

# Evolution of the Bifurcation Phenotype

Phylogenetic Reconstruction and  
Inverse Dynamical Analysis  
of yeast GATA networks

Rainer Machné, James Lu  
Lukas Endler, Stefan Müller, Stefanie Widder, Xtof Flamm,  
Peter Schuster, Heinz Engl, Dougie Murray

TBI, University of Vienna, Austria

Inverse Methods Group, RICAM Linz, Austria

Systems Biology Institute, Tokyo

Bled 2007

# Outline

- ▶ Why yeast GATA factors ?
  - ▶ respiratory/metabolic oscillations in yeast
- ▶ What are yeast GATA factors ?
  - ▶ Nitrogen- and general amino acid metabolism
  - ▶ Homology to metazoan developmental GATAs ?
- ▶ Inverse dynamical analysis
  1. What are *bifurcation phenotypes* ?
  2. bistability and oscillation vs. duplication and mutation
- ▶ Phylogenetic reconstruction
  - ▶ predict *bifurcation phenotypes* in 15 species, 49 proteins

# Outline

- ▶ Why yeast GATA factors ?
  - ▶ respiratory/metabolic oscillations in yeast
- ▶ What are yeast GATA factors ?
  - ▶ Nitrogen- and general amino acid metabolism
  - ▶ Homology to metazoan developmental GATAs ?
- ▶ Inverse dynamical analysis
  1. What are *bifurcation phenotypes* ?
  2. bistability and oscillation vs. duplication and mutation
- ▶ Phylogenetic reconstruction
  - ▶ predict *bifurcation phenotypes* in 15 species, 49 proteins  
AreA 3'UTR:
    - ▶ my second RNA (((...;-)...)))

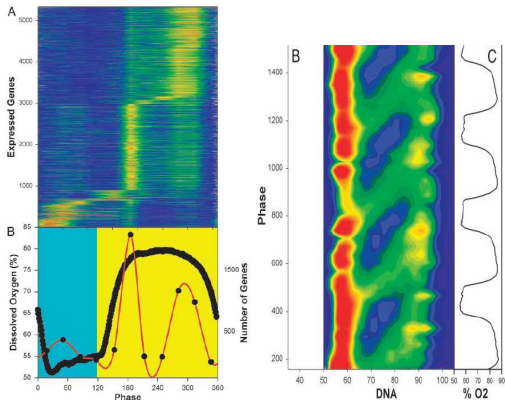
# Outline

- ▶ Why yeast GATA factors ?
  - ▶ respiratory/metabolic oscillations in yeast
- ▶ What are yeast GATA factors ?
  - ▶ Nitrogen- and general amino acid metabolism
  - ▶ Homology to metazoan developmental GATAs ?
- ▶ Inverse dynamical analysis
  1. What are *bifurcation phenotypes* ?
  2. bistability and oscillation vs. duplication and mutation
- ▶ Phylogenetic reconstruction
  - ▶ predict *bifurcation phenotypes* in 15 species, 49 proteins  
AreA 3'UTR:
    - ▶ my second RNA (((...;-)...)))
- ▶ What group I introns (my first) have to do with it ?
- ▶ What miRNA induced DNA mutation has to do with it ?
- ▶ What hyaluronic acid has to do with it ?

# Respiratory oscillations in yeast

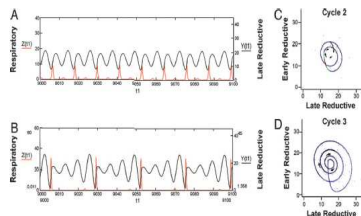
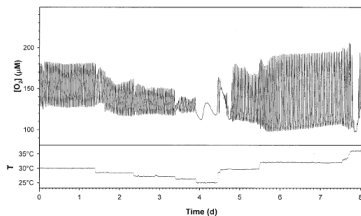
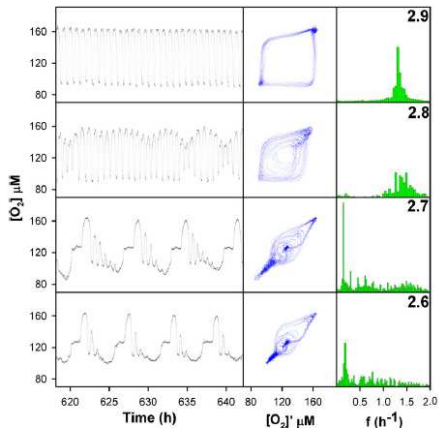
## The redox attractor:

- ▶ 3 genome-wide transcription phases
- ▶ 1 oxidative phase: make ATP (and ROS)
  - ▶ **S metabolism and ribosomes**
- ▶ 2 reductive phases: use ATP
  - ▶ **(re-)build and charge**
- ▶ gating of DNA synthesis and cell cycle !!



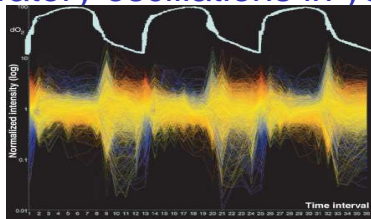
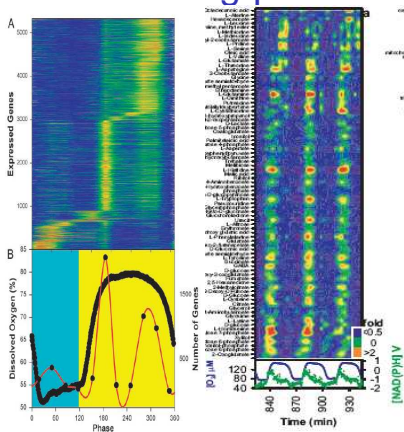
Klevecz et al. 2004 PNAS and (confirmed by) Tu et al. 2005 Science  
Murray et al. 1999 Microbiology, Lloyd et al. 2002 Microbiology  
Murray 2004 Current Genomics, Lloyd and Murray 2005 Trends Biochem Sci  
in rat muscle cells: see papers by Miguel Aon and Sonia Cortassa

# A tunable attractor, e.g. period doubling bifurcations

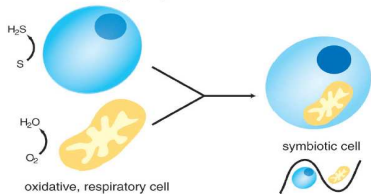


Murray and Lloyd 2006 Biosystems  
 Murray et al. 2001 J Bacteriol  
 Li and Klevecz 2006 PNAS

# Short- and long-period respiratory oscillations in yeast



reductive, non-respiratory cell



Short-period: 40 minutes, cell-cycle gating ?

Long-period: 300 minutes, cell-cycle synchrony ?

