
Figure 1.1.1. Formation of the PIC on the promoter occurs in a series of ordered steps: (A) Activator binds to the promoter and recruits transcription co-activators. (B) TFIID recognizes of a promoter and forms nucleus of transcription initiation complex; TFIIB recognizes the TFIID-promoter complex and form DNA binding (DB) complex; (C) TFIIF/pol II complex is recruited to the promoter. Recruitment of TFIIE and TFIIH completes the PIC formation.

AD: activation domain; DBD: DNA binding domain; UAS: upstream activation sequence; HAT: Histone acetyltransferase; SAGA: spt-ada-gcn5-acetyltransferase.
Figure 1.1.2

- I: Tra1
- II: TAFs
- III: HATs
- IV: TAFs
- V

- Activator interacting
- Histone folding
- TBP interacting
- HAT function
- Activator interacting
Figure 1.1.2. SAGA is a modular complex with five distinct domains. Domain II&IV contain TAFs that may provide histone folding functions. Domain III is a central architectural domain with HAT activity. Domain V may represent TBP interacting surface. Domain I contains a 400kDa protein, Tra1, representing the activator interacting surface.
Figure 1.1.3

A

TRAIL

HEAT  FAT  FRB  PI3K  FATC

B

Gal4

DNA binding  Dimerization  Activation  Activation  Gal80 binding
Figure 1.1.3. (A) Structure of Tra1. Tra1 consists of a HEAT domain (Huntingtin, elongation factor 3, PR65/A, and TOR), a FAT domain (FRAP, ATM, and TRRAP), a FRB domain (FKBP12 rapamycin binding), a PI3K domain and a FATC domain (FAT C-terminal). (B) Gal4 has a DNA binding domain, a linker domain, a dimerization domain and two acidic activation domains.
Figure 1.1.4

Non-inducing condition

Galactose-inducing condition
Figure 1.1.4. In the absence of galactose, Gal4 is inactivated by binding of Gal80 at its activation domain, rendering incapable to interact with transcriptional co-factors. In the presence of galactose, the interaction between Gal3 and Gal80 causes conformational change in the Gal80-Gal4 complex, exposes the Gal4 activation domain to the transcription co-factors and leads to the activation of transcription.
### Table 1.1.1 Histone acetylation enzymes and complexes

<table>
<thead>
<tr>
<th>Acetylation</th>
<th>Subunit(s) and Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hat1</td>
<td>Subunit of the Hat1p-Hat2p HAT complex, acetylates nuclear and cytoplasmic histone H4</td>
</tr>
<tr>
<td>Gcn5</td>
<td>Subunit of the ADA and SAGA HAT complex, modifies histone H2B and H3</td>
</tr>
<tr>
<td>Hpa2</td>
<td>Acetylates histones H3 and H4, autoacetylation</td>
</tr>
<tr>
<td>Esa1</td>
<td>MYST family, essential for viability, acetylates histone H4</td>
</tr>
<tr>
<td>Sas3</td>
<td>MYST family, subunit of NuA3 complex, acetylates histone H3, transcriptional silencing</td>
</tr>
<tr>
<td>Elp3</td>
<td>Subunit of elongator complex, acetylates histone H3 and H4</td>
</tr>
<tr>
<td>TAFII145</td>
<td>Subunit of TAFIID</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Complex</th>
<th>Members</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAGA</td>
<td>Tra1, Ada2, Gcn5, Spt3, Spt7, Spt8, Spt20, Taf5, Taf6, Taf9, Taf10, Taf12, Chd1Sgf11, Sgf29, Sgf73, Ubp8, Ngg1, Hfi1, Sus1</td>
</tr>
<tr>
<td></td>
<td>Interacts with acidic activation domains and TBP. HAT activity stimulates transcription. Acetylates histone H3 and H2B</td>
</tr>
<tr>
<td>SLIK</td>
<td>Tra1, Ada2, Gcn5, Spt3, Spt7(truncated form), Spt20, Taf5, Taf6, Taf9, Taf10, Taf12, Chd1Sgf11, Sgf29, Sgf73, Ubp8, Ngg1, Hfi1, Sus1</td>
</tr>
<tr>
<td></td>
<td>HAT activity stimulates transcription. Acetylates histone H3 and H2B</td>
</tr>
<tr>
<td>NuA4</td>
<td>Tra1, Eaf1, Eaf3, Eaf5, Eaf6, Eaf7, Ynf2, Swc4, Esa1, Epl1, Yaf9</td>
</tr>
<tr>
<td></td>
<td>Interacts with acidic activation domains. HAT activity stimulates transcription in vitro. Acetylates histone H4 and H2A</td>
</tr>
<tr>
<td>NuA3</td>
<td>Eaf6, Taf14, Nto1, Sas3, Yng1</td>
</tr>
<tr>
<td></td>
<td>Acetylates histone H3</td>
</tr>
<tr>
<td>ADA</td>
<td>Ahc1, Ada2, Ahc2, Ngg1, Sgf29, Gcn5</td>
</tr>
<tr>
<td></td>
<td>Acetylates histone H3 and H2B</td>
</tr>
<tr>
<td>Elongator</td>
<td>Elp2, Elp3, Elp4, Iki1, Iki3</td>
</tr>
<tr>
<td></td>
<td>Transcription elongation. Acetylates histones H2A, H2B, H3 and H4</td>
</tr>
</tbody>
</table>
CHAPTER II

Analysis of Gal4-directed transcription activation

using Tra1 mutants selectively defective

for interaction with Gal4
Abstract

Promoter-specific transcriptional activators (activators) stimulate transcription through direct interactions with one or more components of the transcription machinery, termed the “target”. The identification of direct in vivo targets of activators has been a major challenge. Previous studies have provided evidence that the Tra1 subunit of the yeast SAGA (Spt-Ada-Gcn5-acetyltransferase) complex is the target of the yeast activator Gal4. However, several other general transcription factors, in particular the mediator complex, have also been implicated as Gal4 targets. Here, we performed a large-scale genetic screen to derive and characterize tra1 alleles that are selectively defective for interaction with Gal4 in vivo (Gal4 interaction defective (GID) mutants). In contrast to wild-type Tra1, Tra1 GID mutants are not recruited by Gal4 to the promoter and cannot support Gal4-directed transcription, demonstrating the essentiality of the Gal4–Tra1 interaction. In yeast strains expressing a Tra1 GID mutant, binding of Gal4 to the promoter is unexpectedly also diminished, indicating that Gal4 and Tra1 bind cooperatively. Consistent with cooperative binding, we demonstrate that the Gal4–Tra1 interaction occurs predominantly on the promoter and not off DNA. Finally, we show that although Tra1 is targeted by other activators, these interactions are unaffected by GID mutations, revealing an unanticipated specificity of the Gal4–Tra1 interaction.
**Introduction**

Transcription is a highly regulated process. The basal transcription requires minimal protein apparatus for accurate transcription initiation: General transcription factors (GTFs) and the subunits of RNA polymerase II. Transcription initiation by RNA polymerase II involves the assembly of general transcription factors on the core promoter to form a preinitiation complex (PIC). In the cells, genes packaged into chromatin are in general repressed. Therefore, assembly of the GTFs at the promoter must be triggered by activator proteins. Promoter-specific activator proteins (activators) can recruit GTFs to a sequence specific promoter to accelerate PIC formation and induce the expression of genes (Roeder, 1996; Orphanides, 1996; Ptashne, 1997; Lee, 2000). Activators can stimulate transcription at various steps: (1) remove repressors from promoter; (2) recruit GTFs and pol II to the promoter; (3) induce conformational change of PIC; (4) induce covalent modification of proteins in the PIC; and (5) stimulate promoter clearance and elongation. Activators can recruit factors that remodel nucleosomes and make promoter sequences more accessible for GTFs and pol II. H3 and H4 histones are reversibly acetylated. The acetylation of the lysine residues reduces the interaction between histone and DNA by neutralizing their positive charge. Activators that recruit histone acetyltransferase complex stimulate transcription by removing the histones at the promoter. Activators are modular proteins that contain a DNA-binding domain (DBD) and an activation domain (AD). Activator-mediated stimulation of PIC assembly is believed to result
from a direct interaction between the AD and one or more components of the transcription machinery, termed the “target”. The unambiguous identification of the direct \textit{in vivo} targets of activators has been a major challenge in the field.
Results

Isolation of $tra1$ Mutants that Cannot Support Growth on Galactose.

The strategy we used to derive Tra1 mutants that fail to interact with the Gal4 AD is summarized in Figure. 1.2.1.A and described below. We generated a library of random $tra1$ mutants by $in$ $vitro$ hydroxylamine mutagenesis of a low-copy plasmid expressing $TRA1$. The library was transformed into a haploid yeast $tra1$-$\Delta$ strain that was complemented by wild-type (WT) $TRA1$ expressed on a low-copy URA3-containing plasmid. Following eviction of the WT $TRA1$ plasmid on media containing 5-fluoroorotic acid (5-FOA), strains harboring the $tra1$ mutants were analyzed for growth on media containing glucose (YPD), but not galactose (YPG). This approach yielded 13 $tra1$ mutants that were unable to support growth on YPG (Figure. 1.2.1.C).

Sequence analysis revealed that the 13 $tra1$ mutants represented eight distinct alleles (Figure. 1.2.1.B). Mutants 2-7 were found to contain the same mutations and therefore, of these mutants, only $tra1$-$mut2$ was further analyzed. Immunoblot analysis showed that all of the Tra1 mutants were expressed at levels comparable to that of WT Tra1 (Figure 1.2.2.A).

To confirm that the $tra1$ mutants selectively abolished $GAL$ gene expression, we monitored expression of two $GAL$ genes, $GAL1$ and $GAL3$, by quantitative RT-PCR (qRT-PCR). In strains harboring the eight $tra1$ mutants, transcription of $GAL1$ and $GAL3$ in galactose-containing media was severely compromised relative to that observed in a WT $TRA1$ strain (Figure. 1.2.2.B).
contrast, expression of *RPS5* and *RPS0B*, which are SAGA-independent genes, were unaffected by the *tra1* mutants (Figure. 1.2.2.C).

**Development of a Bimolecular Fluorescence Complementation Assay for Detecting Interactions Between Activators and Tra1 in vivo.**

To detect direct interactions between the mutant Tra1 proteins and Gal4 *in vivo*, we developed a bimolecular fluorescence complementation (BiFC) assay, which has been used for the *in vivo* detection of a wide variety of protein–protein interactions (Akman, 2009; Sung, 2007). The BiFC assay is based on the formation of a fluorescent complex comprising two fragments of yellow fluorescent protein (YFP), which are brought together by association of two interacting proteins fused to the fragments (reviewed in Kerppola 2008).

To verify the feasibility of this approach, we first performed a series of control experiments that monitored the interaction between Tra1 and two SAGA-dependent activators, Gal4 and Gcn4. The experimental strategy for detecting activator–Tra1 interactions using the BiFC assay is shown in Figure. 1.2.3.A. We derived a pair of haploid yeast strains, one in which the endogenous Tra1 protein was tagged at the C-terminus with the N-terminal fragment of a YFP variant known as Venus (Nagai, 2002) (Tra1-VN), and a second strain of opposite mating type in which the endogenous activator was tagged at its C-terminus with the C-terminal Venus fragment (Gal4-VC or Gcn4-VC). The two haploid strains
were then mated, and the resulting diploid cells were analyzed by fluorescence microscopy for a nuclear BiFC signal.

As described above, because of Gal80-mediated inhibition, the Gal4–Tra1 interaction is expected to occur in the presence but not the absence of galactose. Figure 1.2.3.B (top panel) shows that a BiFC signal was detected in cells expressing Tra1-VN and Gal4-VC and grown in YPG but not YPD. Identical results were obtained when the Venus fragment was fused to Tra1 at its N-terminus (VN-Tra1) (Figure 1.2.3.B, bottom panel). Thus the BiFC assay detected the Gal4–Tra1 interaction, which, as expected, was galactose dependent.

A Gcn4–Tra1 interaction is expected to occur only under conditions of amino acid starvation. Figure 1.2.3.C shows that a nuclear BiFC signal was not detected in cells expressing Tra1-VN and Gcn4-VC grown in nutrient-rich media (YPD), but was readily detected in amino acid-starved cells grown on histidine (His)-lacking medium in the presence of the competitive inhibitor 3-aminotriazole (3-AT). Collectively, the results of Figure 1.2.3 demonstrate that the BiFC assay can be used to detect interactions between activators and Tra1.

Identification of Tra1 Mutants that are Unable to Interact with Gal4

We used the BiFC assay to analyze the interaction between Gal4 and the Tra1 mutants. In these and subsequent BiFC experiments described below, the endogenous activator was tagged at the C-terminus with VC in a haploid tra1-Δ
strain that was complemented by WT *TRA1* expressed on a URA3-containing plasmid. The strain was then transformed with a low-copy plasmid expressing a Tra1 mutant protein fused to VN at either the C-terminus (Tra1-mut1, -mut2, -mut8, -mut9, -mut12 and -mut13) or N-terminus (Tra1-mut10 and -mut11), and the BiFC signal was monitored following eviction of the WT *TRA1* plasmid. As expected, in cells expressing WT Tra1-VN, a nuclear BiFC signal could be detected upon growth on galactose (Figure. 1.2.4). By contrast, a BiFC signal was not detected in cells expressing any of the Tra1 mutants.

We have previously proposed that Tra1 must be incorporated into an intact SAGA complex for interaction with Gal4 *in vivo* (Bhaumik et al 2004). Therefore, one explanation for the inability of the Tra1 mutants to interact with Gal4 is a failure to be incorporated into SAGA. To address this issue, we analyzed the stable association between Tra1 mutants and the SAGA subunit Spt20 in a co-immunoprecipitation assay. Figure. 1.2.5.A shows that Tra1-mut1 and Tra1-mut8 co-immunoprecipitated with Spt20 at levels comparable to that of WT Tra1. By contrast, the other mutants were completely (Tra1-mut2, -mut9, -mut10, -mut12 and -mut13) or partially (Tra1-mut11) defective for interaction with Spt20. Tra1 is also a subunit of another HAT complex known as NuA4 (Allard et al 1999). Figure. 1.2.5.B shows that all Tra1 mutants co-immunoprecipitated with the NuA4 subunit Eaf1, indicating that they were all efficiently incorporated into the NuA4 complex. Collectively, these results indicate that Tra1-mut1 and Tra1-mut8 are incorporated into the SAGA complex but are unable to interact with
Gal4 and thus can be classified as Gal4 interaction defective (GID) mutants. Tra1-mut8 is of particular importance because it contains only a single amino acid substitution at position 400 (H400Y) (see Figure. 1.2.1.B). The inability of Tra1-mut1 and Tra1-mut8 to interact with Gal4 and support transcription of GAL genes conclusively establishes that the Gal4–Tra1 interaction is essential for Gal4-directed transcription activation.

**Gal4 and Tra1 Bind Cooperatively to the GAL1 Promoter**

The interaction between Gal4 and Tra1 is expected to result in recruitment of the SAGA complex to the promoters of GAL genes. Therefore, in a yeast strain expressing a Tra1 GID mutant, the SAGA complex should not be recruited to GAL genes. To confirm this prediction, we performed a series of chromatin immunoprecipitation (ChIP) assays. As expected, in galactose association of Tra1-mut1 and Tra1-mut8 with the GAL1 promoter was substantially reduced relative to WT Tra1 (Figure. 1.2.6.A). Likewise, association of Spt20 with the GAL1 promoter was comparably reduced in the two Tra1 GID mutant strains. By contrast, recruitment of Tra1-mut1 and Tra1-mut8 to the promoter of RPS0B, a NuA4-dependent gene, was comparable to that of WT Tra1.

Figure. 1.2.6.A also shows that in raffinose Gal4 binding was roughly equivalent in the WT Tra1 and Tra1 GID mutant strains. Unexpectedly, in galactose binding of Gal4 to the GAL1 promoter was substantially reduced in the Tra1 GID mutant strains. We interpret this result to indicate that the Gal4–Tra1
interaction enhances the ability of Gal4 to bind the promoter. Binding of Gal4 to the \textit{GAL1} promoter was also reduced in an \textit{spt20-\Delta} strain (Figure. 1.2.6.\textit{B}), which does not support the Gal4–Tra1 interaction (Bhaumik et al 2002). Collectively, the results of Figures. 1.2.6.\textit{A} and \textit{B} indicate that as a result of the Gal4–Tra1 interaction, Gal4 and Tra1/SAGA bind cooperatively to the promoter.

\textbf{The Gal4–Tra1 Interaction Occurs Predominantly on the Promoter and Not Off DNA.}

An important, unresolved question is whether \textit{in vivo} the interaction between an activator and its target occurs predominantly on the promoter or whether the activator–target interaction is sufficiently stable that it occurs off DNA. To address this issue, we used the BiFC assay to monitor the interaction between Tra1 and a LexA-Gal4 AD fusion-protein that can bind to DNA only in yeast strains engineered to contain LexA-binding sites. A summary of the experimental design is shown in Figure. 1.2.7.\textit{A} and discussed below. We first constructed a haploid \textit{tra1-\Delta} strain harboring plasmids expressing Tra1-VN and a LexA DBD-Gal4 AD fusion-protein tagged with the C-terminal Venus fragment [LexA(DBD)-Gal4(AD)-VC]. This strain was transformed with a series of constructs to derive three strains, the first of which (strain 1) expressed a high-copy plasmid containing four LexA-binding sites located upstream of a \textit{GAL1}-\textit{lacZ} reporter gene (West et al 1984). Strain 2 was identical to strain 1 except that the construct lacked upstream LexA-binding sites (West et al 1984). In strain 3
the \textit{GAL1} core promoter sequence, which is the site at which SAGA and other PIC components are recruited, was deleted from the \textit{GAL1-lacZ} reporter gene. Immunoblot analysis confirmed that the LexA(DBD)-Gal4(AD)-VC fusion-protein was expressed equivalently in the three strains (Figure. 1.2.7.B).

We analyzed the interaction between the Gal4 AD and Tra1 in these three strains using the BiFC assay. Figure. 1.2.8.A shows that a nuclear BiFC signal was detected in strain 1 but not in strain 2, indicating a requirement for LexA-binding sites. As expected, in strain 1, the nuclear BiFC signal was observed only in galactose. The nuclear BiFC signal was also absent from strain 3, which lacked the \textit{GAL1} core promoter. Collectively, these results indicate that the Gal4 AD–Tra1 interaction occurs predominantly on DNA, and is dependent upon both activator-binding sites and the core promoter. Previous report shows that Gal4 binding to the promoter is independent of Gal4-Tra1 interaction by chromatin immunoprecipitation that facilitated by formaldehyde crosslinking (Bauhmik, 2002). By using BiFC assay, we could monitor the real-time interaction between activator and Tra1 that not affected by crosslinking reagent.

As a control, we performed an analogous experiment to monitor the interaction between Gal4 and its negative regulator Gal80. Figure. 1.2.8.B shows that a nuclear BiFC signal was detected in all three yeast strains indicating, as expected, that the Gal4–Gal80 interaction is not dependent upon either activator binding sites or the core promoter. Notably, unlike the Gal4–Tra1 interaction, the Gal4–Gal80 interaction occurred in both raffinose and galactose, which is in
agreement with previous biochemical experiments showing that in galactose Gal80 remains physically associated with Gal4 at a second site (Sil et al 1999).

**The Gal4-Interaction Site on Tra1 is Highly Selective**

We next performed experiments to determine whether the Tra1 GID mutants were selectively defective for interaction with the Gal4 AD or were unable to interact with other activators that also targeted Tra1. Figure 1.2.9.A shows that both Tra1 GID mutant strains grew on His-lacking medium containing 3-AT, indicating that the mutants could support Gcn4-directed transcription. Furthermore, a nuclear BiFC signal could be detected in strains expressing Gcn4-VC and either WT Tra1 or a Tra1 GID mutant (Figure 1.2.9.B). Thus, unlike Gal4, Gcn4 functionally interacts with Tra1 GID mutants.

To further investigate the selectivity of the Gal4 interaction site on Tra1, we sought to identify other activators that interact with Tra1 in the SAGA complex and determine their sensitivity to the Tra1 GID mutations. Toward this end, we first identified Tra1-dependent genes by comparing the mRNA population of a WT TRA1 strain to that of a strain bearing a temperature-sensitive tra1 allele (tra1-2ts; ref. Kulesza et al 2002) under non-permissive conditions. Following inactivation of TRA1, ~3% of yeast genes were down-regulated greater than two-fold (Table 1.2.1), consistent with a previous study that analyzed other tra1 alleles (Mutiu et al 2007).
To identify a set of Tra1-dependent genes that were also SAGA-dependent, we analyzed the 20 genes most affected by \textit{TRA1} inactivation (see Table 1.2.2) for dependence on Spt20. We first mined a published expression profiling study for genes whose transcription is compromised in an \textit{spt20-\Delta} strain (Lee et al 2000) and then confirmed these in silico results by qRT-PCR. This combined analysis identified 11 genes whose transcription was compromised by inactivation of Tra1 or loss of Spt20 (Figure 1.2.10.A). In addition, ChIP analysis showed, as expected, that Tra1 was bound to the promoters of all 11 genes (Figure. 1.2.10.B). Significantly, the two Tra1 GID mutants did not significantly affect transcription of any of these 11 Tra1- and SAGA-dependent genes (Figure 1.2.10.A), suggesting that Tra1 GID mutants might be specific to Gal4 activation.

Next, we tested the ability of the activators that mediate expression of the 11 Tra1- and SAGA-dependent genes to interact with WT Tra1 and the Tra1 GID mutants. To identify activators involved in regulating expression of the Tra1- and SAGA-dependent genes, we searched published genome-wide ChIP-microarray (ChIP-chip) studies (Harbison, 2004; Iyer, 2001; Lee, 2002; Lieb, 2001; Ren, 2000). Using this approach, predicted activators could be identified for 10 of the 11 Tra1- and SAGA-dependent genes analyzed above (Table 1.2.2).

We selected five activators for further analysis: Cbf1, Fkh2, Mcm1, Reb1, and Zap1. Figure 1.2.11 shows a nuclear BiFC signal was detected in yeast strains expressing WT Tra1-VN and either Cbf1-VC, Fkh2-VC, Mcm1-VC, Reb1-VC or Zap1-VC. Thus, as predicted, all five of these activators directly interact
with Tra1. Notably, a nuclear BiFC signal was also observed with all five activators in strains expressing either Tra1 GID mutant. Thus, consistent with the transcription results of Figure. 1.2.10.A, these activators directly interact with Tra1, but this interaction is insensitive to the GID mutations.

Finally, in an independent approach, we attempted to identify genes whose transcription was affected by the Tra1 GID mutants by comparing genome-wide expression profiles of yeast strains harboring either WT TRA1, tra1-mut1 or tra1-mut8 grown in YPD. Remarkably, we found only two genes (GSC2 and HSP30) whose expression was affected more than 2-fold by the tra1-mut1 mutation and no genes that were affected more than 2-fold by the tra1-mut8 mutation (Figure 1.2.12.A and Tables 1.2.3 and 1.2.4). Moreover, even for GSC2 and HSP30, qRT-PCR did not confirm the difference in expression levels observed by microarray analysis and instead revealed that expression of these two genes was comparable in tra1-mut1, tra1-mut8, and the WT TRA1 strains (Figure 1.2.12.B). Collectively, these results indicate that remarkably few, and possibly no other, yeast activators target the same Tra1 region at which Gal4 interacts.
Material and Methods

Plasmid and strain construction

Plasmids pRS414-pDED1-myc-TRA1 and pRS416-pDED1-myc-TRA1 were constructed by cloning the myc-TRA1 fragment from myc-TRA1-YCplac11 (Saleh et al 1998) into pRS414 (CEN TRP1) or pRS416 (CEN URA3), respectively, using restriction enzyme Not I and Sal I.

