Genome Analysis:

Using Inter-phyla Comparisons to Understand Highly Conserved Aspects of the Human Genome

(Even though human, fly and worm are highly divergent, they share many ancient features of transcription.)

Slides freely downloadable from Lectures.GersteinLab.org & “tweetable” (via @markgerstein). See last slide for references & more info.
How might we annotate a human text?

The Semicolon Wars

Every programmer knows there is one true programming language. A new one every week

[I]f you want to be a thorough-going world traveler, you need to learn 6,912 ways to say “Where is the toilet, please?” That’s the number of languages known to be spoken by the peoples of planet Earth, according to Ethnologue.com.

If you want to be the complete polyglot programmer, you also have quite a challenge ahead of you, learning all the ways to say:

```c
printf("hello, world\n");
```

(This one is in C.) A catalog maintained by Bill Kinnersley of the University of Kansas lists about 2,500 programming languages. Another survey, compiled by Diarmuid Piggott, puts the total even higher, at more than 8,500. And keep in mind that whereas human languages have had millennia to evolve and diversify, all the computer languages have sprung up in just 50 years. Even by the more-conservative standards of the Kinnersley count, that means we’ve been inventing one language a week, on average, ever since Fortran.

For ethnologists, linguistic diversity is a cultural resource to be nurtured and preserved, much like biodiversity. A good-enough notation—for expressing an algorithm or defining a data structure.

There are programmers of my acquaintance who will dispute that last statement. I expect to hear from them. They will argue—zealously, ardently, vehemently—that we have indeed found the right programming language, and for me to claim otherwise is willful ignorance. The one true language may not yet be perfect, they’ll concede, but it’s built on a sound foundation and solves the main problems, and now we should all work together to refine and improve it. The catch, of course, is that each of these friends will decide which end of a boiled egg to crack. This famous tempest in an egg cup was replayed 250 years later by designers of computer hardware and communications protocols. When a block of data is stored or transmitted, either the least-significant bit or the most-significant bit can go first. Which way is better? It hardly matters, although life would be easier if everyone made the same choice. But that’s not what has happened, and so quite a lot of hardware and software is needed just to swap ends at boundaries between systems.

This modern echo of Swift’s Endian wars was first pointed out by Danny Cohen of the University of Southern California in a brilliant 1980 memo, “On holy wars and a plea for peace.” The memo, subsequently published in Computer, was widely read and admired; the plea for peace was ignored.

Another feud—largely forgotten, I think, but never settled by truce or treaty—focused on the semicolon. In Algol and Pascal, program statements have to be separated by semicolons. For example, in `x:=0; y:=x+1; z:=2` the semicolons tell the compiler where one statement ends and the next begins. C
Sources of Annotation: Comparative & Functional
Importance of Dark Matter of the Genome

- Non-coding regions contain the control elements for coding regions.
- Some non-coding regions are functional & are pervasively transcribed.
- “Molecular Fossils” in the non-coding genome represent a historical record of the genome.
- Most disease-associated mutations (e.g. GWAS hits) are in non-coding regions.
Only ~1% of personal genome sequences are canonical protein coding genes regions (CDS).

The remaining 99% are non-coding.
Previous studies have compared RNA transcription between closely related organisms (e.g. RNA-seq within mammals, Brawand et al. '11) ...
Previous studies have compared RNA transcription between closely related organisms (e.g. RNA-seq within mammals, Brawand et al. '11) ...
Previous studies have compared RNA transcription between closely related organisms (e.g. RNA-seq within mammals, Brawand et al. '11) …

modENCODE is the first effort to comprehensively integrate diverse data across distantly related species
Inter-phyla Comparison of Regulation & Transcription

• Here, integrated comparison across phyla
  – Of extensive, matched functional genomics data (on human, worm & fly)
  – Revealing ancient “features” of transcription & regulation
  – help with annotating core features of the human genome
Genome Analysis:

Using Inter-phyla Comparisons to Understand Highly Conserved Aspects of the Human Genome

(Even though human, fly and worm are highly divergent, they share many ancient features of transcription.)

