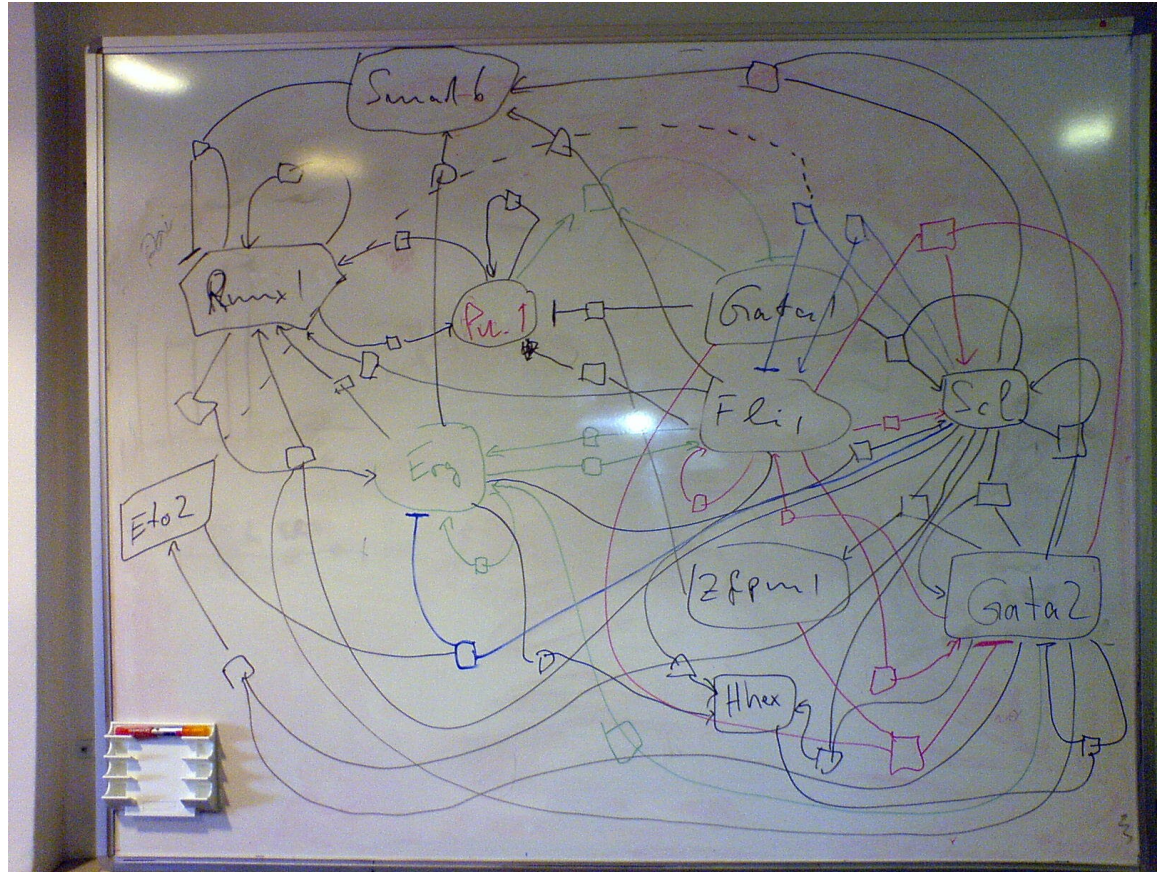


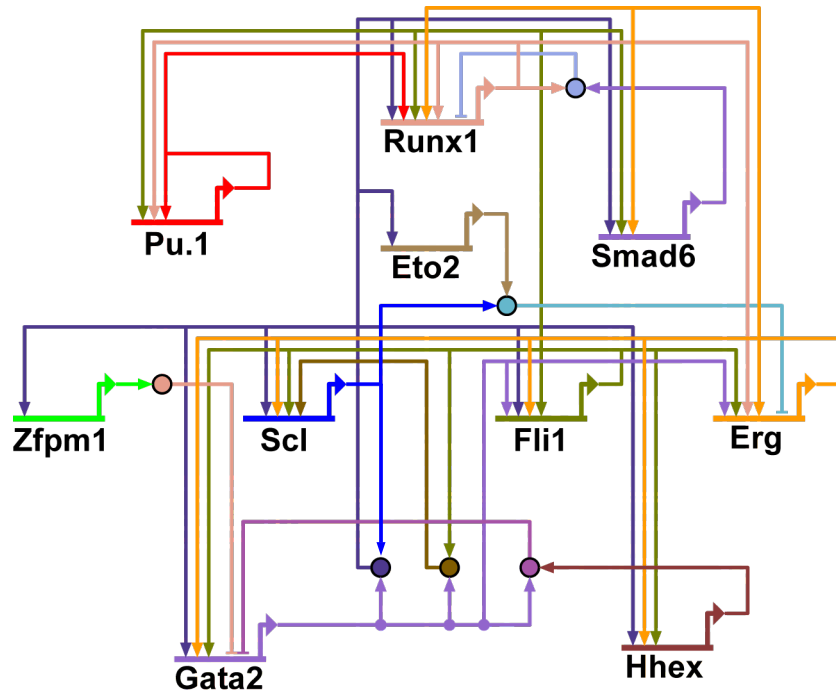
Gene regulation of cellular differentiation

Cis-Regulatory
Data for 10
HSC Gene Loci



Slides courtesy
of Bertie Göttgens
(CIMR Cambridge)