Strain LLY154 is a haploid tra1-∆ strain derived from MDC1 (Kulesza et al 2002) in which the wild-type TRA1 plasmid was replaced with pRS416-pDED1-myc-TRA1. For the qRT-PCR experiments shown in Figure 1.2.2, RNA was prepared from LLY154 transformed with a pRS414-based wild-type or mutant TRA1 plasmid following eviction of pRS416-pDED1-myc-TRA1 on 5-FOA. For the experiment shown in Figure 1.2.10, gene expression was analyzed in LLY154 harboring plasmids expressing TRA1, tra1-mut1 or tra1-mut8, in a haploid tra1-2ts (MDC3) or isogenic wild-type strain (MDC1; (Kulesza et al 2002), or in a haploid spt20-∆ strain (YDA352) or isogenic wild-type strain (FY23; (Winston et al 1995).

To generate Tra1-VN- and VN-Tra1-tagged strains for the BiFC assay in Figure 1.2.3, the N-terminal fragment of Venus was PCR amplified from plasmid pFA6a-VN- His3MX6 and pFA6a-KanMX6-P_{CET1}-VN, respectively (Sung et al 2007), using primers (listed in Table 1.2.6) and transformed into haploid MATa strain BY4741 (Giaever et al 2002). To generate Gal4-VC and Gcn4-VC-tagged strains, the C-terminal fragment of Venus was PCR amplified from plasmid
pFA6a-VC-His3MX6 (Sung et al 2007) and transformed into haploid MATα strain BY4742 (Giaever et al 2002). Pair-wise haploid strain combinations were mated, and the BiFC signal in the resulting diploid cells was examined.

For monitoring interactions between activators and Tra1 mutants, activators were C-terminally tagged at their endogenous locus, as described above, in strain LLY197, a haploid tra1-Δ strain harboring pRS416-pDED1-myc-TRA1. The strains were then transformed with a plasmid expressing wild-type or mutant Tra1-VN (or VN-Tra1) [constructed by PCR amplifying the N-terminal fragment of Venus from plasmid pFA6a-VN-KanMX6 (Sung et al 2007) and inserting it into pRS414-based wild-type and mutant Tra1 plasmids], and the BiFC signal was monitored following eviction pRS416-pDED1-myc-TRA1 on 5-FOA.

To construct LexA(DBD)-Gal4(AD)-VC, the C-terminal Venus PCR fragment from pFA6a-VC-TRP1 (Sung et al 2007) was cloned into plasmid pSH17-4 (Wu et al 2000), which expresses a LexA(DBD)-Gal4(AD) fusion protein. The plasmid expressing LexA(DBD)-Gal4(AD)-VC was co-transformed into strain LLY210 (a haploid tra1-Δ strain harboring a plasmid expressing Tra1-VN) together with pSH18-34 (harboring 4 LexA operators upstream of the GAL1-lacZ reporter gene; (West et al 1984), a derivative of pSH18-34 in which the GAL1 promoter had been deleted; pSH18-34Δ2) or LR1Δ1 (harboring a GAL1-lacZ reporter gene but lacking the LexA operators; (West et al 1984). For Figure 1.2.8, Gal80 was C-terminally tagged at its endogenous locus by PCR amplifying
the N-terminal Venus fragment from pFA6a-VN-TRP1 (Sung et al 2007) and transforming it into W303, generating LLY326, which was then co-transformed with plasmids LexA(DBD)-Gal4(AD)-VC and pSH18-34, LR1Δ1 or pSH18-34Δ2. Strains were grown in galactose or raffinose, and cells were monitored for a BiFC signal.

For the immunoblot experiment of Figure 1.2.2, whole cell extracts were prepared from strain LLY154 transformed with a pRS414-based WT or mutant Tra1 plasmid following eviction of pRS416-pDED1-myc-TRA1 on 5-FOA. For the co-immunoprecipitation experiments of Figure 1.2.5, Spt20 or Eaf1 was C-terminally tagged at its endogenous locus with an HA epitope by PCR amplifying the HA tag from pFA6a-3HA-KanMX6 (Longtine et al 1998) and transforming it into strain LLY154. The strains were then transformed with pRS414-based WT or mutant Tra1 plasmids, and extracts were prepared following eviction of pRS416-pDED1-myc-TRA1.

ChIP was performed using strain LLY154 transformed with pRS414-based WT or mutant Tra1 plasmid or FY23 and YDA352 (Figure 1.2.6).

**TRA1 Mutagenesis Screen**

Plasmid pRS414-pDED1-myc-TRA1 was mutagenized by treatment with hydroxylamine solution [1M hydroxylamine (Sigma), 50 mM sodium pyrophosphate (pH 7.0), 100 mM NaCl, and 2 mM EDTA] at 75°C for 30 mins (Guthrie et al 2004). The mutagenized library was amplified in bacteria and then
transformed into a haploid tra1-Δ strain LLY154 (see Table 1.2.5). Cells were plated on –Trp 5-FOA media, and ~1200 5-FOA-resistant colonies were patched and replica plated onto YPD and YPG media containing 20 µg/ml antimycin A (Sigma). Colonies able to grow on YPD but not YPG were selected, and the plasmid was isolated and sequenced. Strains carrying tra1-mut1 and tra1-mut8 were also analyzed for growth on –His media containing 50 mM 3-AT.

Quantitative RT-PCR
Total RNA was extracted (Schmitt et al 1990), and reverse transcription was performed using SuperScript II Reverse Transcriptase (Invitrogen) followed by qPCR using Fast SYBR Green Master Mix (Applied Biosystems) using the primers listed in Table S6. For the experiments shown in Fig. 1B and S1D, cells were grown in 2% raffinose followed by 2% galactose for 5 mins.

Bimolecular Fluorescence Complementation Assay
The BiFC signal in cells was examined by fluorescence microscopy using a Zeiss AXIO Imager Z2 microscope. A total of 100 cells from at least 7 different fields were counted; representative examples are shown.

Immunoblotting and Co-immunoprecipitation Assays
For Figure 1.2.2, whole cell extracts were prepared as previously described (Brown et al 2001) and blots were probed with an anti-myc (Santa Cruz) or anti-
actin (Abcam) antibody. For Figure 1.2.7, extracts were prepared from strains grown in galactose or raffinose medium, and blots were probed with an anti-LexA (Santa Cruz) or anti-actin (Abcam) antibody. For the co-immunoprecipitation experiments of Figure 1.2.5, Spt20-HA or Eaf1-HA was immunoprecipitated with an anti-HA antibody, and blots were probed with an anti-HA (Abcam) or anti-myc (Santa Cruz) antibody.

**Chromatin Immunoprecipitation**

ChIP was performed as described previously (Harbison et al 2004) using an anti-myc (Abcam), anti-HA (Abcam) or anti-Gal4 (Abcam) antibody. Following reversal of the crosslinks, the DNA was PCR-amplified using gene-specific primers (listed in Table 1.2.6).

**Microarray Analyses**

Strains MDC1 and MDC3 were grown at 30°C and shifted to 37°C for 60, 90 and 120 mins. Haploid TRA1, tra1-mut1 and tra1-mut8 strains were generated as described in Supplemental Methods. RNA was extracted according to standard protocols (Schmitt et al 1990) and hybridized to an Affymetrix YG-S98 array. The tra1-mut1/TRA1 and tra1-mut8/TRA1 experiments were done in duplicate. Statistical analyses were performed using R (Ihaka et al 1996). RMA method (Irizarry et al 2003) in Affy package from Bioconductor (Gentleman et al 2004) was used to summarize the probe level data and normalize the dataset to
remove across array variation. Limma package (Smyth 2004) with randomized block design was used to determine whether a gene’s expression level differs between mutant and WT regardless of time point. Genes with adjusted p-value using B-H method (Benjamini 1995) < 0.05 was considered significant. The microarray data from this publication have been deposited in NCBI’s Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE31391.
Discussion

The results presented in this manuscript definitively establish Tra1 as an essential in vivo target of Gal4 by identifying the site on Tra1 at which Gal4 interacts and demonstrating the necessity of this interaction for Gal4-directed transcription. For Tra1 to interact with Gal4 it first must be incorporated into an intact SAGA complex. The amino acids in Tra1 that we find compromise assembly into the SAGA complex are consistent with the results of a recent study analyzing Tra1 functional domains (Knutson et al 2011). Our results indicate that Tra1 does not have an intrinsic ability to interact with the Gal4 AD but rather requires proper presentation within the SAGA complex. It seems likely that this finding is relevant to the selective interaction of Gal4 with SAGA and not the NuA4 complex, which also contains Tra1.

Several experimental observations have led to the suggestion that an activator will have multiple, functionally redundant targets. For example, in vitro protein–protein interaction experiments have shown that a single activator such as Gal4 (reeves, 2005; Wu, 1996; Melcher, 1995; Neely, 2002; Bryant, 2003; Jeong, 2001; Koh, 1998; Park, 2000) or Gcn4 (Drysdale, 1998; Utley, 1998) can interact with multiple components of the transcription machinery. Likewise, in artificial recruitment experiments, a wide variety of transcription components can stimulate transcription and thus could potentially function as targets (Ptashne,
However, in contrast to this view, we demonstrate that interaction of Gal4 with a single site on Tra1 is required for Gal4-directed transcription.

The fact that many yeast activators contain acidic ADs, with apparently similar sequence features, has suggested that activators have common targets and recognition sites. Consistent with this idea, in vitro protein crosslinking experiments have shown that both the Gal4 and Gcn4 ADs interact with three common proteins; Tra1, Gal11 and Taf12 (Fishburn, 2005; Reeves, 2005). Surprisingly, however, we find that the Gal4-interaction region on Tra1 is remarkably specific and that the Gal4 and Gcn4 ADs recognize Tra1 differentially. Our collective results suggest that at most very few, and likely no other, yeast activators functionally interact with the same region of Tra1 that is recognized by the Gal4 AD.

In this study, we have performed a series of BiFC experiments whose results show that the interaction between the Gal4 AD and Tra1 occurs predominantly on the promoter and not off DNA. These new results explain our previous finding that a Gal4 mutant lacking its DBD failed to interact with Tra1 in vivo (Bhaumik, 2004). Our results help explain how cellular activators avoid a transcription inhibitory process referred to as “squelching”, which occurs following over-expression of a strong AD, such as the herpes simplex virus VP16 AD (Gill, 1988; Triezenberg, 1988). Squelching results from the sequestration of the target by the activator off the promoter; the target is thus unavailable for promoter-bound activators resulting in transcriptional inhibition. By interacting with Tra1
predominantly on the promoter and not off DNA, Gal4 avoids squelching. Squelching is dependent upon both the strength and concentration of the over-expressed AD (Gill, 1988; Triezenberg, 1988). Based upon these considerations, we speculate that cellular activator-target interactions are in general weak, thus ensuring that they occur only on the promoter where they are stabilized by the many other protein–protein and protein–DNA interactions in the PIC.
Figure 1.2.1

A

```
   Mutagenize plasmid with hydroxylamine
   Amplify in E. coli
   Transform into haploid tra1-Δ yeast strain bearing TRA1-URA3 plasmid
   Plate on 5-FOA to evict TRA1-URA3 plasmid
   Plate onto YPD and YPG
   Select colonies able to grow on YPD but not YPG
   Isolate tra1 plasmid
   Sequence to identify mutation(s)
```

B

```
<table>
<thead>
<tr>
<th>tra1 mutant</th>
<th>Mutated residue(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mut1</td>
<td>D397N, S404F, D469N, V1115I</td>
</tr>
<tr>
<td>mut2-7</td>
<td>S1259N, A1299T, V1303I, T1430I</td>
</tr>
<tr>
<td>mut8</td>
<td>H400Y</td>
</tr>
<tr>
<td>mut9</td>
<td>A2471T</td>
</tr>
<tr>
<td>mut10</td>
<td>T1733I</td>
</tr>
<tr>
<td>mut11</td>
<td>A96T, G1465D, E2641K, S2645N, V2942K, A3328T</td>
</tr>
<tr>
<td>mut12</td>
<td>P2590L, E3638K</td>
</tr>
<tr>
<td>mut13</td>
<td>S208F, S627L, E653K, G3027D, P3730S</td>
</tr>
</tbody>
</table>
```

C

```
YPD + Antimycin | YPD + Antimycin | YPD + Antimycin | YPD + Antimycin

gal4-Δ  
TRA1 (WT)  
tra1-mut1  
tra1-mut2  
tra1-mut3  
tra1-mut4  
tra1-mut5  
tra1-mut6  
tra1-mut7
```
Figure 1.2.1. Isolation of tra1 mutants that cannot support growth on galactose. (A) Schematic for mutagenesis of TRA1 and selection of mutants that fail to grow on galactose. (B) Summary of the mutated residues in each of the tra1 mutants. (C) Growth of tra1 mutants 1-13 on YPD and YPG media supplemented with antimycin. Growth of WT TRA1 and gal4-Δ strains are shown as controls. Cells were spotted as 10-fold serial dilutions.
Figure 1.2.2

A

B

C
Figure 1.2.2. Analysis of tra1 mutants that cannot support growth on galactose. (A) Immunoblot analysis showing levels of WT and mutant Tra1 proteins. Actin (Act1) was monitored as a loading control. (B) qRT-PCR analysis monitoring expression of GAL1 and GAL3 in strains expressing WT TRA1 or tra1 mutant grown in raffinose or galactose. (C) qRT-PCR analysis monitoring expression of RPS5 and RPS0B in strains expressing WT TRA1 or each of the tra1 mutants grown in raffinose or galactose. Expression of each gene was normalized to that observed in the WT TRA1 strain grown in raffinose, which was set to 1. The fold induction in galactose in the WT TRA1 strain is indicated. Error bars indicate standard deviation.
Figure 1.2.3

A

No interaction

Interaction

B

Tra1-VN + Gal4-VC

DAPI

Merge

YPD

YPG

C

Tra1-VN + Gcn4-VC

DAPI

Merge

YPD

-His + 3AT

YPD

YPG
Figure 1.2.3. Development of a bimolecular fluorescence complementation assay for detecting interactions between activators and Tra1 in vivo. (A) Schematic diagram depicting the BiFC assay. Tra1 is tagged at the C-terminus with the N-terminal Venus fragment (VN), and the activator (Act) is tagged at the C-terminus with the C-terminal Venus fragment (VC). (B) BiFC assay monitoring interaction between Tra1 and Gal4 in vivo, as evidenced by intense YFP signal (arrowheads) in YPG. The Tra1–Gal4 interaction occurs in the nucleus, as evidenced by co-localization (arrowheads) with the DNA stain 4′,6-diamidino-2-phenylindole (DAPI). Tra1 was tagged at either the C-terminus (top panels) or N-terminus (bottom panels). (C) BiFC assay monitoring interaction between Tra1 and Gcn4 in vivo, as evidenced by intense YFP signal in response to amino acid starvation (–His + 3-AT media).
Figure 1.2.4. Identification of Tra1 mutants that are unable to interact with Gal4. BiFC assay monitoring the interaction between Gal4 and the mutant Tra1 proteins.
Figure 1.2.5. Identification of Tra1 mutants that are unable to interact with HATs complexes. (A) Co-immunoprecipitation assay. Spt20-HA was immunoprecipitated with an anti-HA antibody, and the immunoprecipitate analyzed for the presence of Tra1. The levels of Spt20 and Tra1 in the input extract are shown. (B) The Tra1 mutants are efficiently incorporated into the NuA4 complex. Eaf1-HA was immunoprecipitated with an anti-HA antibody, and the immunoprecipitate analyzed for the presence of Tra1. The levels of Eaf1 and Tra1 in the input extract are shown.
Figure 1.2.6

A

GAL1 promoter

Tra1 (galactose)

Spt20 (galactose)

GAL1 promoter

Gal4 (raffinose)

Gal4 (galactose)

RPS0B promoter

Tra1 (galactose)

B

GAL1 promoter

Gal4 (galactose)

SPT20

spt20Δ

58
Figure 1.2.6. Gal4 and Tra1 bind cooperatively to the \textit{GAL1} promoter. (A) ChIP assay monitoring recruitment of Tra1, Spt20 and Gal4 to the \textit{GAL1} and \textit{RPS0B} promoters in strains expressing WT \textit{TRA1}, \textit{tra1-mut1} or \textit{tra1-mut8} and grown in media containing galactose or raffinose. (B) ChIP assay monitoring recruitment of Gal4 to the \textit{GAL1} promoter in a WT \textit{SPT20} or \textit{spt20-Δ} strain. Error bars indicate standard deviation.
Figure 1.2.7

A

<table>
<thead>
<tr>
<th>Strain</th>
<th>LexA binding sites</th>
<th>GAL1 core</th>
<th>lacZ gene</th>
<th>Predicted BiFC signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strain 1</td>
<td></td>
<td></td>
<td></td>
<td>+++</td>
</tr>
<tr>
<td>Strain 2</td>
<td></td>
<td></td>
<td></td>
<td>+++</td>
</tr>
<tr>
<td>Strain 3</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

B

| Strain: | Tra1-VN | Gal80-VN |  |
|---------|---------|----------|-
|         | Galactose | Raffinose | Galactose | Raffinose |
| 1       | 2        | 3        | 1        | 2        | 3        |
| 2       | 1        | 2        | 1        | 2        | 3        |
| 3       | 1        | 2        | 1        | 2        | 3        |

LexA(DBD)-Gal4(AD)-VC

Act1
Figure 1.2.7. Schematic diagram of the on/off DNA strategy. (A) Schematic diagram of the BiFC-based strategy to detect whether the Gal4–Tra1 interaction occurs predominantly on or off DNA. (B) Immunoblot analysis confirming expression of the LexA(DBD)-Gal4(AD)-VC fusion protein in strains harboring either Tra1-VN or Gal80-VN grown in galactose or raffinose.
Figure 1.2.8. Gal4 and Tra1 interaction occurs on the promoter. (A) BiFC assay monitoring the interaction between Tra1-VN and LexA(DBD)-Gal4(AD)-VC in the three yeast strains grown in galactose or raffinose. (B) BiFC assay monitoring the interaction between Gal80-VN and LexA(DBD)-Gal4(AD)-VC in the three yeast strains grown in galactose or raffinose.
Figure 1.2.9

A

\[
\begin{array}{c|c}
\text{tra1-mut1} & \text{tra1-mut8} \\
\hline
\text{TRA1 (wt)} & \text{gcn4-\Delta} \\
\end{array}
\]

- His

- His + 3AT

B

<table>
<thead>
<tr>
<th>Tra1-VN + Gcn4-VC</th>
<th>DAPI</th>
<th>Merge</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3AT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+3AT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Tra1 (WT)        |      |       |
| -3AT             |      |       |
| +3AT             |      |       |

| Tra1-mut1        |      |       |
| -3AT             |      |       |
| +3AT             |      |       |

| Tra1-mut8        |      |       |
| -3AT             |      |       |
| +3AT             |      |       |
Figure 1.2.9. Gcn4 functionally interacts with the Tra1 GID mutants. (A) Growth of tra1-mut1 and tra1-mut8 on His-lacking media containing or lacking 3-AT. Growth of WT TRA1 and gcn4-Δ strains are shown as controls. (B) BiFC assay monitoring the interaction between Tra1-VN and Gcn4-VC in His-lacking media containing or lacking 3-AT.
Figure 1.2.10

A

qRT-PCR

Relative expression

B

ChIP

Fold enrichment

tra1-2ts

spt20-Δ

tra1-mut1

tra1-mut8
Figure 1.2.10. The Gal4-interaction site on Tra1 is highly selective. (A) qRT-PCR analysis monitoring expression of 11 Tra1- and SAGA-dependent genes in tra1-ts, spt20-Δ, tra1-mut1 and tra1-mut8 strains. Gene expression is presented relative to that observed in a WT strain, which was set to 1 (indicated by the red line). (B) ChIP analysis monitoring binding of Tra1 to the promoters of the 11 Tra1- and SAGA-dependent genes. Error bars indicate standard deviation.
Figure 1.2.11

<table>
<thead>
<tr>
<th></th>
<th>Tra1-VN + Act-VC</th>
<th>DAPI</th>
<th>Merge</th>
<th>Tra1-mut1-VN + Act-VC</th>
<th>DAPI</th>
<th>Merge</th>
<th>Tra1-mut8-VN + Act-VC</th>
<th>DAPI</th>
<th>Merge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cbf1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fkh2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mcm1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reb1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zap1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1.2.11. Tra1 GID mutants can interact with various activators. BiFC analysis monitoring the ability of WT Tra1, Tra1-mut1 or Tra1-mut8 to interact with various activators.
Figure 1.2.12

A

Gene expression in tra1-mut1 vs. Gene expression in TRA1 (WT)

HSP30
GSC2

Gene expression in tra1-mut8

Gene expression in TRA1 (WT)

B

Relative expression

TRA1 (WT)
tra1-mut1
tra1-mut8

GSC2
HSP30
Figure 1.2.12. Microarray data shows that Tra1 GID mutants are Gal4 specific. (A) Scatter plot analyses comparing gene expression in a WT TRA1 strain and tra1-mut1 strain (top) or tra1-mut8 strain (bottom). The red line represents no change in gene expression; the blue dotted line represents 2-fold down-regulation. Gray circles represent all the genes on the array; black circles represent genes with a p-value <0.05; red circles represent genes with a p-value<0.05 and down-regulated more than 2-fold. (B) qRT-PCR analysis monitoring expression of GSC2 and HSP30 in tra1-mut1 and tra1-mut8 strains relative to that observed in a WT TRA1 strain, which was set to 1. Error bars indicate standard deviation.
Table 1.2.1. Microarray data comparing gene expression in a *tra1-2ts* strain relative to a wild-type *TRA1* strain (p<0.05).

