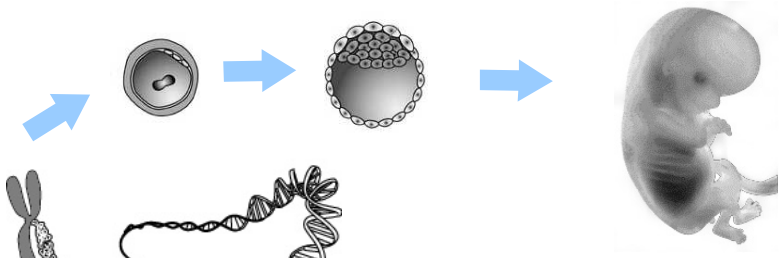
Dissecting the Myc Regulatory Network Using Promoter Microarrays

Bing Ren
Ludwig Institute for Cancer Research
University of California - San Diego

The Human Genome

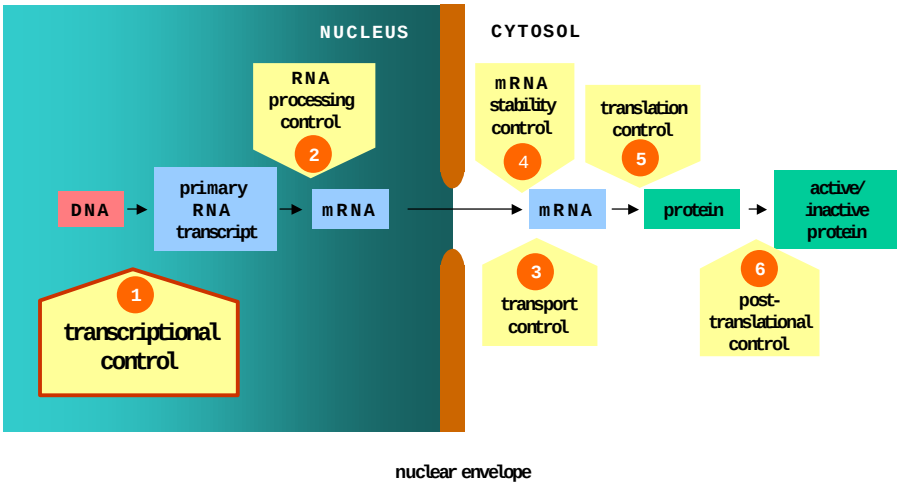


2,900,000,000 Base pairs

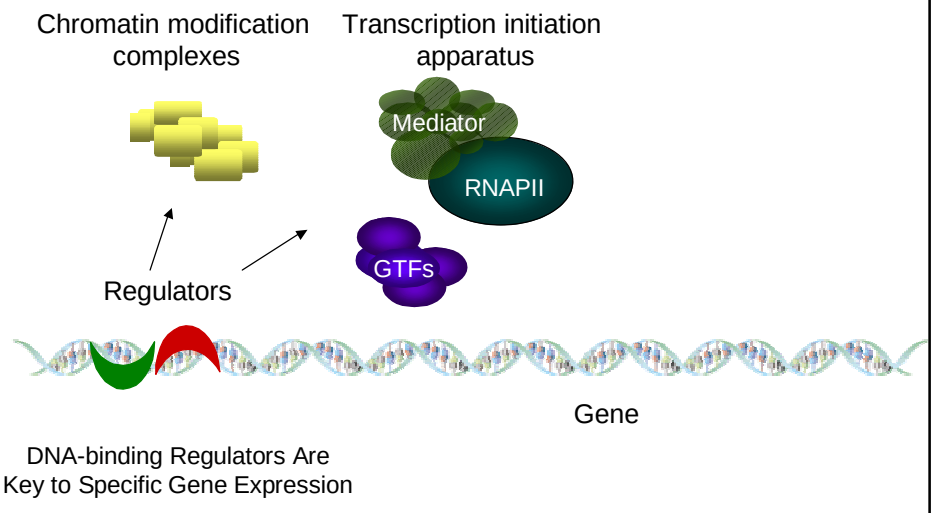
30,000 Genes

2,000 Transcriptional Factors

Regulation of Gene Expression



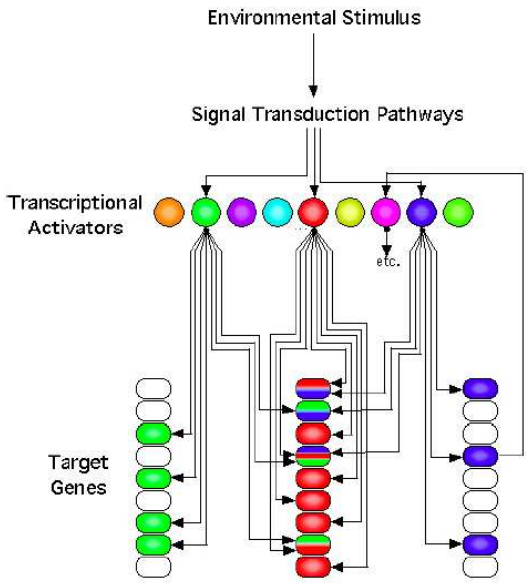
Transcriptional Regulation of Gene Expression



DNA-binding Regulators Are Key to Specific Gene Expression

Transcriptional Regulatory Networks

A transcriptional regulatory Network might describe how a gene expression program is controlled by transcriptional activators across the genome



Cancer Associated Transcription Factors

NAME	CANCER TYPE
AL5Q31	acute lymphoblastic leukemia
ASCL1	small cell lung cancer (SCLC)
BCL3	B-cell leukemia
BCL6	B cell lymphoma
CBFB	myeloid leukemia
CBL	pre-B and pro-B cell lymphomas
CTNNB1	Colon cancer
DEK	Leukemia, acute nonlymphocytic
ERG	Acute myeloid leukemia (AML)
ETS1	erythroblastosis
ETS2	erythroblastosis
ETV6	AML
FOS	murine osteosarcoma
FKHR	rhabdomyosarcoma
GAS41	Glioma
GLI	Glioma
HOX11	Leukemia
IRF2	AML
IRF4	multiple myeloma
JUN	murine osteosarcoma
LMO2	Acute T-cell leukemia
LYL1	T-cell leukemia
MAF	multiple myeloma
MLL	AML
MYB	Leukemia
C-MYC	Lymphoma, Breast Cancer, lung cancer
N-MYC	neuroblastoma

