

# Variance component analysis of polymorphic metabolic networks

Joseph McClay and Edwin van den Oord

# Background:

- Conceptual gulf between molecular biology and quantitative genetics
- Integration of individual differences within systems biology framework

## Questions:

- Does biological network architecture affect the impact of genetic variation on phenotype?
- Is epistasis likely to be prevalent between genes that interact biologically?
- How likely is GxE interaction in general terms?

# Molecular Biology

- Traditionally reductionist
- Often qualitative descriptions
- Recent shift to “systems biology”
- -Omics approaches, etc
- Precise process models
- Example: Von Dassow *et al.*  
Nature (2000)
- Dynamic models
- General descriptions of single  
systems evolving in time

## Quantitative Genetics:

- Individual differences in populations
- Measurement taken at fixed point in time
- Variance components
  - additive
  - dominance
  - interaction (statistical epistasis)

## Conceptual Disparity:

Single Systems	V	Populations
Time		Fixed Point

### Why attempt to reconcile?

- Theory – insight into how metric characters arise (genotype to phenotype, genetic background)
- Practice – complex trait research. How important is epistasis? (Wright, 1980; Moore 2002)

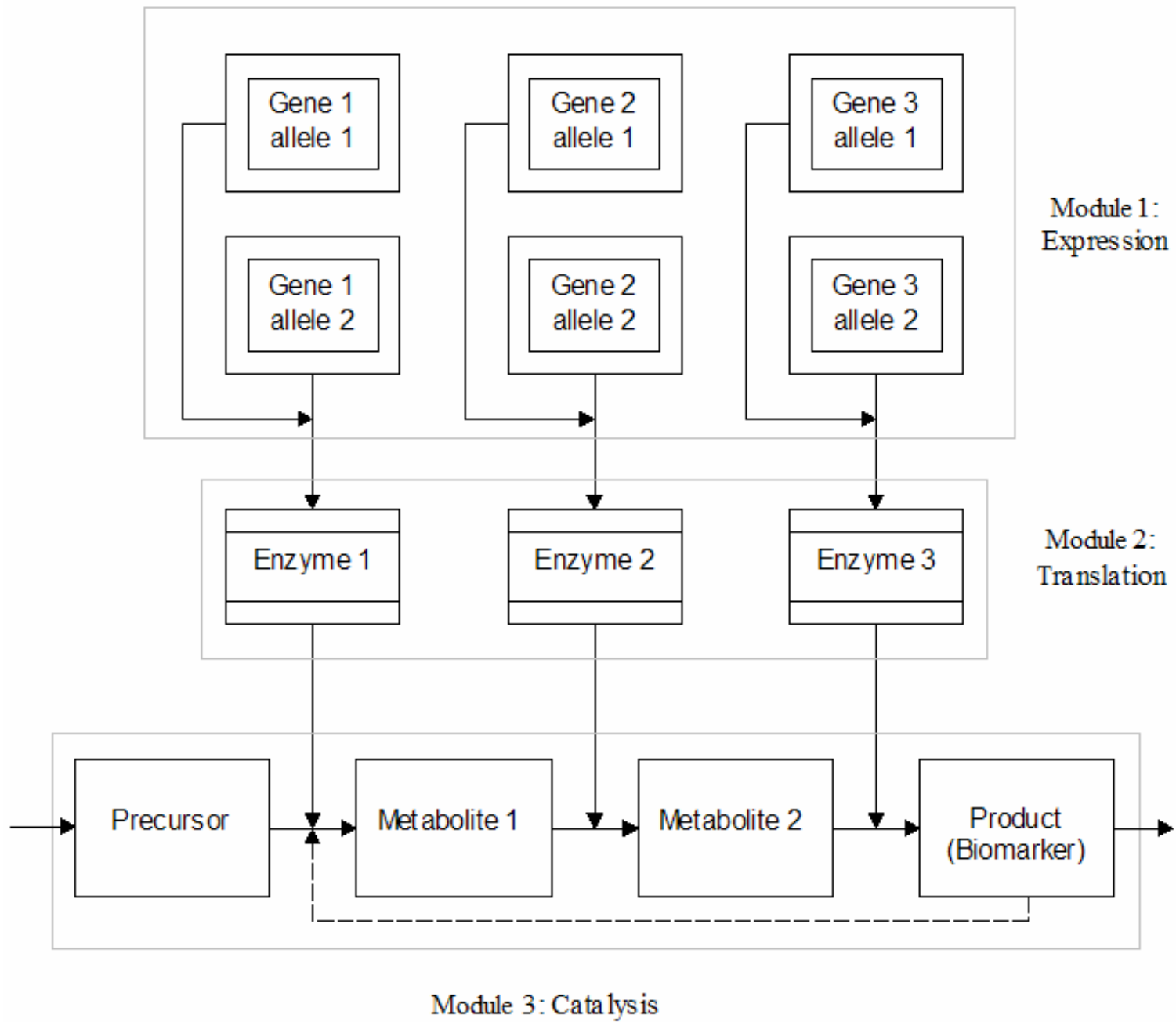
# How can these be reconciled?

- Generate biologically-realistic process models
- Introduce “genetic” variation by alteration of rate constants conferred by different alleles
- Fix frequencies of alleles in infinitely large population
- Solve steady state (equilibrium convergence point) of polymorphic systems, thus eliminating time
- Analyze components of variance in population attributable to each gene

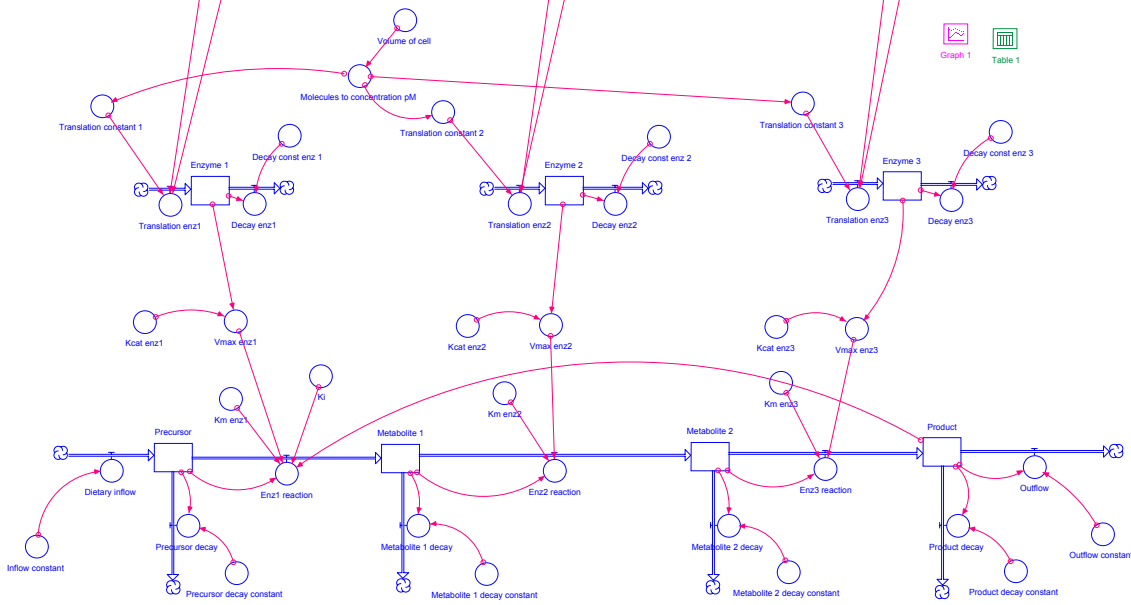
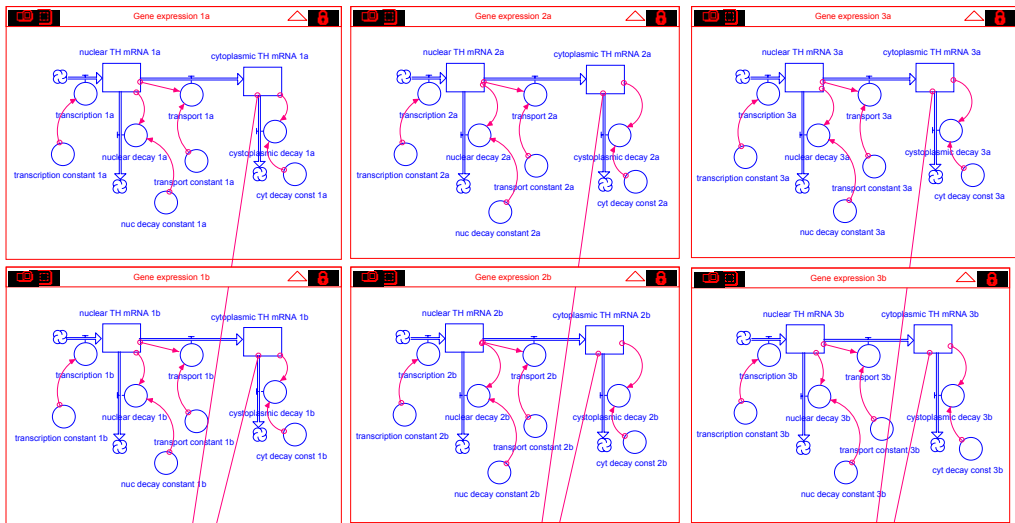
# Summary of our approach:

- 3 enzyme linear pathway (Standard model)
- 3 enzyme linear pathway under control of Competitive Inhibition (CI model) – negative feedback
- Use of generic kinetic constants from biochemical literature (Hargrove 1998, Fell 1997, Stryer 1988, Kakkar et al. 1999, etc)
- Use of accepted and proven biological theory (Michaelis-Menten enzyme kinetics) for underlying mathematics:
- Solve variant Steady States of pathway product ( $[P]^*$ ) – metabolic biomarker (e.g. cholesterol)

Figure 1



Competitive Inhibition (Feedback) model

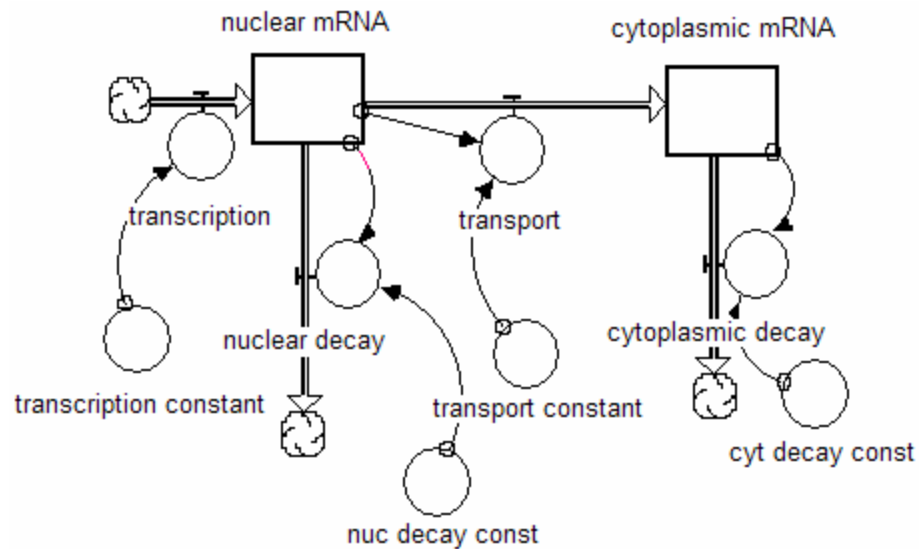


Graph 1  
Table 1

# Forrester Diagrams – “Gene Expression”

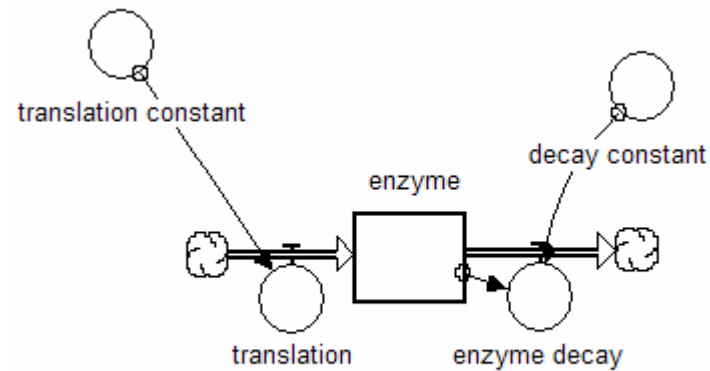
## Sector

### Single allele



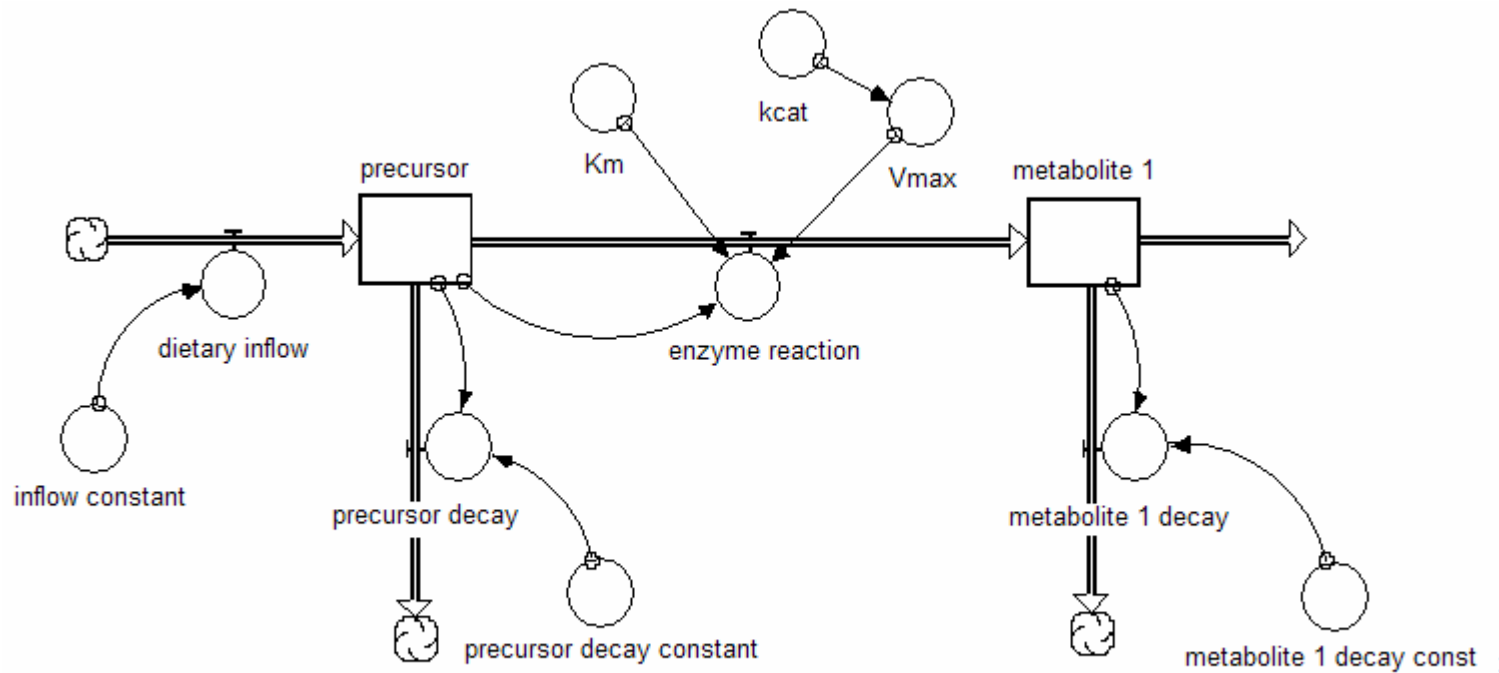
# Forrester Diagrams – “Translation” sector

## Single enzyme



# Forrester Diagrams – “Catalysis” sector

## Single conversion



Expression:

$$\frac{d(nR)}{dt} = \alpha_{nR} - (\lambda_{nR} + \alpha_{cR}) nR \quad (1)$$

$$\frac{d(cR)}{dt} = \alpha_{cR} nR - \lambda_{cR} cR \quad (2)$$

Translation:

$$\frac{d[E]}{dt} = \alpha_E ([cR]_{\text{allele1}} + [cR]_{\text{allele2}}) - \lambda_E [E] \quad (3)$$

