

Mercredi 17 novembre 2004

Introduction à la génomique

***Odile Kalogeropoulos
Université Pierre et Marie Curie
Institut Pasteur***

Conférences du cours

Analyse des génomes

Liste des conférences (Début)

Vendredi 19 novembre

14:30 - 16:00

Protéomique d'expression : méthodes, performances et limites

Thierry RABILLOUD

PROTEOMIQUE

16:15 - 17:45

Génomique structurale et haut débit

Arnaud DUCRUIX

PROTEOMIQUE

Lundi 29 novembre

9:00 - 10:30

L'évolution de la pathogénicité des Mycobactéries :
les apports de la génomique comparative

Stewart COLE

GENOMIQUE COMPARATIVE

10:45 - 12 :15 Organisation structurelle des génomes
et étude de la dynamique de réplication :
une approche sur molécule individuelle d'ADN

Aaron BENSIMON

PHYSIQUE DE L'ADN

Jeudi 2 décembre

9:00 - 10:30

Dynamique cellulaire et approches génomiques

Claude JACQ

TRANSCRIPTOME

10:45 - 12 :15

Archaea et évolution des génomes

Patrick FORTERRE

EVOLUTION

Vendredi 3 décembre

9:00 - 10:30

ADN répété et contrôle épigénétique de l'activité du génome :
le modèle Arabidopsis thaliana

Vincent COLOT

EXPRESSION GENIQUE

10:45 - 12 :15

L'exploration des déficits sensoriels auditifs chez l'homme
et la souris

M.C. SIMMLER

APPLICATION MEDICALE

Liste des conférences (suite)

Lundi 6 décembre

9:00 - 11:00 Le projet génome d'*Anopheles gambiae*

Paul BREY - Karin EIGLMEIER
Isabelle ROSINSKI-CHUPIN

GENOMIQUE STRUCTURALE

Mardi 7 décembre

9:00 - 10:30 Approches génomiques de la biologie des algues
pluricellulaires marines

Bernard KLOAREG

GENOMIQUE STRUCTURALE

10:45 - 12:15 Mécanismes d'évolution du Génome chez les
Picocyanobactéries marines

Frédéric PARTENSKY

GENOMIQUE COMPARATIVE

Mercredi 8 décembre

9:00 - 10:30 Analyse multi-échelle des génomes

Claude THERMES

PHYSIQUE DE L'ADN

10:45 - 12:15 Physique de l'ADN et analyses génomiques

Edouard YERAMIAN

PHYSIQUE DE L'ADN

Jeudi 9 décembre

9:00 - 10:30 Colonisation du génome humain par de l'ADN mitochondrial :
un phénomène en cours

Miria RICCHETTIS

STRUCTURE DE L'ADN

10:45 - 12:15 Application des polymorphismes humains pour l'étude
des retards mentaux de l'enfant

Laurence COLLEAUX

APPLICATION MEDICALE

Vendredi 10 décembre

9:00 - 10:30 Les virus à ARN emprisonnés par leurs génomes

Simon WAIN-HOBSON

GENOMIQUE STRUCTURALE

10:45 - 12:15 Vue des cellules comme des ordinateurs construisant
des ordinateurs

Antoine DANCHIN

REPLICATION DE L'ADN

Comment en est-on arrivé à cette
diversité de sujets abordés par
l'approche génomique ?

Projets génomes

A la date du 16 novembre 2004 :

Génomes séquencés publiés :

- **229 génomes Procaryotes**
- **36 génomes Eucaryotes**

En cours :

- **530 projets de Procaryotes**
- **461 Projets d'Eucaryotes**

QuickTime™ et un
décompresseur TIFF (non compressé)
sont requis pour visionner cette image.

GenomesOnLine
www.genomesonline.org

QuickTime™ et un
décompresseur TIFF (non compressé)
sont requis pour visionner cette image.

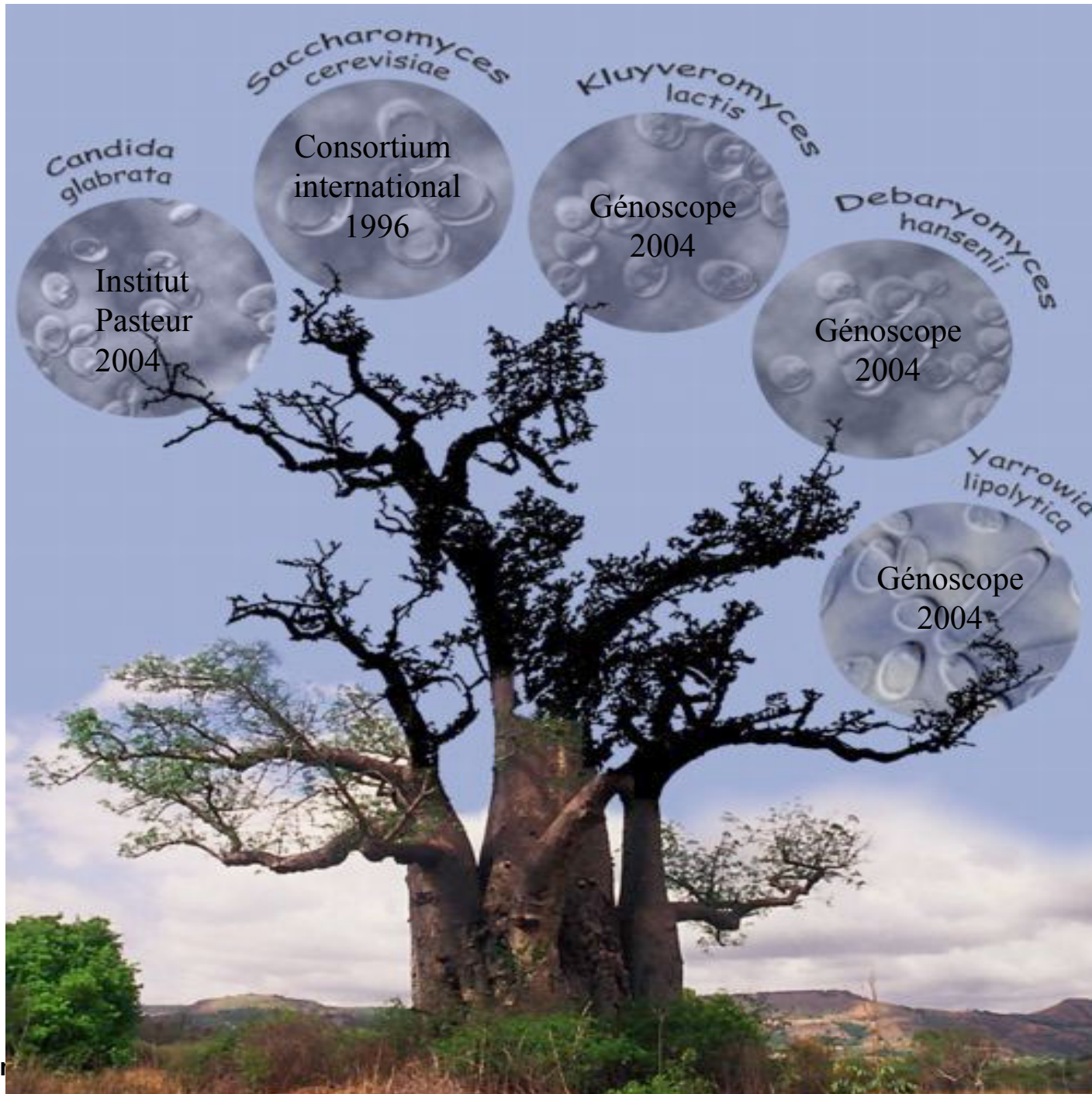
Génomes eucaryotes séquencés

QuickTime™ et un
décompresseur TIFF (non compressé)
sont requis pour visionner cette image.