Regulators	Interaction	Target
Fli1	ACTIVATION	Pu.1
Fli1 + Gata2	ACTIVATION	Scf
Fli1	ACTIVATION	Gata2
Scf + Gata2	ACTIVATION	Gata2
Gata2 + Scf	ACTIVATION	Zfpn1
Gata2 + Scf	ACTIVATION	Scf
Fli1	ACTIVATION	Scf
Fli1	ACTIVATION	Fli1
Gata2	ACTIVATION	Fli1
Fli1	ACTIVATION	Erg
Erg	ACTIVATION	Fli1
Erg	ACTIVATION	Erg
Scf + Eto2	INHIBITION	Erg
Scf + Gata2	ACTIVATION	Fli1
Gata2	ACTIVATION	Erg
Runx1	ACTIVATION	Runx1
Pu.1	ACTIVATION	Runx1
Runx1	ACTIVATION	Pu.1
Gata2 + Scf	ACTIVATION	Runx1
Fli1	ACTIVATION	Runx1
Erg	ACTIVATION	Runx1
Pu.1	ACTIVATION	Runx1
Scf + Gata2	ACTIVATION	Eto2
Runx1 + Smad6	INHIBITION	Runx1
Erg	ACTIVATION	Smad6
Fli1	ACTIVATION	Smad6
Scf + Gata2	ACTIVATION	Smad6
Erg	ACTIVATION	Scf
Scf + Gata2	ACTIVATION	Hhex
Gata2 + Hhex	INHIBITION	Gata2
Fli1	ACTIVATION	Hhex
Erg	ACTIVATION	Hhex
Runx1	ACTIVATION	Erg
Erg	ACTIVATION	Gata2

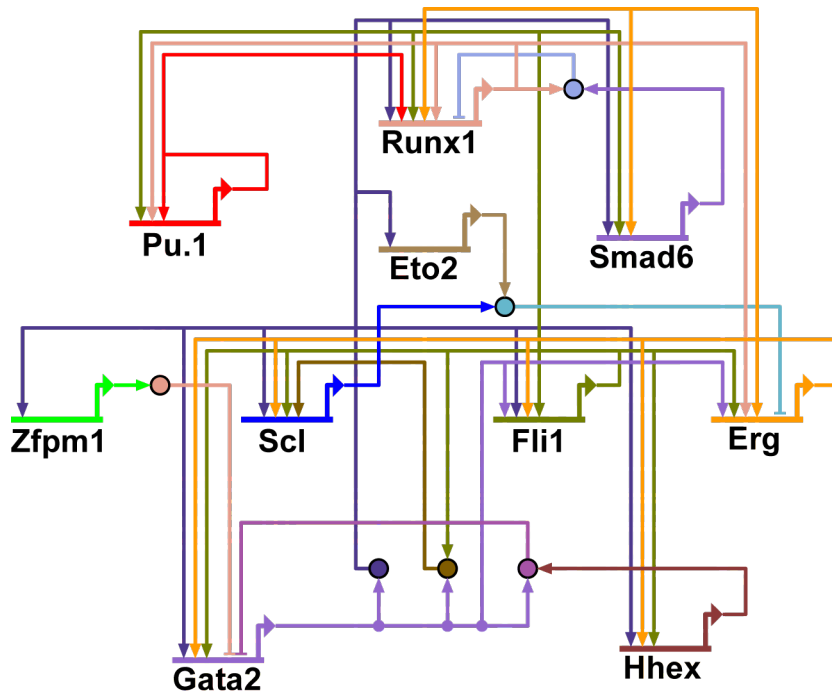


George E. P. Box:

... all models are wrong,

but some are useful

Network Execution – Network States

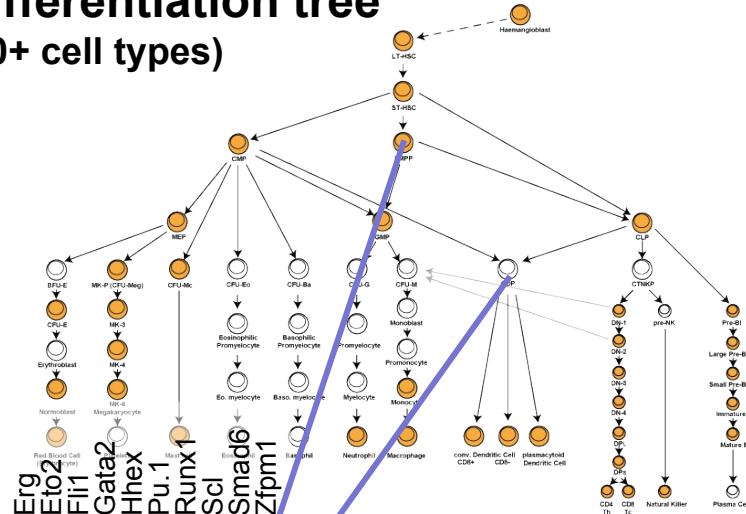


- The state of the whole network is the ensemble of the expression level of its genes
- Simplistically, each gene can be viewed as either expressed (1) or repressed (0)
- The state of the network is then represented by a bit vector
 $\text{State}_x = \{1,0,1,1,1,0,0,1,\dots\}$

10 genes $\rightarrow 2^{10} = 1,024$ expression states

Network Execution – State Transitions

Differentiation tree (50+ cell types)

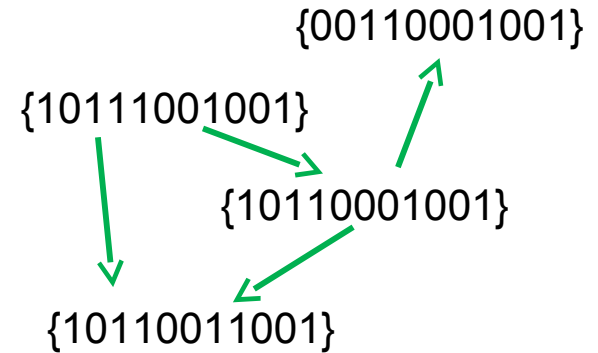


$\{10111001001\}$
 $\{10010011010\}$
 $\{00110000111\}$
 $\{10010001101\}$
 $\{00111001001\}$
 ...

