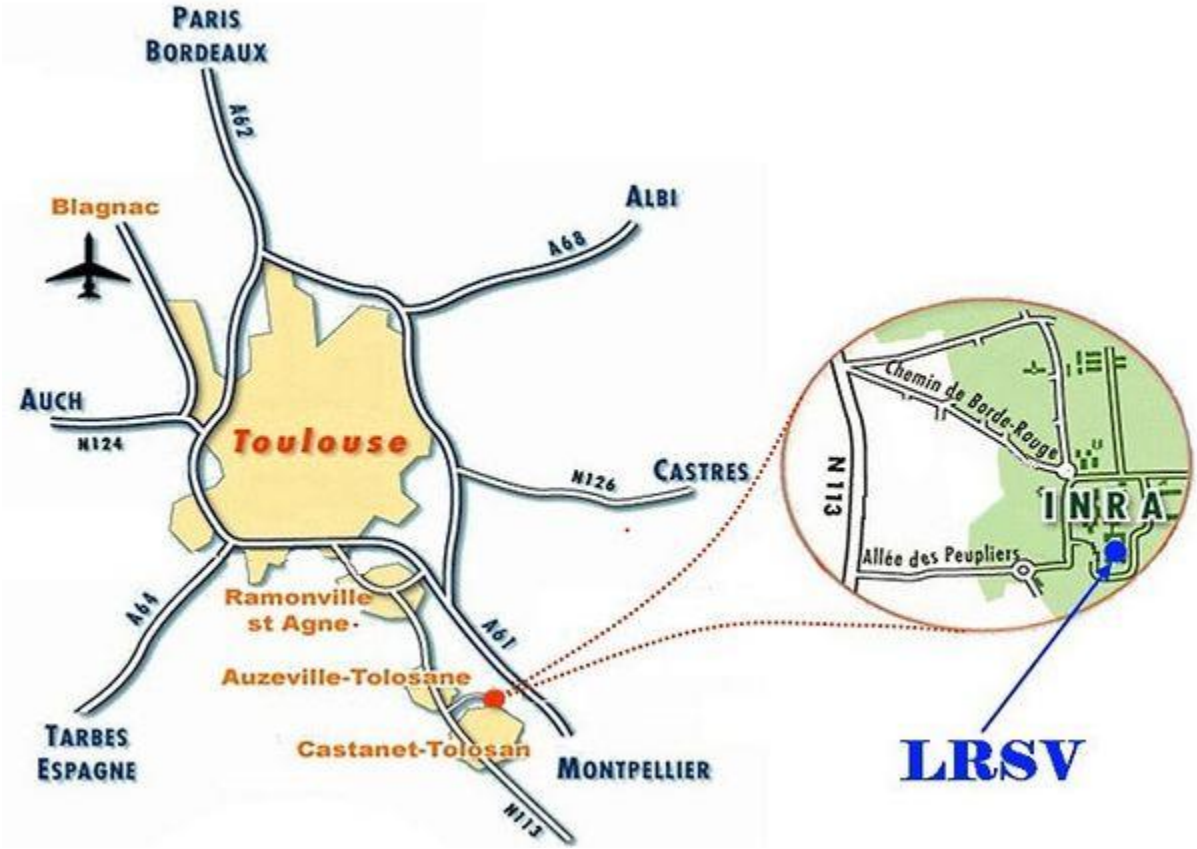
M1 Bioinformatique et Biologie des Systèmes

Jean-Philippe Galaud : galaud@lrsv.ups-tlse.fr

LRSV = Laboratoire de Recherche en Sciences végétales

Site Auzeville (Campus INRA)



Les « om**iques** »

La génom**ique** ?

La transcriptom**ique** ?

La protéom**ique** ?

La métabolom**ique** ?

La chémogénom**ique** ?

La Bioinformat**ique**

Cours : 4h

Faire un point sur vos connaissances
Comment sont générées les données de génomique … ?

1853-1866 **1900** **1953** **1998- ...**

Gregor Mendel
(1822-1884)

- 1ères communications en 1853
- Entre 1856 to 1863, Mendel a cultivé et testé 28,000 plants.

Découverte des lois de l 'hérédité

Mise en application des lois de Mendel

Description de la structure de l 'ADN

Francis Harry Compton Crick

Biochemiker

Francis Harry Compton Crick

* 08. Juni 1916 in Northampton, England

James Dewey Watson

Biochemiker

James Dewey Watson

* 26. April 1928 in Chicago, USA

J. Craig Venter

-Président Celera Genomics

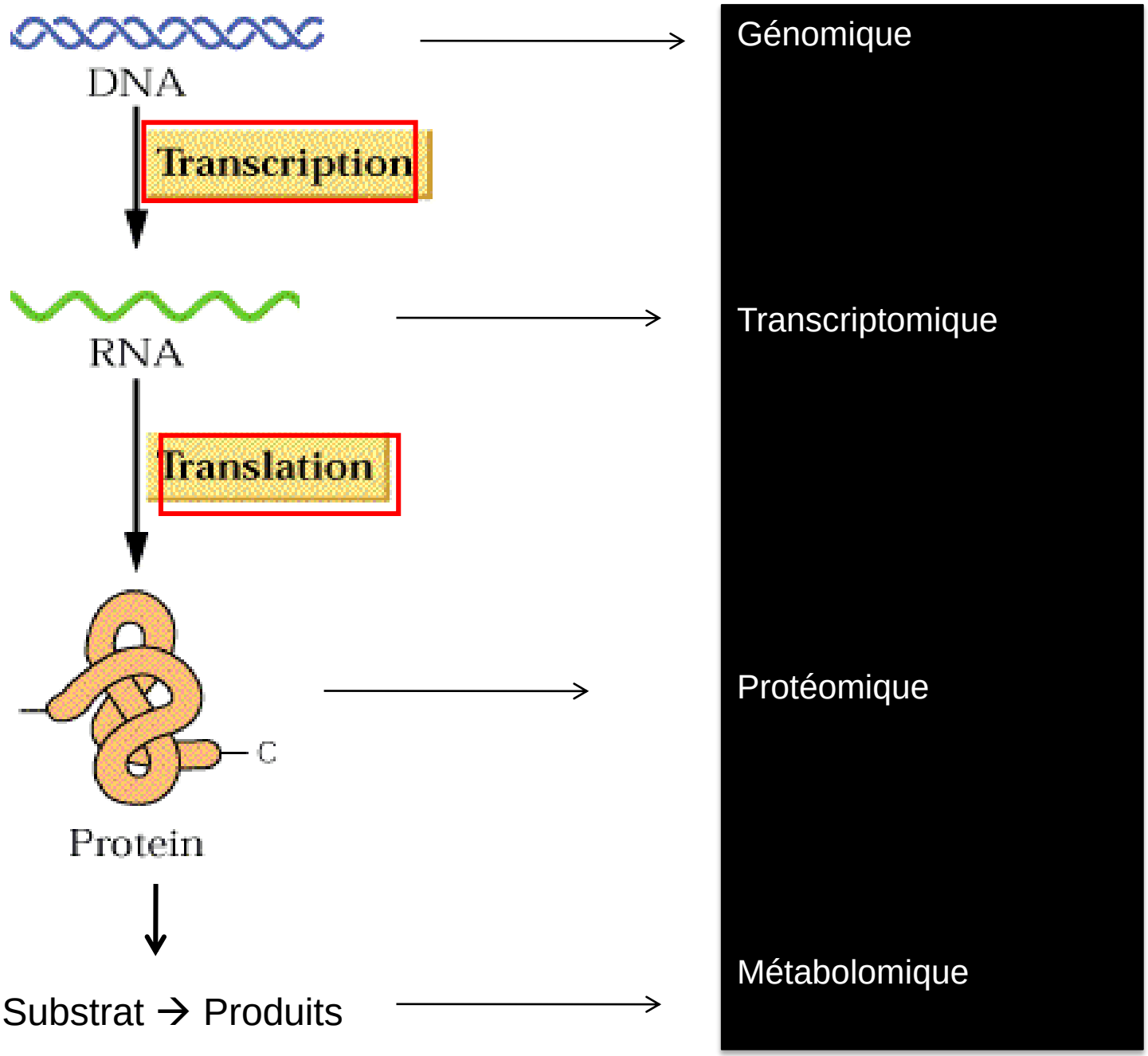
-Sequencage génome
humain
souris
drosophile
Arabidopsis

AFP - Samedi 6 octobre 2007, 19h07
LONDRES (AFP) - Craig Venter est parvenu à réaliser en laboratoire un chromosome de synthèse, premier pas vers la création d'une forme de vie artificielle

La génomique est une biotechnologie qui a pour sujet l'étude des génomes .
Elle travaille au séquençage des chromosomes et à l'étude des fonctions associées aux différents gènes.

La **protéomique** est une science qui étudie, les protéomes ,
c'est à dire l'ensemble des protéines
(leur rôle, leur structure, leur localisation, leurs interactions) dans un organisme , un tissu ou une cellule .

La métabolomique est une science très récente qui
étudie l'ensemble des métabolites (sucres, acides aminés, acides gras,...)
présents dans une cellule , un organe, un organisme.


DNA
Transcription
RNA
Translation
C
Protein
Substrat → Produits
Génomique
Transcriptomique
Protéomique
Métabolomique

CONNAITRE LE GENOME

La génomique

Structure physique

Structure génétique

Les espèces modèles

Plantes

Champignons

Procaryotes

Animaux

↓

Un ancêtre commun ?
LUCA
Last Universal Common Ancestor

Les objectifs

- Localiser et séquencer les gènes

- Étudier la fonction des gènes

Analyse fonctionnelle

Les applications

Construction de génotypes élites

Par :

- la fourniture d'une grande quantité de marqueurs moléculaires
- un meilleur contrôle de la régulation des gènes
- l'identification de gènes candidats pour l'analyse des QTL de caractères majeurs
- l'identification d'allèles favorables

Internet, BLAST et séquencage

Internet : L'accès aux bases de données et aux outils d'analyse devient transparent et facile pour tous

BLAST : rechercher dans les bases de données l'existence de séquences similaires à une séquence donnée c'est la fonction utilisée dans plus de 90% des analyses. BLAST (1990) permet à tout un chacun de le faire, en classant les résultats significatifs

Le séquençage génomique : il donne accès d'un seul coup à toute l'information génétique d'un organisme … mais comment la gérer, la déchiffrer et l'utiliser ?

explosion de nouveaux besoins pour la bioinformatique
explosion de données d'un type nouveau à exploiter

Besoins

Gérer les programmes de séquençage

Annoter les génomes : où sont les gènes? Que font-ils ?

Etablir des ontologies et les peupler : de quoi parle t'on au juste ?

Permettre la comparaison de génomes entre eux

BM → ← Bioinfo

→ Si au départ, c'est la BM qui avait besoin de la Bioinfo

→ aujourd'hui, la bioinfo a aussi besoin des données générées par la BM

Le choix des espèces modèles se fera sur la taille des génomes et sur leur facilite de « culture » en laboratoire

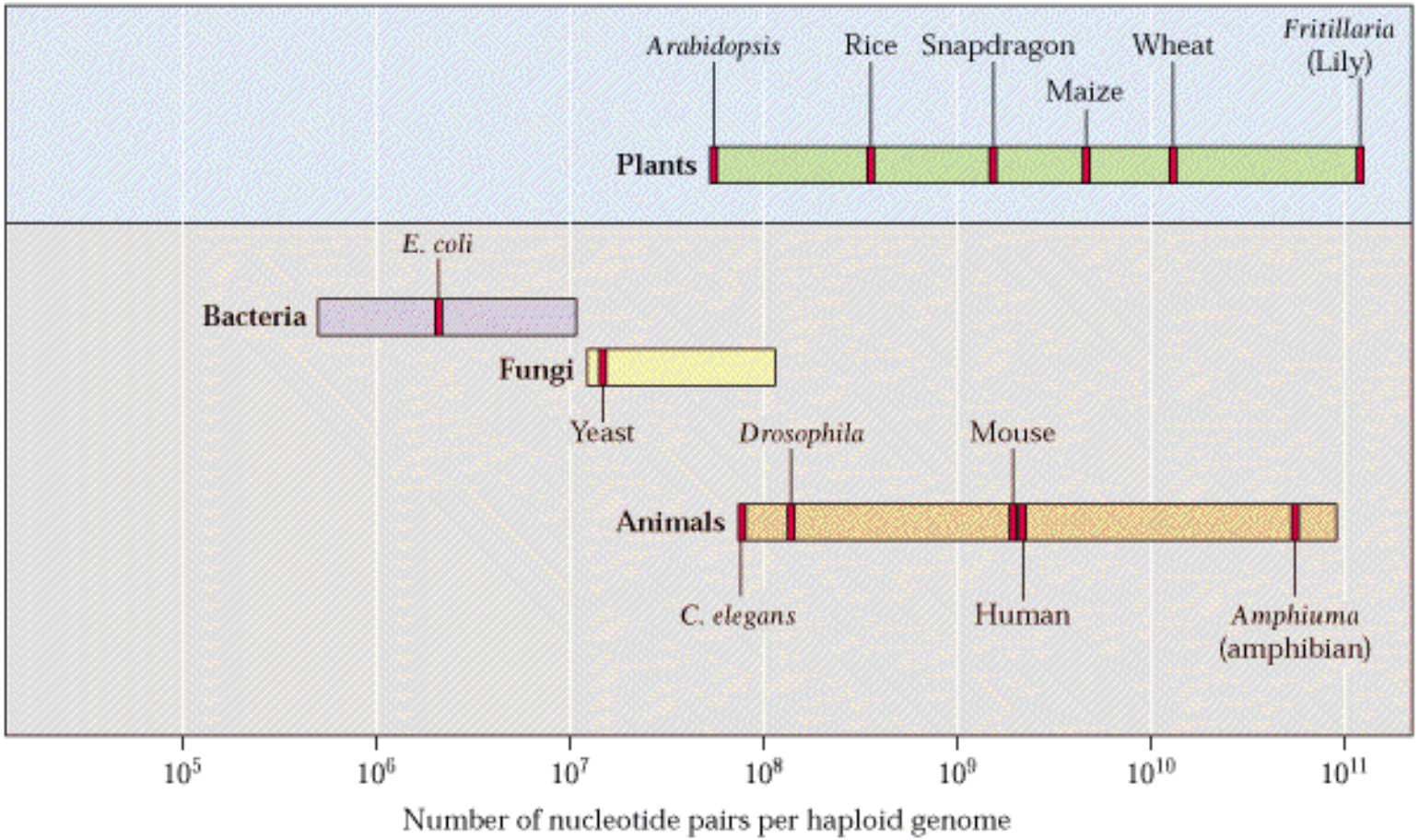
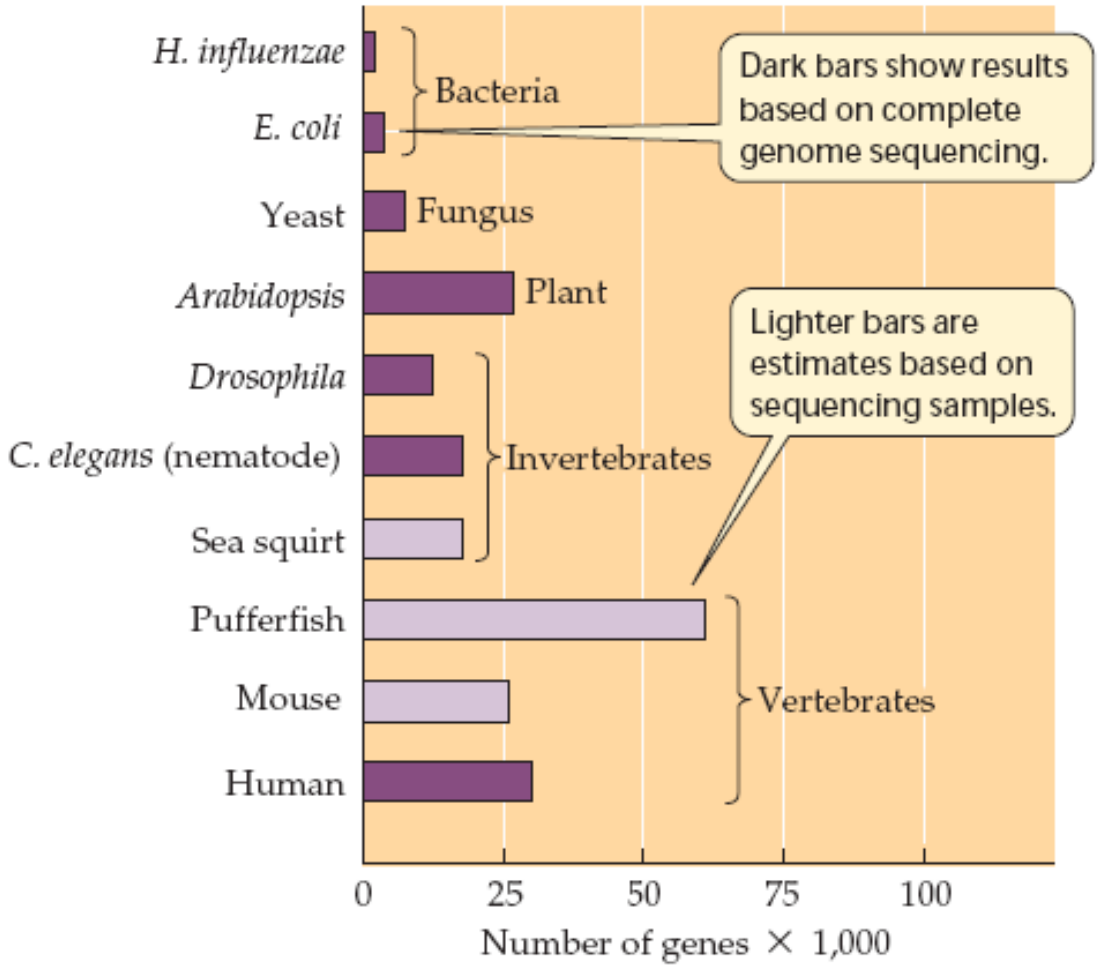


Figure 7.19
C values (haploid genome size in basepairs) from various organisms.