- Lots of Matched Data!
- **Protein Coding Genes & ncRNAs**
  - Uniform density of non-canonical transcription, eg TARs (& eRNAs)
  - Few new ncRNAs similar to existing ones (via incRNA)
- **Cross-species clustering**
  - 16 conserved co-expression modules
  - Hourglass genes in 12 of these
  - Stage alignment of worm & fly development, strongest with hourglass genes
  - Intra-organism hourglass behavior

- **HM models of gene expression**
  - Works for ncRNAs as well as genes
  - Tissue specific
  - Universal cross-species model
- **TF models**
  - Variable importance of regions around genes for chromatin & TFs
  - TF & HM signals are redundant for ‘predicting’ expression
  - Surprisingly, a few TFs are quite predictive
ENCODE/modENCODE data
(encodeproject.org, modencode.org)

Transcriptomes

Regulomes

Also, large-scale chromatin data

575 different experiments containing >67B reads,
modENCODE/ENCODE RNA Resource

- Broad sampling of conditions across transcriptomes of human, worm and fly
- 65 billion reads (585 datasets)
- Comparisons
  - Poly(A)+ RNA
  - embryo & embryonic stem cells
  - developmental time course (worm-fly)
Genome Analysis:

Using Inter-phyla Comparisons to Understand Highly Conserved Aspects of the Human Genome

(Even though human, fly and worm are highly divergent, they share many ancient features of transcription.)

- Lots of Matched Data!
- **Protein Coding Genes & ncRNAs**
  - Uniform density of non-canonical transcription, e.g. TARs (& eRNAs)
  - Few new ncRNAs similar to existing ones (via incRNA)
- **Cross-species clustering**
  - 16 conserved co-expression modules
  - Hourglass genes in 12 of these
  - Stage alignment of worm & fly development, strongest with hourglass genes
  - Intra-organism hourglass behavior
- **HM models of gene expression**
  - Works for ncRNAs as well as genes
  - Tissue specific
  - Universal cross-species model
- **TF models**
  - Variable importance of regions around genes for chromatin & TFs
  - TF & HM signals are redundant for ‘predicting’ expression
  - Surprisingly, a few TFs are quite predictive
Protein-coding gene counts in worm, fly & human have stabilized & have remained fairly constant
Human Genes Are Different in Scale

Comparison of protein coding gene annotations: Human, Worm & Fly

[ENCODE-modencode Transcriptome paper, Nature (in press), doi: 10.1038/nature13424]
Comparison of Protein-Coding Gene Annotation: Similarities in distribution of alternative splicing

- Comparison of protein coding gene annotations: 19,901 (human), 20,377 (worm) & 13,623 (fly)
- No examples of orthologs which share “orthologous” transcript splice isoforms
- Similar overall distribution & distribution of mechanisms
- Some fly genes with extreme numbers of isoforms

Ortholog Example

[ENCODE-modencode Transcriptome paper, Nature (in press), doi: 10.1038/nature13424]
Uniform Annotation of non-coding Elements

- We uniformly processed the RNA-seq expression compendium and for identification of pervasively transcribed regions

[ENCODE-modencode Transcriptome paper, Nature (in press), doi: 10.1038/nature13424]
## Annotated ncRNAs