**Observed
expression
states**

State space:

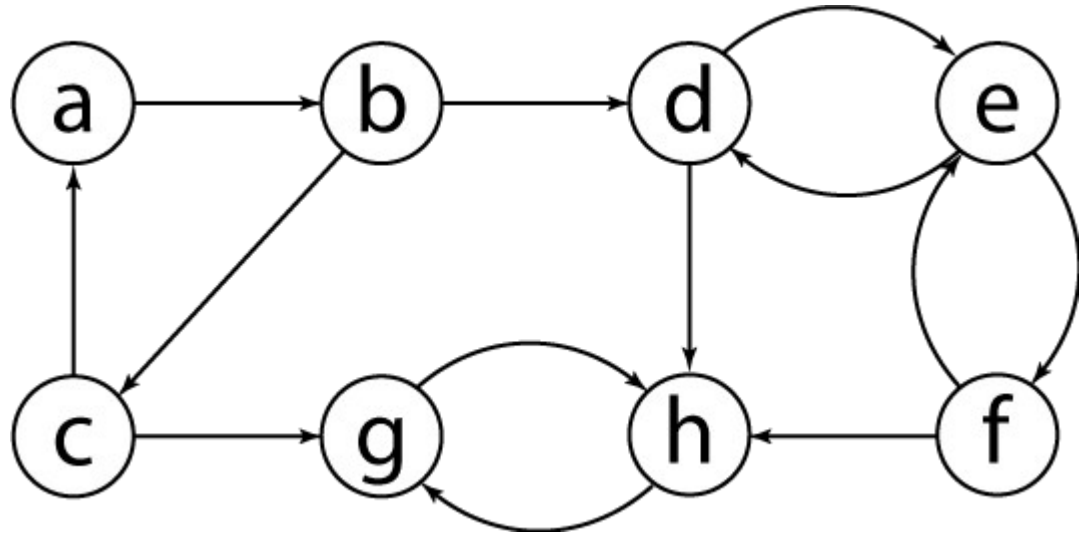
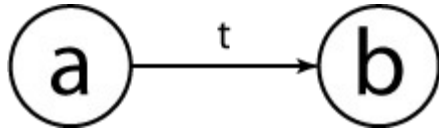
All 1,024 states and their connections permitted by the model



**Movement through state-space
dictated by network topology!**

State Transition Diagram

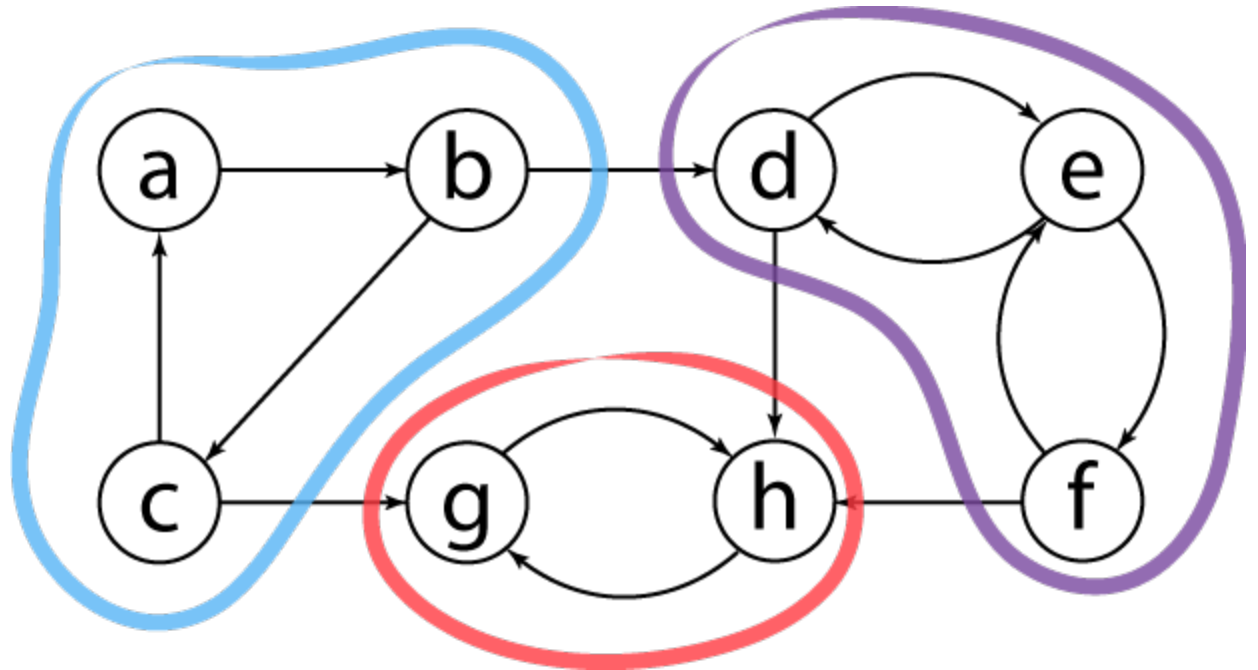
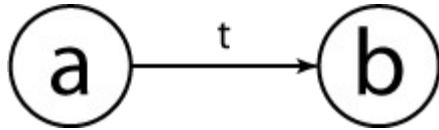
$a = \{1, \mathbf{0}, 1, 1, 1, 0, 0, 1, \dots\}$
 $b = \{1, \mathbf{1}, 1, 1, 1, 0, 0, 1, \dots\}$



