Chapter 19B

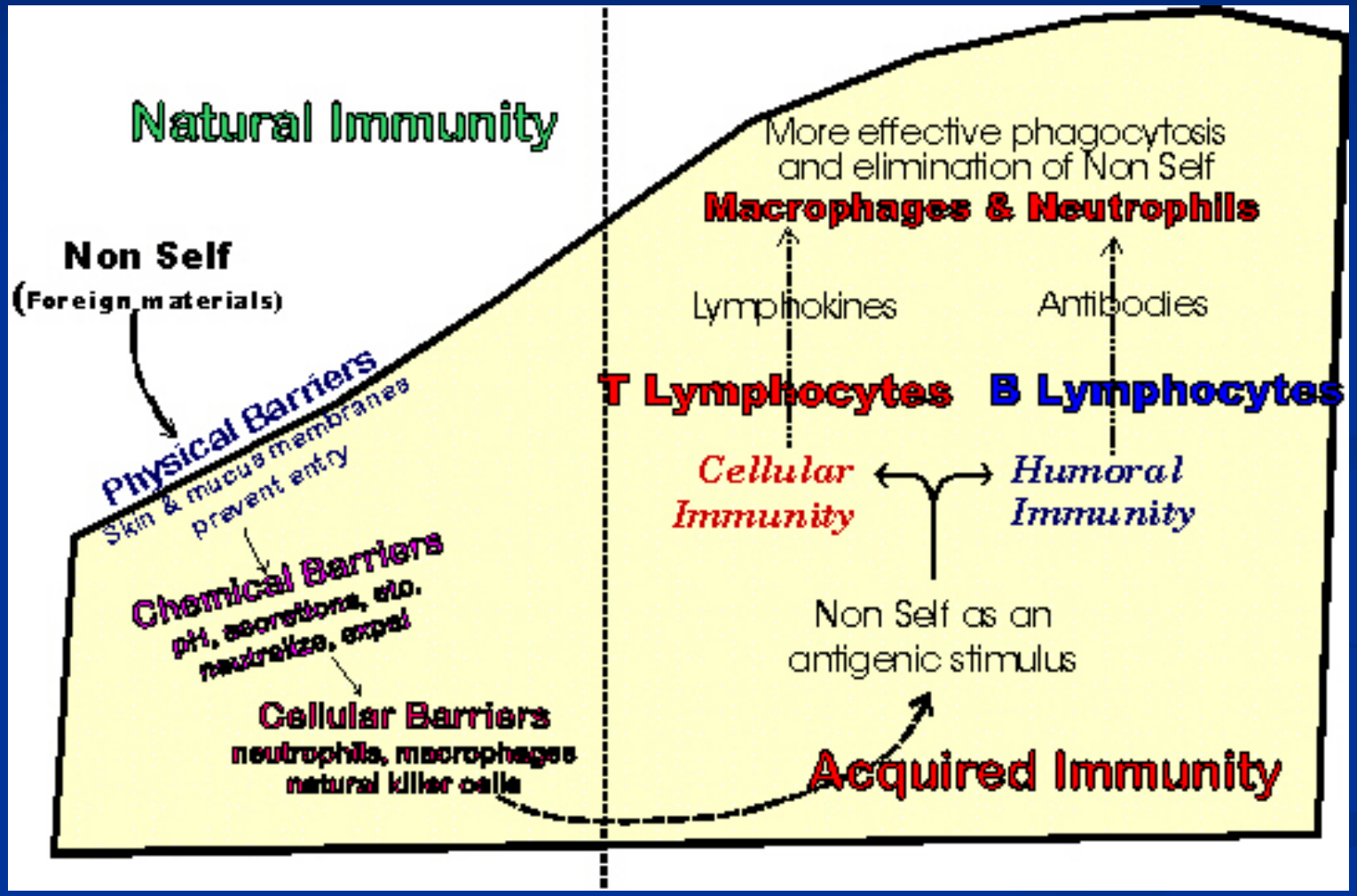
Systems Biology and the Future of Medicine

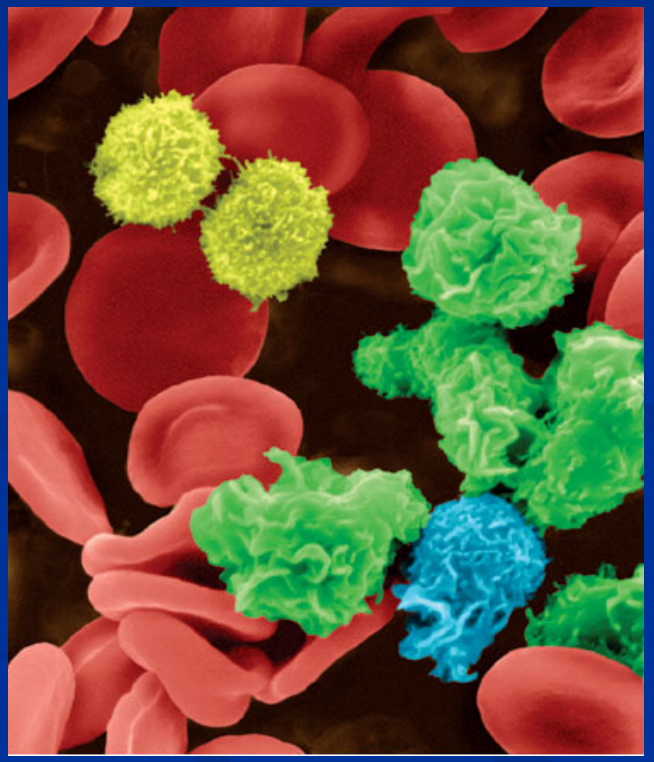
Sections to study

1. What is systems biology?
2. Biology as an informational science
3. The practice of systems biology
4. A systems approach to disease
5. Synthetic biology

1. What is systems biology?

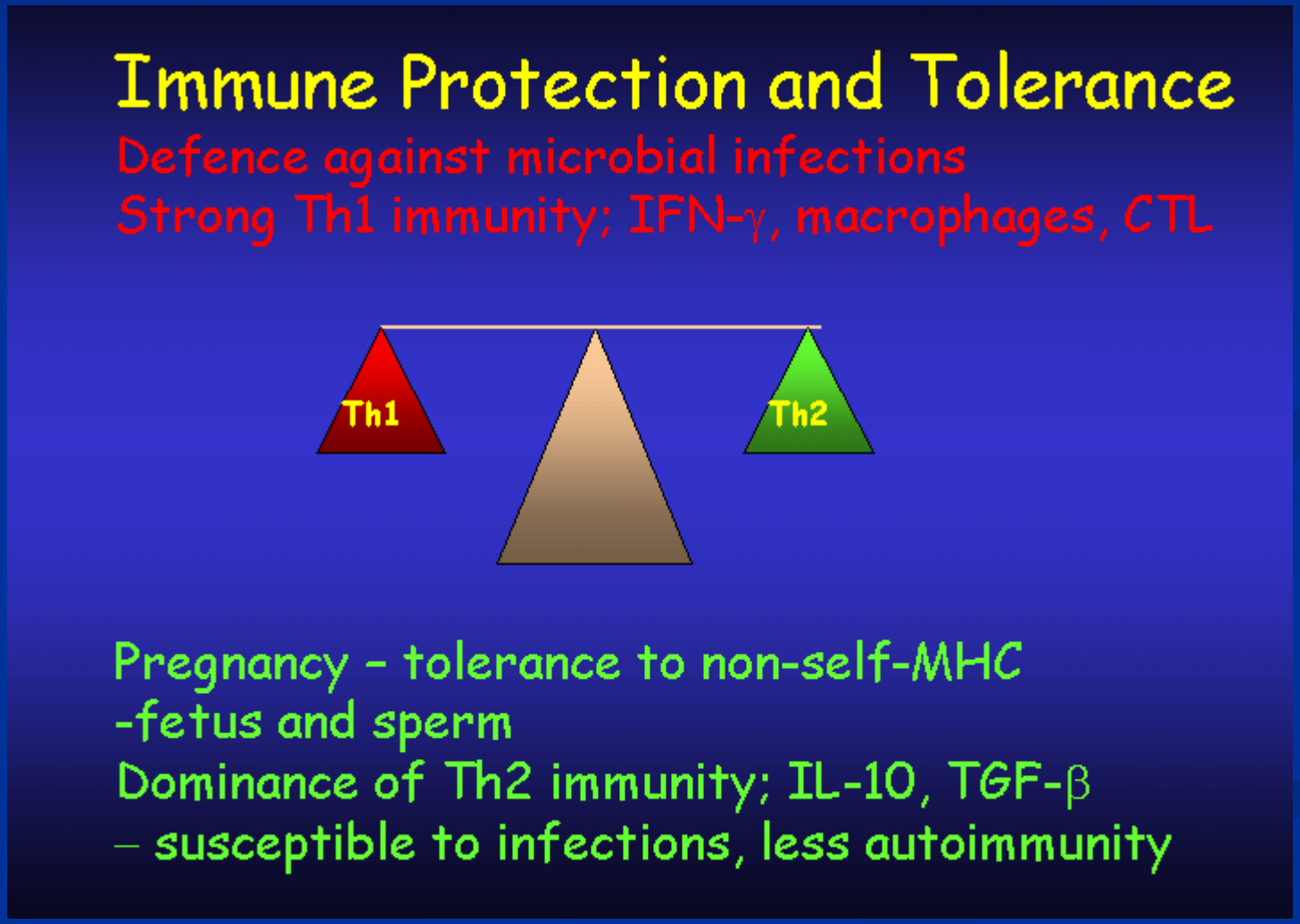
The human immune system





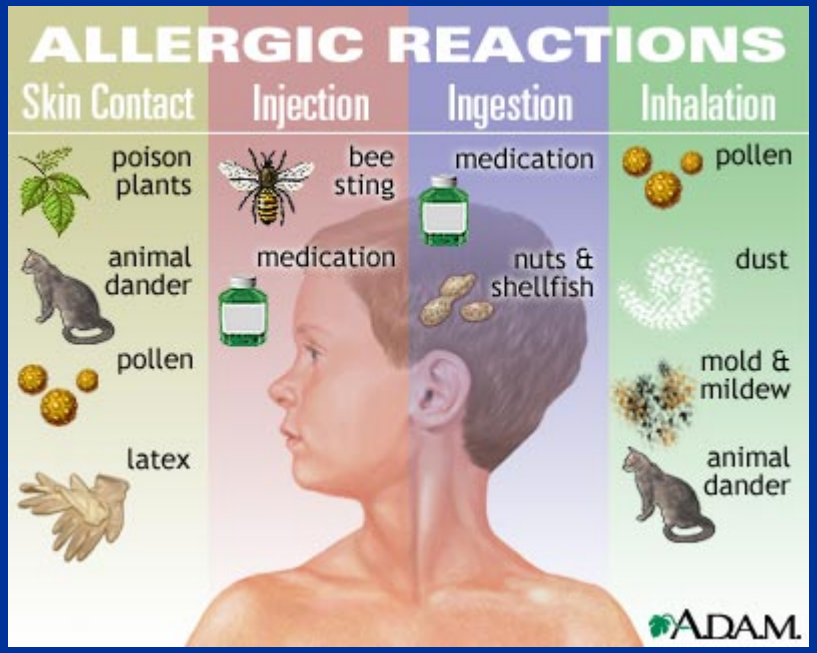
We have known the human immune system so well, but we are still unable to develop vaccines against every infectious disease. Why?

- **We do not understand the immune system's two most fundamental properties: immunity and tolerance.**
  - **Immunologists have studied the components of the immune system one gene and one protein at a time, but not them all together as a system.**



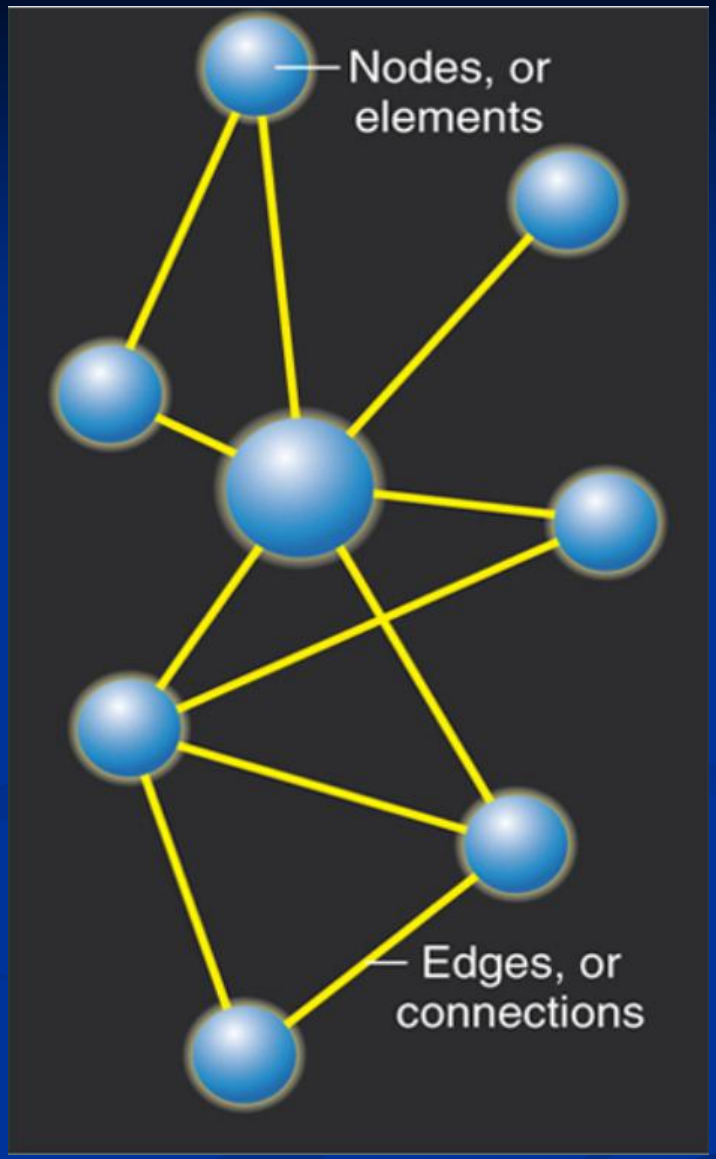
Overreacted immune responses:

- Allergy
  - Pollen, dust mites, peanuts, eggs.
- Overreacted inflammation
  - Inflammatory response in common cold makes it uncomfortable.
- Autoimmune disease: The body's normally protective immune system causes damage to its own tissues. The body responds as if normal tissues are infected or somehow abnormal.
  - Arthritis, diabetes.



