

An integrative, multi-scale, genome-wide model reveals the phenotypic landscape of *Escherichia coli*

Javier Carrera, Raissa Estrela

演讲：孙隽
翻译：熊丹

- Background
- Materials and Methods
- Results
- Discussion

Background

- The development of an integrative genome-scale model : Holy Grail

- Potential {
 - discovery of novel properties and emerging behaviors
 - generating and testing predictable hypotheses
 - guiding experimentation
 - accelerating the in-depth understanding of cellular physiology

Early work

- E-cell

a modular software environment for whole-cell simulation that included organelle sub-models

More recently work

genome-scale simulations were performed to study complex phenomena

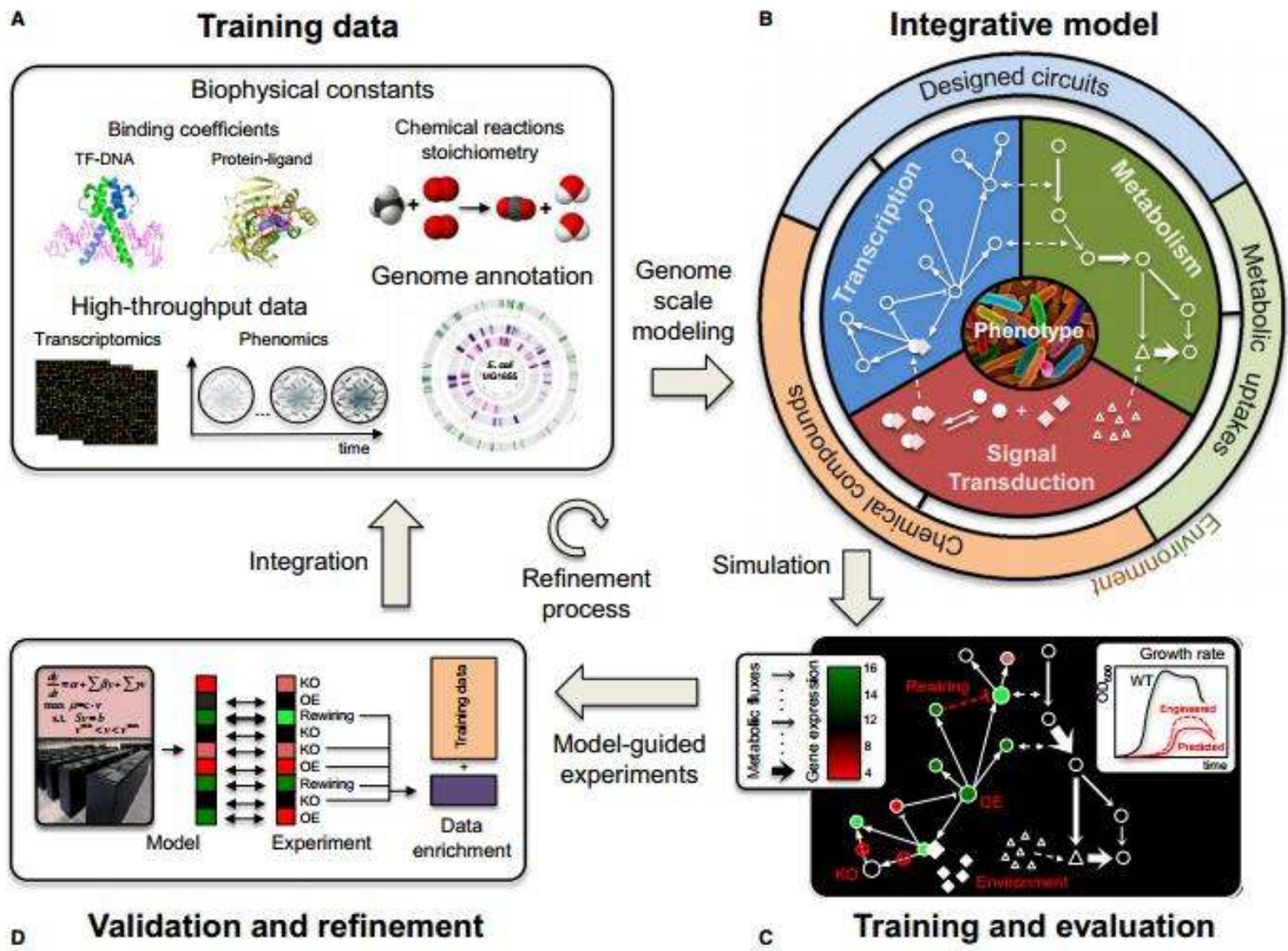
- the emergence of anticipatory behavior during evolution in varying environments
- the noise contributions of an inducible switch
- the effect of stochastic expression to metabolic variability