Liste des génomes eucaryotes publiés

Vertébrés		
Mammifères		
<i>Bos taurus</i> (cow)	Bovin	
<i>Canis familiaris</i> (dog)	Chien	Complet
<i>Felis catus</i> (cat)	Chat	
<i>Homo sapiens</i> (human)	Homme	Complet
<i>Mus musculus</i> (mouse)	Souris	Complet
<i>Ovis aries</i> (sheep)	Mouton	
<i>Rattus norvegicus</i> (rat)	Rat	Complet
<i>Sus scrofa</i> (pig)	Porc	
Autres vertébrés		
<i>Danio rerio</i> (zebrafish)	Poisson	Complet
<i>Ciona intestinalis</i> (sea squirt)	Chordé marin (aspect tube)	Complet
<i>Gallus gallus</i> (chicken)	Poulet	
Invertébrés		
Insectes		
<i>Anopheles gambiae</i> (mosquito)	Moustique	Complet
<i>Apis mellifera</i> (honey bee)	Abeille	
<i>Drosophila melanogaster</i> (fruit fly)	Drosophile	Complet
<i>Bombyx mori</i> (silkworm)		
p50T		
Nématode		
<i>Caenorhabditis elegans</i> (nematode)	Ver	Complet
Protozoaire		
<i>Plasmodium falciparum</i>		Complet
Plantes (liste partielle)		
<i>Arabidopsis thaliana</i> (thale cress)	Arabette	Complet
<i>Avena sativa</i> (oat)	avoine	
<i>Glycine max</i> (soybean)	Soja	
<i>Hordeum vulgare</i> (barley)	Orge	
<i>Lycopersicon esculentum</i> (tomato)	Tomate	
<i>Oryza sativa</i> (rice)	Riz	Complet
<i>Triticum aestivum</i> (wheat)	Blé	
<i>Zea mays</i> (corn)	Maïs	Complet
<i>Cyanidioschyzon merolae</i>	Algue	
Champignons (liste partielle)		
<i>Candida glabrata</i>	Levure	Complet
<i>Debaryomyces hansenii</i>	Levure	Complet
<i>Yarrowia lipolytica</i>	Levure	Complet
<i>Saccharomyces cerevisiae</i> (baker's yeast)	Levure de boulangerie	Complet

QuickTime™ et un
décompresseur TIFF (non compressé)
sont requis pour visionner cette image.



Saccharomyces cerevisiae

Consortium international
1996

Kluyveromyces lactis

Génoscope
2004

Debaryomyces hansenii

Génoscope
2004

Yarrowia lipolytica

Génoscope
2004

Candida glabrata

Institut Pasteur
2004

Immun

En France

La Carte du Réseau national Genopole®





Créé en 1997 sur le site d'Évry, le Genoscope - Centre national de séquençage, dirigé par Jean Weissenbach, constitue l'un des trois Centres nationaux de ressources (avec le Centre national de génotypage et le pôle informatique Infobiogen) autour desquels s'est structuré Genopole®. Aujourd'hui, le campus rassemble autour de la génomique, de la post-génomique et des sciences connexes, des filières complètes d'enseignement supérieur grâce à l'université d'Evry-Val d'Essonne, plus de 1 800 personnes dans vingt-cinq laboratoires de recherche académique et quarante-neuf entreprises innovantes de biotechnologies, implantés sur 71 000 m².
www.genopole.org

Pasteur Genopole® Île-de-France

Pasteur Genopole® Île-de-France regroupe tous les équipements de pointe nécessaires à la réalisation de projets de génomique et de génomique fonctionnelle.

Elle fait partie du Consortium National de Recherche en Génomique (CNRG),
et du Réseau national Genopole®, soutenu par le Ministère Délégué à la Recherche et aux Nouvelles Technologies,
et se spécialise dans l'analyse des microorganismes pathogènes.

Il existe actuellement sept plates-formes à Pasteur Genopole® Île-de-France dont la vocation est de promouvoir la découverte scientifique, notamment dans le domaine du traitement et de la prévention des maladies infectieuses.

- * PF1 : Génomique
- * PF2 : Puces à ADN
- * PF3 : Protéomique
- * PF4 : Intégration et analyse génomiques
- * PF5 : Production de Protéines Recombinantes et Anticorps
- * PF6 : Cristallogénèse et diffraction des rayons X
- * PF7: Synthèse d'Oligonucléotides longs à haut débit