***Biological system*: A collection of interacting elements that carry out a specific biological task.**

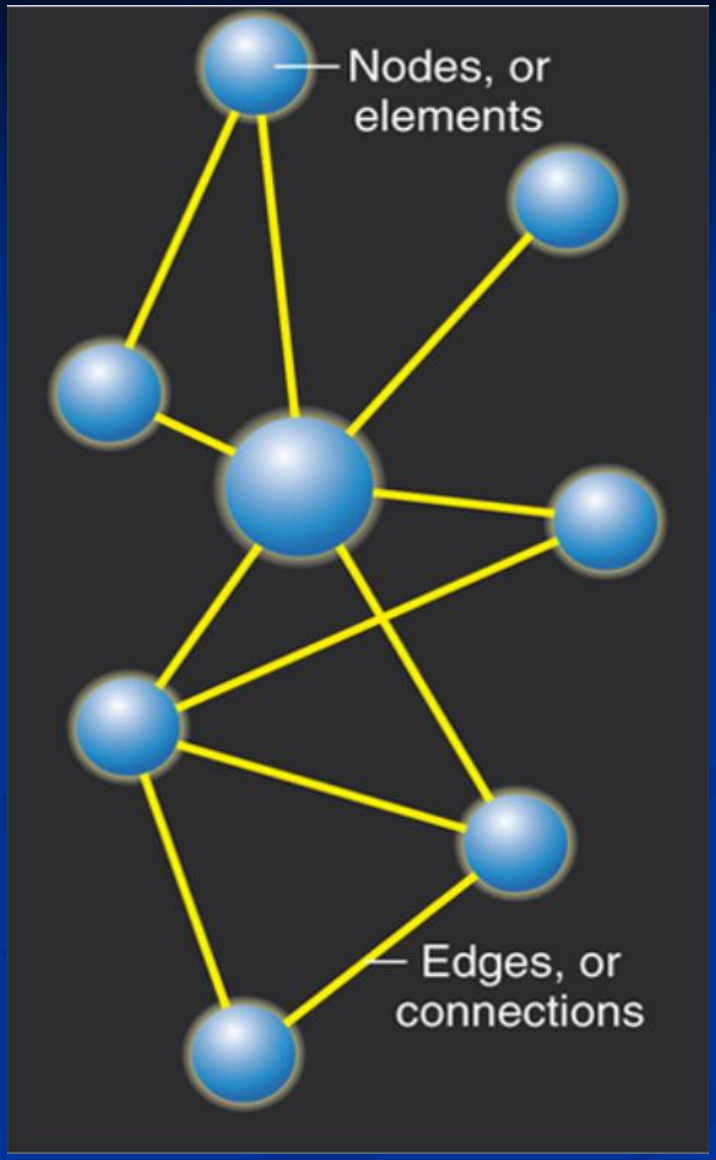
- **Interacting elements**
  - **Cells – such as immune system cells**
  - **Metabolites**
  - **Proteins**
  - **mRNAs**
  - **Control elements of genes**



A hypothetical biological network

Fig. 21.2

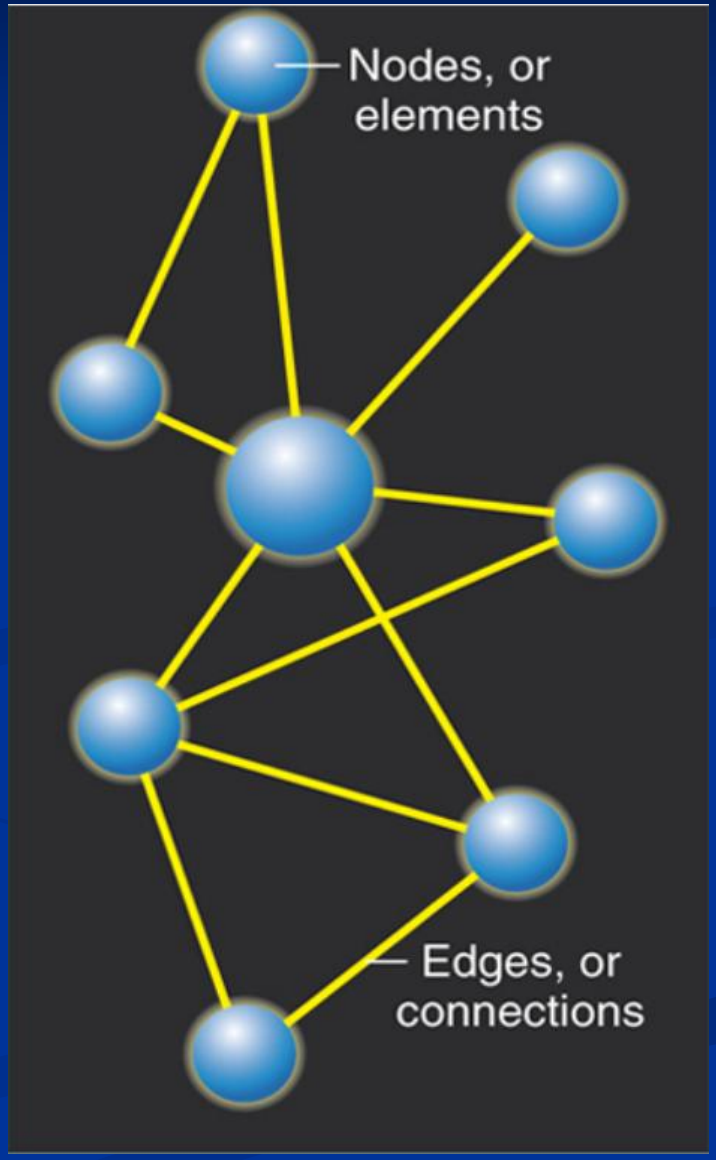
Systems biology **seeks to describe the multiple components of a biological system and analyze the complex interactions of those components within the system and with components of other systems.**



A hypothetical biological network

Four fundamental questions can be asked about a biological system:

- What are the *elements* of the system?
- What are the *physical associations* among the elements?
- How do perturbations affect the system and other systems connected to it?
- How do a system's elements, associations, and relation to changes in the biological context explain its functions?



A hypothetical biological network

2. Biology as an informational science

- **Biological information:**
  - **Digital information in the DNA of the genome**
  - **Environmental information impinges upon and modulates the core digital information.**
- **Their interaction mediates biological activity across three timescales:**
  - **Physiological response to the environment**
  - **Development**
  - **Evolution**
- **Analyzing biological systems in terms of the storage, transmission, and transformation of biological information is central to systems biology.**

Biological information is hierarchical

Ecosystems

↑

Populations

↑

Individual organisms

↑

Networks of cells, tissues, organs

↑

Cells

↑

Molecular machines, networks

↑

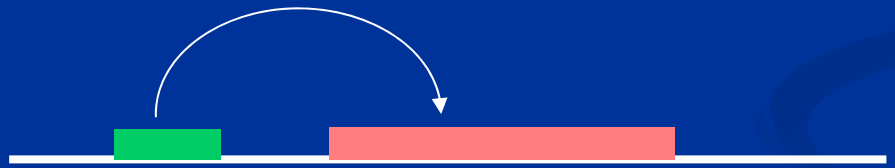
RNAs, proteins

↑

Digital information in genomes

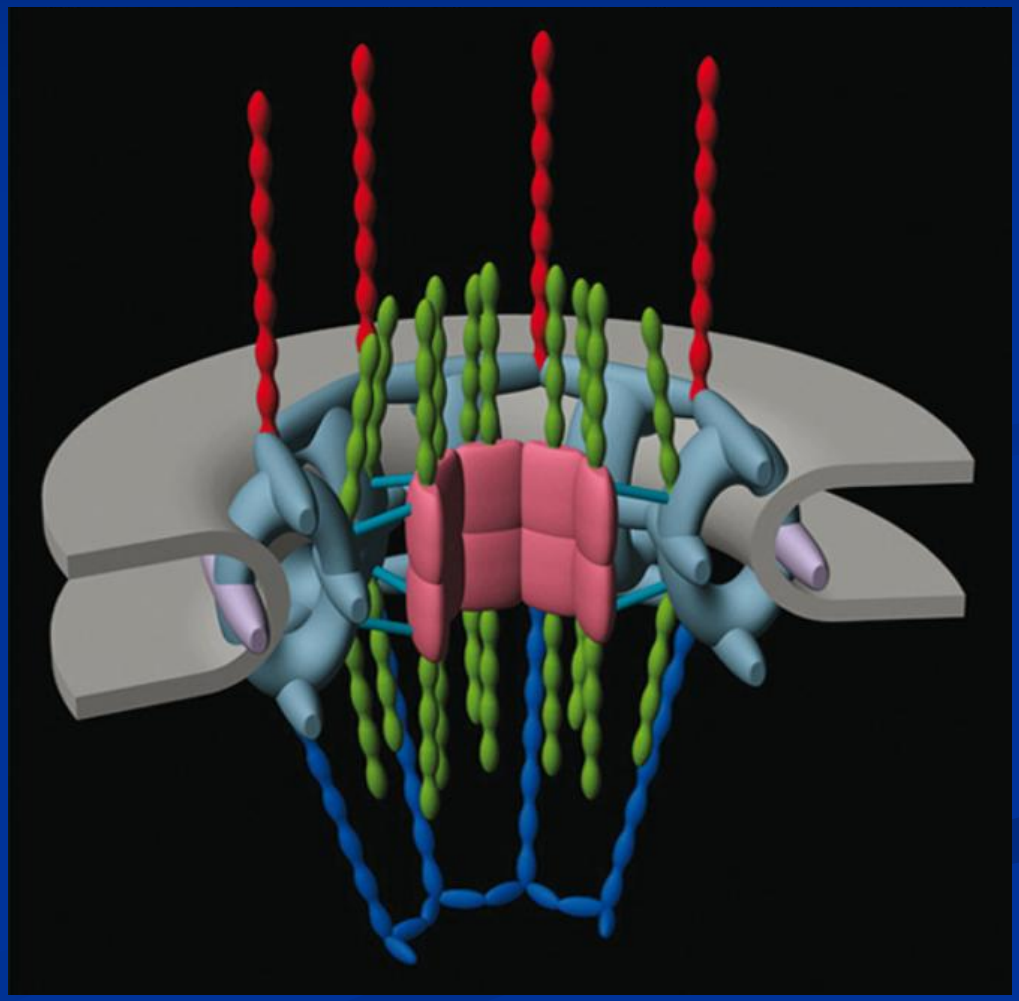
Digital genomic information

- Includes two types of sequences:
  - **Genes** that encode proteins and untranslated RNAs.
  - ***Cis*-control elements**: Short DNA sequences (6-15 bp) that constitute the control elements adjacent (or *cis*) to the genes.



Gene-encoded proteins interact with other proteins and macromolecules to form molecular machines and protein networks.

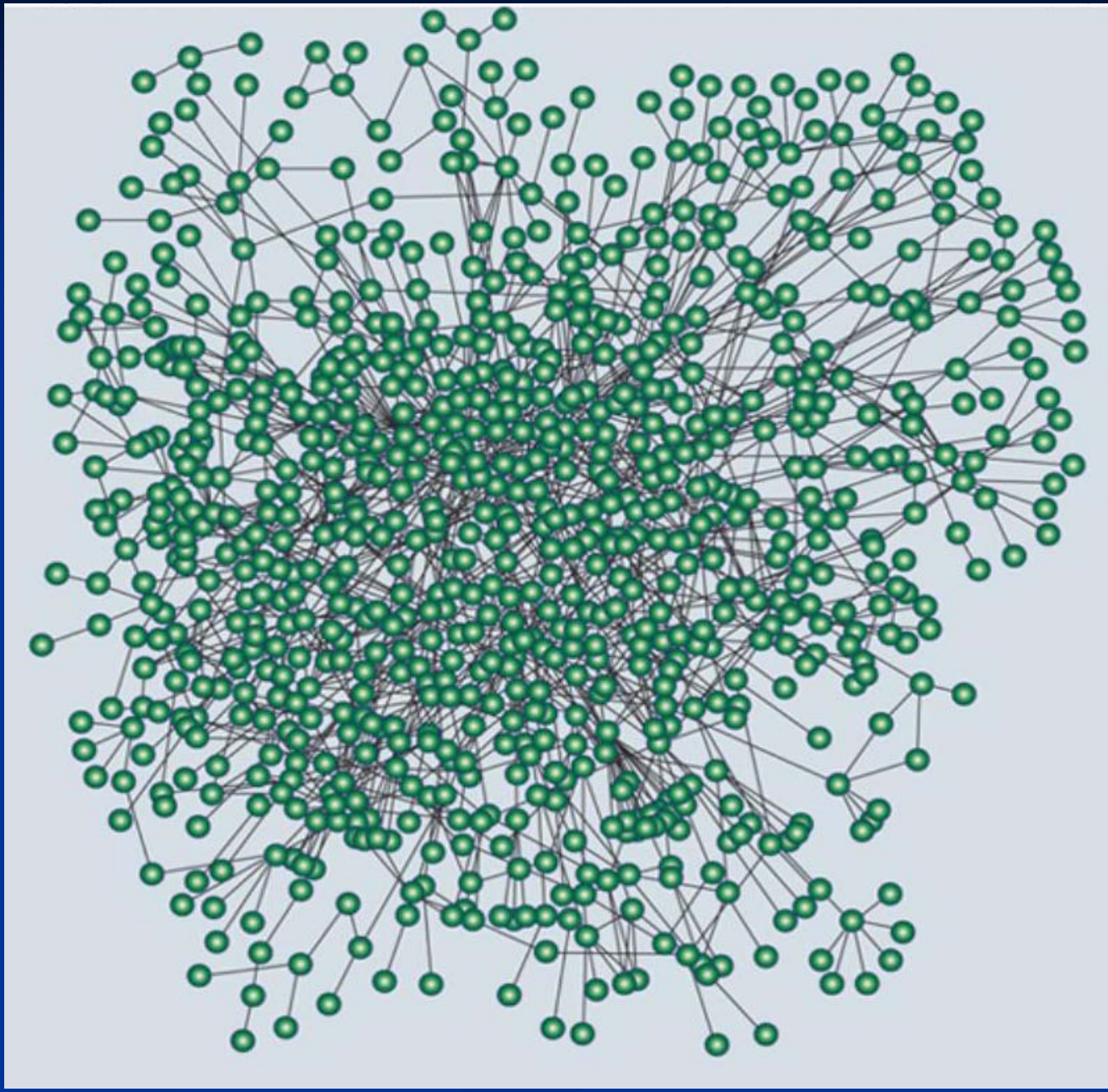
- **Molecular machines:** Nuclear pores, ribosomes, and spliceosomes.