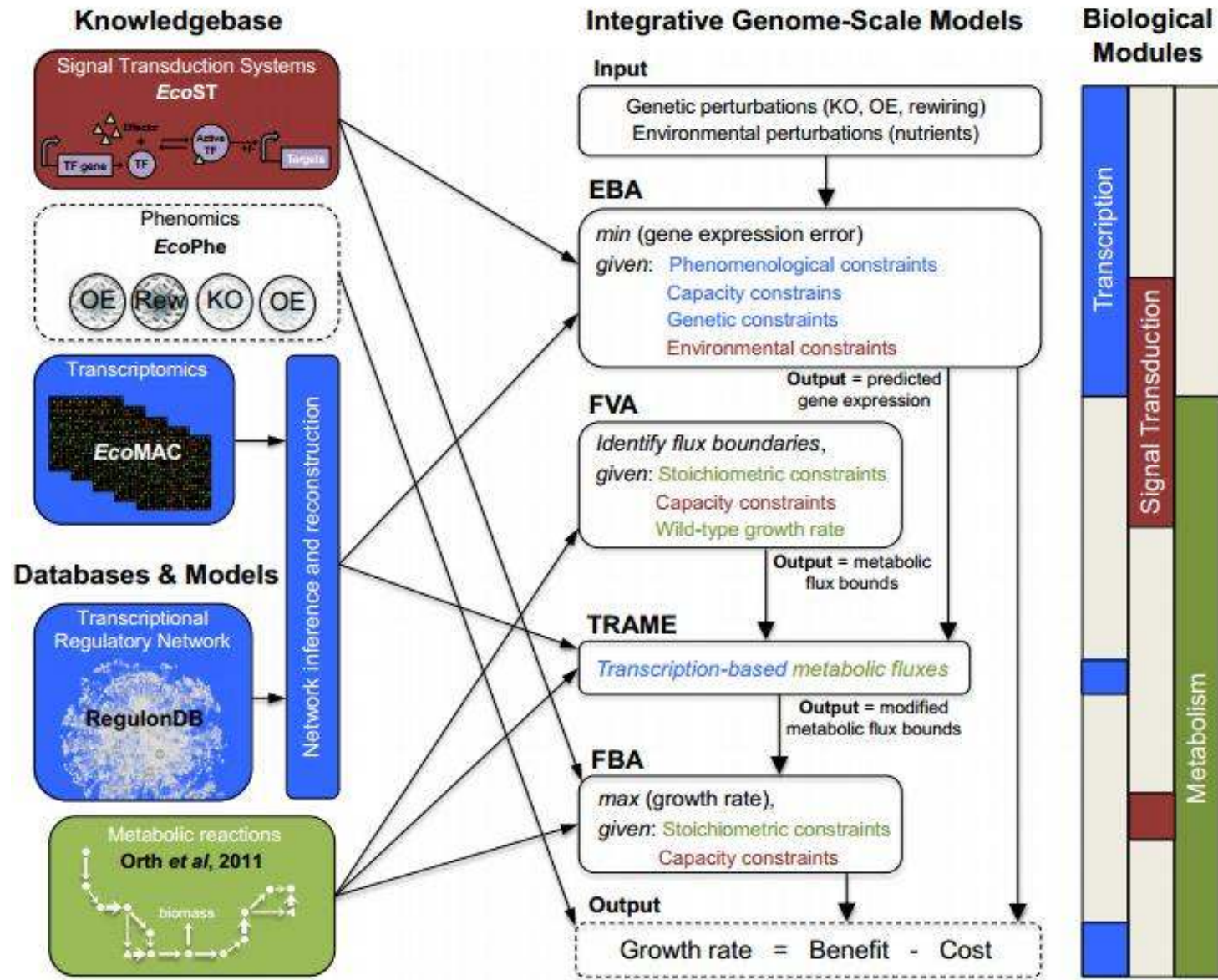
Our aim

To construct a phenomenological model for bacterial organisms that integrates multiple layers of biological organization.

Why *E.coli*

- The wealth of data and knowledge accumulated over the years
- The easiness to culture and manipulate experimentally
- Its importance in medical and biotechnological applications





Materials and Methods

- Data
- Cellular sub-models
- Model integration

Data

- **gene expression**

- from GEO, ASAP database
- constructed a gene expression compendium of 4,189 genes over 2,198 arrays that were collected from 127 scientific articles

- **signal transduction**

- A total of 328 transcription factors (TFs) and 1,357 enzymes were identified by using RegulonDB
- identify 151 instances of signal transduction systems (STSs)