Plate-forme technologique 1 : Génomique



Sequenceurs ABI Prism 3700
Applied Biosystems



Sequenceurs ABI DNA Analyzer 3730
Applied Biosystems



Machines PCR System 9700
Applied Biosystems



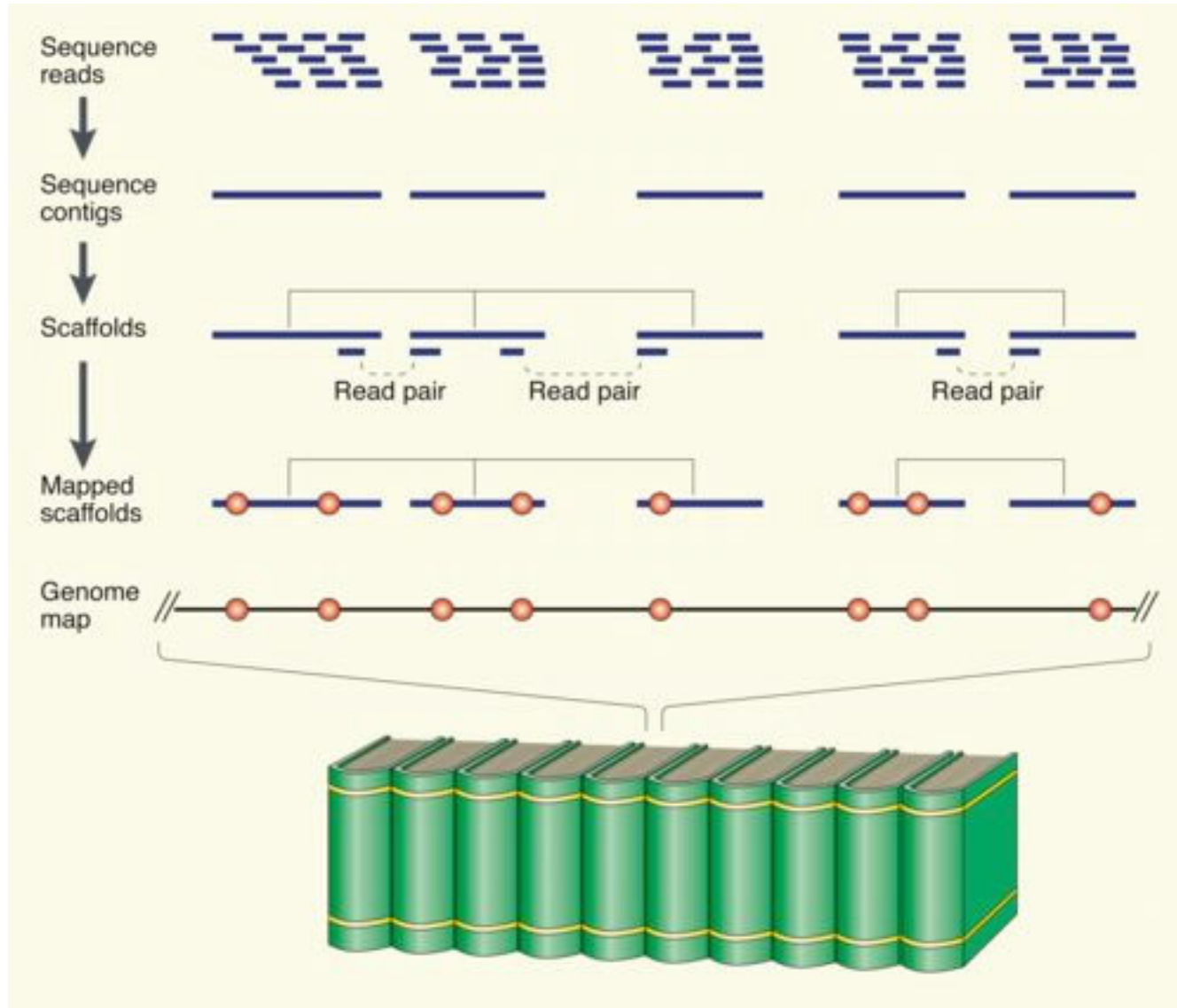


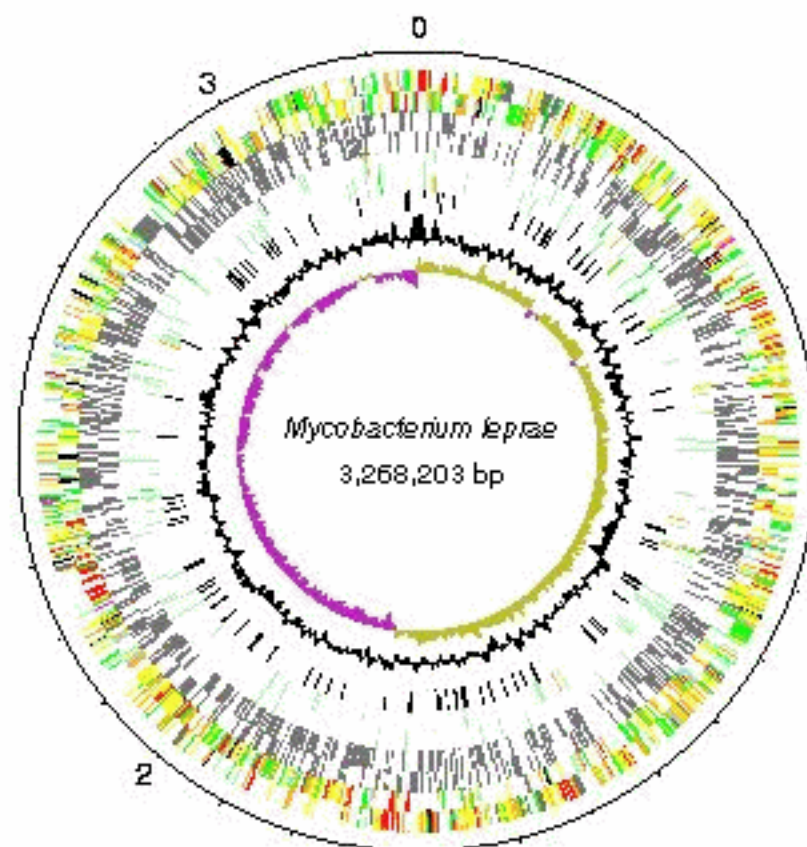
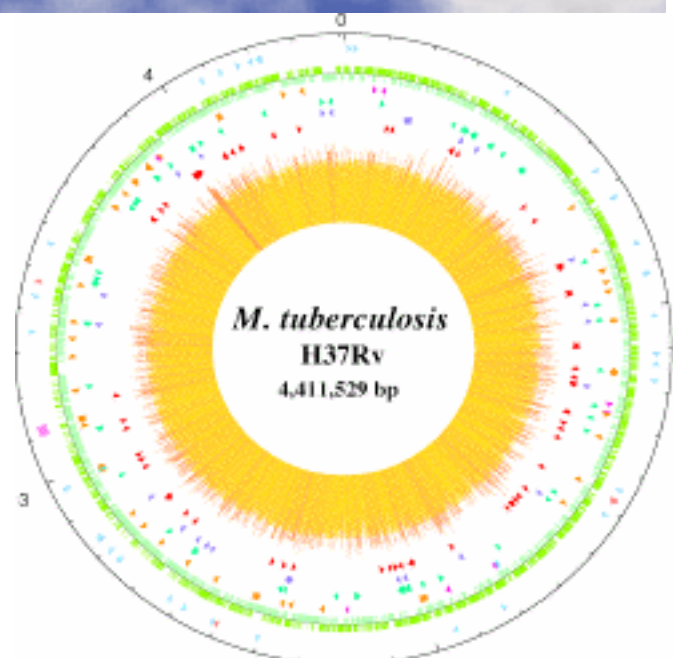
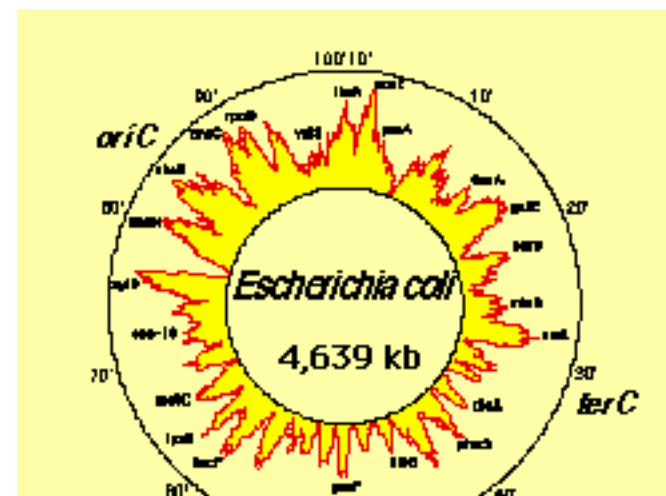
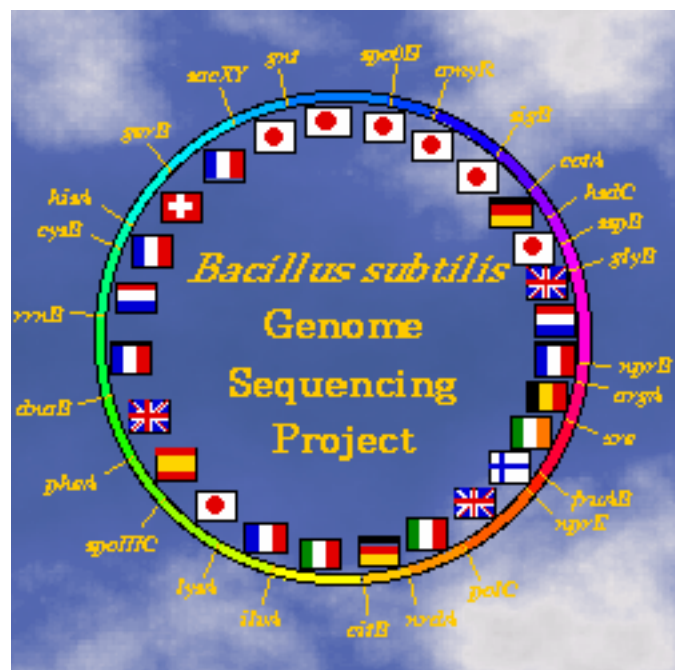
Plate-forme technologique 4 : Intégration et analyse génomiques



Une banque de données locales de l'Institut Pasteur

GenoList genome browser

	<u>SubtiList</u>	<i>Bacillus subtilis</i> strain 168 GenoList browser
	<u>Colibri</u>	<i>Escherichia coli</i> strain K12 GenoList browser
	<u>TubercuList</u>	<i>Mycobacterium tuberculosis</i> strain H37Rv GenoList browser
	<u>Leproma</u>	<i>Mycobacterium leprae</i> strain TN GenoList browser
	<u>BoviList</u>	<i>Mycobacterium bovis</i> strain AF2122/97 GenoList browser
	<u>ListiList</u>	<i>Listeria monocytogenes</i> strain EGD-e and <i>Listeria innocua</i> strain CLIP 11262 Multi-GenoList browser
	<u>LegioList</u>	<i>Legionella pneumophila</i> Paris strain and <i>Legionella pneumophila</i> Lens strain Multi-GenoList browser
		



