Evolution of Temporal Processes in Woodland Yeast

Hoa Giang
Laboratory of Junhyong Kim
Mar 24 2009
Outline

A. Introduction

B. Specific aims
   1. Aim 1 – temperature perturbation
   2. Aim 2 – allelic replacement of CLN3

C. Preliminary results
   1. Methodology
   2. Study of nine strains
   3. Study of temperature perturbation in three strains

D. Research design and methods
A. Introduction – temporal processes

- Organismal function involves temporal processes
  - Circadian rhythm
  - Cardiac cycle of the heart
  - Morphogenesis
  - Development
A. Introduction – temporal processes

• A fundamental aspect of temporal process is event control, i.e. which molecule at which time in which place
  – What is the expression pattern in the morning, afternoon, and evening?
  – Jet lag → What controls the response to changes?

• Mechanisms underlying temporal control involve
  – Autonomous process
  – Inductive process
  – External signals
A. Introduction – cell cycle as a key temporal process

- A key temporal process is the cell division cycle
- Many cell-cycle characteristics are shared among eukaryotes, including cell-cycle phases, cyclin-dependent kinase (cdk) system and checkpoint controls (Rupes 2002)

→ Suggesting that the cell cycle is universally conserved

<table>
<thead>
<tr>
<th>Function</th>
<th>Phase</th>
<th>S. cerevisiae (Budding Yeast)</th>
<th>S. pombe (Fission Yeast)</th>
<th>D. melanogaster (Fruit Fly)</th>
<th>X. laevis (Clawed Toad)</th>
<th>H. sapiens (Human)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyclin-dependent kinases (Cdns)</td>
<td>G1</td>
<td>Cln3 (binds Cdk1)</td>
<td>Puc 1? (binds Cdk1)</td>
<td>Cyclin D (binds Cdk4)</td>
<td>Cyclin D (binds Cdk4)</td>
<td>Cyclin D1, D2, D3 (bind Cdk4 or Cdk8)</td>
</tr>
<tr>
<td></td>
<td>G1/S</td>
<td>Cln1, 2 (bind Cdk1)</td>
<td>Puc1, Cig1? (bind Cdk1)</td>
<td>Cyclin E (binds Cdk2)</td>
<td>Cyclin E (binds Cdk2)</td>
<td>Cyclin E (binds Cdk2)</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>Clb5, 6 (bind Cdk1)</td>
<td>Cig2, Cig1? (bind Cdk1)</td>
<td>Cyclin A (binds Cdk1)</td>
<td>Cyclin A1, A2 (bind Cdk2, Cdk1)</td>
<td>Cyclin A1, A2 (bind Cdk2, Cdk1)</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>Clb1, 2, 3, 4 (bind Cdk1)</td>
<td>Cdc13 (binds Cdk1)</td>
<td>Cyclin B, B3 (bind Cdk1)</td>
<td>Cyclin B1, B2 (bind Cdk1)</td>
<td>Cyclin B1, B2 (bind Cdk1)</td>
</tr>
</tbody>
</table>

Morgan, 2008
A. Introduction – cell cycle as a key temporal process

Budding yeast has been used as a model of cell cycle progression

Cid et al., 2002
A. Introduction – evolution of temporal control

- The change in the temporal process is a mode of developmental evolution, e.g. paedomorphosis
- Heterochrony – change in timing or rate of developmental events, found in earlier or later stages of development
- Heterochrony has been widely studied in terms of morphological evolution
- Heterochrony is genetic-based, e.g. *lin-14* in *C. elegans* or homeotic genes
- Recently, measurement of heterochrony at molecular level has increased our understanding of developmental timing
A. Introduction

- Traditionally, cyclin-dependent kinases and cell-cycle checkpoints have governed cell-cycle progression.

- However, a recent study (Orlando et al., 2008) showed that yeast has continued cell-cycle progression (in terms of transcriptome) without cdk system. The authors introduced a new control mechanism by transcription factor network.

→ Studies of transcriptome dynamics reveal the complexity of cell-cycle progression.

Orlando et al., 2008
A. Introduction – key questions

• How does cell cycle progression constrain evolutionary variation in gene expression?
  – Evolutionary variation across strains and across time

• How do evolutionary forces alter cell cycle-dependent transcriptional regulation?
  – Natural selection: directional/balancing/purifying
  – Mutation

• How do environmental and genetic changes affect transcriptome dynamics?
A. Introduction - “Congruence” hypothesis

The response of an organism to environmental (non-genetic) perturbations is congruent with their response to genetic perturbations (i.e., mutations and recombination), because both perturbations are processed through system-level dynamics.
A. Introduction