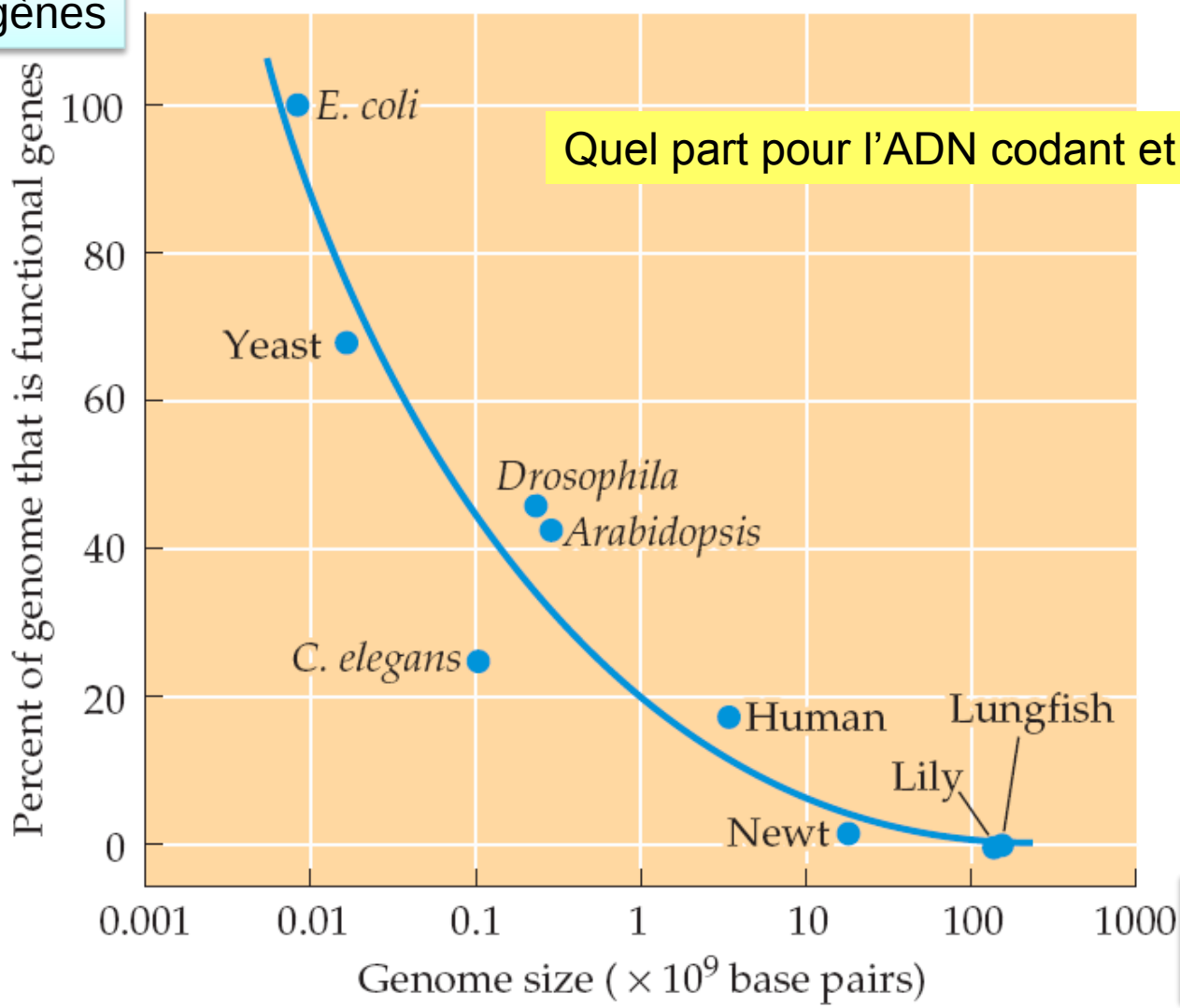
Print Bacl

Combien de gènes par organismes ?



26.7 Complex Organisms Have More Genes than Simpler Organisms Genome sizes have been measured or estimated in a variety of organisms ranging from single-celled prokaryotes to vertebrates.

gènes



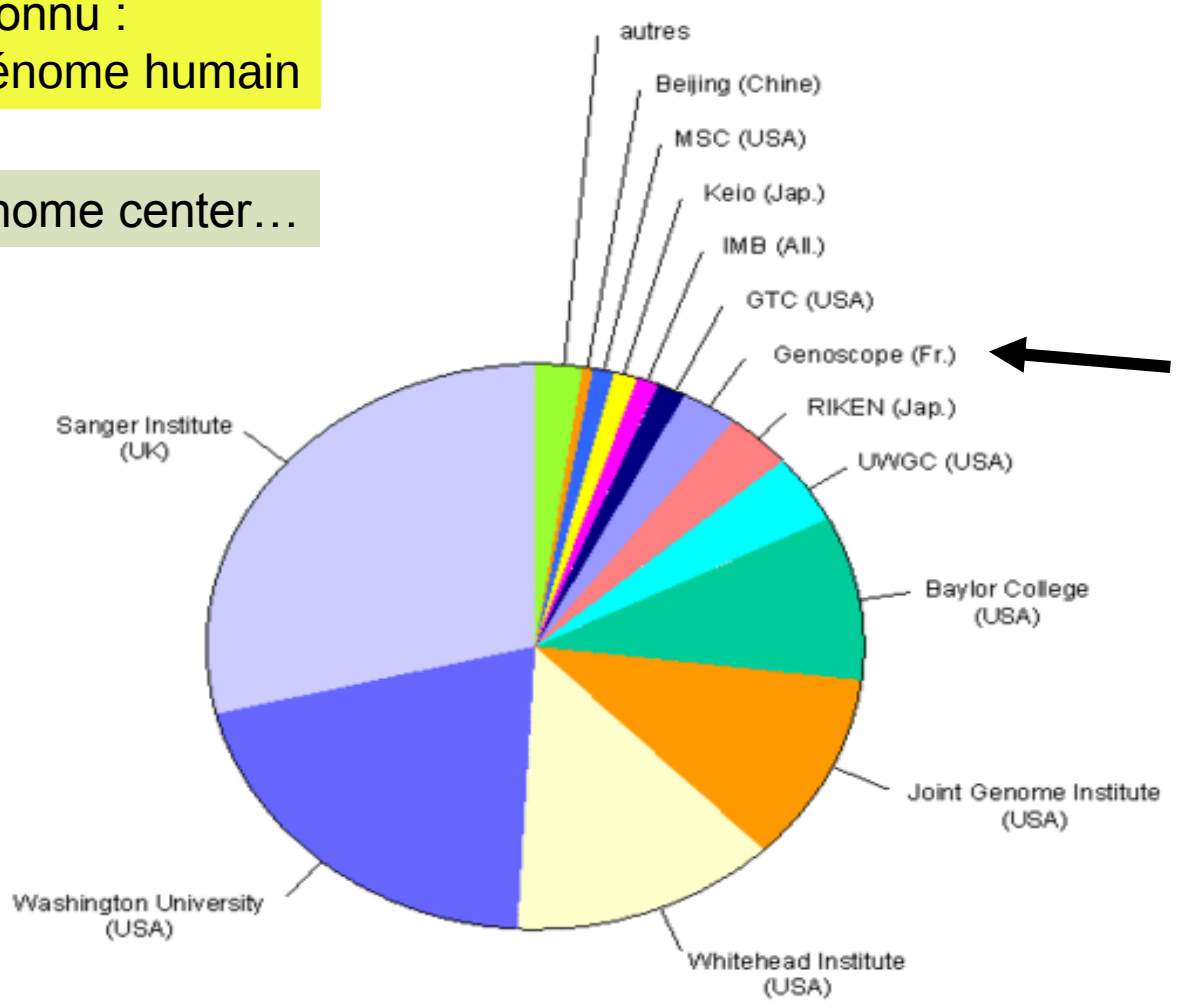
Quel part pour l'ADN codant et non–codant ?

Taille du génome

26.8 A Large Proportion of DNA Is Noncoding Most of the DNA of bacteria and yeasts encodes RNAs or proteins, but most of the DNA of more complex organisms is noncoding. Most noncoding DNA is probably nonfunctional.

L'exemple le mieux connu : Le séquençage du génome humain

Par genome center…

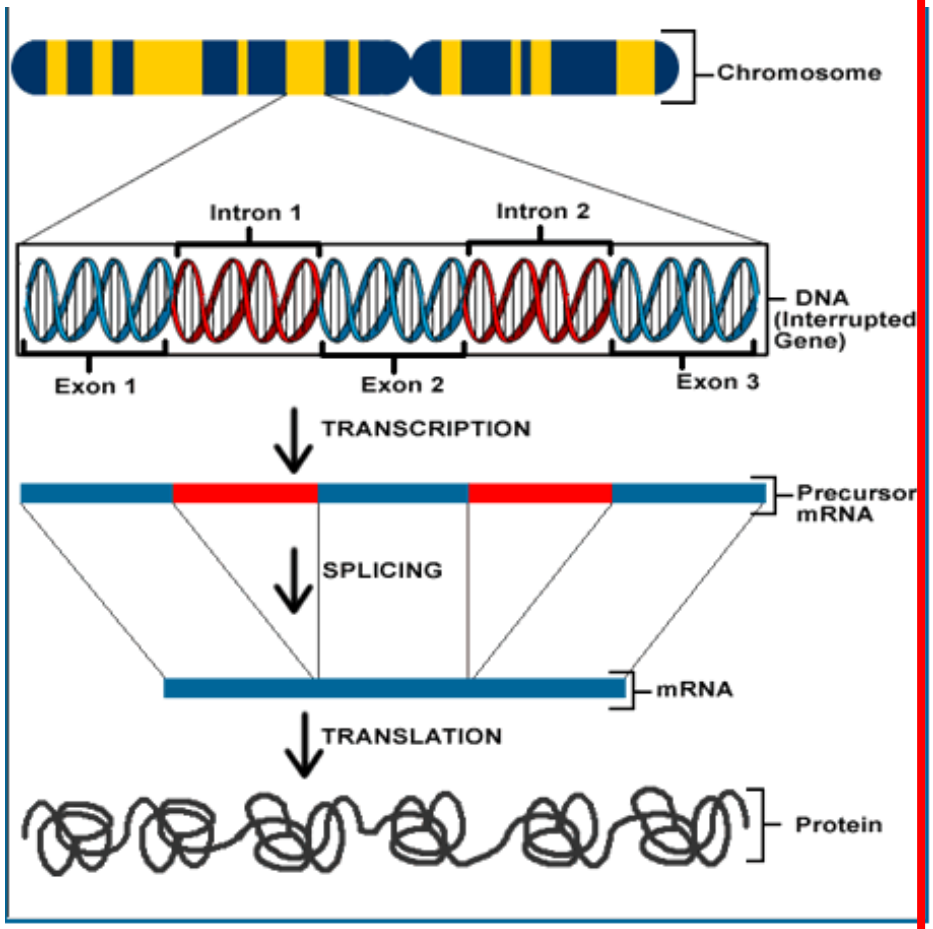


Sur le plan international, les contributions des 6 pays impliqués dans le projet sont les suivantes :

Par pays …

Etats-Unis	60,8 %
Royaume-Uni	28,9 %
Japon	4,9 %
France	2,8 %
Allemagne	1,5 %
Chine	0,7 %

ESTs



igure 1. An overview of the process of protein synthesis.

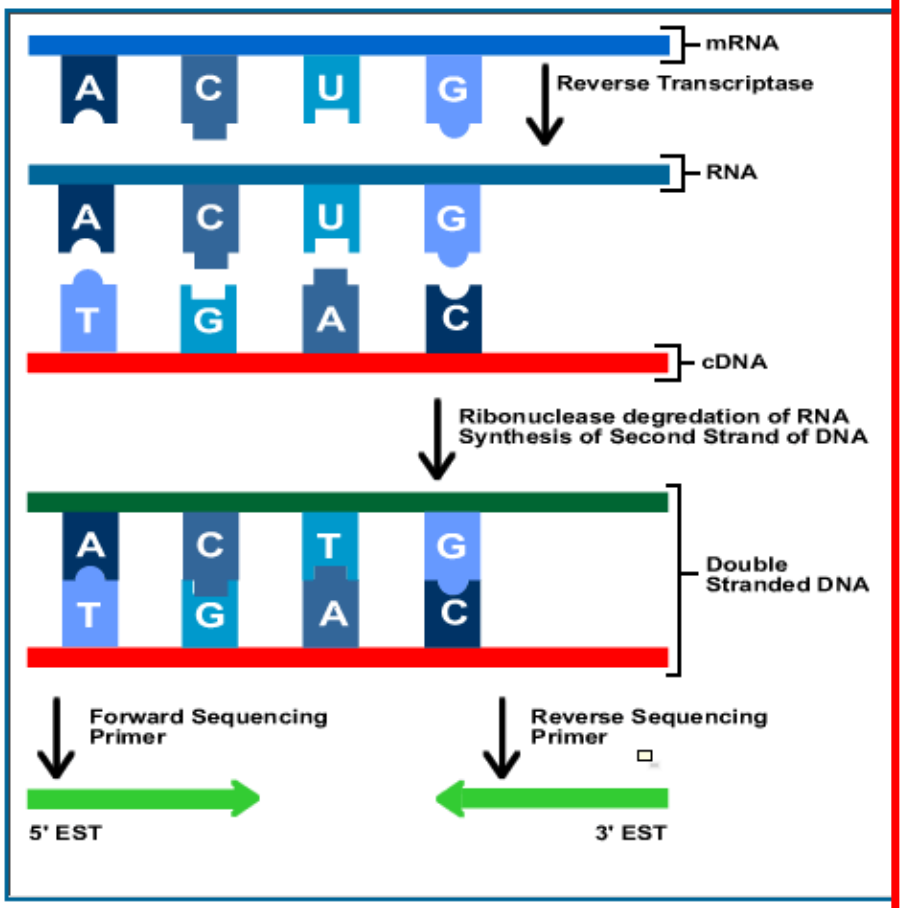
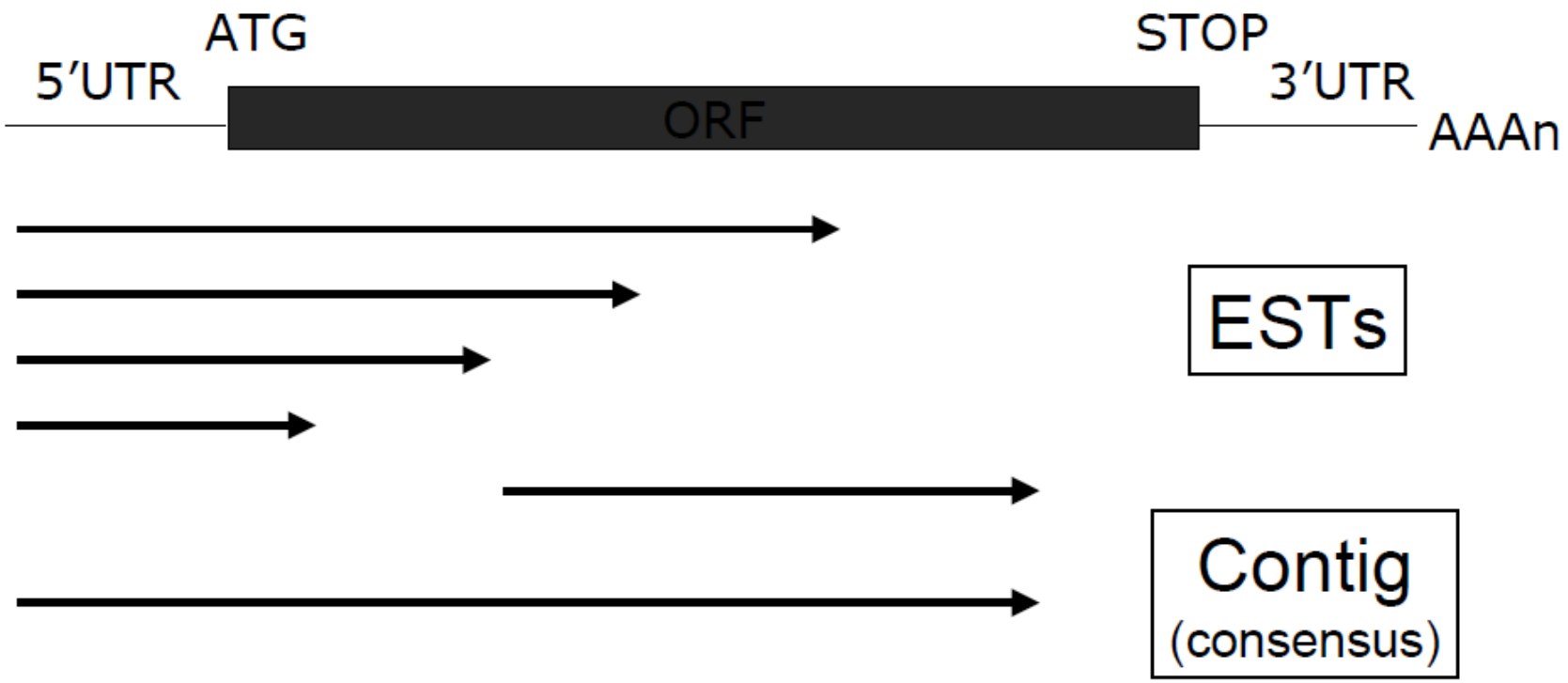


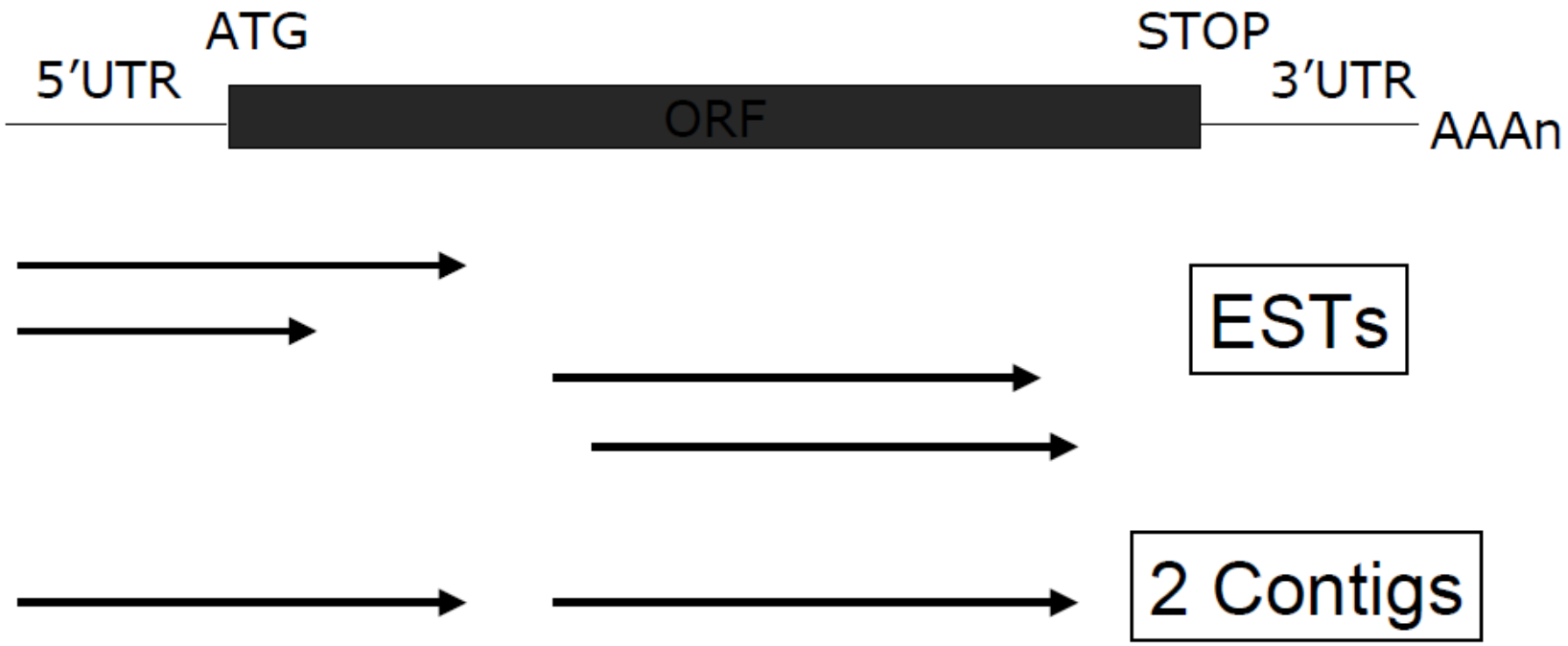
Figure 2. An overview of how ESTs are generated.