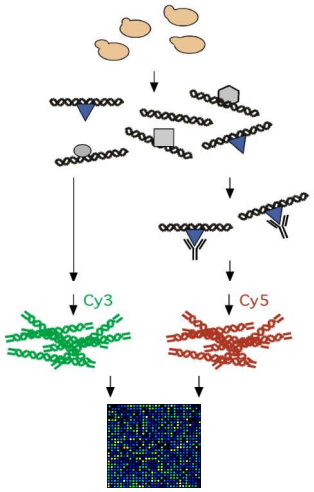
NAME	CANCER TYPE
NFKB1	Acute lymphoblastic leukemia
NFKB2	B-Cell lymphoma
PAX3	alveolar rhabdomyosarcoma
PAX5	B-Cell lymphoma
PAX7	alveolar rhabdomyosarcoma
PBX1	leukemia
RARA	AML
REL	Diffuse large cell lymphoma
RELA	Diffuse large cell lymphoma
RUNX1	Leukemia, acute myeloid
SPI1	acute murine erythroleukemia
STAT3	Leukemia
TAL1/SCL	T cell leukemia
TAL2	T cell leukemia
TCF3	Acute leukemias
BLIMP1	B-cell non-Hodgkin lymphoma
E2F1	Murine Reproductive tract sarcomas
IRF1	AML
MAXI1	prostate adenocarcinoma
PML	Acute promyelocytic leukemia
RB1	retinoblastoma
SMAD3	colorectal cancer
SMAD4/DPC4	pancreatic carcinoma; juvenile polyposis
TFE3	renal cell carcinoma
TP53	Colorectal cancer and other types
WT1	Wilms tumor
ZF9/KLF6	prostate cancer

How to map transcription factor targets?

Rationale & Methodology
Genome wide location analysis

Mapping Transcription factor targets in human cells
- Making Promoter Arrays
- Dissecting the Myc/Max regulatory network
  - What are the target genes?
  - How is specificity determined?

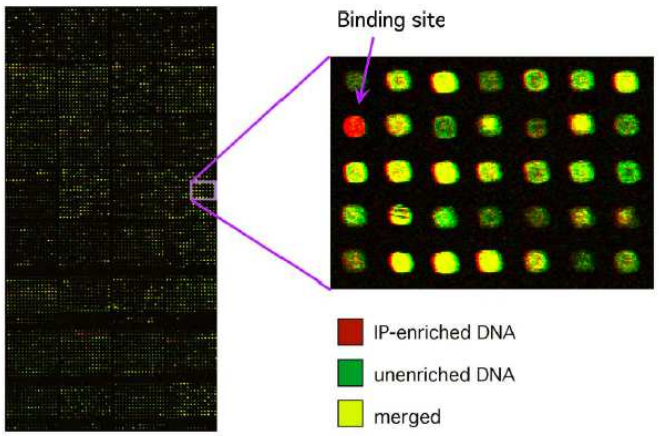
Genome-wide Location Analysis



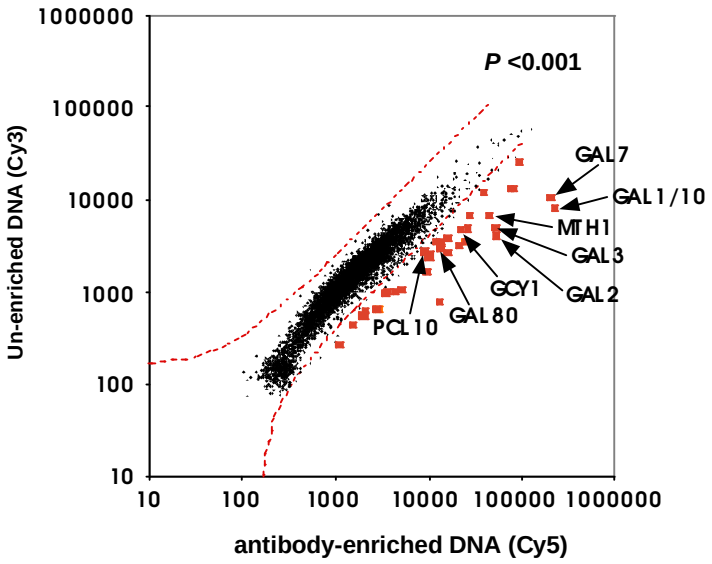
1. Cell cultures grown under appropriate conditions
2. Cross-link with formaldehyde
   Shear chromatin by sonication
3. Immunoprecipitate with specific antibody
   - polyclonal or monoclonal
   - epitope-tag
4. Amplify/label DNA
   - label IP DNA with Cy5
   - label input DNA with Cy3
5. Hybridize to DNA microarray containing all ORFs and intergenic regions and scan

Ren et al. *Science*, 2000; Iyer et al., *Nature* 2001

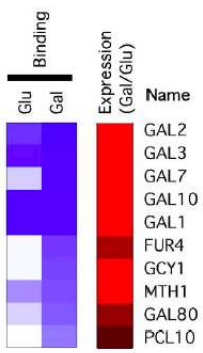
Example of a Scanned Image

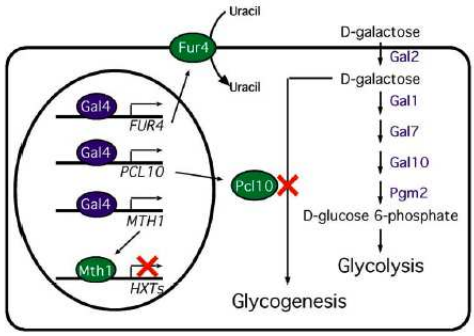


In vivo Binding Sites of Gal4p in yeast



Genome-wide location of the Gal4 activator

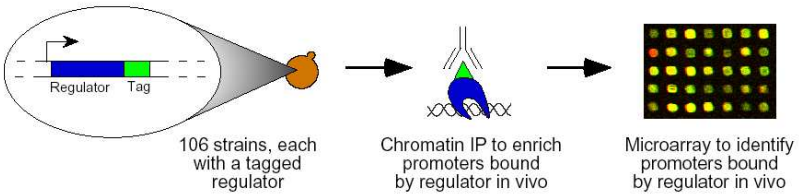




- All previously identified targets confirmed
- New targets confirmed by conventional methods
- All targets depend on wild type Gal4 for expression
- New targets add to understanding of galactose regulation

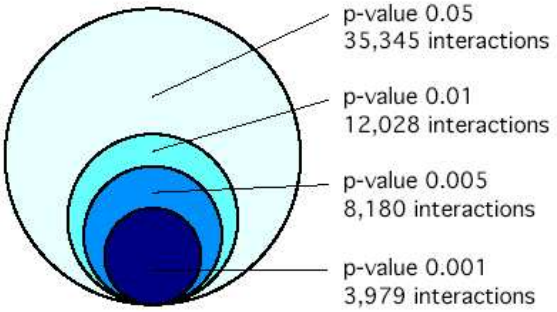
Ren et al., Science 290: 2306 (2000)