<table>
<thead>
<tr>
<th></th>
<th>Elements</th>
<th>Human Genome Coverage</th>
<th>Worm Genome Coverage</th>
<th>Fly Genome Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Kb</td>
<td>%</td>
<td>Kb</td>
</tr>
<tr>
<td>mRNAs (exons)</td>
<td>20,007</td>
<td>86,560</td>
<td>3.0</td>
<td>21,192</td>
</tr>
<tr>
<td>Pseudogenes</td>
<td>11,216</td>
<td>27,089</td>
<td>0.95</td>
<td>881</td>
</tr>
<tr>
<td>Comparative ncRNAs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pri-miRNA</td>
<td>58</td>
<td>1,158</td>
<td>0.04</td>
<td>44</td>
</tr>
<tr>
<td>pre-miRNAs</td>
<td>1,756</td>
<td>162</td>
<td>0.006</td>
<td>221</td>
</tr>
<tr>
<td>tRNAs</td>
<td>624</td>
<td>47</td>
<td>0.002</td>
<td>609</td>
</tr>
<tr>
<td>snoRNAs</td>
<td>1,521</td>
<td>168</td>
<td>0.006</td>
<td>141</td>
</tr>
<tr>
<td>snRNAs</td>
<td>1,944</td>
<td>210</td>
<td>0.007</td>
<td>114</td>
</tr>
<tr>
<td>IncRNAs</td>
<td>10,840</td>
<td>10,581</td>
<td>0.37</td>
<td>233</td>
</tr>
<tr>
<td>Other ncRNAs</td>
<td>5,411</td>
<td>3,268</td>
<td>0.11</td>
<td>40,104</td>
</tr>
<tr>
<td>nc-piRNA loci</td>
<td>88</td>
<td>1,272</td>
<td>0.04</td>
<td>35,329</td>
</tr>
<tr>
<td>Total</td>
<td>22,154</td>
<td>17,770</td>
<td>0.62</td>
<td>41,466</td>
</tr>
</tbody>
</table>

Identify non-canonical transcription in regions of the genome excluding mRNA exons, pseudogenes or annotated ncRNAs.

[ENCODE-modencode Transcriptome paper, Nature (in press), doi: 10.1038/nature13424]
& Non-Canonical Transcription

<table>
<thead>
<tr>
<th></th>
<th>Human</th>
<th>Worm</th>
<th>Fly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Elements</td>
<td>Genome Coverage</td>
<td>Elements</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kb</td>
<td>%</td>
</tr>
<tr>
<td>Total ncRNAs</td>
<td>22,154</td>
<td>17,770</td>
<td>0.62</td>
</tr>
<tr>
<td>Regions Excluding mRNAs, Pseudogenes or Annotated ncRNAs</td>
<td>283,816</td>
<td>2,731,811</td>
<td>95.5</td>
</tr>
<tr>
<td>Transcription Detected (TARs)</td>
<td>708,253</td>
<td>916,401</td>
<td>32.0</td>
</tr>
<tr>
<td>Supervised Predictions</td>
<td>104,016</td>
<td>13,835</td>
<td>0.48</td>
</tr>
</tbody>
</table>

- Similar fraction of non-canonical transcription of non-canonical transcription in human, worm and fly
  - 32-37% of each genome
Integrative ncRNA Finder (*incRNA*) for Worm


![Diagram of ncRNA Finder](image_url)

- **Conserved** DNA alignment
- **Annotations**
- **Candidate ncRNAs**
  - Type I ncRNAs
  - Type II ncRNAs
  - Ambiguous Regions
  - Intergenic Regions (unexpressed)

- **Machine Learning**
  - Sequence Features
  - Structural Features
  - Expression Features
  - Genomic Features

- **Known ncRNAs**
  - CDSs
  - UTRs
  - Unexpressed Intergenic

- **Gold-standard Set from Annotations**

- **Table**

<table>
<thead>
<tr>
<th>ncRNA</th>
<th>2” Struc.</th>
<th>Exp. Class</th>
<th>POL II</th>
<th>TFs</th>
</tr>
</thead>
<tbody>
<tr>
<td>ncRNA 1</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>ncRNA 2</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>ncRNA 3</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

- **Local RNA 2” Structure**
- **Expression Pattern Class**
- **POL II Binding**
- **TFs Binding**
IncRNA: Identification of many candidate ncRNAs through evidence integration

- No single feature (e.g. expr. expts., conservation, or sec. struc.) finds all known ncRNAs => combine features in stat. model
- 90% PPV, 13 of 15 tested validate

TAR Characterization

Non-canonical transcription (TARs):

- Mostly transcribed at lower levels than protein-coding genes.