Example: EST contig



Donc de plusieurs ESTs → 1 contig

Example: nonoverlapping ESTs



Mais aussi … plusieurs ESTs → … plusieurs contigs

De la sequence de l'Est à son annotation et à la base de données

Functional categories: example

01 Metabolism

01.01 Amino acid metabolism

01.02 Nitrogen and sulfur metab

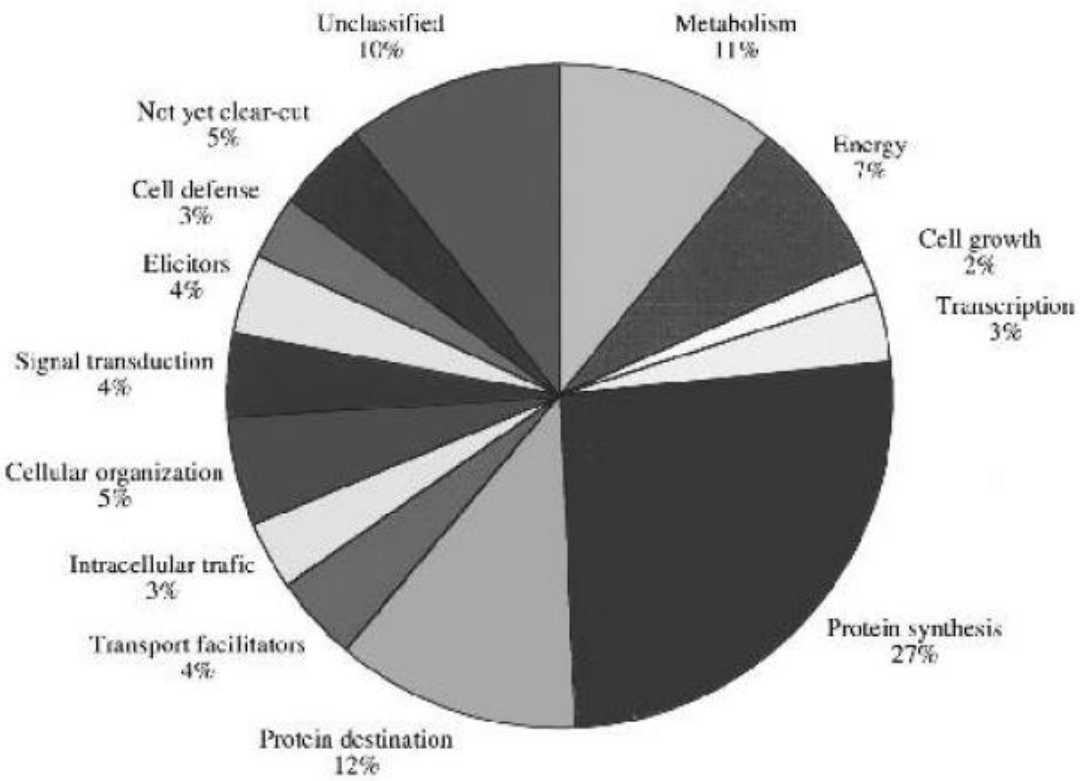
01.03 Nucleotide metabolism

01.04 Phosphate metabolism

01.05 Carbohydrate metabolism

01.06 Lipid and sterol metabolis

01.07 Biosynthesis of vitamins, (
and prosthetic group

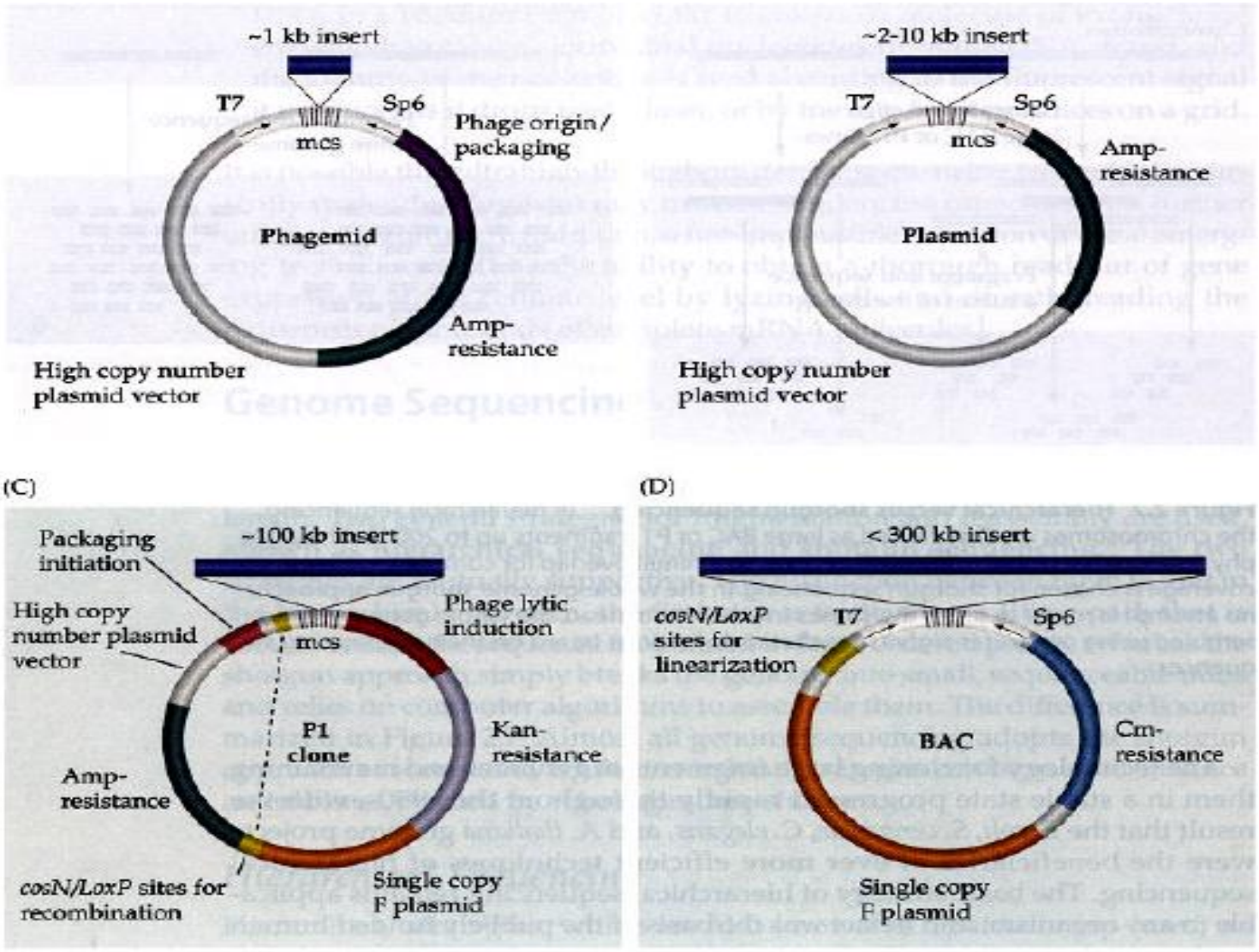


Le séquençage des génomes

Major differences between prokaryote and eukaryote genome

	Genome size	Chromosome	Centromere	Telemere	Organization	Repetitive Sequence
Prokaryotes	Small	Single circular, few linear	No	No	High gene density &lack introns	Non or very low
Eukaryotes	Large	Linear	yes	yes	Low gene density & disrupted by introns	High

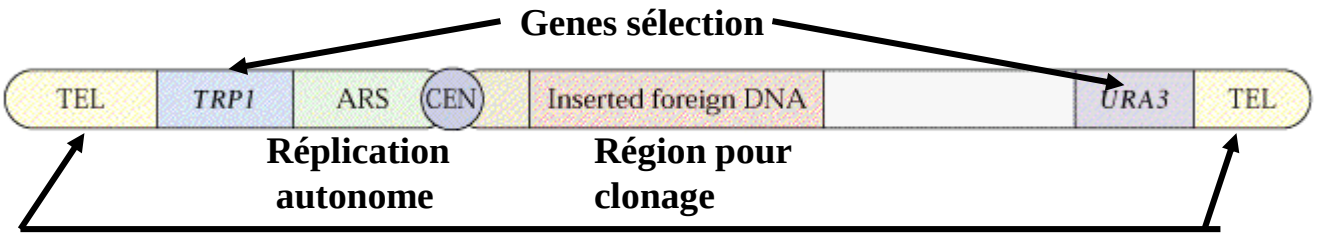
- Le développement de vecteurs pour le séquençage des génomes



YAC : de 250kb à 2900 kb !

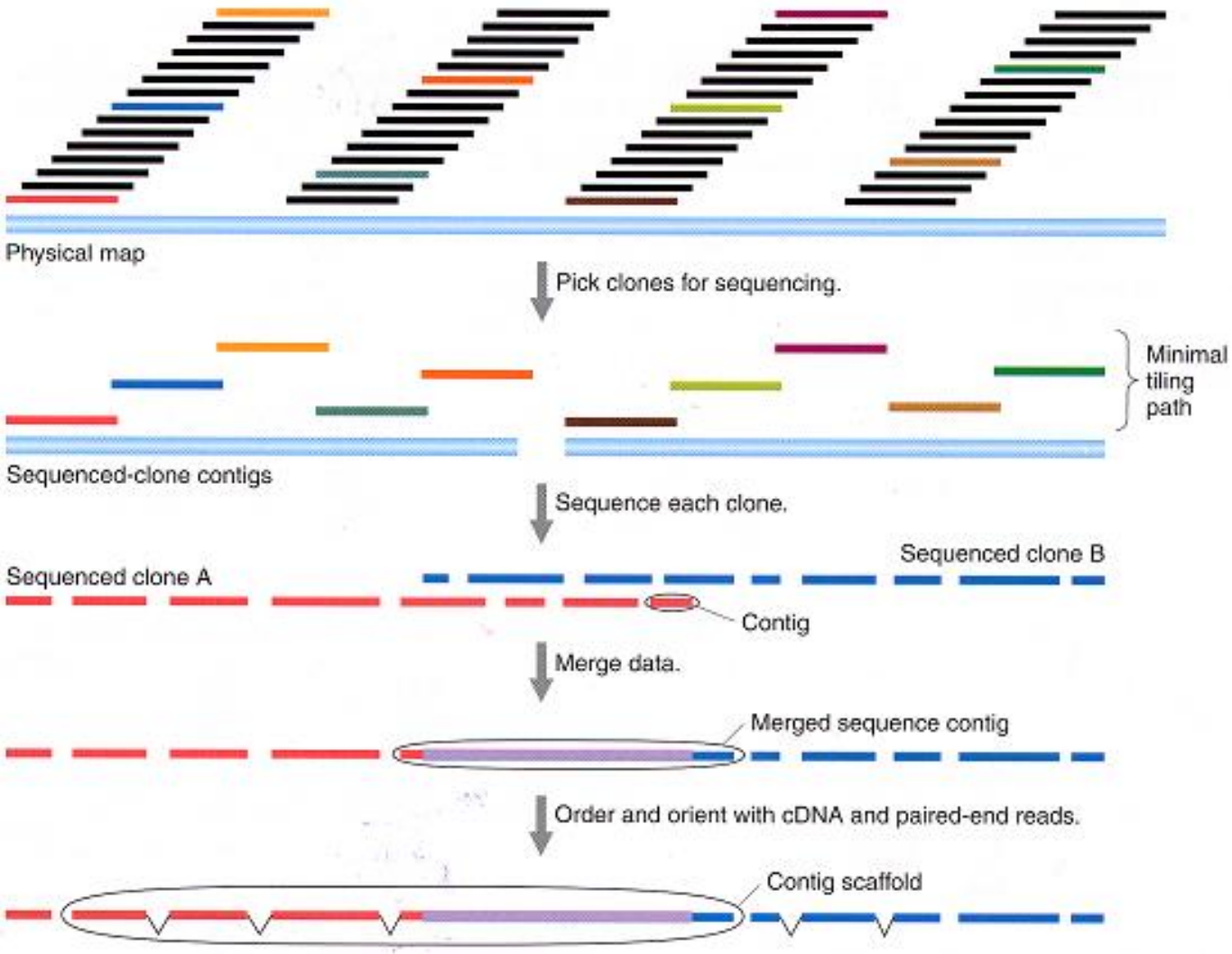
Propagation possible d'ADN exogene dans la levure sous forme d'un chromosome artificiel

→faible efficacite de transfo, pb de recombinaison, clonage de grands fragments



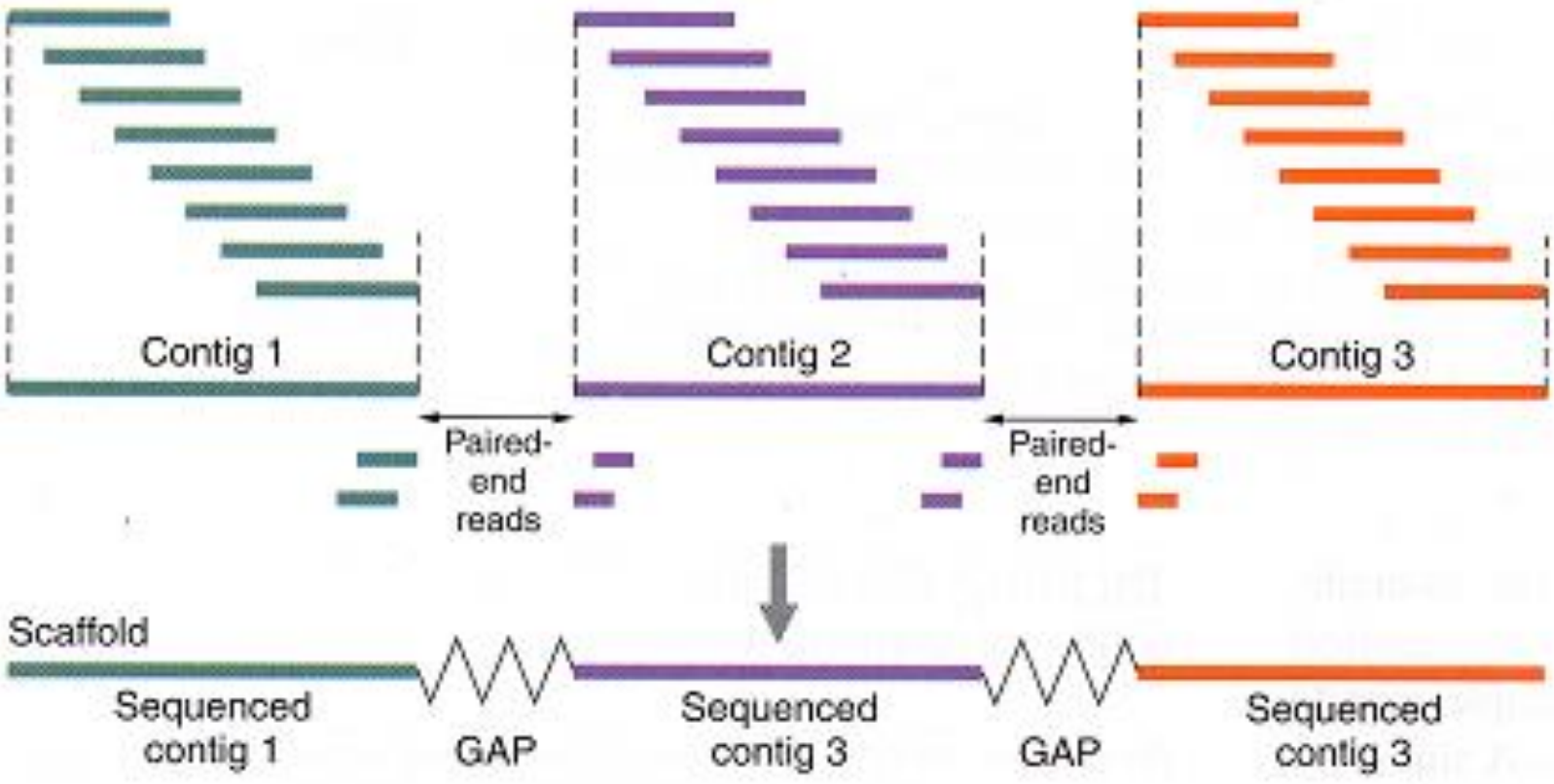
Genome Sequencing by 1st Physical Mapping (Clone à clone)

Extraction A DN, digestion, clonage dans vecteur adapté, assemblage des clones



Genome Sequencing by "Shotgun" Method

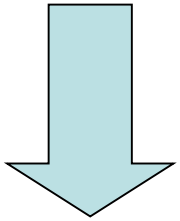
Extraction ADN, digestion en fragment de 1 a 10 kb, sequencage (6-8 fois genome), assemblage



Méthodes de séquençage reposent sur la technique de SANGER

Points -
- **Nécessite d'isoler un clone par PCR, clonage … avant de séquencer**
- **time consuming → donc cher**