Genome-wide Location of Transcriptional Regulators



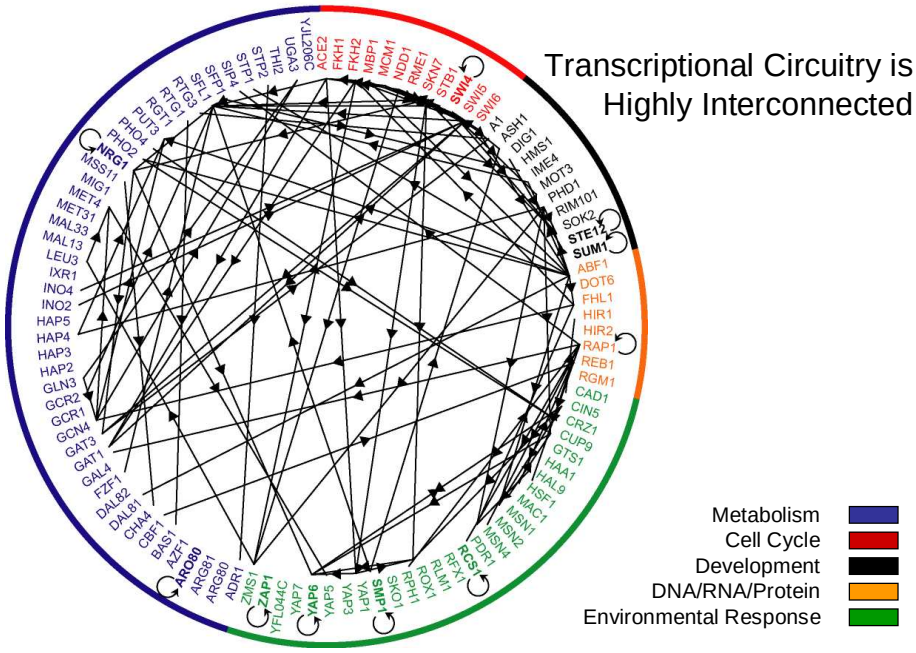
Lee et al., Science 298:799 (2002)

Genome-wide Location of Transcriptional Regulators



At p=0.001:
False positive rate is 5-7%
Regulators identified for 40% of yeast genes (2343)

Transcriptional Circuitry is Highly Interconnected



How to map transcription factor targets in mammalian cells?

Rationale & Methodology

Genome wide location analysis

Mapping Transcription factor targets in human cells

Making promoter microarrays

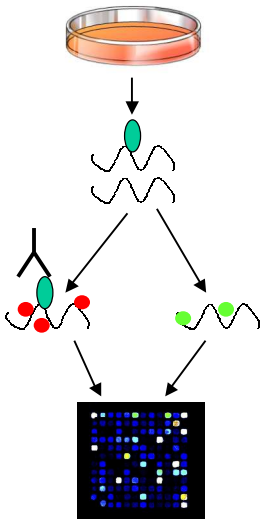
Dissecting the Myc/Max regulatory network

What are the target genes?

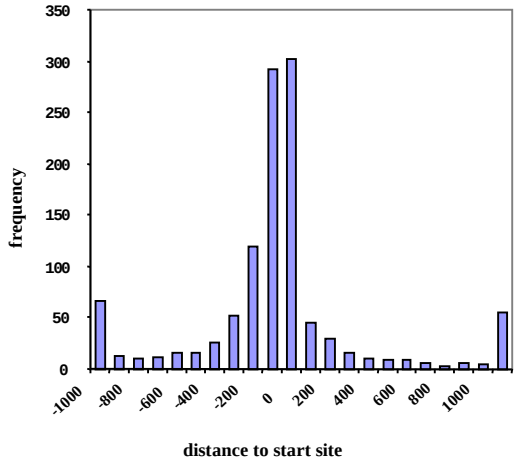
How is specificity determined?

Does genome location analysis work with human cells? —— Challenges

- Genome is 200 times bigger than yeast
- Abundant repetitive sequences
- Annotation of gene structure and function is much less complete
- Many different cell types
- Quality of antibodies