Annotation des génomes

Projet Génolevures



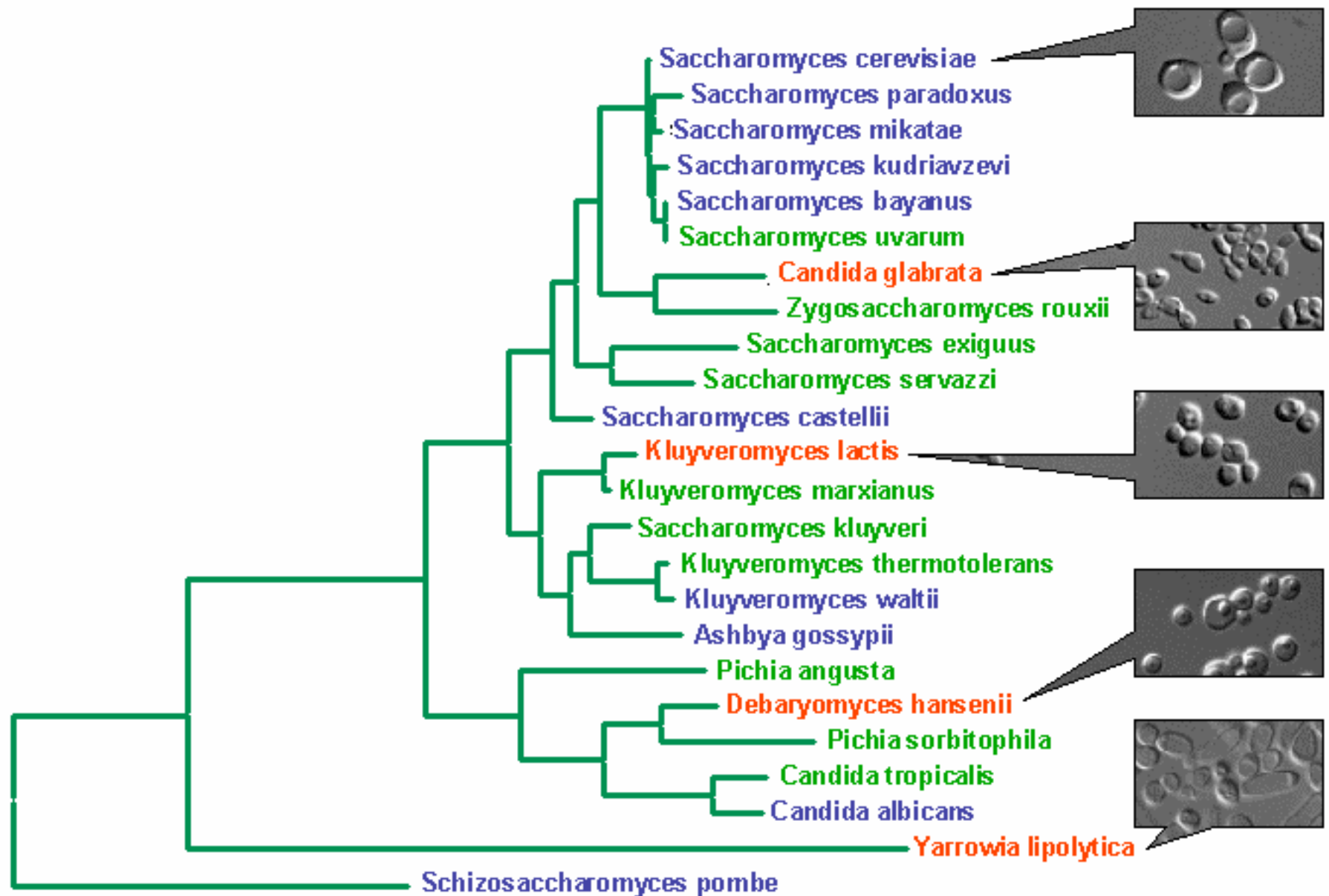


Plate-forme Technologique 2 : Puces à ADN

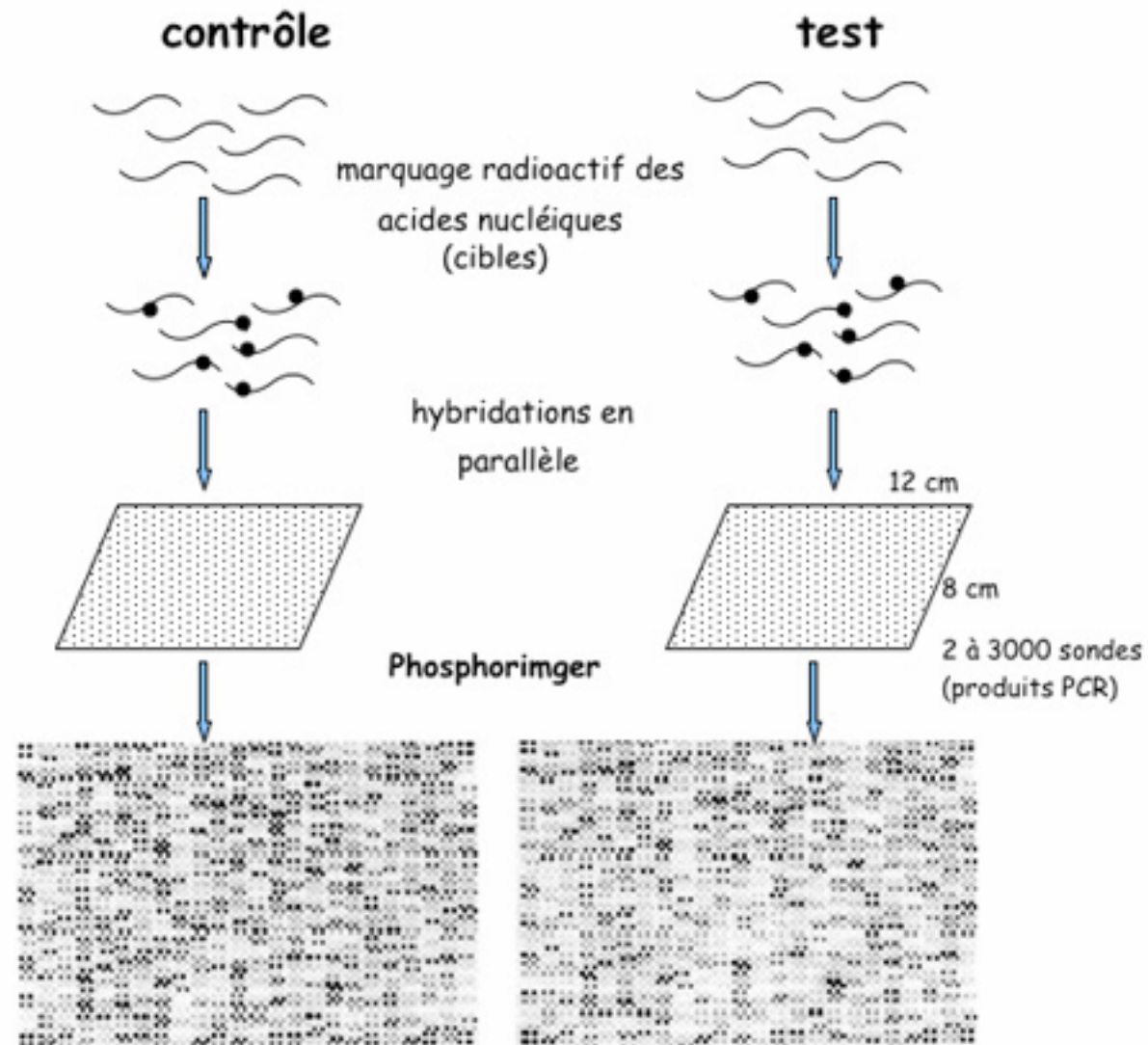


Le principe des puces à ADN est basé sur la technique d'hybridation

Immobilisées sur un support solide (**matrice**), des **sondes** (oligonucléotides longs ou produits PCR) spécifiques de différents gènes permettent de détecter des **cibles complémentaires marquées et présentes dans le mélange à analyser**.

Les signaux d'hybridation sont détectés selon le type de marquage par mesure radiographique ou par fluorescence et quantifiés.

Macroarrays



La technologie des macroarrays

consiste à déposer sur des membranes Nylon des produits PCR à une densité d'environ 40 dépôts / cm².

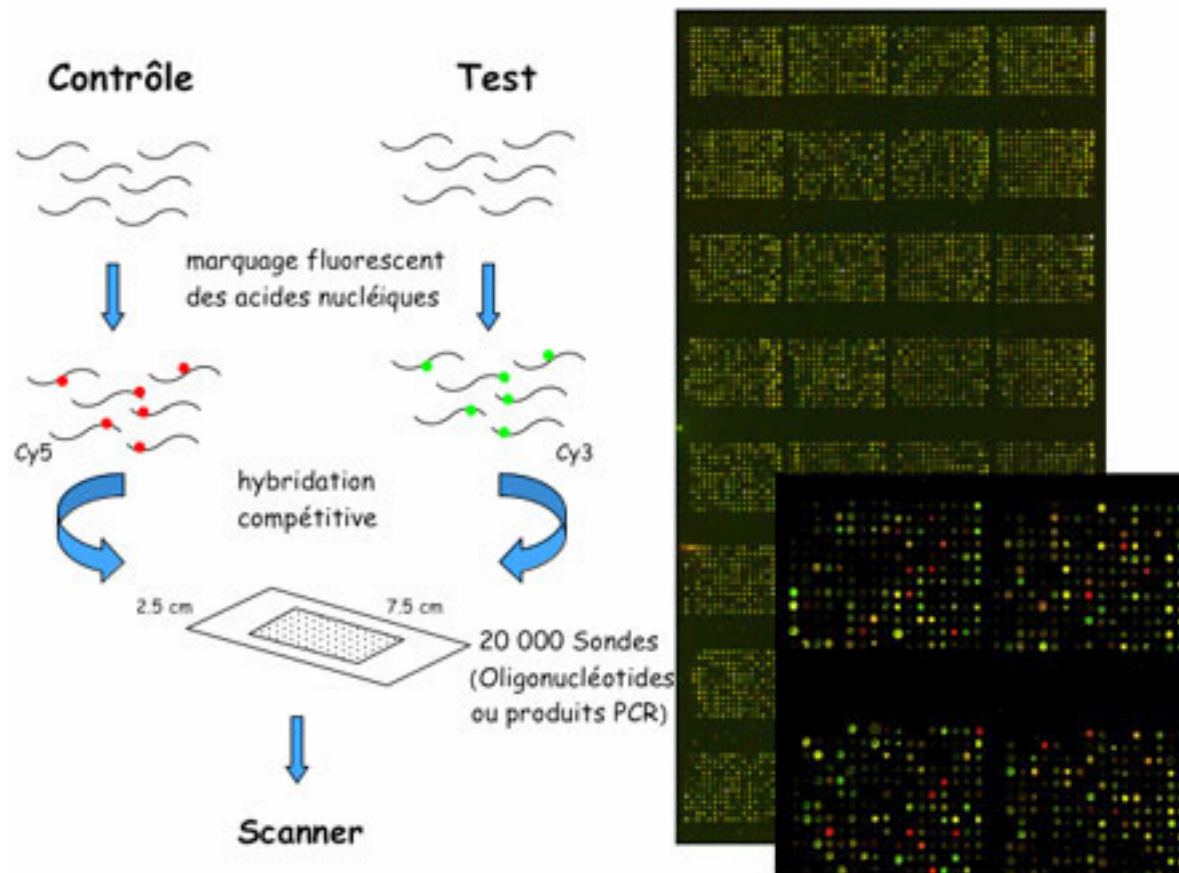
Ces membranes sont ensuite hybridées avec des cibles radioactives marquées au ³³P.

