

Vendredi 4 novembre 2005

Introduction à la génomique

Odile Kalogeropoulos
Université Pierre et Marie Curie
Institut Pasteur
odozier@pasteur.fr

17 Conférences du cours
Analyse des génomes

Début de la liste

Mercredi 9 novembre

13:00-14:30 *Conférence* : L'annotation *in silico* des séquences génomiques **bactériennes** Claudine MEDIGUE
(Génomique, Evry)

Mercredi 16 novembre

9:00-10:30 *Conférence* : Transferts génétiques horizontaux chez les **bactéries** Didier MAZEL
(Institut Pasteur)

10:45-12:15 *Conférence* : Génomique comparative des **levures** : cycles sexuels et régions subtélomériques Cécile FAIRHEAD
(Institut Pasteur)

Mercredi 23 novembre

9:00-10:30 *Conférence* : Analyse des **complexes protéiques** par les approches protéomiques Bertrand SERAPHIN
(CGM CNRS, Gif sur Yvette)

10:45-12:15 *Conférence* : Approches protéomiques pour la mise au point de nouveaux **médicaments anti-herpès** Jean-Jacques DIAZ
(INSERM U369, Faculté de Médecine Lyon)

Vendredi 25 novembre

9:00-10:30 *Conférence* : Récents développements en **génomique structurale** Arnaud DUCRUIX
(UMR 8015 CNRS, Paris)

10:45-12:15 *Conférence* : Approche «**transcriptome**» : bases théoriques et applications Catherine NGUYEN
(INSERM UMR206, Marseille)

Mardi 29 novembre

14:00-15:30 *Conférence* : Le génome de la **paramécie** Linda SPERLING
(CGM CNRS, Gif sur Yvette)

Fin de la liste

Mercredi 30 novembre

9:00-10:30 *Conférence* : La génomique fonctionnelle chez la souris Marie-Christine SIMMLER
(UMR 7622 CNRS, Paris)

10:45-12:15 *Conférence* : Génomique et génétique humaine : les polymorphismes Laurence COLLEAUX
du génome humain comme outils pour l'étude des maladies (INSERM U393 Necker, Paris)

Vendredi 2 décembre

9:00-10:30 *Conférence* : Etudes de la réplication et de l'instabilité génomique sur molécules uniques d'ADN Aaron BENSIMON
(Institut Pasteur)

10:45-12:15 *Conférence* : L'évolution des génomes viraux Simon WAIN-HOBSON
(Institut Pasteur)

Lundi 5 décembre

9:00-10:30 *Conférence* : Phylogénomique des *Archaea* Patrick FORTERRE
(Institut Pasteur)

10:45-12:15 *Conférence* : Reconstruction du génome d'un vertébré ancestral Hugues ROEST-CROLLIUS
(CNRS UMR 8541, ENS, Paris)

Mardi 6 décembre

9:00-10:30 *Conférence* : Biogenèse des mitochondries : une approche génomique Claude JACQ
(CNRS UMR 8541, ENS, Paris)

Vendredi 9 décembre

9:00-10:30 *Conférence* : Comment la génomique peut aider à comprendre le pouvoir pathogène de la bactérie *Helicobacter pylori* ? Hilde DE REUSE
(Institut Pasteur)

10:45-12:15 *Conférence* : Du génome à la cellule : comment prédire la fonction des gènes ? Antoine DANCHIN
(Institut Pasteur)

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

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

GenomesOnLine
www.genomesonline.org

Nombres et Statistiques

Projets génomes

A la date du 1er novembre 2005 :