The nuclear pore in yeast contains about 60 proteins

Fig. 21.3

- **Protein networks:** Sets of interacting proteins or metabolites that execute a particular biological function.

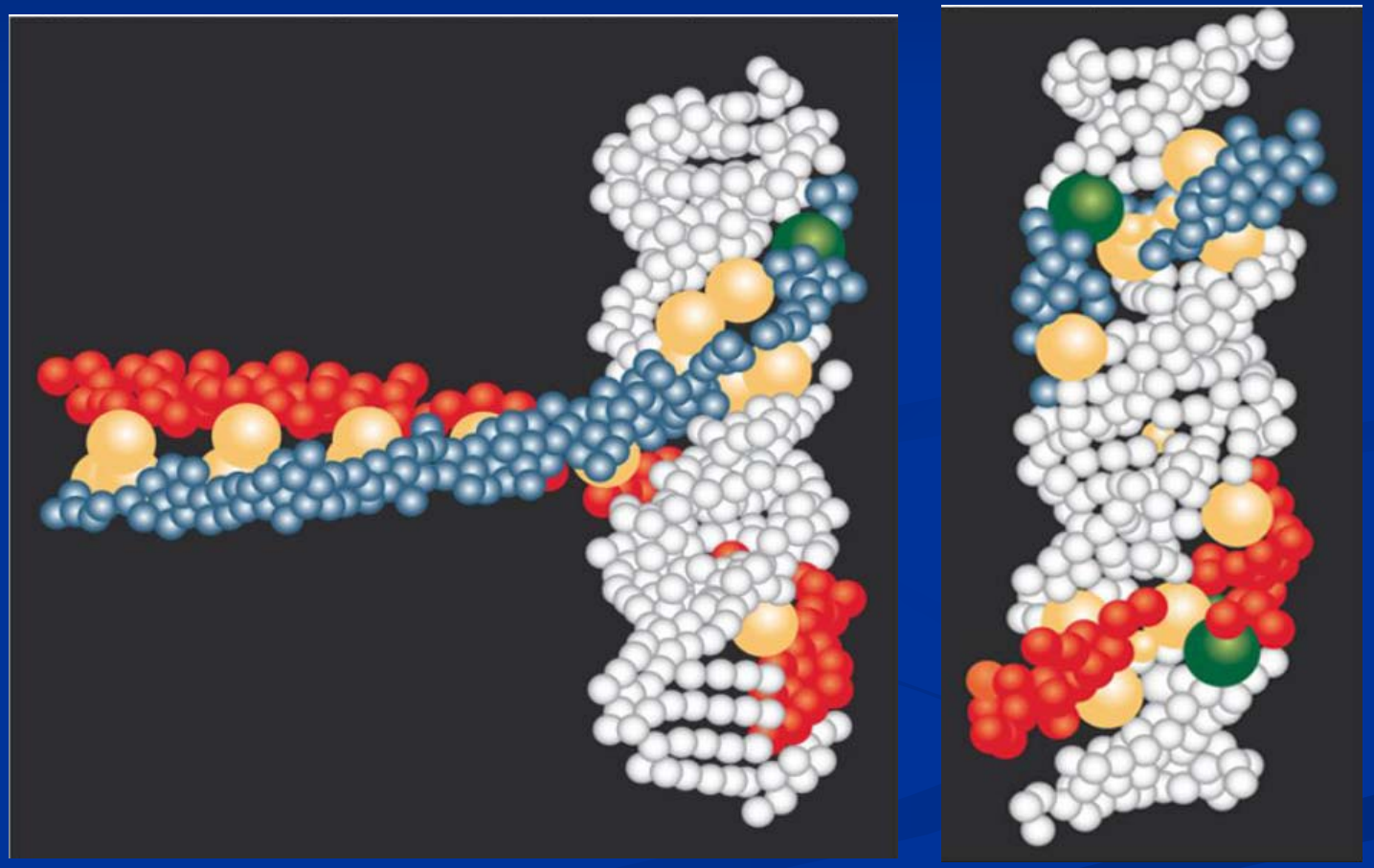


About 2500 proteins and 7000 linkages comprise this yeast protein network

Fig. 21.4

Transcription factors are key players in gene regulatory networks

Transcription factors interact with cognate *cis*-control elements



Transcription factors interact and bind to *cis*-control elements and interact with the basal transcription apparatus