Dans le cas d'expériences d'analyse du transcriptome, ces cibles sont des cDNAs (synthétisés à partir d'ARN extraits de cellules, tissus ou organismes entiers).

Dans le cas d'études de génomique comparée, ces cibles sont des ADNs génomiques.

Les signaux d'hybridation sont détectés à l'aide d'un phosporimager, quantifiés à l'aide logiciels d'analyse d'images puis analysés.

Microarrays



La technologie des microrrays

consiste à déposer sur des lames de verre des produits PCR ou des oligonucléotides longs (50mer-70mer) à des densités pouvant atteindre 6000 dépôts / cm².

Ces lames sont ensuite hybridées avec des **cibles fluorescentes (cDNAs ou ADNs génomique)**. Les molécules fluorescentes les plus couramment utilisées appartiennent à la famille des cyanines, **Cy3 et Cy5**.

Après hybridation, les signaux d'hybridation sont détectés à l'aide d'un scanner de fluorescence.

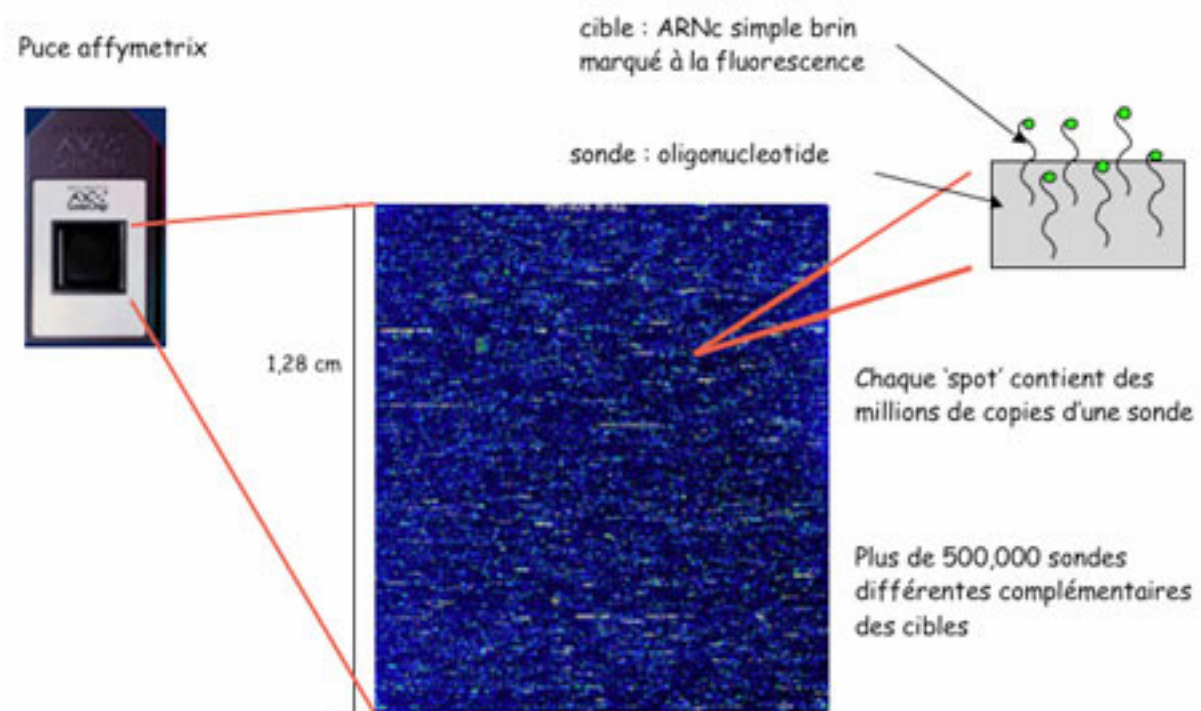
L'utilisation de deux fluorochromes différents permet la détermination des signaux d'hybridat de deux souches distinctes au cours d'une seule expérience.

Une fois les intensités d'hybridation obtenues, **un important travail** d'analyse des données est nécessaire pour en extraire les informations biologiques.

Cette analyse peut être scindée en 4 étapes :

- Ajustement des données
- Tests différentiels
- Classification des gènes
- Stockage et gestion des résultats dans une base de données

Affymetrix



Affymetrix

Contrairement à la technique décrite ci-dessus les oligonucléotides de petite taille (25 mer) ne sont pas déposés sur la lame de verre mais **directement synthétisés in situ**.

Pour chaque séquence nucléotidique, **22 oligos différents sont synthétisés :**

11 perfect-match (PM) et 11 mismatch (MM).

A chaque PM correspond son MM qui diffère par le nucléotide situé en position centrale.

Sur la puce humaine U133A, plus de 22 000 séquences sont représentées soit près de 500 000 oligonucléotides synthétisés.

La synthèse des cibles procède par différentes étapes :

à partir de l'ARN total, du cDNA est synthétisé puis du DNA double brin et enfin du cRNA dans lequel sont incorporés des nucléotides biotinylés.

Après l'hybridation , la révélation se fait par incubation avec de la streptavidine couplée à la phycoérythrine.

Les lames sont ensuite scannées puis analysées par les logiciels MAS 5.0 (analyse absolue résultant de la différence entre PM et MM, analyse comparative entre deux puces) et DMT 3.0 qui permet des analyses différentielles (T-Test, Mann-Whitney, Clustering).

Le processus d'analyse est le même que celui utilisé pour les micro-arrays.

Plate-forme technologique 7 : Synthèse d'Oligonucléotides longs à haut débit



Immunotechnologies 2004 – Module X, 17/12/2004 – Ozier-Kallogéropoulos

Synthèse d'oligodésoxynucléotides pour les puces à ADN :

2 synthétiseurs Mermade IV Bioautomation effectuant les synthèses en format 96 puits,
1 synthétiseur Applied Biosystems 392,
1 synthétiseur Expedite,
1 spectromètre de masse Bruker MALDI-TOF Autoflex 300,
2 électrophorèses capillaires Beckman-Coulter MDQ,
1 HPLC Agilent 1110 avec échantillonnage automatique et un collecteur de fractions,
1 HPLC préparative Delta Prep Waters,
1 HPLC semi-préparative Perkin-Elmer 200,

Environ 700 oligos par jour suivant la longueur des séquences.

Plate-forme Technologique 3 : Protéomique



Immunol

© Shelley D. Spray

Electrophorèse bidimensionnelle :

Ampholine (Millipore) et Immobiline (BioRad), station de traitement, acquisition, et analyse d'image 2D (BioRad).

Deux spectromètres de masse :

- Voyager DE-STR, Maldi-tof et o Sciex API365 triple quadrupole équipé d'une source électrospray (ESI)
- ou d'une source nano-électrospray (NanoSpray).