Catalysis:

$$\frac{d[PC]}{dt} = \alpha_{Pc} - \frac{V_{max(E1)} [PC]}{K_m(E1) + [PC]} - \lambda_{Pc} [PC] \quad (4a)$$

$$\frac{d[PC]}{dt} = \alpha_{Pc} - \frac{V_{max(E1)} [PC]}{K_m(E1) \left( \frac{1 + [Pd]}{K_i} \right) + [PC]} - \lambda_{Pc} [PC] \quad (4b)$$

$$\frac{d[MI]}{dt} = \frac{V_{max(E1)} [PC]}{K_m(E1) + [PC]} - \frac{V_{max(E2)} [MI]}{K_m(E2) + [MI]} - \lambda_{MI} [MI] \quad (5a)$$

$$\frac{d[MI]}{dt} = \frac{V_{max(E1)} [PC]}{K_m(E1) \left( \frac{1 + [Pd]}{K_i} \right) + [PC]} - \frac{V_{max(E2)} [MI]}{K_m(E2) + [MI]} - \lambda_{MI} [MI] \quad (5b)$$

$$\frac{d[M2]}{dt} = \frac{V_{max(E2)} [MI]}{K_m(E2) + [MI]} - \frac{V_{max(E3)} [M2]}{K_m(E3) + [M2]} - \lambda_{M2} [M2] \quad (6)$$

$$\frac{d[Pd]}{dt} = \frac{V_{max(E3)} [M2]}{K_m(E3) + [M2]} - \lambda_{Pd} [Pd] - \omega_{Pd} [Pd] \quad (7)$$

# Table of parameters:

Module components (stocks)	Stock Label	Processes & Constants
<b>EXPRESSION</b> Nuclear mRNA Cytoplasmic mRNA	<i>nR</i> <i>cR</i>	transcription ( $\alpha_{nR}$ ) = $2.42 \times 10^{-4}$ transcripts allele <sup>-1</sup> s <sup>-1</sup> transport ( $\alpha_{cR}$ ) = $8.33 \times 10^{-4}$ s <sup>-1</sup> decay ( $\lambda_{nR}$ ) = $5.78 \times 10^{-4}$ s <sup>-1</sup> decay ( $\lambda_{cR}$ ) = $9.64 \times 10^{-5}$ s <sup>-1</sup>
<b>TRANSLATION</b> Enzyme	E	translation ( $\alpha_E$ ) = $0.028$ s <sup>-1</sup> decay ( $\lambda_E$ ) = $4.81 \times 10^{-5}$ s <sup>-1</sup>
<b>CATALYSIS</b> Precursor Metabolite1 Metabolite2 Product	<i>Pc</i> <i>M1</i> <i>M2</i> <i>Pd</i>	inflow ( $\alpha_{Pc}$ ) = $1.0 \times 10^{-9}$ Ms <sup>-1</sup> $K_m$ (all enzymes) = $1.0 \times 10^{-4}$ M $k_{cat}$ (all enzymes) = $1.0$ s <sup>-1</sup> $K_i$ = $1.0 \times 10^{-6}$ M decay (all metabolites) ( $\lambda_{Pc-Pd}$ ) = $2.78 \times 10^{-5}$ s <sup>-1</sup> outflow ( $\omega_{Pd}$ ) = $2.78 \times 10^{-4}$ s <sup>-1</sup>

Stocks and baseline rate constants used in the models. Most constants obtained from Hargrove (1998).  $K_m$ ,  $k_{cat}$  and  $K_i$  were estimated to be within known ranges (Stryer 1993; Kakkar *et al.* 1999). Note that some processes are equivalent, e.g, transport of nuclear mRNA, which is both the outflow from this stock and the inflow to cytoplasmic mRNA.

## Simple calculation of steady state concentration from rate constants:

$$0 = \alpha_{nR} - (\lambda_{nR} + \alpha_{cR}) nR^*$$

Rearranging gives the steady state of the stock:

$$nR^* = \frac{\alpha_{nR}}{(\lambda_{nR} + \alpha_{cR})}$$

Steady state solution to Michaelis-Menten reactions  
More complex. Steady state solution to feedback  
Scenario solved using Brent's method:  
Brent, R. P. (1973) Algorithms for Minimization  
without Derivatives, Prentice-Hall, New Jersey.

# Baseline model properties

Cannot directly validate, but:

- Standard model  $[P]^* = 2.97\mu\text{M}$
- Feedback (CI) model  $[P]^* = 2.73\mu\text{M}$
- Cellular metabolites:  $1 \times 10^{-6}$  to  $1 \times 10^{-3}\text{M}$
- Model generates acceptable outcome for steady state product concentration

## Introduction of genetic variation:

- Alteration of kinetic constants to simulate changes to transcription rate.
- STELLA not best suited for this purpose
- Model recoded in Splus / R
- Specification of functions to compute  $[P]^*$  for each pathway variant
- Specification of allele frequency in population of pathways
- Calculation of variance in  $[P]^*$
- Available on [www.vipbg.vcu.edu/~edwin](http://www.vipbg.vcu.edu/~edwin)

## Calculation of Variance Components:

- Total variance of  $[P]^*$  ( $V_T$ ) comprised of additive ( $V_A$ ), dominance ( $V_D$ ), and interaction components ( $V_I$ ):

$$V_T = V_A + V_D + V_I$$

- 3 biallelic loci =  $3^3 = 27$  possible genotypes
- Assume Hardy-Weinberg and linkage equilibrium

## Calculation of variance components

- $p$  is the frequency of allele A
- $q = (1 - p)$  the frequency of allele a
- additive ( $V_{Ai}$ ) and dominance variance ( $V_{Di}$ ) explained by gene  $i$  equals (Falconer 1989):

$$V_{Ai} = 2pq[a + d(q-p)]^2$$

$$V_{Di} = (2pqd)^2$$

where the additive score:

$$a = ([Pd]_{AA}^* - [Pd]_{aa}^*)/2$$

and the dominance score:

$$d = [Pd]_{Aa}^* - ([Pd]_{AA}^* + [Pd]_{aa}^*)/2$$

# Calculation of variance components

- Finally, the interaction component:

$$V_I = V_T - (V_A + V_D)$$

# Results I

# Cross-validation of Splus and STELLA models

- All genes set identical and biallelic
- Expression allele 0 (Lo activity)  
0.87 transcripts hour<sup>-1</sup> allele<sup>-1</sup>
- Expression allele 1 (Hi activity)  
2.87 transcripts hour<sup>-1</sup> allele<sup>-1</sup>
- Standard model range: 2.97-3.18μM
- CI model range: 2.73-3.08μM
- All possible genotypes verified in STELLA model to yield identical results – code OK

## Standard Model

Genotypes ranked  
by Steady State  
concentrations  
of Product ( $\mu\text{M}$ )

Genotype G1	Genotype G2	Genotype G3	[P]* $\mu\text{M}$
0	0	0	2.969
0	0	1	3.021
0	1	0	3.021
1	0	0	3.021
0	0	2	3.037
0	2	0	3.037
2	0	0	3.037
0	1	1	3.074
1	0	1	3.074
1	1	0	3.074
0	1	2	3.090
0	2	1	3.090
2	0	1	3.090
1	0	2	3.090
1	2	0	3.090
2	1	0	3.090
0	2	2	3.106
2	0	2	3.106
2	2	0	3.106
1	1	1	3.127
1	1	2	3.144
1	2	1	3.144
2	1	1	3.144
1	2	2	3.161
2	1	2	3.161
2	2	1	3.161
2	2	2	3.177

# Variance Components (%) for standard (No CI) model

Expression level polymorphism – all genes kinetically identical

	Gene 1	Gene 2	Gene 3	Sum Total
% Additive ( $V_A$ )	28.0	27.9	27.9	83.8
% Dominance ( $V_D$ )	5.4	5.4	5.4	16.2
Total per gene	33.3	33.3	33.3	100

$$\text{Gene-Gene Interaction } (V_I) = 100 - 99.998 = 0.002\%$$

## Competitive Inhibition Model

Genotypes ranked  
by Steady State  
concentrations  
of Product ( $\mu\text{M}$ )

Genotype G1	Genotype G2	Genotype G3	[P]* $\mu\text{M}$
0	0	0	2.731
0	0	1	2.775
0	1	0	2.775
0	0	2	2.789
0	2	0	2.789
0	1	1	2.820
0	1	2	2.834
0	2	1	2.834
0	2	2	2.848
1	0	0	2.895
1	0	1	2.944
1	1	0	2.944
2	0	0	2.951
1	0	2	2.959
1	2	0	2.959
1	1	1	2.993
2	0	1	3.001
2	1	0	3.001
1	1	2	3.008
1	2	1	3.008
2	0	2	3.017
2	2	0	3.017
1	2	2	3.023
2	1	1	3.052
2	1	2	3.068
2	2	1	3.068
2	2	2	3.084