Points +
- **mais génère des fragments de 700 à 1000 bp avec peu d'erreurs**



UHTS

ULTRA-HIGH-THROUGHPUT SEQUENCING

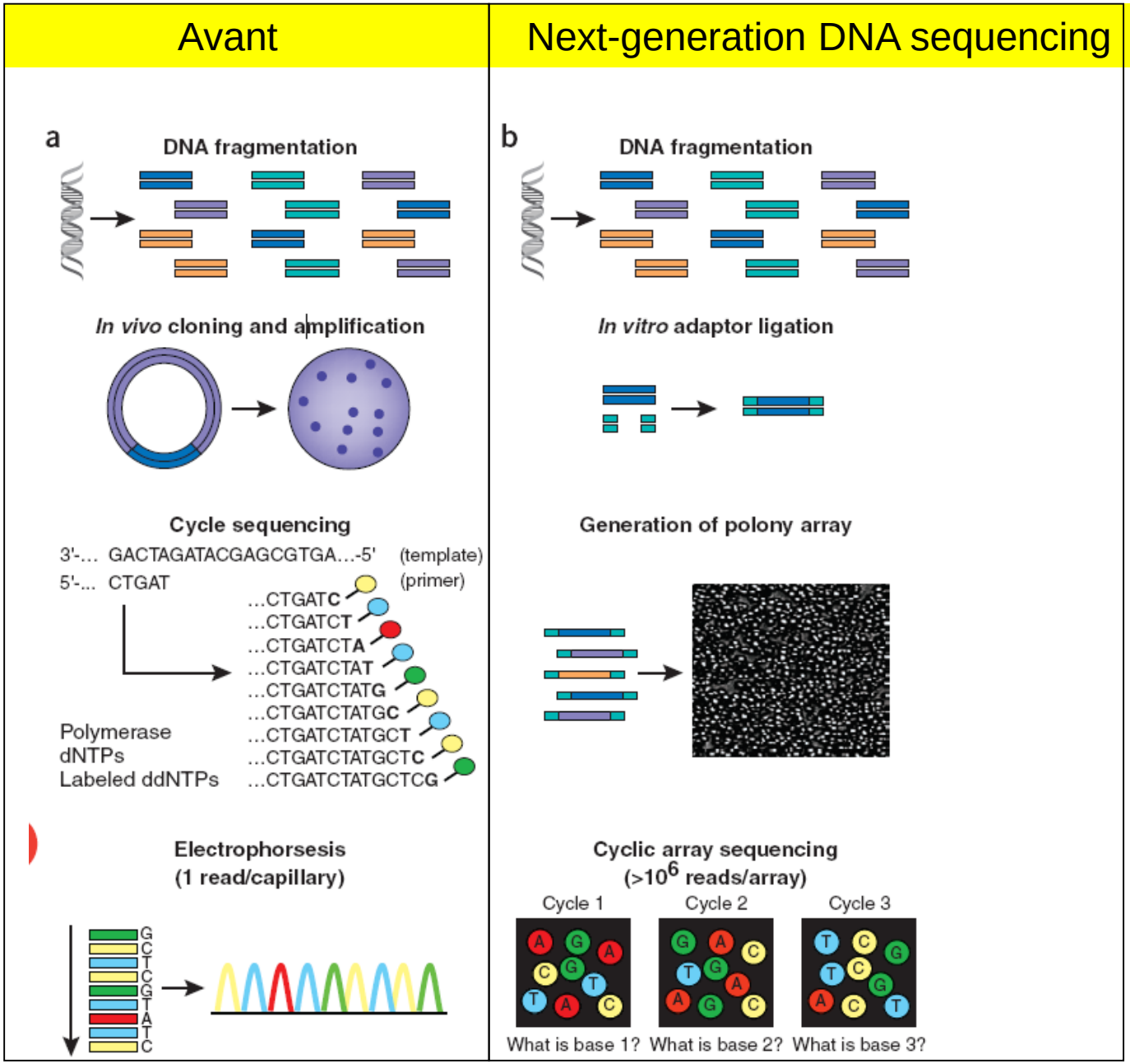


Figure 1 Work flow of conventional versus second-generation sequencing. (a) With high-throughput

doi:10.1038/nature10242

An integrated semiconductor device enabling non-optical genome sequencing

Jonathan M. Rothberg[1], Wolfgang Hinz[1], Todd M. Rearick[1], Jonathan Schultz[1], William Mileski[1], Mel Davey[1], John H. Leamon[1], Kim Johnson[1], Mark J. Milgrew[1], Matthew Edwards[1], Jeremy Hoon[1], Jan F. Simons[1], David Marran[1], Jason W. Myers[1], John F. Davidson[1], Annika Branting[1], John R. Nobile[1], Bernard P. Puc[1], David Light[1], Travis A. Clark[1], Martin Huber[1], Jeffrey T. Branciforte[1], Isaac B. Stoner[1], Simon E. Cawley[1], Michael Lyons[1], Yutao Fu[1], Nils Homer[1], Marina Sedova[1], Xin Miao[1], Brian Reed[1], Jeffrey Sabina[1], Erika Feierstein[1], Michelle Schorn[1], Mohammad Alanjary[1], Eileen Dimalanta[1], Devin Dressman[1], Rachel Kasinskas[1], Tanya Sokolsky[1], Jacqueline A. Fidanza[1], Eugeni Namsaraev[1], Kevin J. McKernan[1], Alan Williams[1], G. Thomas Roth[1] & James Bustillo[1]

Nature 2011, 475, 348-352

DNA sequencing technology in semiconductor able to directly perform non-optical DNA sequencing of genomes.

Sequence data are obtained by directly sensing the ions produced by template-directed DNA polymerase synthesis using all-natural nucleotides on this massively parallel semiconductor-sensing device or ion chip.

The ion chip contains ion-sensitive, field-effect transistor-based sensors in perfect register with 1.2 million wells,
which provide confinement and allow parallel, simultaneous detection of independent sequencing reactions.

We show the performance of the system by sequencing three bacterial genomes, its robustness and scalability
by producing ion chips with up to 10 times as many sensors and sequencing a human genome.

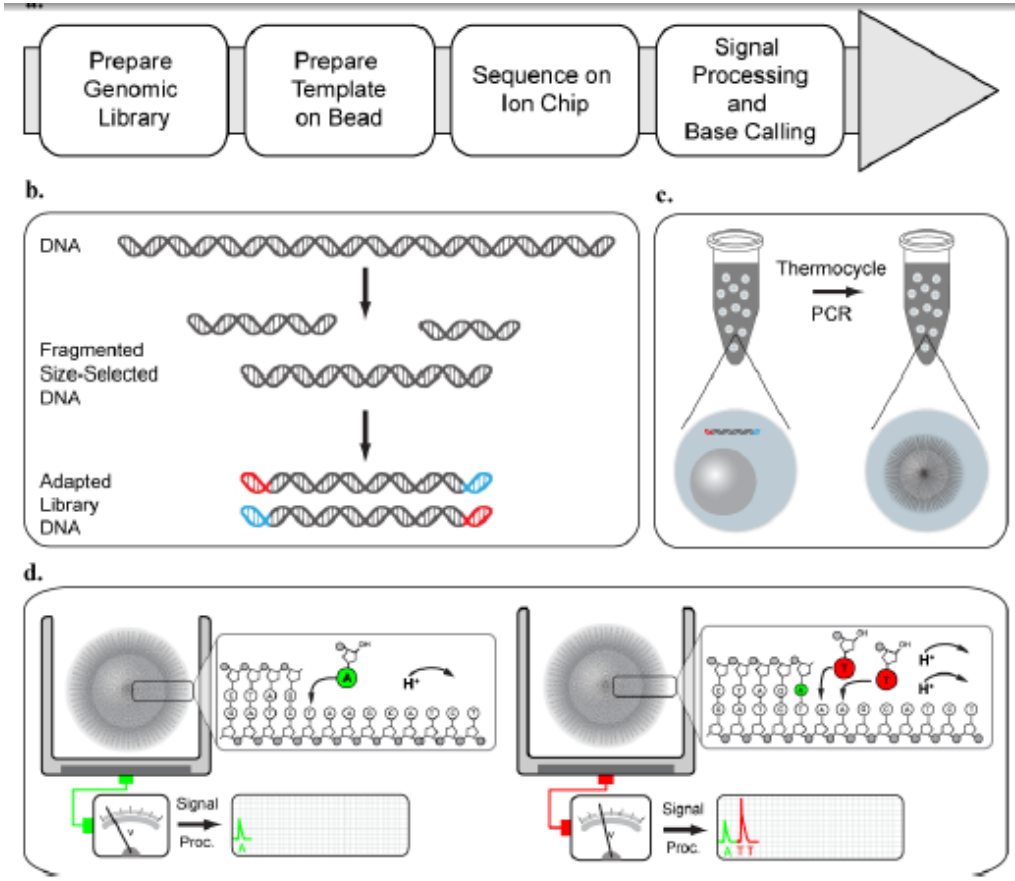
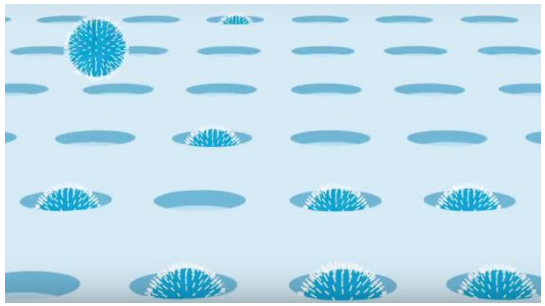


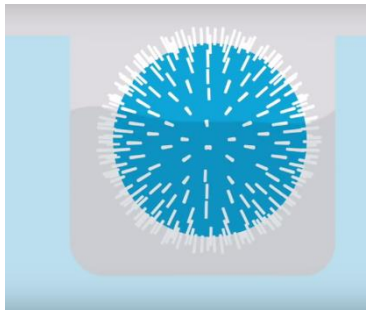
Figure S1 Process overview

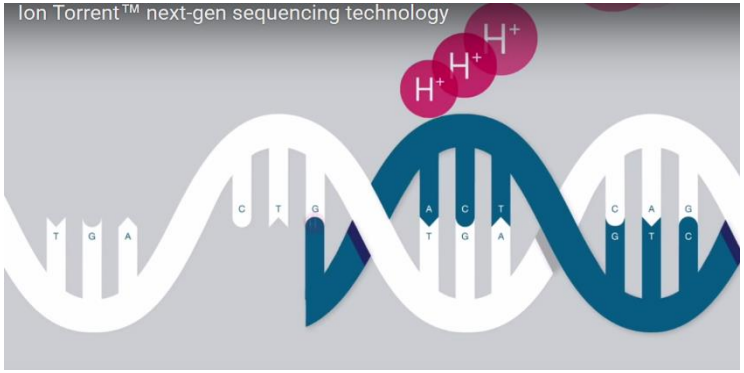
a, Overview of ion sequencing work flow. **b,** Prepare genomic library, DNA is fragmented, sized, and forward and reverse adapters ligated. **c,** Amplify Template on bead, adapter-ligated libraries are clonally amplified onto beads. A magnetic bead-based enrichment process selects template-carrying beads. **d,** Sequence on ion chip, sequencing primers and DNA polymerase are bound to the template-carrying beads, beads are pipetted into the chip's loading port. The chip is installed in the sequencing instrument; all four nucleotides cyclically flowed in an automated 2-hour run. Signal processing, software converts the raw data into measurements of incorporation in each well for each successive nucleotide flow. After bases are called, each read is passed through a filter to exclude low-accuracy reads and per-base quality values are predicted.

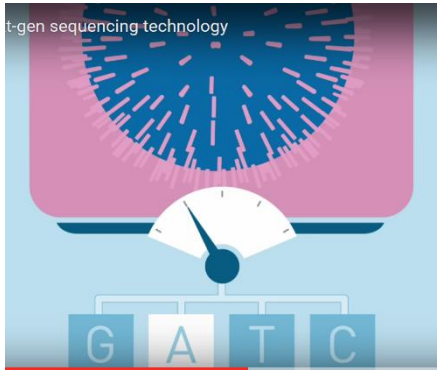
Ion Torrent semiconductor sequencing

Méthode basée sur la detection des ions Hydrogènes qui sont relargués pendant la polymerization de l'ADN
→ Détection par un système "Ion sensors" qui se traduit par la mesure d'un signal electronique



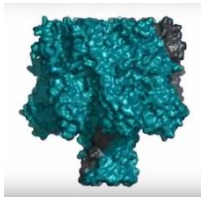


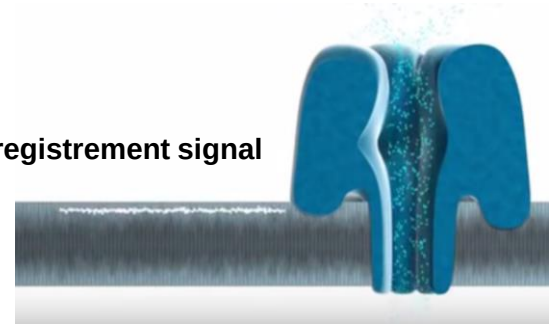


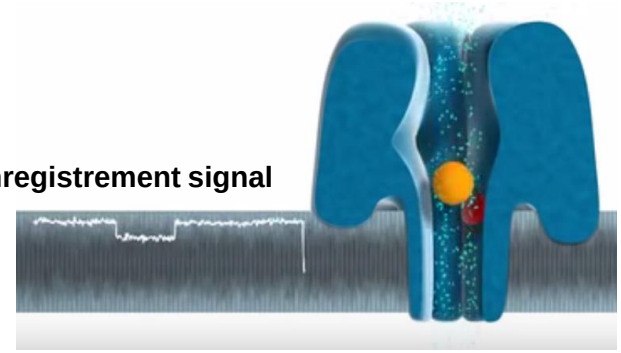


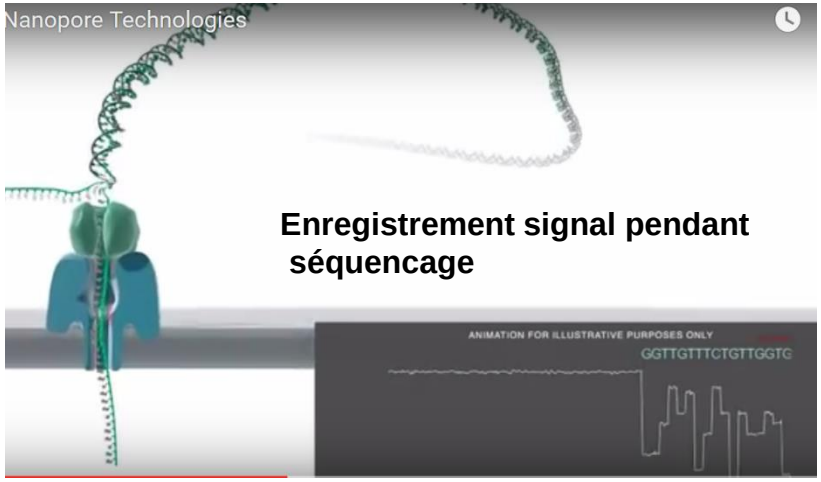
Nanopore DNA sequencing

Méthode basée sur l'émisssion d'un signal electrique qui est généré lorque l'AND passe au travers d'un nanopore.
Le changement depend de la forme, de la taille de l'ADN
Méthode qui ne nécessite pas de nucleotides modifies









Emergence de DNA sequencing sensors (séquenceur de poche)

Minion technology

MinION MkI: portable, real-time biological analyses

Perspective

Biological data sciences in genome research

Michael C. Schatz

Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

Perspective

A vision for ubiquitous sequencing

Yaniv Erlich[1,2]

[1]Department of Computer Science, Columbia University, New York, New York 10027, USA; [2]New York Genome Center, New York, New York 10013, USA

→ Détection pathogenes chez passagers dans les aéroports (*ie* Ebola virus), contrôle identité, Qualité et sécurité alimentaire…)

Single molecule real time sequencing (**SMRT**) – PacBIO

Séquencage en // de nombreux fragments uniques
Une ADN polymerase est fixée au fond du puit avec 1 molecule d'ADN,
Nucléotides couples a un fluorochrome different ce qui permet detection du nucleotide incorporé et de la sequence

1.1 kilobases séquencé en moyenne
Développement → vers 2.5 to 2.9 kilobases et plus (5-20 kilobases)

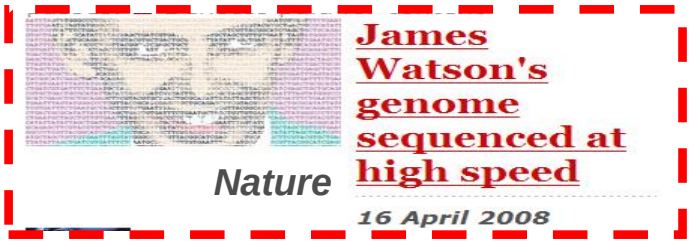
Le séquençage du Génome de Watson , n'est pas le 1er génome humain séquencé (C. Venter en 2007) mais …

nature Vol 452|17 April 2008|doi:10.1038/nature06884

LETTERS

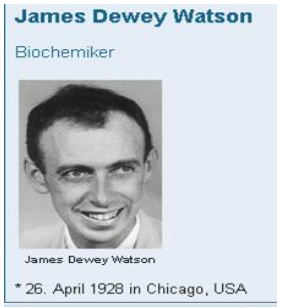
The complete genome of an individual by massively parallel DNA sequencing

David A. Wheeler[1]*, Maithreyan Srinivasan[2]*, Michael Egholm[2]*, Yufeng Shen[1]*, Lei Chen[1], Amy McGuire[3], Wen He[2], Yi-Ju Chen[2], Vinod Makhijani[2], G. Thomas Roth[2], Xavier Gomes[2], Karrie Tartaro[2]†, Faheem Niazi[2], Cynthia L. Turcotte[2], Gerard P. Irzyk[2], James R. Lupski[4,5,6], Craig Chinault[4], Xing-zhi Song[1], Yue Liu[1], Ye Yuan[1], Lynne Nazareth[1], Xiang Qin[1], Donna M. Muzny[1], Marcel Margulies[2], George M. Weinstock[1,4], Richard A. Gibbs[1,4] & Jonathan M. Rothberg[2]†



Stratégie UHTS

J. Craig Venter

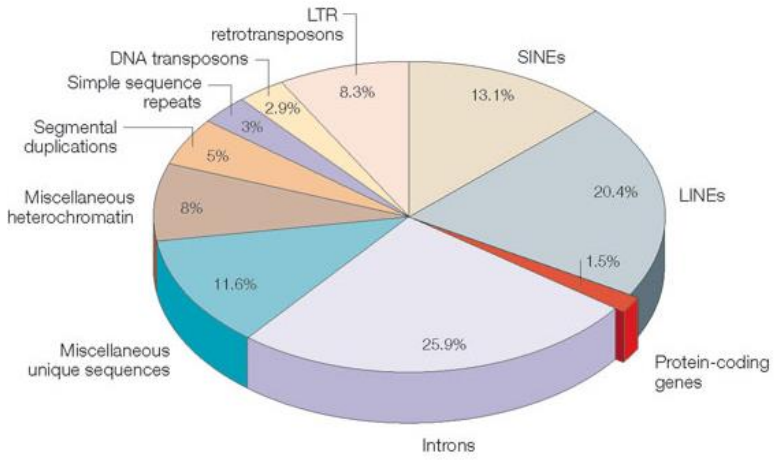


QUICKER, SMALLER, CHEAPER

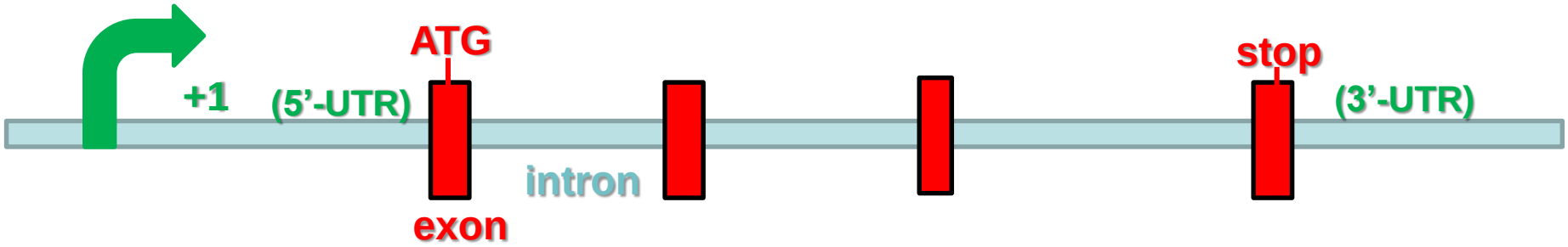
Genome sequenced (publication year)	HGP (2003)	Venter (2007)	Watson (2008)
Time taken (start to finish)	13 years	4 years	4.5 months
Number of scientists listed as authors	> 2,800	31	27
Cost of sequencing (start to finish)	$2.7 billion	$100 million	< $1.5 million
Coverage	8–10 ×	7.5 ×	7.4 ×
Number of institutes involved	16	5	2
Number of countries involved	6	3	1