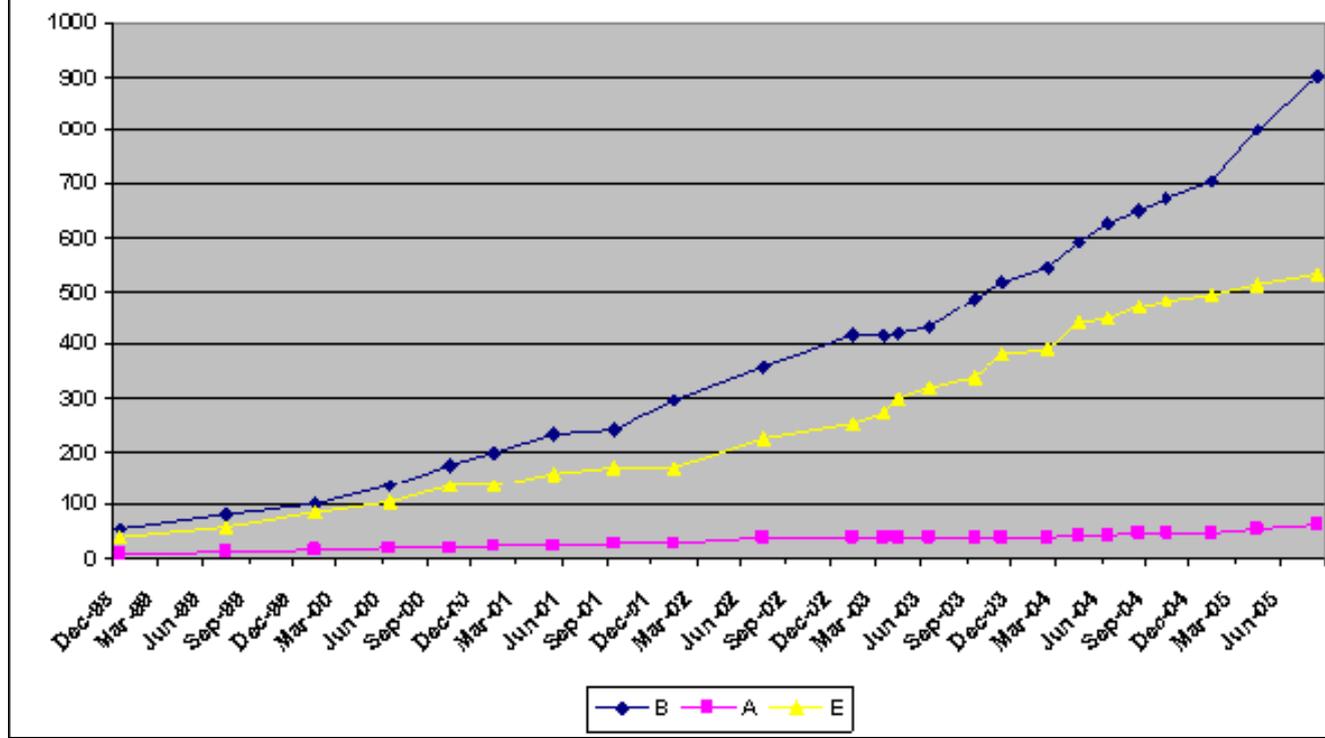
Génomes séquencés publiés :

- **276 génomes Procaryotes**
- **39 génomes Eucaryotes**

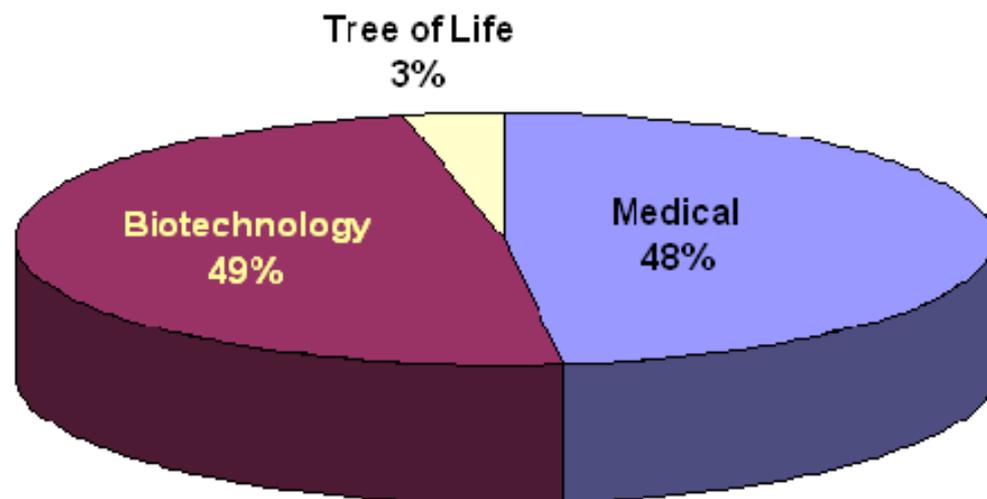
En cours :