# Variance Components (%) for feedback (CI) model

Expression level polymorphism – all genes kinetically identical

	Gene 1	Gene 2	Gene 3	Sum Total
% Additive ( $V_A$ )	75.4	5.5	5.5	86.4
% Dominance ( $V_D$ )	11.4	1.1	1.1	13.6
Total per gene	85.8	6.6	6.6	100

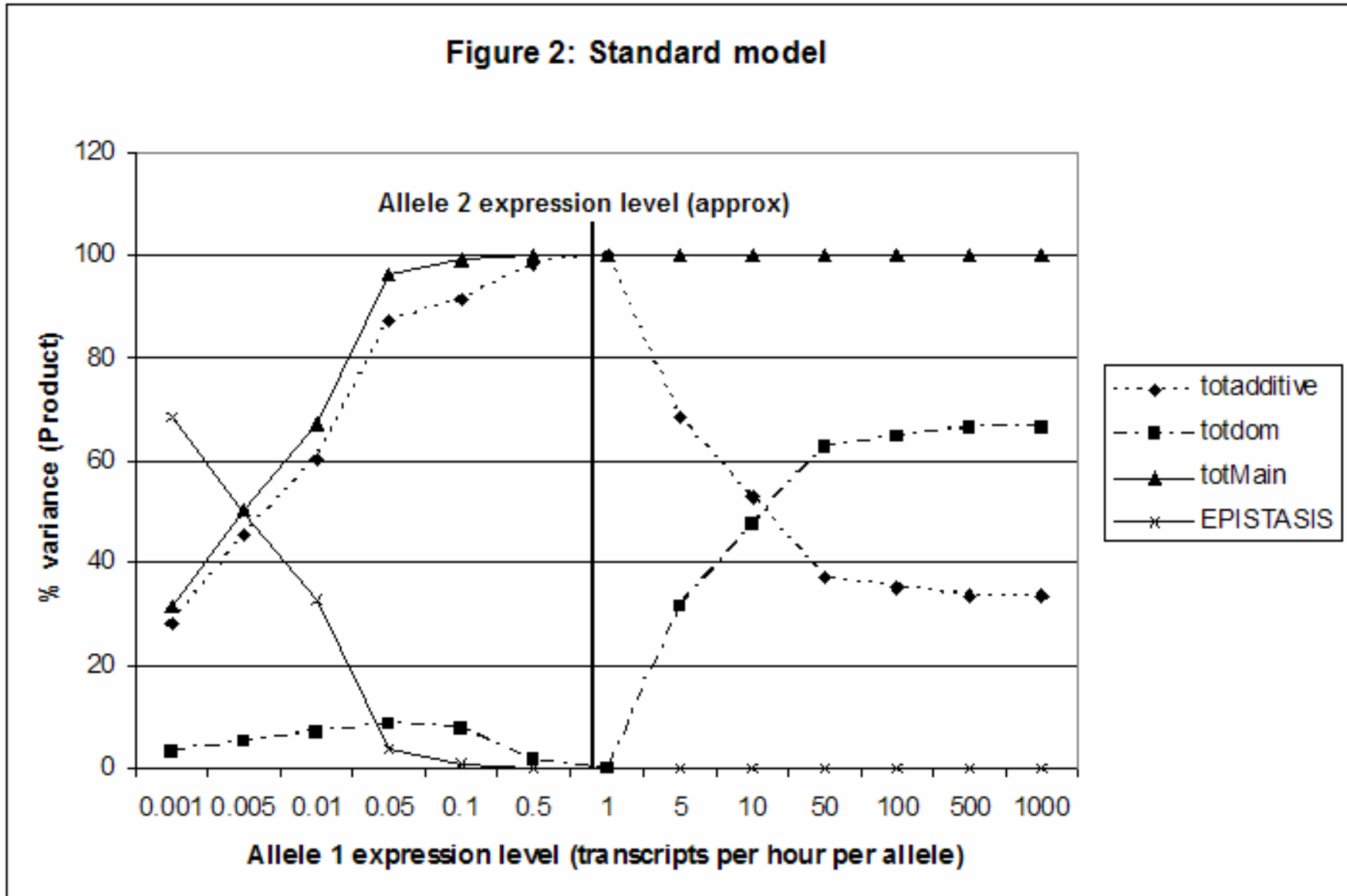
$$\text{Gene-Gene Interaction } (V_I) = 100 - 99.991 = 0.009\%$$

# Parameter changes

- Patterns robust to parameter changes
- Maintaining same Hi / Lo expression alleles, alterations of mRNA transport, mRNA decay, enzyme translation and decay across five orders of magnitude produced limited epistasis
- Maximum epistasis observed:
  - 2.4% (Standard model)
  - 2.0% (CI model)
- Max epi when mRNA decay highest (100 times baseline), similar values when translation lowest (100 times less than baseline).

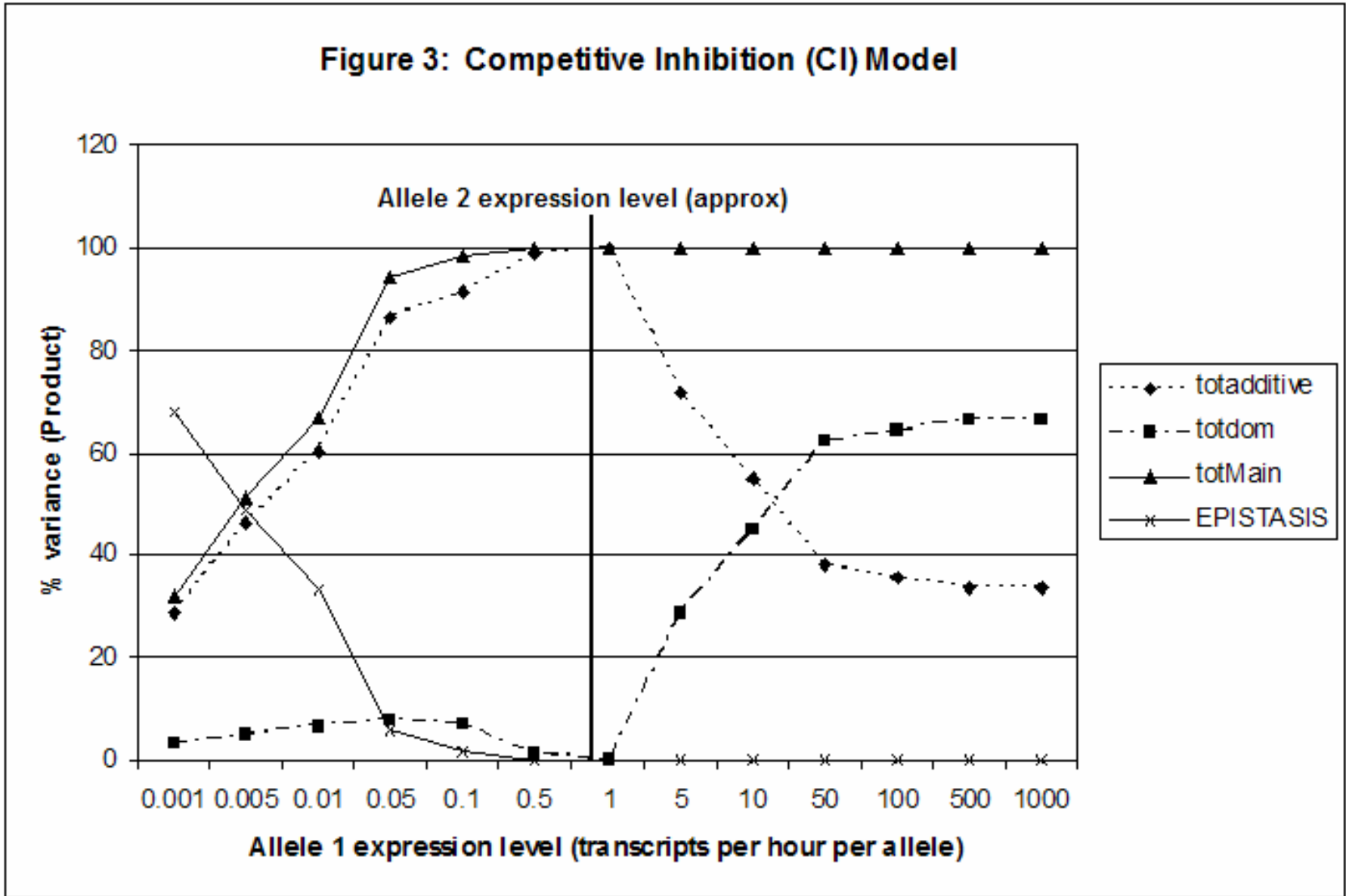
# Variance components across range of allele transcription levels