Human Genome Project

L'ADN non-codant:
Le côté obscur des génomes



autantacagtagatacagtagatagatgtatagtttaggctgagctga
ggcgcgtagatagtcgctgaaggctgagattcgcgttaaatgagatagg
ggcgcgctgagaggggcaggatcggc**chercher**ggggcatgacgtagt
gacaggacatagcagtaagcatagacagtaaaagccggtagcggataga
gaagaacataagagcagaat**une**aaaggataagagaagtatgatagata
tgaatagagaacacagatataagagcagagataagacgaatag**ai**taga
gagacattatgaaccaagatagacaagat**guille**tatagacagataga
gatagaacagtagatagacaggatagacagaatagacagatagaacgta
catataagacagatagaacagatagacaagtagaacagagtatatagag
tatatatagacagatagaatatagatagataaggatttttttttttttt
aagacataagatgatagaatagatagtagatag**dans**gatagataagga
tatagatagatagagagaatataatataatatagagatatatattaggg
gagatataagagtagaaggaatagagaatagagagagaatataaatttt
unegagatacagatagatagaacagatagaacagagatagaacagagat
gataagatagaatatgataagatc**botte**agaatagaatagataagata
gataagatagataagatagatagattttagagataagaagagagatag
tatagatagatagaacagataagacagatagaacagatagaagtaagat
aatataataagag**d**aagatagagacagatagaa**e**agatatagacagata
gagatagatagaatgcaagtaagacagatagaacag**foin**gatagaata



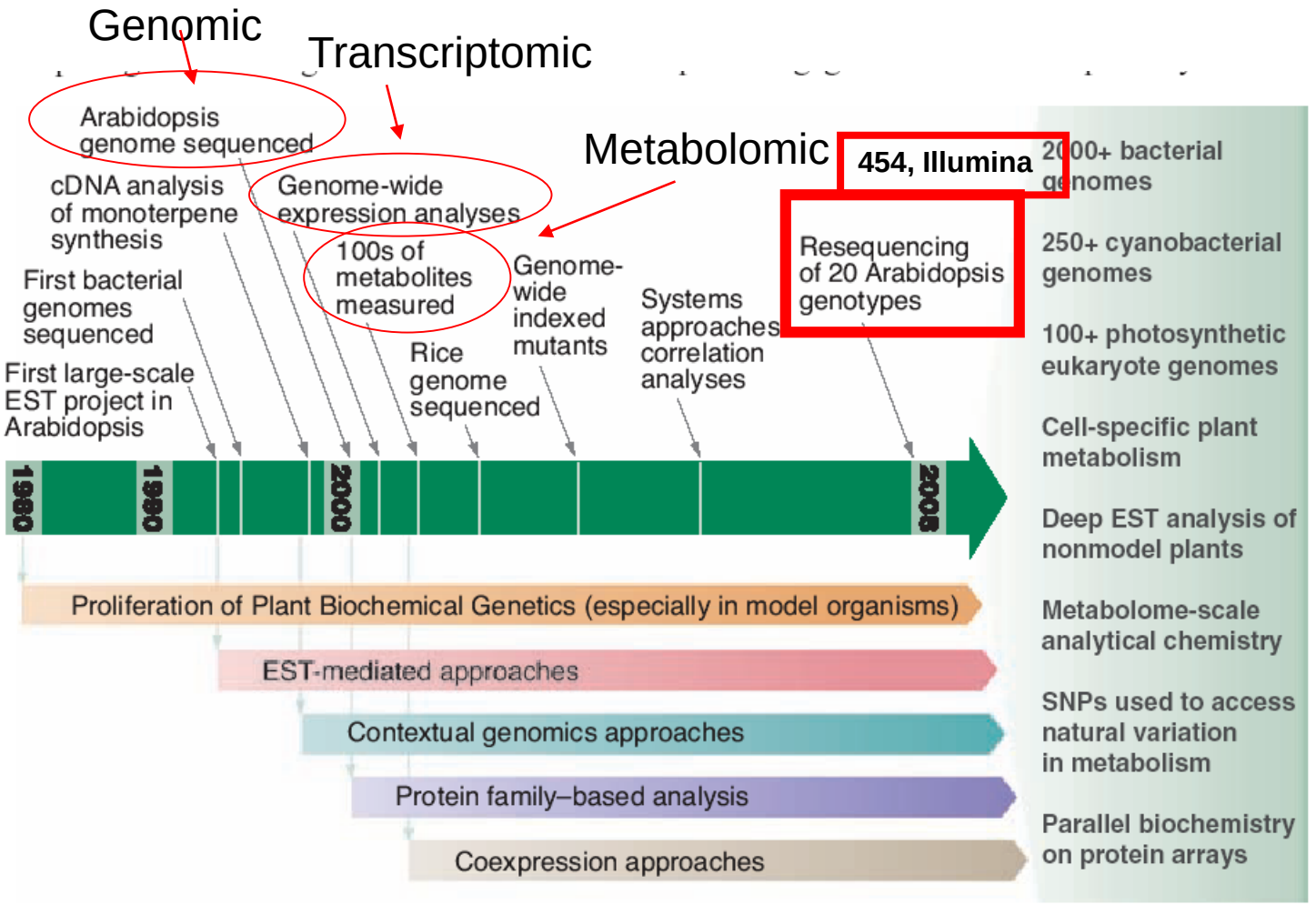


Fig. 1. Time line of genomics-enabled plant biochemistry. Selected major advances to date are indicated above and below the time line. Some approaches and tools likely to further understanding of plant biochemistry during the coming decade are indicated to the right of the time line.

Rappel UHTS

Yamada et al. *Science*, 302, 842-846 , 2003. Empirical Analysis of Transcriptional Activity in the *Arabidopsis* Genome

Arabidospsis : 26828 gènes prédits dont 25540 prédits comme gènes codants

3 classes

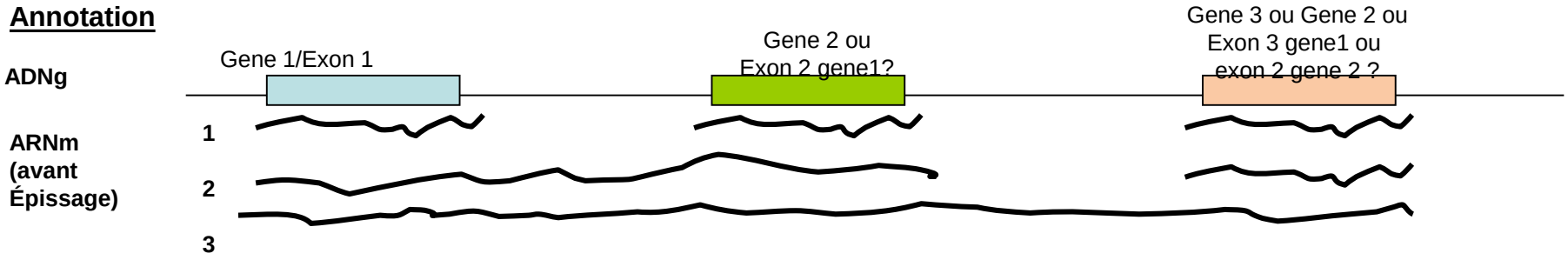
AE
Annotated
expressed
70 % (ESTs)

ANE
Annotated
non
Expressed 30%

NAE
Non
Annotated
expressed ?

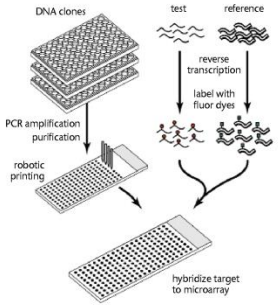
→intérêt des collections d'EST et de cDNA pleine taille (full lenght)

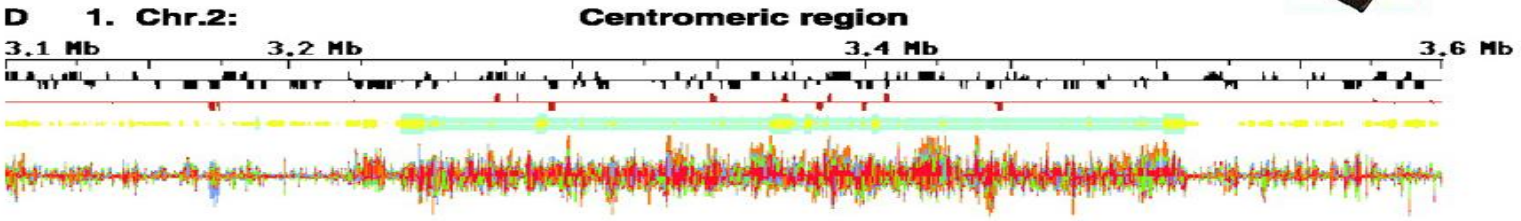
Annotation



Solution ? **High density oligo arrays qui couvre 94 % génome Arabidopsis**

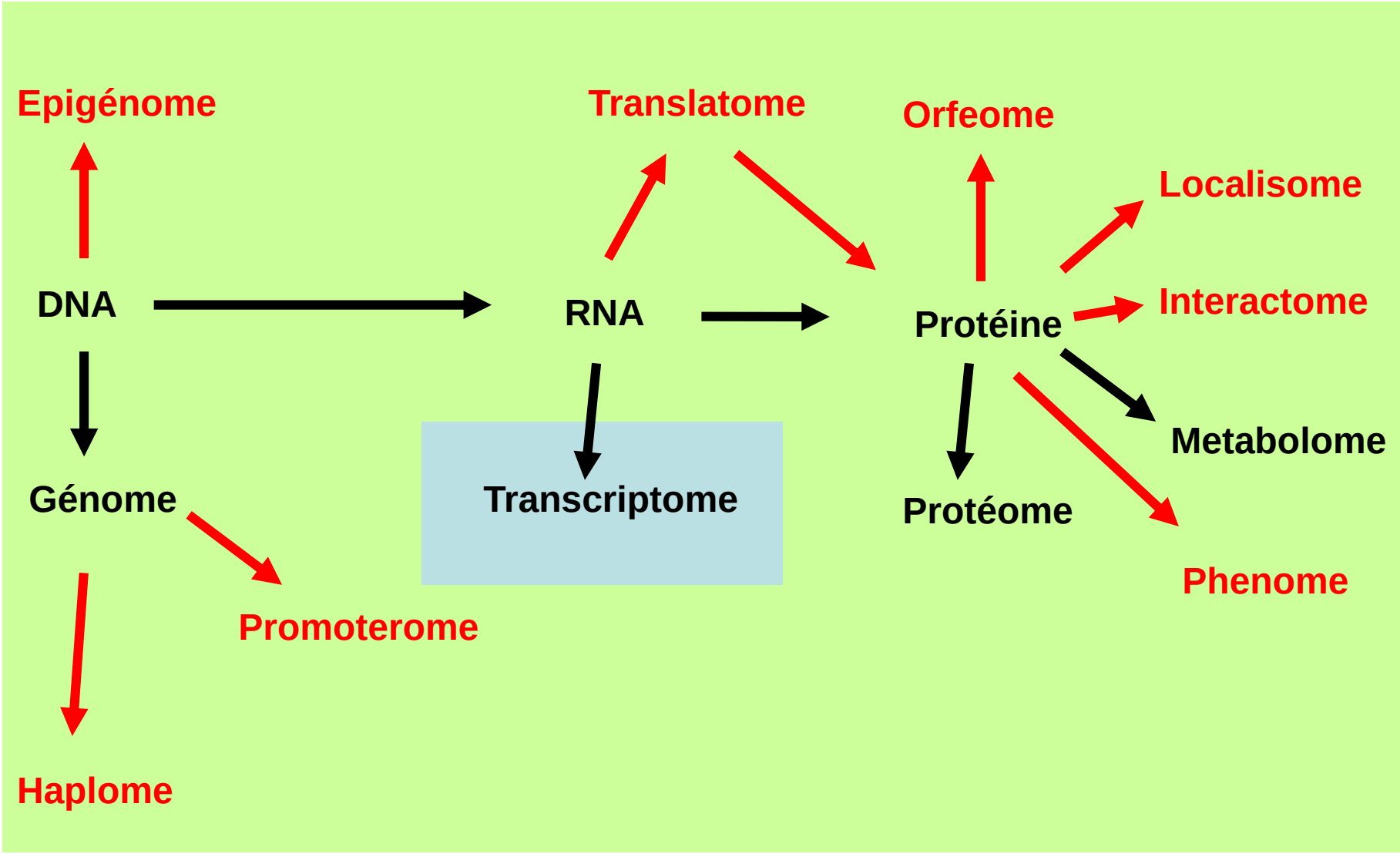
Soit 12 arrays avec 834 000 oligos de 25 mers / arrays
Hybridation avec 4 populations d'ARNm





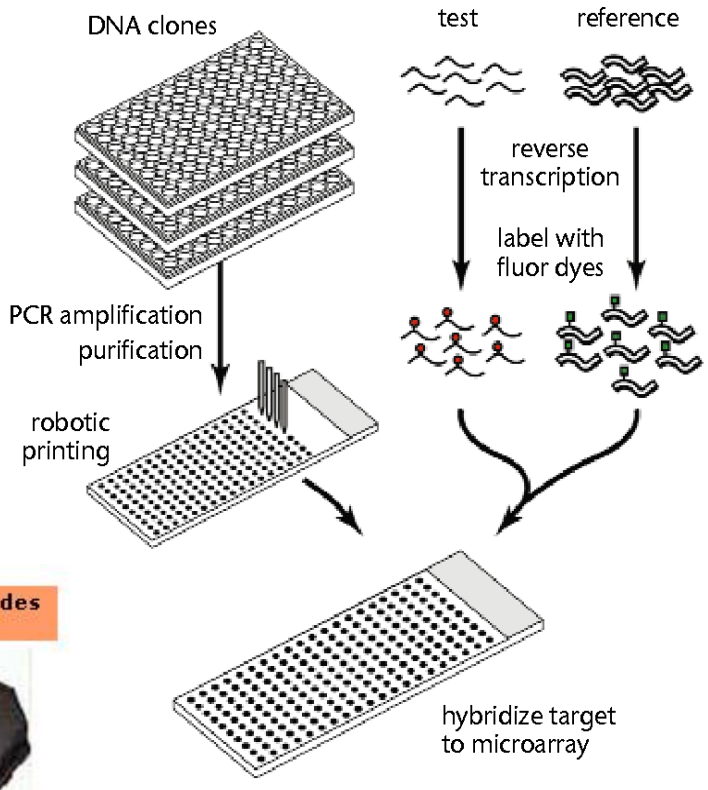
Detection d'activité transcriptionnelle dans région centromérique (40 gènes) et dans régions intergéniques (2000)

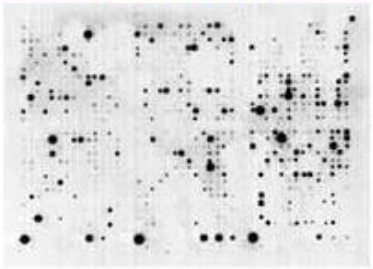
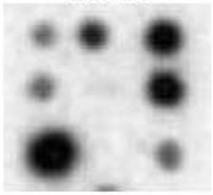
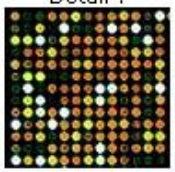
Bilan : 5817 nouvelles unités de transcription soit 30 % de plus que prédit par annotation

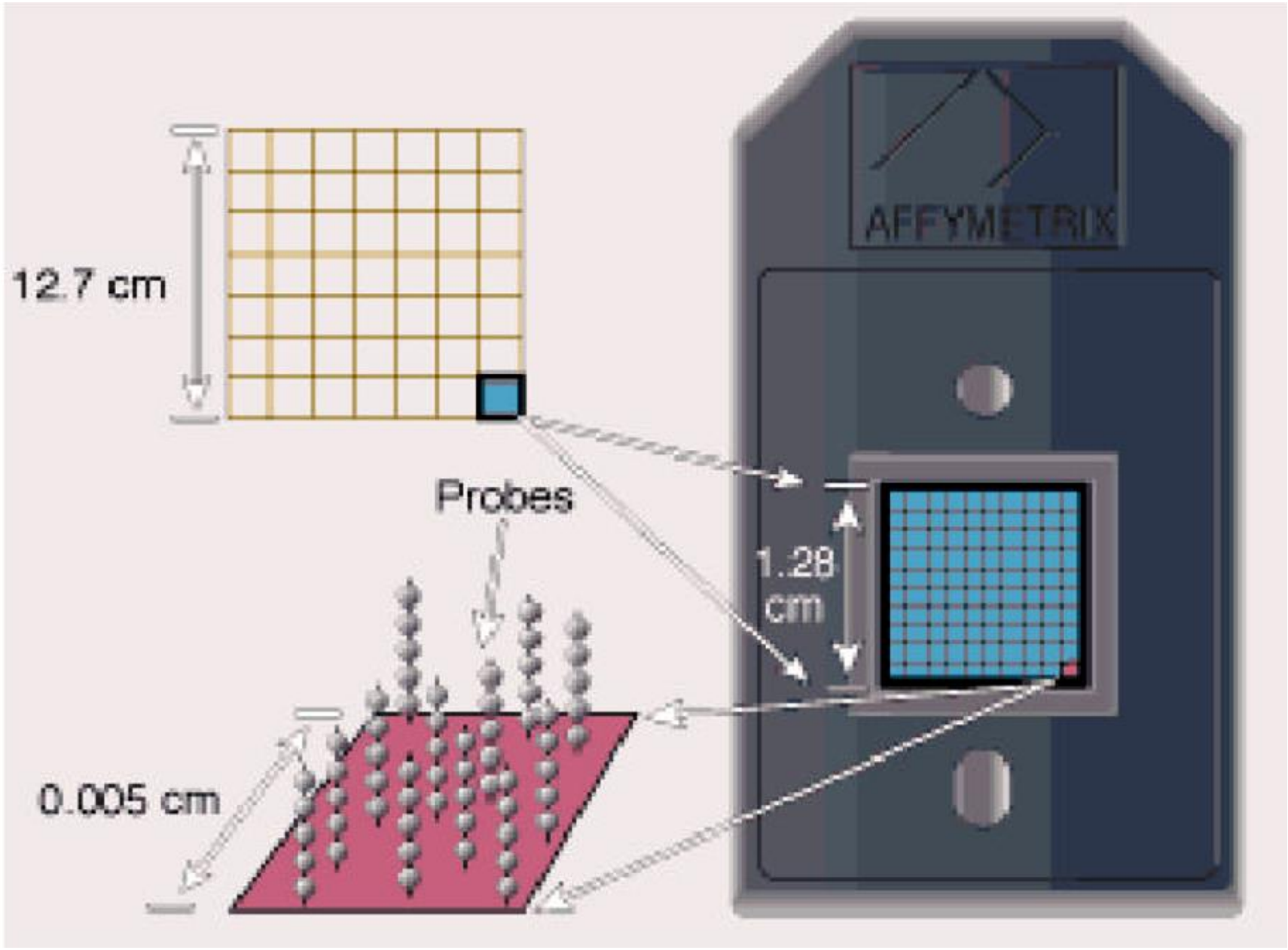

Epigénome
Translatome
Orfeome
Localisome
DNA
RNA
Protéine
Interactome
Transcriptome
Metabolome
Génome
Protéome
Promoterome
Phenome
Haplome