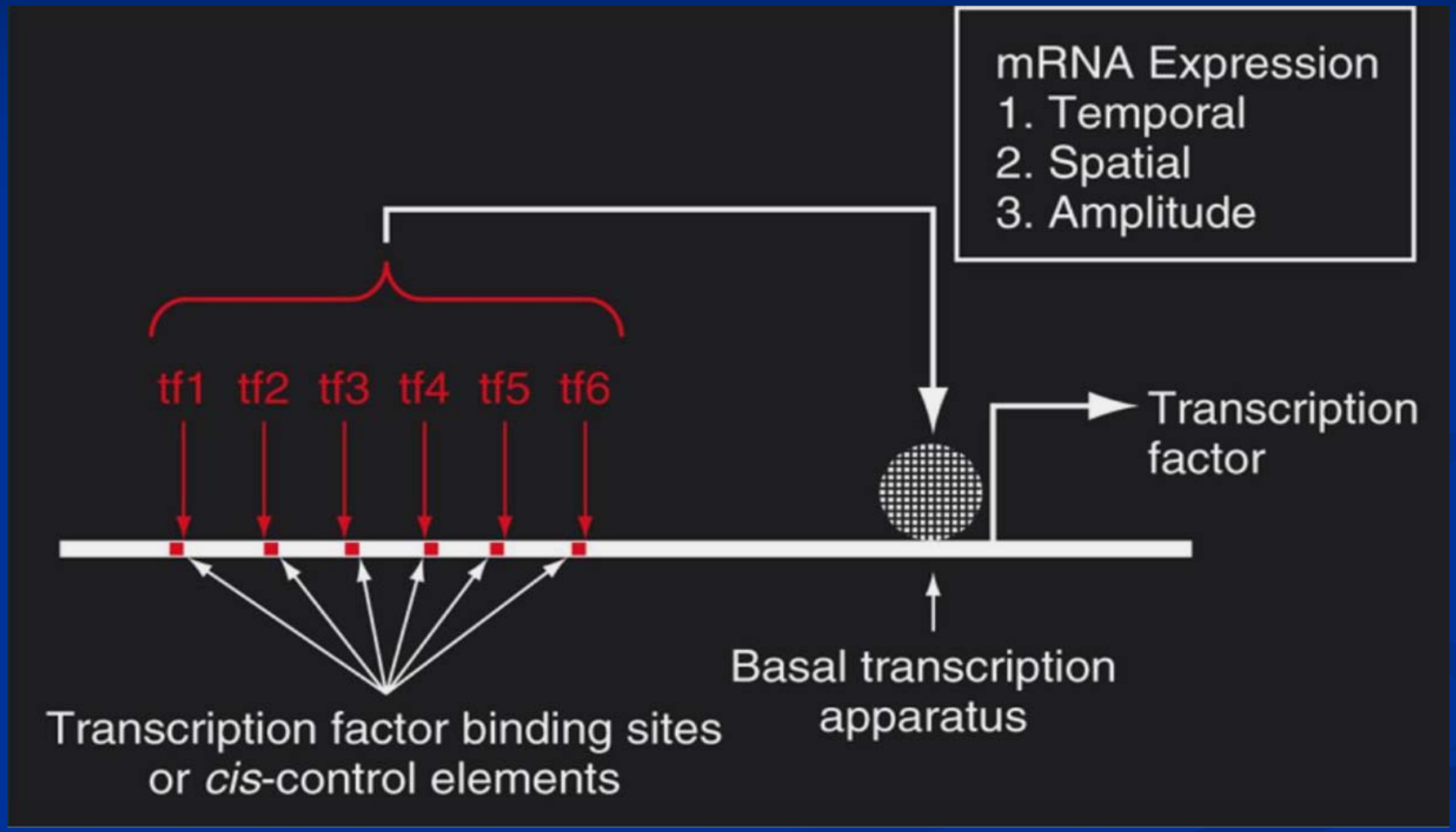


Fig. 21.5

Complex feed-forward and feed-back loops generated by gene regulatory networks

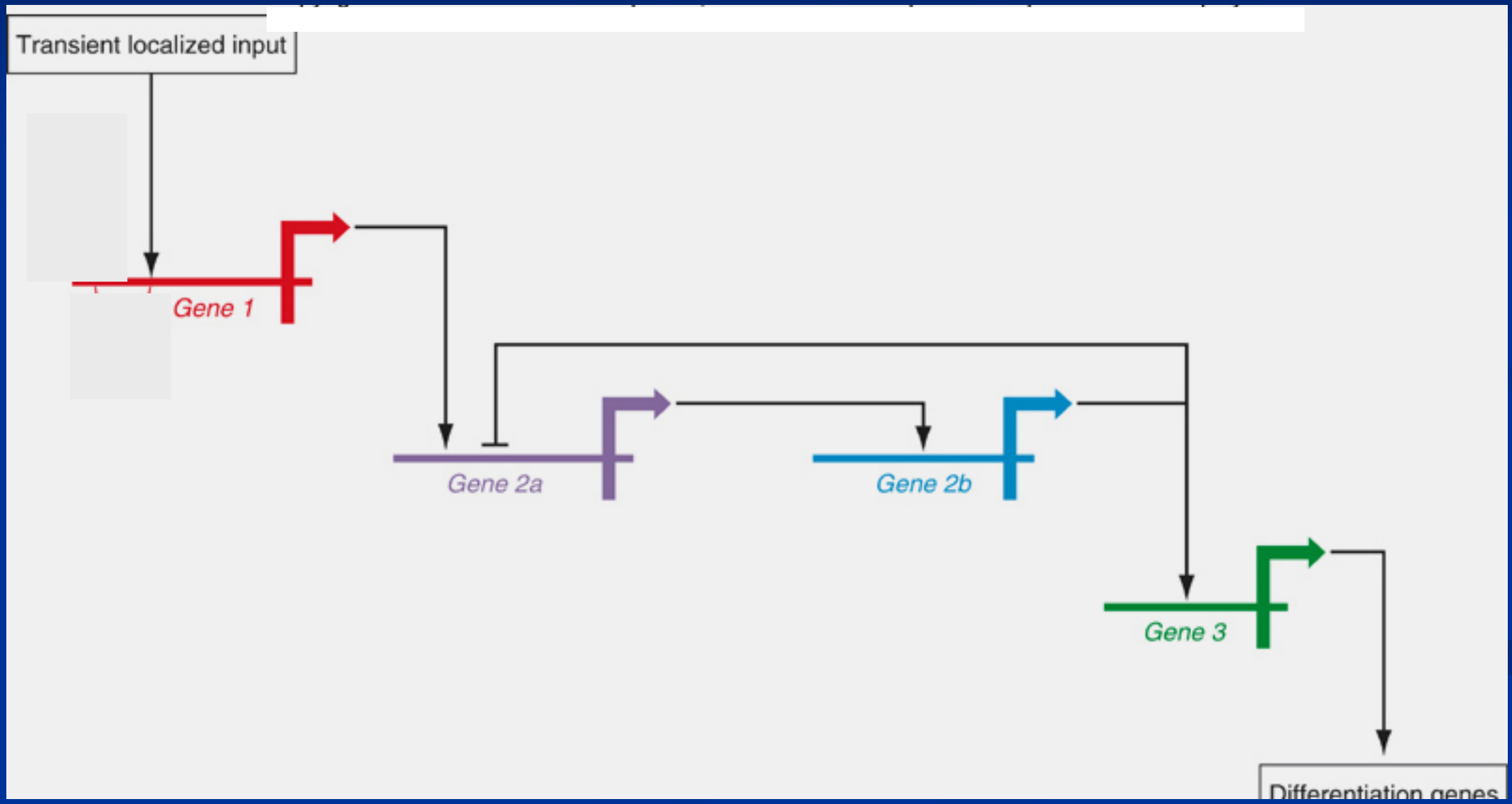


Fig. 21.6

A regulatory network controls gut development in sea urchin larvae

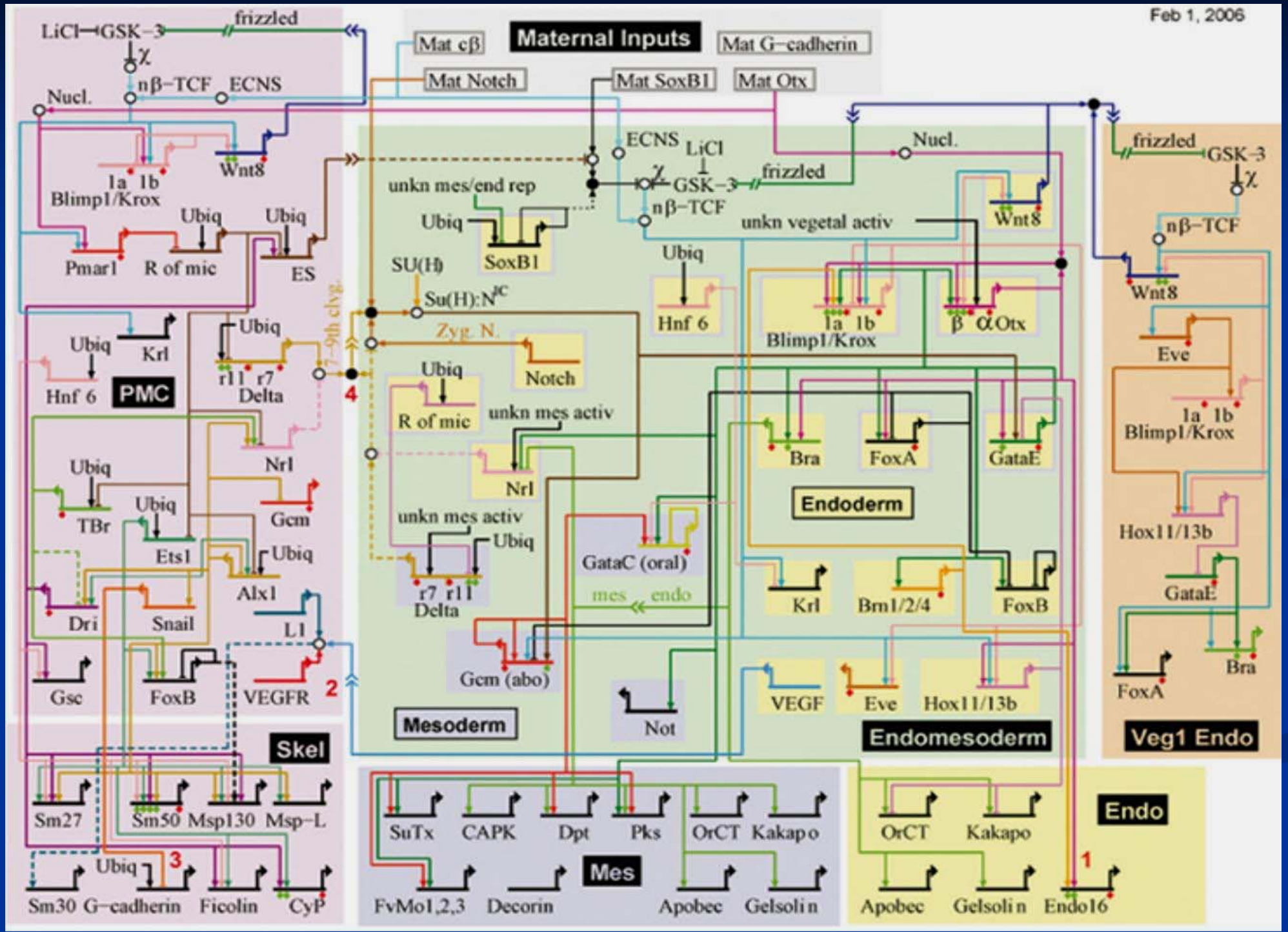
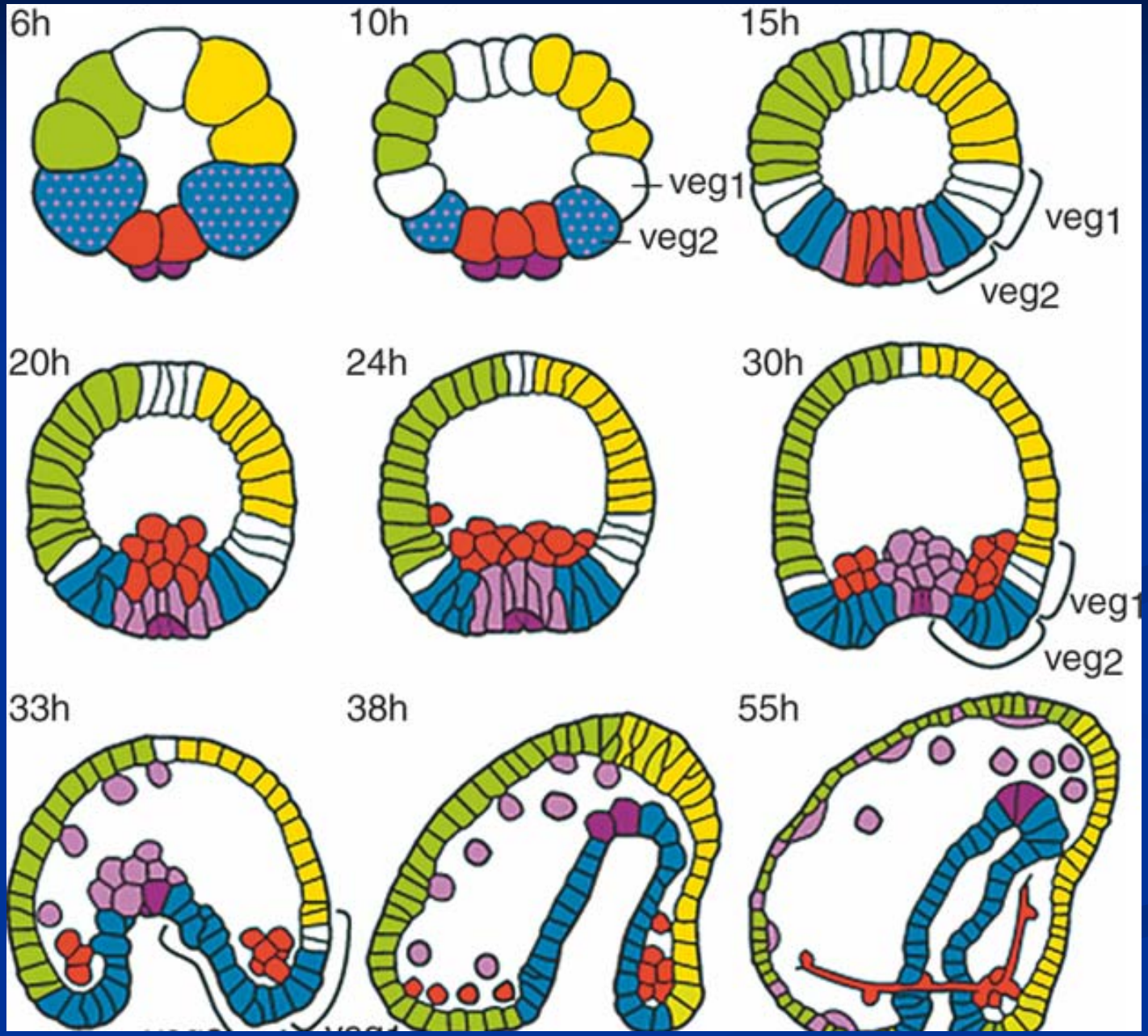


Fig. 21.7

Larval development of the sea urchin



Blue: endo-mesodermal or gut

blastula

gastrula

differentiated larva

Fig. 21.8

Summary

- **Gene regulatory networks integrate and modulate inputs of biological information and transmit the transformed information to protein networks.**
- **Protein networks then use this information to execute specific functions in the organism.**

3. The practice of systems biology

- **The Human Genome Project has provided global data-gathering tools and genomic information on a scale never before available.**
- **These tools and data are central to the current practice of molecular systems biology.**

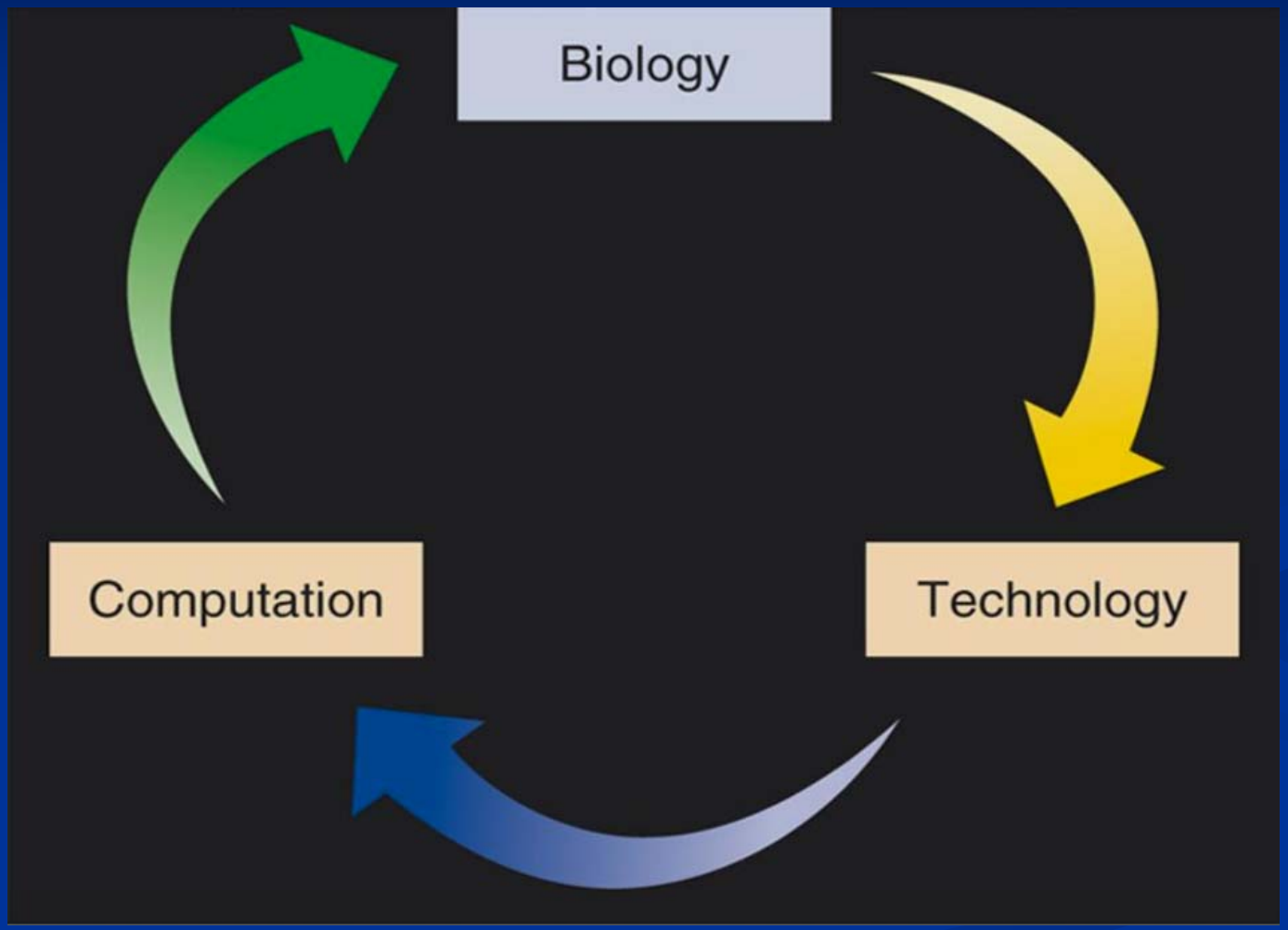
- ***Genome*: The complete set of genetic information in a particular cell or organism.**
  - ***Genomics*: The study of whole genomes.**
- ***Transcriptome*: The complete set of mRNAs in a particular cell or organism.**
- ***Proteome*: The complete set of proteins in a particular cell or organism.**
  - ***Proteomics*: The global analysis of most (or ideally, all) the proteins in a particular cell or organism.**

Foundations for current systems biology

- Complete genome sequences of humans and model organisms
  - List of all genes and predictions of proteins encoded by the genes
- High-throughput platforms for genomics and proteomics
  - Acquisition of global or comprehensive data sets of differing types of biological information (all genes, all mRNAs, all proteins, …)
- Powerful computational tools
  - To acquire, store, analyze, integrate, display, and model information

- **Studies of simple model organisms**
  - **Compile global data sets from experimental manipulations of simple organisms. Provide models for more complex organisms.**
- **Comparative genomics (比较基因组学)**
  - **Determine the logic of life for individual organisms and how it has changed in an evolutionary sense.**

Systems biology is still under development



Several steps in practicing systems biology

1. Scan the biological literature and database for all that is known about the system of interest.
2. Develop a preliminary model about how the system functions.
3. Formulate a hypothesis-driven query about the model.
   - Answer query through genetic or environmental perturbations of the system.
   - Collect comprehensive data sets of different levels of biological information.

4. Integrate different types of data either graphically or mathematically and compare against initial model.
5. Perform iterative perturbations to test the new hypotheses. Repeat steps 3-5 if necessary.
6. Evaluate whether the refined model enables biologists to predict the behavior of the system.

