How much DNA/chip?

Need scatter plot of real experiment

Details

Challenges for human location analysis
Regulation of Genome Expression in Living Cells: Location Analysis

Transcriptional regulatory networks

Location analysis method

Mapping transcriptional regulatory networks
  Carbon source
  Cell cycle
  Multiple information sources useful for mapping
A transcriptional regulatory network might describe how a gene expression program is controlled by transcriptional activators across the genome.
How can transcriptional regulatory networks be mapped?

Expression profiling can reveal gene expression programs

But network maps cannot be derived from expression data alone:

- Transcript populations in expression data are the product of large number of variables
- Noise

Yeast Cell Cycle Gene Expression

Spellman et al. and Cho et al., 1998
Mapping Transcriptional Regulatory Networks: DNA-binding Activators Are Key To Specific Gene Expression
Mapping Transcriptional Regulatory Networks: 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
Example of a Scanned Image

Binding site

- IP-enriched DNA
- unenriched DNA
- merged
Scatter plots

Genome Regulation: Replication of Chromosomes

Location of ORC and MCM binding sites identified 429 replication origins in the yeast genome

- 90% confirmed to have origin activity

John Wyrick, Steve Bell, Oscar Aparicio. Science (2001)
Regulation of Gene Expression by the Gal4 Transcriptional Activator
Genome-wide Location Analysis of Gal4 in Glucose and Galactose Media
Genome-wide Location Analysis of Gal4

<table>
<thead>
<tr>
<th>Name</th>
<th>Binding</th>
<th>Glucose</th>
<th>p-value</th>
<th>Galactose</th>
<th>p-value</th>
<th>Expression (Gal/Glu)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAL2</td>
<td>5.0</td>
<td>8.9E-4</td>
<td>0.29</td>
<td>9.6</td>
<td>1.4E-7</td>
<td>188.6</td>
<td>Galactose permease</td>
</tr>
<tr>
<td>GAL3</td>
<td>5.9</td>
<td>2.9E-4</td>
<td>0.29</td>
<td>8.3</td>
<td>2.5E-7</td>
<td>18.2</td>
<td>Regulatory protein required for rapid induction of galactose pathway</td>
</tr>
<tr>
<td>GAL7</td>
<td>1.5</td>
<td>0.29</td>
<td></td>
<td>8.4</td>
<td>3.8E-7</td>
<td>170.1</td>
<td>UDP-glucose--hexose-1-phosphate uridylyltransferase</td>
</tr>
<tr>
<td>GAL10</td>
<td>8.5</td>
<td>8.5E-5</td>
<td></td>
<td>7.9</td>
<td>7.0E-7</td>
<td>118.8</td>
<td>UDP-glucose 4-epimerase</td>
</tr>
<tr>
<td>GAL1</td>
<td>8.5</td>
<td>8.5E-5</td>
<td></td>
<td>7.9</td>
<td>7.0E-7</td>
<td>271.3</td>
<td>Galactokinase, first step in galactose metabolism</td>
</tr>
<tr>
<td>FUR4</td>
<td>1.1</td>
<td>0.86</td>
<td></td>
<td>4.7</td>
<td>7.1E-6</td>
<td>4.6</td>
<td>Uracile permease (adjacent to GAL1)</td>
</tr>
<tr>
<td>GCY1</td>
<td>1.1</td>
<td>0.73</td>
<td></td>
<td>4.3</td>
<td>1.2E-5</td>
<td>93.1</td>
<td>Galactose-induced oxidoreductase</td>
</tr>
<tr>
<td>MTH1</td>
<td>2.5</td>
<td>0.01</td>
<td></td>
<td>4.1</td>
<td>2.2E-5</td>
<td>21.6</td>
<td>Repressor of hexose transport genes</td>
</tr>
<tr>
<td>GAL80</td>
<td>1.4</td>
<td>0.39</td>
<td></td>
<td>3.7</td>
<td>4.1E-5</td>
<td>4.0</td>
<td>Negative regulator for expression of galactose-induced genes</td>
</tr>
<tr>
<td>PCL10</td>
<td>0.6</td>
<td>1.80</td>
<td></td>
<td>3.0</td>
<td>2.6E-4</td>
<td>2.6</td>
<td>Cyclin that associates with Pho85p, involved in glycogen accumulation</td>
</tr>
</tbody>
</table>
Ambiguity:
Combining Global Location and Expression Data
Consensus Sequences for Transcription Factor Binding Sites

Gal4 binding site (AlignAce)

Gal4-regulated genes
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

Yeast cell cycle

Sister chromatid separation

Growth

DNA replication

Budding

Yeast genome regulation

Simon et al., Cell 106: 697 (2001)
Transcriptional regulatory networks can be mapped by fusing factor binding and gene expression information.