- **804 projets de Procaryotes**
- **547 Projets d'Eucaryotes**

Genome Projects on GOL D according to Phylogenetic Groups ©



Application area of Prokaryotic Genome Projects
© August 2005



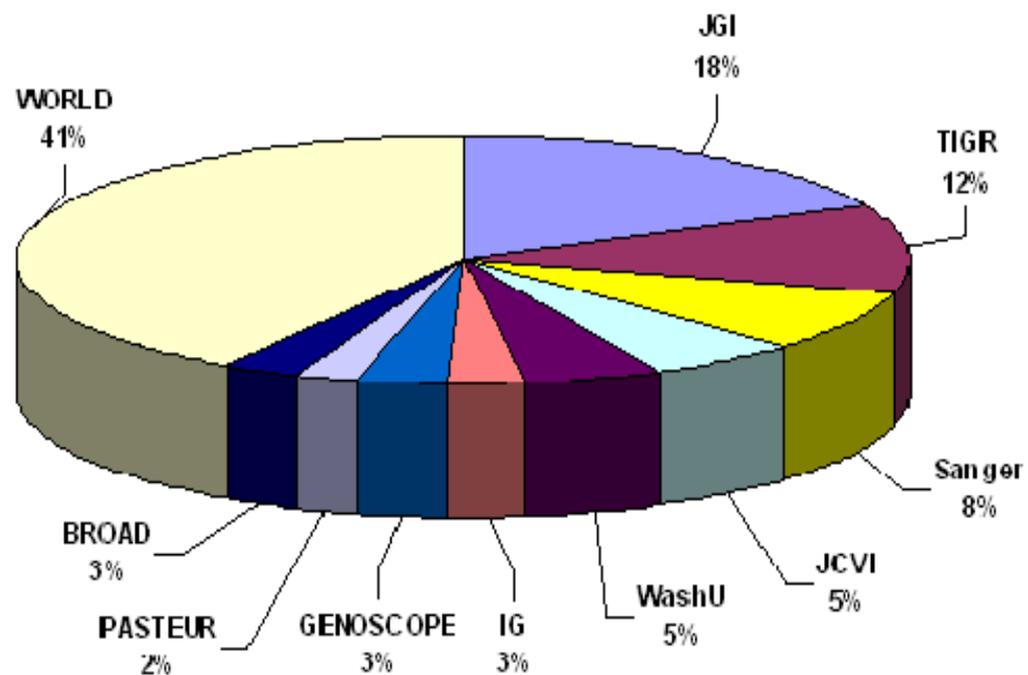
Génomes eucaryotes séquencés

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

Extrait de la 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 (aspectZ 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		
Nematode		
<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)	Mais	Complet
<i>Cyanothrix</i> merolae	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

Major Sequencing Centers August 05: 1496 projects)