Plate-forme technologique 5 : Production de Protéines Recombinantes et Anticorps



Plate-forme technologique 6 : Cristallogénèse et Diffraction des rayons X

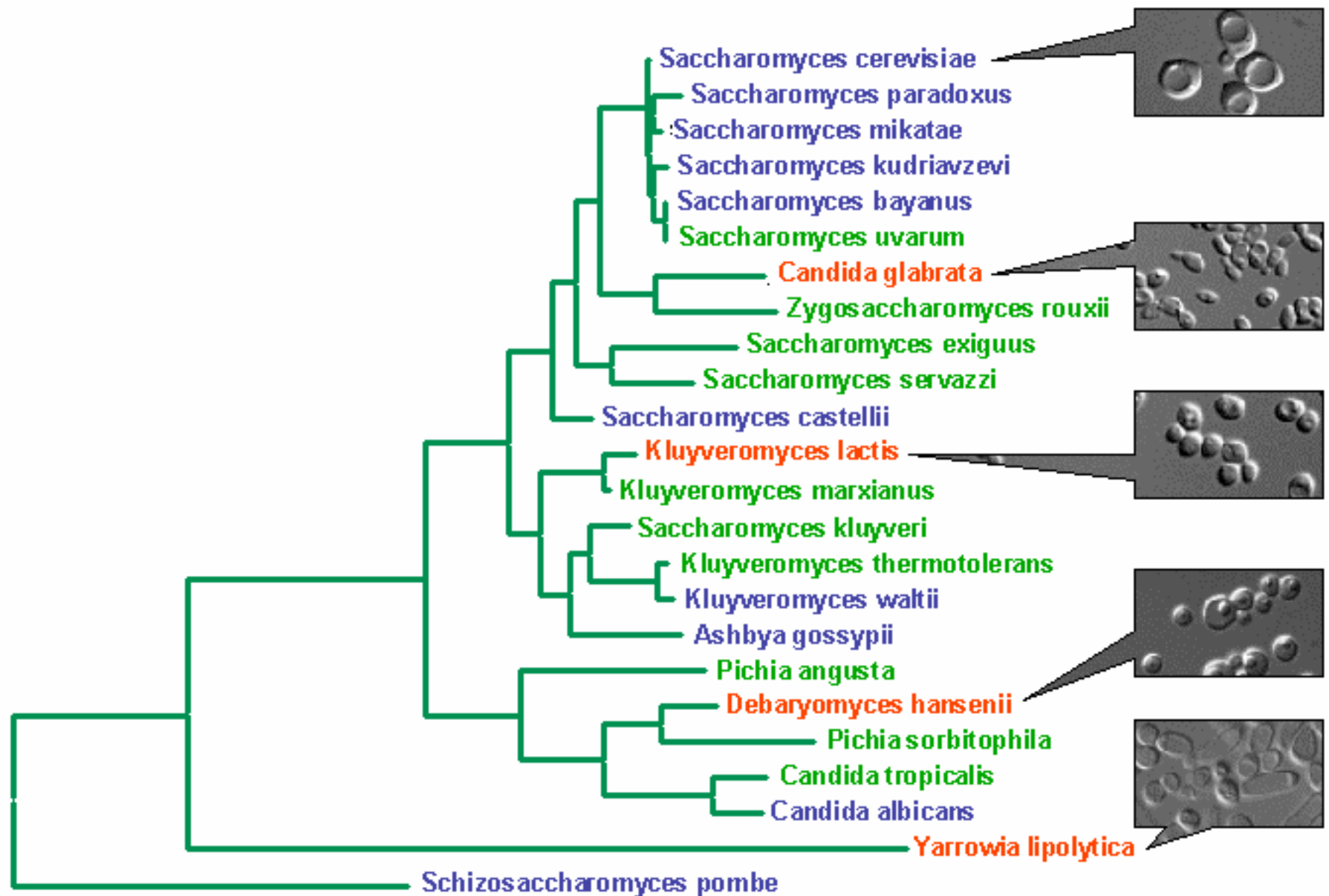


Exemples d'approche génomique chez les levures hémiascomycètes

Génomique comparative et évolution

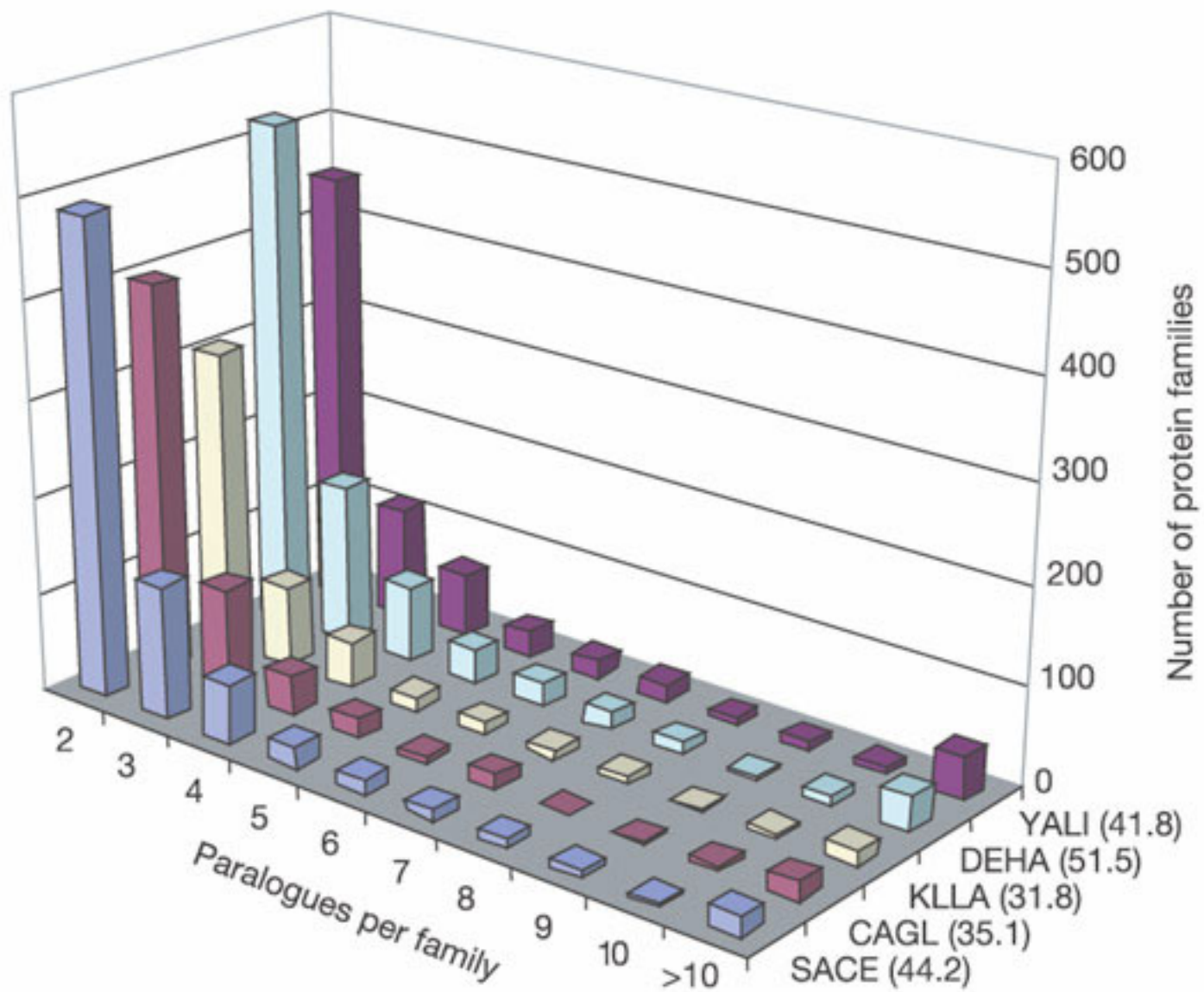
Transcriptomes,

Expression toxique



Genome evolution in yeasts

Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, De Montigny J, Marck C, Neueglise C, Talla E, Goffard N, Frangeul L, Aigle M, Anthouard V, Babour A, Barbe V, Barnay S, Blanchin S, Beckerich JM, Beyne E, Bleykasten C, Boisrame A, Boyer J, Cattolico L, Confanioleri F, De Daruvar A, Despons L, Fabre E, Fairhead C, Ferry-Dumazet H, Groppi A, Hantraye F, Hennequin C, Jauniaux N, Joyet P, Kachouri R, Kerrest A, Koszul R, Lemaire M, Lesur I, Ma L, Muller H, Nicaud JM, Nikolski M, Oztas S, Ozier-Kalogeropoulos O, Pellenz S, Potier S, Richard GF, Straub ML, Suleau A, Swennen D, Tekaia F, Wesolowski-Louvel M, Westhof E, Wirth B, Zeniou-Meyer M, Zivanovic I, Bolotin-Fukuhara M, Thierry A, Bouchier C, Caudron B, Scarpelli C, Gaillardin C, Weissenbach J, Wincker P, Souciet JL.