4. A systems approach to disease

- **Altered cellular networks caused by genetic or environmental perturbations can lead to disease.**

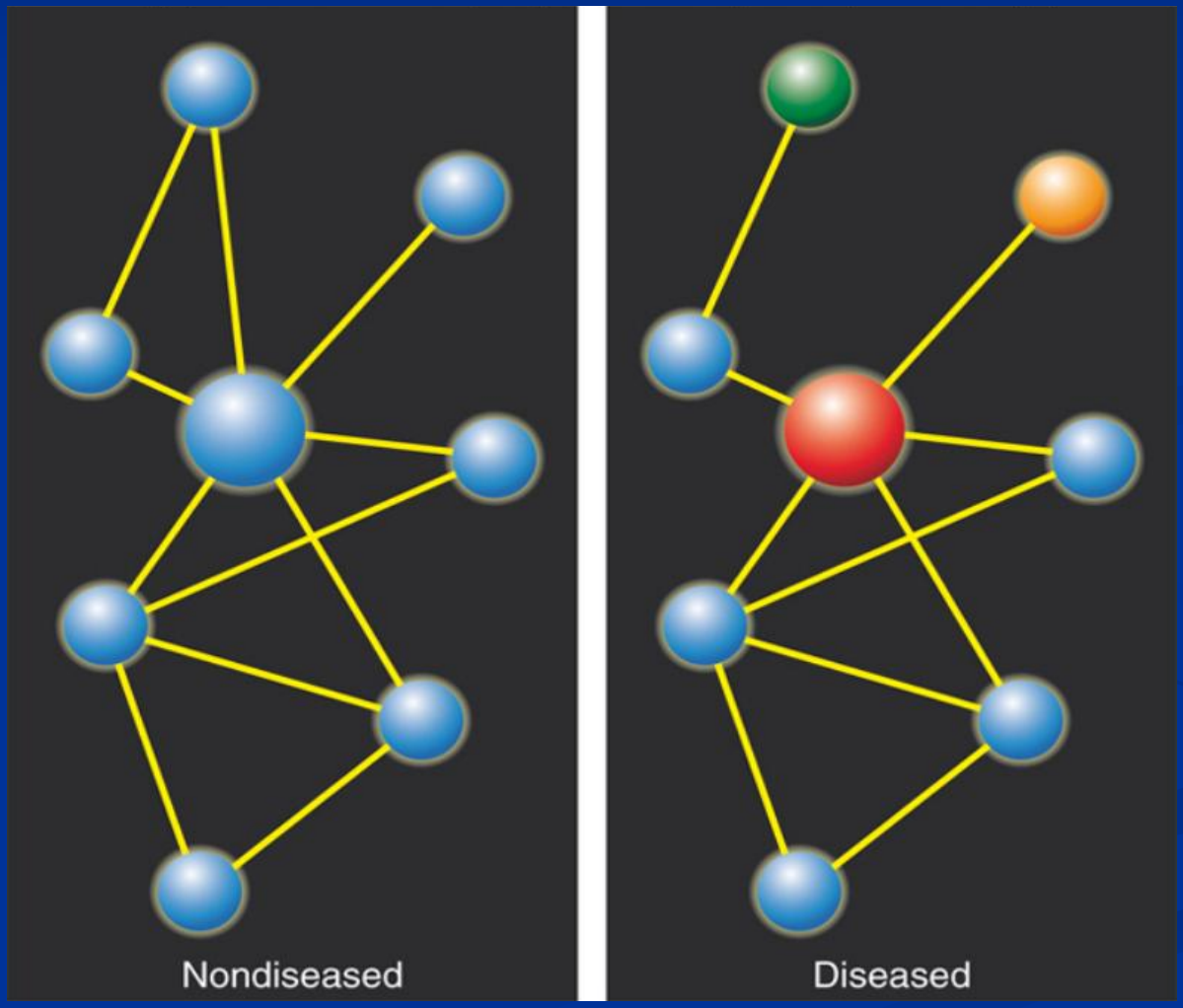


Fig. 21.12

This view opens door for new approaches to diagnostic, therapeutics, and prevention of disease.

- **Identification of biomarkers is a first step.**
  - **Biomarkers: Molecules that are present under specific conditions or when a disease is present.**
  - **Prostate-specific antigen (PSA) as a biomarker. Its level in the blood increases in patients with prostate cancer.**
- **Knowledge of protein interactions can identify drug targets.**
  - **Protein interactions with complimentary drug will kill cell (cancer) or alter function of the network towards normality.**
  - **Will lead to integration of diagnostics and therapy.**

Advances in technology are needed for new medicine

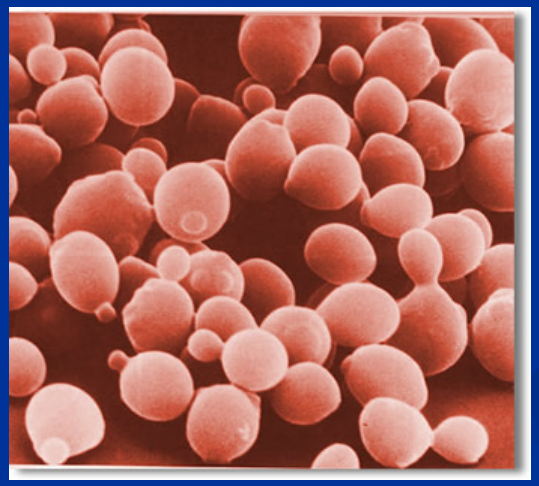
- **Nanotechnology – will sequence genomes rapidly for under $1000.**
- **Molecular imaging improvements – non-invasive visualization of drug activity and function in model organisms and humans.**
- **Microfluidics and nanotechnology – measure, identify, and inexpensively quantify thousands of proteins from a small drop of blood or assess information content of individual cells.**

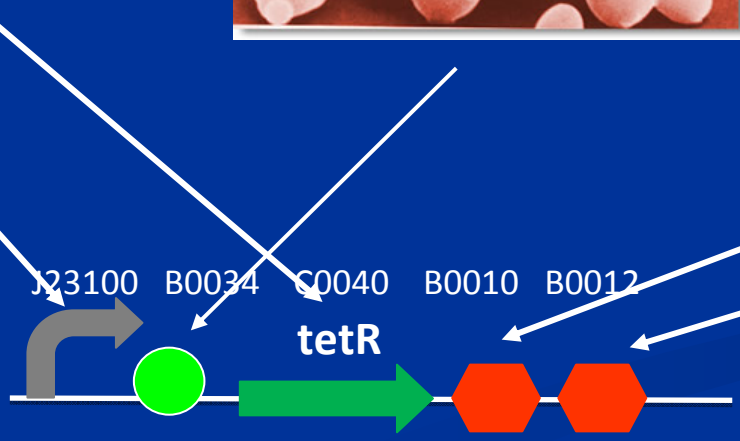
The systems approach leads to predictive, preventive, and personalized medicine

- **Prediction**
  - The genome of a person can be sequenced.
  - 1000-2000 proteins in the blood can be analyzed.
- **Preventive medicine**
  - Development of drugs that can keep networks from being perturbed.
- **Personalized medicine**
  - Apply the power of predictive and preventive medicine to individual needs.

5. Synthetic biology

Synthetic biology (合成生物学): The science of designing and constructing new biological functions and systems not found in nature.



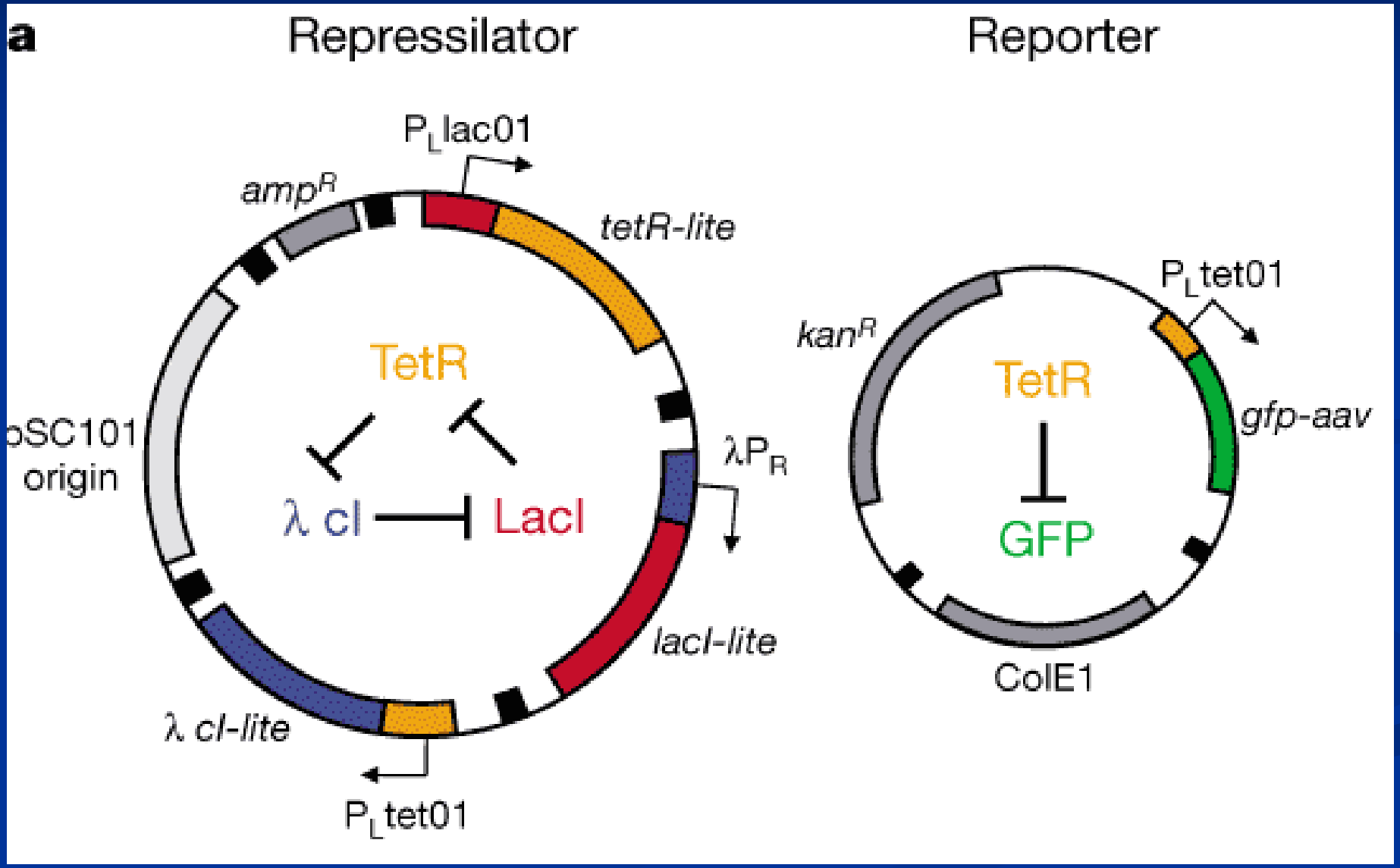


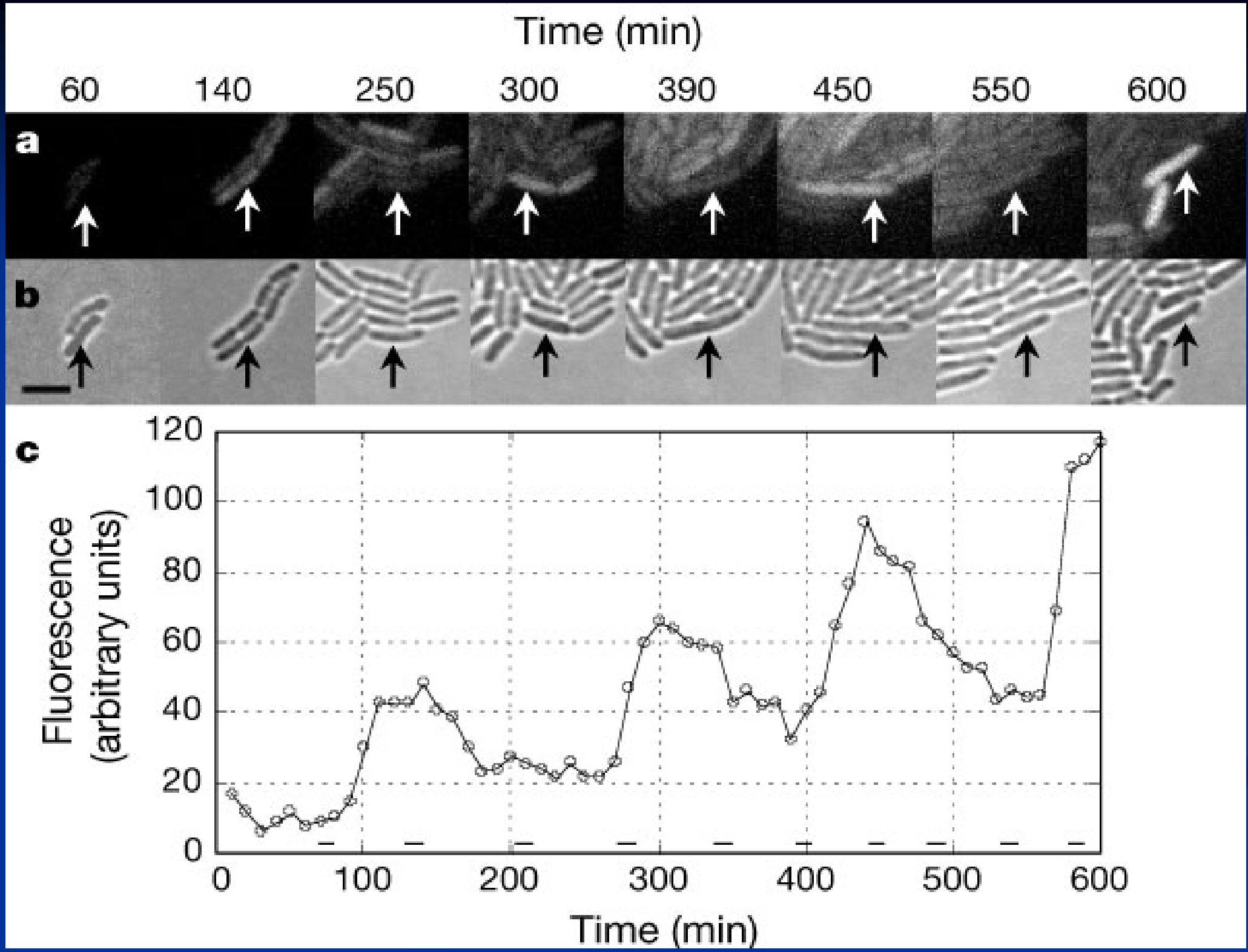
1. **The design and construction of new biological parts, devices, and systems.**
2. **The re-design of existing, natural biological systems for useful purposes.**

- **Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403: 335–338**
- **Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli. Nature* 403: 339–42**

A synthetic oscillatory network of transcriptional regulators

(Michael B. Elowitz and Stanislas Leibler, 2000, *Nature*. 403: 335–338)



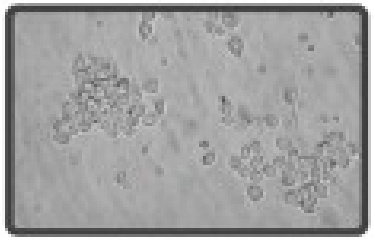

Time (min)
60
140
250
300
390
450
550
600
a
b
c
Fluorescence (arbitrary units)
120
100
80
60
40
20
0
0
100
200
300
400
500
600
Time (min)