Figure 2: Standard model



# Variance components across range of allele transcription levels

Figure 3: Competitive Inhibition (CI) Model



# Variance components across range of allele transcription levels

Figure 4: Competitive Inhibition (CI) model, Increased Feedback

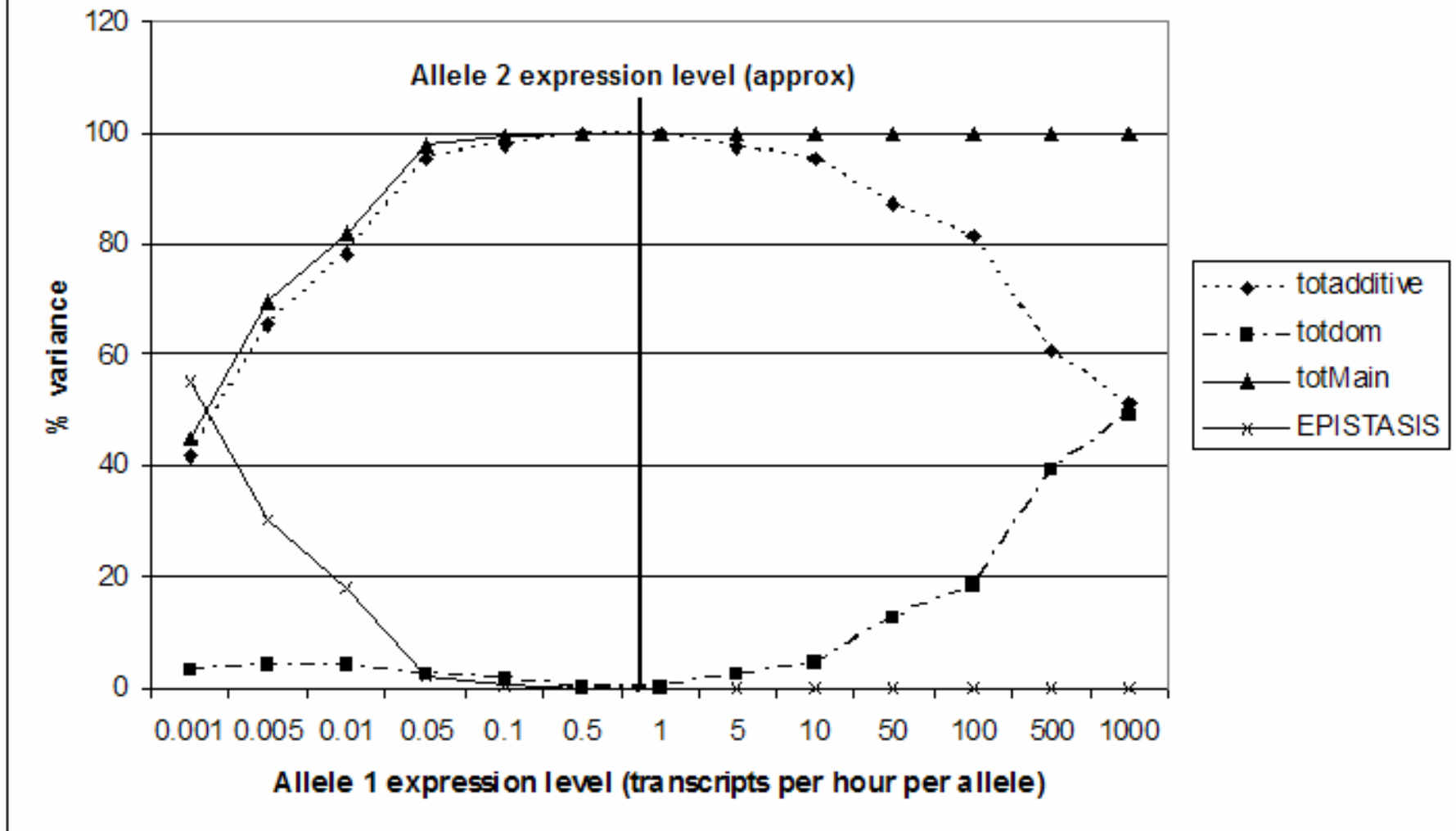
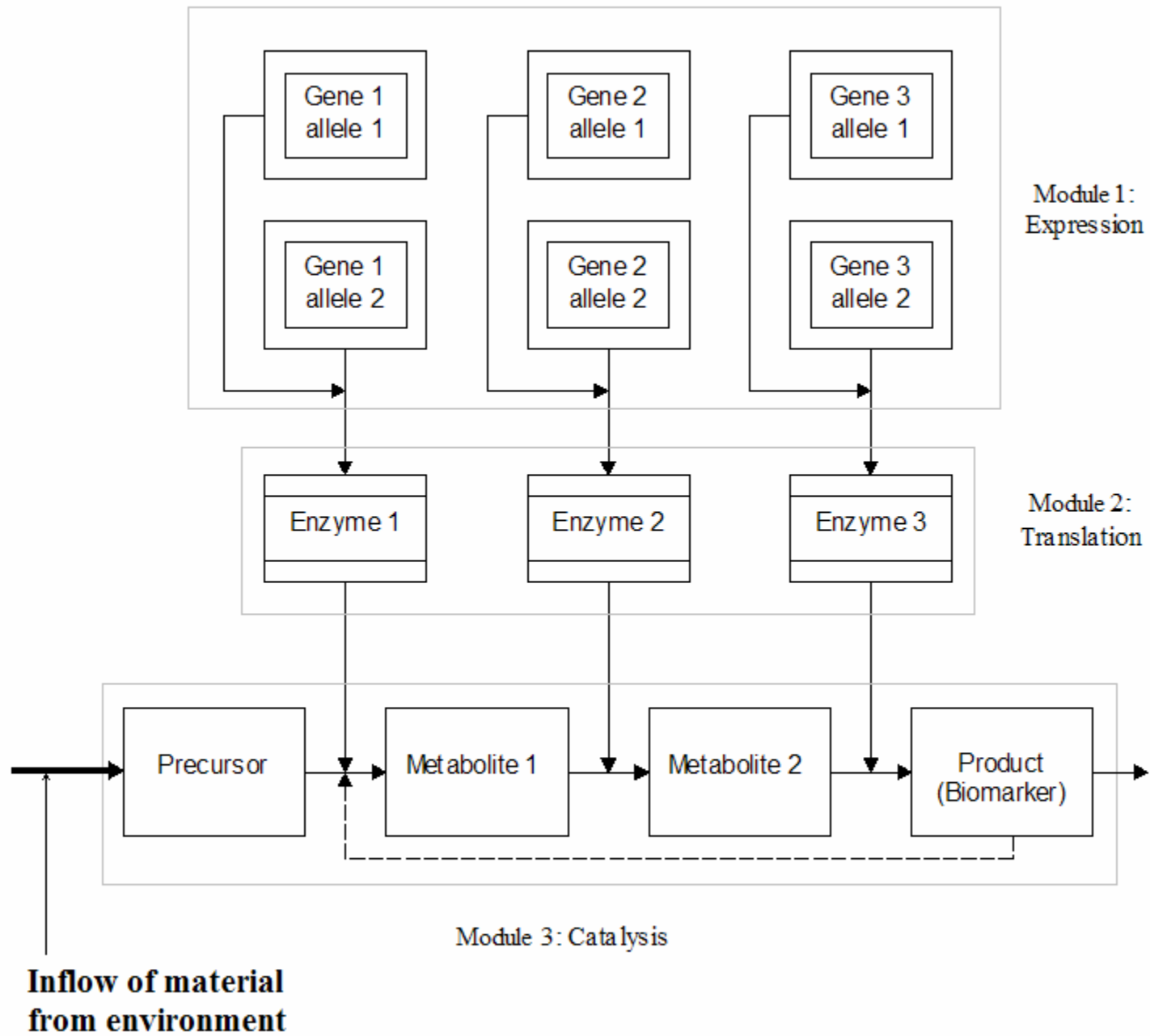
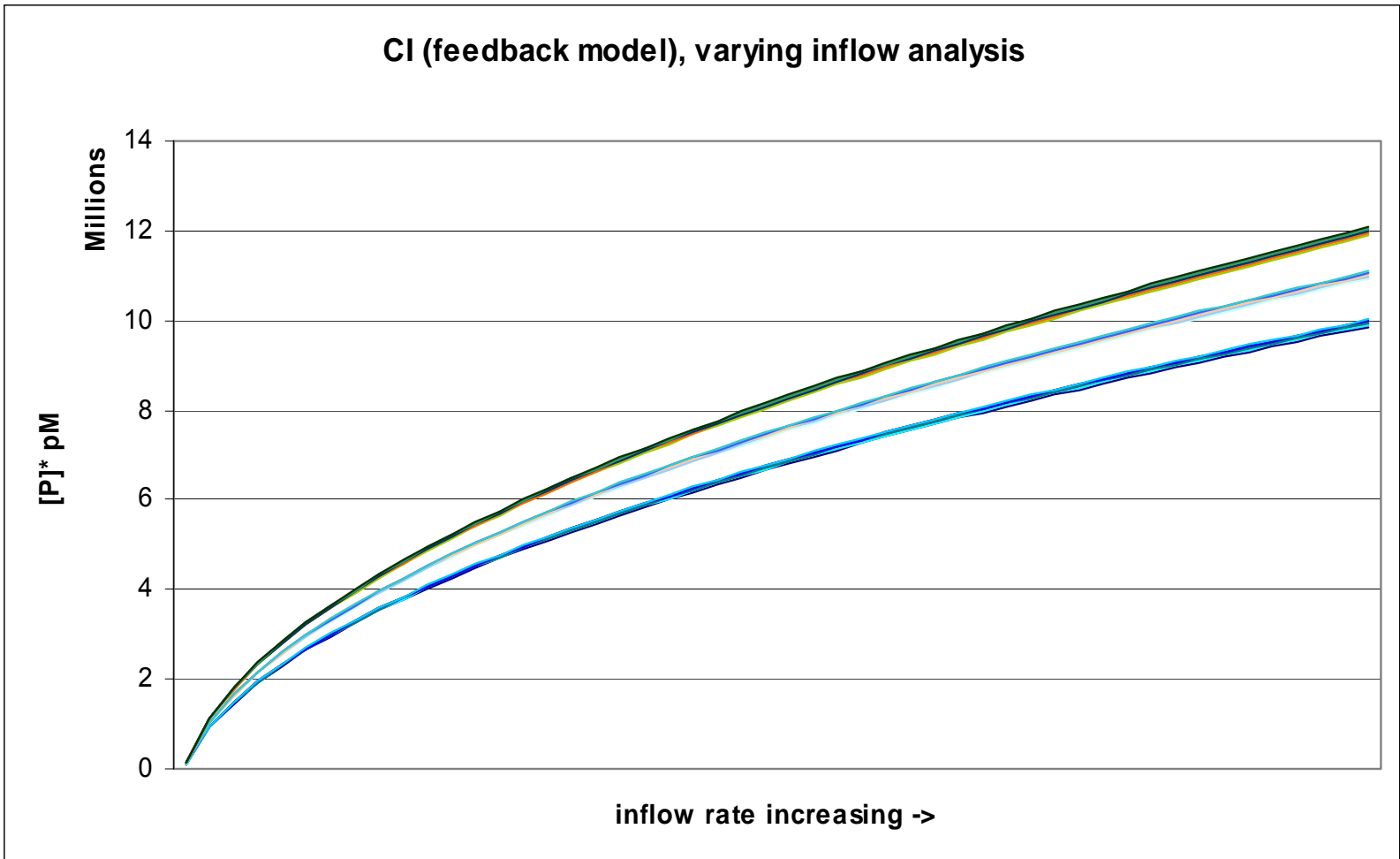