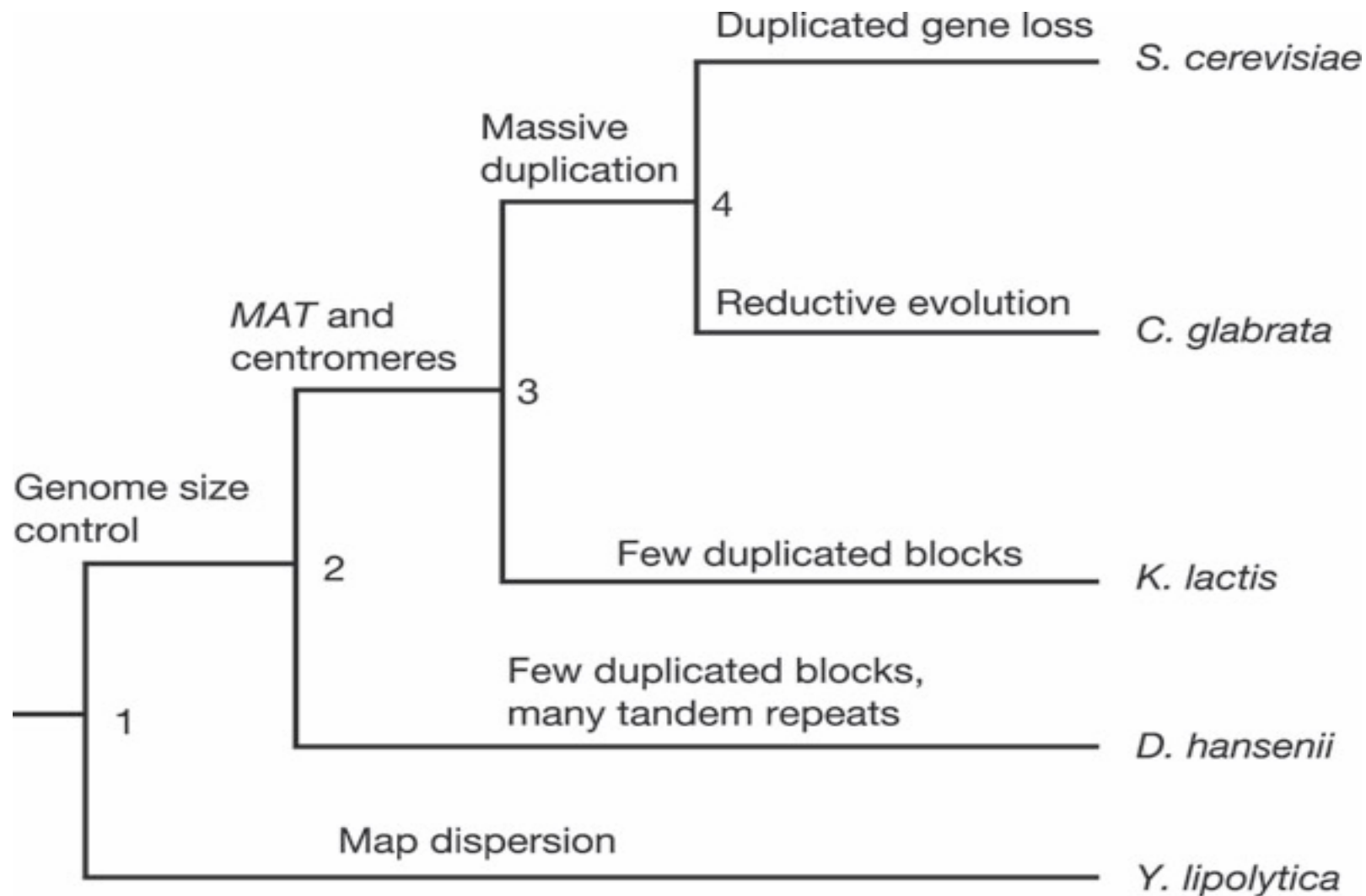


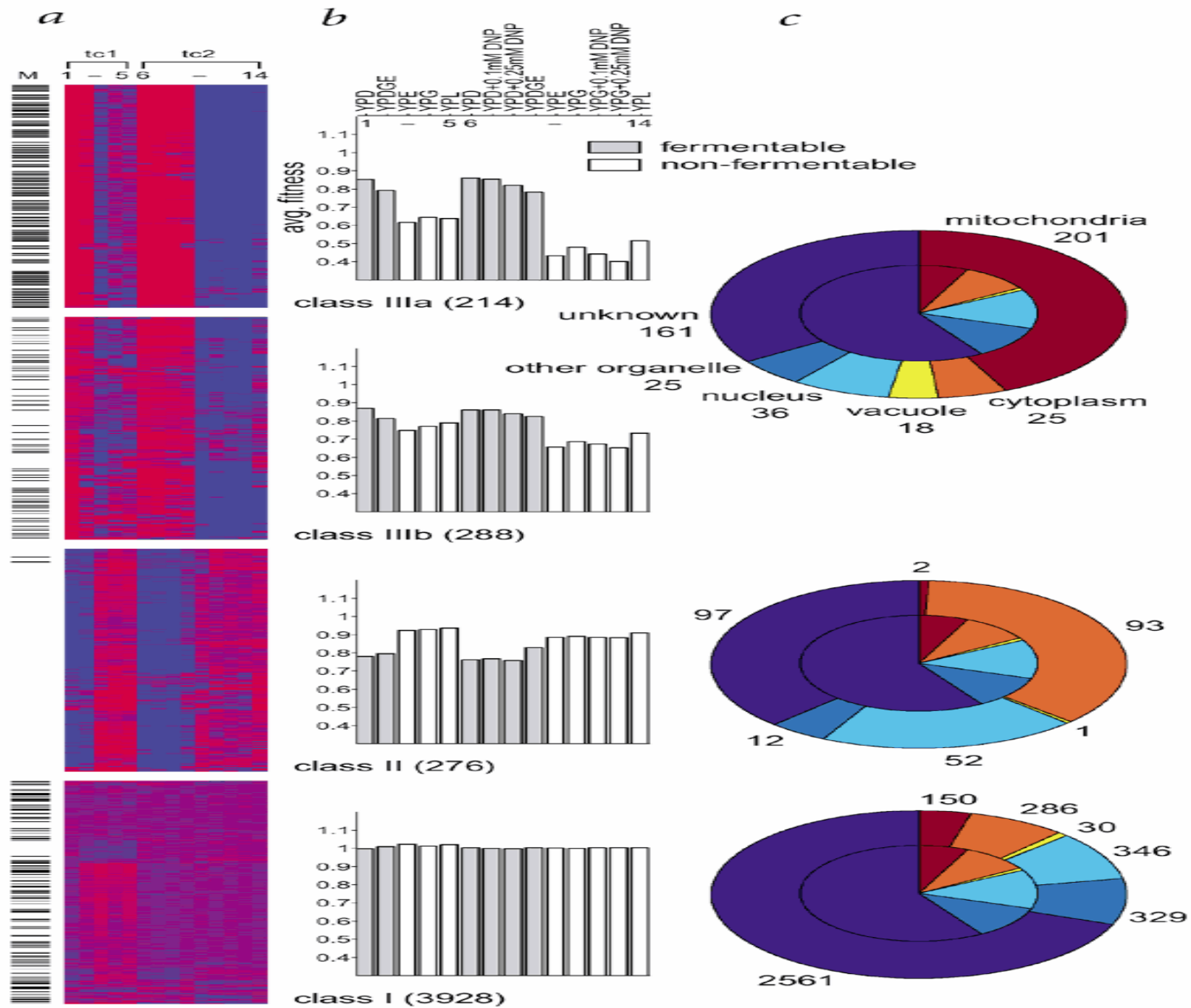
Figure 6 Major evolutionary events in the genomes of hemiascomycetes.

Shown is a cartoon of the evolutionary history of the four sequenced yeasts, plus *S. cerevisiae*.

The tree topology is based on 25S rDNA sequences (see Supplementary Fig. S17).