Building complex systems in synthetic biology

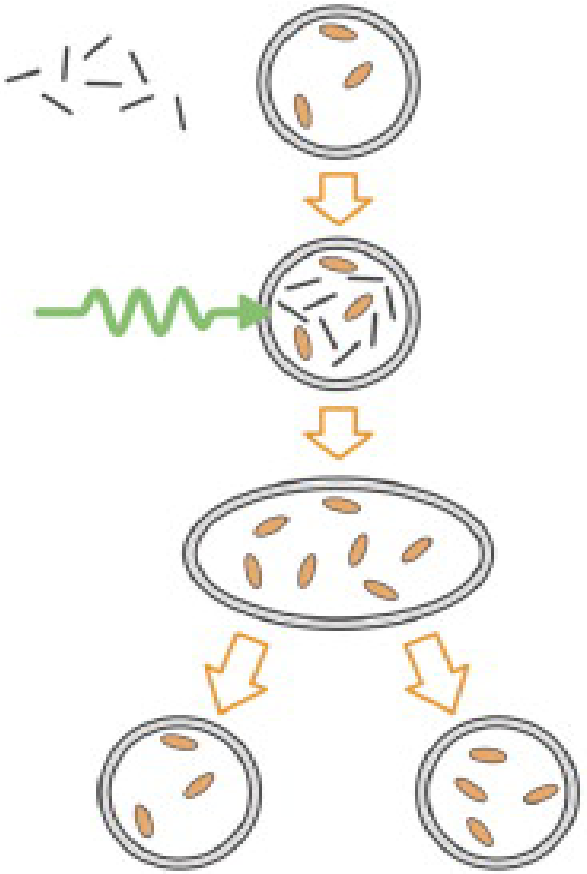
List of genetic building-blocks

Assembled gene

Cells expressing non-natural functions



Biomolecular-design Approach



Encapsulated catalytic molecules and a "fuel"

Acquisition of fuel & addition of light energy

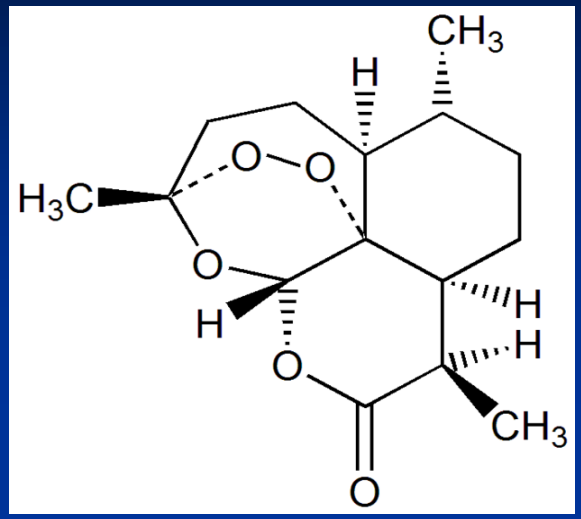
Fuels used to produce more building materials

Reproduction to give two daughter protocells

Protocell-design Approach

Production of *artemisinic acid* (青蒿酸) in yeast

青蒿素

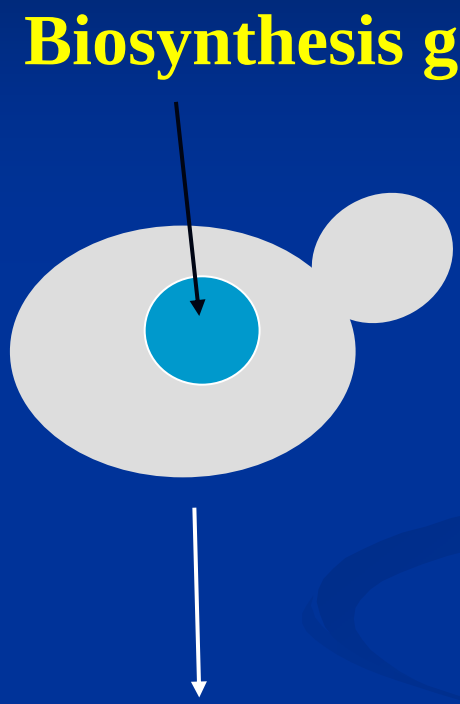


Artemisia annua (黄花蒿)

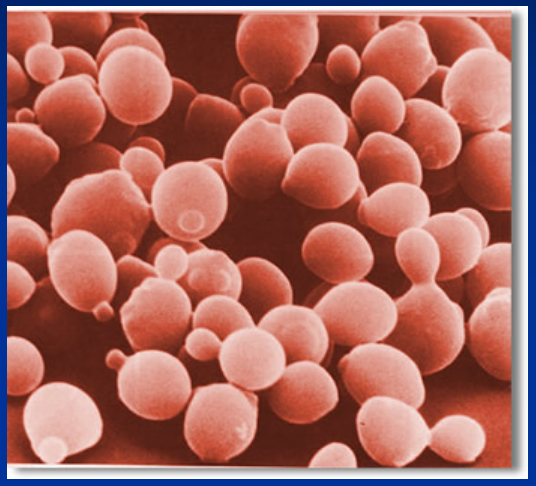
屠呦呦

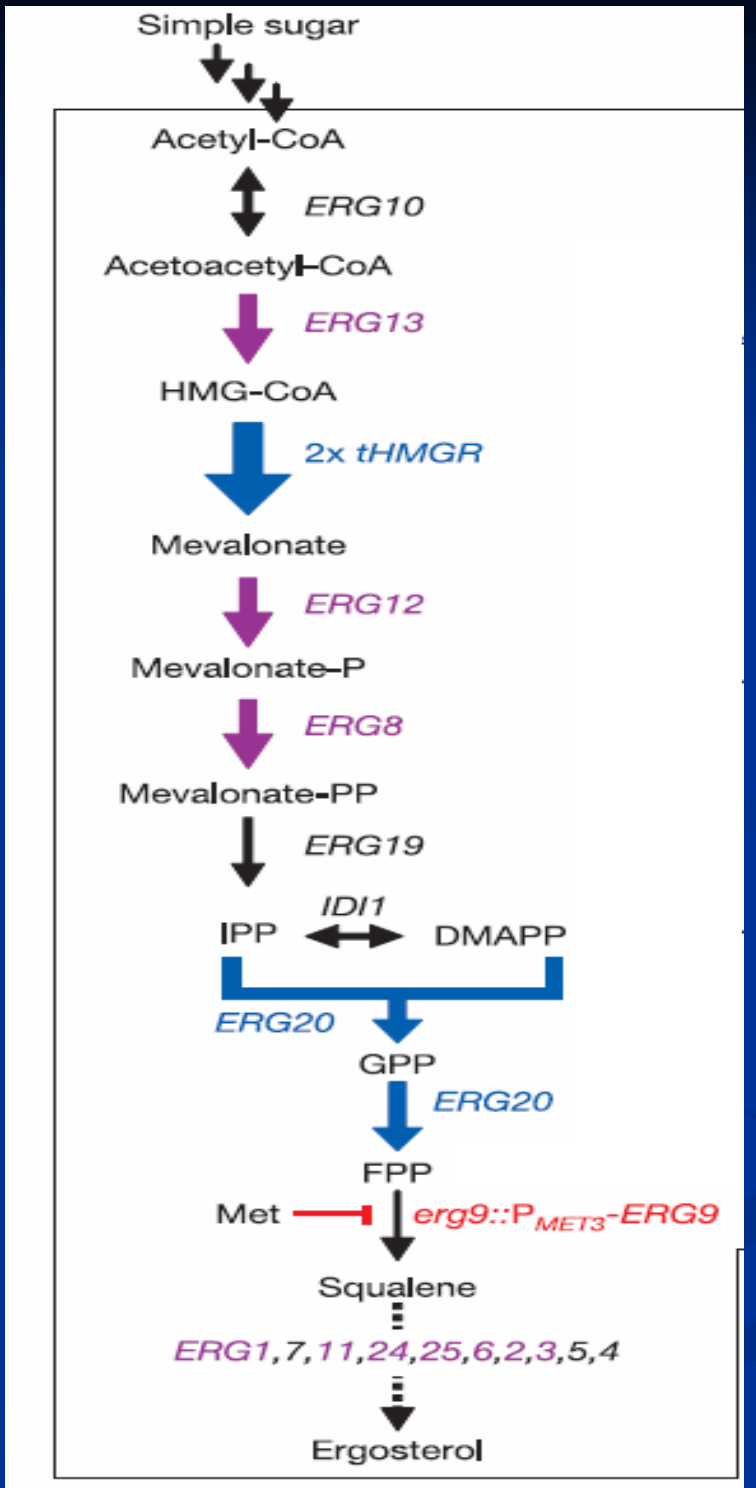
Production of artemisinic acid (青蒿酸) in yeast

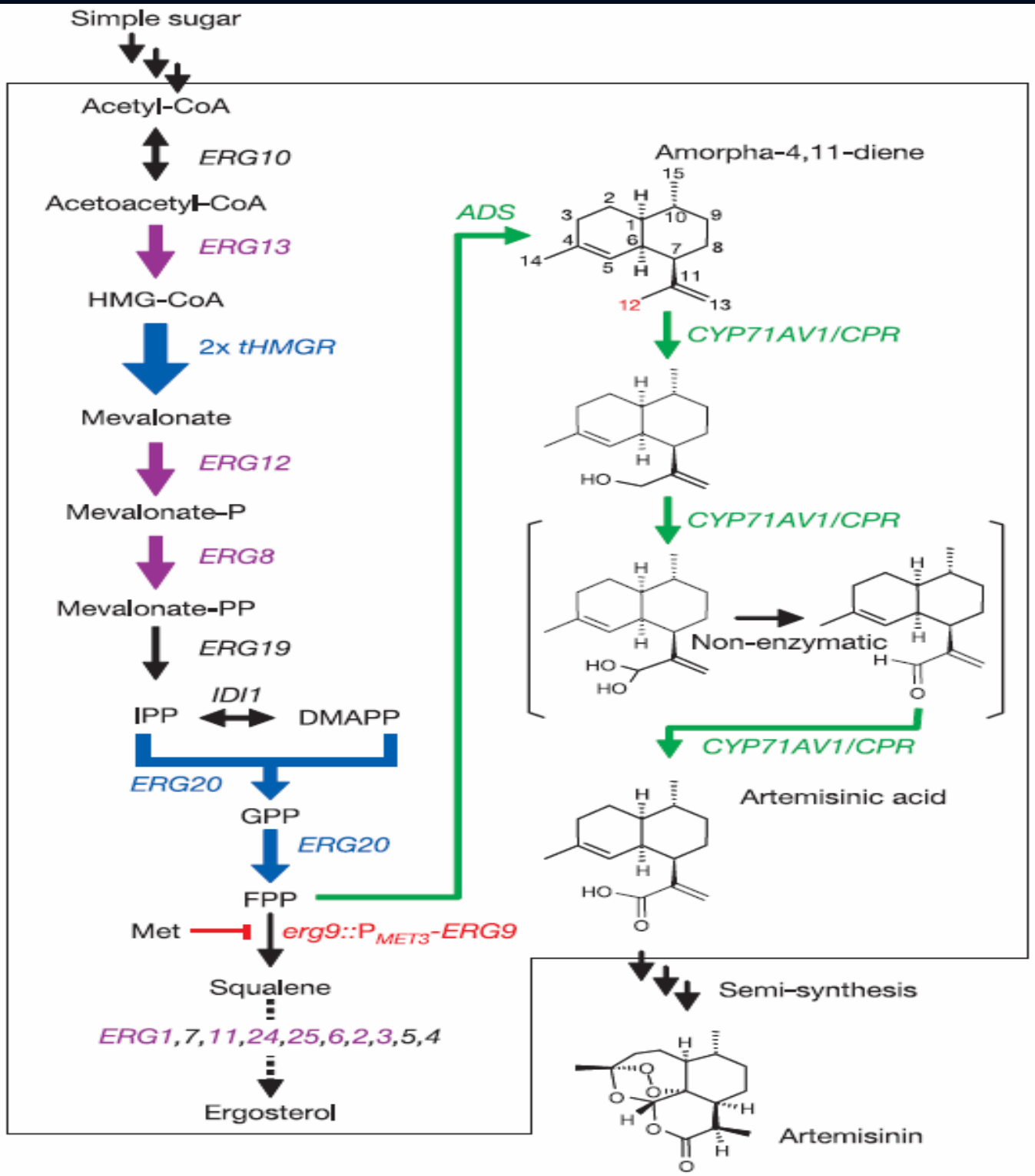
Jay D. Keasling
UC Berkeley



Artemisinic acid → → → **Artemisinin**




Simple sugar
Acetyl-CoA
ERG10
Acetoacetyl-CoA
ERG13
HMG-CoA
2x tHMGR
Mevalonate
ERG12
Mevalonate-P
ERG8
Mevalonate-PP
ERG19
IDI1
IPP
DMAPP
ERG20
GPP
ERG20
FPP
Met
erg9::P MET3 -ERG9
Squalene
ERG1,7,11,24,25,6,2,3,5,4
Ergosterol



1. Engineering the farnesyl pyrophosphate (**FPP**) biosynthetic pathway to increase FPP production and decrease its use for sterols.
2. Introducing amorphadiene synthase gene (**ADS**) from *A. annua.*
3. Introducing a novel cytochrome **P450** gene from *A. annua.*

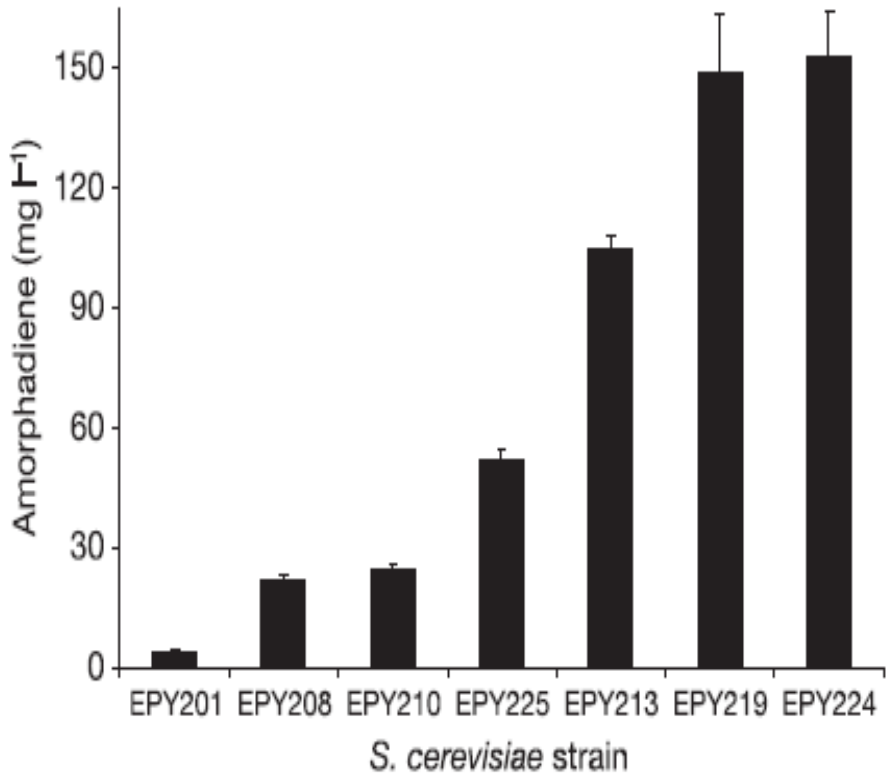


Figure 2 | Production of amorphadiene by *S. cerevisiae* strains. The various *S. cerevisiae* strains are described in the text. Cultures were sampled after 144 h of growth, and amorphadiene levels were quantified. Data, shown as total production, are mean ± s.d. ($n = 3$).

