**Functional genomics**

...is a field of molecular biology that attempts to make use of the vast wealth of data produced by genomic projects (such as genome sequencing projects) to describe gene (and protein) functions and interactions. Unlike genomics, functional genomics focuses on the dynamic aspects such as gene transcription, translation, and protein–protein interactions, as opposed to the static aspects of the genomic information such as DNA sequence or structures. Functional genomics attempts to answer questions about the function of DNA at the levels of genes, RNA transcripts, and protein products. A key characteristic of functional genomics studies is their genome-wide approach to these questions, generally involving high-throughput methods rather than a more traditional “gene-by-gene” approach.

---

**The geneticist’s questions**

- **a)** What is consequence of reduced gene function?
  - 1) gene knockout (deletion, RNAi)
- **b)** What is the consequence of increased gene function?
  - 2) gene overexpression
- **c)** What does the gene (protein) interact with?
  - 3) genetic interactions
    - enhancers (“synthetic lethal”)
    - suppressors
  - 4) protein interaction
    - 2-hybrid
    - pull-down
- **d)** When and where is the gene (protein) expressed?
  - 5) measure gene expression
    - microarrays
    - RNA-Seq

---

**Deleting yeast genes**

[Diagram of yeast genome with genes and markers, showing deletion process]

---

**From Wikipedia, the free encyclopedia**

**Functional genomics**

...is like Chinese food: it looks great and tastes great, but you’re hungry 30 minutes after eating it...
Deletion “cassette” can be made in a PCR

Surveying the \textit{S. cerevisiae} genome for genes required for methylation of lysine 4 of histone H3.

\textit{Dover J et al.} \textit{J. Biol. Chem.} 2002;277:28368-28371

©2002 by American Society for Biochemistry and Molecular Biology

Each deletion strain tagged with two unique 20mers
Bar-coding gene deletions

Hybridize labeled tags to oligonucleotide array containing tag complements

Each tag has unique location

Parallel analysis of deletion strains

Apply Selection

Determine who is missing.

Determine who is here.

One tag → One Deletion Strain

GATTCGATAGCCGCAAGG → 1
CGATTAGGAATGTCATAG → 2
AGCTCATACTAGAATA → 3
... → 6,200

Ron Davis et al., Stanford University
Parallel Analysis of Gene Function

Before selection
Strains are present in equal abundance in the population and produce signals of equal intensity on the chip.

After multiple population doublings under selection
Strains with a growth defect are under-represented in the population and produce a lower intensity signal.

Ron Davis et al., Stanford University

G3 4: 11-18 (2014)
Genome Biol. 11: R60 (2010)
Genome Res. 19: 1836–1842 (2009)

Growth of deletion strains exhibiting reduced fitness in galactose media.


Growth of deletion strains exhibiting reduced fitness in galactose media.

ykl037
Verified putative protein of unknown function

YKL637w
UDP-glucose pyrophosphorylase

Growth of deletion strains exhibiting reduced fitness in galactose media.


yeo990
Dubious open reading frame; unlikely to encode a functional protein

Growth of deletion strains exhibiting reduced fitness in galactose media.


APJ1 YNL077W
Chaperone with a role in SUMO-mediated protein degradation; member of the DnaJ-like family; OE interferes with propagation of the prion; protein is detected in highly purified mitochondria in high-throughput studies; forms nuclear foci upon DNA replication stress

Growth of deletion strains exhibiting reduced fitness in galactose media.


ftr1
Iron permease

gef1
Voltage-gated chloride channel; involved in cation homeostasis

fet3
Ferro-O2-oxidoreductase; oxidizes Fe2+ to Fe3+ for uptake by Ftr1p
Results

Conventional analysis
~25% have growth defect
17.6% dead (~1100 essential genes)
~8% slow growth

Parallel analysis
~40% have growth defect (<98% of wt growth rate)
• Many new genes implicated in key biological processes
• Gene regulation poorly predicts mutant phenotype

Haploinsufficient genes are enriched for metabolic functions.

Haploinsufficient genes are enriched for metabolic functions.
Haploinsufficient genes are highly expressed.


Genetics Society of America

Copyright © 2013 by the Genetics Society of America

CRISPR-Cas9 nuclease

Clustered Regularly Interspaced Short Palindromic Repeats

Copyright © 2013 by the Genetics Society of America

T Wang et al. Science 2014;343:80-84

Published by AAAS

MMR is 6-TG sensitive

Resistance to topoisomerase inhibitor

T Wang et al. Science 2014;343:80-84

Published by AAAS
The geneticist’s questions

a) What is consequence of reduced gene function?
   1) gene knockout (deletion, RNAi)

b) What is the consequence of increased gene function?
   2) gene overexpression

c) What does the gene (protein) interact with?
   3) genetic interactions
   enhancers (“synthetic lethal”)
   suppressors
   4) protein interaction
   2-hybrid
   pull-down

d) When and where is the gene (protein) expressed?
   5) measure gene expression
   microarrays
   RNA-Seq
1) What is consequence of reduced gene function?
   a) gene knockout (deletion, RNAi)

2) What is the consequence of increased gene function?
   a) gene overexpression

3) What does the gene (protein) interact with?
   a) genetic interactions
      enhancers ("synthetic lethal")
      suppressors
   b) protein interaction
      2-hybrid
      pull-down

4) When and where is the gene (protein) expressed?
   a) measure gene expression
      microarrays
      RNA-Seq

---

Parallel pathways revealed in double mutants

Mutations in 1, 2, 3 are "synthetic lethal" with mutations in 4,5

- essential gene
- non-essential gene
- synthetic lethality (SL)
- information flow
- functional pathway

---

The geneticist’s questions
Parallel pathways revealed in double mutants

Mutations in 1, 2, 3 are "synthetic lethal" with mutations in 4, 5

- = essential gene
○ = non-essential gene
-= synthetic lethality (SL)
★ = information flow
□ = functional pathway

The number of genetic interactions averaged 34 in each screen for nonessential genes, with screens that were focused on essential genes exhibiting fivefold more interactions.

dense local neighbourhoods: the immediate neighbours of a gene—its genetic-interaction partners—also tend to interact with one another ~20% of the immediate neighbours of each of three test genes interacted with one another, vs. ~1% for any two genes chosen at random

connectivity distribution follows a power-law distribution:
many genes with few interactions and a few genes with many interactions

Genetic interactions tend to occur among functionally related genes

The fit to a straight line in the log-log plot indicates a power-law degree distribution, a characteristic of a "scale-free" network.

Fig. 4. (A) The degree distribution of SGA array genes not also used as query genes.
Fig. 1 A correlation-based network connecting genes with similar genetic interaction profiles.

M Costanzo et al. Science 2010;327:425-431
Published by AAAS