<table>
<thead>
<tr>
<th>Probe</th>
<th>Gene symbol</th>
<th>120 min</th>
<th>90 min</th>
<th>60 min</th>
</tr>
</thead>
<tbody>
<tr>
<td>10012_at</td>
<td>NA</td>
<td>-1.226372245</td>
<td>-1.207768354</td>
<td>-1.088297812</td>
</tr>
<tr>
<td>10154_at</td>
<td>IDP2</td>
<td>-1.062025648</td>
<td>-1.397346472</td>
<td>-2.097301878</td>
</tr>
<tr>
<td>10204_at</td>
<td>ZRT2</td>
<td>-1.693546497</td>
<td>-1.694102453</td>
<td>-1.9383917</td>
</tr>
<tr>
<td>10238_at</td>
<td>BUD20</td>
<td>-1.750530576</td>
<td>-1.795024945</td>
<td>-1.565246542</td>
</tr>
<tr>
<td>10268_at</td>
<td>REX2</td>
<td>-1.255752195</td>
<td>-1.164401017</td>
<td>-1.224383784</td>
</tr>
<tr>
<td>10282_at</td>
<td>NA</td>
<td>-1.531035433</td>
<td>-1.974016187</td>
<td>-1.59479357</td>
</tr>
<tr>
<td>10377_at</td>
<td>NA</td>
<td>1.440655369</td>
<td>1.653805439</td>
<td>1.459506989</td>
</tr>
<tr>
<td>10378_at</td>
<td>AQY2</td>
<td>1.301360348</td>
<td>1.601900901</td>
<td>1.258684398</td>
</tr>
<tr>
<td>10415_at</td>
<td>MMP1</td>
<td>-1.322554037</td>
<td>-1.608306622</td>
<td>-1.401658201</td>
</tr>
<tr>
<td>10478_at</td>
<td>PCK1</td>
<td>-1.293644718</td>
<td>-1.282967759</td>
<td>-1.252414073</td>
</tr>
<tr>
<td>10497_at</td>
<td>UBP11</td>
<td>-1.300677363</td>
<td>-1.27364203</td>
<td>-1.013231909</td>
</tr>
<tr>
<td>10497_at</td>
<td>SIS2</td>
<td>-1.252818834</td>
<td>-1.231091171</td>
<td>-1.300768796</td>
</tr>
<tr>
<td>10501_at</td>
<td>ECM4</td>
<td>-1.001039214</td>
<td>1.642771735</td>
<td>1.345510776</td>
</tr>
<tr>
<td>10601_at</td>
<td>SFK1</td>
<td>-1.05344171</td>
<td>-1.275306044</td>
<td>-1.171416271</td>
</tr>
<tr>
<td>10852_at</td>
<td>DAN1</td>
<td>1.758400415</td>
<td>2.591673792</td>
<td>1.547578953</td>
</tr>
<tr>
<td>10858_s_at</td>
<td>THI11</td>
<td>1.214394359</td>
<td>1.395405163</td>
<td>1.350608614</td>
</tr>
<tr>
<td>10954_at</td>
<td>LIA1</td>
<td>-1.375695121</td>
<td>-1.642771735</td>
<td>-1.345510776</td>
</tr>
<tr>
<td>11127_at</td>
<td>ALB1</td>
<td>-1.325608385</td>
<td>-1.34043061</td>
<td>-1.079525381</td>
</tr>
<tr>
<td>11133_at</td>
<td>NCA3</td>
<td>-2.016193799</td>
<td>-2.416885115</td>
<td>-2.163019521</td>
</tr>
<tr>
<td>11177_at</td>
<td>FMP33</td>
<td>-1.535237707</td>
<td>-1.29764351</td>
<td>-1.396321813</td>
</tr>
<tr>
<td>11226_at</td>
<td>ACO2</td>
<td>1.692816942</td>
<td>1.418572159</td>
<td>1.657043875</td>
</tr>
<tr>
<td>11228_at</td>
<td>PHO90</td>
<td>-1.132894771</td>
<td>-1.102049894</td>
<td>-1.165011219</td>
</tr>
<tr>
<td>11260_at</td>
<td>OPT1</td>
<td>-1.598875666</td>
<td>-1.628435295</td>
<td>-1.204435409</td>
</tr>
<tr>
<td>11262_s_at</td>
<td>PHO11</td>
<td>-1.49630264</td>
<td>-2.016522899</td>
<td>-2.104202627</td>
</tr>
<tr>
<td>11342_at</td>
<td>MAK16</td>
<td>-1.054861274</td>
<td>-1.432647125</td>
<td>-1.05656865</td>
</tr>
<tr>
<td>11370_at</td>
<td>CYC3</td>
<td>-1.369717708</td>
<td>-1.259428941</td>
<td>-1.08470426</td>
</tr>
<tr>
<td>11389_at</td>
<td>ECM1</td>
<td>-1.634960648</td>
<td>-1.714500625</td>
<td>-1.435056855</td>
</tr>
<tr>
<td>3189_i_at</td>
<td>COS8</td>
<td>1.673309659</td>
<td>2.025539479</td>
<td>1.884785592</td>
</tr>
<tr>
<td>3259_at</td>
<td>NA</td>
<td>1.316981342</td>
<td>1.612948588</td>
<td>1.406599971</td>
</tr>
<tr>
<td>3316_f_at</td>
<td>NA</td>
<td>1.184817697</td>
<td>1.347626442</td>
<td>1.392113641</td>
</tr>
<tr>
<td>3358_f_at</td>
<td>NA</td>
<td>1.250214504</td>
<td>1.045026222</td>
<td>1.015595932</td>
</tr>
<tr>
<td>3361_f_at</td>
<td>NA</td>
<td>1.415801349</td>
<td>1.677852533</td>
<td>1.188172238</td>
</tr>
<tr>
<td>3368_f_at</td>
<td>NA</td>
<td>1.192894452</td>
<td>1.131515</td>
<td>1.167811189</td>
</tr>
<tr>
<td>3371_f_at</td>
<td>NA</td>
<td>1.893949049</td>
<td>1.848916649</td>
<td>1.074171085</td>
</tr>
<tr>
<td>3372_f_at</td>
<td>NA</td>
<td>1.252031885</td>
<td>1.372837615</td>
<td>1.53997973</td>
</tr>
<tr>
<td>3418_f_at</td>
<td>NA</td>
<td>1.650029693</td>
<td>1.783416026</td>
<td>1.129191255</td>
</tr>
<tr>
<td>3438_f_at</td>
<td>NA</td>
<td>1.153403486</td>
<td>1.474104928</td>
<td>1.406412669</td>
</tr>
<tr>
<td>3476_f_at</td>
<td>NA</td>
<td>1.219314979</td>
<td>1.770233827</td>
<td>1.237017432</td>
</tr>
<tr>
<td>3479_f_at</td>
<td>NA</td>
<td>1.13734296</td>
<td>1.160289726</td>
<td>1.200642625</td>
</tr>
<tr>
<td>3482_f_at</td>
<td>NA</td>
<td>1.697940527</td>
<td>1.857024486</td>
<td>1.246425058</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td>Row_1_VALUE</td>
<td>Row_2_VALUE</td>
<td>Row_3_VALUE</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>3491_f_at</td>
<td>NA</td>
<td>2.01701709</td>
<td>1.80687404</td>
<td>1.283049356</td>
</tr>
<tr>
<td>3508_f_at</td>
<td>NA</td>
<td>1.199963494</td>
<td>1.10408395</td>
<td>1.007203658</td>
</tr>
<tr>
<td>3514_f_at</td>
<td>NA</td>
<td>1.4641883</td>
<td>1.789969773</td>
<td>1.260490591</td>
</tr>
<tr>
<td>3548_s_at</td>
<td>NA</td>
<td>1.649591911</td>
<td>1.831869171</td>
<td>1.628764795</td>
</tr>
<tr>
<td>3563_s_at</td>
<td>NA</td>
<td>1.404690762</td>
<td>1.35792858</td>
<td>1.270205922</td>
</tr>
<tr>
<td>3594_f_at</td>
<td>NA</td>
<td>1.687428704</td>
<td>1.691750573</td>
<td>1.144042998</td>
</tr>
<tr>
<td>3598_f_at</td>
<td>NA</td>
<td>1.302151446</td>
<td>1.464566019</td>
<td>1.327377947</td>
</tr>
<tr>
<td>3645_f_at</td>
<td>NA</td>
<td>1.058435983</td>
<td>1.122632103</td>
<td>1.085881394</td>
</tr>
<tr>
<td>3687_f_at</td>
<td>NA</td>
<td>1.195607967</td>
<td>1.082197877</td>
<td>1.089775105</td>
</tr>
<tr>
<td>3695_f_at</td>
<td>NA</td>
<td>1.123799689</td>
<td>1.050095761</td>
<td>1.08656152</td>
</tr>
<tr>
<td>3707_f_at</td>
<td>NA</td>
<td>1.196868317</td>
<td>1.84150666</td>
<td>1.214973265</td>
</tr>
<tr>
<td>3709_f_at</td>
<td>NA</td>
<td>1.583758884</td>
<td>1.839268552</td>
<td>1.184649556</td>
</tr>
<tr>
<td>3716_f_at</td>
<td>NA</td>
<td>1.517062975</td>
<td>1.328377265</td>
<td>1.509759953</td>
</tr>
<tr>
<td>3758_f_at</td>
<td>NA</td>
<td>1.583597896</td>
<td>1.781075676</td>
<td>1.193950851</td>
</tr>
<tr>
<td>3764_s_at</td>
<td>NA</td>
<td>2.481400306</td>
<td>2.14278409</td>
<td>1.068471933</td>
</tr>
<tr>
<td>3767_s_at</td>
<td>NA</td>
<td>1.1195607967</td>
<td>1.082197877</td>
<td>1.098775105</td>
</tr>
<tr>
<td>3769_s_at</td>
<td>NA</td>
<td>1.232606256</td>
<td>2.234937299</td>
<td>1.216994846</td>
</tr>
<tr>
<td>3794_f_at</td>
<td>NA</td>
<td>1.016381195</td>
<td>1.784031442</td>
<td>1.131791624</td>
</tr>
<tr>
<td>3803_f_at</td>
<td>NA</td>
<td>1.225986232</td>
<td>1.54525296</td>
<td>1.449226313</td>
</tr>
<tr>
<td>3813_f_at</td>
<td>NA</td>
<td>2.058513645</td>
<td>1.716623277</td>
<td>1.36665844</td>
</tr>
<tr>
<td>3814_f_at</td>
<td>NA</td>
<td>1.343798436</td>
<td>1.034031015</td>
<td>1.192923085</td>
</tr>
<tr>
<td>3821_f_at</td>
<td>NA</td>
<td>2.058513645</td>
<td>1.716623277</td>
<td>1.342865844</td>
</tr>
<tr>
<td>3827_f_at</td>
<td>NA</td>
<td>1.15195968</td>
<td>1.485584325</td>
<td>1.320938084</td>
</tr>
<tr>
<td>3840_f_at</td>
<td>NA</td>
<td>1.31947678</td>
<td>1.899737437</td>
<td>1.166644167</td>
</tr>
<tr>
<td>3867_s_at</td>
<td>NA</td>
<td>1.530869411</td>
<td>1.728419633</td>
<td>1.514660293</td>
</tr>
<tr>
<td>3884_f_at</td>
<td>NA</td>
<td>1.7511441</td>
<td>1.632430561</td>
<td>1.227354629</td>
</tr>
<tr>
<td>3924_f_at</td>
<td>NA</td>
<td>1.04279452</td>
<td>1.878273994</td>
<td>1.158755642</td>
</tr>
<tr>
<td>4070_at</td>
<td>LYS1</td>
<td>1.650724845</td>
<td>1.979292924</td>
<td>1.901143433</td>
</tr>
<tr>
<td>4109_at</td>
<td>BAR1</td>
<td>2.269907778</td>
<td>1.632377873</td>
<td>1.51711513</td>
</tr>
<tr>
<td>4385_at</td>
<td>SPL2</td>
<td>-2.963481901</td>
<td>-3.967203495</td>
<td>-3.216817334</td>
</tr>
<tr>
<td>4386_at</td>
<td>ARO9</td>
<td>1.881736181</td>
<td>1.938215698</td>
<td>1.815695735</td>
</tr>
<tr>
<td>4390_at</td>
<td>NA</td>
<td>-1.07971541</td>
<td>-1.545567189</td>
<td>-1.026451368</td>
</tr>
<tr>
<td>4425_at</td>
<td>GAR1</td>
<td>-1.039890823</td>
<td>-1.380629771</td>
<td>-1.142678777</td>
</tr>
<tr>
<td>4426_at</td>
<td>HXT4</td>
<td>-1.152168479</td>
<td>-1.062658621</td>
<td>-1.22241608</td>
</tr>
<tr>
<td>4506_at</td>
<td>NA</td>
<td>-1.782036847</td>
<td>-1.653810522</td>
<td>-1.49785359</td>
</tr>
<tr>
<td>4561_s_at</td>
<td>NA</td>
<td>1.318713765</td>
<td>1.35859904</td>
<td>1.283820774</td>
</tr>
<tr>
<td>4583_at</td>
<td>SNF6</td>
<td>-1.160956469</td>
<td>-1.327787652</td>
<td>-1.123321353</td>
</tr>
<tr>
<td>4584_at</td>
<td>RIM4</td>
<td>1.83492348</td>
<td>1.851553314</td>
<td>1.036397929</td>
</tr>
<tr>
<td>4801_at</td>
<td>HIP1</td>
<td>-1.573368309</td>
<td>-1.378326261</td>
<td>-1.065595404</td>
</tr>
<tr>
<td>4813_at</td>
<td>NA</td>
<td>1.455183565</td>
<td>1.576188988</td>
<td>1.392299632</td>
</tr>
<tr>
<td>4832_at</td>
<td>ATF2</td>
<td>1.204774527</td>
<td>1.078300867</td>
<td>1.067023028</td>
</tr>
<tr>
<td>4842_at</td>
<td>BTN2</td>
<td>2.087277044</td>
<td>1.61624537</td>
<td>1.335855976</td>
</tr>
<tr>
<td>4850_at</td>
<td>CYS4</td>
<td>-1.13283771</td>
<td>-1.44113781</td>
<td>-1.287177133</td>
</tr>
<tr>
<td>4884_i_at</td>
<td>TPO2</td>
<td>1.649076161</td>
<td>1.617706435</td>
<td>1.610188624</td>
</tr>
<tr>
<td>4961_at</td>
<td>NA</td>
<td>-2.408236232</td>
<td>-3.175988828</td>
<td>-2.509424446</td>
</tr>
<tr>
<td>4994_at</td>
<td>NA</td>
<td>-1.407296483</td>
<td>-1.332803497</td>
<td>-1.34130352</td>
</tr>
<tr>
<td>gene</td>
<td>expression level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGA2</td>
<td>1.565750381</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DBP3</td>
<td>-1.753055449</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>1.222144693</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BUR6</td>
<td>1.302064445</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSA2</td>
<td>-1.71226706</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAD51</td>
<td>-1.996859608</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SER3</td>
<td>2.023708952</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCY2</td>
<td>-1.553462513</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRB1</td>
<td>1.062843303</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>1.114135951</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STP4</td>
<td>-1.229512878</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMI2</td>
<td>-1.291517468</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MFA1</td>
<td>1.898381709</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UTP5</td>
<td>-1.162193042</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARO10</td>
<td>2.308300344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VID21</td>
<td>-1.182472328</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUM1</td>
<td>-1.272076063</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHM6</td>
<td>-1.032407245</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVM1</td>
<td>-1.008429008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHP1</td>
<td>-1.411653576</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KNH1</td>
<td>-1.275016056</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LYS21</td>
<td>1.116486658</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STF1</td>
<td>-1.656991103</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>-1.602812585</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>-1.530738573</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFD2</td>
<td>-1.445552695</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REI1</td>
<td>-1.401543984</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHO89</td>
<td>-2.503977896</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWC5</td>
<td>-1.048700677</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZTA1</td>
<td>-1.17062068</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>-1.251903099</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FUR4</td>
<td>-1.078044562</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDS1</td>
<td>-1.67052184</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMT1</td>
<td>-1.391186039</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>URA7</td>
<td>-1.053633004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>1.085936664</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRO77</td>
<td>1.18558893</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DPM1</td>
<td>-1.152418059</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOC4</td>
<td>-1.437844264</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIF5</td>
<td>-1.244558998</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPA135</td>
<td>-1.124629988</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VTC3</td>
<td>-1.032394156</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOG1</td>
<td>-1.037288051</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRM3</td>
<td>1.009103304</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBD2</td>
<td>-1.263894032</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>1.325978148</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRE10</td>
<td>-1.103490156</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Gene Name</td>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>8186_at</td>
<td>NA</td>
<td>-1.212722769</td>
<td>-1.518279607</td>
<td>-1.308317842</td>
</tr>
<tr>
<td>8203_at</td>
<td>VTS1</td>
<td>-1.352692379</td>
<td>-1.523772776</td>
<td>-1.243778792</td>
</tr>
<tr>
<td>8241_at</td>
<td>MCH5</td>
<td>-1.979615186</td>
<td>-1.857201354</td>
<td>-1.98965766</td>
</tr>
<tr>
<td>8294_at</td>
<td>TPO4</td>
<td>1.05295604</td>
<td>1.415287096</td>
<td>1.115166897</td>
</tr>
<tr>
<td>8380_at</td>
<td>GAC1</td>
<td>-1.117067866</td>
<td>-1.254129824</td>
<td>-1.399384206</td>
</tr>
<tr>
<td>8387_at</td>
<td>GSP2</td>
<td>-1.011384844</td>
<td>-1.205741659</td>
<td>-1.155062716</td>
</tr>
<tr>
<td>8392_at</td>
<td>PNO1</td>
<td>-1.218703267</td>
<td>-1.608700345</td>
<td>-1.328932841</td>
</tr>
<tr>
<td>8445_at</td>
<td>LEU9</td>
<td>-1.259761381</td>
<td>-1.553940362</td>
<td>-1.195761111</td>
</tr>
<tr>
<td>8459_at</td>
<td>BUD21</td>
<td>-1.256649017</td>
<td>-1.24377915</td>
<td>-1.0291905</td>
</tr>
<tr>
<td>8473_at</td>
<td>TMA46</td>
<td>-1.306530465</td>
<td>-1.713773351</td>
<td>-1.463901523</td>
</tr>
<tr>
<td>8505_at</td>
<td>HMS1</td>
<td>1.193320725</td>
<td>1.688812373</td>
<td>1.12109662</td>
</tr>
<tr>
<td>8522_at</td>
<td>RSB1</td>
<td>-1.329883044</td>
<td>-1.429946965</td>
<td>-1.258654409</td>
</tr>
<tr>
<td>8550_at</td>
<td>TRM10</td>
<td>-1.274527458</td>
<td>-1.639366951</td>
<td>-1.290153472</td>
</tr>
<tr>
<td>8633_at</td>
<td>GRE2</td>
<td>-1.119489139</td>
<td>-1.363424375</td>
<td>-1.258363188</td>
</tr>
<tr>
<td>8714_at</td>
<td>NA</td>
<td>2.212995366</td>
<td>2.274634949</td>
<td>2.957207553</td>
</tr>
<tr>
<td>8808_at</td>
<td>LYS9</td>
<td>2.244712898</td>
<td>2.274634949</td>
<td>2.957207553</td>
</tr>
<tr>
<td>8811_at</td>
<td>NOG2</td>
<td>-1.146876988</td>
<td>-1.314533053</td>
<td>-1.096270779</td>
</tr>
<tr>
<td>8929_at</td>
<td>NA</td>
<td>-2.277537975</td>
<td>-2.086489943</td>
<td>-2.406972879</td>
</tr>
<tr>
<td>8981_at</td>
<td>RPS7B</td>
<td>-1.475543515</td>
<td>-1.565764084</td>
<td>-1.12391523</td>
</tr>
<tr>
<td>9023_at</td>
<td>MFA2</td>
<td>1.509525881</td>
<td>1.389804511</td>
<td>1.318109629</td>
</tr>
<tr>
<td>9027_at</td>
<td>AAH1</td>
<td>-2.286343878</td>
<td>-2.467192932</td>
<td>-2.07385226</td>
</tr>
<tr>
<td>9032_at</td>
<td>IPI3</td>
<td>-1.070736902</td>
<td>-1.891546001</td>
<td>-1.60922076</td>
</tr>
<tr>
<td>9068_at</td>
<td>DUG3</td>
<td>-1.142812885</td>
<td>-1.390121488</td>
<td>-1.027935402</td>
</tr>
<tr>
<td>9150_at</td>
<td>PCL1</td>
<td>1.318949524</td>
<td>1.749086757</td>
<td>1.643209794</td>
</tr>
<tr>
<td>9193_at</td>
<td>NA</td>
<td>2.34707451</td>
<td>2.294445897</td>
<td>1.026687324</td>
</tr>
<tr>
<td>9309_at</td>
<td>ADH2</td>
<td>-1.600901037</td>
<td>-2.012313222</td>
<td>-1.740172961</td>
</tr>
<tr>
<td>9339_at</td>
<td>HAS1</td>
<td>-1.35700158</td>
<td>-1.321860344</td>
<td>-1.19658027</td>
</tr>
<tr>
<td>9452_at</td>
<td>RGM1</td>
<td>-2.300959152</td>
<td>-2.199347063</td>
<td>-1.83996093</td>
</tr>
<tr>
<td>9515_at</td>
<td>NA</td>
<td>-1.150737949</td>
<td>-1.009873126</td>
<td>-1.276454584</td>
</tr>
<tr>
<td>9545_at</td>
<td>NA</td>
<td>-1.186881424</td>
<td>-1.077135163</td>
<td>-1.37896075</td>
</tr>
<tr>
<td>9637_at</td>
<td>ERG5</td>
<td>1.322483728</td>
<td>1.219407014</td>
<td>1.038951437</td>
</tr>
<tr>
<td>9760_at</td>
<td>PGA3</td>
<td>-1.131001678</td>
<td>-1.36540303</td>
<td>-1.43342088</td>
</tr>
<tr>
<td>9762_at</td>
<td>PHO84</td>
<td>-1.426957046</td>
<td>-1.312845821</td>
<td>-1.849082623</td>
</tr>
<tr>
<td>9764_at</td>
<td>GTR1</td>
<td>-1.078301943</td>
<td>-1.335501187</td>
<td>-1.25391772</td>
</tr>
<tr>
<td>9774_at</td>
<td>DAT1</td>
<td>-1.461455646</td>
<td>-1.498636278</td>
<td>-1.247005569</td>
</tr>
<tr>
<td>9853_at</td>
<td>FPR4</td>
<td>-1.213093437</td>
<td>-1.283012062</td>
<td>-1.060745604</td>
</tr>
<tr>
<td>9868_at</td>
<td>COX19</td>
<td>-1.82045205</td>
<td>-2.044719937</td>
<td>-1.49866797</td>
</tr>
<tr>
<td>9885_at</td>
<td>CAR2</td>
<td>1.705506715</td>
<td>1.77349023</td>
<td>1.34087261</td>
</tr>
<tr>
<td>9898_at</td>
<td>NA</td>
<td>-1.247535322</td>
<td>-1.57140508</td>
<td>-1.131618907</td>
</tr>
</tbody>
</table>
Table 1.2.2. List of the top 20 genes most affected by Tra1 inactivation, as assessed by expression profiling. For those genes also affected by SAGA inactivation, the predicted activators are listed.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Fold down-regulation</th>
<th>Putative activators</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPL2</td>
<td>Suppressor of plc1 deletion; downregulates low-affinity phosphate transport during phosphate limitation</td>
<td>9.30</td>
<td>Aro80, Cbf1, Pho4, Cbf1, Pho4, Pho2</td>
</tr>
<tr>
<td>PHO89</td>
<td>Involved in phosphate metabolism; Na+Pi co-transporter; transcription regulated by Pho4</td>
<td>7.26</td>
<td>Pho4, Pho4, Aft2</td>
</tr>
<tr>
<td>YGR035C</td>
<td>Unknown; transcription activated by Yrm1 and Yrr1</td>
<td>5.69</td>
<td></td>
</tr>
<tr>
<td>YNL058C</td>
<td>Unknown</td>
<td>5.30</td>
<td>Ndd1, Mcm1, Fkh1, Mcm1, Ndd1</td>
</tr>
<tr>
<td>NCA3</td>
<td>Nuclear control of ATPase; functions with Nca2 to regulate mitochondrial expression of F0-F1 ATP synthase subunits</td>
<td>4.48</td>
<td></td>
</tr>
<tr>
<td>PHO11</td>
<td>Involved in phosphate metabolism; acid phosphatase; phosphate starvation-induced transcription coordinately regulated by Pho4 and Pho2</td>
<td>4.30</td>
<td>Pho4</td>
</tr>
<tr>
<td>IDP2</td>
<td>Cytosolic NADP-specific isocitrate dehydrogenase;</td>
<td>4.28</td>
<td></td>
</tr>
<tr>
<td>AAH1</td>
<td>Adenine aminohydrolase; transcriptionally regulated by nutrient levels and growth phase</td>
<td>4.21</td>
<td></td>
</tr>
<tr>
<td>MCH5</td>
<td>Riboflavin transporter</td>
<td>3.97</td>
<td>Put3</td>
</tr>
<tr>
<td>ZRT2</td>
<td>Zinc-regulated transporter; transcription induced under low-zinc conditions by Zap1</td>
<td>3.83</td>
<td>Zap1</td>
</tr>
<tr>
<td>PHO84</td>
<td>Involved in phosphate metabolism; high-affinity inorganic phosphate transporter; regulated by Pho4 and Spt7</td>
<td>3.60</td>
<td>Pho4, Cbf1, Pho4</td>
</tr>
<tr>
<td>RGM1</td>
<td>Putative transcriptional repressor with proline-rich zinc fingers</td>
<td>3.58</td>
<td>Reb1</td>
</tr>
<tr>
<td>RAD51</td>
<td>Strand exchange protein involved in recombinational repair of double-strand breaks</td>
<td>3.50</td>
<td>Mbp1, Swi6, Swi6</td>
</tr>
<tr>
<td>ADH2</td>
<td>Glucose-repressible alcohol dehydrogenase; transcriptionally regulated by Adr1</td>
<td>3.34</td>
<td></td>
</tr>
<tr>
<td>SOL4</td>
<td>Suppressor of Los1-1; 6-phosphogluconolactonase</td>
<td>3.23</td>
<td></td>
</tr>
<tr>
<td>ZTA1</td>
<td>Zeta-crystallin; NADPH-dependent quinone reductase</td>
<td>3.20</td>
<td></td>
</tr>
<tr>
<td>STF1</td>
<td>Stabilizing factor; involved in regulation of the mitochondrial F1-F0 ATP synthase</td>
<td>3.19</td>
<td></td>
</tr>
<tr>
<td>EMI2</td>
<td>Early meiotic induction; required for transcriptional induction of early meiotic-specific transcription factor IME1</td>
<td>3.12</td>
<td></td>
</tr>
<tr>
<td>RFU1</td>
<td>Inhibits Doa4 deubiquitinating activity</td>
<td>3.02</td>
<td></td>
</tr>
<tr>
<td>BUD20</td>
<td>Involved in bud-site selection</td>
<td>2.96</td>
<td></td>
</tr>
</tbody>
</table>
Table 1.2.3. Microarray data comparing gene expression in a *tra1-mut1* strain relative to a wild-type *TRA1* strain (p<0.05).