Puces à ADN :

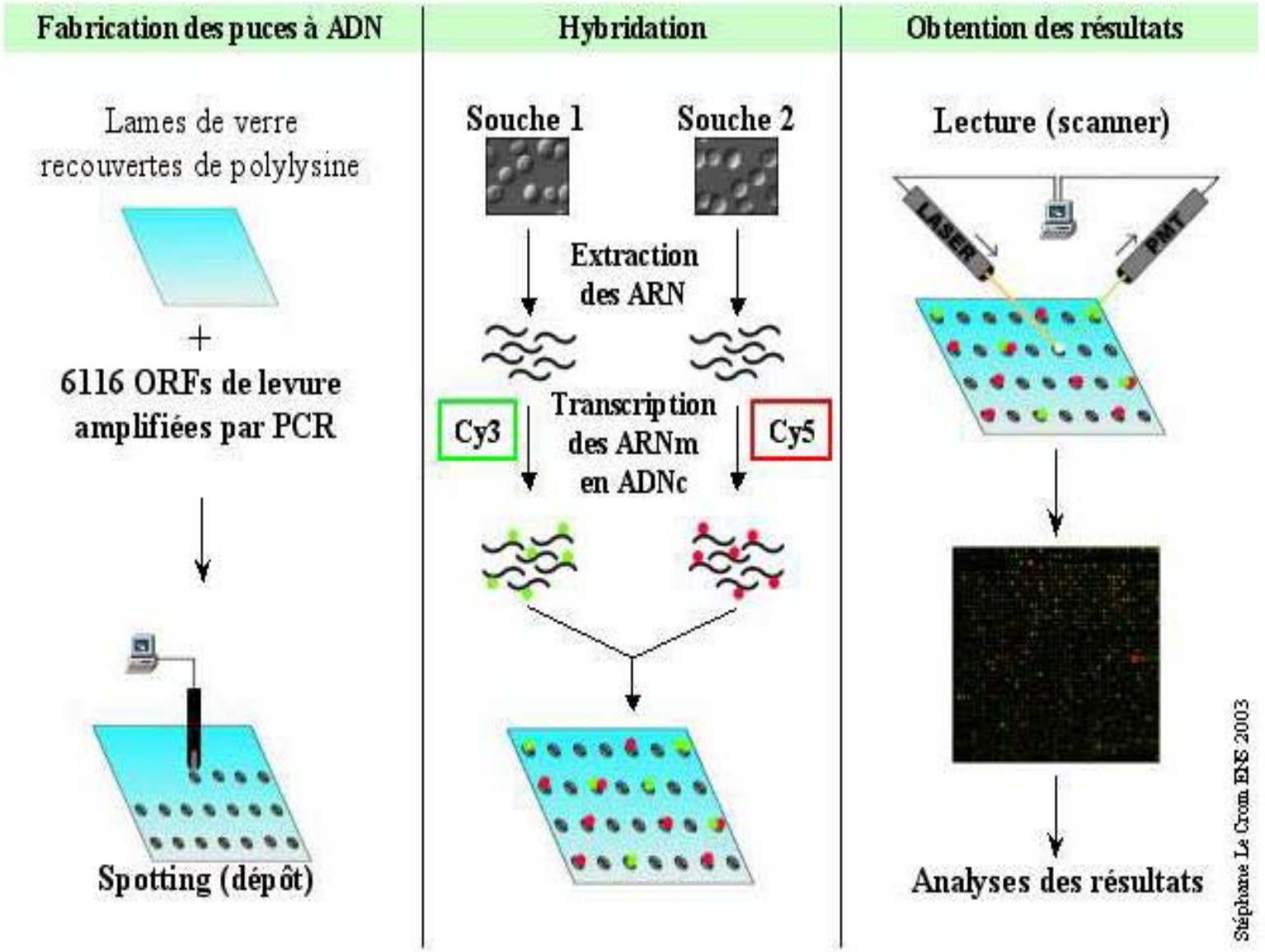
- des cDNAs à la puce ?
- Quels supports ?



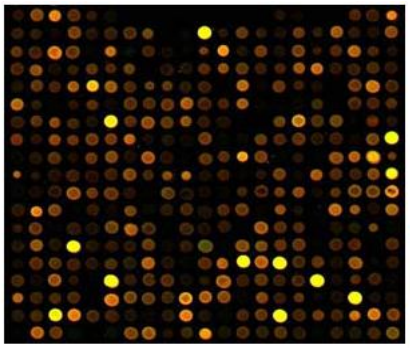
Filtres haute densité (macroarrays)	Lames de verre (microarrays)	Puces à oligonucléotides
		
Détail : 	Détail : 	Détail : 
Taille : 12cm x 8cm	Taille : 5,4cm x 0,9cm	Taille : 1,28cm x 1,28cm
◆ 2400 clones par membrane ◆ marquage radioactif ◆ 1 condition expérimentale par membrane	◆ 10000 clones par lame ◆ marquage fluorescent ◆ 2 conditions expérimentales par lame	◆ 300000 oligonucléotides par lame ◆ marquage fluorescent ◆ 1 condition expérimentale par lame


12.7 cm
Probes
1.28 cm
0.005 cm
AFFYMETRIX

Puces à ADN : Principe



Puces à ADN: du résultat à la signification biologique ?

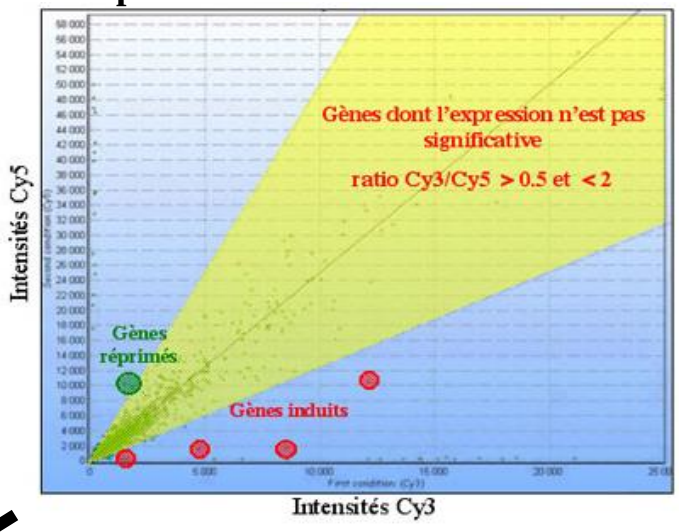


Rouge : Spé A
Vert : Spé B
Jaune : Commun

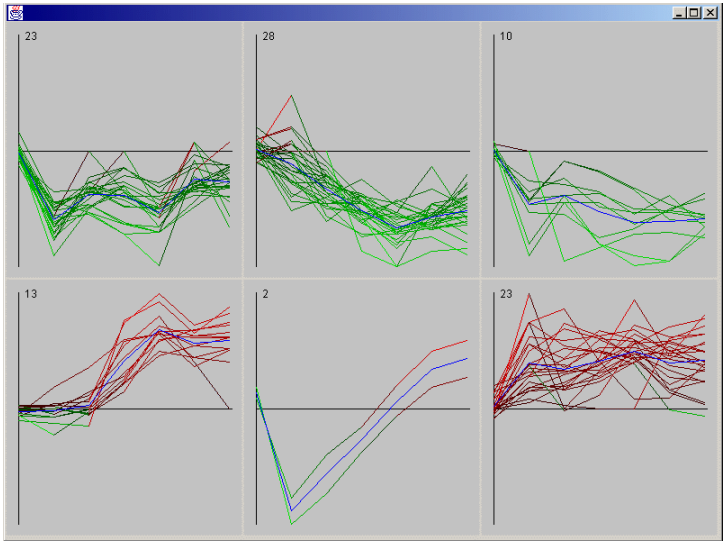
Sonde A

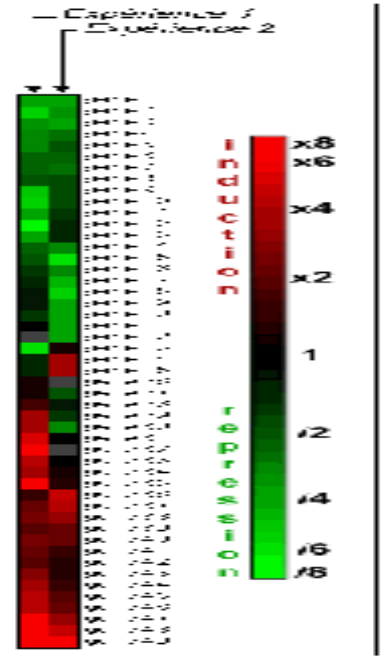
Sonde B

Quels sont les gènes dont l'expression est réprimée ? Induite ? Invariable ?



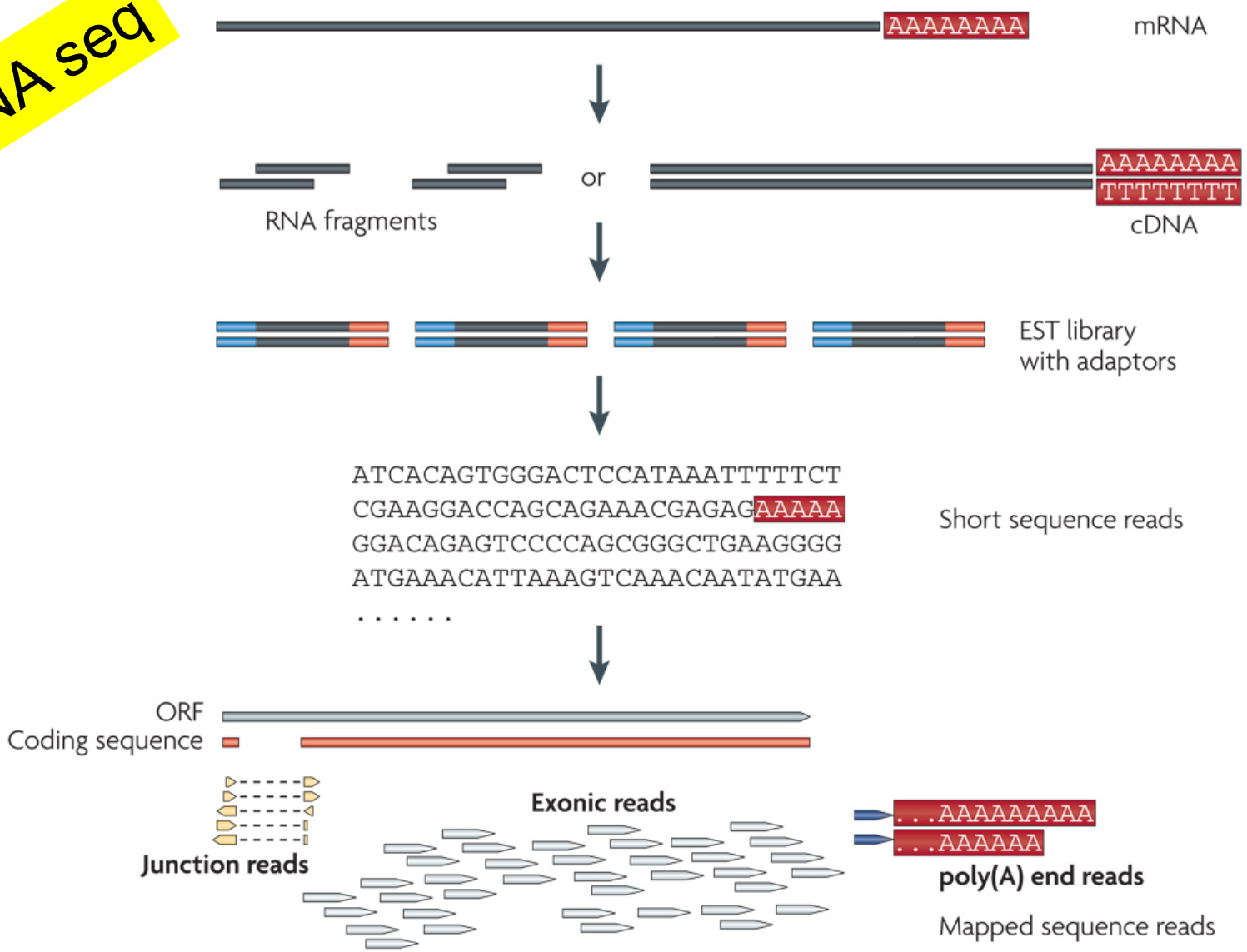
Quels sont les gènes dont le profil d'expression est identique ?





Le clustering d'expression

RNA seq


AAAAAAAA
mRNA
or
RNA fragments
AAAAAAAA
TTTTTTTT
cDNA
EST library
with adaptors
ATCACAGTGGGACTCCATAAATTTTTCT
CGAAGGACCAGCAGAAACGAGAGAAAAA
GGACAGAGTCCCCAGCGGGCTGAAGGGG
ATGAAACATTAAAGTCAAACAATATGAA
. . . . . .
Short sequence reads
ORF
Coding sequence
Junction reads
Exonic reads
. . .AAAAAAAAA
. . .AAAAAA
poly(A) end reads
Mapped sequence reads

ChIP on chip

Chromatin ImmunoPrecipitation on Chip

ChIP and sequencing high throughput

utilisées pour repérer des sites de fixation de facteurs de transcription

localiser ces sites et d'étudier les séquences d'ADN correspondantes

1.Liaison covalente in vivo des protéines à l'ADN
2.Extraction de l'ADN
3.Découpage de l'ADN par sonication
4.Sélection des fragments grâce à un anticorps
5.Précipitation des complexes ADN-protéine-anticorps,
6. Séparation du complexe ADN-protéine pour garder ADN (proteinase K)
7. On obtient une collection de fragments d'ADN qui interagissent avec une protéine d'intérêt .

combinaison de la technique de Chromatin Immunoprécipitation avec la méthode des puces à ADN.

8.On utilise une puce à ADN pour identifier les fragments

8.On séquence (illumina)… pour identifier les fragments

Nécessité de valider données haut débit :

RT_Q-PCR, Northern, lignées promoteur::gène rapporteur…

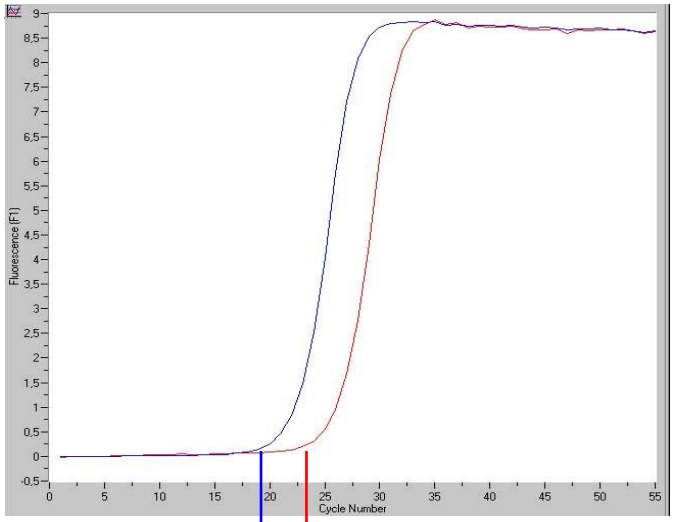
Etude d'expression des gènes : La qPCR

La PCR en temps réel

- fluorochrome, excitation 470 nm, émission 530 nm

- s'intercale dans ADN db en cours de synthèse mesure à la fin de l'étape d'élongation de chaque cycle : fluo proportionnelle à quantité d'ADN db formé

fluo



← *phase stat* : limitation (primers)

nbre de cycles

Ct : Crossing Threshold / point *:* moment ou la fluo passe au dessus du seuil (BDF)

Validation données Micro-arrays

→QPCR
→ Fluidigm Technology

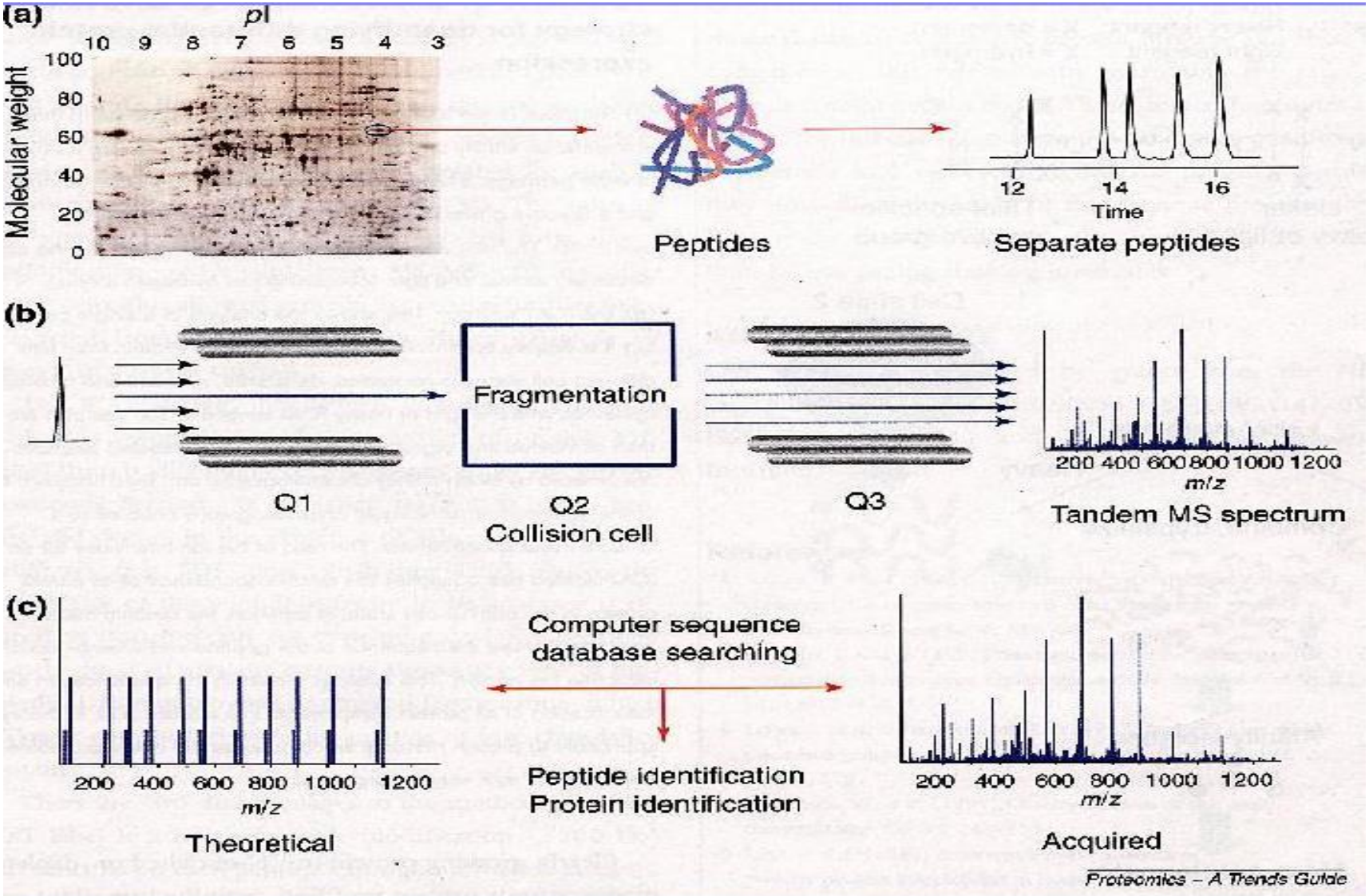
Système microfluidique

Echantillons d'ADNc (jusqu'à 96)