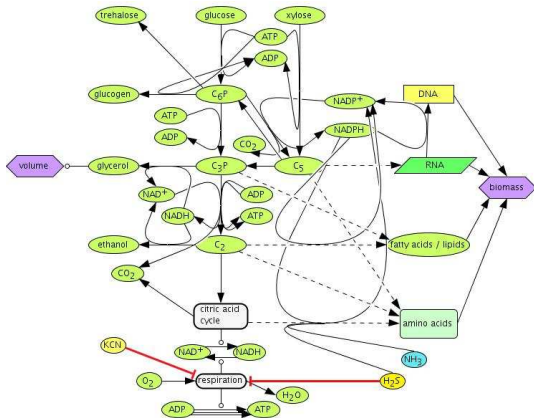
Klavec et al. 2004 PNAS

Murray, Beckmann, Kitano 2007 PNAS

Tu et al. 2005 Science

Murray et al. 1999 Microbiology, Lloyd et al. 2002 Microbiology  
 Murray 2004 Current Genomis, Lloyd and Murray 2005 Trends Biochem Sci  
 in rat muscle cells: see papers by Miguel Aon and Sonia Cortassa

# Metabolic models - glycolytic oscillations

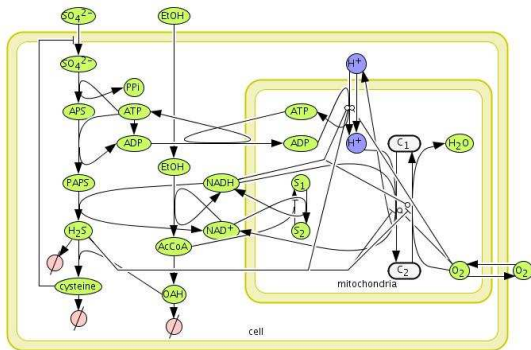


## Similarities:

- ▶ energy (ATP) and redox (NAD(P)H) oscillations
- ▶ *acetaldehyde* (C<sub>2</sub>) acts as population synchronizer
- ▶ H<sub>2</sub>S is the major synchronizer for respiratory oscillation
- ▶ H<sub>2</sub>S and KCN both inhibit the electron transport chain
- ▶ Roughly an order of magnitude different period!



# Metabolic models - sulfate assimilation



## metabolic feedback in sulfate assimilation

glycolytic oscillator

TCA/respiration

gene regulatory network

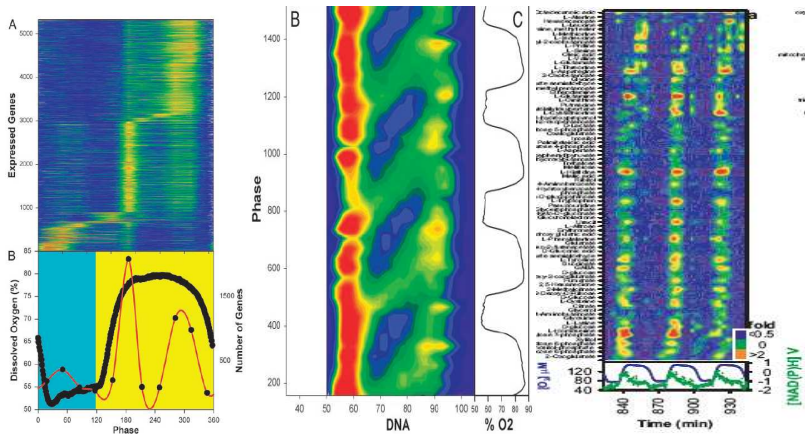
⇒ coupled oscillators : different timescales

⇒ multistability of TCA : respiration vs. amino acids

⇒ coupled oscillators : cell cycle gating

Wolf, Sohn, Heinrich, Kuriyama, 2001 FEBS Letters

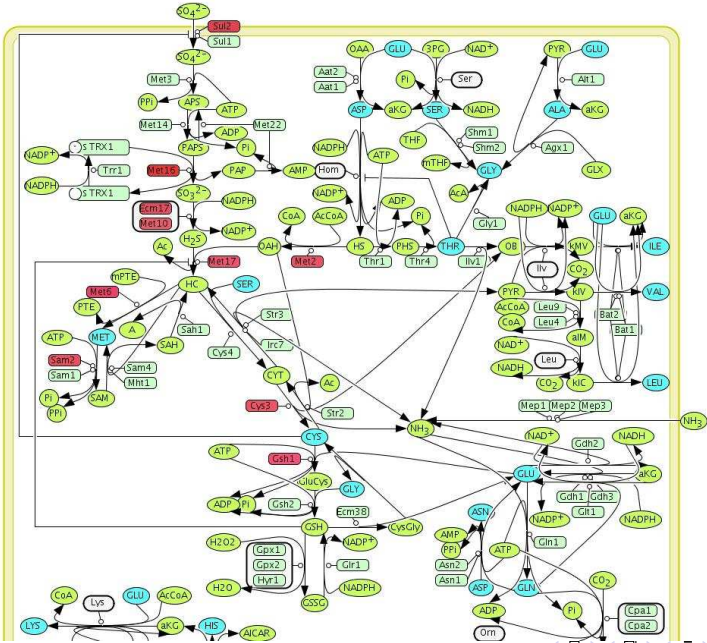
# A cell-wide oscillation



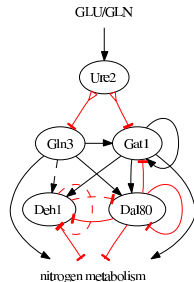
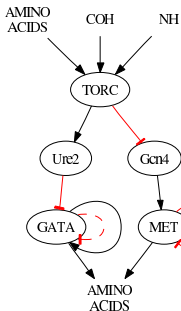
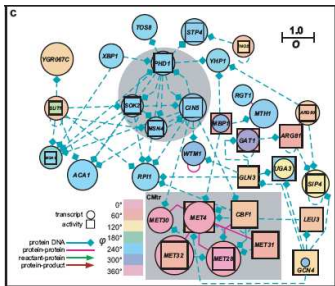
a stable cell-wide oscillation in continuous yeast culture without any nutrient deficiency etc. : **default state !!(?)**

Klavec et al. 2004 PNAS  
Murray, Beckmann, Kitano 2007 PNAS

# Metabolic models - amino acid synthesis

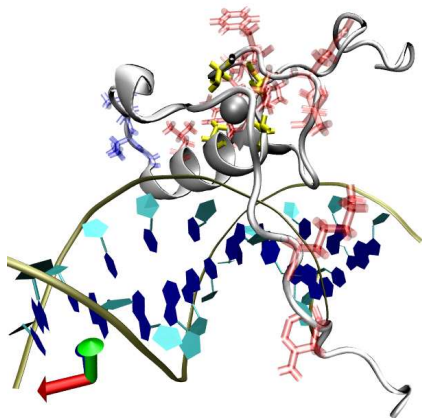


# Gene regulatory models: MET and GATA networks



- ▶ TOR kinase complex: general nutrient **starvation** response
- ▶ GCN4 controls general amino acid synthesis
- ▶ GATA network controls nitrogen metabolism (GLU, GLN)
- ▶ MET network controls sulfur metabolism (MET, CYS)
  - ▶ **GLU + CYS + GLY** → **glutathione**
  - ▶ **MET + ATP** → **S-adenosylmethionine**

# What are GATA factors ?

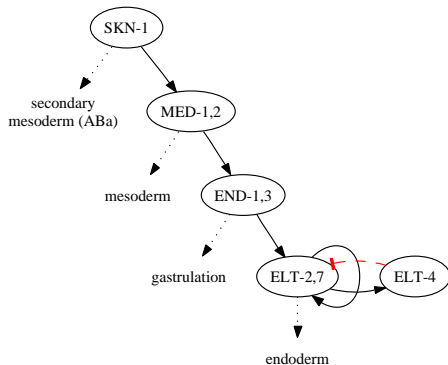
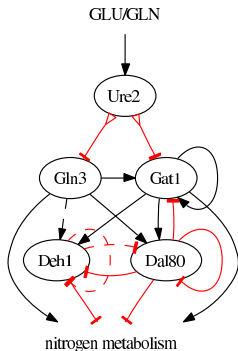