<table>
<thead>
<tr>
<th>ID</th>
<th>Gene symbol</th>
<th>Fold change</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2595</td>
<td>3310_s_at</td>
<td>-1.207130626</td>
<td>0.001779359</td>
</tr>
<tr>
<td>1545</td>
<td>2207_s_at</td>
<td>-1.19256344</td>
<td>0.012666811</td>
</tr>
<tr>
<td>3189</td>
<td>3913_s_at</td>
<td>-1.013557778</td>
<td>0.005681208</td>
</tr>
<tr>
<td>6130</td>
<td>6854_at</td>
<td>-1.005050451</td>
<td>0.000142935</td>
</tr>
<tr>
<td>4174</td>
<td>4898_at</td>
<td>-0.882156111</td>
<td>0.000596586</td>
</tr>
<tr>
<td>2272</td>
<td>2944_g_at</td>
<td>-0.874190892</td>
<td>0.003812645</td>
</tr>
<tr>
<td>1544</td>
<td>2206_g_at</td>
<td>-0.872542119</td>
<td>0.00423667</td>
</tr>
<tr>
<td>8980</td>
<td>9704_at</td>
<td>-0.852927463</td>
<td>0.001119004</td>
</tr>
<tr>
<td>6217</td>
<td>6941_s_at</td>
<td>-0.851942748</td>
<td>0.00722219</td>
</tr>
<tr>
<td>5203</td>
<td>5927_at</td>
<td>-0.84978062</td>
<td>0.002388645</td>
</tr>
<tr>
<td>2504</td>
<td>3187_s_at</td>
<td>-0.84015179</td>
<td>0.073767104</td>
</tr>
<tr>
<td>3468</td>
<td>4192_at</td>
<td>-0.827467305</td>
<td>0.007772781</td>
</tr>
<tr>
<td>3225</td>
<td>3949_i_at</td>
<td>-0.817149365</td>
<td>0.01218005</td>
</tr>
<tr>
<td>6832</td>
<td>7556_at</td>
<td>-0.812754741</td>
<td>0.000632008</td>
</tr>
<tr>
<td>2298</td>
<td>2970_s_at</td>
<td>-0.804534336</td>
<td>0.029141472</td>
</tr>
<tr>
<td>70</td>
<td>10069_f_at</td>
<td>-0.798251776</td>
<td>0.017974932</td>
</tr>
<tr>
<td>4698</td>
<td>5422_s_at</td>
<td>-0.783115415</td>
<td>0.001070263</td>
</tr>
<tr>
<td>6041</td>
<td>6765_at</td>
<td>-0.772422209</td>
<td>0.005972537</td>
</tr>
<tr>
<td>69</td>
<td>10068_i_at</td>
<td>-0.762286462</td>
<td>0.089983193</td>
</tr>
<tr>
<td>5600</td>
<td>6324_g_at</td>
<td>-0.739842079</td>
<td>0.016506907</td>
</tr>
<tr>
<td>9144</td>
<td>9868_at</td>
<td>-0.718216349</td>
<td>0.098525152</td>
</tr>
<tr>
<td>2164</td>
<td>2836_g_at</td>
<td>-0.704534724</td>
<td>0.085787013</td>
</tr>
<tr>
<td>4896</td>
<td>5620_at</td>
<td>-0.700606395</td>
<td>0.030026657</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Gene Name</td>
<td>Expression Value</td>
<td>Fold Change</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
<td>------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>2948</td>
<td>3672_f_at</td>
<td>YBL005W-B</td>
<td>-0.67181392</td>
</tr>
<tr>
<td>8622</td>
<td>9346_at</td>
<td>YMR295C</td>
<td>-0.670228666</td>
</tr>
<tr>
<td>7107</td>
<td>7831_at</td>
<td>YPL058C</td>
<td>-0.667407491</td>
</tr>
<tr>
<td>757</td>
<td>10756_at</td>
<td>YKL163W</td>
<td>-0.665228757</td>
</tr>
<tr>
<td>2292</td>
<td>2964_at</td>
<td>YKL182W</td>
<td>-0.66111955</td>
</tr>
<tr>
<td>3668</td>
<td>4392_at</td>
<td>YHR143W</td>
<td>-0.65723394</td>
</tr>
<tr>
<td>7836</td>
<td>8560_at</td>
<td>YOL002C</td>
<td>-0.643088742</td>
</tr>
<tr>
<td>4310</td>
<td>5034_at</td>
<td>YGL028C</td>
<td>-0.639647293</td>
</tr>
<tr>
<td>8129</td>
<td>8853_at</td>
<td>YNR005C</td>
<td>-0.637146572</td>
</tr>
<tr>
<td>1999</td>
<td>2666_s_at</td>
<td>YNL054W</td>
<td>-0.63380518</td>
</tr>
<tr>
<td>4820</td>
<td>5544_at</td>
<td>YER150W</td>
<td>-0.630562676</td>
</tr>
<tr>
<td>3856</td>
<td>4580_at</td>
<td>YHL028W</td>
<td>-0.627175751</td>
</tr>
<tr>
<td>4119</td>
<td>4843_at</td>
<td>YGR143W</td>
<td>-0.621938131</td>
</tr>
<tr>
<td>3191</td>
<td>3915_s_at</td>
<td>YAR009C</td>
<td>-0.619467852</td>
</tr>
<tr>
<td>5331</td>
<td>6055_at</td>
<td>YDR404C</td>
<td>-0.618070705</td>
</tr>
<tr>
<td>1976</td>
<td>2643_s_at</td>
<td>YOR153W</td>
<td>-0.617999736</td>
</tr>
<tr>
<td>8172</td>
<td>8896_at</td>
<td>YNL046W</td>
<td>-0.614783895</td>
</tr>
<tr>
<td>1124</td>
<td>11123_at</td>
<td>YJL079C</td>
<td>-0.613590464</td>
</tr>
<tr>
<td>3611</td>
<td>4335_at</td>
<td>YHR175W</td>
<td>-0.610404855</td>
</tr>
<tr>
<td>5193</td>
<td>5917_at</td>
<td>YDR538W</td>
<td>-0.606220723</td>
</tr>
<tr>
<td>5148</td>
<td>5872_at</td>
<td>YDR374W-A</td>
<td>-0.604194051</td>
</tr>
<tr>
<td>3519</td>
<td>4243_at</td>
<td>YIL158W</td>
<td>-0.603320387</td>
</tr>
<tr>
<td>7798</td>
<td>8522_at</td>
<td>YOR049C</td>
<td>-0.59964098</td>
</tr>
<tr>
<td>3455</td>
<td>4179_at</td>
<td>YIL087C</td>
<td>-0.595911291</td>
</tr>
<tr>
<td>3340</td>
<td>4064_at</td>
<td>YIR028W</td>
<td>-0.592191668</td>
</tr>
<tr>
<td>5756</td>
<td>6480_at</td>
<td>YDR019C</td>
<td>-0.590975313</td>
</tr>
</tbody>
</table>
Table 1.2.4. Microarray data comparing gene expression in a *tra1-mut8* strain relative to a wild-type *TRA1* strain (p<0.05).

<table>
<thead>
<tr>
<th>ID</th>
<th>Gene symbol</th>
<th>Fold change</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2595</td>
<td>3310_s_at</td>
<td>YDR034C-A</td>
<td>-1.097067224</td>
</tr>
<tr>
<td>9145</td>
<td>9869_at</td>
<td>YLR264C-A</td>
<td>-0.922627325</td>
</tr>
<tr>
<td>6130</td>
<td>6854_at</td>
<td>YCR021C</td>
<td>-0.895472494</td>
</tr>
<tr>
<td>6217</td>
<td>6941_s_at</td>
<td>YER190C-A</td>
<td>-0.804113967</td>
</tr>
<tr>
<td>4160</td>
<td>4884_i_at</td>
<td>YGR138C</td>
<td>-0.768574006</td>
</tr>
<tr>
<td>3189</td>
<td>3913_s_at</td>
<td>YAR010C</td>
<td>-0.729111841</td>
</tr>
<tr>
<td>5148</td>
<td>5872_at</td>
<td>YDR374W-A</td>
<td>-0.666964525</td>
</tr>
<tr>
<td>7836</td>
<td>8560_at</td>
<td>YOL002C</td>
<td>-0.666094003</td>
</tr>
<tr>
<td>3340</td>
<td>4064_at</td>
<td>YIR028W</td>
<td>-0.651389945</td>
</tr>
<tr>
<td>123</td>
<td>10122_at</td>
<td>YLR230W</td>
<td>-0.63627118</td>
</tr>
<tr>
<td>5203</td>
<td>5927_at</td>
<td>YDL048C</td>
<td>-0.613805619</td>
</tr>
<tr>
<td>431</td>
<td>10430_at</td>
<td>YKL183C-A</td>
<td>-0.609352353</td>
</tr>
<tr>
<td>388</td>
<td>10387_at</td>
<td>YLL044W</td>
<td>-0.601495203</td>
</tr>
<tr>
<td>5297</td>
<td>6021_f_at</td>
<td>YDR461W</td>
<td>-0.587224505</td>
</tr>
<tr>
<td>Strain</td>
<td>Genotype</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------------</td>
<td></td>
</tr>
<tr>
<td>FY23</td>
<td>MATα leu2Δ1 trp1Δ63 ura3-52</td>
<td>Winston 1995</td>
<td></td>
</tr>
<tr>
<td>YDA352</td>
<td>MATα leu2Δ1 trp1Δ63 ura3-52 spt20::URA3</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>BY4741</td>
<td>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0</td>
<td>Giaever 2002</td>
<td></td>
</tr>
<tr>
<td>LLY074</td>
<td>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 TRA1-VN</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY071</td>
<td>MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 VN-TRA1</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>BY4742</td>
<td>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</td>
<td>Giaever 2002</td>
<td></td>
</tr>
<tr>
<td>LLY080</td>
<td>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 GAL4-VC</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY082</td>
<td>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 GCN4-VC</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>MDC1</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1</td>
<td>Kulesza 2002</td>
<td></td>
</tr>
<tr>
<td>MDC3</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1</td>
<td>Kulesza 2002</td>
<td></td>
</tr>
<tr>
<td>LLY154</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY129</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY131</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY197</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY210</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY240</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY241</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY242</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY238</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>Strain</td>
<td>Genotype</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>---------------------------------------------------------------------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>LLY245</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1 tr1::LEU2 REB1-VC+[myc-TRA1/CEN/URA3]</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY246</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1 tr1::LEU2 MCM1-VC+[myc-TRA1/CEN/URA3]</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY247</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1 tr1::LEU2 RAP1-VC+[myc-TRA1/CEN/URA3]</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY278</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1 tr1::LEU2 GCN4-VC+[myc-TRA1/CEN/URA3]</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>LLY279</td>
<td>MATa ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1 tr1::LEU2 FKH2-VC+[myc-TRA1/CEN/URA3]</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>W303</td>
<td>MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15</td>
<td>Veal 2003</td>
<td></td>
</tr>
<tr>
<td>LLY326</td>
<td>MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 GAL80-VN</td>
<td>This study</td>
<td></td>
</tr>
</tbody>
</table>
Table 1.2.6. Primer sequences for quantitative real-time RT-PCR (qRT-PCR) analysis, chromatin immunoprecipitation (ChIP) and yeast strain construction.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward or Reverse</th>
<th>Sequence (5' → 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAH1</td>
<td>Forward</td>
<td>CGGTCGAGGAATTAAACGAA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CCACGCCAAATCAAAGAAAT</td>
</tr>
<tr>
<td>ADH2</td>
<td>Forward</td>
<td>ACTACGCGGGTATCAAATGG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TCGTGGGTTGTAACCAGACAA</td>
</tr>
<tr>
<td>GAL1</td>
<td>Forward</td>
<td>GGATCGAGCACGAATAAAGG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>ATTCAATATCGCCGTTCCAG</td>
</tr>
<tr>
<td>GAL3</td>
<td>Forward</td>
<td>ACCCTTATGAGGTGCGCAGA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GCGCAAGTAATGCGAGATGA</td>
</tr>
<tr>
<td>GSC2</td>
<td>Forward</td>
<td>GGTACAGCACCAGTTAGGA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>AAACAAAATCCGCTGCAAAC</td>
</tr>
<tr>
<td>HSP30</td>
<td>Forward</td>
<td>CTAATTGAGTCAGCCGTGA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TCCGTAGCATGGTGATGAGA</td>
</tr>
<tr>
<td>MCH5</td>
<td>Forward</td>
<td>TCGCTACCAGCACTCTACT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TAGTGGCGCCGAAACACTTAC</td>
</tr>
<tr>
<td>PHO11</td>
<td>Forward</td>
<td>GTGATTTGCCGGAAGTGTG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TTTTGCTTTGCTGACAGTG</td>
</tr>
<tr>
<td>PHO84</td>
<td>Forward</td>
<td>TCTGCAAACCAGCTGGCTC</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CACCACCAGTACCAATACCC</td>
</tr>
<tr>
<td>PHO89</td>
<td>Forward</td>
<td>AGGCACAAGAGTTCCTTGTA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>ACCAACAAGACAGCCACAC</td>
</tr>
<tr>
<td>RAD51</td>
<td>Forward</td>
<td>AGATATTGAGGCCACCAAC</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>AGGCAGCTTCATCGTGATGCT</td>
</tr>
<tr>
<td>Gene</td>
<td>Forward or Reverse</td>
<td>Sequence (5' → 3')</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>RPS5</td>
<td>Forward</td>
<td>ACTGACCAAAAAACCAATCCA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TTGACGTCTAGCAGCACCAC</td>
</tr>
<tr>
<td>RPS0B</td>
<td>Forward</td>
<td>CAATCGCCCGGTAGATTCACT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GGCCTGAGCGTCTAATCTTG</td>
</tr>
<tr>
<td>RGM1</td>
<td>Forward</td>
<td>CTCCTGCTAGCCCAACAGAC</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GCAATGGTATGGGAGCAGAT</td>
</tr>
<tr>
<td>SPL2</td>
<td>Forward</td>
<td>GGTCACCAGCATAGGGAAG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TTGAGCCTCCTGCACTCTTT</td>
</tr>
<tr>
<td>YNL058C</td>
<td>Forward</td>
<td>CCAAGTCCCTTCCATCGTTA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CTGCTGTAGATGGGGGTGTT</td>
</tr>
<tr>
<td>ZRT2</td>
<td>Forward</td>
<td>AAAAAAGAACCCGCTCAAT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CATGGGAACTGGTGGAAACT</td>
</tr>
</tbody>
</table>

### ChIP

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward or Reverse</th>
<th>Sequence (5' → 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAH1</td>
<td>Forward</td>
<td>TTGCAAATTTCATCGCTTA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>ATGCGATGAGGTCAAGTTCC</td>
</tr>
<tr>
<td>ADH2</td>
<td>Forward</td>
<td>AACGGCTTTCGCTCATAAAA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CCAAAAAGTGAACCCCATTT</td>
</tr>
<tr>
<td>GAL1</td>
<td>Forward</td>
<td>CGTTCCTGAAACGCAGATGT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GCCAGGTTACTGCAATTTT</td>
</tr>
<tr>
<td>MCH5</td>
<td>Forward</td>
<td>CCAAGCGCATATTCAACC</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TTCCATCTTTCCTGCTGT</td>
</tr>
<tr>
<td>PHO11</td>
<td>Forward</td>
<td>TACTGGGAGCGAGGGTGCTT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>AAAAAACAAACGCATCACC</td>
</tr>
<tr>
<td>PHO84</td>
<td>Forward</td>
<td>TGAAGTCATCTCCGGGTAGTG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GAGGCGGTATAATCGAAAA</td>
</tr>
<tr>
<td>PHO89</td>
<td>Forward</td>
<td>AAGGATCCAAAGTGCTGTG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CAGCACGTGGGAGACAAATA</td>
</tr>
<tr>
<td>Gene</td>
<td>Forward or Reverse</td>
<td>Sequence (5' → 3')</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------</td>
<td>-----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>RAD51</td>
<td>Forward</td>
<td>TGCACCACTACCGTTCTTCCA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GGGTAACGAAGTGGCCTA</td>
</tr>
<tr>
<td>RGM1</td>
<td>Forward</td>
<td>AGAATTTGCTCAAGGCGCTA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CGCCTTTTATAGTGCCCATC</td>
</tr>
<tr>
<td>RPS0B</td>
<td>Forward</td>
<td>TGGGCATCTACCTTTCAACA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TGTATGAATGCGATGGGTTT</td>
</tr>
<tr>
<td>SPL2</td>
<td>Forward</td>
<td>TGCCCCATGGGAGTCTAGT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CTCTCTTTTCCCACGTGCTC</td>
</tr>
<tr>
<td>YNL058C</td>
<td>Forward</td>
<td>TTGACAAGGTATGAGTCTAGTTGATT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CATGAAACCCTTAATGATTTCC</td>
</tr>
<tr>
<td>ZRT2</td>
<td>Forward</td>
<td>GGTATATGCGACTCTTCTCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CCGTCTAACCACAGGGTGT</td>
</tr>
</tbody>
</table>

**Yeast strain construction**

**Genomic tagging**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward or Reverse</th>
<th>Sequence (5' → 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBF1-VC</td>
<td>Forward</td>
<td>GAAACAAGAGAAGAAAGAGAGCCTAGGAGCGATAAATCCACATGAGGCTTGTTACAAACGCAGATGACAGGCGATCCCGCGGTT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>ACGCACTACATAGGGAGAAGCTCAAGGAGATACATTATATCTACTCTTAAACTCTCGATGAAATCCAGCTCGT</td>
</tr>
<tr>
<td>CIN5-VC</td>
<td>Forward</td>
<td>AAACAGTGAATTGAAAGAAAATGATTGAATCTAAAGCAGATAATTAAAAAGAAGATACATAGGGAGAAGCTCAAGGCGATCCCGGGTT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TGTGCTTTGAAAACCTTAAAGATGTATACCTACTACTACTATAATTACATTATTTCTATGGAATTCCAGCTCGT</td>
</tr>
<tr>
<td>FKH2-VC</td>
<td>Forward</td>
<td>CAAGGAACATTAAACTAGATACGGATGTCAGAAGATCAGTATTACAAACAGAAAGGCACGATGACAAAGGTCAGGATCCCGGGTT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>AGCCATTCTCATTATTTCTATTATGCATTGATGATTGATTCCTTTGTTTGGTCATGGAATCCAGCTCGT</td>
</tr>
<tr>
<td>Gene</td>
<td>Primer Set</td>
<td>Forward Sequence</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>GAL4-VC</td>
<td>Forward</td>
<td>ATGGATGATGATATATAACTATCTATTCGATGA TGAAGATACCCCCACCAACAAAAACAAAAAAAAGAG GGTGACCGGATCCCCGGGTT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TGAGATGGGTGCACGATGCACAGTGGAAGTTG AACTTGCGGGGGTTTTTCGATATCTACGATTC ATTTGATGAATTGAGCTCAGTT</td>
</tr>
<tr>
<td>GAL80-VN</td>
<td>Forward</td>
<td>TTAATCGAGAGCGGTATTTTATAAAGTAGAACATGAG TGGGCTCCACATTTAAACGTTAGCAAAATTCTG CATTATAGATTTTGGGCAGGATTCAGCTCAGTT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TTACATAGATATATACTCATTAATCGTTTTTTTA TAACGTTCGCTGACTGGGGGCAAGCAGCA GCCAAGATGCTTTGATGAAATTGAGCTCAGTT</td>
</tr>
<tr>
<td>GCN4-VC</td>
<td>Forward</td>
<td>TCGAAAAATTATCACTTGGAAAATGAGGTTG CCAGATTAAAGAAATAGTTGGCGAAGCC GGTATCCCCGGGTTAATTAA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TCGAAAAATTATCACTTGGAAAATGAGGTTG CCAGATTAAAGAAATAGTTGGCGAAGCC GGTATCCCCGGGTTAATTAA</td>
</tr>
<tr>
<td>MCM1-VC</td>
<td>Forward</td>
<td>TGCTGCTACCCACAAAATTCATTTCGATACGAACTTGGCGAAGCC GGTATCCCCGGGTTAATTAA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>AGCTTTTTTCTCTTTTAATGCTGTTTCTATGAAATT ATATACGGAAAATCGATACGATTTAACGAGCTCGGTT</td>
</tr>
<tr>
<td>RAP1-VC</td>
<td>Forward</td>
<td>TGGAAACTGGTAGAATGGGAAATCGGAGGAAAAAG ATTTTTGGAGAAGGACCTTTGAATTACAGGAGCGAGTACGAGGATCCCCGGGTT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TGAATAAAAGGAGTTAAAAATAGTTAAAAACATG GGTGTTACTTCAATTACTCATCGATGAAATC GAGCTCGGTT</td>
</tr>
<tr>
<td>REB1-VC</td>
<td>Forward</td>
<td>GCTAGTTGATTATTATTGGCTCACAATTTTTCAA TGAAAGACAGAAATGAAGAAATTTGATCTCG TAAGAGAAGACTTTGGGTGGCGAGGATTTCACTGAGGATCATCCTC GGTGTTACTTCAATTACTCATCGATGAAATC GAGCTCGGTT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CAAAACATTATGGAGTTTTTCGCTTTTCGATTTAACGAAT TATATTTCGGGAAATCGATGAAATTGAGCTCAGTT</td>
</tr>
<tr>
<td>TRA1-VN</td>
<td>Forward</td>
<td>CCTAGACTGTATTTGGCGAGCGTGTCTGTCAGTCC</td>
</tr>
<tr>
<td>Gene</td>
<td>Forward or Reverse</td>
<td>Sequence (5’ → 3’)</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>CBF1-VC</td>
<td>Forward</td>
<td>GCAGGAAGAAGCTGGGAATAG</td>
</tr>
<tr>
<td>CIN5-VC</td>
<td>Forward</td>
<td>AAAATCGAAGCGGACAAAG</td>
</tr>
<tr>
<td>FKH2-VC</td>
<td>Forward</td>
<td>ATAGCGGAGGCCAACAGAG</td>
</tr>
<tr>
<td>GAL4-VC</td>
<td>Forward</td>
<td>CACAACCAATTGCCTCCTCT</td>
</tr>
<tr>
<td>GAL80-VN</td>
<td>Forward</td>
<td>CTAGAGCAACGACTTCCC</td>
</tr>
<tr>
<td>GCN4-VC</td>
<td>Forward</td>
<td>AATCCAGTGATCCTGCTGCT</td>
</tr>
<tr>
<td>MCM1-VC</td>
<td>Forward</td>
<td>CAGCAACAAATGTCACAGCA</td>
</tr>
<tr>
<td>RAP1-VC</td>
<td>Forward</td>
<td>TGTTAATCCTTCCTCCAAG</td>
</tr>
<tr>
<td>REB1-VC</td>
<td>Forward</td>
<td>TGAAGCCTGGCATGAAATTA</td>
</tr>
<tr>
<td>TRA1-VN</td>
<td>Forward</td>
<td>GGTCAACCAACGCTTGAC</td>
</tr>
<tr>
<td>VN-TRA1</td>
<td>Forward</td>
<td>CCAGGAGCGACCATCTTCTCAA</td>
</tr>
</tbody>
</table>