- **Main subject**: temporal control architectures of transcriptome dynamics

- **Guiding hypothesis**: Evolutionary variation mostly involves timing changes in transcriptome dynamics, a proxy of temporal control architectures; and The response characteristics of expression dynamics to perturbations can predict evolutionary patterns

- **Methodology**: comparative transcriptome analysis of woodland yeast under global and local perturbations
B. Specific aims

• **Aim 1:** *Comparative genomic analysis of temporal trajectory across different strains of yeast under temperature perturbation*

• **Aim 2:** *Comparative genomic analysis of temporal trajectory across different strains of yeast under allelic replacement of Cln3, a key G1/S control gene in cell cycle*
B. Specific aims – models of time-domain transformation by perturbation

Temporal expression trajectories

uniform dilation

phase shift

non-uniform dilation
B. Specific aims – control architectures

- Transcriptome dynamics is a proxy of temporal control architectures
- Preliminary studies demonstrated transcriptome trajectory differences by strains and perturbations
- Trajectory changes may express either uniform dilation (namely, robust architecture) or phase shift/non-uniform dilation (so-called tunable architecture) under perturbation
- Comparison of time-course expression with and without perturbations will help place genes into the two architectures
B. Specific aims – Aim 1 rationale

• Temperature has a global effect on kinetics

• Preliminary studies showed temperature decrease induces both time-dilations and shifts in timing of peak in expression trajectories

• Temporal trajectory changes by temperature perturbations will help dissect the genes into two categories of control architecture: robust vs. tunable

• Genes with different time-domain transformation probably show similar time-domain changes across strains

• Increase the density of time-course will avoid the artifact of interpolation in hypothesis testing
B. Specific aims – **Aim 1 design**

- **Approach:** I will measure **time-course** transcriptome dynamics of 5 natural strains of *S. cerevisiae* under 3 different temperature regimes (14, 18, and 22 °C) using custom 2-channel glass arrays. Each time-course will include **36 time-points** covering one and a third of cell cycle.

- **Expected outcome:** Microarray data will consist of a time series of transcriptome dynamics for 5 strains under 3 temperature regimes. Analysis will provide two categories of genes according to their control architecture types.
B. Specific aims – Aim 2 rationale

• Cyclin Cln3 is a key protein in cell cycle regulation at G1/S phase

• Constructed S288c strains, into which natural variants of Cln3 have been transformed, displayed phenotypic changes

• Local gene-specific perturbations can propagate to larger systemic effects and control architectures might express divergent responses

• Comparison of temporal trajectory changes in these constructs may reveal control architectures with respect to local perturbations
B. Specific aims – Aim 2 design

- **Approach:** I will measure *time-course* transcriptome dynamics of 4 constructed S288C-derived strains, in which the *CLN3* allele has been replaced by a natural allele of *CLN3*, using custom microarrays as in Aim 1. Each time-course will include 18 time-points covering one and a third of cell cycle.

- **Expected outcome:** Microarray data will consist of a time series of transcriptome dynamics for 4 constructed strains. Analysis will provide two categories of genes according to their control architecture types.
C. Preliminary results - methodology

• Natural collection of Saccharomyces:
  – Collection of 45 *S. cer* and 90 *S. par*, including woodland strains and Italian vineyard strains (in laboratory of Dr. Paul Sniegowski)
  – Heterothallic haploid, MATa, and HO-knockout derivatives for previous nine strains are available.

• Sampling method:
  – Synchronization using α mating factor
  – 18 time-points across 1.3 cell division cycles with respect to cell cycle length
C. Preliminary results - methodology

• Transcriptome measurement:
  – mRNA is extracted from cell culture with RNeasy kit
  – Fluorescence-labeled cDNA is generated from 15ug of mRNA
  – cDNAs are hybridized to custom microarrays, containing ~6360 70-mer yeast oligonucleotides
  – Unsynchronized S288c-derived strain is used as common reference
  – 2 dye-swapped technical replicates for each sample
  – Manually quality-controlled for image processing, and normalized for dye bias using lowess regression
C. Preliminary results - a study of 9 strains (Simola)

• 8 woodland strains of *S.cer* and 1 derivative of S288c (named YPS183)
• Heterothallic, haploid, MATa, and *HOΔ:kanMX* isolates
• Synchronization with α mating factor
• Sampling for 18 time-points in minimal SD medium at 18°C
• Hybridize to 2-channel spotted microarrays with 2 dye-swap replicates for each sample
• Each slide was controlled for good quality and normalized for dye bias
• Analysis of temporal transcriptome variation and covariance
C. Preliminary results - a study of 9 strains (Simola)

Cell size and cell cycle length variation suggested statistically significant genetic variability (ANOVA test at 5% level)
C. Preliminary results - a study of 9 strains (Simola)

Gene expression variability of 9 strains

- Average trajectory
- Expression derivative
- Relative line variance
- Proportion of genes under purifying selection
- Budding index
C. Preliminary results - a study of 9 strains (Simola): major findings