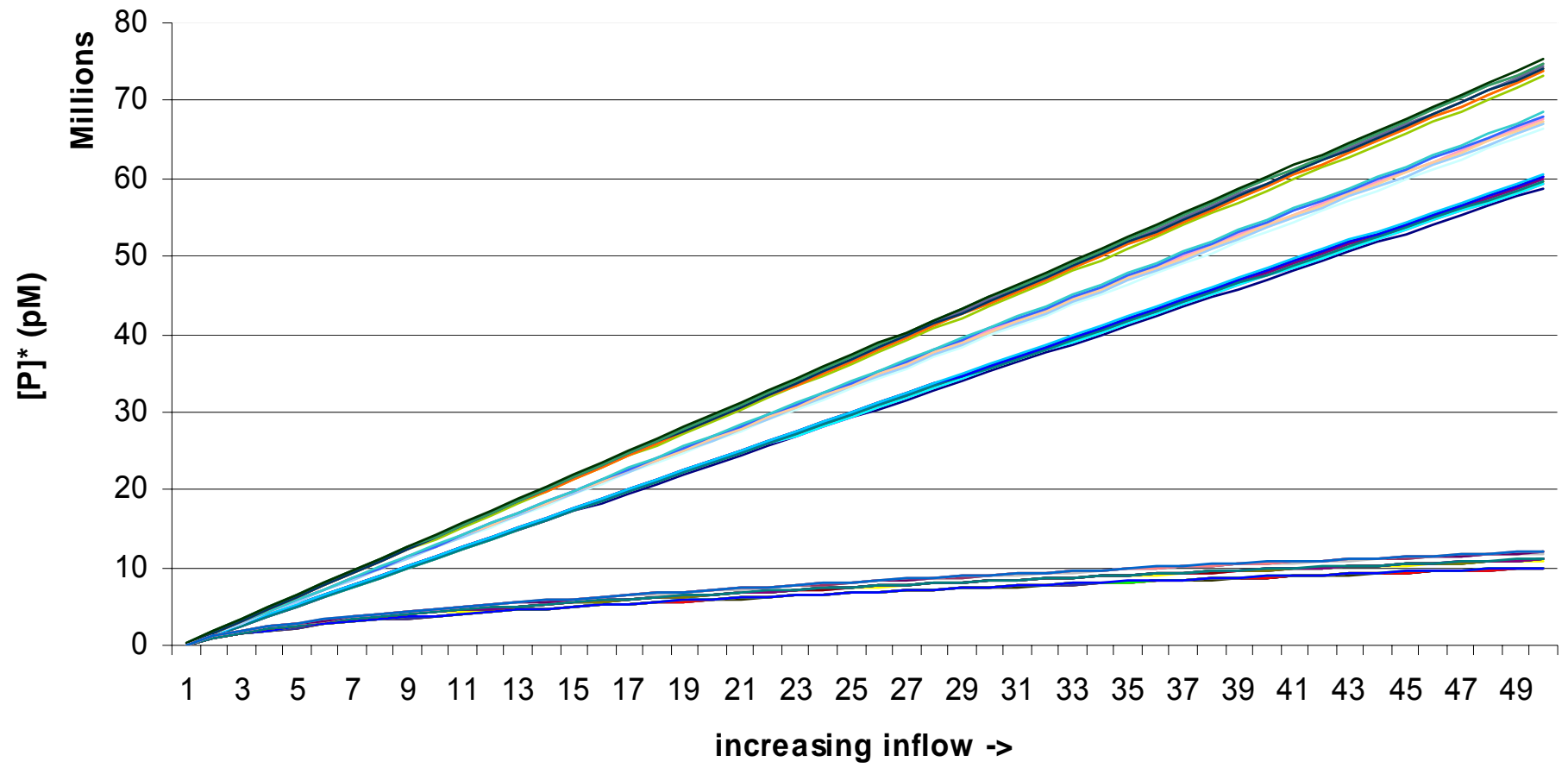
Figure 1





Stepwise inflow increase in 50 increments from 10 – 5000 pMs<sup>-1</sup>

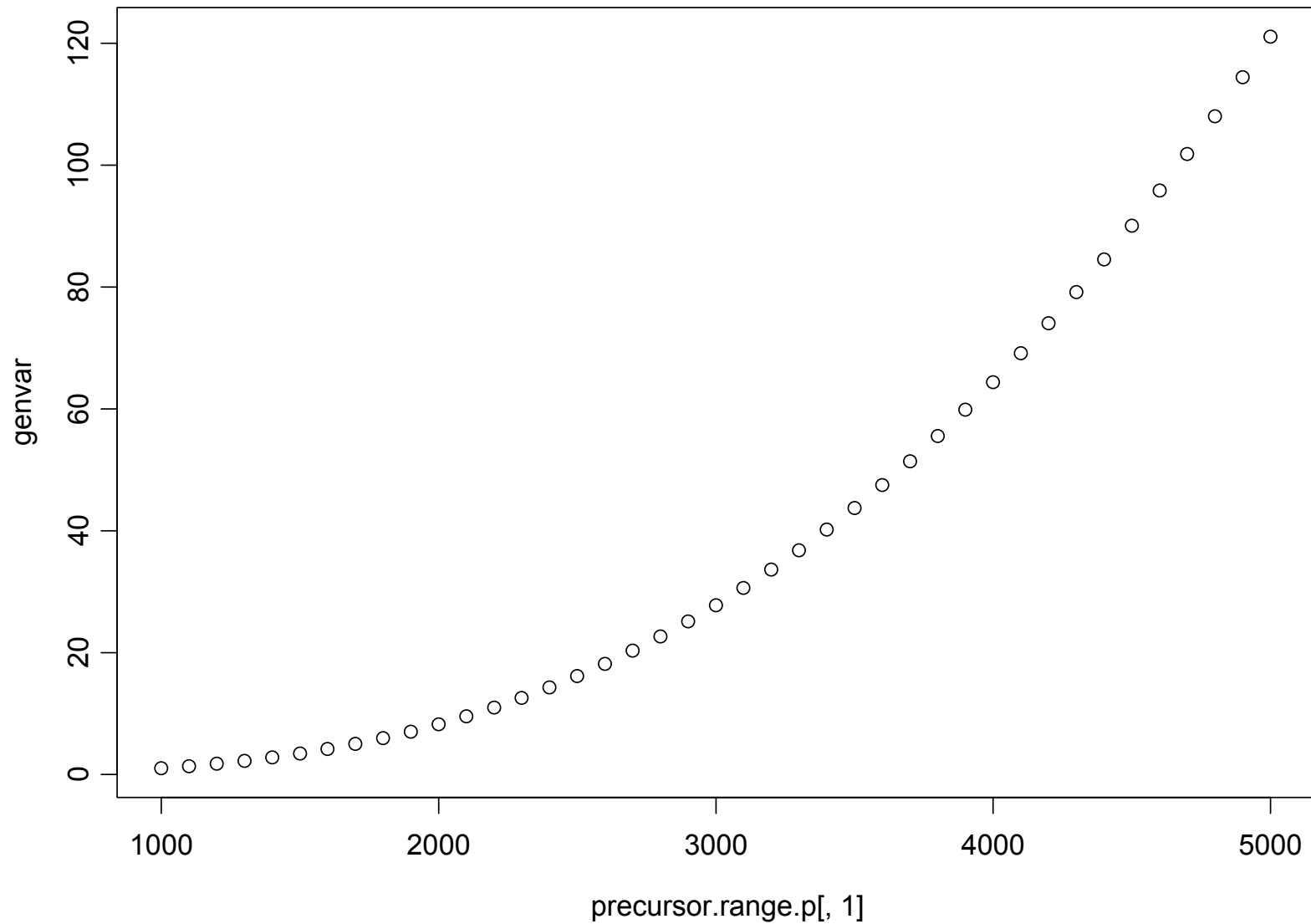
### Comparison of response to changing inflow rates between Standard and CI models



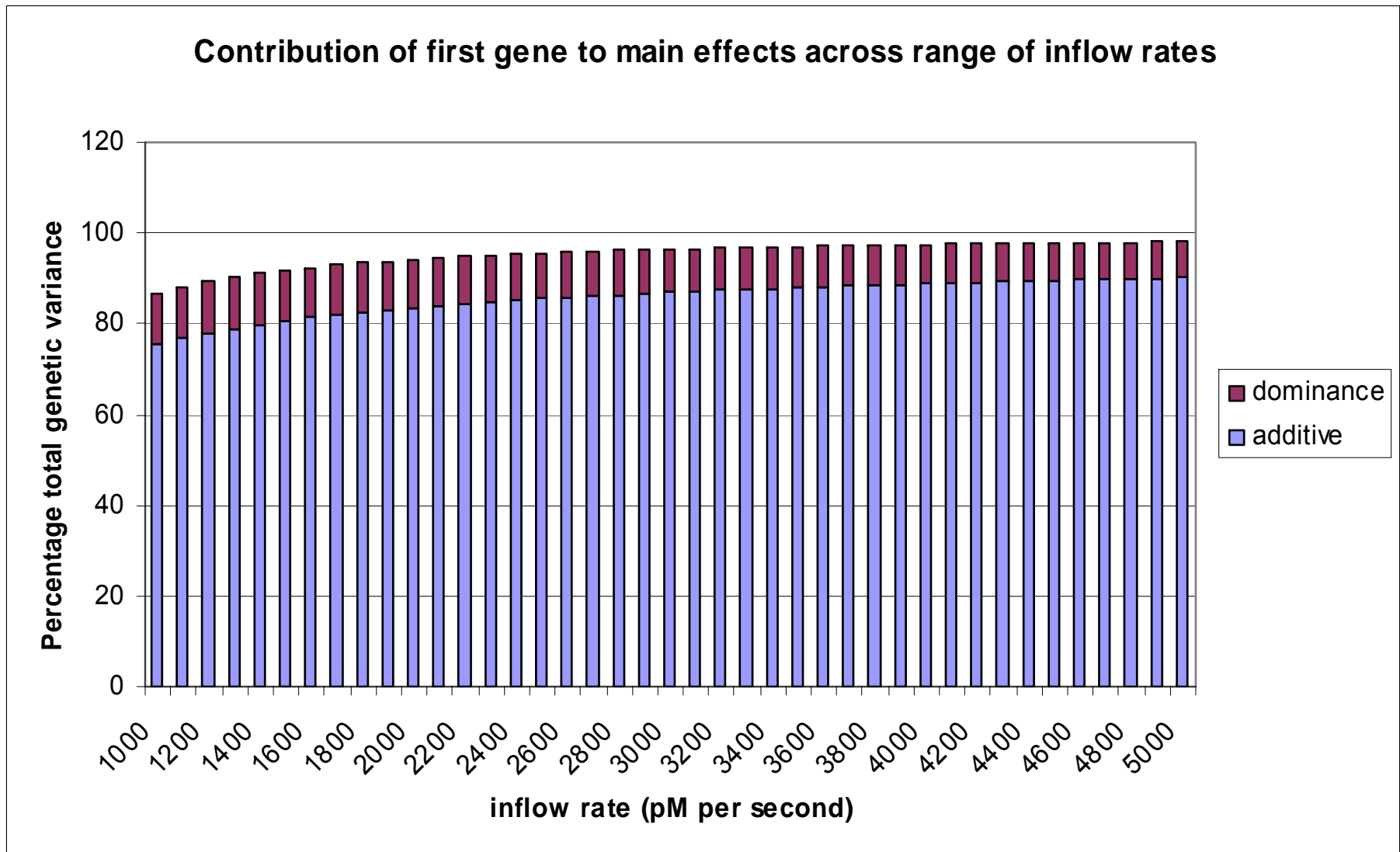
# Quantifying environment

- Genetic case
  - 1. (relatively) precise knowledge of biological processes
  - 2. Kinetic constants
  - 3. Population genetics
  
  - Environment we cannot model with same precision, so don't know the distribution and can't compute proportion of GxE compared to total variance.
  - However, we can compute the increase in genetic variance as a function of environment