**Confirmation of genomic tagging**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward or Reverse</th>
<th>Sequence (5’ → 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAGAAACTTAGCAAGAACAGACGTGAACTTCATGCCATGGTTCGGTCGACGGATCCCCGGGGTT</td>
<td>Reverse</td>
<td>ATTTTTTGAGGCTTTCTCTACCTTCATTAATACAGAGAAAGGCCGGAAAACCTAACCCTAGTATCACGCTTCGATGAAATCCGAGCTCGTTCGTTTTAAAC</td>
</tr>
<tr>
<td>CATAAAGAAGATTTAATTCAGAGAAGATCTATCACACGATATCTCCATAGTTACAAGAATACCGCATTTCAGGGAATTCGAGCTCGTTCGTTTTAAAC</td>
<td>Reverse</td>
<td>ATGTCGATCAGTCAGCTGCTCAGTGAGTGATATCACGACGATGACAGATCCCCGGGTAC</td>
</tr>
<tr>
<td>AAAATTGAACTCAGAAATGGAATGCGCAATTGAAAGAAAGTTGCGAATTTGATTACCAAGAGACGCGATGACAAAGGTCGACGGATCCCCGGTT</td>
<td>Reverse</td>
<td>TTTATGTCATGCTTCAAGAACCAGCCTTTCACGTTGGGACCATGATGACATCAGACGATGACCCGAGCTCGTGTTTAAAC</td>
</tr>
<tr>
<td>GCAGGAAGAACTGGGAATACTGGGAATAG</td>
<td>Reverse</td>
<td>GCAGGAAGAAGCTGGGAATAG</td>
</tr>
<tr>
<td>AAAATCGAAGCGGACAAAGAAG</td>
<td>Reverse</td>
<td>AAAATCGAAGCGGACAAAGAAG</td>
</tr>
<tr>
<td>ATAGCGGAGGCAAAACAGAGAAG</td>
<td>Reverse</td>
<td>ATAGCGGAGGCAAAACAGAGAAG</td>
</tr>
<tr>
<td>CACAACCAATTGCCTCCTCCTC</td>
<td>Reverse</td>
<td>CACAACCAATTGCCTCCTCCTC</td>
</tr>
<tr>
<td>CTAGAGCAACGACTTCCCG</td>
<td>Reverse</td>
<td>CTAGAGCAACGACTTCCCG</td>
</tr>
<tr>
<td>AATCCAGTGATCCTGCTGCT</td>
<td>Reverse</td>
<td>AATCCAGTGATCCTGCTGCT</td>
</tr>
<tr>
<td>CAGCAACAAATGTCACAGCA</td>
<td>Reverse</td>
<td>CAGCAACAAATGTCACAGCA</td>
</tr>
<tr>
<td>TGTTAATCCTTCCTCCAAG</td>
<td>Reverse</td>
<td>TGTTAATCCTTCCTCCAAG</td>
</tr>
<tr>
<td>TGAAGCCTGGCATGAAATTA</td>
<td>Reverse</td>
<td>TGAAGCCTGGCATGAAATTA</td>
</tr>
<tr>
<td>GGTCAACCAACGCTTGACC</td>
<td>Reverse</td>
<td>GGTCAACCAACGCTTGACC</td>
</tr>
<tr>
<td>CCAGGAGCGCACCATCTTCTTCAA</td>
<td>Reverse</td>
<td>CCAGGAGCGCACCATCTTCTTCAA</td>
</tr>
<tr>
<td>Gene</td>
<td>Forward or Reverse</td>
<td>Sequence (5' → 3')</td>
</tr>
<tr>
<td>------------</td>
<td>--------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ZAP1-VC</td>
<td>Reverse</td>
<td>TCGTGATTTCACCAGATCGTC</td>
</tr>
<tr>
<td>VC</td>
<td>Reverse</td>
<td>TTCTCGTTGGGGTCTTTGCTCAG</td>
</tr>
<tr>
<td>VN</td>
<td>Reverse</td>
<td>TTGAAGAAGATGGTGCGCTCCTG</td>
</tr>
</tbody>
</table>

**Plasmid tagging**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward or Reverse</th>
<th>Sequence (5' → 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAL4(AD)-VC</td>
<td>Forward</td>
<td>ATGGGATGATGTATATAACTATCTATTTCGATGATGA AGATACCCCAACCAAAACAAAAAAAGAGGGGTGAAGGGATGGA CCGATCCCAGGTT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>AATGAAAGAAATTGAGATGGTGACGATGACAGATGA TTGAAGTGAACCTTGCGGGGTGTTTTCAGATTACCTACGATTACCTCGATGGAACGATGGA CCGATCCCAGGTT</td>
</tr>
<tr>
<td>TRA1-VN</td>
<td>Forward</td>
<td>TTCAGACTGTATTGGCAGCGTGCTGTCAGTCCAAG AAACCTTAGGAAGACGAGCTGAACCTTATGCGCAATAGCGGGATGGAACGATGGA CCGATCCCAGGTT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>ATGACCATGATTACGCGCAATAGCAGTCAGTCCAAG AAACCTTAGGAAGACGAGCTGAACCTTATGCGCAATAGCGGGATGGAACGATGGA CCGATCCCAGGTT</td>
</tr>
<tr>
<td>VN-TRA1</td>
<td>Forward</td>
<td>GCCGCCGCGCATGGTGAGCAAGGGCGAGG GCCGCCGCGCATGGTGAGCAAGGGCGAGG GCCGCCGCGCATGGTGAGCAAGGGCGAGG GCCGCCGCGCATGGTGAGCAAGGGCGAGG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GCCGCCGCGCATGGTGAGCAAGGGCGAGG GCCGCCGCGCATGGTGAGCAAGGGCGAGG GCCGCCGCGCATGGTGAGCAAGGGCGAGG GCCGCCGCGCATGGTGAGCAAGGGCGAGG</td>
</tr>
</tbody>
</table>

**Confirmation of plasmid tagging**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward or Reverse</th>
<th>Sequence (5' → 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAL4)AD)-VC</td>
<td>Forward</td>
<td>CACAACCAA TTGCCTTCCTCT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TTCTCGTTGGGGTCTTTGCTCAG</td>
</tr>
<tr>
<td>TRA1-VN</td>
<td>Forward</td>
<td>CCAGGAGCGCAACTTCTTCTCA</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GGTGGCGATGGATCTGTTA</td>
</tr>
<tr>
<td>VN-TRA1</td>
<td>Forward</td>
<td>GAGCTCAGATCCATGCACCACATGG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GAGCTCAGATCCATGCACCACATGG</td>
</tr>
</tbody>
</table>


Cirillo LA, Zaret KS. 1999. An early developmental transcription factor complex that is more stable on nucleosome core particles than on free DNA. Mol. Cell 4:961–69


Weiss, S. & Gladstone, L. A mammalian system for the incorporation of cytidine triphosphate into ribonucleic acid. J. Am. Chem. Soc. 81, 4118–4119 (1959)


PART II

IDENTIFICATION OF NEW TUMOR SUPPRESSOR GENES THROUGH A LARGE-SCALE SHRNA SCREEN FOR MOUSE TUMORIGENESIS
CHAPTER I

GENERAL INTRODUCTION
2.1.1 Cancer

Cancer formation is the transformation of normal cells into highly malignant invasive derivatives, via a series of premalignant states (Fearon & Vogelstein, 1990). It involves dynamic genetic changes in the genome in a stepwise process, and each step enables precancerous cells to acquire one type of growth advantage. The theory that cancer is a genetic disease was first proposed by David von Hansemann in 1890. Later, in the early twentieth century, a zoologist named Theodor Boveri brought up a bold and visionary hypothesis that provided the foundation for viewing cancer as a genetic disease. He believed that cancer origins from a single cell with genetic alterations and the genetic alterations caused by environment stimulation could endow unlimited cell growth. Boveri also proposed the concepts of cell cycle checkpoint, tumor suppressor genes and oncogenes. Many of his predictions are now explained by the molecular mechanisms that discovered using modern techniques.

Although cancer is a complex disease, its initiation and progression have several characteristic hallmarks that are shared by almost all types of cancers. Weinberg, Pouyssegur, Elledge and their colleagues, summarized the hallmarks of cancer and their function in transformation process (Hanahan & Weinberg, 2000, 2011; Kroemer & Pouyssegur, 2008; Luo et al, 2009). Hallmarks of cancer can be categorized into two classes: “emerging hallmarks”, which initiate or
promote transformation progress, and “enabling characteristics”, that are resulted from “emerging hallmarks” and are the phenotypes shared by many cancer types. “Emerging hallmarks” include: (1) deregulation of growth promoting or inhibiting signals, (2) evasion of apoptosis or senescence, (3) unlimited replication (bypass cellular crisis), (4) metabolism reprogramming, (5) deregulation of angiogenesis switch, (6) avoidance of immune surveillance, (7) tissue invasion and metastasis (Hanahan & Weinberg, 2000, 2011; Kroemer & Pouyssegur, 2008; Luo et al, 2009). “Enabling characteristics” are the results of genetic alterations and cellular changes from the “emerging hallmarks”. They include: (1) genome instability (DNA damage and DNA replication stress, mitotic stress), (2) cellular stress (proteotoxic stress, metabolic stress, oxidative stress), (3) inflammatory responses. Transformation of human cells is a long process of accumulation of these cellular alterations, which are caused by alteration or deregulation of many genes (Summarized in Figure 2.1.1). The accumulation of these changes is more important than their order with respect to one another. These genes may have functions that promote the transformation process or functions that prevent cells from transformation. Gaining oncogenic functions and losing tumor suppression functions of a normal cell lead to the transformation to cancer cell.
2.1.2 Oncogene

History of oncogene discovery

One hundred years ago, the discovery of Rous sarcoma virus by Dr. Peyton Rous opened the field of tumor virology and led to the discovery of first oncogene Src and molecular mechanisms of cancer formation (Rous, 1911). In 1970s, in an effort to understand the function of viral oncogene Src and Abl, Bishop and Hunter groups found that protein phosphorylation was important in the transformation process (Levinson et al 1978; Eckhart et al 1979). The next year, the tyrosine kinase receptor oncogene, EGFR (epidermal growth factor receptor), was found by Cohen, followed by the discovery of PDGF (platelet-derived growth factor) by Waterfield and Antoniades (Ushiro and Cohen 1980; Waterfield et al 1983; Doolittle et al 1983). In 1982, Weinberg, Wigler and Barbacid groups cloned the first cellular oncogene RAS and discovered a single amino-acid change that could constitutively active RAS protein and lead to transformation (Shih and Weinberg 1982; Goldfarb et al 1982; Pulciani et al 1982; Tabin et al 1982; Reddy et al 1982; Taparowsky et al 1982). Discoveries of RAS, EGFR, PDGF together with v-erbB suggest that oncogenes are components of the normal growth regulatory machinery. Later on, many other oncogenes were found having functions upstream or downstream of these kinases signal transducers.

The role of oncogene in cancer development
Oncogenes initiate and promote transformation during cancer development by regulating the “emerging hallmarks” of cancer. Oncogenes (1) allow cells to proliferate in an uncontrolled and unlimited manner that is resistant to programmed cell death, (2) adapt the cell metabolism to anaerobic, acidic and high stress environment, (3) signal surrounding non-malignant cells to secret growth factors, (4) change environment to favor cancer cell growth, (5) protect cells from killing by immune response, and (6) promote metastasis to distant location.

**Activation of oncogenes**

Proto-oncogenes are genes with normal cellular functions. They can be activated and converted to cancerous oncogenes by mutations or elevated expression, which includes epigenetic regulation, gene amplification and chromosome translocation. For example, mutations in RAS, which is a small GTPase, can constitutively activate the signaling pathways that lead to cell growth and division, independent of extracellular signals (Tabin et al 1982; Reddy et al 1982; Taparowsky et al 1982). Amplification of fibroblast growth factor receptor 1 (FGFR1) in lung squamous carcinoma activates MAPK and AKT pathways and results in cell proliferation independent of ligand binding (Dutt et al 2011; Weiss et al 2010). BCR-ABL oncogene, resulted from translocation of chromosomes 9 and 22, can enhance cell division, inhibit DNA repair and cause chronic myeloid leukemia (CML) (Nowell, 1960). Promoter hypomethylation of
BCL-2 gene causes its overexpression in B-cell chronic lymphocytic lymphomas (Hanada et al 1993).

**Classification of oncogenes**

Oncogenes can be grouped into seven categories based on their cellular functions: (1) growth factors (e.g. c-Sis); (2) tyrosine kinase receptors (e.g. EGFR, FGFR); (3) transcription factors (e.g. MYC); (4) signal transducer, which includes cytoplasmic tyrosine kinases (e.g. Src family), cytoplasmic serine/threonine kinases (e.g. Raf) and signal regulatory proteins (e.g. Ras); (5) chromatin remodelers (e.g. ALL1); (6) apoptosis regulators (e.g. Bcl-2 family); and (7) microRNAs (e.g. miR-17).

**Receptor tyrosine kinases**

Receptor tyrosine kinase (RTK) is a large family of cell surface receptors. It is so called because RTK can catalyze the tyrosines on target proteins by transfer the γ phosphate of ATP to hydroxyl groups of tyrosines. RTK has very important cellular functions in regulating cell cycles, migration, metabolism, survival, proliferation and differentiation. The structure, activation and signaling of members in RTK family are similar (Hunter, 1998). Here we use FGFR as an example of receptor tyrosine kinases.

**2.1.3 FGFR1**

**Structure** There are five members of the FGF receptors: FGFR1(flg), FGFR2(bek, KGFR), FGFR3, FGFR4 and FGFR1L(has no kinase domain). First
FGF receptor was designated as fms-like gene (flg) because it was isolated from a human endothelial cell cDNA library by hybridizing at relaxed stringency using the v-fms oncogene as a probe (Ruta et al 1988). Further analysis of its structure and function revealed that flg was the receptor for acidic FGF (Ruta et al 1989). As illustrated in Figure 2.1.2, the structure of FGFR1 consists of an extracellular domain (EC), a transmembrane (TM) stretch, a juxtamembrane (JM) domain, a kinase domain (KD) and a C-terminal (CT) (Klint, 1999). Three phosphorylation sites, Y653, Y654 and Y730, appear to be involved in the kinase regulation. Phosphorylation of Y766 binds to phospholipase Cr (PLCr).

**Activation**

FGFs are secreted glycoproteins that are sequestered by heparin sulphate proteoglycans to the extracellular matrix (HPSGs). To signal, FGFs are liberated from extracellular matrix and bind to cell surface HPSG and FGF receptors (Harmer et al 2004, Mohammadi et al 2005). Crystal structure shows that they form a stable 2:2 FGF: FGFR complex stabilized by heparin (Schlessinger, 2000). Dimerization of FGFR1 leads to trans-autophosphorylation of their cytoplasmic domains (Schlessinger, 1988). Binding of FGF to FGFR stimulates tyrosine kinase activity was first discovered in NIH3T3 cells (Coughlin et al 1988). In addition, FGFR can also be activated by inhibition of protein tyrosine phosphatases, receptor amplification or mutations, independent of ligand binding.

**Signaling**
Under normal conditions, FGF binding leads to kinase activation, tyrosine phosphorylation of residues in the kinase domain and tyrosine phosphorylation of the C-terminal tail of FGFR. The major effector of FGFR is the adaptor protein FRS2 (FGFR substrate 2). FRS2 constitutively associates with FGFR juxtamembrane domain independent of its activation and gets phosphorylated on several Tyr and Ser/Thr residues following FGFR activation (Ong et al 2000). pFRS2 functions as a docking site for Grb2 and activates MAPK and AKT pathway through SOS and Gab1 respectively, and in turn regulates cell proliferation and cell survival (Eswarakumar et al 2005; Altomare et al 2005). The C-terminal phosphorylation of FGFR phosphorylates and activates Phospholipase Cr (PLCr) through its Src homology 2 (SH2) domain (Peters et al 1992). Activated PLCr hydrolyses phosphatidylinosito-4,5-biphosphate (PIP2) to phosphoatidylinositol-3,4,5-triphosphate (PIP3) and diacylglycerol (DAG), and then activate protein kinase C (PKC) and MAPK pathway (Klint et al 1999).

**Oncogenic functions**

Several types of genetic changes of FGFR1 can cause its oncogenic function: gene amplification, activating mutations, chromosomal translocations, single nucleotide polymorphisms and aberrant splicing. However, constitutive activation by amplification or mutation is the most common in cancer. Oncogenic FGFR1 signaling can cause proliferation, survival, migration, invasion and angiogenesis, and has been found in many types of cancers (Table 2.1.1 and 2.1.2). FGFR1 amplification, results of chromosome amplification, is the most
common genetic alteration in lung squamous carcinoma (~20%) (Weiss et al 2010; Dutt et al 2011), and one of the most common focal amplification in estrogen receptor (ER)-positive breast cancer (10%) (Courjal et al 1997; Jacquemier et al 1994; Reis-Filho et al 2006). Amplification of FGFR1 has also been found in oral squamous carcinoma (Freier et al 2007), ovarian cancer (Gorringe et al 2007), bladder cancer (Simon et al 2001) and rhabdomyosarcoma (Missiaglia et al 2009). Constitutive active FGFR TDII (thanatophoric dysplasia type II) like mutation (FGFR1K656E, FGFR3K650E, FGFR4K645E) can activate Stat1 and Stat3 (Hart et al 2000).

2.1.4 Tumor suppressor genes

Mechanisms of tumor suppressor

In the 1970s and early 1980s, more and more evidence hinted that there is a second type of growth-controlling genes that function to constrain or suppress cell proliferation, so called tumor suppressor genes (TSGs). TSGs negatively regulate cancer development by several different mechanisms: (1) repressing cell growth signaling pathways and inhibiting cell division, (2) blocking cell cycle in response to DNA or cellular damage, (3) initiating programmed cell death, (4) promoting cell adhesion to prevent metastasis, (5) functioning in DNA repair to reduce the risk of genomic instability. TSGs can be classified into three categories based on their functions: “caretaker”, “gatekeeper” and “landscaper” (Kinzler and Vogelstein, 1997; Michor et al., 2004; Ashworth et al, 2011).
“Caretaker” genes maintain genome stability and prevent mutation acquisition. For example, BRCA1 and BRCA2 play critical roles in DNA repair. “Gatekeeper” genes control cell growth, for example, p53 can induce cell cycle arrest or cell death. “Landscaper” genes control the microenvironment for tumor growth. For example, TSP1 has anti-angiogenesis function therefore cutoff the nutrient supply for tumor growth (Volpert et al, 2002).

History of tumor suppressor gene discovery

Due to the recessive nature of TSGs, TSGs discovery is more difficult than oncogenes. In early years, people discovered tumor suppressing elements using transfer approaches, which deliver intact or fragmented chromosomes, such as yeast artificial chromosome (YAC) or bacterial artificial chromosome (BAC), into cancer cells to monitor the inhibition of cell growth (Harris et al 1969; Fournier et al 1977; Stanbridge, 1976, 1992; Murakami 1995,1998, Todd 1996). The discovery of RB gene, the first tumor suppressor, confirmed the “two-hit hypothesis”: cancer is caused by at least two successive mutations, the first hit inactivates one allele of a TSG and the second hit results in the loss of the other allele (Friend et al 1986, Knudson, 1971). In cancer cells, loss of heterozygosity (LOH) is considered the second hit of the “two-hit hypothesis”. Single nucleotide polymorphism (SNP)-based array was widely used to detect genome-wide copy number and LOH status to identify new TSGs. In early 1980s, fluorescence in situ hybridization (FISH) was developed to detect and localize the presence or absence of specific DNA sequences, which provided an alternative way to
identify TSGs. However, FISH is only useful to identify genes with known sequences and can only test one gene at a time. In 1990s, comparative genomic hybridization (CGH) was developed to analyze the copy number changes of TSGs. However, this approach cannot detect chromosome aberrations, such as translocations or inversions. Many genes are silenced by epigenetic events, such as methylation, irrespective of their LOH status in tumors. During last decades, gene expression arrays and promoter methylation arrays were widely used in discovering genes that are differentially expressed in cancer cells and normal cells, which could be potential TSGs.

**Genome-wide RNAi screen**

The elucidation of the mechanism of RNAi by Mello and Fire in 1998 spurred an industry of RNAi biotechnology. RNAi is a naturally occurring gene expression regulation mechanism. In eukaryotes, double stranded interfering RNA targets and degrades complementary mRNA, and results in selective gene silencing. By delivering short hairpin RNA (shRNA), the precursor of siRNA, into cultured cells, we can knock down a specific gene in very short time and at much lower cost, comparing to traditional targeted gene inactivation or knock-out techniques. The development of genome-wide RNA interference (RNAi) library made it possible to conduct high throughput loss of function screens for specific phenotypes with assistance from statistics and bioinformatics (Paddison et al 2004; Berns et al 2004). Pools of shRNA expressing vectors are introduced into cells by transfection or infection. Cells are then selected for certain phenotypes.
In positive screens, cells are selected for ability to transform, or induced drug resistance, etc. Transformed cells or surviving cells were then used for DNA isolation. shRNA sequence were amplified by PCR and sequenced. In negative screens, e.g. drug sensitivity, DNA from cell populations from before and after drug treatment are isolated, sequenced and compared for gene identification (Figure 2.1.3).

Taking advantage of this technique, many new TSGs were identified. For example, REST is found as a tumor suppressor in breast cancer (Westbrook et al 2005); Sfrp1, Numb, Mek1 and Angpt2 are tumor suppressors in lymphoma (Bric et al 2009); XPO4 is a tumor suppressor for liver cancer (Zender et al 2008, etc). From our study, we identified 24 new tumor suppressors for lung squamous carcinomas cancer.

2.1.5 Targeting Drug Development

Besides surgical removal, cancer therapies usually target cancer cells, cancer vasculature, the immune system and even the bone marrow. After decades of extensive studies of the molecular basis underpinning cancer development, therapeutics targeting the hallmarks of cancer cells have been rapidly developed. In cancer cells, signal transduction pathways are frequently altered. Therefore, many drugs, such as tyrosine and threonine kinase inhibitors, are developed directly targeting signal transducting molecules. Imatinib (Glivec), a drug that treats chronic myelocytic leukemia (CML) by inhibiting BCR-ABL, is
the first tyrosine kinase inhibitor approved for clinical use (Druker et al 2001). Gefitinib and Erlotinib, EGFR inhibitors, are subsequently approved for non-small cell lung cancer (NSCLC) treatment (Lynch et al 2004). However, lung squamous cell carcinomas (LSCC) bare a different genetic alternation that is not responsive to EGFR inhibitors. FGFR1 has been found amplified in more than 20% of LSCC patients. It suggests that FGFR1 can be a potential drug target for LSCC treatment.

Drug development research on FGFR signaling pathway are majorly using small-molecule tyrosine kinase inhibitors, but blocking antibodies and ligand-trap approaches are also being developed (Table 2.1.3). Tyrosine kinase inhibitors (TKIs) such as PD173074, ponatinib or TKI258 block FGFR signals and reduce FRS2 tyrosine phosphorylation, causing the dissociation of Grb2 and therefore inhibit ERK and AKT activities. Blocking antibodies can decrease tumor cell proliferation and induce apoptosis same as TKIs, however, it may potentially recruit immune effector cells to the tumor and may lead to antibody-dependent cellular cytotoxicity. Many small molecule inhibitors of FGFR tyrosine kinase activity have been discovered and described in the literature. Many of these are currently in clinical trials in various types of cancers (Table 2.1.3). Our study suggests that the application of FGFR tyrosine kinase inhibitors is not restricted in LSCC. Patients missing FGFR negative regulating TSGs can be selected for treatment with FGFR inhibitors.
Figure 2.1.1

Genetic/epigenetic alteration

Normal cells

Mutations in TSGs/Oncogenes

self-sufficiency in growth signals; insensitivity to growth-inhibitory signals

Hyperplasia

Mutations in TSGs/oncogenes; LOH

Evasion of apoptosis; DNA damage/replication stress; proteotoxic stress; mitotic stress, metabolic stress; oxidative stress

Metaplasia

Deregulation of telomerase

Limitless replication;

Dysplasia

Mutations in cellular metabolism regulator; Deregulation of angiogenesis switch

Glycolysis metabolism; microenvironment change

Carcinoma

eg. Mutations in genes involved in cell adhesion

Immune surveillance; Invasion and metastasis

Metastasis
Figure 2.1.1. Transformation is a multistep process.