- **Phenomics compendium**

- bacterial growth information for 616 of the arrays in EcoMAC by EcoPhe

Cellular sub-models

- Signal transduction model

$$y_{\text{TF}} = y_{\text{TF}}^{\text{wt}} + \Omega(C_{\text{TF}}^{\text{max}} - C_{\text{TF}}^{\text{min}}) \chi_{\text{TF}}^E \frac{\Delta n_E}{\Delta n_E^{\text{max}}}$$

- Transcriptional model and EBA

$$E = \frac{1}{2} [\bar{y}_{\text{TF}} \quad \bar{\varepsilon}_{\text{TF}}] \bar{H} \begin{bmatrix} \bar{y}_{\text{TF}} \\ \bar{\varepsilon}_{\text{TF}} \end{bmatrix} + \bar{f} \begin{bmatrix} \bar{y}_{\text{TF}} \\ \bar{\varepsilon}_{\text{TF}} \end{bmatrix}$$

- Metabolic model and Transcription-based Flux Enrichment

$$PV_{\min} \leq v \leq PV_{\max}, \text{ where } P_e = \left(\frac{y_e}{y_e^{\text{wt}}} \right)^n$$

Model integration

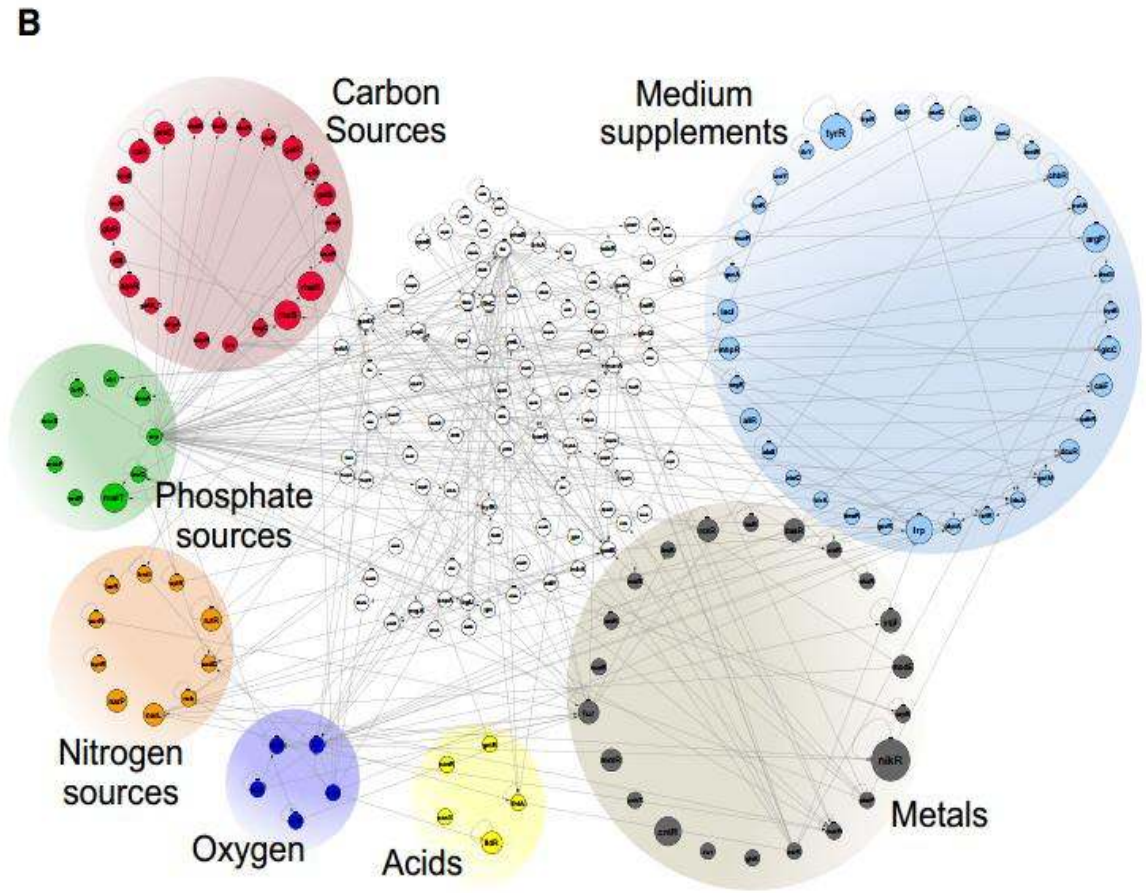
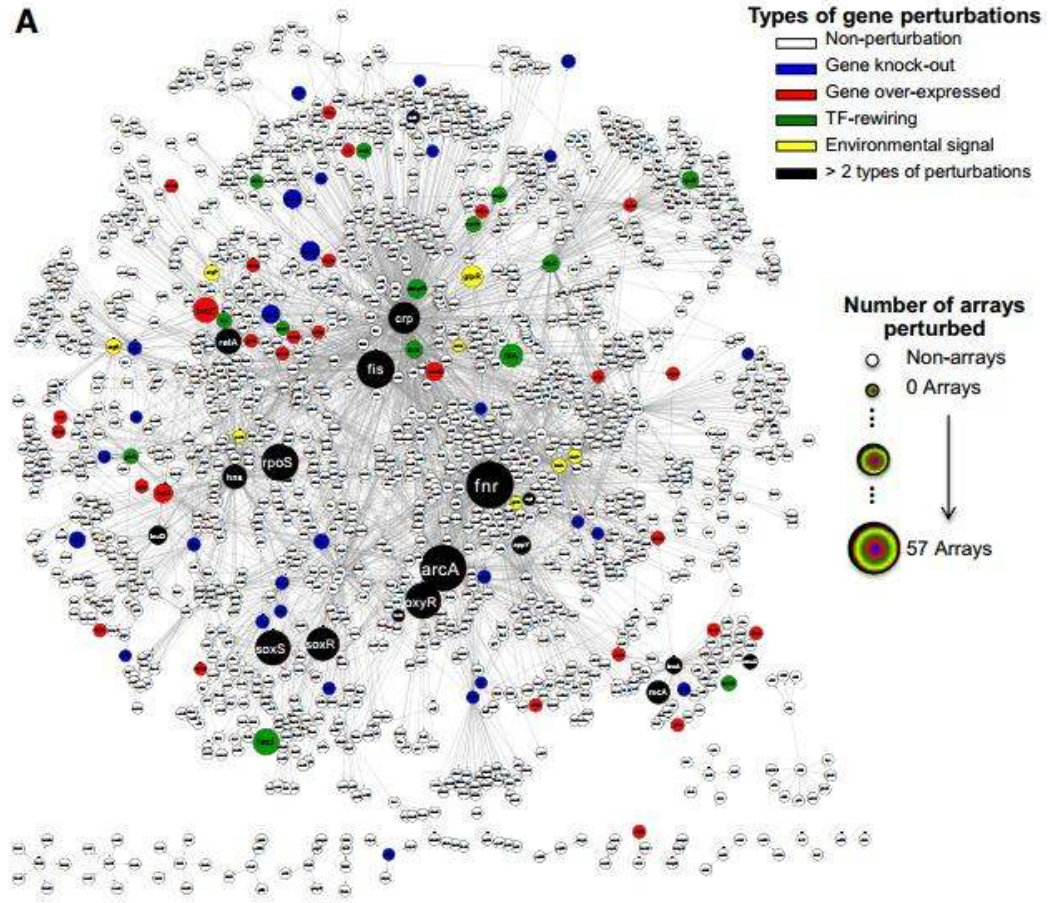
$$c = \frac{1}{N_G} \sum_g \left| \frac{\bar{y}_g - y_g^{\text{WT}}}{y_g^{\text{WT}}} \right|$$