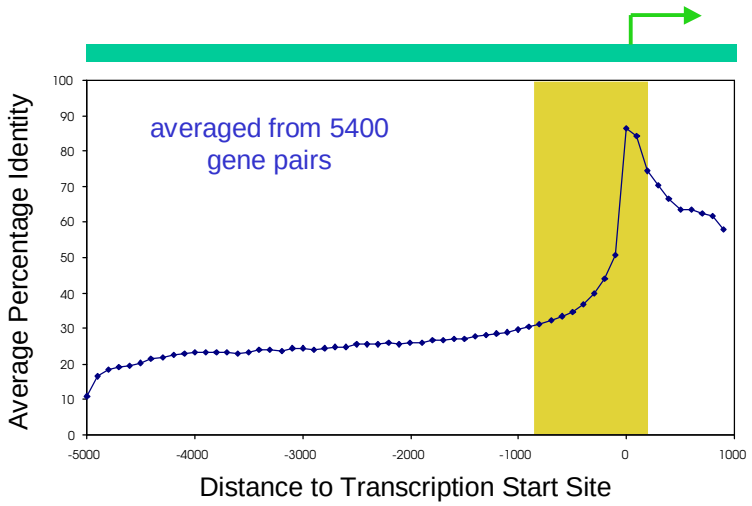


Transcription Factors Tend to Bind to Proximal Promoter Regions

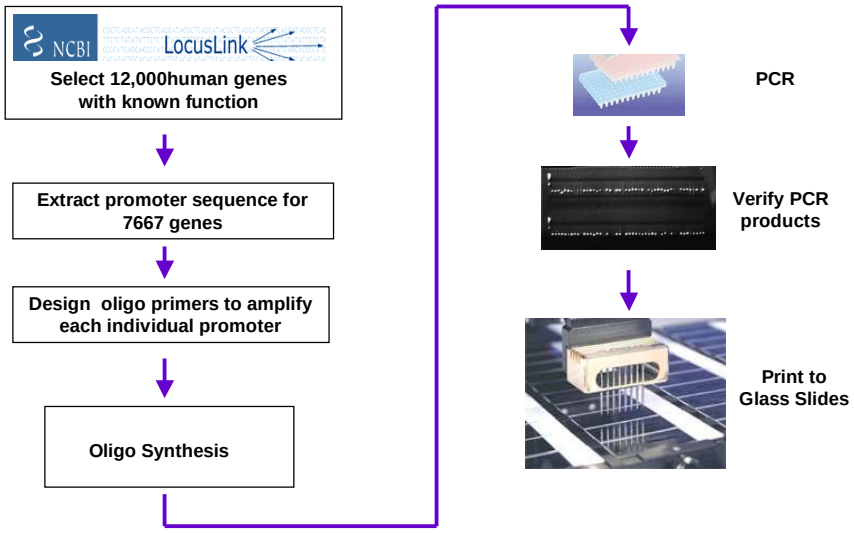


Source:TransFAC

Gene promoter regions are highly conserved between human and mouse orthologs



Making the Human Promoter Arrays



Mapping transcription factor targets

Rationale & Methodology

- Genome wide location analysis

Mapping Transcription factor targets in human cells

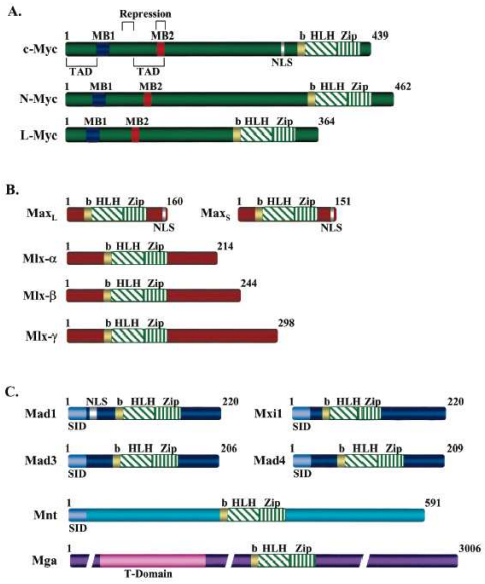
- Making Promoter Arrays
- Dissecting the Myc/Max regulatory network
  - What target genes?
  - How is specificity determined?

Translocation of c-myc in Human Cancers

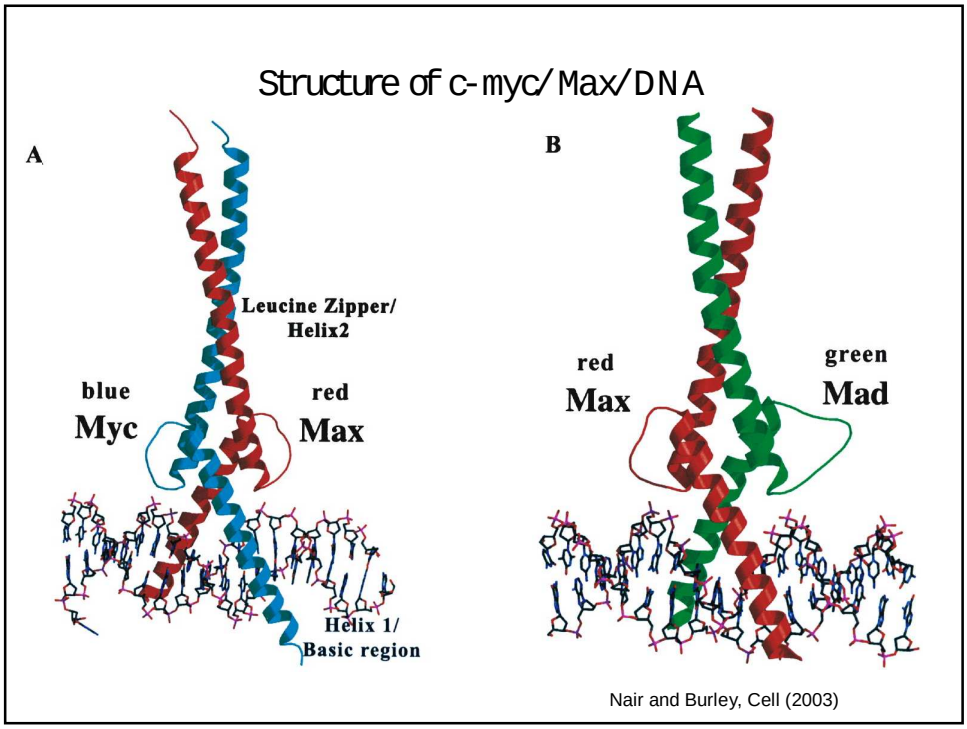
Translocation	*Genes*	*Disease*
t(8;14)(q24;q32)	c-*myc*/IgH	Burkitt's lymphoma
t(2;8)(p12;q24)	Igκ/c-*myc*	Burkitt's lymphoma
t(8;22)(q24;q11)	c-*myc*/Igλ	Burkitt's lymphoma
t(8;14)(q24;q32)	c-*myc*/IgH	Diffuse large cell lymphoma
t(8;14)(q24;q11)	c-*myc*/TCRα,β	T-cell acute lymphoblastic leukemia
t(8;14)(q24;q32)	c-*myc*/IgH	Multiple myeloma
t(2;8)(p12;q24)	Igκ/c-*myc*	Multiple myeloma
t(8;22)(q24;q11)	c-*myc*/Igλ	Multiple myeloma

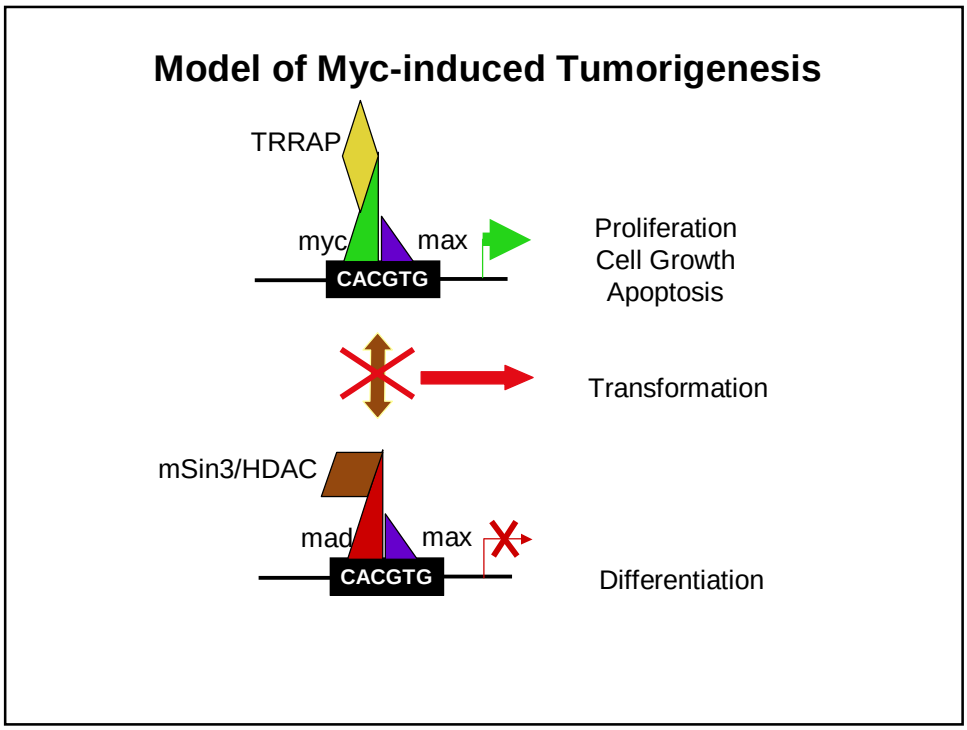
Boxer and Dang, Oncogene (2001)