Amorces PCR (jusqu'à 96)

Soit 9216 réactions PCR suivies en temps réel

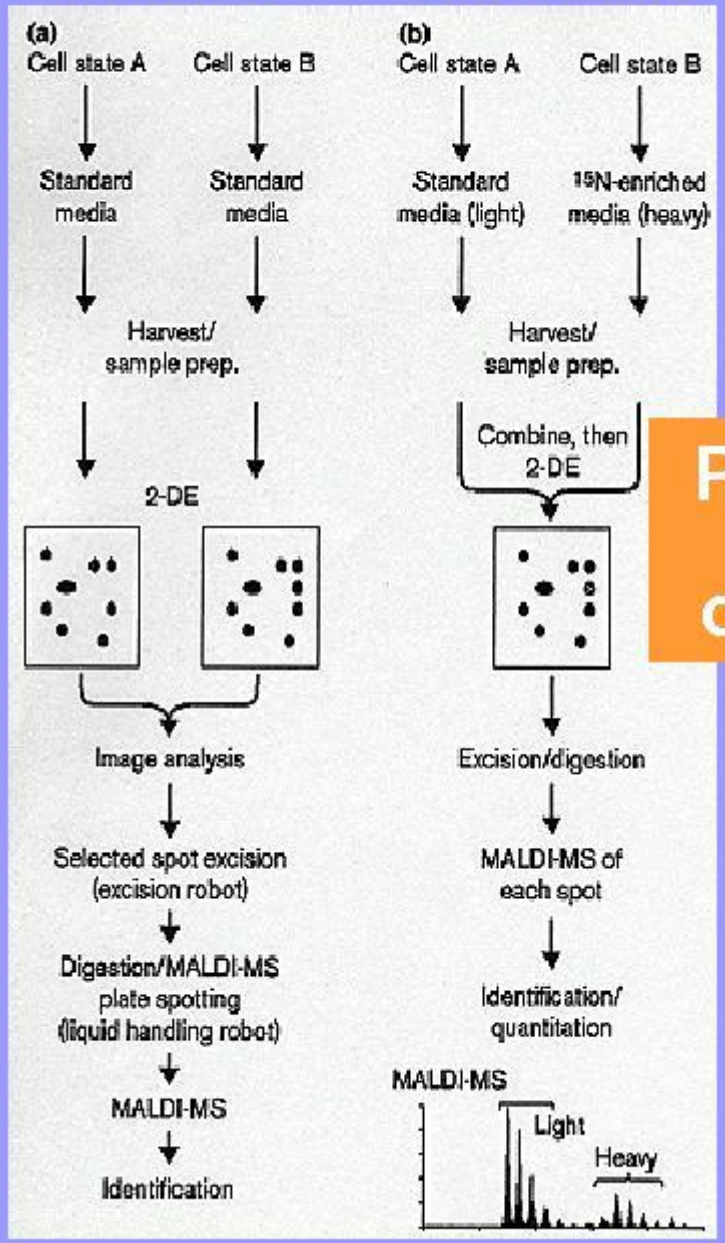
La spectrométrie de masse ...



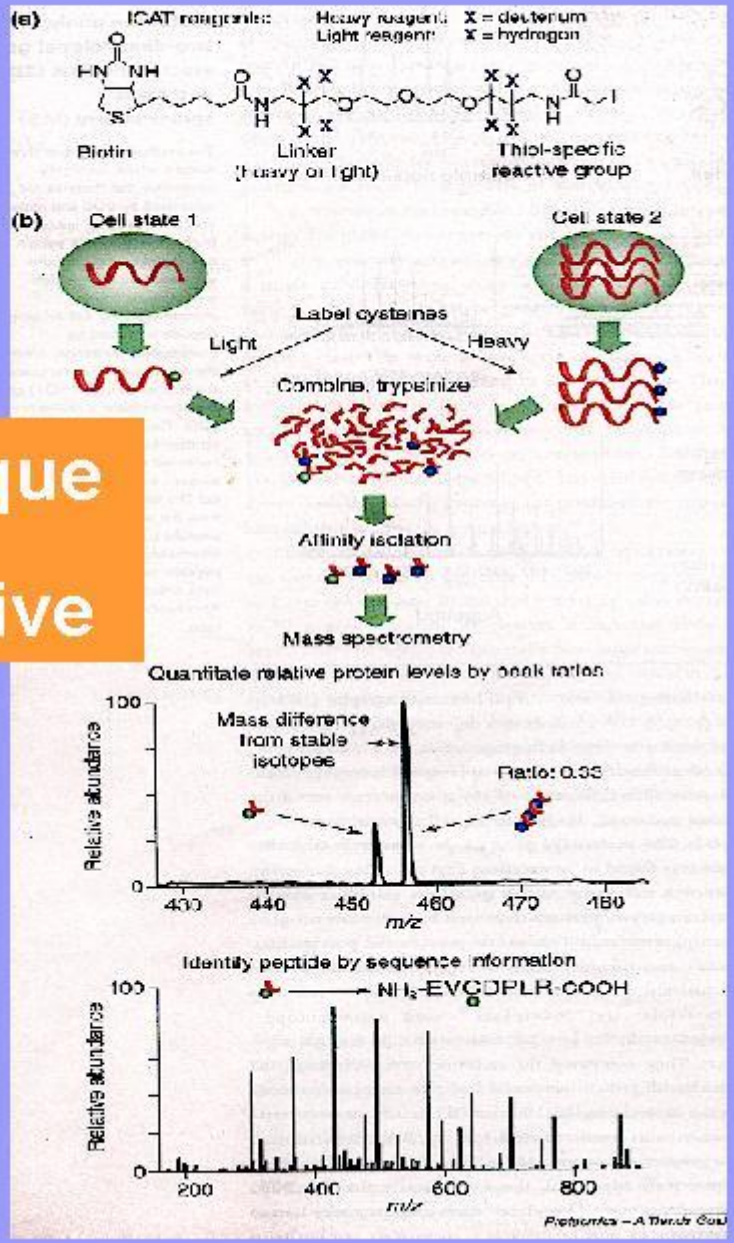
Protéomique comparative

Méthode "standard"

Marquage différentiel

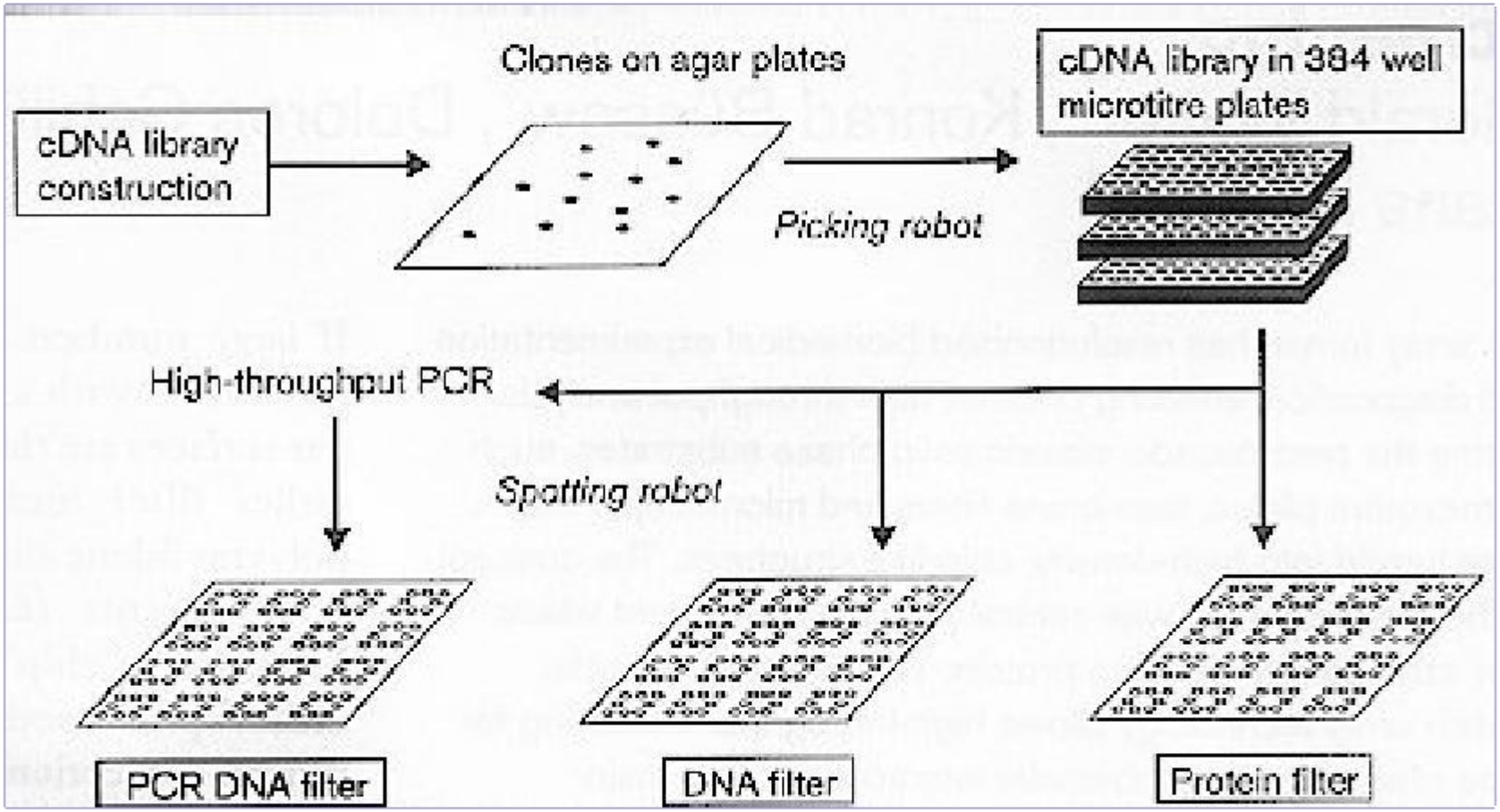


Isotope-Coded Affinity Tag (ICAT)



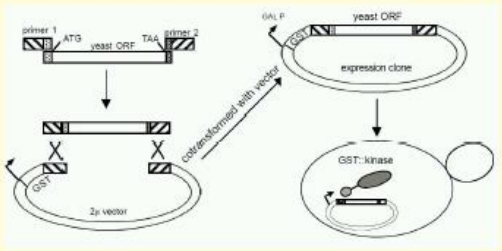
Des approches à haut débit pour la recherche de partenaires protéiques ?

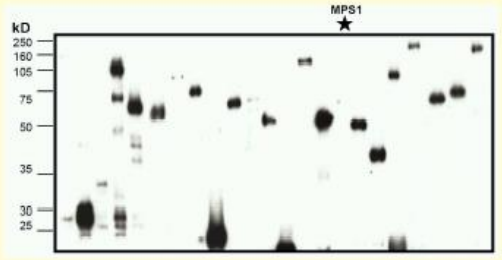
Protein arrays

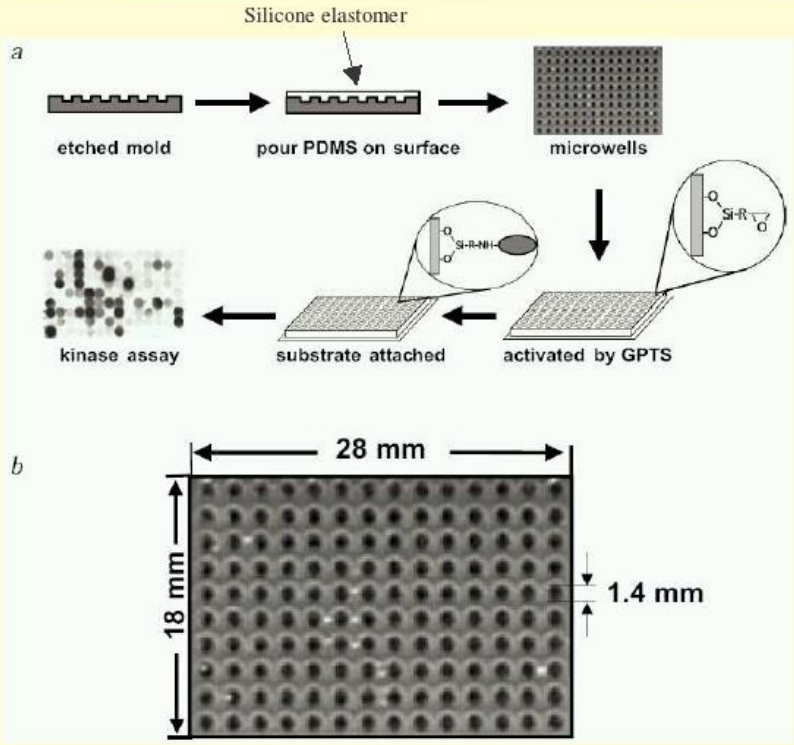


Ex: utilisation de protein chips pour la recherche de substrat d'une kinase

Analysis using protein chips







- Very good signal to noise ratio (10x better than microtiter plate)
- small amount of material needed (1/20th of what is needed for a 384-well plate
- assays are extremely sensitive (even proteins not detectable on immunblot can be assayed)
- inexpensive
- widely applicable

Nat Genet. 2000 Nov;26(3):283-9.
Analysis of yeast protein kinases using protein chips.
Zhu H, Klemic JF, Chang S, Bertone P, Casamayor A, Klemic KG, Smith D, Gerstein M, Reed MA, Snyder M.

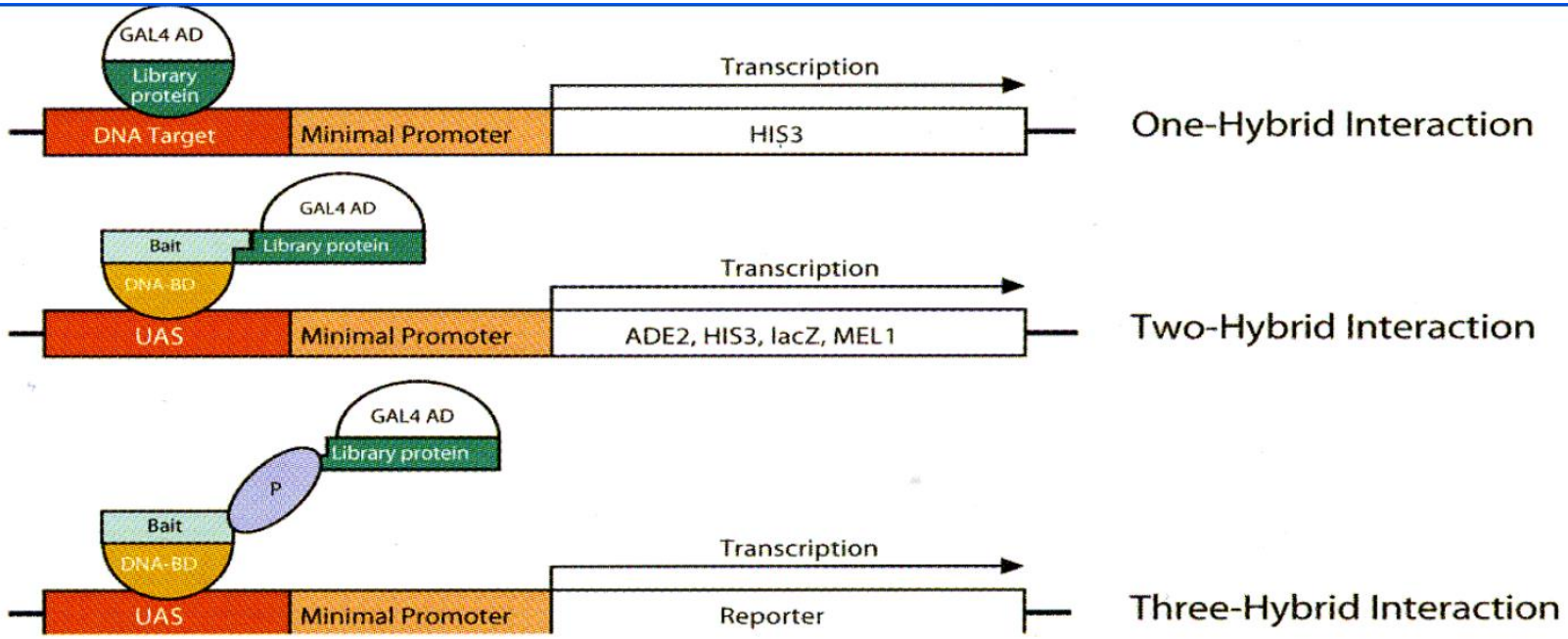


Figure 1. Yeast hybrid technology. Matchmaker systems use a sensitive transcriptional assay to detect one-hybrid, two-hybrid, and three-hybrid interactions. A yeast reporter strain transformed with both a bait and library plasmid will express the plasmids' gene inserts as fusions to either the GAL4 DNA-BD (DNA-binding domain) or AD (transcription-activating domain), depending on the plasmid (see below). If a library protein interacts with a bait protein (two-hybrid) or DNA target (one-hybrid) the host strain actively expresses the reporter gene located downstream of the promoter. Three-hybrid interactions, or protein interactions that occur via a third protein (P), can be detected using our pBridge Vector (Cat. No. 630404). UAS = GAL4-responsive upstream activating sequence.