**Full state space generated by a specific network topology
executing the model from all possible initial configurations**

Strongly Connected Components

$a = \{1, \mathbf{0}, 1, 1, 1, 0, 0, 1, \dots\}$
 $b = \{1, \mathbf{1}, 1, 1, 1, 0, 0, 1, \dots\}$



In a strongly connected component there is a path from each vertex (state) to every other vertex of the same SCC.

Strongly Connected Components


$a = \{1, 0, 1, 1, 1, 0, 0, 1, \dots\}$

Terminal strongly connected components in the state space graph represent the steady states of the biological system (attractors)

A strongly connected component is terminal if there are no outgoing arcs to any other SCC. Once g or h are reached, the computation is “trapped”.

Expression Patterns and Stable States

Scf	Gata2	Fog1	Fli1	Erg	Pu.1	Eto2	Hhex	Runx1	Smad6	
Red	Blue	Red	Blue	Red	Blue	Red	Red	Red	Blue	Erythroid
Blue	Blue	Red	Red	Blue	Blue	Blue	Blue	Blue	Blue	CD8 activated
Blue	Blue	Red	Red	Blue	Blue	Blue	Blue	Red	Blue	CD4 activated
Blue	Blue	Blue	Red	Blue	Red	Red	Red	Red	Blue	Monocyte
Blue	Blue	Red	Red	Blue	Red	Red	Red	Red	Blue	B Cell
Blue	Blue	Red	Red	Blue	Blue	Blue	Red	Red	Blue	CD8
Blue	Blue	Red	Red	Blue	Blue	Blue	Red	Red	Blue	CD4
Blue	Blue	Red	Red	Blue	Blue	Blue	Red	Red	Blue	NK
Red	Blue	Red	Red	Blue	Blue	Blue	Red	Red	Blue	Granulocyte
Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	HSC



Dynamic Expression

All Expressed in HSCs

Expression Patterns and Stable States

ScI	Gata2	Fog1	Fli1	Erg	Pu.1	Eto2	Hhex	Runx1	Smad6	
Red	Blue	Red	Blue	Red	Blue	Red	Red	Red	Blue	Erythroid
Blue	Blue	Red	Red	Blue	Blue	Blue	Blue	Blue	Blue	CD8 activated
Blue	Blue	Red	Red	Blue	Blue	Blue	Blue	Red	Blue	CD4 activated
Blue	Blue	Blue	Red	Blue	Red	Red	Red	Red	Blue	Monocyte
Blue	Blue	Red	Red	Blue	Red	Red	Red	Red	Blue	B Cell
Blue	Blue	Red	Red	Blue	Blue	Blue	Red	Red	Blue	CD8
Blue	Blue	Red	Red	Blue	Blue	Blue	Red	Red	Blue	CD4
Blue	Blue	Red	Red	Blue	Blue	Blue	Red	Red	Blue	NK
Red	Blue	Red	Red	Red	Red	Blue	Red	Red	Blue	Granulocyte
Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	HSC
Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	S1