The Myc, Max and Mad Families

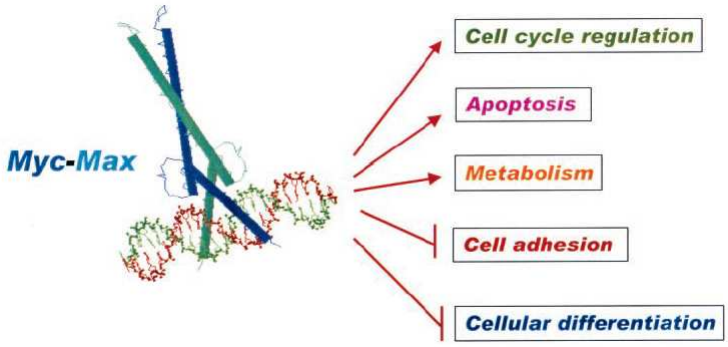


Baudino and Cleveland, MCB (2001)


Structure of c-myc/Max/DNA
A
Leucine Zipper/
Helix2
blue
Myc
red
Max
Helix 1/
Basic region
B
red
Max
green
Mad
Nair and Burley, Cell (2003)


Model of Myc-induced Tumorigenesis
TRRAP
myc
max
CACGTG
Proliferation
Cell Growth
Apoptosis
Transformation
mSin3/HDAC
mad
max
CACGTG
Differentiation

c-Myc affects a variety of cellular functions

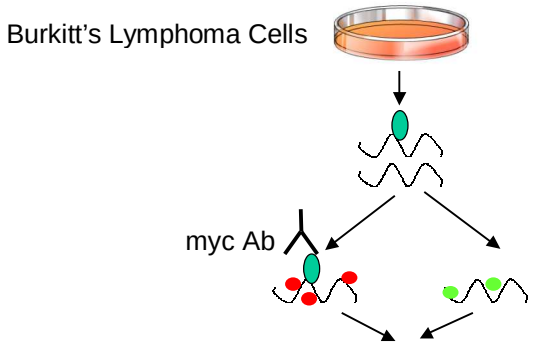


Boxer and Dang, Oncogene (2001)

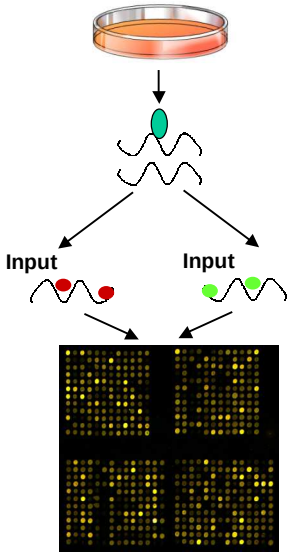
Questions Regarding the Myc Regulatory Network

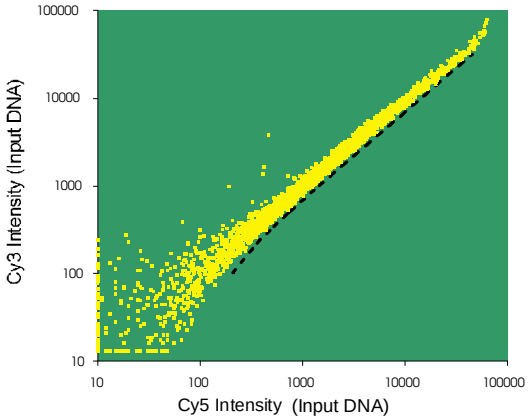
- What are the target genes?
- How are the target sites recognized in vivo?
- What are the mechanisms of c-myc activation or repression?

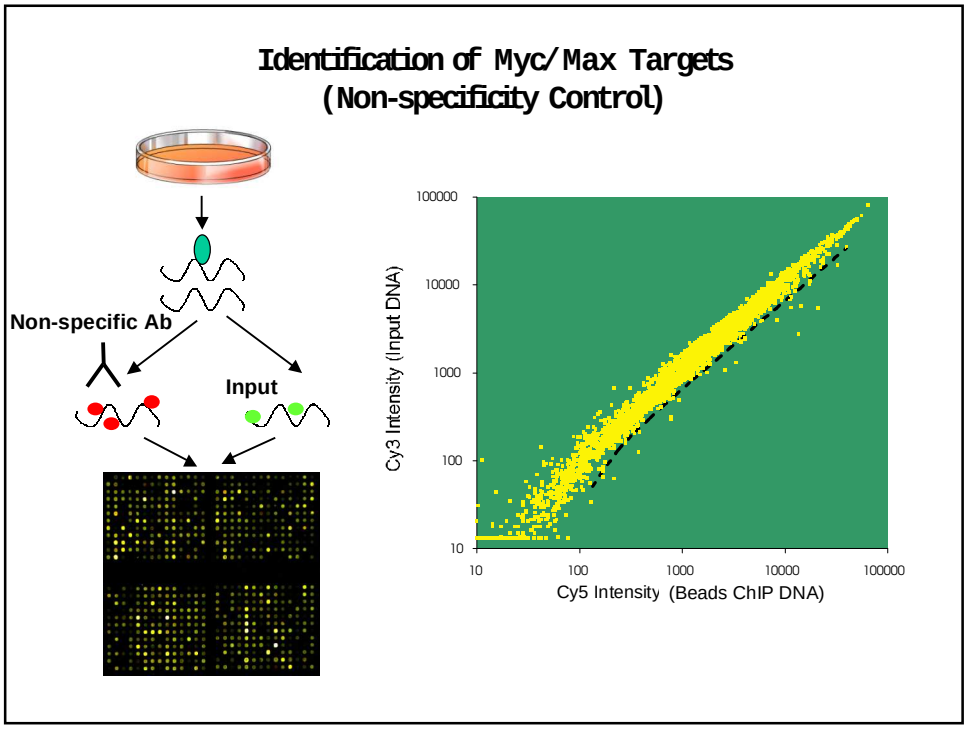
Identification of Myc/Max Targets - Strategy