- Enrichment for overlap of TARs with ENCODE enhancers and distal HOT regions $\rightarrow$ potential enhancer RNAs (eRNAs).

**Human, Worm & Fly**

[ENCODE-modencode Transcriptome paper, Nature (in press), doi: 10.1038/nature13424]

HOT Regions = High TF Co-occupancy
Genome Analysis:

Using Inter-phyla Comparisons to Understand Highly Conserved Aspects of the Human Genome

(Even though human, fly and worm are highly divergent, they share many ancient features of transcription.)

- Lots of Matched Data!
- **Protein Coding Genes & ncRNAs**
  - Uniform density of non-canonical transcription, eg TARs (& eRNAs)
  - Few new ncRNAs similar to existing ones (via incRNA)
- **Cross-species clustering**
  - 16 conserved co-expression modules
  - Hourglass genes in 12 of these
  - Stage alignment of worm & fly development, strongest with hourglass genes
  - Intra-organism hourglass behavior

- **HM models of gene expression**
  - Works for ncRNAs as well as genes
  - Tissue specific
  - Universal cross-species model
- **TF models**
  - Variable importance of regions around genes for chromatin & TFs
  - TF & HM signals are redundant for ‘predicting’ expression
  - Surprisingly, a few TFs are quite predictive
A toy example

Every node $i$ is assigned with a spin value $\sigma_i$ (labels of modules: $1,2,\ldots,q$).

$$H = \sum_{i,j} \left( -W_{ij}^{(A)} + p_{ij}^{(A)} \right) \delta_{\sigma_i \sigma_j} + \sum_{i,j} \left( -W_{ij}^{(B)} + p_{ij}^{(B)} \right) \delta_{\sigma_i \sigma_j} - \kappa \sum_{(i,j') \in Ortho} \delta_{\sigma_i \sigma_{j'}}$$

reward a co-expressed pair with the same value
punish a non co-expressed pair with the same value
reward an orthologous pair with the same value
A toy example

Species A

2

2

2

2

Species B

4

4

4

4

co-expressed

orthologs

species A specific

conserved modules

species B specific

The ground state configuration correspond to three modules: 1, 2, 4.
Cross-Species Co-expression Clustering

Use Potts model (generalized Ising model) to simultaneously cluster co-expressed genes within an organism as well as orthologs shared between organisms.

Conservation of modules across # of species
- □ >30
- □ □ 2-30
- □ □ □ □ 1

[ENCODE-modencode Transcriptome paper, Nature (in press), doi: 10.1038/nature13424]
16 Conserved Modules

Conservation of module across number of species

[ENCODE-modencode Transcriptome paper, Nature (in press), doi: 10.1038/nature13424]
ncRNAs associated with modules

Non-canonical transcription (TARs):

- Identify TARs that are significantly correlated and anti-correlated with genes in the 16 modules.

[ENCODE-modencode Transcriptome paper, Nature (in press), doi: 10.1038/nature13424]
Alignment of Developmental Time-Course

For worm & fly find stage-specific genes

We can align developmental stages using fraction of shared orthologs between worm and fly amongst these

Reuse of genes from LE in worm in fly pupa

[ENCODE-modencode Transcriptome paper, submitted]
Hourglass Behavior

**Canonical Inter-organism Behavior**

- “Hourglass hypothesis”: all organisms go through a particular stage in embryonic development ("phylotypic" stage) where inter-organism expression differences of orthologous genes are smallest.
- We identify modules (12 out of 16) which have this behavior at the phylotypic stage.