EPY201： ADS

EPY208： tHMGR

EPY210： *UPC2-1*, tHMGR

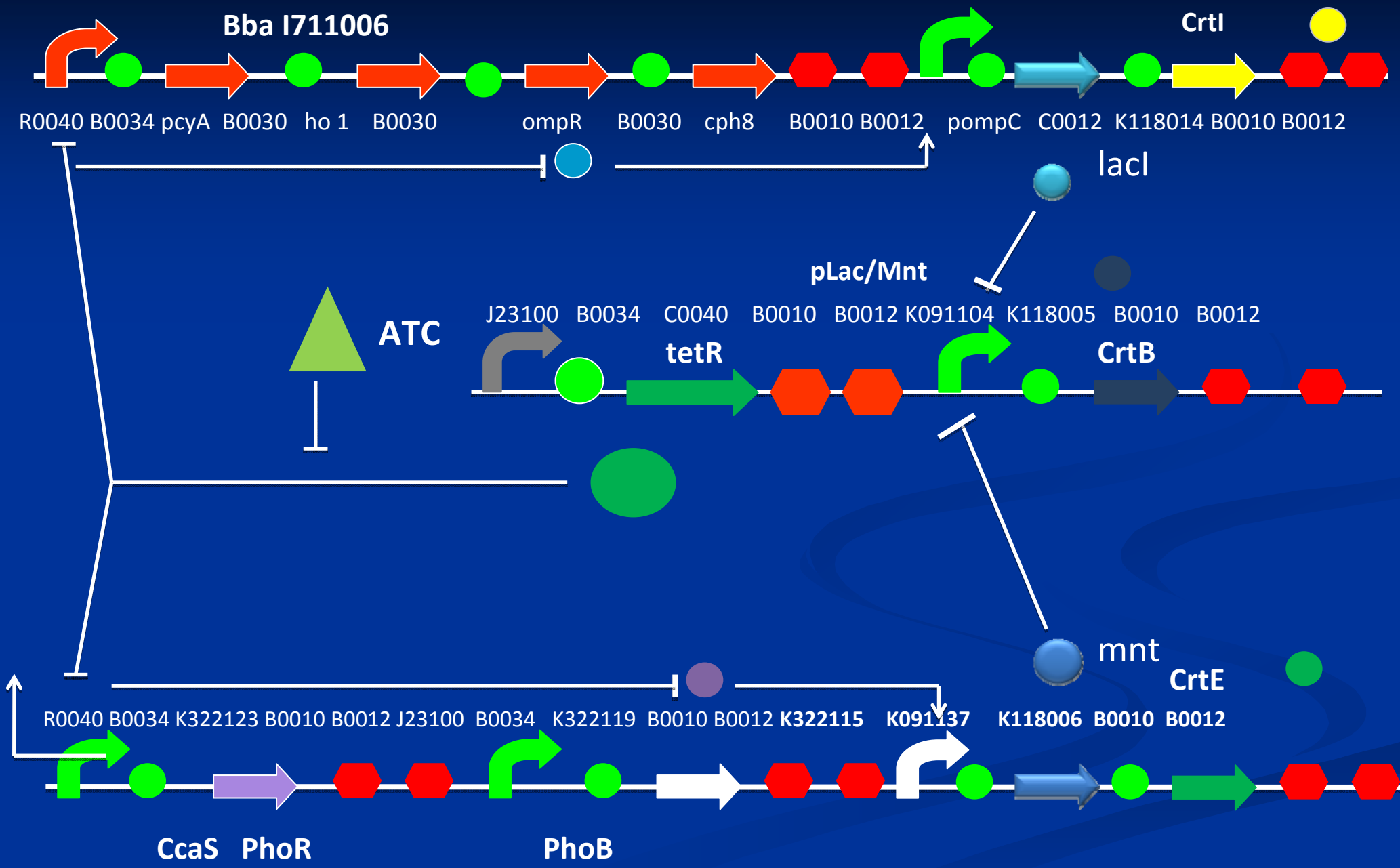
EPY225： ERG9

EPY213： ERG9，*UPC2-1*，tHMGR

EPY219： ERG9，*UPC2-1*, tHMGR×2

EPY224： ERG9，*UPC2-1*, tHMGR×2，EMG20

A nearly 500-fold increase!


Bba I711006
CrtI
R0040 B0034 pcyA B0030 ho 1 B0030 ompR B0030 cph8 B0010 B0012 pompC C0012 K118014 B0010 B0012
lacI
pLac/Mnt
ATC
J23100 B0034 C0040 B0010 B0012 K091104 K118005 B0010 B0012
tetR
CrtB
mnt
CrtE
R0040 B0034 K322123 B0010 B0012 J23100 B0034 K322119 B0010 B0012 K322115 K091137 K118006 B0010 B0012
CcaS PhoR
PhoB

The International Genetically Engineered Machine (iGEM) competition

Log in

Synthetic Biology
based on standard parts

Home | iGEM Competition | Registry | Community | Impact | Press | Sponsors | About

High School Division

iGEM 2013 HS

Happening Now: Registration Closed

- project descriptions
- iGEM forums
- late registration closed
- 3A Assembly and Transformation Efficiency Kits

more...

2013hs.igem.org

Collegiate Division

iGEM 2013

Happening Now: Registration now open!

- partner offer: Mathworks
- iGEM 2013 registration open!
- project ideas
- Registry Help pages
- Start a Team

more...

2013.igem.org

Entrepreneurship Division

iGEME

Happening Now: Stay Tuned

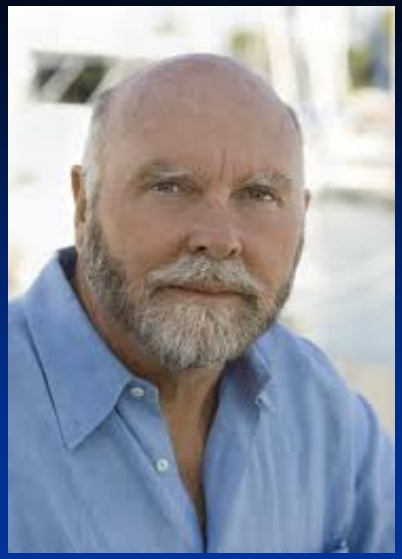
Stay Tuned for More Information on iGEM E

- **The competition provides standard biological parts (BioBricks), and asking student teams to design and build genetic machines with them. Student teams can also submit their own BioBricks.**
- **Successful projects produce cells that exhibit new and unusual properties by engineering sets of multiple genes together with mechanisms to regulate their expression.**

iGEM competition results

	Winner	Finalist	Finalist	Finalist	Finalist	Finalist	Complete Results
2011	Washington	Imperial (2nd)	ZJU China (3rd)	MIT (4th)			iGEM 2011
2010	Slovenia	Peking (2nd)	BCCS Bristol (3rd)	Cambridge	Imperial	TU Delft	iGEM 2010
2009	Cambridge	Heidelberg (2nd)	Valencia (3rd)	Freiburg Bioware	Groningen	Imperial	iGEM 2009
2008	Slovenia	Freiburg (2nd)	Caltech (3rd)	Harvard	NYMU Taipei	UC Berkeley	iGEM 2008
2007	Peking	Paris	Slovenia	UC Berkeley	UCSF	USTC	iGEM 2007 [note 1]
2006	Slovenia	Imperial (2nd)	Princeton (3rd)				iGEM 2006
2005	Years prior to 2006 had no specific winners						iGEM 2005
2004							IAP 2004, SBC 2004
2003							IAP 2003

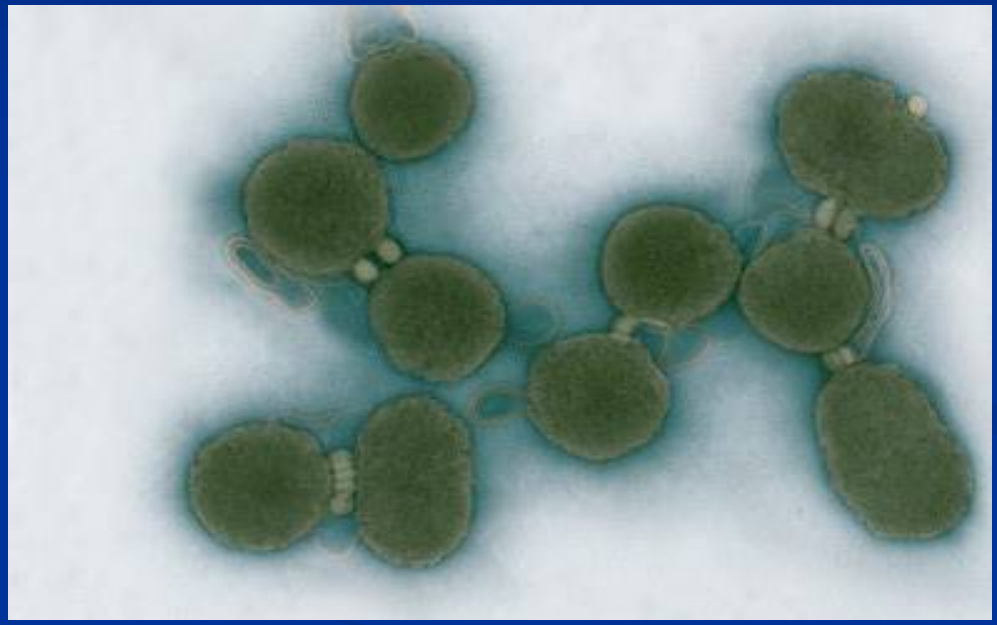
J. Craig Venter and synthetic biology



- Founder of Celera Genomics and The Institute for Genomic Research.
  - In 1995, published two bacterial genomes: *Haemophilus influenzae* Rd and *Mycoplasma genitalium*.
  - In 2003, published human genome.
  - In 2007, published his personal genome.
- President of the J. Craig Venter Institute and co-founder of Synthetic Genomics
  - Dedicated to using modified microorganisms to produce clean fuels and biochemicals.
  - In 2010, created the first bacteria with a synthesized genome.

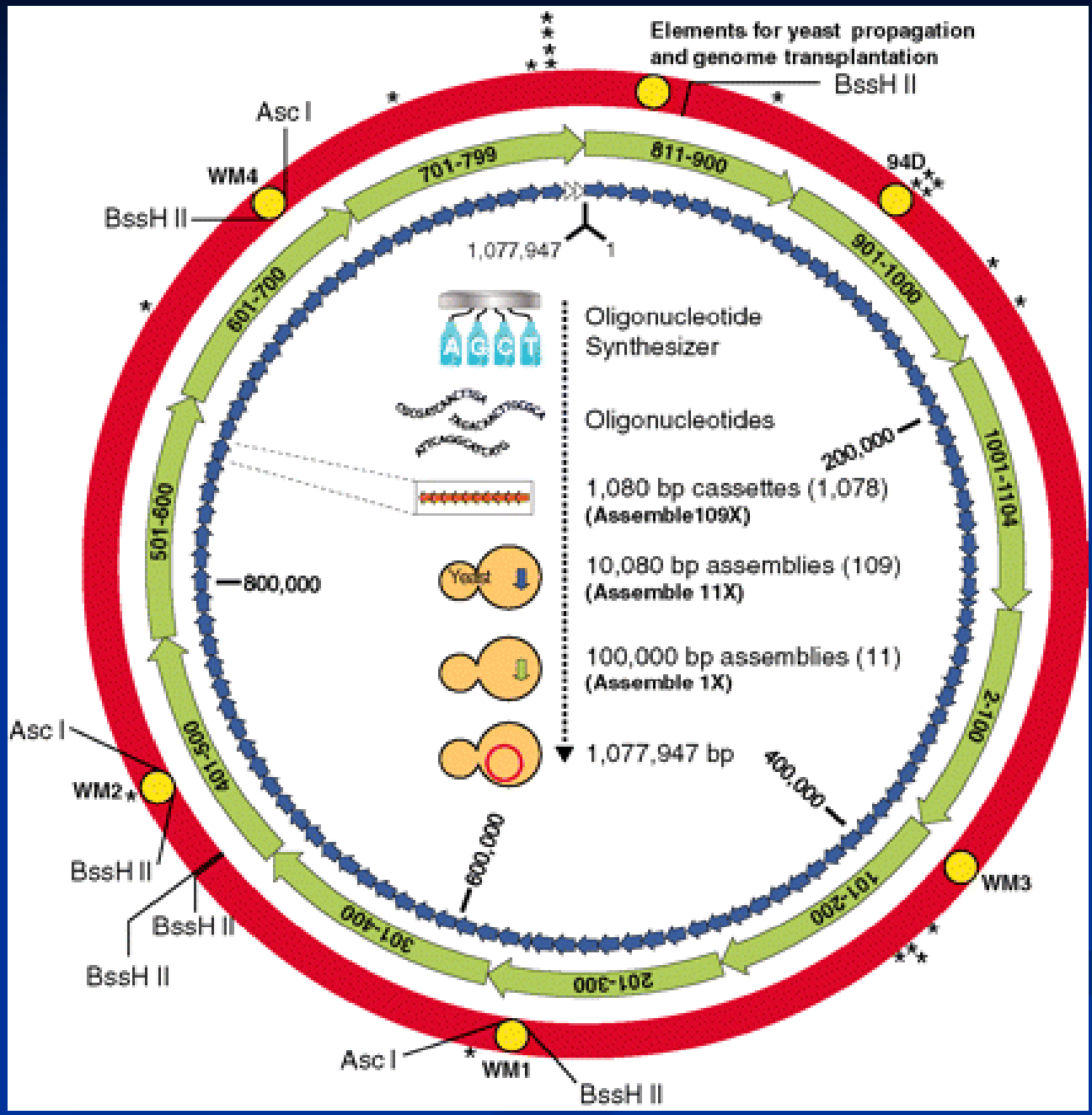
***Mycoplasma mycoides* JCVI-syn1.0 (丝状支原体): The first single-celled organism with a completely artificially synthesized genome.**

- **J. Craig Venter Institute**
- **Gibson DG *et al*. (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*. 329: 52-56.**



The negatively stained transmission electron micrographs of aggregated *M. mycoides*.