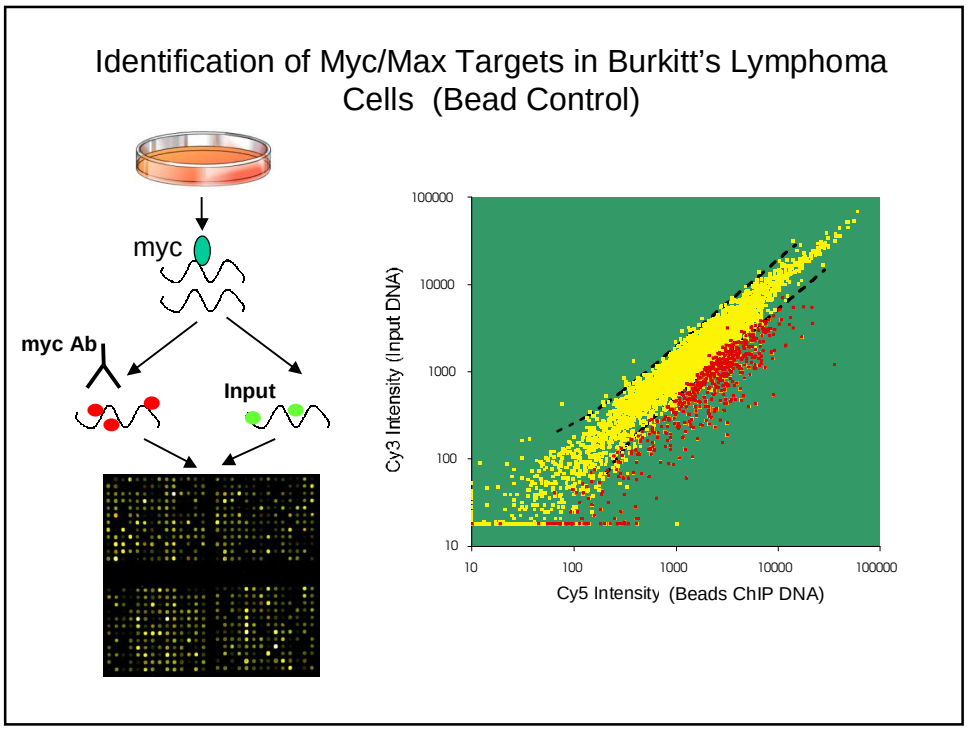


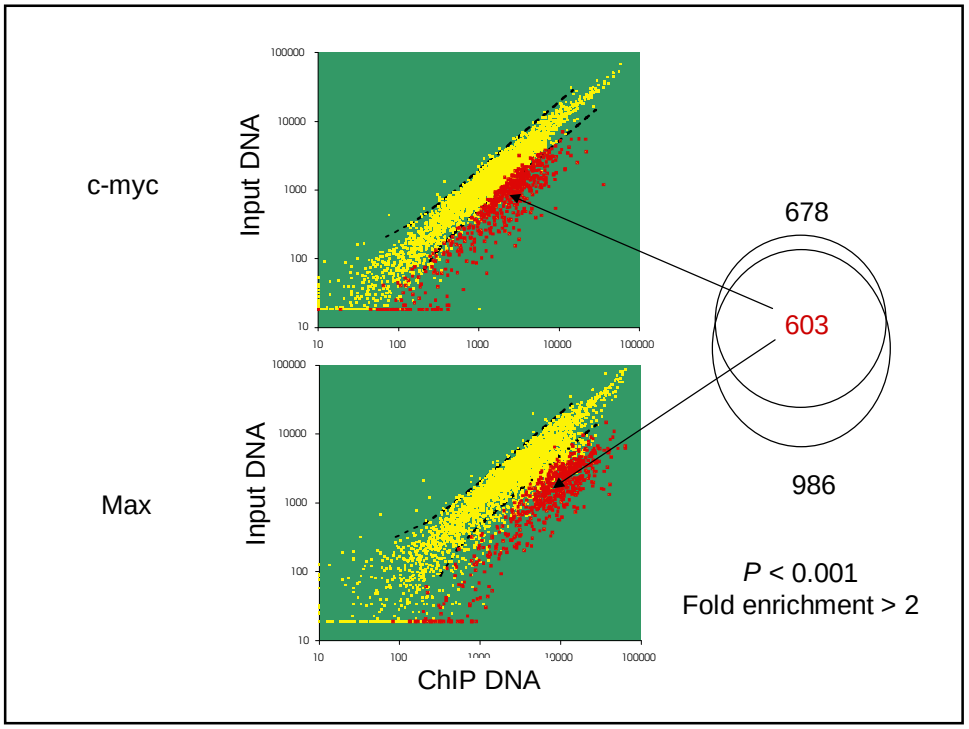
Identification of Myc/Max Targets (control 1)