Systematic screen for human disease genes in yeast

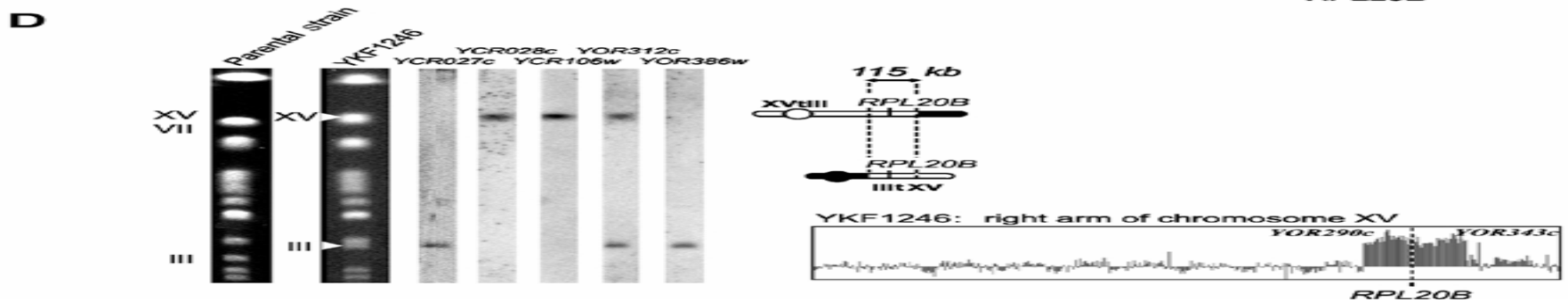
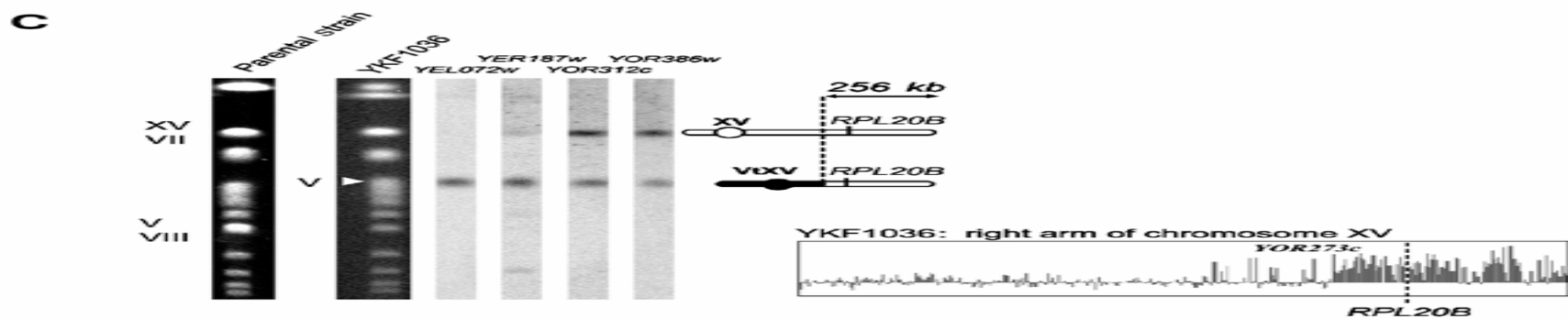
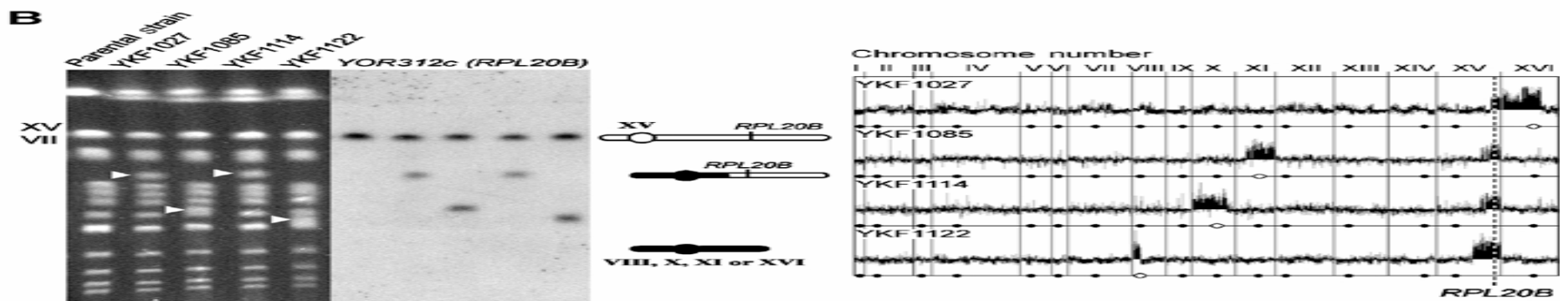
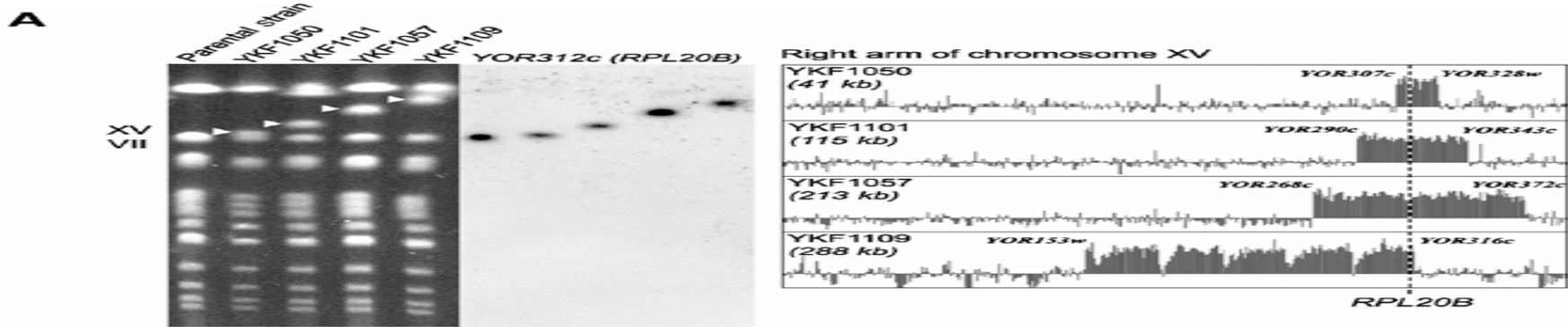
Steinmetz *et al*, 2002, nature genetics,
vol 31, p400, 2002



QDP	Yeast gene	Human gene	OMIM	Pathophysiological defect
	<i>SDH1</i>	<i>SDHA</i>	600857	RCC II subunit
	<i>SDH2</i>	<i>SDHB</i>	185470	RCC II subunit
	<i>BCS1</i>	<i>BCS1L</i>	603647	RCC III assembly
	<i>SHY1</i>	<i>SURF1</i>	185620	RCC IV assembly
	<i>SCO1</i>	<i>SCO1</i>	603644	RCC IV assembly
	<i>SCO1</i>	<i>SCO2</i>	604272	RCC IV assembly
	<i>COX10</i>	<i>COX10</i>	602125	RCC IV assembly
	<i>LAT1</i>	<i>PDX1</i>	245349	pyruvate DH
	<i>PDA1</i>	<i>PDHA1</i>	312170	pyruvate DH
	<i>PDA1</i>	<i>BCKDHA</i>	248600	AA catabolism (DH)
	<i>PDB1</i>	<i>BCKDHB</i>	248611	AA catabolism (DH)
	<i>KGD2</i>	<i>DBT</i>	248610	AA catabolism (DH)
	<i>LPD1</i>	<i>DLD</i>	246900	pyruvate/AA/ α -KGDH
	<i>MIS1</i>	<i>MTHFD1</i>	172460	AA metabolism
	<i>GCV3</i>	<i>GCSH</i>	238330	AA metabolism
	<i>YHM1</i>	<i>SLC25A15</i>	603861	small-molecule transport
	<i>PET8</i>	<i>SLC25A15</i>	603861	small-molecule transport
	<i>FUM1</i>	<i>FH</i>	136850	TCA-cycle enzyme
	<i>MIP1</i>	<i>POLG</i>	174763	maintenance of mtDNA
	<i>HEM14</i>	<i>PPOX</i>	600923	heme biosynthesis
	<i>YTA12</i>	<i>SPG7</i>	602783	ATP-dependent protease
	<i>CCC2</i>	<i>ATP7B</i>	277900	copper-transport ATPase

Eucaryotic genome evolution through the spontaneous duplication of large chromosomal segments

**Romain Koszul, Sandrine Caburet, Bernard Dujon and
Gilles Fischer**

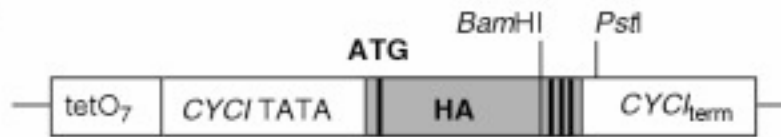


[illegible]

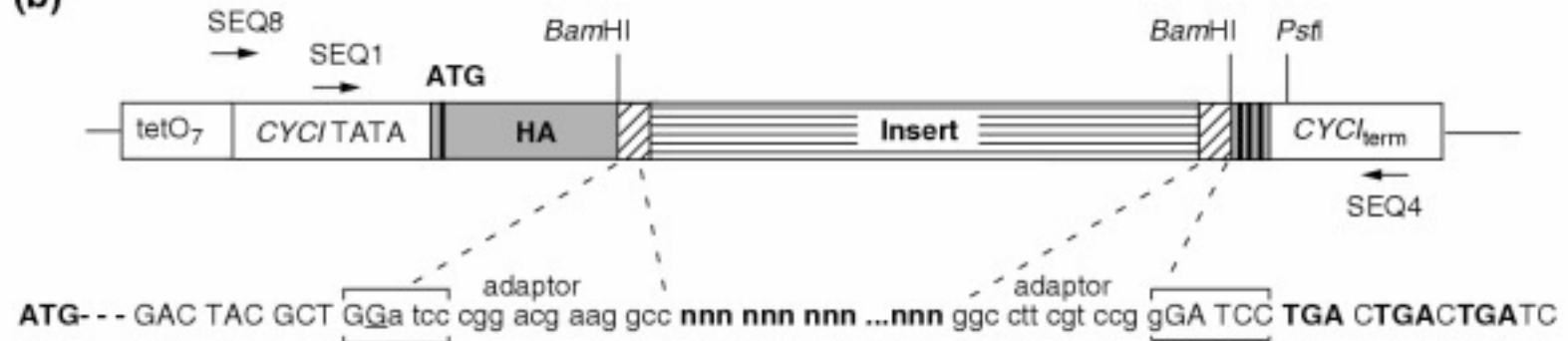
**Large-scale exploration of growth inhibition caused by
overexpression of genomic fragments in *Saccharomyces cerevisiae*.**

**Boyer J, Badis G, Fairhead C, Talla E, Hantraye F, Fabre E, Fischer G,
Hennequin C, Koszul R, Lafontaine I, Ozier-Kalogeropoulos O, Ricchetti M,
Richard GF, Thierry A, Dujon B.**

(a)



(b)



(c)