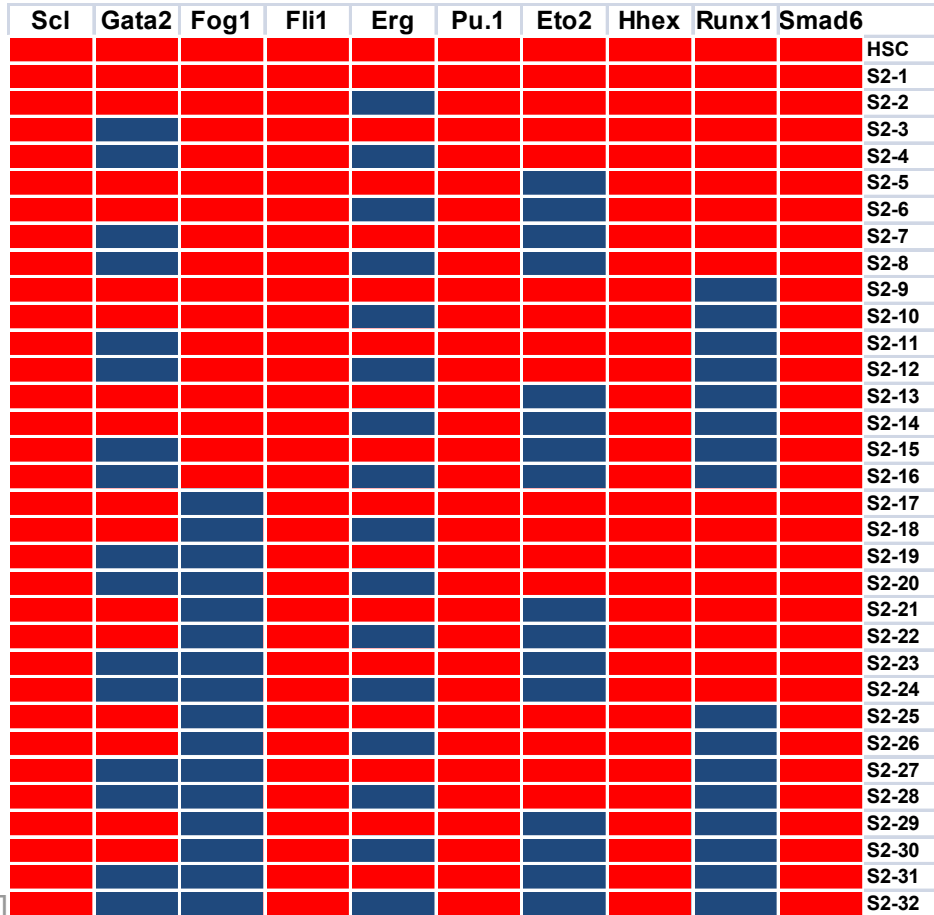
S1 ~ non-Blood Cell

Expression Patterns and Stable States

Scf	Gata2	Fog1	Fli1	Erg	Pu.1	Eto2	Hhex	Runx1	Smad6	
Red	Blue	Red	Blue	Red	Blue	Red	Red	Red	Blue	Erythroid
Blue	Blue	Red	Red	Blue	Blue	Blue	Blue	Blue	Blue	CD8 activated
Blue	Blue	Red	Red	Blue	Blue	Blue	Blue	Red	Blue	CD4 activated
Blue	Blue	Blue	Red	Blue	Red	Red	Red	Red	Blue	Monocyte
Blue	Blue	Red	Red	Blue	Red	Red	Red	Red	Blue	B Cell
Blue	Blue	Red	Red	Blue	Blue	Blue	Red	Red	Blue	CD8
Blue	Blue	Red	Red	Blue	Blue	Blue	Red	Red	Blue	CD4
Blue	Blue	Red	Red	Blue	Blue	Blue	Red	Red	Blue	NK
Red	Blue	Red	Red	Blue	Blue	Blue	Red	Red	Blue	Granulocyte
Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	HSC
Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	S1
Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	S2-1

S2-1 = HSC, but

Expression Patterns and Stable States



**S2-1 = HSC, but
it is 1 of 32 substates of a
“Terminal Steady State”**

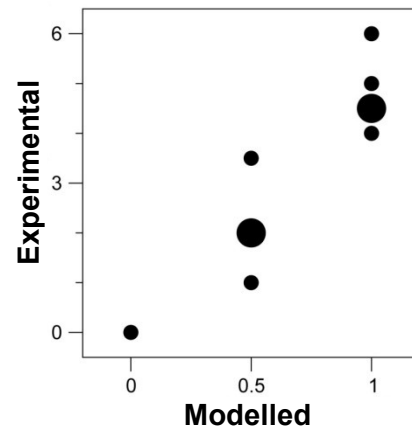
→ Stem Cell Heterogeneity?

Expression Patterns and Stable States

Goodell Lab Single Cell Expression Data:

	Scl	Gata2	Fli1	Zfp1	Gata1	Pu.1	Eto2	Hhex	Runx1	Smad6
1	P	P	P	P	A	A	P	A	A	A
2	P	P	P	A	A	A	A	A	A	A
3	P	A	A	A	A	P	A	P	A	A
4	A	A	A	A	A	P	A	P	A	A
5	P	A	A	P	A	A	A	P	A	P
6	A	A	A	A	A	P	A	A	A	A
7	A	A	A	A	A	A	A	A	A	P
8	A	A	A	A	A	P	A	A	P	A
9	P	P	A	A	A	A	A	P	A	P
10	A	A	P	A	A	A	A	A	A	A
11	P	P	P	A	A	A	P	P	A	P
12	A	A	P	A	A	A	A	A	A	P