*A. nidulans* AreA  
DNA-binding domain

## Transcription factors

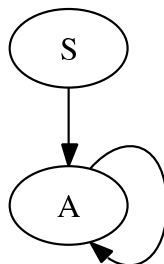
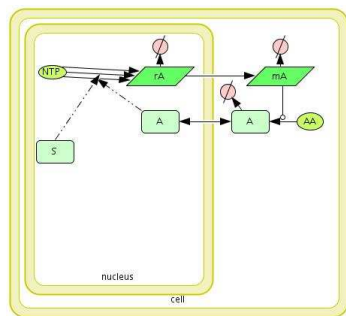
- ▶ bind (**multiple adjacent**) HGATAR sites
- ▶ act as transcriptional activators **and / or** repressors
- ▶ (also) involved in chromatin remodelling
- ▶ yeast: competitive inhibitors - w/o activating domain
- ▶ **often auto-regulatory!**

# Yeast and *C. elegans* GATA networks



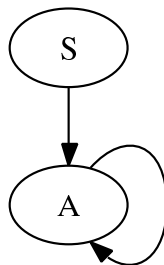
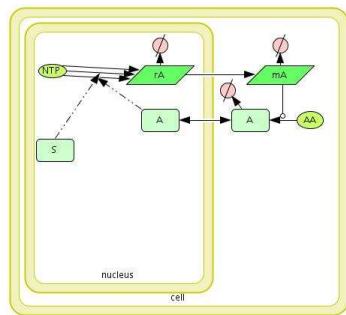
- ▶ Gln3/Gat1 regulates central nitrogen metabolism: GLU/GLN
- ▶ activates general a.a. synthesis, e.g. **amino acid permeases**
- ▶ Elt-2 is the terminal transcription factor of a GATA cascade
- ▶ activates intestine specific genes, e.g. **amino acid permeases**

# GATA network: model and parameters



1. minimal model assumptions:  
all linear, except for transcriptional activation:  $V_{max} * (A / (K_d + A))^2$
2. parameters and variables:  
chosen to accord to recent quantitative data (in orders of magnitude)  
 $K_d = 18nM \approx 10$  molecules/femtolitre  
max. transcription rate : 6-8 transcripts/minute
3. Inverse Dynamical Analysis: how can duplication events and subsequent mutations change the dynamic repertoire of the system ?

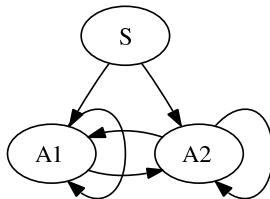
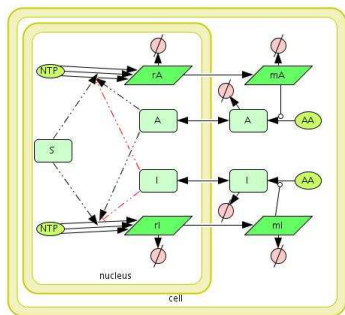
# GATA network: model and parameters



1. minimal model assumptions:  
all linear, except for transcriptional activation:  $V_{max} * (A / (K_d + A))^2$
2. parameters and variables:  
chosen to accord to recent quantitative data (in orders of magnitude)  
 $K_d = 18nM \approx 10$  molecules/femtolitre  
max. transcription rate : 6-8 transcripts/minute  
< **1 RNA/cell** (..1998.., Holland MJ 2002 JBC)
3. Inverse Dynamical Analysis: how can duplication events and subsequent mutations change the dynamic repertoire of the system ?

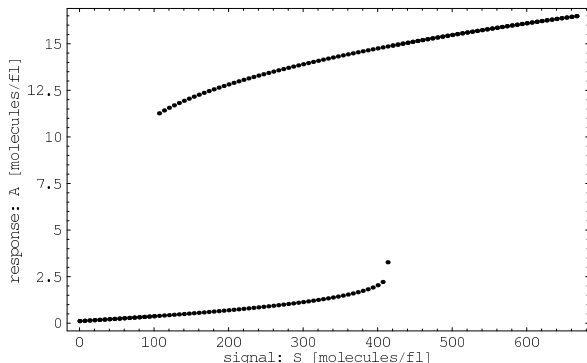
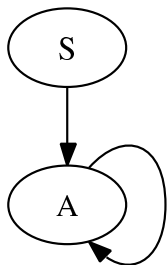


# GATA network: model and parameters



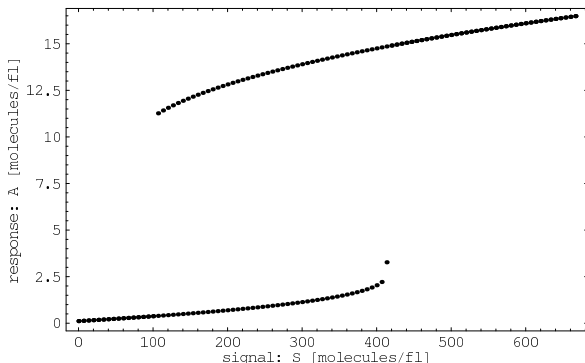
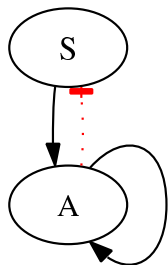
1. minimal model assumptions:  
all linear, except for transcriptional activation:  $V_{max} * (A / (K_d + A))^2$
2. parameters and variables:  
chosen to accord to recent quantitative data (in orders of magnitude)  
 $K_d = 18nM \approx 10$  molecules/femtolitre  
max. transcription rate : 6-8 transcripts/minute
3. **Inverse Dynamical Analysis: how can duplication events and subsequent mutations change the dynamic repertoire of the system ?**

# Auto-activator: hysteresis



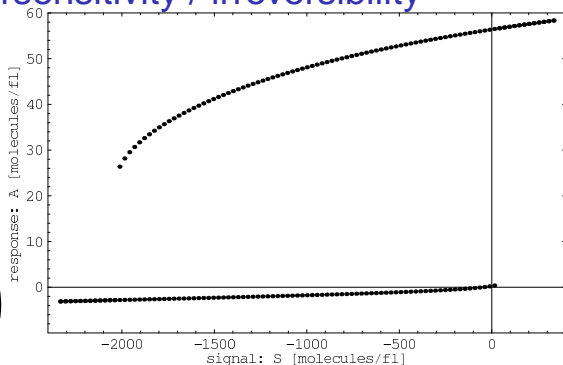
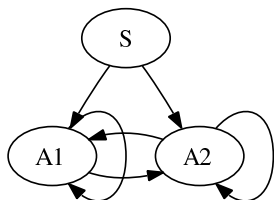
- ▶ **S = low GLN concentration** (GLN3 in yeast)
- ▶ A recovers GLN concentration by synthesis and transport
- ▶ lower S does not mean that A is not needed anymore!!
- ▶ S should really decrease significantly to switch-off A
  - ▶ ⇒ avoid oscillating GLN deficiency

# Auto-activator: hysteresis (... but ... why?)



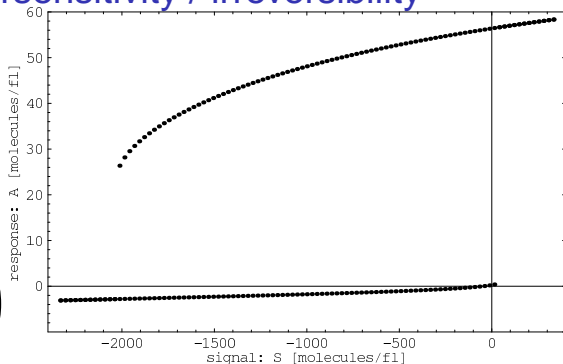
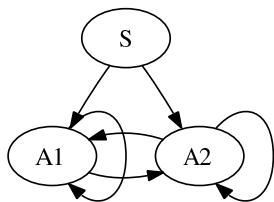
- ▶ S = low GLN concentration (GLN3 in yeast)
- ▶ A recovers GLN concentration by synthesis and transport
- ▶ lower S does not mean that A is not needed anymore!!
- ▶ S should really decrease significantly to switch-off A
  - ▶ ⇒ **avoid oscillating GLN deficiency**