Genetic or epigenetic alterations in tumor suppressor genes or oncogenes lead to the acquiring cancer hallmarks. The accumulation of these changes, rather than their order, is most important.
Figure 2.1.2

(Klint et al, 1999)
Figure 2.1.2. Schematic structure of FGFR1. FGFR1 has 6 domains: (1) Extracellular domain contains three immunoglobulin (Ig)-like domains. Some splicing variants have two (II and III) Ig-like domains. (2) Transmembrane domain. (3) Juxtamembrane domain. (4) Kinase domain with a short kinase insert. (5) C-terminal tail.
Figure 2.1.3

Pools of shRNA virus infection Select for certain phenotype

Isolate DNA
PCR amplify the shRNA sequence

Positive screen

negative screen

Sequence identify shRNA
Figure 2.1.3. Strategy of Pooled RNAi Library Screen. Pools of shRNA expressing vectors are introduced into cells by transfection or infection. Cells are then selected for certain phenotypes. In positive screens, cells are selected for the ability to transform, drug resistance etc. Transformed cells or survived cells were then used to isolate DNA. shRNA sequence were amplified by PCR and sequenced. In negative screens, e.g. drug sensitivity screen, DNA from cell populations from before and after drug treatment are isolated, sequenced and compared for genes identification.
Table 2.1.1. Genetic alterations in FGFR1 and related cancer

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplification</td>
<td>Lung squamous carcinoma (~20%), breast cancer (10%), ovarian cancer (~5%), bladder cancer (3%) and rhabdomyosarcoma (3%)</td>
<td>Courjal 1997, Jacquemier 1994, Gorringe 2007, Simon 2001, Missiaglia 2009</td>
</tr>
<tr>
<td>Mutation</td>
<td>P252S: Melanoma</td>
<td>Lin, 2008; Ruhe, 2007</td>
</tr>
<tr>
<td></td>
<td>P252T, V664: Lung cancer</td>
<td>Davies, 2005; Greenman, 2007</td>
</tr>
<tr>
<td></td>
<td>K656E: Glioblastoma</td>
<td>Network, 2008</td>
</tr>
<tr>
<td></td>
<td>S125L: Breast</td>
<td>Greenman, 2007</td>
</tr>
<tr>
<td>Translocation</td>
<td>Stem cell leukaeia and lymphoma syndrome (or EMS), chronic myeloid leukaemia (rare)</td>
<td>Xiao 1998, Roumiatsev 2004</td>
</tr>
</tbody>
</table>
Table 2.1.2 Altered expression of FGFR1 in cancer

<table>
<thead>
<tr>
<th>Cancer type</th>
<th>Up-regulated (references)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain</td>
<td>Kroemer 2008; Luo 2009; Colotta 2009; Stratton 2009</td>
</tr>
<tr>
<td>Head and neck</td>
<td>Olsen 2004; Zhang 2006; Ornitz 1996; Schlessinger 2000; Lew 2009; Chen 2007; Mohammadi 1996</td>
</tr>
<tr>
<td>Sarcoma</td>
<td>Lu 2008; Parsa 2008; Murakami 2008</td>
</tr>
<tr>
<td>Soft tissue sarcoma</td>
<td>Beenkan 2009; Chin 2006; Toyokawa 2009</td>
</tr>
<tr>
<td>Thyroid</td>
<td>Ray 2004; Bernard-Pierrot 2008</td>
</tr>
<tr>
<td>Breast</td>
<td>Grand 2004; Trudel 2006; Keats 2003; zingone 2010; Cha 2008; Ezzat 2002; Itoh 1994; Cha 2009</td>
</tr>
<tr>
<td>Liver</td>
<td>Jang 2001</td>
</tr>
<tr>
<td>Stomach</td>
<td>Martinez-Torrecuadrada 2008; Hernanda 2008</td>
</tr>
<tr>
<td>Testis</td>
<td>Dellacono 1997</td>
</tr>
<tr>
<td>Colon</td>
<td>Memarzadeh 2007</td>
</tr>
<tr>
<td>Uterus</td>
<td>Joyce 2009; Abdel-Rahman 2008</td>
</tr>
<tr>
<td>Ovary</td>
<td>Darby 2006; Murphy 2010</td>
</tr>
<tr>
<td>Skin</td>
<td>Cho 2004; Tannheimer 2000</td>
</tr>
<tr>
<td>Leukaemia/MPD/lymphoma</td>
<td>Cho 2007; Fogarty 2007; Knights 2010; Byron 2010; Claudio 2007</td>
</tr>
<tr>
<td>Drug name</td>
<td>Company</td>
</tr>
<tr>
<td>-----------</td>
<td>---------</td>
</tr>
<tr>
<td><strong>Small molecular tyrosine kinase inhibitors</strong></td>
<td></td>
</tr>
<tr>
<td>SUS402</td>
<td></td>
</tr>
<tr>
<td>PD173074</td>
<td></td>
</tr>
<tr>
<td>Ponatinib(AP24534)</td>
<td>Ariad Pharmaceuticals</td>
</tr>
<tr>
<td>Cediranib(AZD2171)</td>
<td>AstraZeneca</td>
</tr>
<tr>
<td>Dovitinib(TKI258)</td>
<td>Novartis</td>
</tr>
<tr>
<td>Intedanib(BIBF1120)</td>
<td>Boehringer Ingelheim</td>
</tr>
<tr>
<td>Brivanib(BMS-582,664)</td>
<td>Bristol-Myer Squibb</td>
</tr>
<tr>
<td>E7080</td>
<td>Eisai</td>
</tr>
<tr>
<td>TSU-68 (Orantinib)</td>
<td>Taiho Pharmaceutical</td>
</tr>
<tr>
<td>AZD4547</td>
<td>Astra Zeneca</td>
</tr>
<tr>
<td>BGJ398</td>
<td>Novartis</td>
</tr>
<tr>
<td>Masitinib (AB1010)</td>
<td>AB Science</td>
</tr>
<tr>
<td>SU6668</td>
<td>Sugen</td>
</tr>
<tr>
<td>Pazopanib (GW786034)</td>
<td>Glaxo Smith Kline</td>
</tr>
<tr>
<td>RO438596</td>
<td>Roche</td>
</tr>
<tr>
<td><strong>FGFR antibodies and FGF ligand traps</strong></td>
<td></td>
</tr>
<tr>
<td>IMC-A1</td>
<td>ImClone</td>
</tr>
<tr>
<td>PRO-001</td>
<td>ProChon Biotech</td>
</tr>
<tr>
<td>R3Mab</td>
<td>Genentech</td>
</tr>
<tr>
<td>TA6</td>
<td>Genentech</td>
</tr>
<tr>
<td>FP-1039</td>
<td>Five Prime Therapeutics</td>
</tr>
</tbody>
</table>

(ref: Greulich & Pollock, 2011; Knights & Cook, 2010)
CHAPTER II

Identification of Squamous Cell Lung Cancer Tumor Suppressor Genes that Negatively Regulate FGFR1 Signaling
Abstract

A goal of contemporary cancer research is to identify the genes responsible for neoplastic transformation. Here we describe a functional genomics approach to discover new tumor suppressor genes (TSGs). Cells that are immortalized but non-tumorigenic were stably transduced with pools of short hairpin RNAs (shRNAs) and tested for their ability to form tumors in mice. ShRNAs in any resulting tumors were identified by sequencing to reveal candidate TSGs, which were then validated experimentally and by expression analysis of human tumor samples. Using this approach, we identified and validated 24 TSGs that were down-regulated in >70% of human lung squamous cell carcinomas (LSCCs). Amplification of fibroblast growth factor receptor 1 (FGFR1), which aberrantly increases FGFR signalling, is the most common known genetic alteration in LSCC. Remarkably, we found that 17 of the 24 TSGs encode negative regulators of FGFR signalling. Knockdown of 11 TSGs in immortalized human bronchial epithelial cells results in increased total or phosphorylated levels of FGFR1 leading to transformation and sensitivity to FGFR1 pharmacological inhibition. Our results indicate that many LSCCs without FGFR1 amplification or activating mutations may respond to FGFR1 inhibitors and in a predictable manner.
Introduction

The conversion of a normal cell to a cancer cell is a stepwise process that involves the activation of genes that promote cancer, oncogenes, and the inactivation of genes that protect cells from cancer, such as tumor suppressor genes (TSGs). A goal of contemporary cancer research is to identify all genes responsible for neoplastic transformation. For example, the identification of new TSGs has important implications for both diagnosing and treating cancer.

TSGs are defined by a series of characteristic properties (Weinberg 2006). First, TSGs are inactivated in certain human tumor types either by deletion, mutation or a transcriptional repression mechanism including epigenetic silencing. Epigenetically silenced TSGs bear several hallmarks including repressive histone modifications and hypermethylated CpG-rich promoter regions (Esteller 2007). Second, inactivation of a TSG in an appropriate cell type promotes one or more growth-related properties such as immortalization, proliferation, or tumorigenesis. Conversely, over-expression of a TSG in an appropriate cell type has one or more growth-suppressive effects such as inhibition of cell cycle progression or the induction of apoptosis or senescence.

One of the ongoing efforts to identify the genes responsible for tumor formation is partial or complete sequencing of cancer genomes (Mardis and Wilson 2009). Although this approach has and will continue to reveal genes involved in tumorigenesis, there are specific challenges for discovery of TSGs solely by sequence analysis of cancer genomes. First, the genomes of cancer
cells are error-prone (Loeb 2001) and thus many of the mutations found in cancer cells are not causally related to the transformed phenotype. Second, as mentioned above, in many instances TSGs are inactivated by epigenetic silencing rather than a mutational event, and thus would not be identified by conventional genome-sequencing methods.

Fibroblast growth factors receptors (FGFR) are cell surface tyrosine kinase receptors for fibroblast growth factors. They play crucial roles in development, cell proliferation, survival and migrations. FGFRs family consists of four members (FGFR1-4) and one non-catalytic functional protein (FGFR1L). FGFR signaling can be activated by (1) ligand induction (2) FGFR amplification or activating mutations. Activation of FGFR1 leads to kinase activation, tyrosine phosphorylation of residues in the kinase domain and the C-terminal tail. The major effector of FGFR signaling is the adaptor protein FRS2 (FGFR substrate 2). FRS2 is constitutively associated with FGFR independent of its activation and gets phosphorylated on several Tyr residues following FGFR activation (Ong et al 2000). pFRS2 functions as a docking site for Grb2 and activates MAPK and AKT pathway through SOS and Gab1 respectively, and in turn regulates cell proliferation and cell survival (Eswarakumar et al 2005; Altomare et al 2005). The C-terminal phosphorylation of FGFR activates protein kinase C (PKC) and MAPK pathways by phosphorylating and activating Phospholipase C r (PLCr) through its Src homology 2 (SH2) domain (Peters et al 1992; Klint et al 1999). In cancer, FGFR signaling pathway is usually constitutively activated by amplification or
mutation. Oncogenic FGFR1 signaling can cause proliferation, survival, migration, invasion and angiogenesis. FGFR1 amplification is the most common genetic alteration in lung squamous carcinoma (~20%) (Weiss et al 2010; Dutt et al 2011), and one of the most common focal amplifications in estrogen receptor (ER)-positive breast cancers (10%) (Courjal et al 1997; Jacquemier et al 1994; Reis-Filho et al 2006). Amplification of FGFR1 has also been found in oral squamous carcinoma (Freier et al 2007), ovarian cancer (Gorringe et al 2007), bladder cancer (Simon et al 2001) and rhabdomyosarcoma (Missiaglia et al 2009). Amplification of FGFR1 activates MAPK but not AKT pathway (Weiss et al 2010).

Drug development research on inhibiting FGFR signaling pathway are primarily using small-molecule tyrosine kinase inhibitors, but blocking antibodies and ligand-trap approaches are also being developed. Tyrosine kinase inhibitors (TKIs) such as PD173074, ponatinib or TKI258 block FGFR signaling and reduce FRS2 tyrosine phosphorylations, cause the dissociation of Grb2 and therefore decrease in ERK and AKT activity.

Functional genomics approaches for TSG discovery are highly complementary to cancer genome sequencing efforts. In this regard, several recent studies have described in vivo RNA interference (RNAi) screens in mice that have successfully identified TSGs involved in liver cancer (Zender et al 2008) and genes that suppress lymphoma progression (Bric 2009, Meacham 2009). These studies were performed using a limited collection of shRNAs
targeting genes that had been previously implicated or suspected to be involved in cancer. Here we describe and demonstrate a large-scale RNAi-based functional approach for TSG discovery that is unbiased with regard to either the genes being targeted or cancer type. From the screen, we identified 24 TSGs in lung squamous carcinoma, suggesting that they can be used as new biomarkers in cancer diagnosis. Among these 24 TSGs, 17 can negatively regulate FGFR1 signaling pathway and 12 regulate expression or activation levels of the receptor. Lung small air way cells are sensitive to FGFR1 inhibitors when knocking down individual TSGs. Therefore, we propose that the application of FGFR1 inhibitor can be expanded beyond FGFR1 amplification patients. Expression level of TSGs can be used as marker for drug selection and to predict drug responsiveness.
RESULTS

A large-scale shRNA screen for the identification of candidate TSGs

To identify TSG candidates, we performed a large-scale shRNA screen for
tumorigenesis in mice, which is summarized in Figure. 2.2.1.A and is discussed
below. We used NIH 3T3 fibroblasts, which are immortalized but not transformed
and can be rendered tumorigenic by a wide range of oncogenic events (Stacey
and Kung, 1984; Ikawa et al 1988, Smith et al 1990; Velu et al 1987; Di Fiore et
al 1987).

Before initiating the primary screen, we performed a reconstruction
experiment to determine a minimal number of transformed NIH 3T3 cells required
for tumorigenesis in our xenograft system. In this approach, 1000, 100 or no
Kras-transformed NIH 3T3 cells were added to non-transformed NIH 3T3 cells to
achieve a total of 1x10^6 cells, which were then injected into the flanks of nude
mice. As expected, no tumors formed in the sample lacking Kras NIH 3T3 cells.
However, tumors were observed in samples containing either 100 or 1000 Kras
NIH 3T3 cells (Figure. 2.2.1.B), indicating that 100 transformed NIH 3T3 cells
(and likely fewer) are sufficient to initiate a detectable tumor in a mouse
xenograft.

For the primary screen, a mouse shRNA library comprising 62,400
shRNAs directed against ~28,000 genes (Silva et al 2005) was divided into 10
pools, which were packaged into retrovirus particles and used to stably transduce
mouse NIH 3T3 cells. For each of the 10 stable cell populations, 1x10^6 cells were
injected subcutaneously into the flank of a nude mouse. Thus, each of the 6240 shRNAs in each pool was present in ~160 cells, which is greater than the 100 transformed NIH 3T3 cells found in the reconstruction experiment to be required for tumor formation. We found that seven of the 10 shRNA pools gave rise to detectable tumors. After two weeks, tumors were dissected and genomic DNA was extracted. To identify the candidate shRNAs, the shRNA region of the transduced virus was PCR amplified, cloned and sequenced (Gazin et al 2007). We recovered 41 independent shRNAs from the seven tumors (Table 2.2.1).

To validate the candidates, single shRNAs directed against each gene were stably transduced into NIH 3T3 cells, and the knockdown (KD) cell lines were generated. We tested the ability of the NIH 3T3 KD cells lines to form tumors in vivo following subcutaneous injection into nude mice. Figure 2.2.1.C shows that all 33 shRNAs promoted tumor formation, whereas a control non-silencing shRNA, as expected, did not. Quantitative real-time RT-PCR (qRT-PCR) confirmed in all cases that expression of the target gene was decreased in the corresponding KD cell line (Figure 2.2.2).

Of the 33 candidate TSGs, we could identify unambiguous human homologs in 28 cases, which are summarized in Table 2.2.2. These 28 human TSG candidates are involved in diverse biological processes including signal transduction, transcriptional regulation, cell growth/metabolism and DNA/RNA metabolism. Significantly, several of the genes such as IGF2R (Probst et al 2009, O’Gorman et al 1999; Motyka et al 2000; Nam et al 2010), SEMA3B (Castro-
Rivera et al 2008; Rolny et al 2008; Nair et al 2007), and STK11 (also known as LKB1) (Mehenni et al 2005; Mehenni et al 2005) are well documented TSGs, validating the overall experimental strategy.

**Down-regulation of candidate TSGs in human lung squamous cell carcinoma samples**

To gain insight into the tumor types in which the candidate TSGs may play a role, we performed a series of cancer database searches. Analysis of the 28 human TSG candidates using the NCBI SKY/M-FISH and CGH database (National Center for Biotechnology Information Spectral Karyotyping, Multiplex Fluorescence In Situ Hybridization and Comparative Genomic Hybridization; ref. 23) revealed that all the genes have been found to harbor deletions in either one or both copies in multiple cancer types (Table 2.2.3). Furthermore, a search of the Oncomine cancer profiling database (Rhodes et al 2007) revealed that, with the exception of PKD1L3, all the candidate TSGs are significantly down-regulated in a variety of cancers compared to normal tissues (Table 2.2.4).

Notably, a large fraction of the candidate TSGs were down-regulated in lung cancers. To confirm and extend this finding, we analyzed expression of the 28 genes in a series of human lung squamous cell carcinoma (hLSCC), lung adenocarcinoma and normal lung samples. The qRT-PCR analysis of Figure 2.2.3 shows that 24 of the 28 candidate TSGs were down-regulated greater than 2-fold in ≥ 70% of the hLSCC samples analyzed. In several
instances, the candidate TSGs were down-regulated greater than 10-fold in all (or almost all) hLSCC samples (e.g., \textit{ANGPT1}, \textit{FLNA}, \textit{IGF2R}, \textit{NME4}, \textit{SEMA3B} and \textit{STK11}). Six genes were also down-regulated greater than 2-fold in \(\geq 70\%\) of the lung adenocarcinoma samples (Figure 2.2.4).

**Knocking down TSGs activates FGFR1 signaling pathway**

It has been shown that FGFR1 amplification is the most common genetic alteration in LSCC (~20%) (Weiss, 2010; Dutt, 2011). In order to test whether knocking down TSGs can activate FGFR1 signaling pathway, we first knocked down our candidate TSGs using gene specific shRNAs in SA (small airway) cells, which are immortalized but non-transformed human lung small airway epithelia cells. Then we examined the activation levels of FRS2, which is phosphorylated upon FGFR1 activation, in 24 TSG KD cell lines. As shown in Figure 2.2.5, we found 17 out of 24 TSG KD cell lines increased pFRS2 to the level comparable to NCI-H520, while the total FRS2 expression levels were same as non-silencing control. We further analyzed the phospho-FGFR1 and total-FGFR1 levels. We found that knocking down 7 genes (FLNA, GAPVD1, MYD88, PIGH, PTGIS, SFRS9, SPAST) increased the total FGFR1 levels, and knocking down 5 genes (ANGPT1, DAPP1, FPR3, NME4, PTPN4) increased pFGFR1 levels without elevated tFGFR1. However, Y653 phosphorylation of FGFR1 was not increased in FLNA and MYD88 KD cell lines, although the total FGFR1 is increased. 6 genes (CDK5R1, DDX52, SDF2L1, SPOP, STK11, TXNRD1) increased pFRS2
level without elevated phospho- or total-FGFR1 levels. It is known that pFRS2 functions as a docking site for Grb2 and activates SOS and Gab1, and leads to activation of MAPK and AKT signaling pathways (Eswarakumar et al 2005; Altomare et al 2005). We checked the levels of activated Akt and Erk, and found that, except PTGIS and SPAST, all TSGs KD cell lines with increased pFRS2 levels had increased pErk levels. DAPP1 could affect both MAPK and AKT signaling pathways. Four genes (IGF2R, LSM8, MAP1A, SEMA3B) inhibited pErk level independent of FGFR1 signaling. KD IGF2R didn’t activate FGFR signaling validated our results because IGF2R is a tyrosine kinase receptor independent of FGFR signaling. Based on these results, we proposed a model in Figure 2.2.6 showing possible mechanisms that the TSGs function by negatively regulating FGFR1 signaling at different steps. To further understand the relationship between our TSGs and FGFR1 signaling pathway, we knocked down FGFR1 in NCI-H520 cell line, using FGFR1 specific shRNA, and analyzed the mRNA levels of 24 TSGs. We found de-repression of 3 genes (ANGPT1, DAPP1, TXNRD1) after FGFR1 knock down (Figure 2.2.7). These results suggest that, these genes may have negative feedback effects on regulating FGFR1 signaling pathway.

Knocking down TSGs induces transformation of lung epithelia cells

Next, we tested whether knocking down TSGs has the same transformation function as FGFR1 amplification in lung epithelia cells. First of all,
we tested whether overexpression FGFR1 could transform mouse and human cell lines. As shown in Figure 2.2.8.A, 1 million FGFR1 overexpressed NIH3T3 cells formed a tumor when subcutaneously injected into nude mice, while no tumor formed in the empty vector control. This result is consistent with a previous report that FGFR1 can transform NIH3T3 (Hart et al 2000), and further confirmed that our TSG candidates from NIH3T3 screen might be related to FGFR1 oncogenic pathway. Then we tested whether overexpression of FGFR1 can transform human SA cells. We found 5 million FGFR1 overexpressed SA cells formed a visible tumor in nude mice, while no tumor formed in the empty vector control (Figure 2.2.8.A).

Then we want to test whether the activation of FGFR1 pathway by knocking down our TSG candidates can cause transformation of human SA lung cells. In Figure 2.2.8.B, we performed in vitro soft agar colony formation assay, and found 14 out of 17 FGFR1 regulating genes and 4 FGFR1 signaling independent genes can form colonies in soft agar (Figure 2.2.8.B). Then we injected knock down cell lines, that formed colonies in soft agar, into the flanks of nude mice and found they can all form tumors in vivo (Figure 2.2.8.C). Our results suggest that inhibiting these TSGs in normal cells can induce transformation by activating FGFR1 oncogenic signaling.
SA cells with TSGs knocking down are sensitive to FGFR inhibitor

FGFR1 is a tyrosine kinase receptor. Many small molecules have been generated to inhibit the FGFR1 tyrosine kinase (PD173074, ponatinib, BIBF1120 etc) and they have been shown to inhibit cancer cell growth in vivo and in vitro (Weiss et al 2010; Joseph et al 2012; Hilberg et al 2008). Some of these inhibitors are in the process of clinical trials for cancer treatment. Ponatinib is an oral tyrosine kinase inhibitor, and it is now in a pivotal phase II trial in patients with chronic myeloid leukemia (CML) and acute lymphoblastic leukemia (ALL), and it has been shown to inhibit kinase activity of FGFR1. We wanted to test whether our TSGs KD cell lines are sensitive to Ponatinib. We treated cells with 500nM ponatinib for three days and found TSG KD cell lines that have increased pFGFR1 levels are sensitive to Ponatinib compared to non-silencing control, shown in Figure 2.2.9.A. Our results suggest that the application of Ponatinib in lung squamous carcinoma might not be restricted to the patients with FGFR1 amplification. Lung cancer patients with repressed levels of these TSGs may also be good candidates for Ponatinib treatment. To further confirm this, we first tested the Ponatinib sensitivity of A427, which is a lung squamous carcinoma cell line with normal copy of FGFR1. A427 showed similar sensitivity to Ponatinib compare to FGFR1 amplified cell line NCI-H520, while a control cell line A549 with normal level of FGFR1 showed more resistance (Figure 2.2.9.C). Then, we compared the mRNA level of 24 TSG candidates in A427 and SA or HBEC cell lines. SA and HBEC are both immortalized but non-transformed lung epithelial
cells. And we found 3 genes (STK11, MYD88, DAPP1) are highly repressed in A427 compare to SA or HBEC (Figure 2.2.9.B). So we confirmed that patients with normal FGFR1 status but TSG deficient might be sensitive to FGFR1 inhibitor. Therefore, by examining the expression levels of TSGs in lung cancer patients, we could predict drug responsiveness and treatment outcome.