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



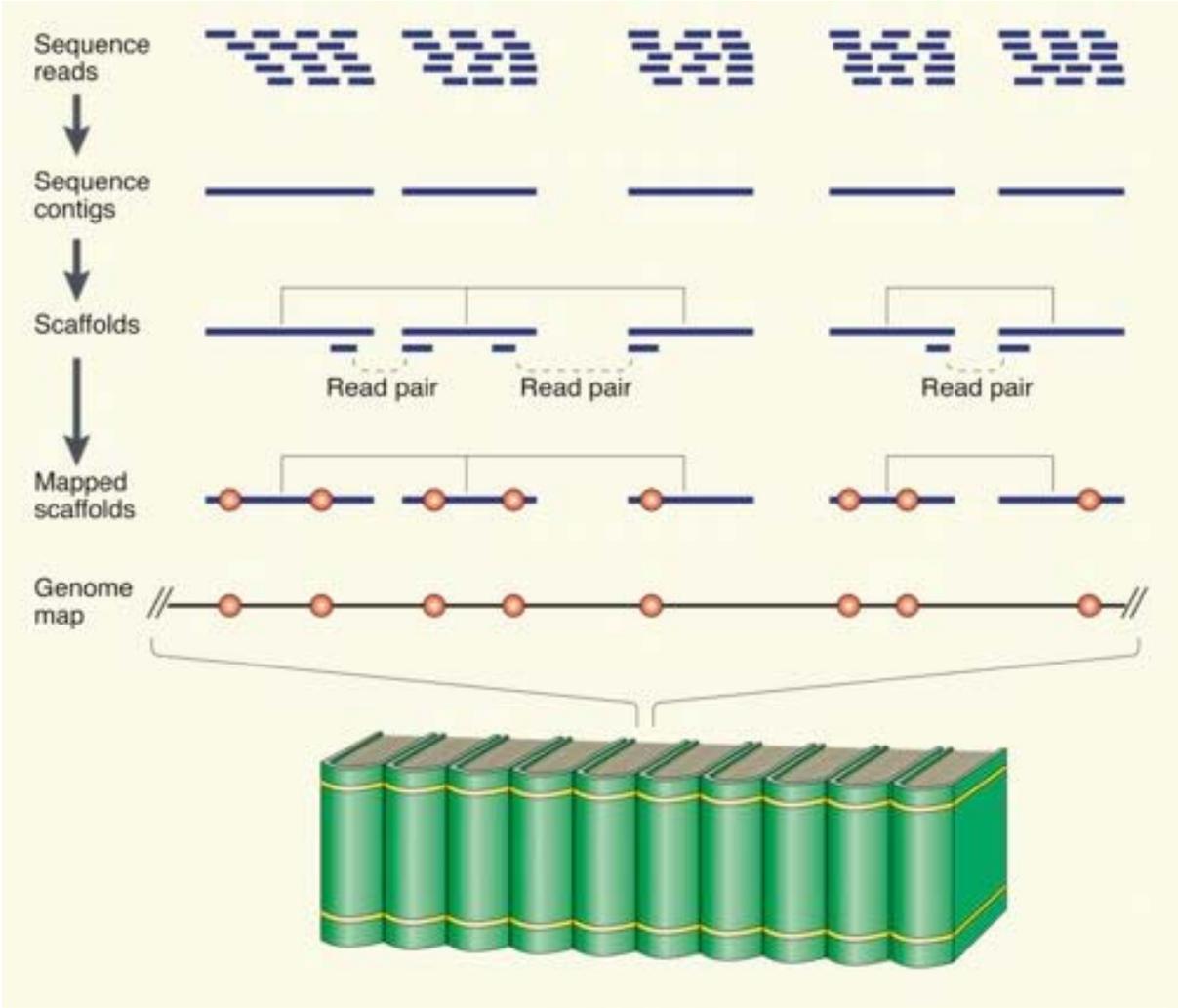
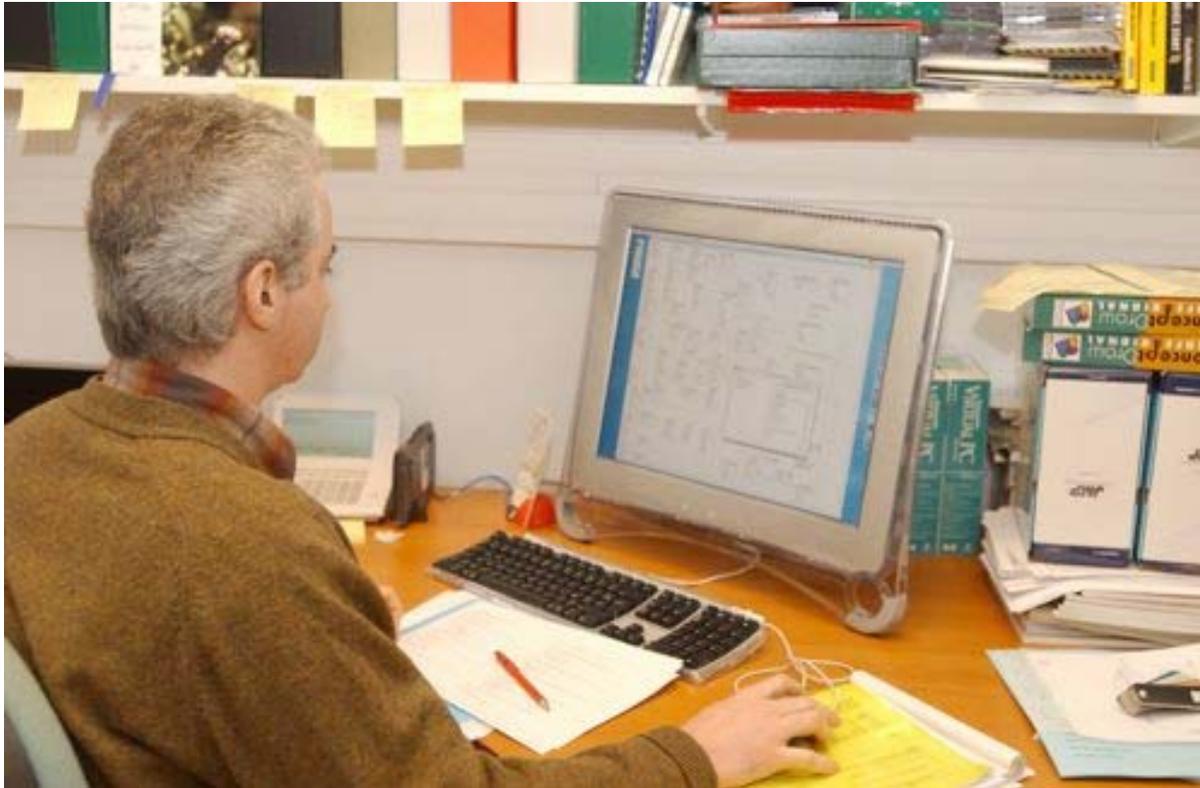