[ENCODE-modencode Transcriptome paper, submitted]
Using only orthologs in 12 "hourglass" modules show stronger alignment except for absence of genes at the phylotypic stage

- By definition genes in hourglass modules are not phylotypic stage specific, hence the gap
Hourglass Behavior

Intra-organism Behavior also Present

• We observe that the expression of genes across 12 modules are the most tightly coordinated at the phylotypic stage (fly).

• Strongly correlated correlation at phylotypic stage (worm).

[ENCODE-modencode Transcriptome paper, submitted]
Genome Analysis:

Using Inter-phyla Comparisons to Understand Highly Conserved Aspects of the Human Genome

(Even though human, fly and worm are highly divergent, they share many ancient features of transcription.)

• Lots of Matched Data!
  • **Protein Coding Genes & ncRNAs**
    • Uniform density of non-canonical transcription, eg TARs (& eRNAs)
    • Few new ncRNAs similar to existing ones (via incRNA)
  • **Cross-species clustering**
    • 16 conserved co-expression modules
    • Hourglass genes in 12 of these
    • Stage alignment of worm & fly development, strongest with hourglass genes
    • Intra-organism hourglass behavior
  • **HM models of gene expression**
    • Works for ncRNAs as well as genes
    • Tissue specific
    • Universal cross-species model
  • **TF models**
    • Variable importance of regions around genes for chromatin & TFs
    • TF & HM signals are redundant for ‘predicting’ expression
    • Surprisingly, a few TFs are quite predictive
Inputs v Outputs: Upstream Binding/Modification v Expression

[ENCODE-modencode Transcriptome paper, Nature (in press), doi: 10.1038/nature13424]
Histone Modification (HM) model

Chromatin features: Histone modifications

Predictors

RNA-Seq data

Prediction target: Gene expression level

Gene k

HM1, 2, 3, ......

Bin 1

Bin 2

......

Bin160

[Cheng et al. (11), Genome Biol. 12: R15]
His. mods around TSS & TTS are clearly related to level of gene expression, in a position-dependent fashion.

[Science 330:6012]  [Related work: Ouyang et al. ('09) PNAS; Karlic et al. ('10) PNAS]
Human Results

Pearson’s $r = 0.9$ (p-value < $2.2 \times 10^{-16}$)
RMSE = 1.9
Classification: AUC = 0.95
Regression: $r = 0.77$ (RMSE = 2.3)

Relative contribution to $R^2$:
- H3K79me2
- H3K36me3
- DNase I
- H3K9ac
- H3K4me3
- H3K27ac
- H3K4me2
- H2A.Z
- H4K20me1
- H3K4me1
- H3K27me3
- H3K9me1
- Control
- Normalized CpG
- H3K9me3
Integrate all histone modifications to predict gene expression levels

Classify H/L genes (SVM)

Magnitude of Prediction from a “bin” around the TSS

Predict expression values

R=0.75

R=0.60

* = LOG_{10} RPKM

[Cheng et al. (11), Genome Biol. 12: R15]
HM models are tissue specific

HM model-- Best prediction is achieved by using histone modification and expression data from the same developmental stage

[Cheng et al. ('11) Genome Biol. 12: R15]
Comparison of Models for Gene Expression

Human, Worm & Fly

[ENCODE-modencode Transcriptome paper, Nature (in press), doi: 10.1038/nature13424]
Application of chromatin model in 5 species: Consistent Performance

>50% of variation of expression levels can be explained by HMs
Universal, cross-organism Model

We construct single universal HM model (one set of parameters)
  • works almost as well as species specific models
  • works for both mRNAs and ncRNAs

<table>
<thead>
<tr>
<th>Model Trained in</th>
<th>Human</th>
<th>Worm</th>
<th>Fly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>.82</td>
<td>.66</td>
<td>.69</td>
</tr>
<tr>
<td>Worm</td>
<td>.66</td>
<td>.74</td>
<td>.70</td>
</tr>
<tr>
<td>Fly</td>
<td>.69</td>
<td>.68</td>
<td>.84</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prediction Accuracy of Universal Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein coding</td>
</tr>
<tr>
<td>.80</td>
</tr>
<tr>
<td>ncRNA</td>
</tr>
</tbody>
</table>
Genome Analysis:

Using Inter-phyla Comparisons to Understand Highly Conserved Aspects of the Human Genome

(Even though human, fly and worm are highly divergent, they share many ancient features of transcription.)