# A duplication: hypersensitivity / irreversibility



- ▶ duplication leads to **hypersensitivity** or even **irreversibility** wrt to S
- ▶ immunological/pathogenic context, e.g.
  - ▶ IL?  $\rightarrow$  STAT6  $\rightarrow$  GATA-3  $\rightarrow$  GATA-3  $\rightarrow$   $T_{h2}$  activation
- ▶ **haploinsufficiency** ? move bif. to the right
  - ▶ development of haploid lifestyle in yeasts
  - ▶ haploinsufficiency diseases known for both GATA-3 and GATA-6

# A duplication: hypersensitivity / irreversibility



- ▶ duplication leads to **hypersensitivity** or even **irreversibility** wrt to S
- ▶ immunological/pathogenic context, e.g.
  - ▶ IL?  $\rightarrow$  STAT6  $\rightarrow$  GATA-3  $\rightarrow$  GATA-3  $\rightarrow$   $T_{h2}$  activation
- ▶ **haploinsufficiency** ? move bif. to the right
  - ▶ development of haploid lifestyle in yeasts
  - ▶ haploinsufficiency diseases known for both GATA-3 and GATA-6

# Inverse Dynamical Analysis of *bifurcation phenotypes*

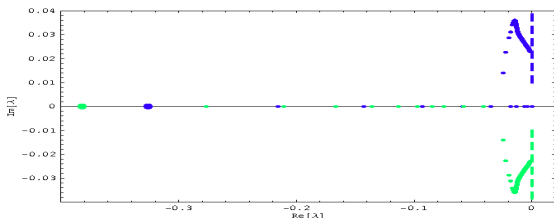
## 1. Inverse eigenvalue analysis

- ▶ Probe the possibility for multistability/oscillations
- ▶ Identify minimal dynamic-changing parameter combinations using sparsity-promoting regularization

## 2. Forward bifurcation analysis (w.r.t identified parameters)

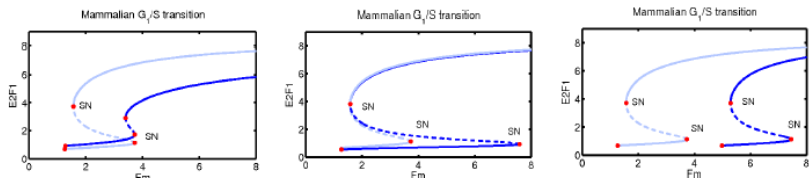
## 3. Inverse bifurcation analysis

- ▶ Design the system to obtain desired dynamical behavior
- ▶ Identify influential parameters: propose experiments for verification



⇒ MathSBML/Mathematica package by James Lu,  
might go into SOSlib extension later

# Relieving hypersensitivity / irreversibility



James Lu, Machne R, Schuster P, Engl H 2007 BIRD07 conference paper

- ▶ use IEA/IBA to find relieving mutations
  - ▶ scaled model to allow for individual mutations:

$$K_{a1}^{A1} = bA1 \cdot ba1 \cdot K_d$$

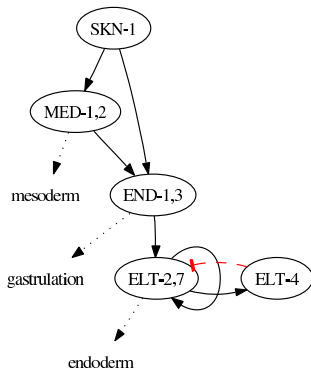
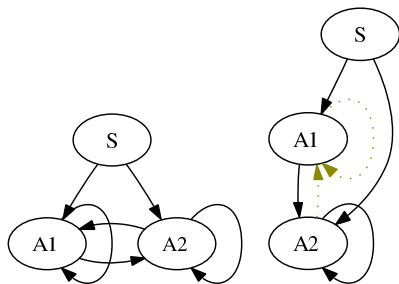
$$K_{a2}^{A1} = bA1 \cdot ba2 \cdot K_d$$

$$K_{a1}^{A2} = bA2 \cdot ba1 \cdot K_d$$

$$K_{a2}^{A2} = bA2 \cdot ba2 \cdot K_d$$

- ▶ e.g. mutation of hGATAr binding site in gene a1: *ba1*
- ▶ e.g. mutation of hGATAr binding domain in protein A1: *bA1*

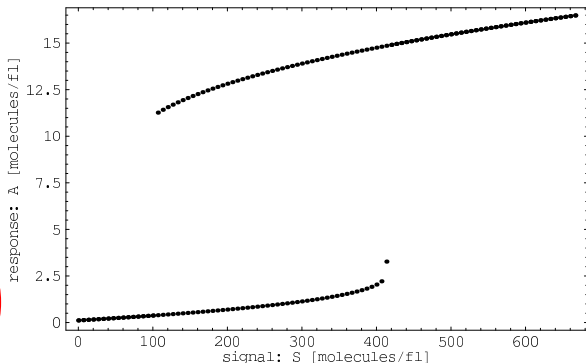
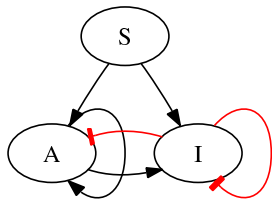
# Generation of a cascade ?



1. duplication of A
  - ▶ hypersensitivity stress!
2. diverse mutation possible to relieve stress, e.g.  $ba1 \rightarrow 0$
3. only some of these mutations allow further ev. (devo.) pathways (*Fluchtlinien*)
4. diversification of A1 and A2 target genes, e.g.  $\Delta bA1$



# Duplication of DNA binding domain: oscillations ?



- ▶ more complex bifurcation diagrams:  
co-existence of bistability (saddle-node bif.) and oscillations (Hopf bif.)
- ▶ Why oscillations: embedding into global oscillator
  - ▶ cell cycle (Swi 5 binding sites in GAT1!)
  - ▶ respiratory oscillations

# Phylogenetic reconstruction I: proteins

## Pipeline:

1. 4 GAT1-like and 6(?) other Zn-f. from *S. cerevisiae*
2. psiblast Zn-finger in full genomes
  - ▶ collected 107 proteins
3. filter by dialign/clustal
  - ▶ 49 proteins with GAT1-like Zn-f. : same DNA binding site!
4. characterize proteins by clustal/dialign/probcons
  - ▶ extreme size variations, very little conservation, but...  
some small domains conserved:
  - ▶ Zn-finger variations correspond to other domains
  - ▶ NES or not!
  - ▶ activating domain, activator-characteristic C-term
  - ▶ inhibitor-characteristic Leucine zipper, no AD
  - ▶ Fe-binding element in 2GATA set
  - ▶ many N, Q, SPT, KR rich regions

# Phylogenetic reconstruction I: proteins

## Pipeline:

1. 4 GAT1-like and 6(?) other Zn-f. from *S. cerevisiae*
2. psiblast Zn-finger in full genomes
  - ▶ collected 107 proteins
3. filter by dialign/clustal
  - ▶ 49 proteins with GAT1-like Zn-f. : same DNA binding site!
4. characterize proteins by clustal/dialign/probcons
  - ▶ extreme size variations, very little conservation, but...  
some small domains conserved:
  - ▶ Zn-finger variations correspond to other domains
  - ▶ NES or not!
  - ▶ activating domain, activator-characteristic C-term
  - ▶ inhibitor-characteristic Leucine zipper, no AD
  - ▶ Fe-binding element in 2GATA set
  - ▶ many N, Q, SPT, KR rich regions

# Conserved Zn finger

## Hemiascomycota

*S. cerevisiae*

*Ashbya gossypii*

*Candida glabrata*

*Kluyveromyces lactii*

*Debaromyces hansenii*

*Yarrowia lipolytica*

*Candida albicans*

...?...