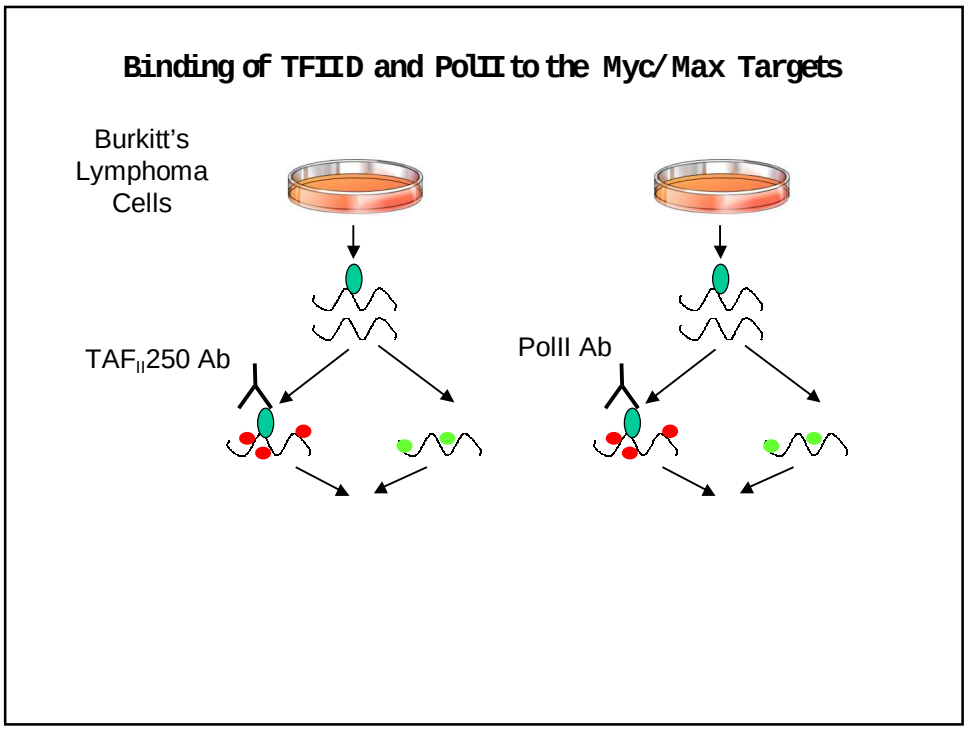


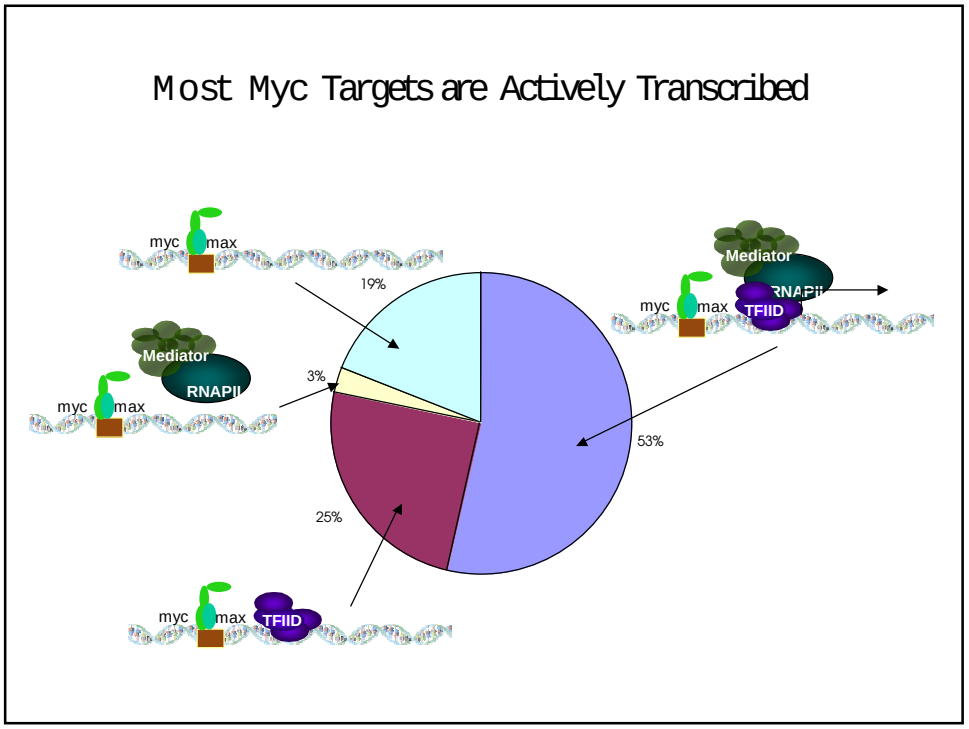


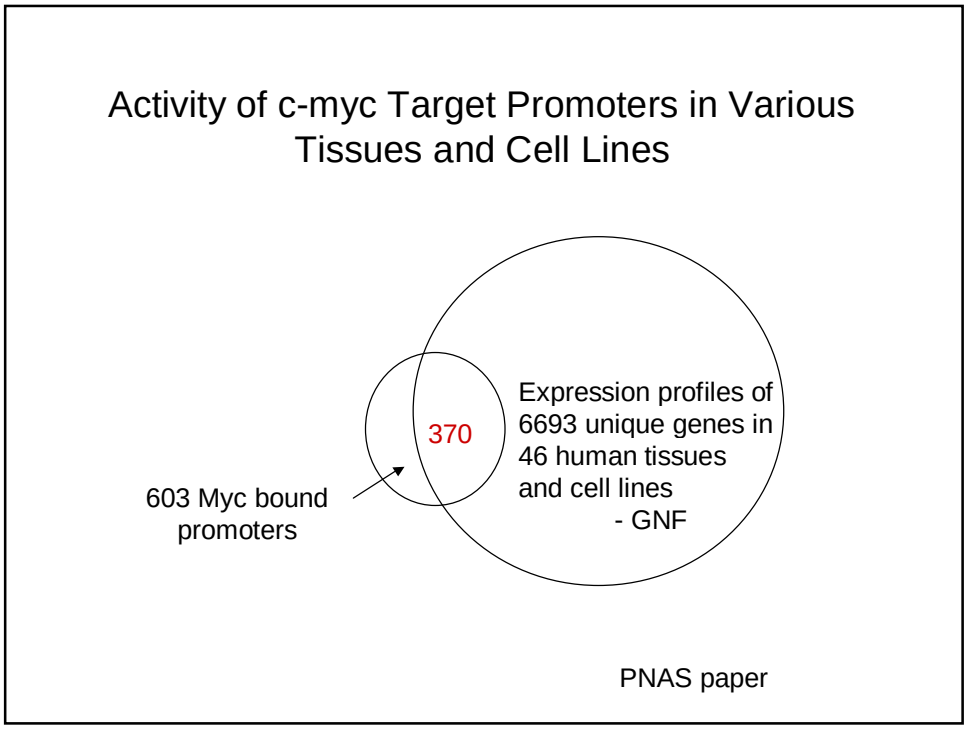

Identification of Myc/Max Targets
(Non-specificity Control)
Non-specific Ab
Input
Cy3 Intensity (Input DNA)
Cy5 Intensity (Beads ChIP DNA)
100000
10000
1000
100
10


Identification of Myc/Max Targets in Burkitt's Lymphoma Cells (Bead Control)
myc
myc Ab
Input
Cy3 Intensity (Input DNA)
Cy5 Intensity (Beads ChIP DNA)
100000
10000
1000
100
10