- Most genes at most times are under strong purifying selection
  → Entire temporal trajectory is likely functionally important

- Time-point 111 near the end of S phase shows lowered proportion of genes under purifying selection
  → Expression near the end of cell division is not as critical as other times

- There are significant changes in the covariance pattern across different time-points
  → Regulatory architectures at different time-points are differentiated from each other
C. Preliminary results - study of temperature perturbation in 3 strains

- 2 woodland strains (YPS2073 and YPS2055) and the lab strain (YPS183)
- Sampling for 18 time-points in minimal SD medium at 14°C, following similar procedures as at 18°C
- Cell-cycle length increases approximately 2-fold for all 3 strains
  → Calibration of cell-cycle length is necessary for temperature perturbation
- Microarray has been processed as described in methodology
C. Preliminary results - study of temperature perturbation in 3 strains

Expression trajectories of a SBF-regulated gene CLN2 (YPL256C) and a MBF-regulated RNR1 (YER070W)
C. Preliminary results - study of temperature perturbation in 3 strains

• Non-linear regression for transformation

\[ f_j^i(t) \sim a f_k^i(\alpha t^2 + \beta t + \gamma) + b + \delta \]

• Estimate (\(\alpha, \beta, \gamma\)) parameters for each gene across temperatures or strains

• Hypothesized models of hierarchical test
  – Model 1: no significant relationship of the time trajectories
  – Model 2: only the uniform dilation fit (\(\beta\) is non-zero)
  – Model 3: both uniform dilation and phase shift fit (\(\beta, \gamma\) are non-zero)
C. Preliminary results - study of temperature perturbation in 3 strains

- Number of genes with a significant time effect (model 2 or 3) in evolutionary comparisons

<table>
<thead>
<tr>
<th>Strain</th>
<th># Genes with significant time effect</th>
<th># Remaining genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>YPS183</td>
<td>4775 (79%)</td>
<td>1307 (21%)</td>
</tr>
<tr>
<td>YPS2055</td>
<td>4841 (80%)</td>
<td>1241 (2%)</td>
</tr>
<tr>
<td>YPS2060</td>
<td>5185 (85%)</td>
<td>897 (15%)</td>
</tr>
<tr>
<td>YPS2066</td>
<td>5213 (86%)</td>
<td>869 (14%)</td>
</tr>
<tr>
<td>YPS2067</td>
<td>5177 (85%)</td>
<td>905 (15%)</td>
</tr>
<tr>
<td>YPS2073</td>
<td>5204 (86%)</td>
<td>878 (14%)</td>
</tr>
<tr>
<td>YPS2079</td>
<td>4985 (82%)</td>
<td>1097 (18%)</td>
</tr>
<tr>
<td>YPS3060</td>
<td>5000 (82%)</td>
<td>1082 (18%)</td>
</tr>
<tr>
<td>YPS3395</td>
<td>5101 (84%)</td>
<td>981 (16%)</td>
</tr>
</tbody>
</table>
### C. Preliminary results - study of temperature perturbation in 3 strains

<table>
<thead>
<tr>
<th>Evolutionary Comparison</th>
<th>Temperature Comparisons</th>
<th>Model I</th>
<th>Model II</th>
<th>Model III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>YPS183</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evolutionary Comparison</td>
<td>Non-Time-Domain</td>
<td>527 (+70.9)</td>
<td>515 (+4.7)</td>
<td>608 (-75.7)</td>
</tr>
<tr>
<td></td>
<td>Time-Domain</td>
<td>1154 (-70.9)</td>
<td>1366 (-4.7)</td>
<td>1912 (+75.7)</td>
</tr>
<tr>
<td></td>
<td><strong>YPS2055</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evolutionary Comparison</td>
<td>Non-Time-Domain</td>
<td>586 (+17.2)</td>
<td>600 (+22.4)</td>
<td>721 (-39.7)</td>
</tr>
<tr>
<td></td>
<td>Time-Domain</td>
<td>1228 (-17.2)</td>
<td>1242 (-22.4)</td>
<td>1705 (+39.7)</td>
</tr>
<tr>
<td></td>
<td><strong>YPS2073</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evolutionary Comparison</td>
<td>Non-Time-Domain</td>
<td>514 (+14.3)</td>
<td>531 (+10.4)</td>
<td>720 (-24.7)</td>
</tr>
<tr>
<td></td>
<td>Time-Domain</td>
<td>1208 (-14.3)</td>
<td>1263 (-10.4)</td>
<td>1846 (+24.7)</td>
</tr>
</tbody>
</table>
C. Preliminary results - study of temperature perturbation in 3 strains: major findings