$$\bar{\mu} = B - c$$

Results

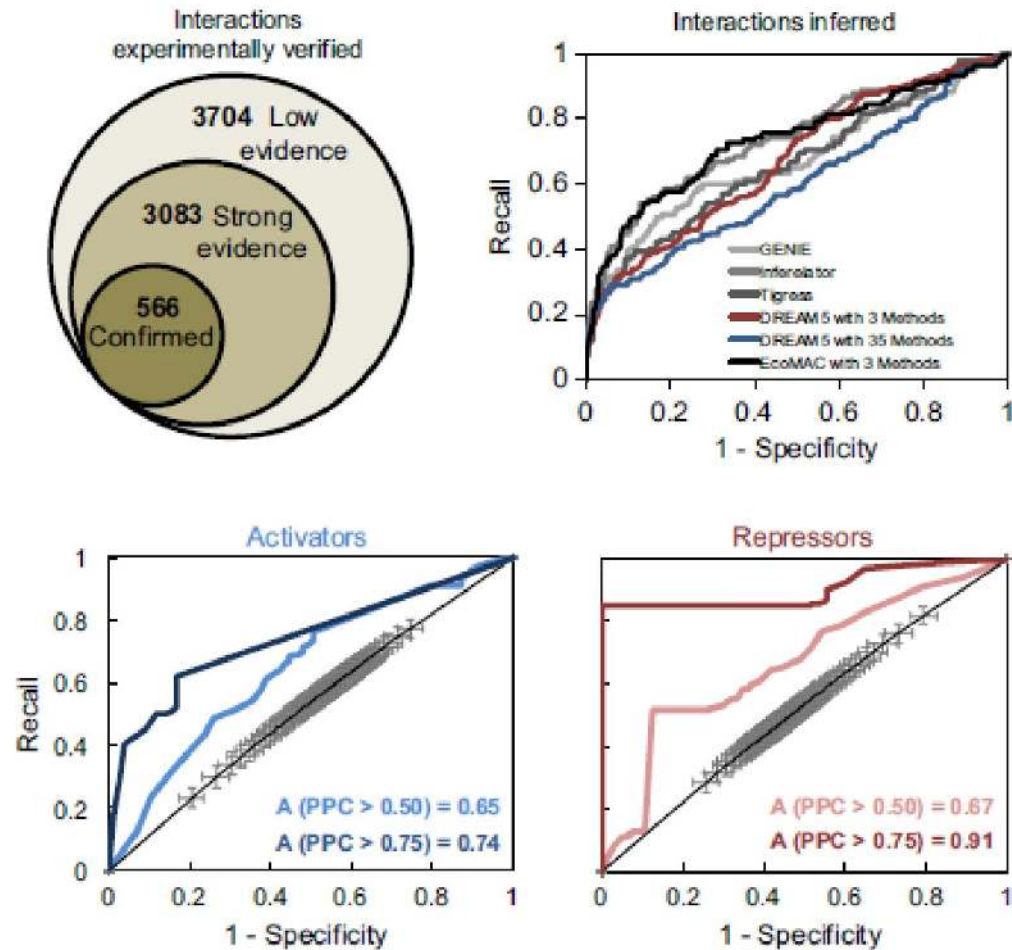
- Genetic and environmental gene expression diversity
- An integrative knowledgebase as a base to regulatory network enrichment
- Expression Balance Analysis
- Phenotypic predictions in an integrated model
- Model enrichment through targeted experimentation