- **Artificially synthesized *Mycoplasma mycoides* genome with modifications. Deleted 4000 bp (2 genes), replaced 10 genes with four “watermark” sequences, each over 1000 bp.**
- **1.08 Mb genome, assembled *in vitro* and in yeast cells.**
- **Recipient: *Mycoplasma capricolum* (山羊支原体).**



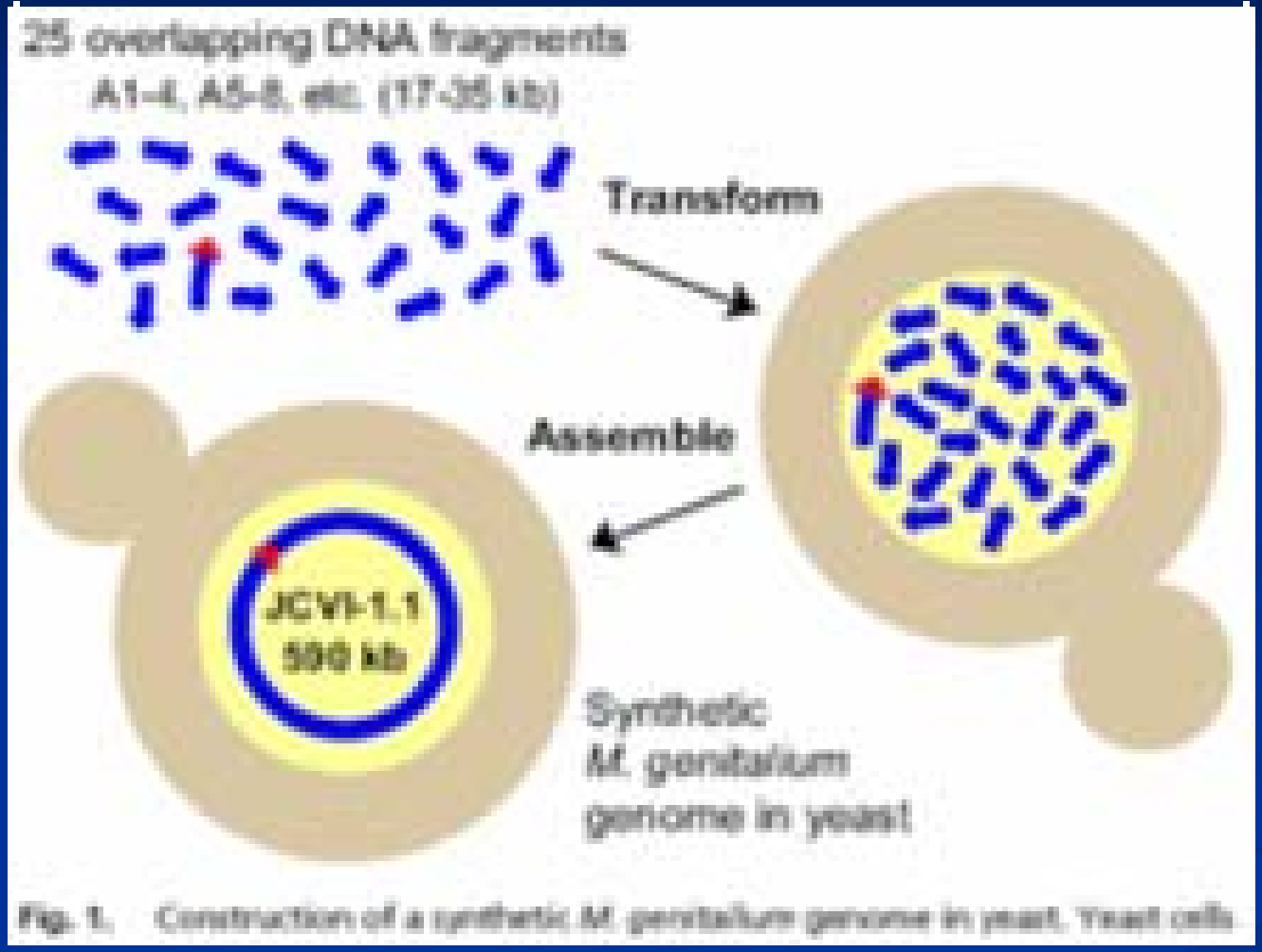
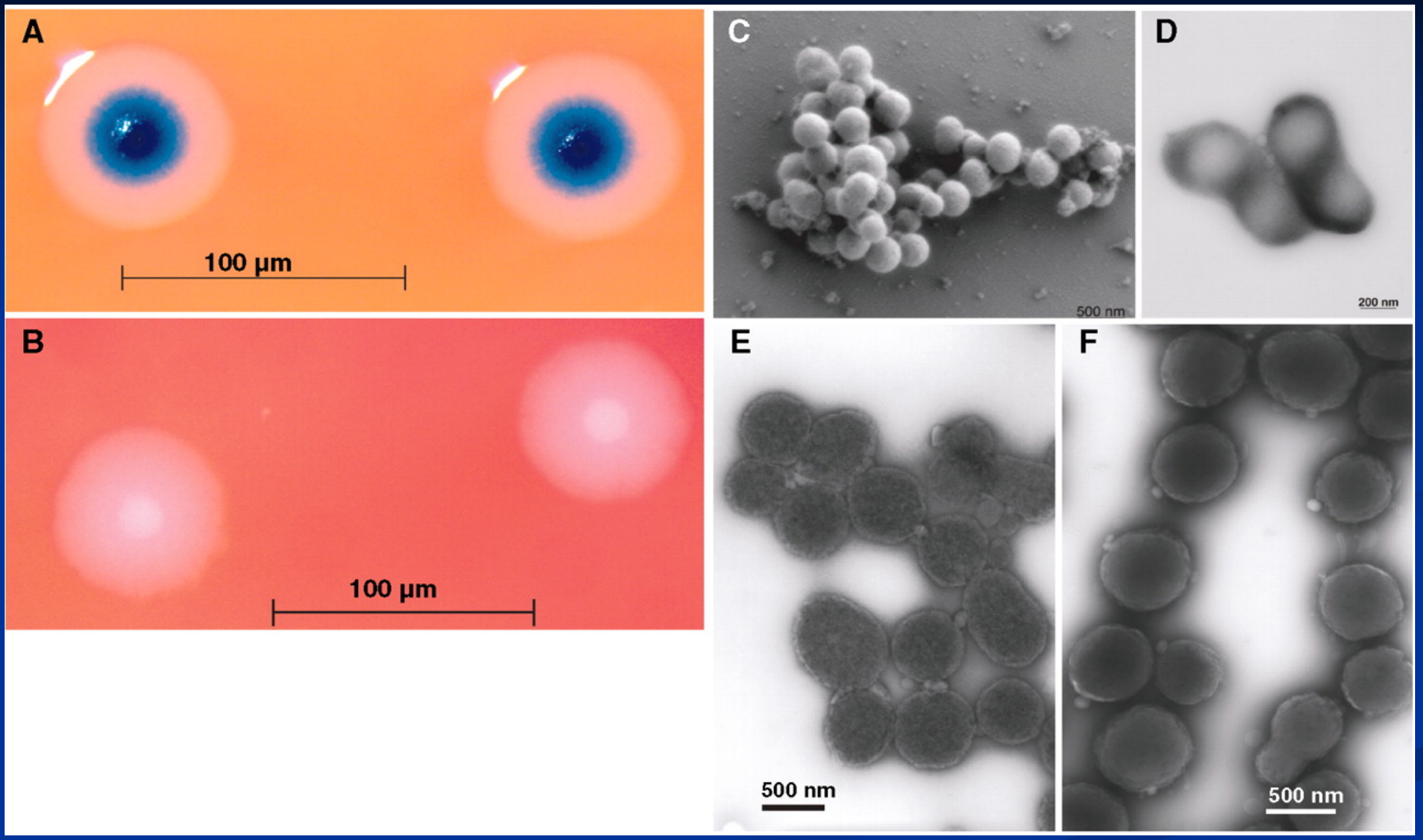
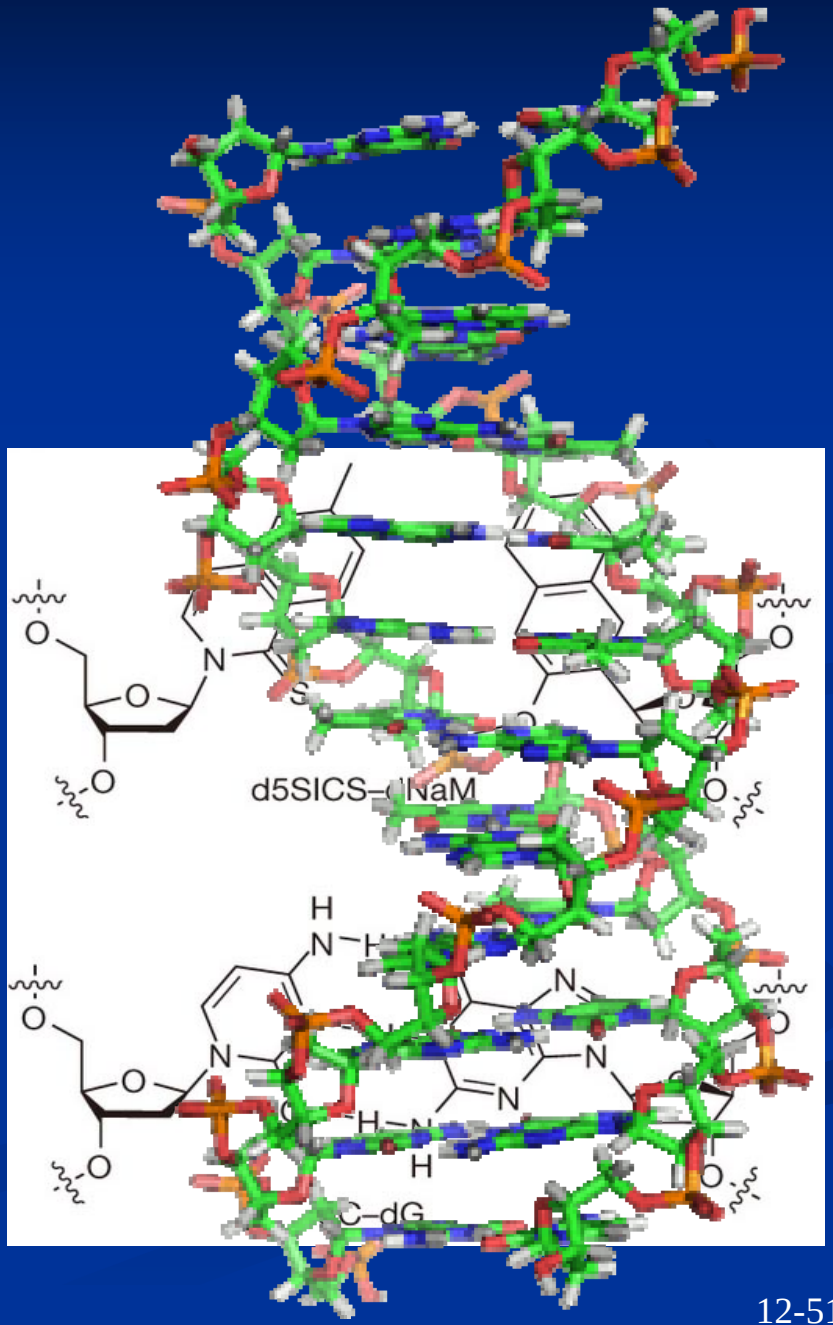


Fig. 1. Construction of a synthetic *M. genitalium* genome in yeast. Yeast cells


A
100 µm
B
100 µm
C
500 nm
D
200 nm
E
500 nm
F
500 nm

Introducing unnatural base pairs (UBP) into DNA

- One **d5SICS-dNaM** pair was introduced into an *E. coli* plasmid in vitro by DNA Pol.I.
- Plasmids replicate well in *E. coli* cells express algal NTP transporter in the presence of UBP.



Malyshev DA, Dhami K, Lavergne T *et al*. 2014 A semi-synthetic organism with an expanded genetic alphabet. *Nature* 509: 385-388.