*S. pombe*

## Ascomycota

*Aspergillus nidulans*

*A. fumigatus*

*A. terreus*

*Neurospora crassa*

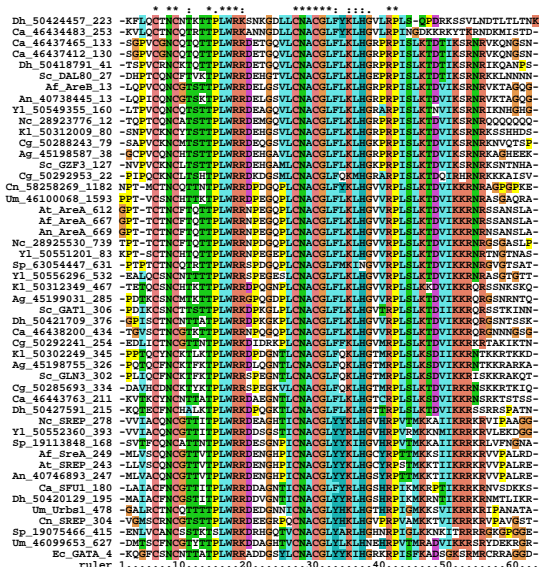
## Basidiomycota

*Cryptococcus neoformans*

*Ustilago maydis*

...?...

*Encephalitozoon cuniculi*



# Activation domain

```
* ***** *:*****. : :****: * ..***** *:*:*:*:  
Kl_50302249_133 LNLONGEIAQLWDFNVVHEFMTPSDHS---DSATISAPSSFNSEQFTPGI  
Ag_45198755_114 LNLONGEIDQLWHFNVDEFMTPSEQSGMSDGTATISQPNSFTSDAAL---  
Sc_GLN3_124 SNQONGEIAQLWDFNVDFNMTPSNSS---GSATISAPNSFTSDIPQYNH  
Cg_50285693_143 SNQONGEIAQLWDFNVDTLNIITPSNNS---GSATISAPNSYNSIPLGVH  
ruler 1.....10.....20.....30.....40.....50
```



## GLN3 specific, $\alpha$ -helix

Svetlov and Cooper 1997

```
: : * . *: **:**:  
Cn_58258269_587 VWKAVARARDTLPNGQRMENLTWRMMHLTLKKKEEQAAKEKEEREKEDKE---  
Um_46100068_872 VWKMYAKQRSQLPNGARMENLTWRMMAMTLRKKKEQAVDAKLALAEQSN---  
Af_Area_101 IWKLYSRTKAQLPNQERMENLTWRMMAMSLRKRERERLQQNNRPPSTKNSGS---  
An_Area_105 IWKLYSRTKAQLPNQERMENLTWRMMALSLRRQERERACQARASSQKSPV---  
At_Area_58 IWKLYSRTKAQLPNQERMENLTWRMMAMSLKRQEQORLQQ-SKSSPNPSCGIA--  
Nc_28925530_169 LWKYFAKTKLALPNQERMENLTWRMMAKPLQTYRROMETDRTHRFSESAPO---  
Kl_50312349_22 LWRMYSKAKASLPYRERMSLTWRMMGMKHLKHASSGGSDKNTGTEASGAN---  
Sp_63054447_41 LWKMFSAKASELSNGRRVENLTWRLLMSINLQKNLTPNGDSNTLTPDTFSDP---  
Ag_45199031_27 VWSMYSRARTTLAHRRTVNLAWRMMGVAARRRDTARAERGVRSAAQSGS---  
Ca_46438200_1 ---MYYRARHSLPYQKR MENLTWRMMYINNKSIFTNTNNAKKEIFEQSLDFEID  
Sc_GAT1_58 IWDLYSSAQKILPDSNRILNLSWRLLHNRISFHRINRIMQHSNSIMDFSSAF---  
ruler 1.....10.....20.....30.....40.....50.....
```



## GAT1 specific, putative activation domain

⇒ check charge patterns, 2<sup>ndary</sup> structure

# Activator-specific C-terminus

```
          :*
Dh_50421709_549 EKSDNDWLGMA-- 11
Ca_46438200_677 NNQNDWLNIN-- 12
  Sc_GAT1_499  GNLSLDWLNIN-- 12
Yl_50556296_684 EKDEVDWIKLNPS 13
Ca_46443763_671 DYKDLWIKKDI-- 12
Dh_50427591_691 VAKDLWIKFDL-- 12
Kl_50302249_770 VMDDLWIKFQGN-- 12
Ag_45198755_796 VMEDLDWIKFSI-- 12
  Sc_GLN3_719  LADELWIKFQGI-- 12
Cg_50285693_812 LEQDLWIKFQGI-- 12
  At_AreA_806  ASQENWLTMSI-- 12
  An_AreA_865  ASQENWLTMSI-- 12
  Af_AreA_870  ASQENWLTMSI-- 12
Nc_28925530_1025 TGAENWLTMSI-- 12
Kl_50312349_681 DVTSWLTLSI-- 12
Ag_45199031_457 NTATWLTLSI-- 12
Yl_50551201_227 RDQENWLSMAL-- 12
Dh_50424457_455 QGPNNWLDVSN 13
Ca_46434483_459 TNNWNNLDSPA 13
sp_63054447_845 VSKSMDWYSVM-- 11
      ruler 1.....10...
```



probably an  $\alpha$ -helix

⇒ check 2<sup>ndary</sup> structure

# Inhibitor-specific

```
*:***:**: : *:: : : .  
Kl_50312009_462 LKTRISELELVTDLYKRHFVFLDSRCKALEQEL  
Sc_DAL80_225 LKTKISELELVTDLYKKHIFQLNEKCKQLEVEL  
Cg_50288243_426 LRTRINELELVTDLYKRHIYELDERCKLEKEL  
Sc_GZF3_507 LKTRINELELVTDLYRRHINELDKCRALERL  
Ag_45198587_272 LKTRINELELINDLYKRHFELNDRCKMLEMKH  
Cg_50292953_148 LKTRVNELESITRLYKNHITRLEQRCQILESKL  
Ca_46437465_547 LKTRISELELVNDLYRTRIMELEAMEQAARLRE  
Ca_46437412_541 LKTRISELELVNDLYRTRIMELEAMEQAARLRE  
Dh_50418791_471 LKTRISELELVNDLYRTRIMELEAMEQAARLRE  
Af_AreB_208 LKTRVSELELINGLFRGRVAELEQSDATARRSE  
An_40738445_203 LKTRVSELDLINGLFRGRVAELEQSDATARRSE  
Yl_50549355_483 LKTRLSELELVNDLFRSRVAEVEAAEQARRSE  
Nc_28923776_183 LKTRVSELEVIQELYRGRHLHOLETEENIRQASE  
ruler 1.....10.....20.....30...
```