**Candidate TSG promoter hypermethylation in hLSCC samples**

To determine whether down-regulation of any of the candidate TSGs was a result of epigenetic silencing due to promoter hypermethylation, we performed comprehensive DNA methylation analysis to quantify promoter DNA methylation (Bibikova et al 2009). Genomic DNA was isolated from paired primary hLSCC and normal lung squamous cell samples from 18 individuals, subjected to bisulfite conversion, and hybridized to a genome-wide methylation array containing >27,500 individual CpG sites spanning ~14,500 promoters. To verify the quality of the methylation analysis, we first confirmed that well-established tumor suppressor genes, whose promoters are known to be hypermethylated in lung cancers. PTGIS (Stearman et al 2007; Gray et al 2009), showed significantly higher promoter hypermethylation in lung cancer relative to normal samples. The methylation heat map of Figure 2.2.10.A and histogram of Figure 2.2.10.B show that among the 26 candidate TSGs analyzed (GAPVD1 and PKD1L3 were absent from the array), nine displayed a statistically significant increase in promoter hypermethylation in lung cancer samples relative to normal samples.
samples. These results indicate that epigenetic silencing due to promoter hypermethylation represents one mechanism by which the candidate TSGs are down-regulated in human lung cancers.

**Over-expression of candidate TSGs inhibits growth of hLSCC cells**

To test whether ectopic expression of these candidate TSGs could inhibit cell proliferation, a representative subset were cloned into a retroviral vector co-expressing green fluorescent protein (GFP) and transduced into NCI-H520 hLSCC cells. Cells were selected with puromycin for 4 weeks and analyzed in a colony suppression assay. The results, shown in Figure 2.2.11.A and quantified in Figure 2.2.11.B, demonstrate that over-expression of all 16 genes tested reduced colony formation to varying extents. We then asked whether ectopic expression of the candidate TSGs could inhibit tumor growth in mouse xenografts. Retroviral vectors expressing 12 of the genes were transduced into NCI-H520 cells, and 24 h later GFP-positive cells were isolated by fluorescence-activated cell sorting and injected subcutaneously into nude mice. Figure 2.2.11.C shows that ectopic expression of all of the genes tested markedly suppressed tumor growth compared to the vector control.

TSGs have been found to inhibit tumor growth through a variety of mechanisms including negatively regulating cell cycle progression, inducing senescence or promoting apoptosis (Sheer et al 2004). To gain insight into the mechanism-of-action of the candidate TSGs, GFP-positive NCI-H520 cells
ectopically expressing 16 of the TSG candidates were analyzed for their ability to induce apoptosis or senescence. We found that ectopic expression of 12 genes induced apoptosis in NCI-H520 cells as evidenced by cleavage of Caspase 3. (Figure 2.2.12.A and B). Furthermore, ectopic expression of six candidate TSGs induced senescence in NCI-H520 cells as evidenced by senescence-associated β-galactosidase staining (Figure 2.2.12.C and D).
Material and Methods

Cell lines and culture

NIH 3T3, K:Molv NIH 3T3 (referred to here as Kras NIH 3T3) and NCI-H520 cells (ATCC) were grown in DMEM or RPMI medium, respectively, supplemented with 10% fetal bovine serum (Atlanta Biologicals) and 1X penicillin/streptomycin (Invitrogen). SA cells are obtained from Scott Randell at University of North Carolina at Chapel Hill School of Medicine, cells were grown in BEGM medium from Lonza (CC-3170).

Large-scale in vivo RNAi screen

To determine the number of transformed cells that could potentially seed a tumor in our experimental system, 0, 100 or 1000 Kras-transformed NIH 3T3 cells were mixed with non-transformed NIH 3T3 cells to achieve a final population of 1x10^6 cells, which were then injected into the flanks of nude mice (n=3). Tumor dimensions were measured every 1-2 days from the time of appearance of the tumors, and tumor volume was calculated using the formula \( \frac{\pi}{4} \times \text{length} \times \text{width}^2 \).

The mouse shRNA\(^{\text{mir}}\) library (release 2.16; Open Biosystems) was obtained through the University of Massachusetts Medical School (UMMS) RNAi Core Facility. Ten retroviral pools, each comprising ~6240 shRNA clones, were generated with titers of ~2.6x10^5 pfu/ml, as previously described (Gazin et al 2004). Briefly, 1x10^6 NIH 3T3 cells were transduced at a multiplicity of infection...
(MOI) of 0.2 with the retroviral stocks in 100 mm plates, and 2 days later selected for resistance to puromycin (1.5 µg/ml) for 7 days. For each pool, 1x10^6 cells were then injected subcutaneously into the flank of a BALB/c nu/nu mouse (Taconic Farms). Tumors were excised and dissected, and genomic DNA was extracted. To identify the candidate shRNAs, the shRNA region of the transduced virus was PCR amplified, cloned and sequenced (Gazin et al 2004). To validate the candidates, 2x10^6 NIH3T3 cells with individual shRNA KD (Table 2.2.5) were injected subcutaneously into nude mice (n=2); mice were sacrificed and photographed when the tumor size reached 1.5-2 cm. All animal experiments were performed in accordance with institutional and national guidelines and regulations and approved by the Institution Animal Care and Use Committee.

**In vitro transformation assay**

SA cells were infected with a lentivirus carrying an individual shRNA at high MOI. 1x10^4 cells were seeded in 6-well plates in soft agar. Colonies are stained with crystal violet and scored after 4 weeks.

**Identification of human homologs**

Human homologs were identified using NCBI’s HomoloGene (www.ncbi.nlm.nih.gov/homologene). Human genes were considered to be homologs if they shared greater than 50% identity at the protein level with the mouse gene.
Analysis of gene expression in human lung cancer samples

Total RNA from 10 normal lung, 10 adenocarcinoma and 27 hLSCC samples were obtained from the UMMS Cancer Center Tissue Bank. qRT-PCR was performed using primers listed in Table 2.2.6.

Analysis of promoter methylation in human lung cancer samples

Genomic DNA was extracted from ~200-300 mg of fresh-frozen primary hLSCC tissue and matching normal lung squamous cell tissue from 18 individuals (obtained from The Prince Charles Hospital and University of Queensland Thoracic Research Center). High molecular-weight DNA was purified using the DNeasy Tissue Kit (Qiagen). Cleaned genomic DNA (1 µg) was bisulfite-converted using the EZ DNA Methylation Kit (Zymo Research), hybridized to multisample Illumina HumanMethylation27 v1.0 BeadChips. Images were processed, data extracted and control probes checked using the BeadStudio Methylation Module (Illumina) using default settings. All CpGs located on the X-chromosome were removed to avoid gender-specific bias.

Percent methylation was calculated by measuring the intensity ratio of methylated to unmethylated DNA, giving a $\beta$-value between 0 (100% methylation) and 1 (0% methylation). For each gene, the fold change was calculated for each individual by dividing the percent methylation in the tumor sample by the normal control. A gene with a fold change $\geq 1.5$ was considered
hypermethylated (i.e., ≥50% increase in methylation) in hLSCC compared to normal lung, whereas a gene with a fold change <1.5 was considered to have no change in methylation status. To evaluate the significance of any observed number of hypermethylation events n for each gene, we estimated the probability of obtaining the value n or more in random data drawn according to a null model (i.e., individuals are independent and the hypermethylation rate is uniformly distributed among genes). Statistical analyses were performed using R (Ihaka and Gentleman 1996). The p-value for each gene was estimated based on binomial distribution.

**Ectopic expression of TSGs and analysis of tumor suppressor activity**

TSG cDNAs (Table 2.2.7) were cloned, by PCR (using primers listed in Table 2.2.6) followed by restriction enzyme digestion, into MSCV PIG (Puro-IRES-GFP) (Addgene). For some genes, a 3xFlag tag sequence was incorporated into the primers for cloning in-frame with the target gene. Murine stem cell viruses (MSCVs) carrying TSGs were packaged in 293T cells and used to infect NCI-H520 cells. For the colony suppression assay, NCI-H520 cells (5x10^3) were plated in 6-well plates, infected with retroviruses at a multiplicity of infection (MOI) of 2 (or, in the case of IGF2R, transfected with a plasmid expressing IGF2R or empty vector), puromycin selected for 4 weeks and stained with crystal violet. For the mouse tumorigenesis assays, NCI-H520 cells were infected with retroviruses and 2 d later FACS sorted for GFP-positive cells. 2x10^6
viable GFP-positive cells were injected subcutaneously into nude mice, and
tumor dimensions were measured and calculated every 3-4 d [using the formula
\((\text{length} \times \text{width}^2) \times (\pi/4)\)]. For apoptosis assays, NCI-H520 cells were infected
with retroviruses for 24 h, puromycin selected for 24 h, and recovered in fresh
medium without puromycin for 3 d. Immunoblots were probed using an antibody
against cleaved caspase-3 (Asp175) (Cell Signaling Technology) and quantified
using Image J software (NIH). For cellular senescence assays, NCI-H520 cells
were infected with retrovirus at a MOI of 2, and the medium was changed 24 h
after infection. Cells were stained for \(\beta\)-galactosidase and visualized as
previously described (Wajapeyee et al 2008).
DISCUSSION

In this report, we have described an experimental strategy for the systematic identification of candidate TSGs. Based upon this approach, and subsequent functional analyses, we have identified human genes that have all the expected properties of TSGs including: promotion of tumorigenesis following inactivation, frequent deletion or down-regulation in human tumor samples, and inhibition of colony formation and/or tumor growth upon over-expression. Significantly, several of the genes we identified are well documented TSGs (e.g., STK11 and SEMA3B), thus validating the experimental approach. Although we performed the primary screen by analyzing tumor formation in mice, a viable alternative would have been to first analyze growth in soft agar. We felt that a compelling advantage of the tumorigenesis assay is that it is a more stringent test of the transformed phenotype, and would therefore reduce the number of false positives obtained in the primary screen. However, an advantage of using the soft agar assay as a primary screen is that it is likely to be more economical. The relative advantages of the two approaches should be considered when designing other similar, large-scale screens.

Our reconstruction experiment indicates that that, in principal, the design of the primary screen was sufficient to analyze all 62,400 shRNAs in the library. However, for several reasons we are reluctant to conclude that the screen was saturating and that all possible candidate TSGs were identified. In particular, transformed cells may have different growth rates depending upon the particular
shRNA present in the cell. Therefore, some shRNAs that induce transformation events may also cause cells to grow slowly and, as a consequence, those shRNAs may be relatively under-represented in the resulting tumors. In addition, we have found that not all shRNAs are equally efficacious and thus some candidate TSGs may be missed due to sub-optimal knockdown efficiency.

Nonetheless, our experiments successfully identified a large number of candidate TSGs and therefore we were not compelled to search more exhaustively for further candidates. However, additional candidate TSGs could be identified by, for example, injecting each pool into more mice and obtaining a greater number of shRNA sequences from each tumor using either conventional approaches or massively parallel short sequencing. In any case, the possibility that there are additional candidate TSGs that were not identified in our screen does not diminish the fact that the approach successfully resulted in the discovery of a large number of new TSGs.

As discussed above, there have been several previous studies in which TSGs have been identified through in vivo RNAi screens using a limited collection of shRNAs that were selected based upon their previous implication in cancer (Bric et al, 2009; Meacham et al, 2009; Zender et al, 2008). The substantially smaller number of shRNAs screened in these studies increases the likelihood of complete coverage of the collection used. However, a significant disadvantage of this smaller scale approach is that the collection of shRNAs is biased and limited, precluding unexpected results from being obtained.
In this regard, several of the genes we found were previously unrecognized TSGs, with well-established roles in processes other than growth suppression. For example, \textit{MYD88} is a gene best known for its role in innate immunity (reviewed in Arancibia et al 2007) and \textit{SFRS9} encodes a pre-mRNA splicing factor (Screaton et al 1995). Furthermore, the TSGs we identified that negatively regulate FGFR1 signaling could provide new hints for functional studies of these TSGs. For example, PTPN4 is a known tyrosine phosphatase (Wu et al 2006) and can dephosphorylate the immunoreceptor tyrosine-based activation motifs (ITAMs). We showed that knocking down PTPN4 increased phospho level of FGFR1, suggesting that FGFR1 might be a potential substrate of PTPN4. GAPVD1 has been shown to mediate ubiquitination and degradation of EGFR (Xiong, 2007). From our experiments, we showed that KD GAPVD1 increased level of total FGFR1, therefore we think GAPVD1 might also mediate ubiquitination and degradation of FGFR1. It is widely accepted that FGFR signals promote angiogenesis, and people found FGFR1 regulates angiogenesis through cytokines interleukin-4 and pleiotrophin in embryos. However, little is known about FGFR1 induced angiogenesis in tumors. We found that Angpt1, a known anti-angiogenesis protein, is repressed by FGFR1 overexpression and knocking down ANGPT1 can further activate FGFR1 signaling, thus form a negative feedback regulation of FGFR1. It might be one of the mechanisms of angiogenesis in FGFR1 amplified lung cancer.
Our results highlight the power of unbiased, large-scale functional screening for cancer gene discovery. The shRNA screening approach we describe is a general strategy that can be used to discover new TSGs in any tumor type. A wide range of immortalized but non-transformed cell lines derived from different cell types are available and would presumably lead to the identification of distinct TSGs. Appropriate cell lines have been isolated from human tissue, such as MCF10A mammary epithelial cells, or experimentally derived from primary human cells using defined oncogenes (see, for examples, Hahn et al 1999, Gupta et al 2005, Sato et al 2006, Radulovich et al 2008, Qian et al 2005).
Figure 2.2.1

A

Mouse NIH 3T3 fibroblasts
Transduce with mouse shRNA library pools
Puromycin select for 7 days
Inject cells subcutaneously into nude mice
Allow tumors to form for 2 weeks
Excise tumor and isolate genomic DNA
PCR, clone, sequence to identify shRNA
Confirm candidates by soft agar assay and tumor formation

B

Tumor volume (mm³)

Days

0 Kras NIH 3T3 cells
100
1000

C

NS  Angpt1  Cdk5r1  Cr1l  Dapp1  Ddx52  Dnajc12  Flna  Fpr3  Gapvd1  Gzma  Igf2r  Lsm8  Mtap1a  Myd88  Nme4  Nup205  Orc1l  Pigh  Pkd1l3  Prl7a2  Ptgis  Ptpn4  Sdf2l1  Sema3b  Sfrs9  Slfn4  Spast  Spop  Stk11  Txnrtd  Wap  Zfp422
Figure 2.2.1. Large-scale shRNA screen for the identification of candidate TSGs. (A) Schematic summary of the screen. (B) 0, 100 or 1000 Kras-transformed NIH 3T3 cells were mixed with non-transformed NIH 3T3 cells to achieve a final population of 1x10^6 cells, which were then injected into the flanks of nude mice (n=3) (C) Tumor formation assay. The NIH 3T3 KD cell lines were subcutaneously injected into nude mice, and tumors were photographed at various time points following injection.
Figure 2.2.2. shRNA knock down efficiency. qRT-PCR analysis monitoring shRNA-mediated knockdown efficiency for each TSG (error bars represent SD, n=3). Values are given relative to expression of each gene following treatment with a non-silencing (NS) shRNA, which was arbitrarily set to 1.
Figure 2.2.3
Figure 2.2.3. Down-regulation of candidate TSGs in human lung squamous cell carcinoma (hLSCC) samples. qRT-PCR analysis monitoring expression of each TSG in 27 hLSCC samples (error bars represent SD, n=3). Values were normalized to the expression of the TSG in 10 normal lung samples, the average of which was set to 1. The red line indicates 2-fold down-regulation. Asterisks indicate samples whose fold down-regulation vastly exceeds that shown in the graph. Samples have been re-ordered from least to most down-regulated gene.
Figure 2.2.4
Figure 2.2.4. Down-regulation of candidate TSGs in human lung adenocarcinoma samples. qRT-PCR analysis monitoring expression of each TSG in 10 human lung adenocarcinoma samples (error bars represent SD, n=3). Values were normalized to the expression of the TSG in 10 normal lung samples, the average of which was set to 1. The red line indicates 2-fold down-regulation.
Figure 2.2.5. Knocking down TSGs in Small Airway (SA) cell line activates FGFR signaling pathway. (A) Western blots show the levels of key proteins in the FGFR1 signaling pathway after treatment with non-silencing shRNA or shRNA targeting candidate TSGs. Genes labeled in red have increased pFRS2 level, which indicates activation of FGFR1 signaling pathway. (B,C) Western blots show the tFGFR1 and pFGFR1 levels of KD cell lines that activate FGFR1 signaling pathway. Genes labeled in red have increased tFGFR1 level. Genes labeled in green have increased pFGFR1 without increase tFGFR1 level.
Figure 2.2.6

- FLNA, GAPVD1, MYD88, PTGIS, SFRS9, SPAST

- ANGPT1, DAPP1, FPR3, NME4, PTPN4

- IGF2R, LSM8, MAP1A, PIGH, SEMA3B

- CDK5R1, DDX52, SDF2L1, SPOP, STK11, TXNRD1

- FGFR1

- FRS2

- Y436

- Y653P/Y654P

- Y766P

- DAPP1

- pErk

- pAkt

- ANGPT1, DAPP1, TXNRD1
**Figure 2.2.6. Model of TSGs negatively regulating FGFR1 signaling.** 7 genes (FLNA, GAPVD1, MYD88, PIGH, PTGIS, SFRS9, SPAST) inhibit total FGFR1 levels; 5 genes (ANGPT1, DAPP1, FPR3, NME4, PTPN4) inhibit pFGFR1 levels without elevated tFGFR1. Y653 phosphorylation of FGFR1 was not increased in FLNA and MYD88 KD cell lines, although the total FGFR1 is increased; 6 genes (CDK5R1, DDX52, SDF2L1, SPOP, STK11, TXNRD1) inhibit pFRS2 levels without elevated phospho- or total-FGFR1 levels; Four genes (IGF2R, LSM8, MAP1A, SEMA3B) inhibit pErk level independent of FGFR1 signaling.
Candidate TSGs

Fold de-repression after FGFR1 KD

Figure 2.2.7
Figure 2.2.7. Candidate TSGs repressed by FGFR1 signaling at the transcriptional level. Three candidate TSGs are de-repressed after FGFR1 knock down in NCI-H520, monitored by qRT-PCR.
Figure 2.2.8

A

Vector  FGFR1 OE

NIH3T3

SA

B

Fold increase of colony number compared to NS

FGFR1 pathway

C

ANGPT1  CDK5R1  DAPP1  DDX52  FLNA

FPR3  GAPVD1  MYD88  PTGIS  PTPN4

SDF2L1  SFRS9  SPOP  STK11  NS
Figure 2.2.8. Knocking down TSGs induces transformation of lung epithelia cells. (A) 1 million NIH3T3 cells overexpressing FGFR1 and 5 million SA cells overexpressing FGFR1 formed tumors when subcutaneously injected into the flank of nude mice, while no tumor formed in empty vector controls. (B) In vitro soft agar colony formation assay showed 14 out of 17 FGFR1 regulating genes and 4 FGFR1 signaling independent genes can form colonies in soft agar. (C) KD cell lines that formed colonies in soft agar were injected into the flanks of nude mice and formed tumors in vivo.
**Figure 2.2.9.** SA cells with TSGs knocked down are sensitive to FGFR inhibitor. (A) After treatment with 500nM ponatinib for three days, TSG KD cell lines that have increased pFGFR1 levels (see figure 2.2.5) were sensitive to Ponatinib compared to non-silencing control. (B) 3 genes (STK11, MYD88, DAPP1) are highly repressed in A427 compared to SA or HBEC. (C) A427 showed similar sensitivity to Ponatinib (500nM) compared to FGFR1 amplified cell line NCI-H520, while a control cell line A549 with normal level of FGFR1 showed more resistance.
Figure 2.2.10

TSGs

A

B

Percent of individuals with >50% increase in promoter hypermethylation in tumor samples relative to normal samples.
Figure 2.2.10. Candidate TSG promoter hypermethylation in hLSCC samples. (A) Left, methylation heat map for candidate TSGs showing hypomethylation (red) or hypermethylation (green) in the tumor sample. Genes that are significantly hypermethylated (defined as a $\geq 50\%$ increase in methylation) in tumor samples (p-value $< 0.05$) are indicated by an asterisk. Right, color key. (B) Percent of individuals with significant hypermethylation in nine candidate TSGs. The p-value for each gene is shown.
Figure 2.2.10

(A) A panel showing various cell culture images with different gene expressions.

(B) A bar graph illustrating colony number with error bars for different gene expressions.

(C) A line graph showing tumor size over days for various gene expressions.
Figure 2.2.11. Over-expression of candidate TSGs inhibits growth of hLSCC cells. (A) Colony suppression assay. NCI-H520 cells infected with a retrovirus expressing each candidate TSG or empty vector were stained with crystal violet. (B) Quantification of the colony suppression assay shown in (A) (error bars represent SD, n=3). (C) Tumor formation assay. NCI-H520 cells ectopically expressing a TSG were injected subcutaneously into nude mice and tumor growth was monitored (error bars represent SD, n=3).
Figure 2.2.12. Mechanisms of candidate TSGs inhibiting growth of hLSCC cells. (A) Apoptosis assays. NCI-H520 cells ectopically expressing each TSG were analyzed by immunoblot for cleaved caspase 3. (B). Cleaved caspase 3 protein levels visualized on western blots were quantified and normalized to actin levels, and then to cleaved caspase 3 levels in the vector control, which was set to 1. The red line denotes a 2-fold increase. (C) Cellular senescence assays. NCI-H520 cells expressing each TSG were stained for senescence-associated β-galactosidase. (D) The percent positive cells was quantified by counting the blue stained cells (error bars represent SD, n=3). The red line denotes a 2-fold increase relative to the vector control.
<table>
<thead>
<tr>
<th>Pool number</th>
<th>ShRNAs recovered from the tumor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dnajc12, Gapvd1, Mrps18c, Mtap1a, Sfrs9, Svs6, Zfp422</td>
</tr>
<tr>
<td>2</td>
<td>Ddx52</td>
</tr>
<tr>
<td>3</td>
<td>Angpt1, Cdk5r1, Dapp1, Efna3, Nme4, Nup205, Pkd1I3, Ptpn4, Spast, Stk11</td>
</tr>
<tr>
<td>4</td>
<td>Flna, Fpr3, Gzma, Orc1I, Prl7a2, Gm11541, Sdf2I1, Sip1, Riken4632401L01</td>
</tr>
<tr>
<td>6</td>
<td>Mad2l1, Slfn4</td>
</tr>
<tr>
<td>8</td>
<td>Dhrs9, Pigh</td>
</tr>
<tr>
<td>9</td>
<td>Cr1l, Igf2r, Lsm8, Myd88, Ptgis, Sema3b, Spop, Txnrd1, Wap, Gm5058</td>
</tr>
</tbody>
</table>
Table 2.2.2. List of 32 candidate TSGs that validate in colony formation and mouse tumorigenesis assays.