Genetic and environmental gene expression diversity

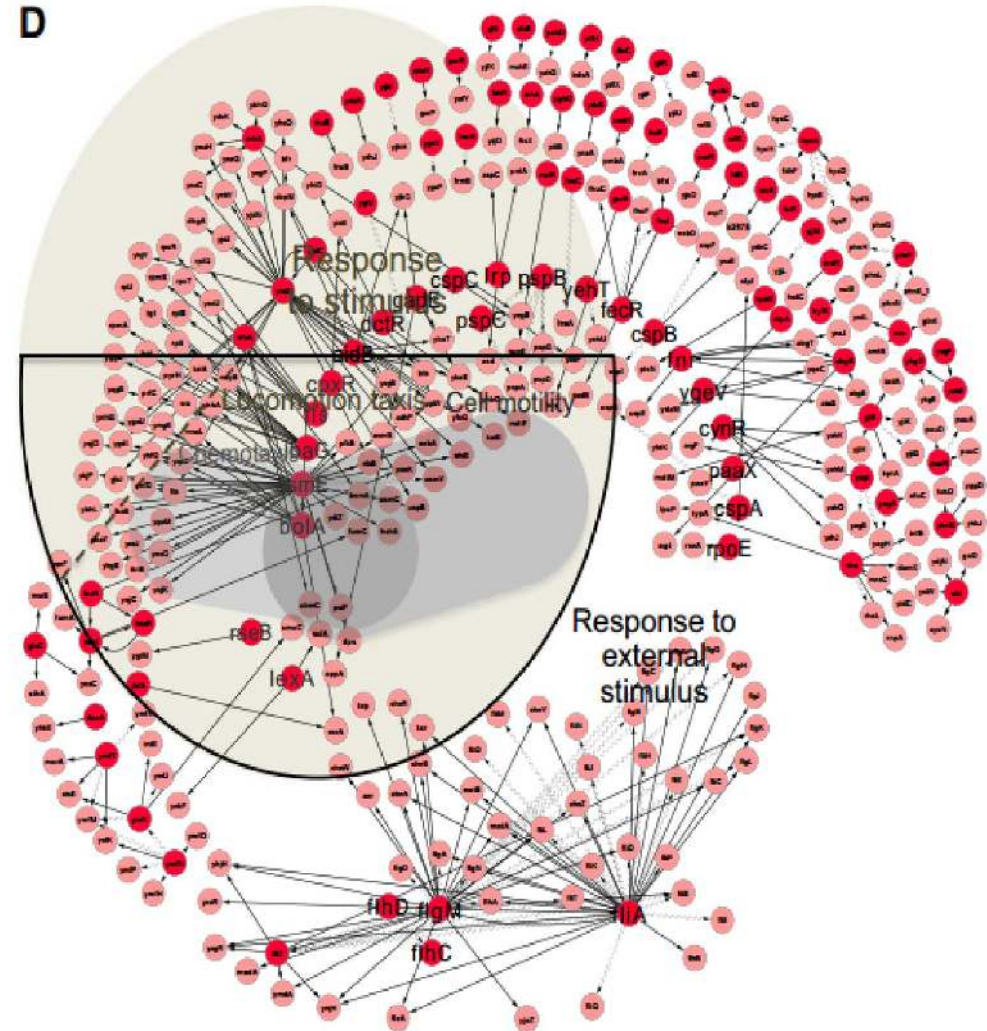


An integrative knowledgebase as