Heterogeneous Expression

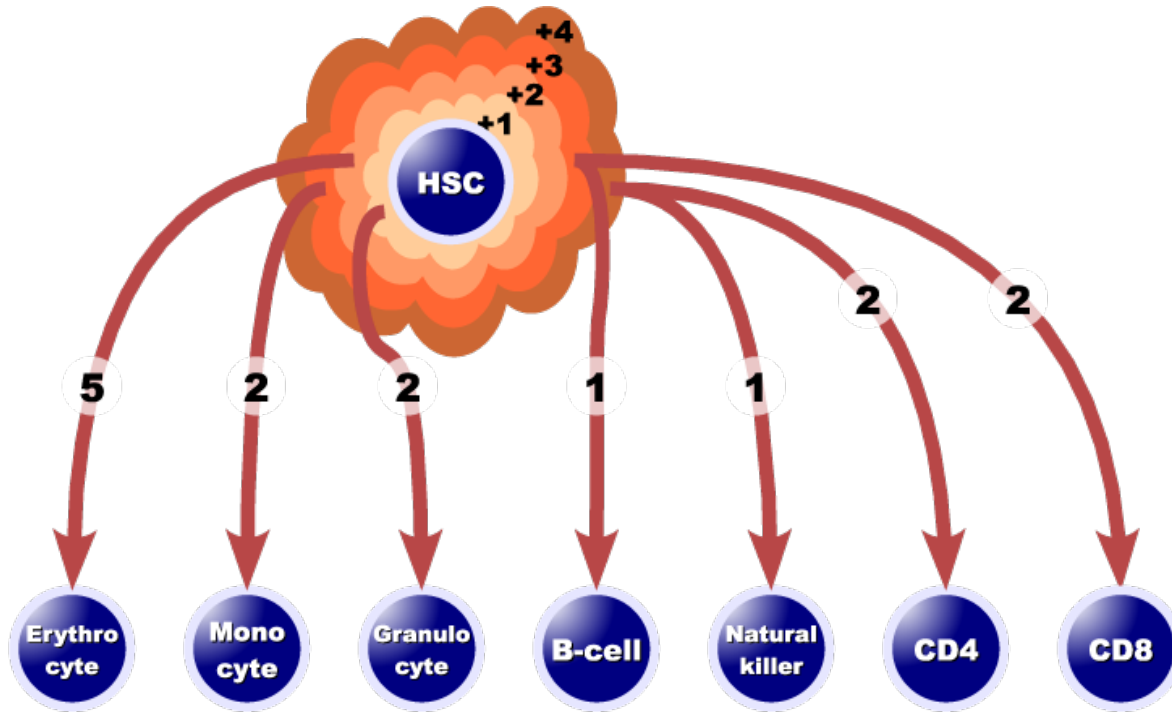


- There is some evidence for heterogeneous expression
- Correlation between experimental data and model predictions
- **BUT: Discretisation (0/1) clearly simplistic**
- **However: Model qualitatively the same with graded expression levels (0/0.5/1)**

Was this Modelling Exercise Useful?

- **Stem cell heterogeneity widely viewed as key mechanism for differentiation**
- **Generally assumed to be caused by non-specific noise**
- **BUT: Our model built on cis-regulatory data suggests “hard-wired” noise**
- **Discrete stem cell substates would have important implications:**
 - i) Why are lineage relationships of Stem/Progenitor cells so complex?**
 - ii) How do the same mutations cause different flavours of leukaemia?**

Analysis of State Transitions



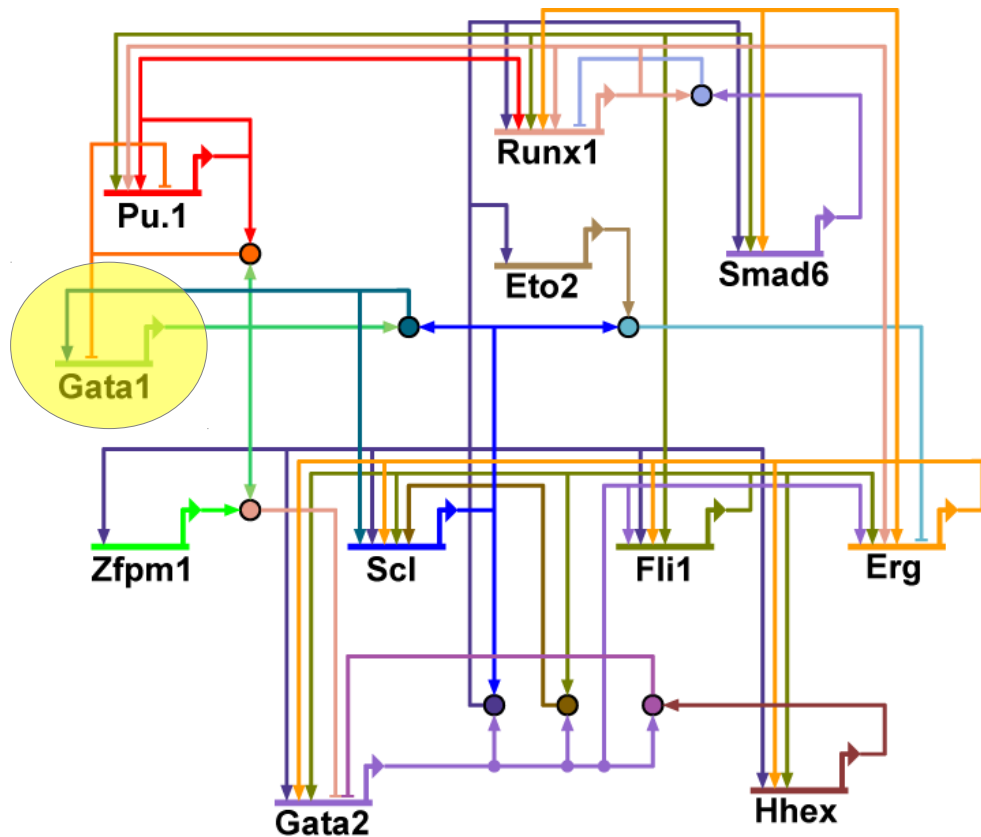
State Transitions:

- 11 transition paths follow the developmental tree but need an initial 'push' out of the HSC (up to 4 'pushes')
- 11 transition paths in the reverse direction reach the HSC without requiring a 'push'