## leucine zipper

... Cooper 1997

```
*:* *.* * **:* ** . * ** : ** :  
Af_SreA_178 GSCPGGGNCNGTGGAEAGDGC PAYNNRVYKS  
An_40746893_178 GSCPGGGNCNGTGGAEAGDGC PAYNNRVYKS  
At_SREP_171 GSCPGGGSCNGTGGAEAGDGC PAYNNRVYKS  
Nc_SREP_202 GSCPGGGRCNGTGGAEAGCGCC PAYNNRVYKS  
Yl_50552360_332 GSCPGDGHCNGTGGSAACSGCLAYNNRINKA  
Um_Urbs1_409 GSCPGDGLCNGTGGTASCSCGCPAYNNNLSHA  
Ca_SFU1_114 GSCKGDGRCNGTGGSAACKGCPAYNNRIVAK  
Dh_50420129_128 GSCKGDGRCNGTGGSMACKGCPAYNNRVVVK  
Sp_19113848_68 GTCAGDGFRCNGTGGSACTGCPALNNRIRSL  
Cn_SREP_163 GTCPGDGRNGAGGKAGCEGCPYNNSIASG  
ruler 1.....10.....20.....30.
```



## iron-responsive element between 2 Zn fingers

# Phylogenetic reconstruction I: domain structures

```
2GATA      -----ZnN-FeB-ZnC-----H
           -----ZnN--ZnC-----H
C.n.       -----FeB-ZnCe-----H

G          ---A1-----ZnCe-----H
A          ---A2-----ZnCe-----H
I          -----ZnC-----L

Y.l.       -----ZnCe-----H
U.m./C.n. --A1-/1000/-----ZnCe--/1000/----
```

- ▶ *S. cerevisiae*, *C. glabrata*: 2 A, 2 I, 0 2GATA
  - ▶ pattern corresponds to known yeast genome evolution
- ▶ *Aspergillus sp.* et al.: 1 A, 1 I, 1 2GATA
  - ▶ *A. terreus*: lost I, and see below
- ▶ new inhibitor in *Yarrowia lipolytica* ?
- ▶ new iron-responsive activator in *S. pombe* ?

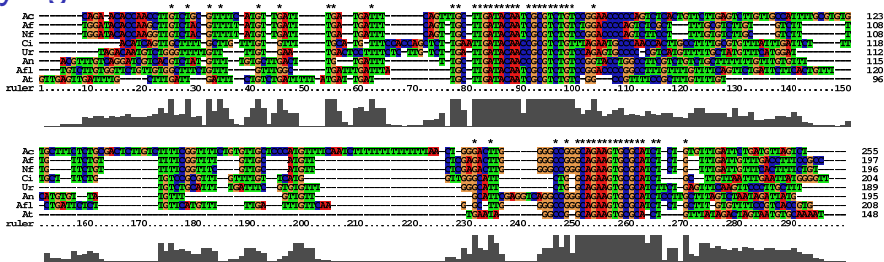


# Phylogenetic reconstruction II: hGATAr sites

```
>Sc_DAL80_upstream_1000
-----c-----c--c-----T-----
---T-----A--TT---g-T-----T-----
>Sc_GAT1_upstream_1000
-----A-----
-----A-----A-AAAA-----
>At_AreA_upstream_1000
-----cT--c--c---
>Yl_50549355_upstream_1000
---T--T-Tc---gTT-----T-c-----
-----T--c-----A-gg-----
```

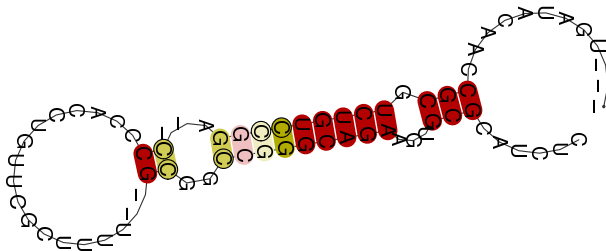
- ▶ hGATAr sites
  - ▶ multiple adjacent binding sites in some but not all
  - ▶ different distances, correspond to genome-size constraints
  - ▶ distances between adjacent sites ? relation to GATA protein size
  - ▶ **construct networks for 15 species!**
- ▶ again *A. terreus*: weak auto-activation

# Phylogenetic reconstruction II: 3'UTR



1. dialign 1000 nt. downstream of 49 coding sequences
2. AreA 3'UTR found in 3 species: known feedback function!
  - ▶ half-life w/o 3'UTR: ca 23 min
  - ▶ half-life with 3'UTR, -GLN: ca. 40 min
  - ▶ half-life with 3'UTR, +GLN: ca. 7 min
3. blast fungal genomes with *A. nidulans* sequence
  - ▶ retrieved sequences from 5 additional species
4. again *A. terreus*: very small T-rich loop

## Phylogenetic reconstruction II: 3'UTR



1. dialign 1000 nt. downstream of 49 coding sequences
2. AreA 3'UTR found in 3 species: known feedback function!
  - ▶ half-life w/o 3'UTR: ca 23 min
  - ▶ half-life with 3'UTR, -GLN: ca. 40 min
  - ▶ half-life with 3'UTR, +GLN: ca. 7 min
3. blast fungal genomes with *A. nidulans* sequence
  - ▶ retrieved sequences from 5 additional species
4. again *A. terreus*: very small T-rich loop

# Outlook

1. finish above, add some of the detected new models and potential new bifurcation parameters
2. maybe collect closer yeast relatives to better define protein domains and find more UTR RNA structures
3. do experiments (anyone, please)

Thanks:

Dougie Murray  
Lukas Endler  
Stefan Müller  
Stefanie Widder  
Xtof Flamm  
Peter Schuster  
Heinz Engl

gesundheit and a well recovery of all  
your livers and lungs

# The group I intron conspiracy

- ▶ Mobile group I introns (mgli) frequently appear in redox-related genes (phages, cyanobacteria, unicellular photosynthetic organisms, chloroplasts, mitochondria)
- ▶ One mgli is splice-inhibited by  $\text{NADP}^+$ , two mgli regulated by redox machinery of respiratory and photosynthetic  $e^-$  transport chains, respectively
- ▶ 7 cyanobacteria (origin of both mgli and  $\text{O}_2$  using  $e^-$  transport chains) contain mgli in fMet-tRNA (translation initiation)
- ▶ the only fMet-tRNA of an organism is really a bad bad place for a *selfish* intron

# The group I intron conspiracy

- ▶ Mobile group I introns (mgli) frequently appear in redox-related genes (phages, cyanobacteria, unicellular photosynthetic organisms, chloroplasts, mitochondria)
- ▶ One mgli is splice-inhibited by  $\text{NADP}^+$ , two mgli regulated by redox machinery of respiratory and photosynthetic  $e^-$  transport chains, respectively
- ▶ 7 cyanobacteria (origin of both mgli and  $\text{O}_2$  using  $e^-$  transport chains) contain mgli in fMet-tRNA (translation initiation)
- ▶ the only fMet-tRNA of an organism is really a bad bad place for a *selfish* intron

# The group I intron conspiracy

- ▶ Mobile group I introns (mgli) frequently appear in redox-related genes (phages, cyanobacteria, unicellular photosynthetic organisms, chloroplasts, mitochondria)
- ▶ One mgli is splice-inhibited by  $\text{NADP}^+$ , two mgli regulated by redox machinery of respiratory and photosynthetic  $e^-$  transport chains, respectively
- ▶ 7 cyanobacteria (origin of both mgli and  $\text{O}_2$  using  $e^-$  transport chains) contain mgli in fMet-tRNA (translation initiation)
- ▶ the only fMet-tRNA of an organism is really a bad bad place for a *selfish* intron