a base to regulatory network enrichment

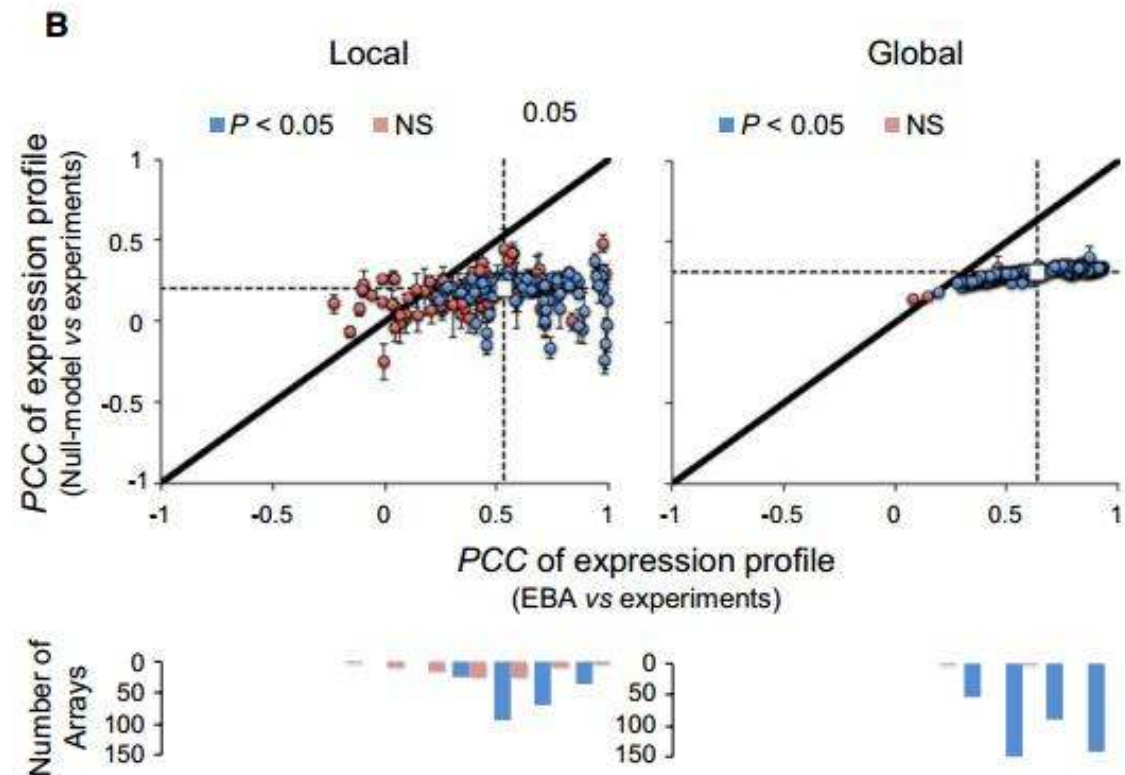
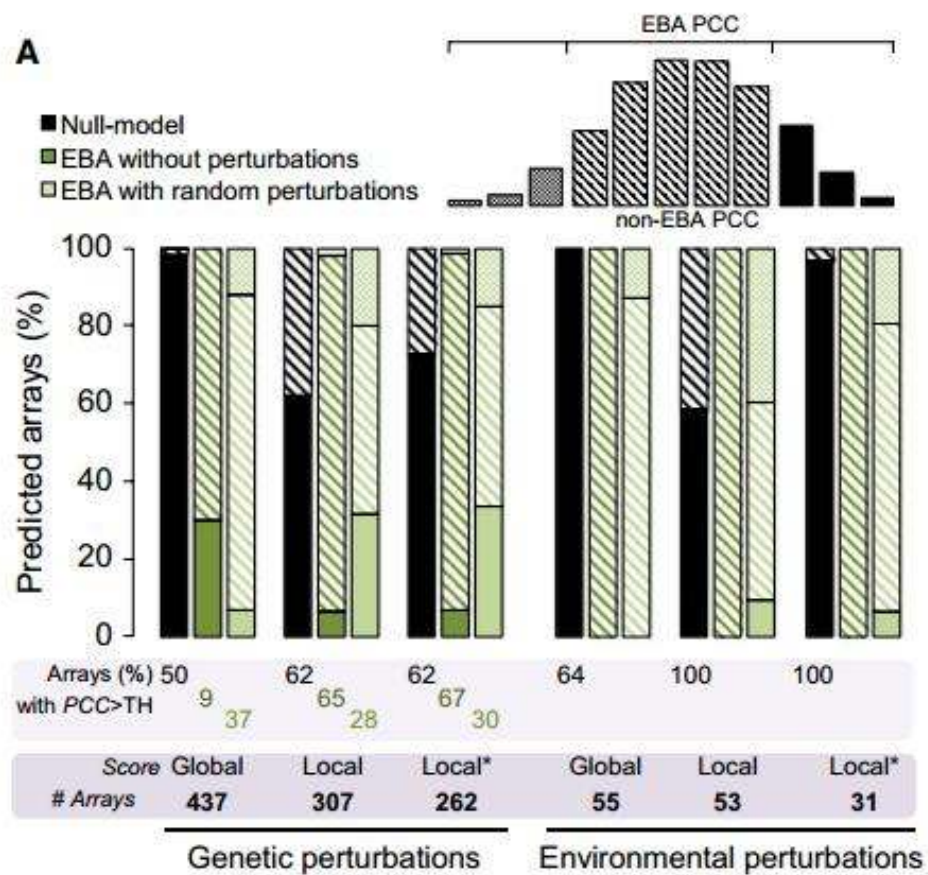
C