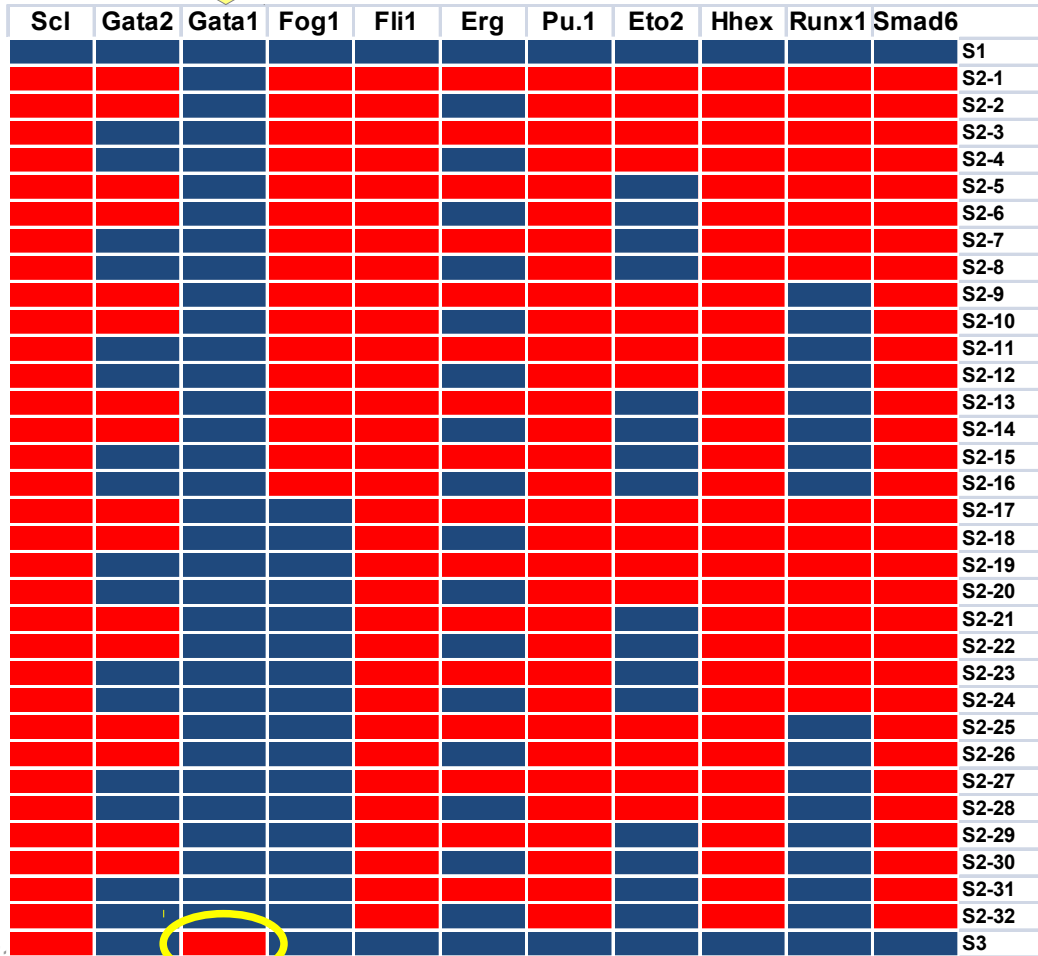
- 1) Stem Cell State has inherent stability
- 2) Exit requires a push
- 3) Reversal needs to be blocked by a lock

**Analysis of a potential 'push':
Gata1 /erythroid differentiation**

Regulators	Interaction	Target
Pu.1 + Gata1	INHIBITION	Gata1
Gata1 + Pu.1	INHIBITION	Pu.1
Gata1 + Zfp1	INHIBITION	Gata2
Scl + Gata1	ACTIVATION	Scl
Scl + Gata1	ACTIVATION	Gata1
Gata2 + Scl	ACTIVATION	Scl
Fli1	ACTIVATION	Scl
Fli1	ACTIVATION	Fli1
Gata2	ACTIVATION	Fli1
Fli1	ACTIVATION	Erg
Erg	ACTIVATION	Fli1
Erg	ACTIVATION	Erg
Scl + Eto2	INHIBITION	Erg
Scl + Gata2	ACTIVATION	Fli1
Gata2	ACTIVATION	Erg
Runx1	ACTIVATION	Runx1
Pu.1	ACTIVATION	Runx1
Runx1	ACTIVATION	Pu.1
Gata2 + Scl	ACTIVATION	Runx1
Fli1	ACTIVATION	Runx1
Erg	ACTIVATION	Runx1
Pu.1	ACTIVATION	Pu.1
Scl + Gata2	ACTIVATION	Eto2
Runx1 + Smad6	INHIBITION	Runx1
Erg	ACTIVATION	Smad6
Fli1	ACTIVATION	Smad6
Scl + Gata2	ACTIVATION	Smad6
Erg	ACTIVATION	Scl
Scl + Gata2	ACTIVATION	Hhex
Gata2 + Hhex	INHIBITION	Gata2
Fli1	ACTIVATION	Hhex
Erg	ACTIVATION	Hhex
Runx1	ACTIVATION	Erg
Erg	ACTIVATION	Gata2



Stable States



S1: still the same

S2: still the same 32 substates

→ Gata1 off in all “HSC substates”

S3: New stable state

→ ~ mature erythroid

Can we now simulate erythroid differentiation by “ectopically expressing” Gata1 in HSCs?