Annu. Rev. Plant Biol. 2003. 54:669–89
doi: 10.1146/annurev.arplant.54.031902.135014

METABOLOMICS IN SYSTEMS BIOLOGY

Wolfram Weckwerth
Max-Planck-Institut für Molekulare Pflanzenphysiologie, 14424 Potsdam, Germany;
email: weckwerth@mpimp-golm.mpg.de

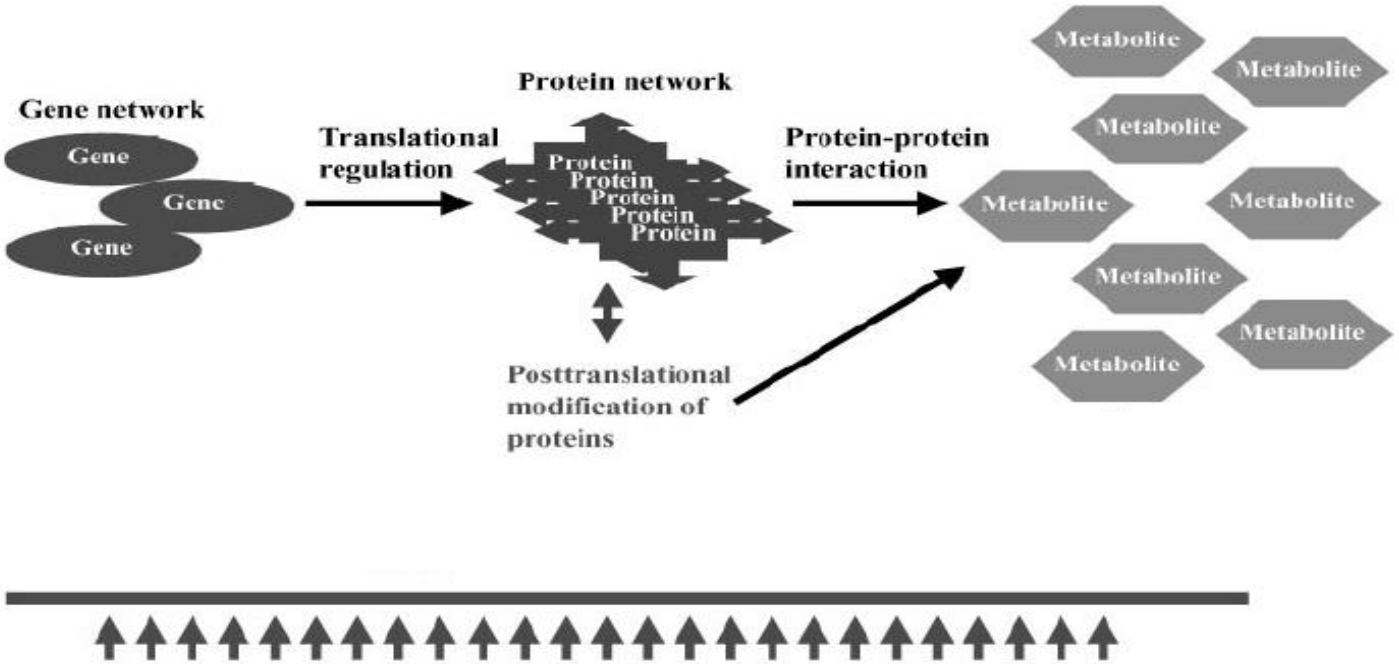


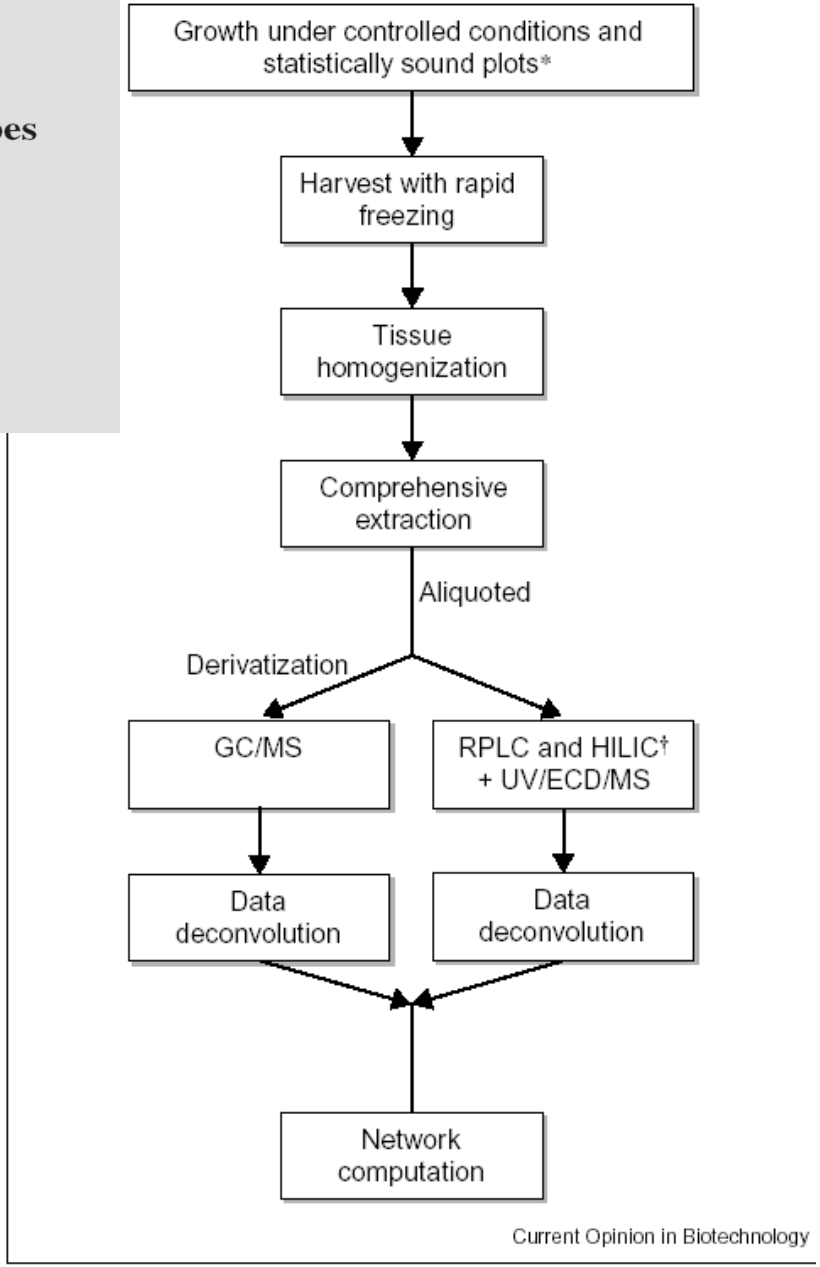
Figure 1 Amplification of a metabolic network and feedback regulation in response to developmental and environmental conditions.

Plant Molecular Biology **48:** 155–171, 2002.

Metabolomics – the link between genotypes and phenotypes

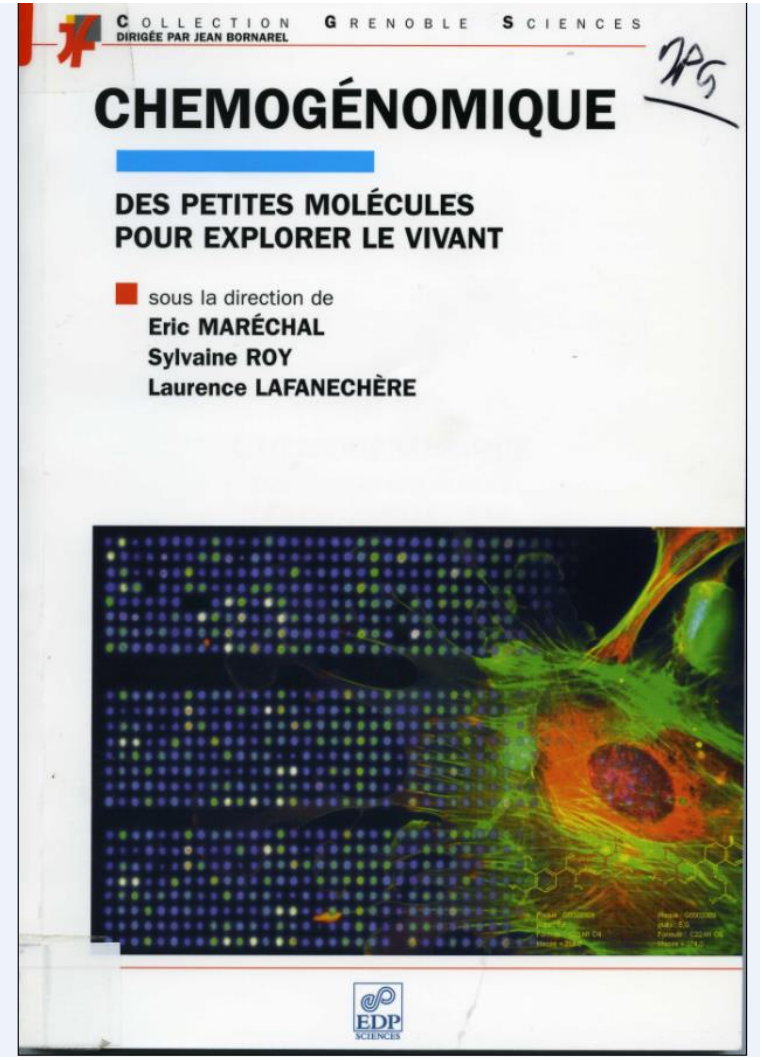
Oliver Fiehn
Max-Planck Institute of Molecular Plant Physiology, 14424 Potsdam, Germany
(e-mail fiehn@mpimp-golm.mpg.de)

Key words: functional genomics, mass spectrometry, metabolism, metabolite profiling



Proposed scheme for comprehensive metabolomic data acquisition.

La chémogénomique



= criblage à haut débit de molécules naturelles ou de synthèses pour découvrir de nouvelles cibles thérapeutiques et de nouveaux médicaments

…

Ou de nouvelles molécules interférant avec Fonctions cellulaires ou des molécules d'intérêt en agronomie et microbiologie (pesticides, herbicides, etc …)

- Les chimiothèques

- les criblages de molécules et la mesure de l'effet biologique

- Intérêt en post-génomique

Les chimiothèques :

La chimiothèque nationale … … … en attendant la chimiothèque européenne

http://chimiotheque-nationale.enscm.fr

Regroupe les collections des labos français

Mise à jour le 8 Septembre 2010

Composés repertoriés : **44093**
Composés en plaque : **32267**
Extraits naturels : **12838**

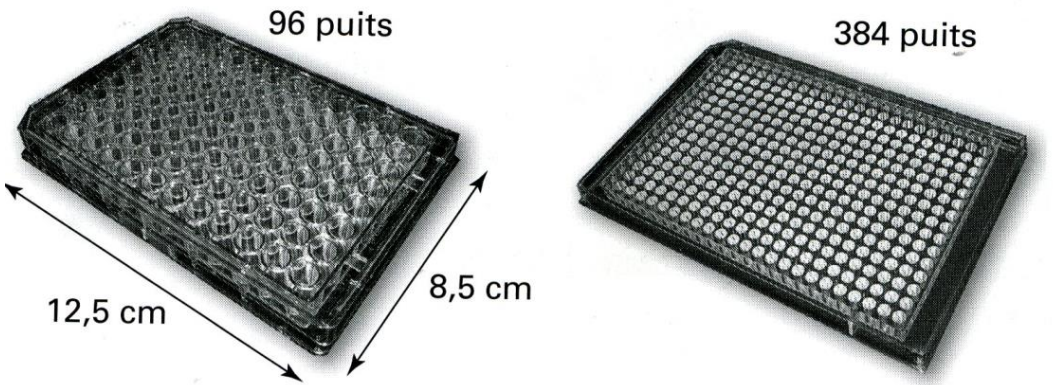
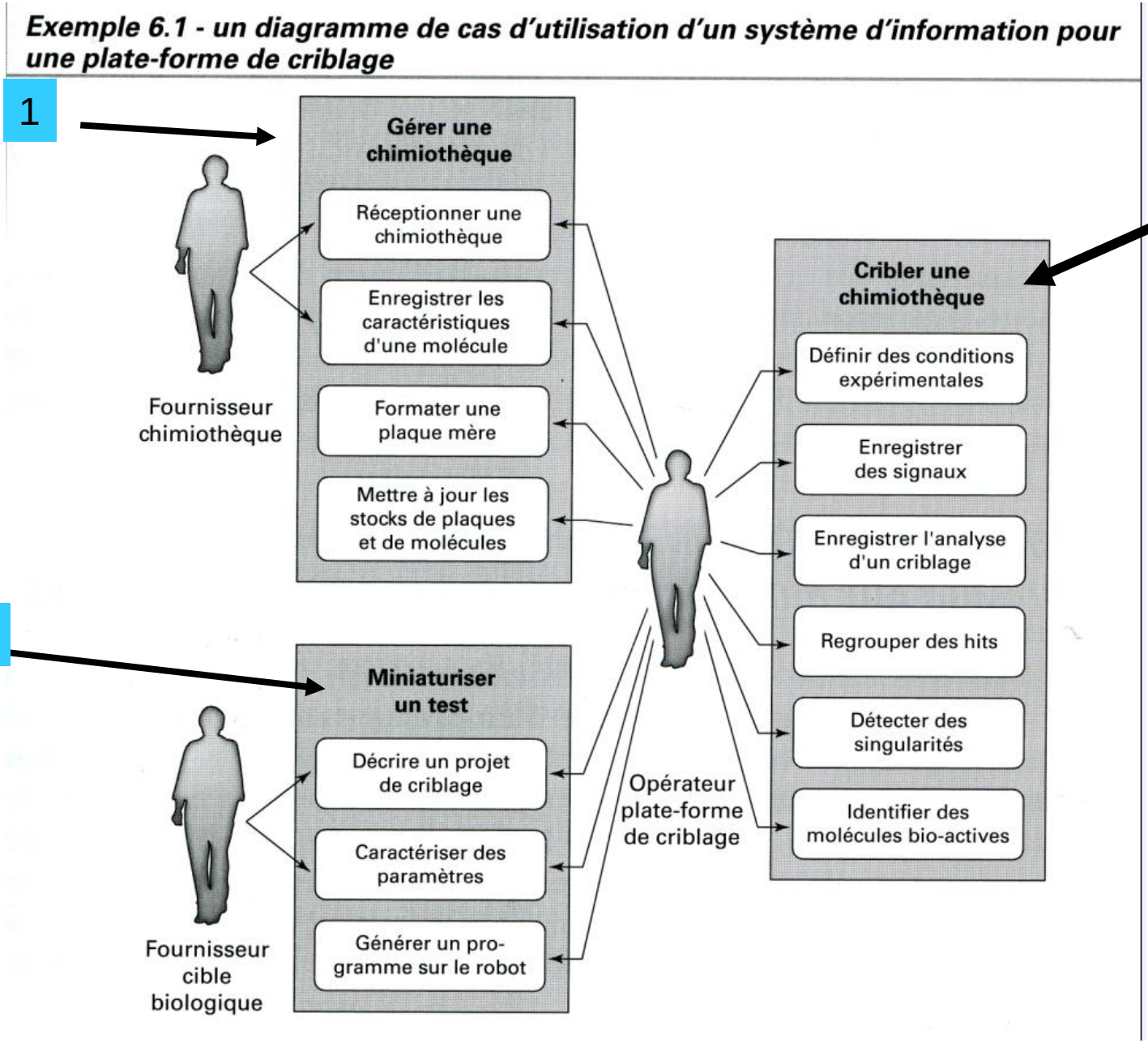
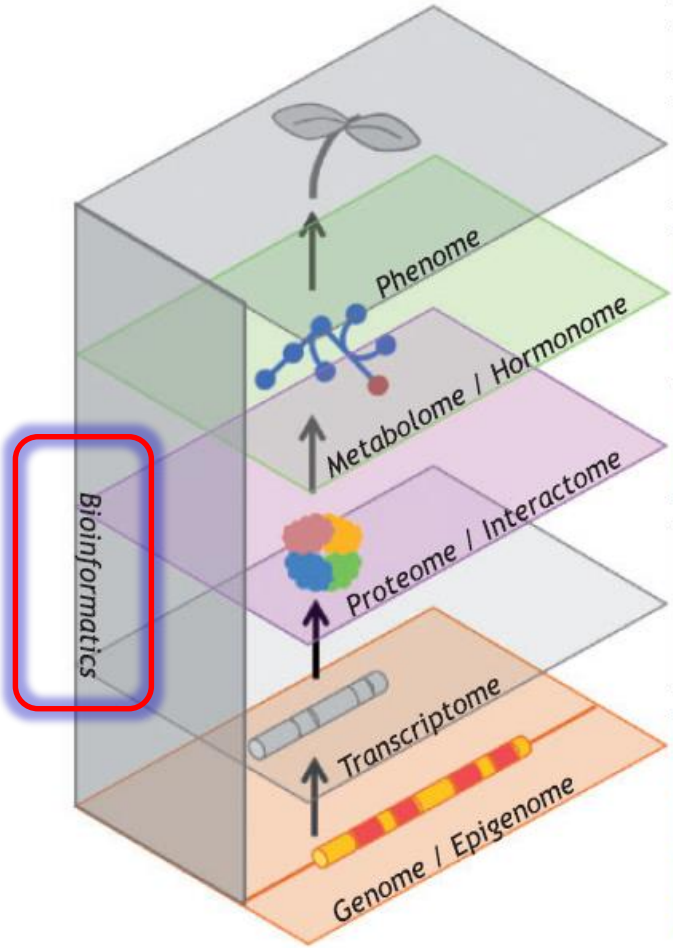


Fig. 1.2 - Plaques multipuits de 96 et 384 puits. (85 x 125 mm)

Substances naturelles (ex : Pierre Fabre)
Chimie combinatoire : synthèses rapide de composés à partir d'1 châssis moléculaire donné

La stratégie





Omics instances	Resources
Integrated database	TAIR[1]
Mutant lines	FOX line[2], Ac/Ds tag line[3], T-DNA tag line[4], TILLING[5]
Natural variations	NASC[6], ABRC[7]
Metabolic map	PMN AraCyc[8], Reactome[9]
Metabolome profiles	PRIMe[10], Golm Metabolome Database[11]
Hormonome profile	RiceFOX DB[2]
Proteome / modificome profiles	RIPP-DB[12], PPDB[13], PhosPhAt[14]
Subcellular localization	PODB2[15], SUBAII[16], NASC proteome database[17]
Interactome maps	TAIR[1], AtPID[18], Plant Interactome Database[19]
Full-length cDNA clones, ESTs	RAFL clones[20], RARGE[21]
Expression profiles	AtGenExpress[22], Genevestigator[23]
Non-coding RNA profiles	Arabidopsis MPSS[24]
Co-expression network	ATTEDII[25]
Genome sequence, gene annotation	TAIR[1]
Re-sequencing	Arabidopsis 1001 genome project[26]
Focused gene family database (eg. Transcription factor)	RARTF[27], AGRIS[28], DATF[29]
DNA methylome	SIGnAL[30]
Chromatin epigenome	EPIC web site[31]

Fig. 1 An updated omic space with emerging omics layers: epigenome, interactome and hormonome added to each of the closely related layers with illustrative resources for *Arabidopsis* available on the web. [1]http://www.arabidopsis.org/, [2]http://ricefox.psc.riken.jp/, [3]http://rarge.gsc.riken

K. Mochida and K. Shinozaki
Plant Cell Physiol. 52(12): 2017–2038 (2011)