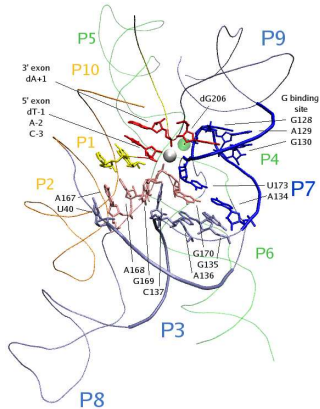
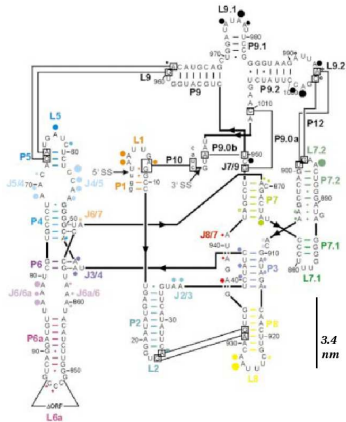


# The group I intron conspiracy

- ▶ Mobile group I introns (mgli) frequently appear in redox-related genes (phages, cyanobacteria, unicellular photosynthetic organisms, chloroplasts, mitochondria)
- ▶ One mgli is splice-inhibited by  $\text{NADP}^+$ , two mgli regulated by redox machinery of respiratory and photosynthetic  $e^-$  transport chains, respectively
- ▶ 7 cyanobacteria (origin of both mgli and  $\text{O}_2$  using  $e^-$  transport chains) contain mgli in fMet-tRNA (translation initiation)
- ▶ the only fMet-tRNA of an organism is really a bad bad place for a *selfish* intron

**Could mobile introns with splicing sensitivity to redox state spread a simple adaptation mechanism (emergency switch for cellular redox reactors), e.g. during transition from the anaerobic to the aerobic ( $\text{O}_2$ ) world?**

# Group I self-splicing introns

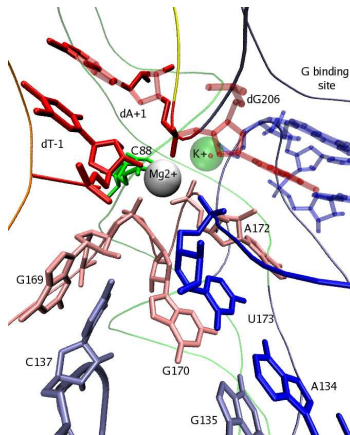
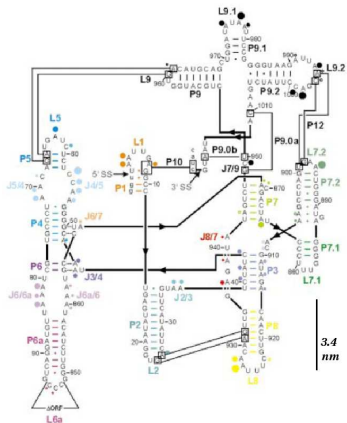


Nobel Prize 1989, Thomas R. Cech (with Sidney Altman)

- ▶ group I and II introns are self-splicing *ribozymes*

Cech 1990 Annu Rev Biochem  
 Waldsich et al. 2002 EMBO J  
 Adams et al. 2004 RNA/Nature

# Group I self-splicing introns



Nobel Prize 1989, Thomas R. Cech (with Sidney Altman)

- ▶ group I and II introns are self-splicing *ribozymes*

Cech 1990 Annu Rev Biochem  
Waldsich et al. 2002 EMBO J  
Adams et al. 2004 RNA/Nature

# A function of mobile introns?

- ▶ Positive *td* splice regulators
  - ▶ **guanosines**
  - ▶ **specific splicing factor Cyt-18 (Tyrosyl-tRNA synthetase)**
  - ▶ RNA chaperones, e.g. StpA
  - ▶ HEG derived maturases
- ▶ Negative *td* splice regulators
  - ▶ competitive, guanosine analog
    - ▶ **deoxyguanosine** and dideoxyguanosine
    - ▶ **amino acid arginine**
    - ▶ **coenzyme flavin FMN**
    - ▶ **coenzyme thiamin pyrophosphates**
    - ▶ antibiotics streptomycin, viomycin, capreomycin
  - ▶ competitive, no guanosine analog
    - ▶ antibiotic lysinomycin
  - ▶ non-competitive, mixed
    - ▶ aminoglycoside antibiotics kanamycin, tobramycin, tetracycline, pentamidine, spectinomycin
    - ▶ **NAD<sup>+</sup>**
    - ▶ **NADP<sup>+</sup>**
- ▶ Balance replication / transcription / translation with redox conditions?
  - ▶ Emergency switch for cellular redox reactors?
  - ▶ Involved in **redox clocks**?

# A function of mobile introns?

- ▶ Positive *td* splice regulators
  - ▶ **guanosines**
  - ▶ **specific splicing factor Cyt-18 (Tyrosyl-tRNA synthetase)**
  - ▶ RNA chaperones, e.g. StpA
  - ▶ HEG derived maturases
- ▶ Negative *td* splice regulators
  - ▶ competitive, guanosine analog
    - ▶ **deoxyguanosine** and dideoxyguanosine
    - ▶ **amino acid arginine**
    - ▶ **coenzyme flavin FMN**
    - ▶ **coenzyme thiamin pyrophosphates**
    - ▶ antibiotics streptomycin, viomycin, capreomycin
  - ▶ competitive, no guanosine analog
    - ▶ antibiotic lysinomycin
  - ▶ non-competitive, mixed
    - ▶ aminoglycoside antibiotics kanamycin, tobramycin, tetracycline, pentamidine, spectinomycin
    - ▶ **NAD<sup>+</sup>**
    - ▶ **NADP<sup>+</sup>**
- ▶ Balance replication / transcription / translation with redox conditions?
- ▶ Emergency switch for cellular redox reactors?
- ▶ Involved in **redox clocks**?

# A function of mobile introns?

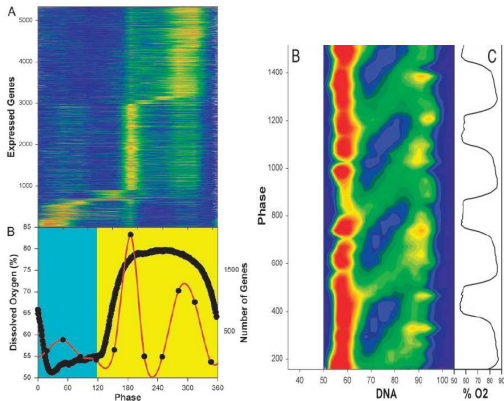
- ▶ Positive *td* splice regulators
  - ▶ **guanosines**
  - ▶ **specific splicing factor Cyt-18 (Tyrosyl-tRNA synthetase)**
  - ▶ RNA chaperones, e.g. StpA
  - ▶ HEG derived maturases
- ▶ Negative *td* splice regulators
  - ▶ competitive, guanosine analog
    - ▶ **deoxyguanosine** and dideoxyguanosine
    - ▶ **amino acid arginine**
    - ▶ **coenzyme flavin FMN**
    - ▶ **coenzyme thiamin pyrophosphates**
    - ▶ antibiotics streptomycin, viomycin, capreomycin
  - ▶ competitive, no guanosine analog
    - ▶ antibiotic lysinomycin
  - ▶ non-competitive, mixed
    - ▶ aminoglycoside antibiotics kanamycin, tobramycin, tetracycline, pentamidine, spectinomycin
    - ▶ **NAD<sup>+</sup>**
    - ▶ **NADP<sup>+</sup>**
- ▶ Balance replication / transcription / translation with redox conditions?
- ▶ Emergency switch for cellular redox reactors?
- ▶ Involved in **redox clocks**?