Gata1 might be 1-switch differentiation trigger:

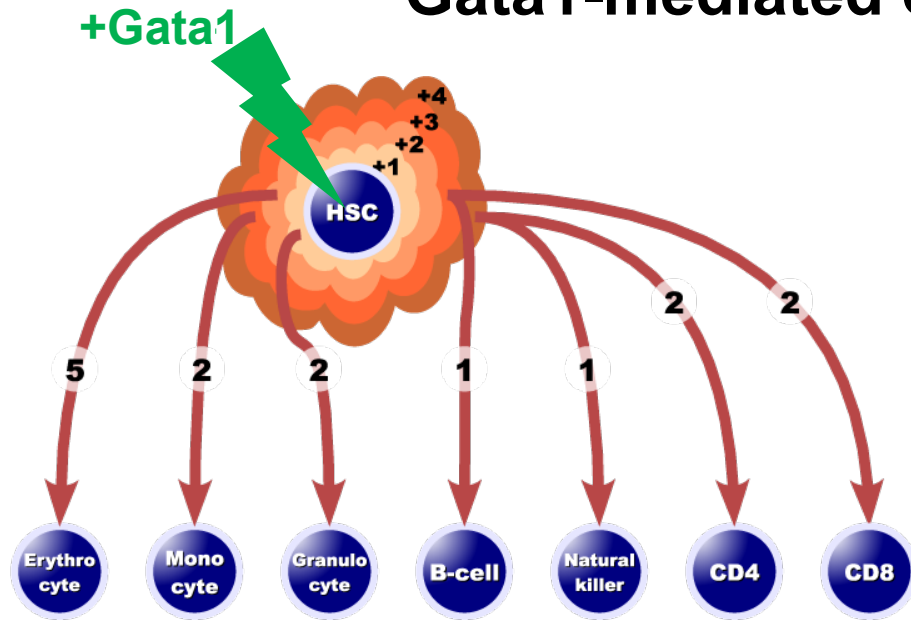
Visvader JE, Elefanty AG, Strasser A, Adams JM
EMBO J 1992

Kulesa H, Frampton J and Graf T
Genes Dev 1995

Heyworth C, Pearson S, May G, Enver T
EMBO J 2002



Gata1-mediated State Transitions



Simulated forced expression of Gata1 in “HSC substates” not enough to trigger progression to erythroid state S3

→ Missing action of Gata1 outside core network?

→ Missing action of Gata1 inside core network?

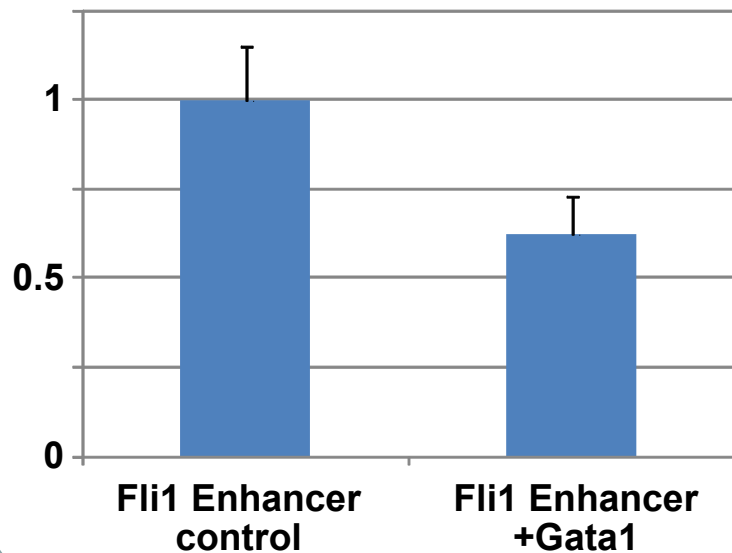
→ Investigated 2nd possibility

→ Could there be another Gata switch?

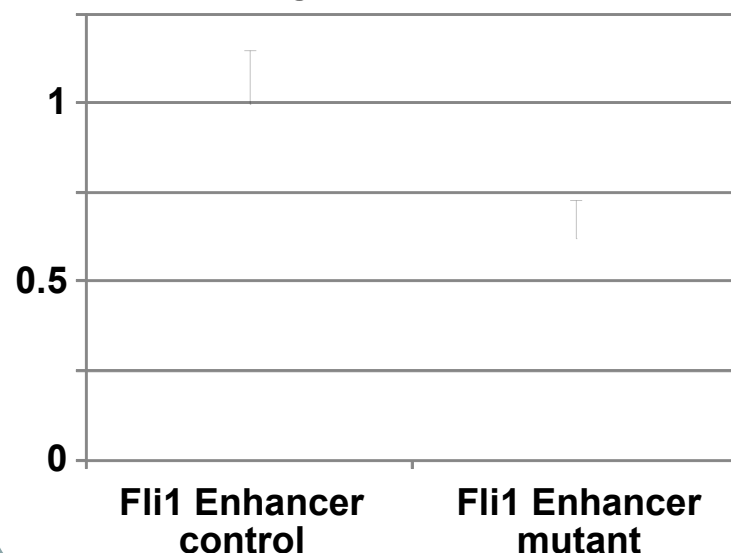
Gata switch first characterised on Gata2 -3.8 enhancer

Grass JA, Boyer ME, Pal S, Wu J, Weiss MJ, Bresnick EH *PNAS* 2003

Progenitor Cells



Erythroid Cells



Consistent with Gata1 mediated repression of the Fli1 +12 Enhancer

Models built on cis-regulatory experimental data:

Boolean models of 11 TF core network

- **Stem cell expression state stabilised**
- **Hard-wired heterogeneity in modelled stem cell state**
- **Implications for normal differentiation and leukaemia**
- **→ models differentiation ‘pushes’ and ‘locking-in’ of mature cells**