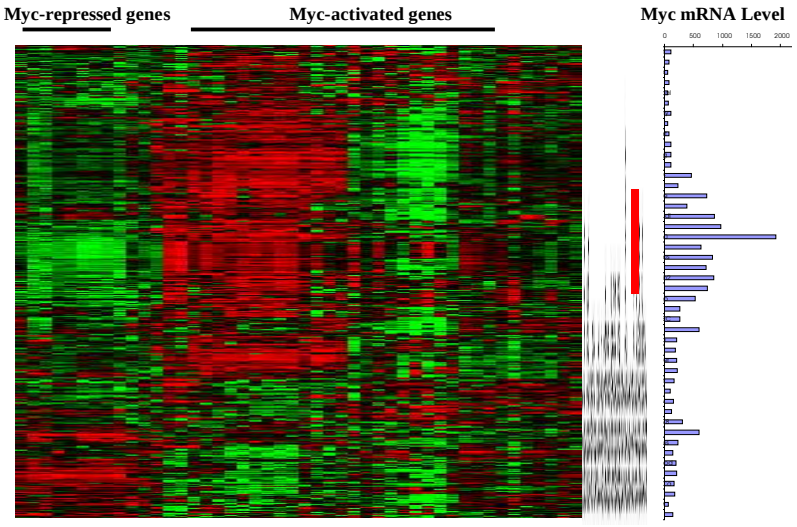

c-myc
Max
Input DNA
Input DNA
ChIP DNA
678
603
986
P < 0.001
Fold enrichment > 2


Binding of TFIID and PolII to the Myc/Max Targets
Burkitt's Lymphoma Cells
TAFII250 Ab
PolII Ab


Most Myc Targets are Actively Transcribed
myc
max
19%
Mediator
RNAPII
3%
myc
max
53%
Mediator
myc
max
TFIID
25%
myc
max
TFIID


Activity of c-myc Target Promoters in Various Tissues and Cell Lines
370
603 Myc bound promoters
Expression profiles of 6693 unique genes in 46 human tissues and cell lines
- GNF
PNAS paper

Expression of Myc Target Genes correlates with myc transcription in 47 human tissues and cell lines



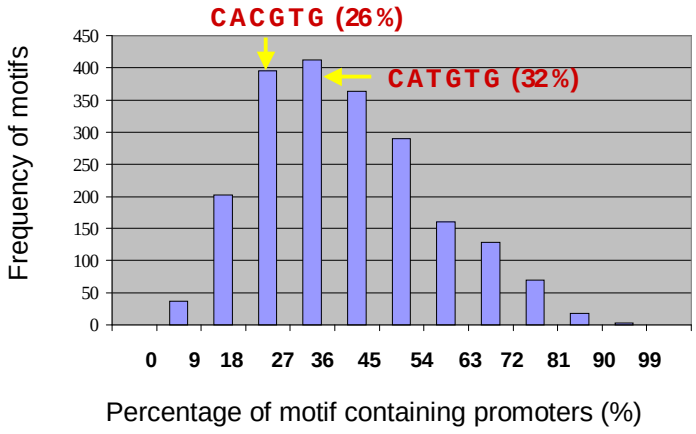
Summary

- Genome wide location analysis is a general method for transcription factor target identification
- Myc/Max bind to a large number of gene promoters in Burkitt's lymphoma cells
- A majority of the c-myc target genes appear to be actively transcribed in the Burkitt's lymphoma cells
- A small number of c-myc target genes appear to be negatively regulated by this protein

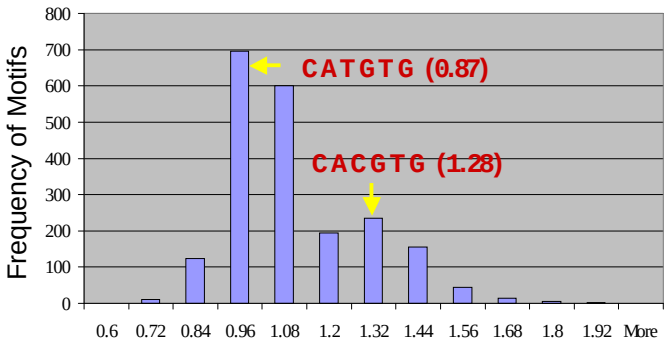
Sequences that specify c-myc binding to its target promoters

- Is E-box alone responsible for c-myc's specificity?
- Are there other DNA elements responsible for c-myc binding specificity?
- Other mechanisms that determine c-myc binding specificity in vivo?

Frequency Distribution of All Possible 6-mer Motifs in c-myc-bound Promoters

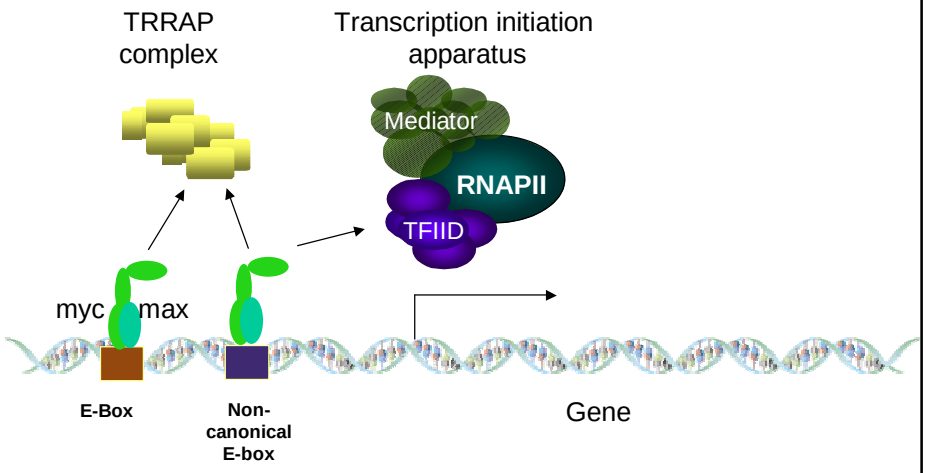


Is E-box Preferentially Located in c-myc-bound Promoters?



Ratio between motif frequency in c-myc-bound promoters and random promoters

Model for c-myc Action in Cancer Cells



Summary

- E-box alone can not explain the specificity of c-myc binding to target promoters
- In vivo binding of c-myc to gene promoters may involve c-myc/non-canonical E-box interactions
- C-myc binding to gene promoter may also involve mechanisms other than specific protein/DNA interactions

Future work

- Capturing a snap shot of the transcriptional regulatory network in mammalian cells
- Identifying the mechanisms that govern in vivo binding specificity of transcription factors
- Integrating protein-protein, protein-DNA, mRNA data

Acknowledgements
Sara Van Calcar
Zirong Li
Chunxu Qu