Yeast cell cycle

Factor Binding

Gene Expression

Yeast genome regulation

Activators that function during one stage of the cell cycle regulate activators that function during the next stage. This serial regulation of transcriptional activators forms a connected, circular regulatory network.

Target Gene Key:
Activators
Cyclins

Yeast genome regulation
Cell Cycle Regulatory Network

Pre-replication complex formation
Mcm3, Mcm6, Cdc6, Cdc46

Mating
Ste2, Ste6, Far1, Mfa1, Mfa2, Aga1, Aga2

Budding
Cln1, Cln2, Gic1, Gic2, Msb2, Rsr1, Bud9, Mnn1, Och1, Psa1, Gin4, Gas1, Exg1, Gil1, Kre6, Crh1, Cwp1, Cis3, Cwp2, Scw4, Rax2

Cytokinesis
Cts1, Egt2, Mcm1, Swi5, Acc2

DNA replication & repair
Clb5, Clb6, Rnr1, Rad27, Cdc21, Dur1, Rad51, Cdc45, Ir1, Pds5, Mcd1, Mcm2

Mitosis control
Clb2, Ace2, Swi5, Cdc20, Spo12, Apc1, Tem1

Chromatin
Htb1, Htb2, Hta1, Hta2, Hta3, Hho1, Hhf1, Hht1, Tel2, Ctf18, Hos3, Arp7

Yeast genome regulation
Information needed to more fully model transcriptional regulatory networks

Expression profiles for major biological processes:
  cell cycle
  nutrition
  environmental response
  genome maintenance
  development

Factor Location:
  activators and repressors
  coactivators and corepressors
  transcription apparatus
  chromatin modifying factors

Genome Sequence
  conserved cis-acting sequences
What are we learning from transcriptional regulatory networks?

How biological processes are coordinated through gene expression

How genes for multisubunit complexes are coregulated

How expression of individual genes is regulated by many different transcription factors

How modification of nucleosomes and the transcription apparatus contributes to regulation

How to annotate genes of unknown function
Human Transcriptional Regulatory Networks: Transcription Factors Implicated in Human Disease

Cancer
AML1, p53, PLZF, PML, Rb, WT1

Developmental Defects
GATA1, VDR, CRX, CBP, MeCP2

Diabetes
IPF1, HNF4a, TCF/HNF1, TCF2

Immunological Defects
RFX5, WHN

Obesity
PPARgamma, SIM1

Hypertension
NR3C2, GCCR

Genome Location Analysis

- Human cells

- Factor binding sites enriched with anti-factor antibody

- Factor-enriched DNA labeled with Cy5, total genomic DNA with Cy3

- Microarray with human promoters

Bing Ren
Hieu Cam
Brian Dynlacht
Model For Cell Cycle Regulation by E2F Transcription Factors

Human genome regulation
Regulation of cell cycle by E2Fs
Location analysis in living human cells: E2F4 binding to target genes

Human genome regulation
Regulation of cell cycle by E2Fs
Selected E2F Target Genes

Cell Cycle Regulation
- RB Family: RP•B, p107
- Cyclins: cyclin A, Cdc2, cdc25A
- E2F Family: E2F2, E2F3

DNA Replication
- Initiation: Orc1, Mcm3, Mcm5, Mcm6, Cdc6
- Replication factors: Rfc2, Rfc3, Rfc4, Prim2A, PolA2, Top2A

DNA Repair
- Mismatch repair: MSH2, MLH1
- Base excision repair: UNG
- Nucleotide excision repair: RAD1, RAD54

Checkpoint controls
- DNA damage checkpoint: p53, Chk1
- Mitotic spindle checkpoint: CENPE, Bub3, Mad2

Human genome regulation
Regulation of cell cycle by E2Fs
Summary

Now possible to map genomic targets of regulators in living human cells

E2F binds directly to tumor suppressors and other genes whose mutation results in genomic instability and cancer

E2F links cell cycle progression with regulation of genes involved in DNA repair, replication and G2/M checkpoints

Human genome regulation
Regulation of cell cycle by E2Fs

Bing Ren et al., Genes & Development 16: 245 (2002)