Plate-forme technologique 4 : Intégration et analyse génomiques



Une banque de données locales de l'Institut Pasteur

GenoList genome browser



[SubtiList](#)

Bacillus subtilis strain 168 GenoList browser



[Colibri](#)

Escherichia coli strain K12 GenoList browser



[TubercuList](#)

Mycobacterium tuberculosis strain H37Rv GenoList browser



[Leproma](#)

Mycobacterium leprae strain TN GenoList browser



[BoviList](#)

Mycobacterium bovis strain AF2122/97 GenoList browser



[ListiList](#)

Listeria monocytogenes strain EGD-e and *Listeria innocua* strain CLIP 11262
Multi-GenoList browser



[LegioList](#)

Legionella pneumophila Paris strain and *Legionella pneumophila* Lens strain
Multi-GenoList browser



Plate-forme Technologique 2 : Puces à ADN

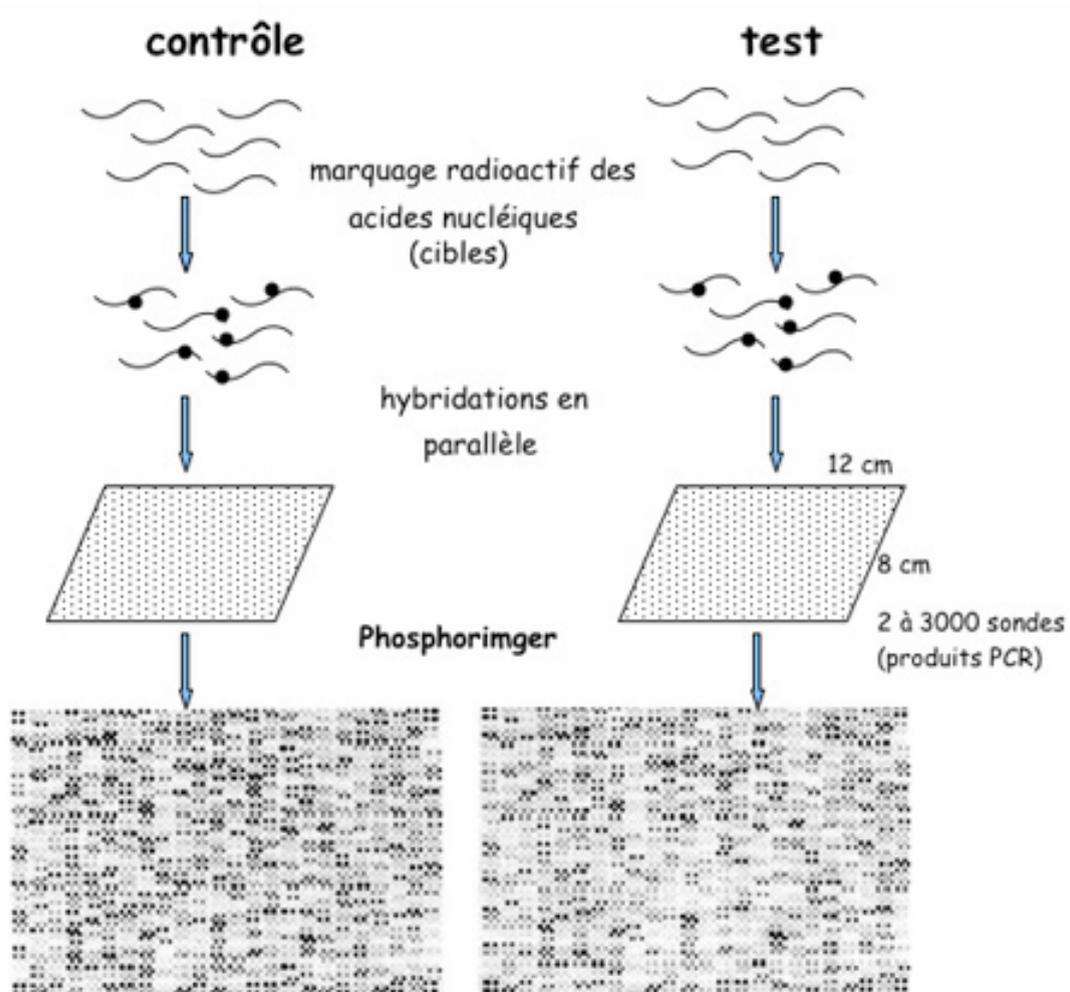


Le principe des puces à ADN est basé sur la technique d'hybridation

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.