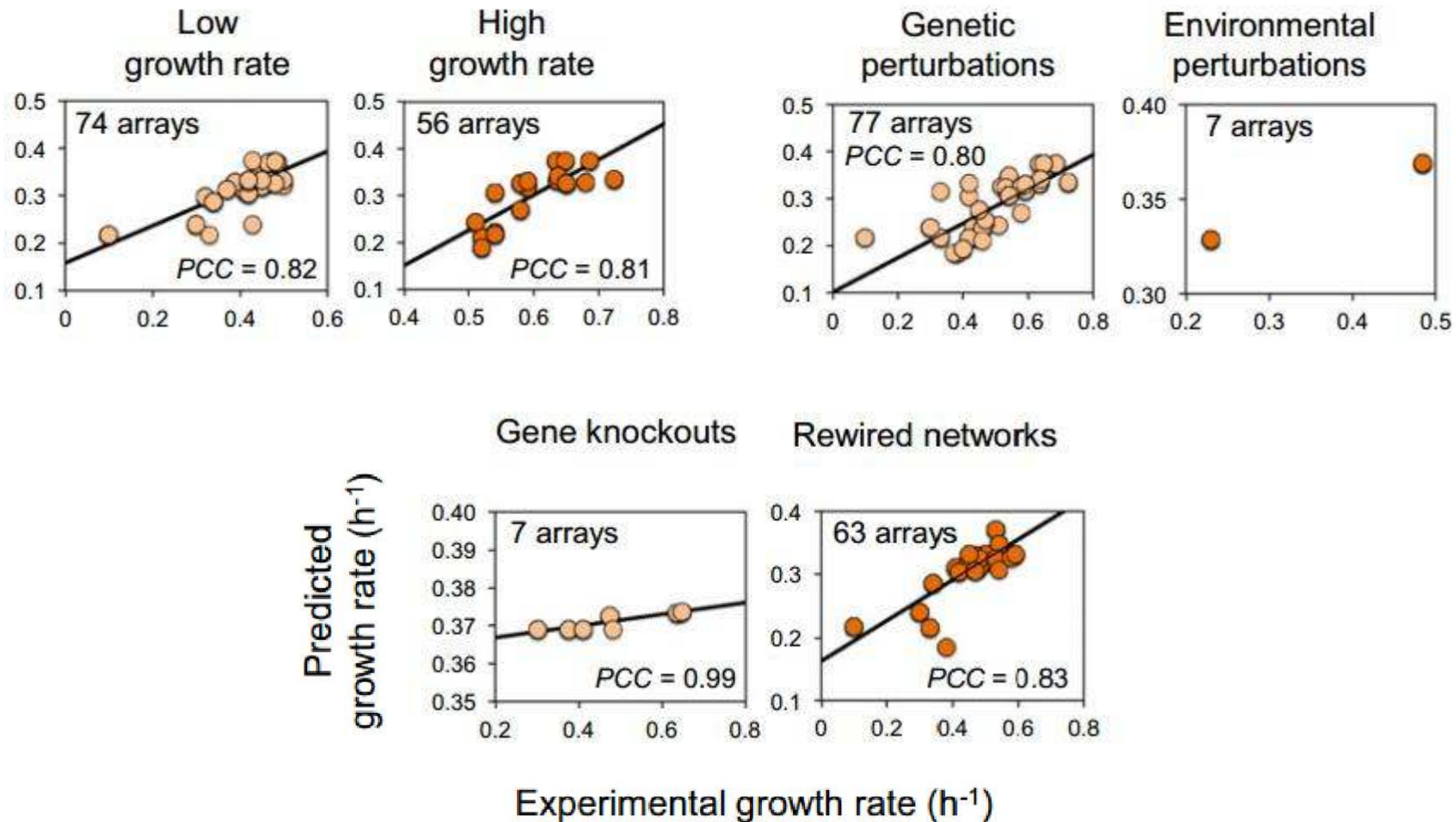
D



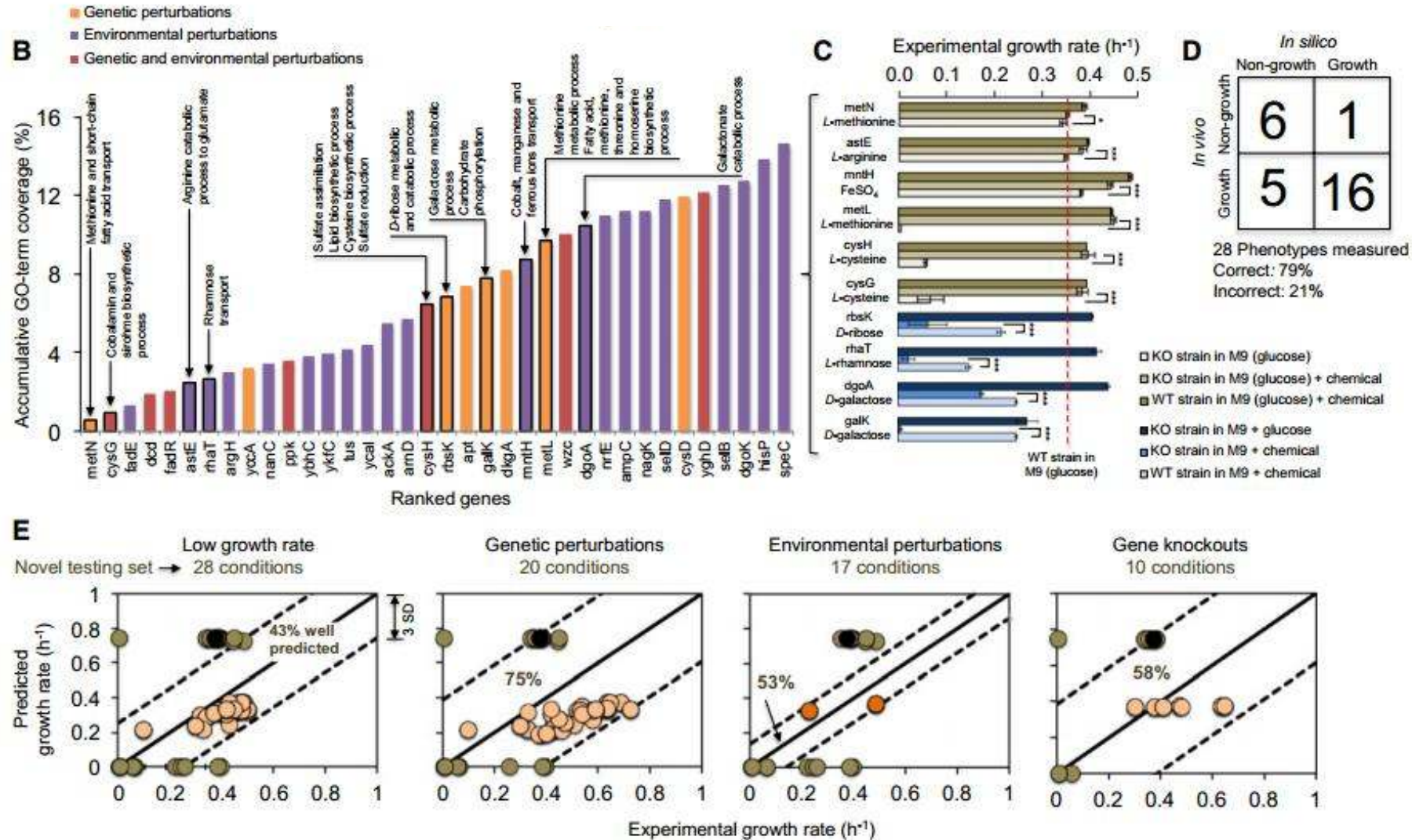
Expression Balance Analysis



Phenotypic predictions in an integrated model



Model enrichment through targeted experimentation



Discussion

- Advantage

- the creation of a signal transduction network (EcoST)
- its integration to the transcriptional and metabolic network through constraint modeling

- Disadvantage

- Coverage
- the severe bias to negative samples in the ground truth

Inspiration

- Binding site
- Funtion

Thanks for your attention