• Lots of Matched Data!
• **Protein Coding Genes & ncRNAs**
  • Uniform density of non-canonical transcription, eg TARs (& eRNAs)
  • Few new ncRNAs similar to existing ones (via incRNA)
• **Cross-species clustering**
  • 16 conserved co-expression modules
  • Hourglass genes in 12 of these
  • Stage alignment of worm & fly development, strongest with hourglass genes
  • Intra-organism hourglass behavior
• **HM models of gene expression**
  • Works for ncRNAs as well as genes
  • Tissue specific
  • Universal cross-species model
• **TF models**
  • Variable importance of regions around genes for chromatin & TFs
  • TF & HM signals are redundant for ‘predicting’ expression
  • Surprisingly, a few TFs are quite predictive
Doing a Model with TFs: Positive and negative regulators from correlating TF signal at TSS with gene expression

[Cheng et al. (’11) *PLOS CB*]
Predictor v2: 2-levels, now with TFs

[Cheng et al. NAR ('11)]
Human Results

![Graph showing measured and predicted expression levels for CAGE PolyA+ K562 Whole Cell with Pearson's r=0.81; RMSE=2.57. Classification: AUC = 0.89. Regression: r = 0.62; RMSE = 3.06.]

Prediction of Differential Expression

TF model—differential TF binding signals are predictive of differential expression levels between two human cell lines

TFs active in a particular cell type are the strongest predictors in that cell

Models Illuminates Different Regions of Influence for TFs vs HMs

- Datasets
  - ChIP-Seq for 12 TFs (Chen et al. 2008)
  - ChIP-Seq for 7 HMs (Meissner et al.’08; Mikkelsen et al. ’07)
  - RNA-Seq (Cloonan et al. 2008)

A TF+HM model that combine TF and HM features does NOT improve accuracy!

Comparison of TF & HM model across 3 phyla (Human, Worm & Fly)

Comparison of TF model and HM models around TSS (numbers are correlations)
TF model accuracy only needs a small number of TFs for high accuracy (>90%)
Conservation of regulatory networks rewiring

<table>
<thead>
<tr>
<th></th>
<th>Human</th>
<th>Worm</th>
<th>Fly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of TFs (interactions)</td>
<td>155 (1331)</td>
<td>79 (496)</td>
<td>25 (50)</td>
</tr>
<tr>
<td>Fraction of feedback edges</td>
<td>30%</td>
<td>22%</td>
<td>4%</td>
</tr>
<tr>
<td>% of nodes in Bottom, Middle, Top</td>
<td>30, 37, 33</td>
<td>32, 55, 13</td>
<td>46, 47, 7</td>
</tr>
</tbody>
</table>

[ENCODE-modencode Regulation paper, submitted]
Optimally arrange TFs into 3 levels by simulated annealing, maximizing downward-pointing edges

Hierarchy height distribution approximated by 3 levels
Probing direction framed as an optimization problem

Integration of TF hierarchy with other ‘omic information: more influential & connected TFs on the top

Avg. correlation between binding signal of TF & gene expr. of its target

Integration of TF hierarchy with other ‘omic information: more influential & connected TFs on the top

Genome Analysis:

Using Inter-phyla Comparisons to Understand Highly Conserved Aspects of the Human Genome

(Even though human, fly and worm are highly divergent, they share many ancient features of transcription.)