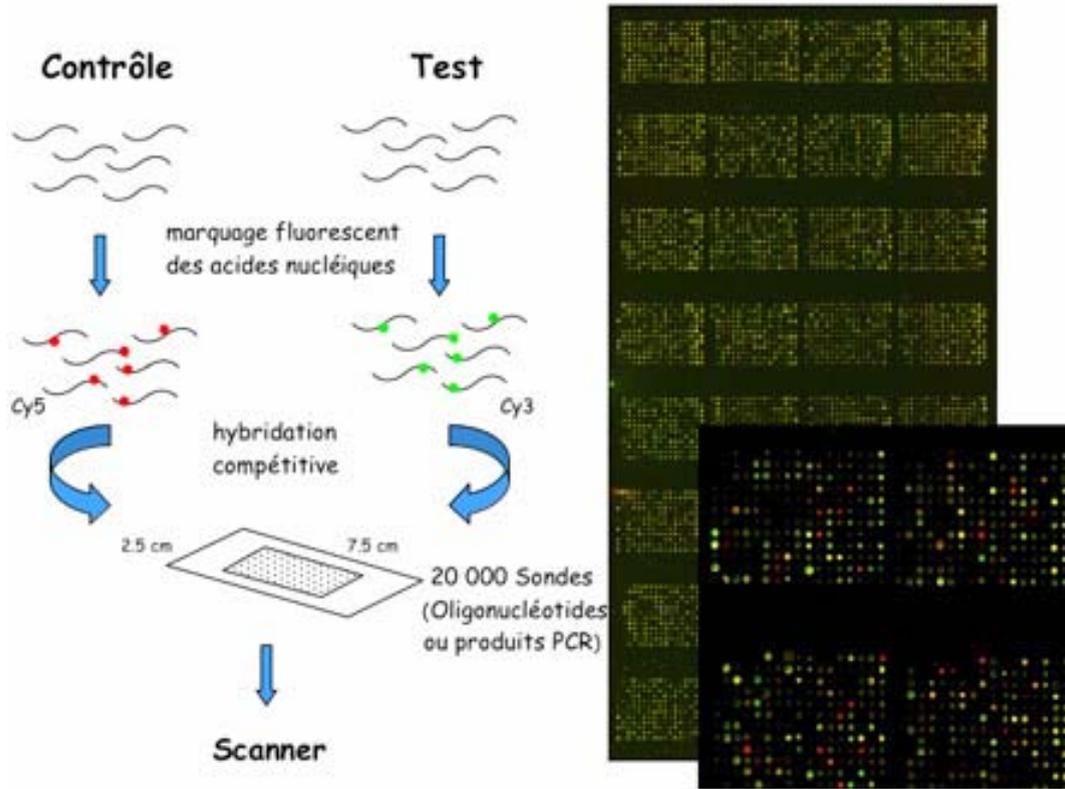
Ces membranes sont ensuite hybridées avec des cibles radioactives marquées au ³³P.

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 phosphorimager, 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.

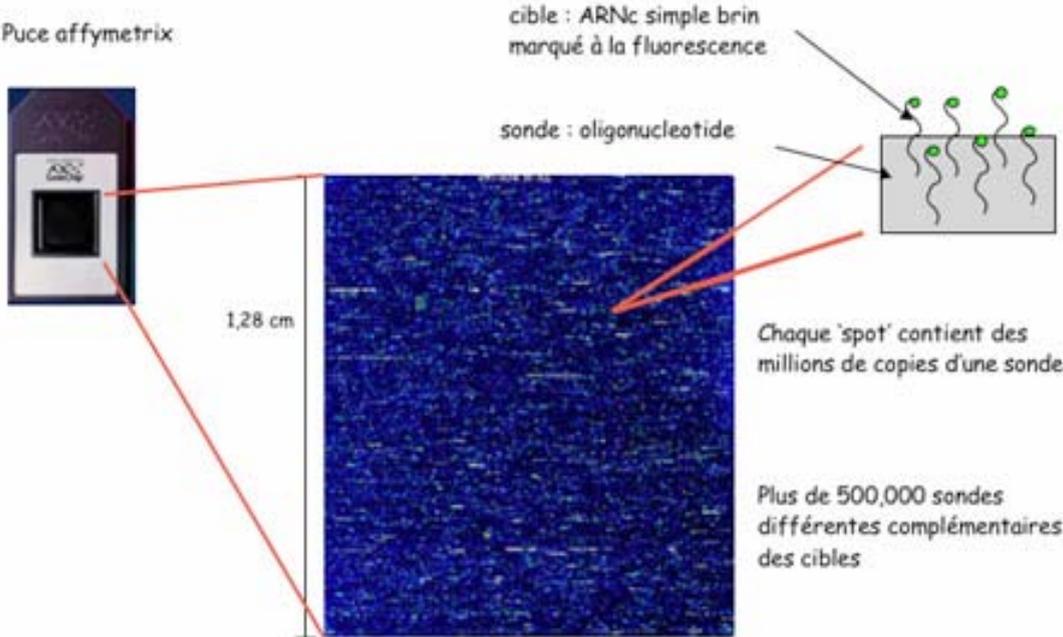
L'utilisation de deux fluorochromes différents permet la détermination des signaux d'hybridation de deux souches distinctes au cours d'une seule expérience.

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



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