<table>
<thead>
<tr>
<th>Biological process</th>
<th>Gene symbol</th>
<th>Gene name</th>
<th>Human homolog</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell signaling</td>
<td>Angpt1</td>
<td>angiopoietin 1</td>
<td>ANGPT1</td>
</tr>
<tr>
<td></td>
<td>Dapp1</td>
<td>dual adaptor for phosphotyrosine and 3-phosphoinositides 1</td>
<td>DAPP1</td>
</tr>
<tr>
<td></td>
<td>Fpr3</td>
<td>formyl peptide receptor 3</td>
<td>FPR3</td>
</tr>
<tr>
<td></td>
<td>Gapvd1</td>
<td>GTPase activating protein and VPS9 domains 1</td>
<td>GAPVD1</td>
</tr>
<tr>
<td></td>
<td>Igf2r</td>
<td>insulin-like growth factor 2 receptor</td>
<td>IGF2R</td>
</tr>
<tr>
<td></td>
<td>Myd88</td>
<td>myeloid differentiation primary response gene 88</td>
<td>MYD88</td>
</tr>
<tr>
<td></td>
<td>Pkd1l3</td>
<td>polycystic kidney disease 1 like 3</td>
<td>PKD1L3</td>
</tr>
<tr>
<td></td>
<td>Pr7a2</td>
<td>prolactin family 7, subfamily a, member 2</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Ptpn4</td>
<td>protein tyrosine phosphatase, non-receptor type 4</td>
<td>PTPN4</td>
</tr>
<tr>
<td></td>
<td>Stk11</td>
<td>serine/threonine kinase 11</td>
<td>STK11</td>
</tr>
<tr>
<td>Cell growth/migration/metabolism</td>
<td>Cdk5r1</td>
<td>cyclin-dependent kinase 5, regulatory subunit 1 (p35)</td>
<td>CDK5R1</td>
</tr>
<tr>
<td></td>
<td>Flna</td>
<td>filamin, alpha</td>
<td>FLNA</td>
</tr>
<tr>
<td></td>
<td>Mtap1a</td>
<td>microtubule-associated protein 1 A</td>
<td>MAP1A</td>
</tr>
<tr>
<td></td>
<td>Nme4</td>
<td>non-metastatic cells 4, protein expressed in</td>
<td>NME4</td>
</tr>
<tr>
<td></td>
<td>Sema3b</td>
<td>sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (emaphoring) 3B</td>
<td>SEMA3B</td>
</tr>
<tr>
<td></td>
<td>Sfrs9</td>
<td>splicing factor, arginine/serine rich 9</td>
<td>SFRS9</td>
</tr>
<tr>
<td></td>
<td>Spast</td>
<td>spastin</td>
<td>SPAST</td>
</tr>
<tr>
<td>DNA/RNA metabolism</td>
<td>Ddx52</td>
<td>DEAD (Asp-Glu-Ala-Asp) box polypeptide 52</td>
<td>DDX52</td>
</tr>
<tr>
<td></td>
<td>Lsm8</td>
<td>LSM8 homolog, U6 small nuclear RNA associated (S. cerevisiae)</td>
<td>LSM8</td>
</tr>
<tr>
<td></td>
<td>Orc1l</td>
<td>origin recognition complex, subunit 1-like (S. cerevisiae)</td>
<td>ORC1L</td>
</tr>
<tr>
<td></td>
<td>Sfrs9</td>
<td>splicing factor, arginine/serine rich 9</td>
<td>SFRS9</td>
</tr>
<tr>
<td>Immunity</td>
<td>Cr1l</td>
<td>complement component (3b/4b) receptor 1-like</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Gzma</td>
<td>granzyme A</td>
<td>GZMA</td>
</tr>
<tr>
<td>Lipid metabolism</td>
<td>Pigh</td>
<td>phosphatidylinositol glycan anchor biosynthesis, class H</td>
<td>PIGH</td>
</tr>
<tr>
<td></td>
<td>Ptgis</td>
<td>prostaglandin I2 (prostacyclin) synthase</td>
<td>PTGIS</td>
</tr>
<tr>
<td>Nuclear transport</td>
<td>Nup205</td>
<td>nucleoporin 205</td>
<td>NUP205</td>
</tr>
<tr>
<td>Transcriptional regulation</td>
<td>Spop</td>
<td>speckle-type POZ protein</td>
<td>SPOP</td>
</tr>
<tr>
<td></td>
<td>Zfp422</td>
<td>zinc finger protein 422</td>
<td>ZNF22</td>
</tr>
<tr>
<td>Protein folding/export/metabolism</td>
<td>Dnajc12</td>
<td>DnaJ (Hsp40) homolog, subfamily C, member 12</td>
<td>DNAJC12</td>
</tr>
<tr>
<td></td>
<td>Sdf2l1</td>
<td>stromal cell-derived factor 2-like 1</td>
<td>SDF2L1</td>
</tr>
<tr>
<td></td>
<td>Txnrd1</td>
<td>thioredoxin reductase 1</td>
<td>TXNRD1</td>
</tr>
<tr>
<td>Unknown</td>
<td>Wap</td>
<td>whey acidic protein</td>
<td>–</td>
</tr>
</tbody>
</table>
Table 2.2.3. Summary of NCBI SKY/MFISH & CGH data for candidate TSGs.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosome location</th>
<th>Bladder</th>
<th>Brain&amp;CNS</th>
<th>Breast</th>
<th>Cervical</th>
<th>Colorectal</th>
<th>Kidney</th>
<th>Leukemia</th>
<th>Liver</th>
<th>Lung</th>
<th>Lymphoma</th>
<th>Melanoma</th>
<th>Myeloma</th>
<th>Ovarian</th>
<th>Pancreatic</th>
<th>Prostate</th>
<th>Sarcoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANGPT1</td>
<td>8q22.3-q23</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>CDK5R1</td>
<td>17q11.2</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>DAPP1</td>
<td>4q25-q27</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>DDX52</td>
<td>17q21.1</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>DNAJC12</td>
<td>10q22.1</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>FLNA</td>
<td>Xq28</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>FPR3</td>
<td>19q13.3-q13.4</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>GAPVD1</td>
<td>9q33.3</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>GZMA</td>
<td>5q11-q12</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>IGFR2</td>
<td>6q26</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>LSM8</td>
<td>7q31.1-q31.3</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>MAP1A</td>
<td>15q13-qter</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>MYD88</td>
<td>3p22</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>NME4</td>
<td>16p13.3</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>NUP205</td>
<td>7q33</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>ORC1L</td>
<td>1p32</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>PIGH</td>
<td>14q11-q24</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>PKD1L3</td>
<td>16q22.2</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>PTGIS</td>
<td>20q13.13</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>PTPN4</td>
<td>2q14.2</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>SDF2L1</td>
<td>22q11.21</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>SEMA3B</td>
<td>3p21.3</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>SFRS9</td>
<td>12q24.31</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>SPAST</td>
<td>2p24-p21</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>SPOP</td>
<td>17q21.33</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>STK11</td>
<td>19p13.3</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>TXNRD1</td>
<td>12q23-q24.1</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>ZNF22</td>
<td>10q11</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

● indicates a reported deletion of the gene in the cancer type.
Table 2.2.4. Summary of Oncomine data analysis querying whether candidate TSGs are down-regulated in cancer versus normal tissue.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Bladder</th>
<th>Brain&amp;CNS</th>
<th>Breast</th>
<th>Cervical</th>
<th>Colorectal</th>
<th>Esophageal</th>
<th>Gastric</th>
<th>Head&amp;Neck</th>
<th>Kidney</th>
<th>Lymphoma</th>
<th>Melanoma</th>
<th>Myeloma</th>
<th>Other</th>
<th>Ovarian</th>
<th>Pancreatic</th>
<th>Prostate</th>
<th>Sarcoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANGPT1</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>CDK5R1</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>DAPP1</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>DDX52</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>DNAJC12</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>FLNA</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>FPR3</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>GAPVD1</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>GZMA</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>IGF2R</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>LSM8</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>MAP1A</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>MYD88</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>NME4</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>NUP205</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>ORC1L</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>PIGH</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>PKD1L3</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>PTGIS</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>PTPN4</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>SDF2L1</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>SEMA3B</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>SFRS9</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>SPAST</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>SPOP</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>STK11</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>TXNRD1</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>ZNF22</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

● indicates one or more reports of significant down-regulation of the gene in cancer versus normal tissue.
Table 2.2.5. List of catalog numbers for shRNAs obtained from Open Biosystems.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Catalog number for 1(^{st}) shRNA</th>
<th>Catalog number for 2(^{nd}) shRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angpt1</td>
<td>RMM1766-96738501</td>
<td>RMM1766-96739079</td>
</tr>
<tr>
<td>Cdk5r1</td>
<td>RMM1766-96739750</td>
<td>RMM1766-96738443</td>
</tr>
<tr>
<td>Cr1l</td>
<td>RMM1766-97042824</td>
<td>RMM3981-98063631</td>
</tr>
<tr>
<td>Dapp1</td>
<td>RMM1766-96738885</td>
<td>RMM1766-96880910</td>
</tr>
<tr>
<td>Ddx52</td>
<td>RHS1764-9687389</td>
<td>RMM1766-96889298</td>
</tr>
<tr>
<td>Dnajc12</td>
<td>RMM1766-9106520</td>
<td>RMM1766-96743200</td>
</tr>
<tr>
<td>Flna</td>
<td>RMM1766-96741212</td>
<td>RMM1766-9352033</td>
</tr>
<tr>
<td>Fpr3</td>
<td>RMM1766-96746297</td>
<td>RMM1766-98468326</td>
</tr>
<tr>
<td>Gapvd1</td>
<td>RMM1766-9336338</td>
<td>RMM1766-9336322</td>
</tr>
<tr>
<td>Gzma</td>
<td>RMM1766-96745254</td>
<td>RMM1766-9106902</td>
</tr>
<tr>
<td>Igf2r</td>
<td>RMM1766-96746297</td>
<td>RMM1766-98467508</td>
</tr>
<tr>
<td>Lsm8</td>
<td>RMM1766-97044512</td>
<td>RMM1766-96881489</td>
</tr>
<tr>
<td>Mtap1a</td>
<td>RMM1766-9341607</td>
<td>RMM1766-96874116</td>
</tr>
<tr>
<td>Myd88</td>
<td>RMM1766-97042719</td>
<td>RMM3981-97065530</td>
</tr>
<tr>
<td>Nme4</td>
<td>RMM1766-96740301</td>
<td>RMM1766-96740215</td>
</tr>
<tr>
<td>Nup205</td>
<td>RMM1766-9673936</td>
<td>RMM1766-9355153</td>
</tr>
<tr>
<td>Orc1l</td>
<td>RMM1766-96744068</td>
<td>RMM1766-96744205</td>
</tr>
<tr>
<td>Pigh</td>
<td>RMM1766-96891459</td>
<td>RMM1766-96737813</td>
</tr>
<tr>
<td>Pkd1l3</td>
<td>RMM1766-9353837</td>
<td>RMM1766-9354529</td>
</tr>
<tr>
<td>Prl7a2</td>
<td>RMM1766-96744814</td>
<td>RMM3981-98069754</td>
</tr>
<tr>
<td>Ptgis</td>
<td>RMM1766-97042638</td>
<td>RMM1766-96738598</td>
</tr>
<tr>
<td>Ptpn4</td>
<td>RMM1766-96740179</td>
<td>RMM1766-96740341</td>
</tr>
<tr>
<td>Sdf21</td>
<td>RMM1766-96743359</td>
<td>RMM3981-99012961</td>
</tr>
<tr>
<td>Sema3b</td>
<td>RMM1766-97042710</td>
<td>RMM3981-9621999</td>
</tr>
<tr>
<td>Sfrs9</td>
<td>RMM1766-9106052</td>
<td>RMM3981-97059906</td>
</tr>
<tr>
<td>Slfn4</td>
<td>RMM1766-96878856</td>
<td>RMM1766-97043261</td>
</tr>
<tr>
<td>Spast</td>
<td>RMM1766-96737853</td>
<td>RMM1766-96871825</td>
</tr>
<tr>
<td>Spop</td>
<td>RMM1766-97044276</td>
<td>RMM3981-98497970</td>
</tr>
<tr>
<td>Stk11</td>
<td>RMM1766-96740219</td>
<td>RMM3981-9591552</td>
</tr>
<tr>
<td>Txnrd1</td>
<td>RMM1766-97042622</td>
<td>RMM1766-97042953</td>
</tr>
<tr>
<td>Wap</td>
<td>RMM1766-97043220</td>
<td>RMM1766-97044227</td>
</tr>
<tr>
<td>Zfp422</td>
<td>RMM1766-9336451</td>
<td>RMM1766-96739222</td>
</tr>
</tbody>
</table>
Table 2.2.6. List of primers used for quantitative real-time RT-PCR and for cloning TSGs.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward primer (5’ → 3’)</th>
<th>Reverse primer (5’ → 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>qRT-PCR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angpt1 (mouse)</td>
<td>GTGCCGATTTCAGCAGAAGAGG</td>
<td>CATGATTGCTCCGCAGAAGG</td>
</tr>
<tr>
<td>ANGPT1 (human)</td>
<td>GACGAGACAGGAAACACAGAGC</td>
<td>GGGCACATTGCAGACATACAG</td>
</tr>
<tr>
<td>CDk5r1 (mouse)</td>
<td>TGCTGACATGGCTTACCTGCTCC</td>
<td>TGTGACAGAGAGGACAGCCAGG</td>
</tr>
<tr>
<td>CDK5R1 (human)</td>
<td>CACCGCGATTTGAATGCTTCCGG</td>
<td>CCGAGCGCATGCTTCCCCAGC</td>
</tr>
<tr>
<td>Cr11 (mouse)</td>
<td>GTCGGCCTCTGCTTCCGCTTCC</td>
<td>GGTGACTGCTGCTTACAGTCT</td>
</tr>
<tr>
<td>Dapp1 (mouse)</td>
<td>TGGTGAAGAGGAGGAAAGA</td>
<td>CCGAAGAGGAGGAAAGAAGA</td>
</tr>
<tr>
<td>DAPPA1 (human)</td>
<td>CCAAGAGGAGGAGGAAAGA</td>
<td>CCAAGAGGAGGAAAGAAGA</td>
</tr>
<tr>
<td>Ddx52 (mouse)</td>
<td>TCTGAGCAAGGAGGATGATTGC</td>
<td>TGTGACAGAGAGGACAGCCAGG</td>
</tr>
<tr>
<td>DDX52 (human)</td>
<td>TTCTGAGGAACGCTTTGCTGTC</td>
<td>TGTGACAGAGAGGACAGCCAGG</td>
</tr>
</tbody>
</table>
| Dnajc12 (mouse) | CTGGCTGAGGGAGCGTCACCA  | TTCTGAGGAACGCTTTGCTGTC |}

```
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Mouse Sequence</th>
<th>Human Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sfrs9</td>
<td>CGGAATGGGGATGGTTGAAT</td>
<td>CAGACCAGCGACCGGTAGTAG</td>
</tr>
<tr>
<td>SFRS9</td>
<td>ATATGCCCTGCCTAAGCTGG</td>
<td>AGCTGGTGCTTCTCTCAGGA</td>
</tr>
<tr>
<td>Sfln4</td>
<td>GAATGGGTGAAGCGCCAGAT</td>
<td>TCTTGAAGCCACATACCACCT</td>
</tr>
<tr>
<td>Spast</td>
<td>TGGAAATCTCAGGGATGGCTAG</td>
<td>TTTGGAAGCAGATACATACCCT</td>
</tr>
<tr>
<td>SPAST</td>
<td>ATATGCCCTGCCTAAGCTGG</td>
<td>GCTGACGCTGCTTCTCAGGA</td>
</tr>
<tr>
<td>Stk11</td>
<td>GCACGAGGTAAGCGCAGA</td>
<td>CCAAAGTCCCGAAGTGAGTG</td>
</tr>
<tr>
<td>STK11</td>
<td>CATGACCTGCTGGCTCCAGCT</td>
<td>TGTGACTGCCCTCTCTCTTCT</td>
</tr>
<tr>
<td>Txnrd1</td>
<td>AGACGCTTCAGTCCTCAC</td>
<td>TTCTGGACACAGTACCTAC</td>
</tr>
<tr>
<td>TXNRD1</td>
<td>AGTAGGTTCCACATGCAGAC</td>
<td>TTTGATCAAGTTCTCAAGCACAACAAAA</td>
</tr>
<tr>
<td>Wap</td>
<td>CCAGCAAGGAAAGAACCTTATG</td>
<td>TTCTCGTCCACGTCTTCT</td>
</tr>
<tr>
<td>Zfp422</td>
<td>CCAGCAAGGAAAGAACCTTATG</td>
<td>TTCTCGTCCACGTCTTCT</td>
</tr>
<tr>
<td>ZNF22</td>
<td>CTCTCATCTGAGCAGCACA</td>
<td>GAGACCACACACAGACTTCC</td>
</tr>
</tbody>
</table>

**Cloning**

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Mouse Sequence</th>
<th>Human Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANGPT1</td>
<td>GGAAGATCTTCCATGACAGTTT</td>
<td>CCCTCGAGGTCAGGGAATCC</td>
</tr>
<tr>
<td>CDK5R1</td>
<td>CCCTCAGGGGTGACGAGAGGTGGCA</td>
<td>CCCTCGAGGGTGATTACAG</td>
</tr>
<tr>
<td>DAPP1</td>
<td>GGAAGATCTGGATGCAGGGG</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>DDX52</td>
<td>GGAGATCTGCTCGACAGAAG</td>
<td>CCCTCGAGGTCAGGGAATCC</td>
</tr>
<tr>
<td>FPR3</td>
<td>GAAGATCTGCTCGACAGAAG</td>
<td>CCCTCGAGGTCAGGGAATCC</td>
</tr>
<tr>
<td>GZMA</td>
<td>GAAGATCTGCTCGACAGAAG</td>
<td>CCCTCGAGGTCAGGGAATCC</td>
</tr>
<tr>
<td>IGF2R</td>
<td>CCCTCAGGGGTGACGAGAGGTGGCA</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>LSM8</td>
<td>GGAAGATCTGCTCGACAGAAG</td>
<td>CCCTCGAGGTCAGGGAATCC</td>
</tr>
<tr>
<td>MYD88</td>
<td>GAAGATCTGCTCGACAGAAG</td>
<td>CCCTCGAGGTCAGGGAATCC</td>
</tr>
<tr>
<td>PIGH</td>
<td>GGAAGATCTGCTCGACAGAAG</td>
<td>CCCTCGAGGTCAGGGAATCC</td>
</tr>
<tr>
<td>PTGIS</td>
<td>GAAGATCTGCTCGACAGAAG</td>
<td>CCCTCGAGGTCAGGGAATCC</td>
</tr>
<tr>
<td>SDF2L1</td>
<td>GAAGATCTGCTCGACAGAAG</td>
<td>CCCTCGAGGTCAGGGAATCC</td>
</tr>
<tr>
<td>SEMA3B</td>
<td>CCCTCAGGGGTGACGAGAGGTGGCA</td>
<td>CCCTCGAGGTCAGGGAATCC</td>
</tr>
</tbody>
</table>

**Cloning**

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Mouse Sequence</th>
<th>Human Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloning</td>
<td>ANGPT1</td>
<td>CCCTCGAGGTCAGGGAATCC</td>
</tr>
<tr>
<td>Cloning</td>
<td>CDK5R1</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>Cloning</td>
<td>DAPP1</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>Cloning</td>
<td>DDX52</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>Cloning</td>
<td>FPR3</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>Cloning</td>
<td>GZMA</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>Cloning</td>
<td>IGF2R</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>Cloning</td>
<td>LSM8</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>Cloning</td>
<td>MYD88</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>Cloning</td>
<td>PIGH</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>Cloning</td>
<td>PTGIS</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>Cloning</td>
<td>SDF2L1</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
<tr>
<td>Cloning</td>
<td>SEMA3B</td>
<td>CCCTCGAGGTCATTTTAT</td>
</tr>
</tbody>
</table>

183
<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence 1</th>
<th>Sequence 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFRS9</td>
<td>CCGCTCGAGATGGACTACAAAG ACCATGACCGGTGATTATAAAGAT CATGACATCGATTACAAGGATG ACGATGACAAGATGTCGGGCTG GCCGGAC</td>
<td>CCGGAATTCTCAGTAGGGCCTGAAAGGAG</td>
</tr>
<tr>
<td>SPOP</td>
<td>GAAGATCTATGGACTACAAAGA CCATGACCGGTGATTATAAAGATCATGACATCGATTACAAGGATGCGATGACAAGATGTCAGGGTTCCAAGTCC</td>
<td>CCGCTCGAGCGGTTAGGATTGCTTCAGGGCTT</td>
</tr>
<tr>
<td>STK11</td>
<td>GGAAGATCTATGGACTACAAAG ACCATGACCGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGATGGAAGGTGGTGCCGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCGCTCGAGCGGTCAGTGCTGCTTGAGGCC</td>
</tr>
</tbody>
</table>
Table 2.2.7. Source and catalog numbers for human cDNAs used to clone TSGs.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Source</th>
<th>Catalog number</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANGPT1</td>
<td>Open Biosystems</td>
<td>EHS1001-99608061</td>
</tr>
<tr>
<td>CDK5R1</td>
<td>Open Biosystems</td>
<td>MHS1010-58341</td>
</tr>
<tr>
<td>DAPP1</td>
<td>Open Biosystems</td>
<td>MHS1010-7429735</td>
</tr>
<tr>
<td>DDX52</td>
<td>Open Biosystems</td>
<td>MHS1011-9199110</td>
</tr>
<tr>
<td>DNAJC12</td>
<td>Open Biosystems</td>
<td>MHS1010-58074</td>
</tr>
<tr>
<td>FPR3</td>
<td>Open Biosystems</td>
<td>MHS1010-98684799</td>
</tr>
<tr>
<td>GZMA</td>
<td>Open Biosystems</td>
<td>MHS1011-76468</td>
</tr>
<tr>
<td>IGF2R</td>
<td>ATCC</td>
<td>95661</td>
</tr>
<tr>
<td>LSM8</td>
<td>Open Biosystems</td>
<td>MHS1011-60450</td>
</tr>
<tr>
<td>MYD88</td>
<td>Open Biosystems</td>
<td>MHS1010-73828</td>
</tr>
<tr>
<td>NUP205</td>
<td>Open Biosystems</td>
<td>EHS1001-99865050</td>
</tr>
<tr>
<td>PIGH</td>
<td>Open Biosystems</td>
<td>MHS1011-60796</td>
</tr>
<tr>
<td>PTGIS</td>
<td>Open Biosystems</td>
<td>MHS1010-98052790</td>
</tr>
<tr>
<td>SDF2L1</td>
<td>Open Biosystems</td>
<td>MHS4426-99240278</td>
</tr>
<tr>
<td>SEMA3B</td>
<td>Open Biosystems</td>
<td>MHS1010-74089</td>
</tr>
<tr>
<td>SFRS9</td>
<td>Open Biosystems</td>
<td>MHS1010-98052451</td>
</tr>
<tr>
<td>SPOP</td>
<td>Open Biosystems</td>
<td>MHS1010-57540</td>
</tr>
<tr>
<td>STK11</td>
<td>Open Biosystems</td>
<td>MHS1011-62254</td>
</tr>
<tr>
<td>TXNRD1</td>
<td>Open Biosystems</td>
<td>MHS1010-58057</td>
</tr>
</tbody>
</table>
REFERENCES


Chesi M, Nardini E, Brents LA, Schrock E, Ried T, Kuehl WM and Bergsagel PL. Frequent translocation t(4;14)(p16.3;q32.3) in multiple myeloma is associated with increased expression and activating mutations of fibroblast growth factor receptor 3. Nat. Genet. 16, 260-264, (1997).


C.M Counter, A.A Avilion, C.E LeFeuvre, N.G Stewart, C.W Greider, C.B Harley, S Bacchetti. Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. EMBO J., 11 (1992), pp. 1921–1929


Doolittle, R. F. et al. Simian sarcoma virus onc gene, v-sis, is derived from the gene (or genes) encoding a platelet-derived growth factor. Science 221, 275–277 (1983)


Goldstein, M., Meller, I. and Orr-Urtreger, A. (2007) FGFR1 over-expression in primary rhabdomyosarcoma tumors is associated with hypomethylation of a 5' CpG island and abnormal expression of the AKT1, NOG, and BMP4 genes. Genes Chromosomes Cancer 46, 1028-1038


invasive carcinoma of the breast is a molecular entity distinct from luminal invasive ductal carcinoma. Eur. J. Cancer 46, 2399-2407


Murphy, T., Darby, S., Mathers, M. E. and Gnanapragasam, V. J. (2009) Evidence for distinct alterations in the FGF axis in prostate cancer progression to an aggressive clinical phenotype. J. Pathol. 220, 452-460


Neve, R. M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F. L., Fevr, T., Clark, L.,


J. Schlessinger. Signal transduction by allosteric receptor oligomerization

J. Schlessinger, A.N. Plotnikov, O.A. Ibrahimi, A.V. Eliseenkova, B.K. Yeh, A.
Yayon, R.J. Linhardt, M. Mohammadi. Crystal structure of a ternary FGF-FGFR-
heparin complex reveals a dual role for heparin in FGFR binding and

Screaton, G.R., et al. Identification and characterization of three members of the


Shih, C. & Weinberg, R. A. Isolation of a transforming sequence from a human

Shin, E. Y., Lee, B. H., Yang, J. H., Shin, K. S., Lee, G. K., Yun, H. Y., Song,
Oncol. 126, 519-528

Shingu, K., Fujimori, M., Ito, K., Hama, Y., Kasuga, Y., Kobayashi, S., Itoh, N.
growth factor receptor-1 in thyroid diseases: difference between neoplasms and
hyperplastic lesions. Endocr. J. 45, 35-43

tissue microarray analysis of 3p25 (RAF1) and 8p12 (FGFR1) copy number alterations in urinary bladder cancer. Cancer Res. 61, 4514-4519