# mRNA, tRNA and rRNA introns

Consequences of splice inhibition:

- ▶ **T4 phage dNTP synth**: adapt to available reductive potential
  - ▶ **e<sup>-</sup> transport chain**: avoid oxidative stress
  - ▶ **fMet-tRNA** (7 cyanobacteria):  
**block translation**
  - ▶ **Leu-tRNAs** (conserved from cyanob. to chloroplasts):  
**attenuate or bias translation**
  - ▶ **rRNA** (div. bacteria, nucleus, chloroplasts, mitochondria):  
**block translation**
- ⇒ is there a global regulation of translation?
- ⇒ is it redox regulated?

# Respiratory oscillations in yeast



## The redox attractor:

- ▶ 3 genome-wide gene expression phases
- ▶ 2 reductive phases:
  - ▶ **rebuild and charge**
- ▶ 1 oxidative phase:
  - ▶ **S metabolism and ribosomes**
- ▶ gating of DNA synthesis and cell cycle

### *S. cerevisiae*:

involvement of COX1 and COB introns in respiratory phase?

Redox dependent global translation phases:

involvement of rRNA, tRNA intron splice modulation by redox?

Klavec et al. 2004 PNAS and (confirmed by) Tu et al. 2005 Science  
Murray et al. 1999 Microbiology, Lloyd et al. 2002 Microbiology  
Murray 2004 Current Genomics, Lloyd and Murray 2005 Trends Biochem Sci



# Summary & Conclusion

- ▶ *td* intron: *in vitro* non-competitive **splice inhibition by NADP<sup>+</sup>**, a by-product of reactions catalyzed by the *frd* and *td* gene products
- ▶ T4 dNTP synthesis as a **NADPH consuming process**: key enzymes of reductive steps **contain introns**
- ▶ T4-like **cyanophages carry a photosynthetic group I intron** in variable regions in the vicinity of *td* and *nrdB* genes
- ▶ Group I (also II, III) intron abundance in **e<sup>-</sup> transport chains of photosynthesis and respiration**, 2 examples of regulation
- ▶ **Cyanobacterial origin of ... all this!**
- ▶ Cyanobacteria and cyanophages form an ecological 'metaspecies' with **dynamic genomes** of individual species; a perfect environment for mobile introns

# Summary & Conclusion

- ▶ *td* intron: *in vitro* non-competitive **splice inhibition by NADP<sup>+</sup>**, a by-product of reactions catalyzed by the *frd* and *td* gene products
- ▶ T4 dNTP synthesis as a **NADPH consuming process**: key enzymes of reductive steps **contain introns**
- ▶ T4-like **cyanophages carry a photosynthetic group I intron** in variable regions in the vicinity of *td* and *nrdB* genes
- ▶ Group I (also II, III) intron abundance in **e<sup>-</sup> transport chains of photosynthesis and respiration**, 2 examples of regulation
- ▶ **Cyanobacterial origin of ... all this!**
- ▶ Cyanobacteria and cyanophages form an ecological 'metaspecies' with **dynamic genomes** of individual species; a perfect environment for mobile introns

# Summary & Conclusion

- ▶ *td* intron: *in vitro* non-competitive splice inhibition by NADP<sup>+</sup>, a by-product of reactions catalyzed by the *frd* and *td* gene products
- ▶ T4 dNTP synthesis as a NADPH consuming process: key enzymes of reductive steps contain introns
- ▶ T4-like cyanophages carry a photosynthetic group I intron in variable regions in the vicinity of *td* and *nrdB* genes
- ▶ Group I (also II, III) intron abundance in e<sup>-</sup> transport chains of photosynthesis and respiration, 2 examples of regulation
- ▶ Cyanobacterial origin of ... all this!
- ▶ Cyanobacteria and cyanophages form an ecological 'metaspecies' with dynamic genomes of individual species; a perfect environment for mobile introns

# Summary & Conclusion

- ▶ *td* intron: *in vitro* non-competitive splice inhibition by NADP<sup>+</sup>, a by-product of reactions catalyzed by the *frd* and *td* gene products
- ▶ T4 dNTP synthesis as a NADPH consuming process: key enzymes of reductive steps contain introns
- ▶ T4-like cyanophages carry a photosynthetic group I intron in variable regions in the vicinity of *td* and *nrdB* genes
- ▶ Group I (also II, III) intron abundance in e<sup>-</sup> transport chains of photosynthesis and respiration, 2 examples of regulation
- ▶ Cyanobacterial origin of ... all this!
- ▶ Cyanobacteria and cyanophages form an ecological 'metaspecies' with dynamic genomes of individual species; a perfect environment for mobile introns

# Summary & Conclusion

- ▶ *td* intron: *in vitro* non-competitive splice inhibition by  $\text{NADP}^+$ , a by-product of reactions catalyzed by the *frd* and *td* gene products
- ▶ T4 dNTP synthesis as a NADPH consuming process: key enzymes of reductive steps contain introns
- ▶ T4-like cyanophages carry a photosynthetic group I intron in variable regions in the vicinity of *td* and *nrdB* genes
- ▶ Group I (also II, III) intron abundance in  $e^-$  transport chains of photosynthesis and respiration, 2 examples of regulation
- ▶ Cyanobacterial origin of ... all this:
- ▶ Cyanobacteria and cyanophages form an ecological 'metaspecies' with dynamic genomes of individual species; a perfect environment for mobile introns

# Summary & Conclusion

- ▶ *td* intron: *in vitro* non-competitive splice inhibition by  $\text{NADP}^+$ , a by-product of reactions catalyzed by the *frd* and *td* gene products
- ▶ T4 dNTP synthesis as a NADPH consuming process: key enzymes of reductive steps contain introns
- ▶ T4-like cyanophages carry a photosynthetic group I intron in variable regions in the vicinity of *td* and *nrdB* genes
- ▶ Group I (also II, III) intron abundance in  $e^-$  transport chains of photosynthesis and respiration, 2 examples of regulation
- ▶ Cyanobacterial origin of ... all this:
- ▶ Cyanobacteria and cyanophages form an ecological 'metaspecies' with dynamic genomes of individual species; a perfect environment for mobile introns

# Summary & Conclusion

- ▶ *td* intron: *in vitro* non-competitive splice inhibition by  $\text{NADP}^+$ , a by-product of reactions catalyzed by the *frd* and *td* gene products
- ▶ T4 dNTP synthesis as a NADPH consuming process: key enzymes of reductive steps contain introns
- ▶ T4-like cyanophages carry a photosynthetic group I intron in variable regions in the vicinity of *td* and *nrdB* genes
- ▶ Group I (also II, III) intron abundance in  $e^-$  transport chains of photosynthesis and respiration, 2 examples of regulation
- ▶ Cyanobacterial origin of ... all this:
- ▶ Cyanobacteria and cyanophages form an ecological 'metaspecies' with dynamic genomes of individual species; a perfect environment for mobile introns

**Could mobile introns with splicing sensitivity to redox state spread a simple adaptation mechanism, e.g. during transition from the anaerobic to the aerobic ( $\text{O}_2$ ) world?**

Thanks:

Chris K. Mathews

Renee Schroeder, Herbert Wank

Christoph Flamm, Lukas Endler, Stefan Müller

Ladislav Nedbal, Georg Schmetterer

Douglas Murray, Noriko Hiroi, Akira Funahashi