• Environmental (temperature) and genetic (strain) perturbations result in changes in temporal trajectories
• Most of the genes show significantly better explanation by a time-domain transformation across at least one strain pair
• Genes with non-uniform dilation by temperature are more likely to show evolutionary time-domain changes
D. Research design and methods

Aim 1: comparative genomic analysis of temperature perturbation

• Rationale:
  – Temporal trajectories of gene expression changes follow different time dilation models under temperature perturbation
  – Lack of a fine-scale trajectory leads to the application of interpolation methods, which introduces artifacts into the results
  – Small number of strains reduces the power of statistical test
  – Is there any association with gene functional class (e.g. using GO terms)?
Aim 1: comparative genomic analysis of temperature perturbation

• Experimental design: transcriptome measurement
  – 36 time-points covering 1.3 cell cycles
  – 5 strains: 4 woodland strains and the lab strain
  – 3 temperatures: 14, 18, and 22°C
  – 2 dye-swap replicates

→ Total: 2(rep)x5(strain)x36(time)x3(temp)=1080 slides
D. Research design and methods

Aim 1: comparative genomic analysis of temperature perturbation

• Data analysis:
  
  (A) Temporal variation analysis
  
  (B) Temporal covariance analysis
  
  (C) Hierarchical test for time-domain transformation models
  
  (D) Functional association analysis using GO terms
D. Research design and methods
Aim 1: comparative genomic analysis of temperature perturbation

(A) Temporal variation analysis:
- Compute expression trajectory of ~6000 genes, all strains
- Estimate derivative of expression trajectories
- Test for stabilizing selection using F-ratio to mutational variance (estimated from mutation accumulation lines in a collaborative study with Clifford Zeyl at Wake Forrest)

(B) Temporal covariance analysis:
- Reduce number of variation dimension by Singular Value Decomposition (SVD) method
- Estimate covariance matrices across time-points and across species
- Matrix comparison using methods listed in (Steppan et al, 2002)
D. Research design and methods

Aim 1: comparative genomic analysis of temperature perturbation

(C) Hierarchical test of time-domain transformation:

- Non-linear regression for transformation: \( f_j^i(t) \sim af_k^i(\alpha t^2 + \beta t + \gamma) + b + \delta \)
- Estimate \((\alpha, \beta, \gamma)\) parameters for each gene across temperatures
- Categorize into 3 models of hierarchical test
- Test hypothesis that genes in model 3 are likely to be more evolutionarily labile (time-domain transformation across strains)

(D) Functional association analysis analysis using GO terms:

- Assess association of the 3 models with functional inferences using GO terms
- Test association of model 2 with periodic genes
D. Research design and methods

Aim 2: comparative genomic analysis of *CLN3* allelic replacement

• **Rationale:**
  - Temporal trajectories of gene expression changes follow different time dilation models among strains
  - Cyclin Cln3 is a key regulator of cell cycle, specifically at G1/S transition phase
  - Transformation of natural Cln3 into the lab strain significantly affects cell size of the lab strain
  - Local perturbation can reveal the control architectures of gene dynamics, which might be a small effect
D. Research design and methods

Aim 2: comparative genomic analysis of $CLN3$ allelic replacement

• Experimental design: **allelic replacement using adaptamers**
  
  – Subclone natural $CLN3$ and $URA3$ (*K. lactis*) using adaptamers and PCR
  
  – Fuse $CLN3$ and $URA3$ fragments by matching the adaptamers
  
  – Transform into YPS183 and select for Ura3 marker
  
  – Counter-select with 5-FOA for single copy alleles
  
  – 4 constructed strains are made with 4 natural $CLN3$ alleles from 4 woodland strains
D. Research design and methods

Aim 2: comparative genomic analysis of $CLN3$ allelic replacement

- Experimental design: **transcriptome measurement**
  - 18 time-points covering 1.3 cell cycles, 12 of which is near G1/S phase
  - 4 constructed strains: $CLN3$ alleles will be replaced with 4 natural $CLN3$ alleles from 4 woodland strains
  - 4 dye-swap replicates

$\Rightarrow$ Total: $4(rep) \times 4(strain) \times 18(time) = 288$ slides
D. Research design and methods

Aim 2: comparative genomic analysis of \textit{CLN3} allelic replacement

- Data analysis: identical to that of Aim 1
  
  (A) Temporal variation analysis
  
  (B) Temporal covariance analysis
  
  (C) Hierarchical test for time-domain transformation models
  
  (D) Functional association analysis using GO terms
Acknowledgement

• My advisor: Dr. Junhyong Kim

• Kim lab members:
  – Stephen Fisher
  – Chantal Francis
  – Daniel Simola
  – Miler Lee
  – Sheng Guo
  – Shreedhar Natarajan
“how development in its entirety has evolved”

Wolpert, 1999