• Lots of Matched Data!
• **Protein Coding Genes & ncRNAs**
  • Uniform density of non-canonical transcription, eg TARs (& eRNAs)
  • Few new ncRNAs similar to existing ones (via incRNA)
• **Cross-species clustering**
  • 16 conserved co-expression modules
  • Hourglass genes in 12 of these
  • Stage alignment of worm & fly development, strongest with hourglass genes
  • Intra-organism hourglass behavior

• **HM models of gene expression**
  • Works for ncRNAs as well as genes
  • Tissue specific
  • Universal cross-species model
• **TF models**
  • Variable importance of regions around genes for chromatin & TFs
  • TF & HM signals are redundant for ‘predicting’ expression
  • Surprisingly, a few TFs are quite predictive
Acknowledgements

modENCODE/ENCODE Transcriptome Group

Joel Rozowsky, Koon-Kiu Yan, Daifeng Wang, Chao Cheng,
James B. Brown, Carrie A. Davis, LaDeana Hillier, Cristina Sisu, Jingyi Jessica Li,
Baikang Pei, Arif O. Harmanci, Michael O. Duff, Sarah Djebari, Roger P. Alexander, Burak H.
Alver, Raymond K. Auerbach, Kimberly Bell, Peter J. Bickel, Max E. Boeck, Nathan P. Boley,
Benjamin W. Booth, Lucy Cherbas, Peter Cherbas, Chao Di, Alex Dobin, Jorg Drenkow, Brent
Ewing, Gang Fang, Megan Fastuca, Elise A. Feingold, Adam Frankish, Guanjun Gao, Peter J.
Good, Phil Green, Roderic Guigó, Ann Hammonds, Jen Harrow, Roger A. Hoskins, Cédric
Howald, Long Hu, Haiyan Huang, Tim J. P. Hubbard, Chau Huynh, Sonali Jha, Dionna Kasper,
Masaomi Kato, Thomas C. Kaufman, Rob Kitchen, Erik Ladewig, Julien Lagarde, Eric Lai, Jing
Leng, Zhi Lu, Michael MacCoss, Gemma May, Rebecca McWhirter, Gennifer Merrihew,
David M. Miller, Ali Mortazavi, Rabi Murad, Brian Oliver, Sara Olson, Peter Park, Michael J.
Pazin, Norbert Perrimon, Dmitri Pervouchine, Valerie Reinke, Alexandre Reymond, Garrett
Robinson, Anastasia Samsonova, Gary I. Saunders, Felix Schlesinger, Anurag Sethi, Frank J.
Slack, William C. Spencer, Marcus H. Stoiber, Pnina Strasbourger, Andrea Tanzer, Owen A.
Thompson, Kenneth H. Wan, Guilin Wang, Huaien Wang, Kathie L. Watkins, Jiayu Wen, Kejia
Wen, Chenghai Xue, Li Yang, Kevin Yip, Chris Zaleski, Yan Zhang, Henry Zheng, Steven E.
Brenner, Brenton R. Graveley, Susan E. Celniker, Thomas R Gingeras,
Robert Waterston

Hiring Postdocs. See gersteinlab.org/jobs!
Default Theme

- Default Outline Level 1
  - Level 2
More Information on this Talk

SUBJECT: Networks

DESCRIPTION:

NOTES:
This PPT should work on mac & PC. Paper references in the talk were mostly from Papers.GersteinLab.org.

PERMISSIONS: This Presentation is copyright Mark Gerstein, Yale University, 2010. Please read permissions statement at http://www.gersteinlab.org/misc/permissions.html . Feel free to use images in the talk with PROPER acknowledgement (via citation to relevant papers or link to gersteinlab.org).

PHOTOS & IMAGES. For thoughts on the source and permissions of many of the photos and clipped images in this presentation see http://streams.gerstein.info . In particular, many of the images have particular EXIF tags, such as kwpotppt , that can be easily queried from flickr, viz: http://www.flickr.com/photos/mbgmbg/tags/kwpotppt .