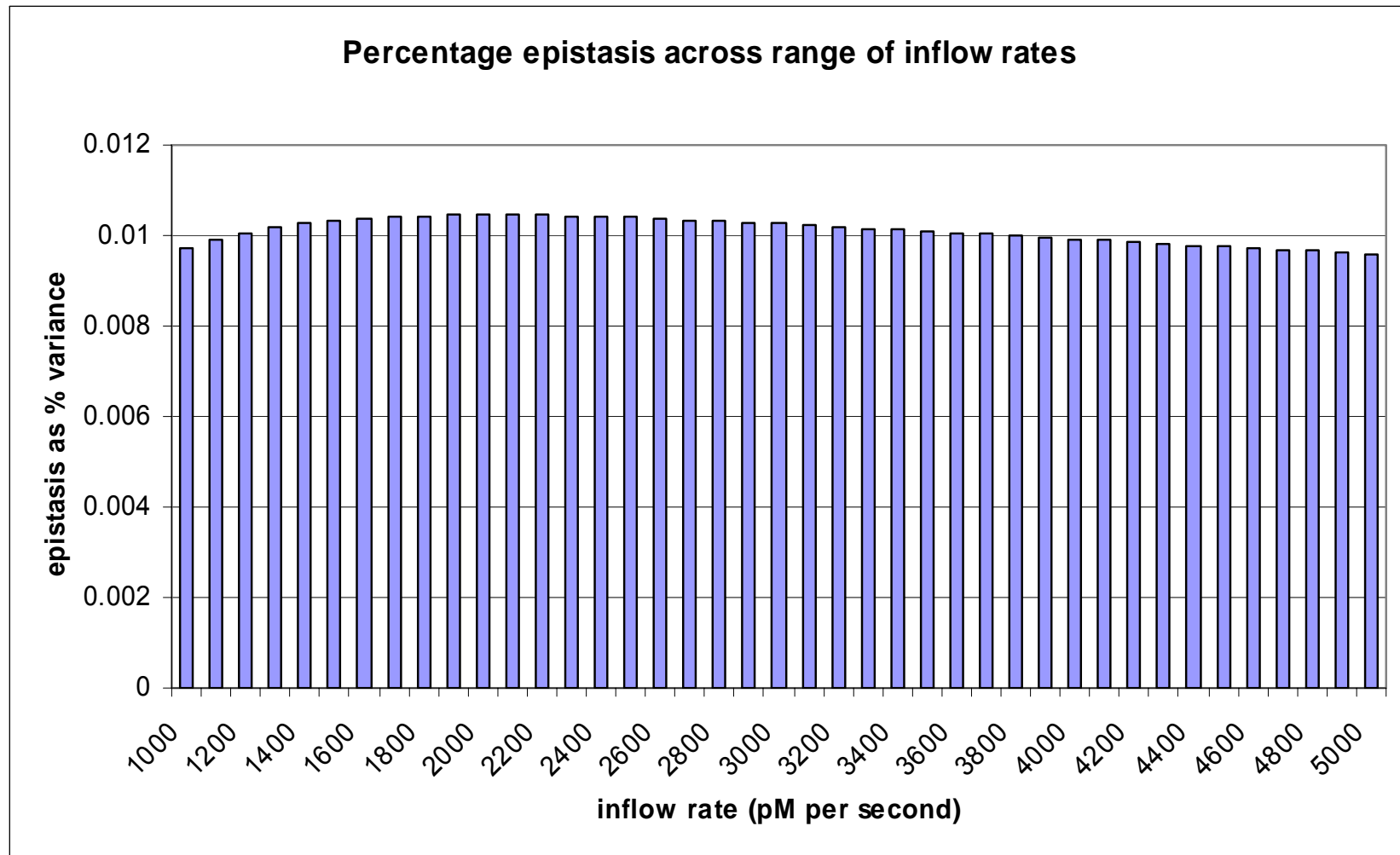
# Proportional change in genetic variance as function of flow of precursor from environment



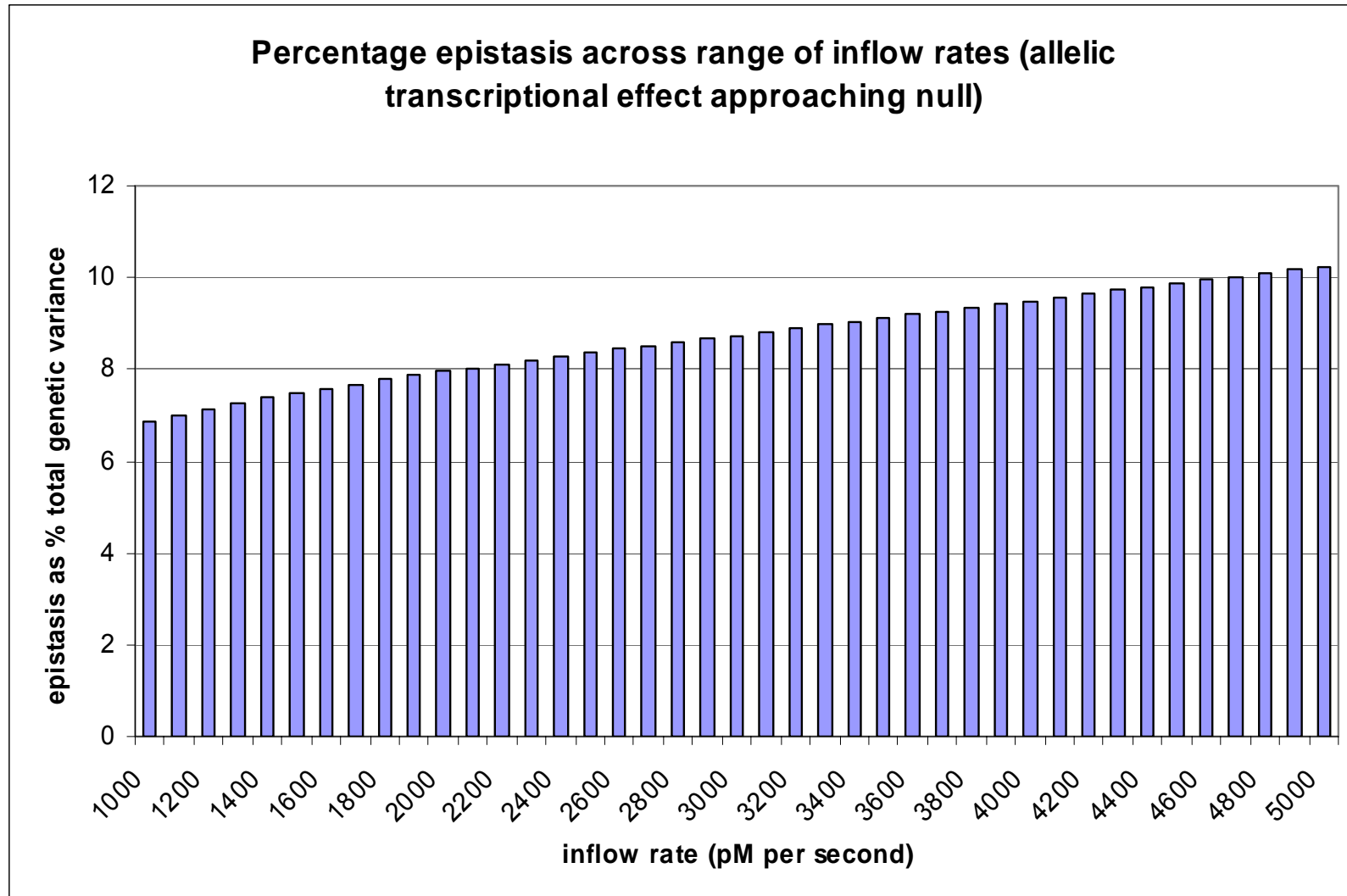
# Change in contribution of first gene to overall genetic variance, allelic effect standard Hi,Lo



Change in epistatic variance (as proportion of total genetic effects) as function of flow of precursor from environment, allelic effect standard Hi,Lo



Change in epistatic variance (as proportion of total genetic effects) as function of flow of precursor from environment, allelic effect approaching null



# Discussion

- Mostly main effects
- Large additive contribution in such a complex system
- To be expected? (Cheverud & Routman 1995)
- Dominance in metabolic pathways, in accordance with Kacser & Burns (1981) and Bagheri & Wagner (2004)
- Variance component outcome from Standard model agrees well with enzyme flux models by Keightley (1989)

# Discussion (contd)

- General lack of epistasis except when allelic effect approaches null
- Biological interaction  $\neq$  statistical interaction
- *How* gene products interact may be more important than merely the fact that they do
- Physiological epistasis predicted when enzymes approach saturation (Bagheri-Chaichian *et al.*, 2003)
- Could this explain the pattern of epistasis in our models? That is:
  - maximal with alleles approaching null
  - increasing with inflow

# Discussion (contd)

- Variance components largely robust to parameter changes, except allelic effects
- When feedback (CI) is introduced, 1<sup>st</sup> gene in pathway accounts for greatest share of main effects
- Could modeling of real systems help us to identify those genes most crucial in complex phenotypes?

# Discussion (contd)

- GxE? Differential (nonequivalent) response of polymorphic systems to changes in flow rates from E
- As material flow from E increased, genetic variance increased dramatically
- Should we expect a main effect of E in all instances of GxE, at least in systems of these types?

# A thought experiment:

- Mill metaphor
  - genes as modifiers of environmental flows



FIGURE CXIII.

