

MASTER Biologie-Santé

HMBS362

Vieillessement cérébral

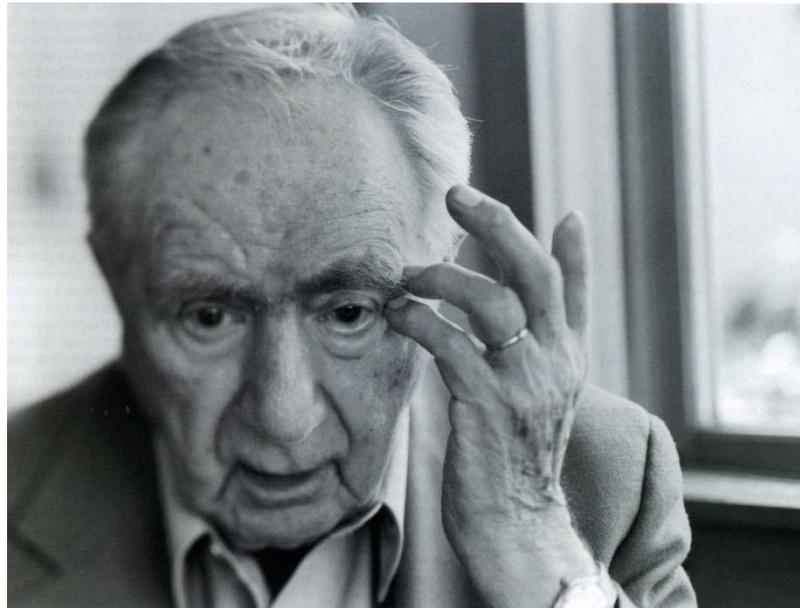
Approche transcriptomique

Gina Devau

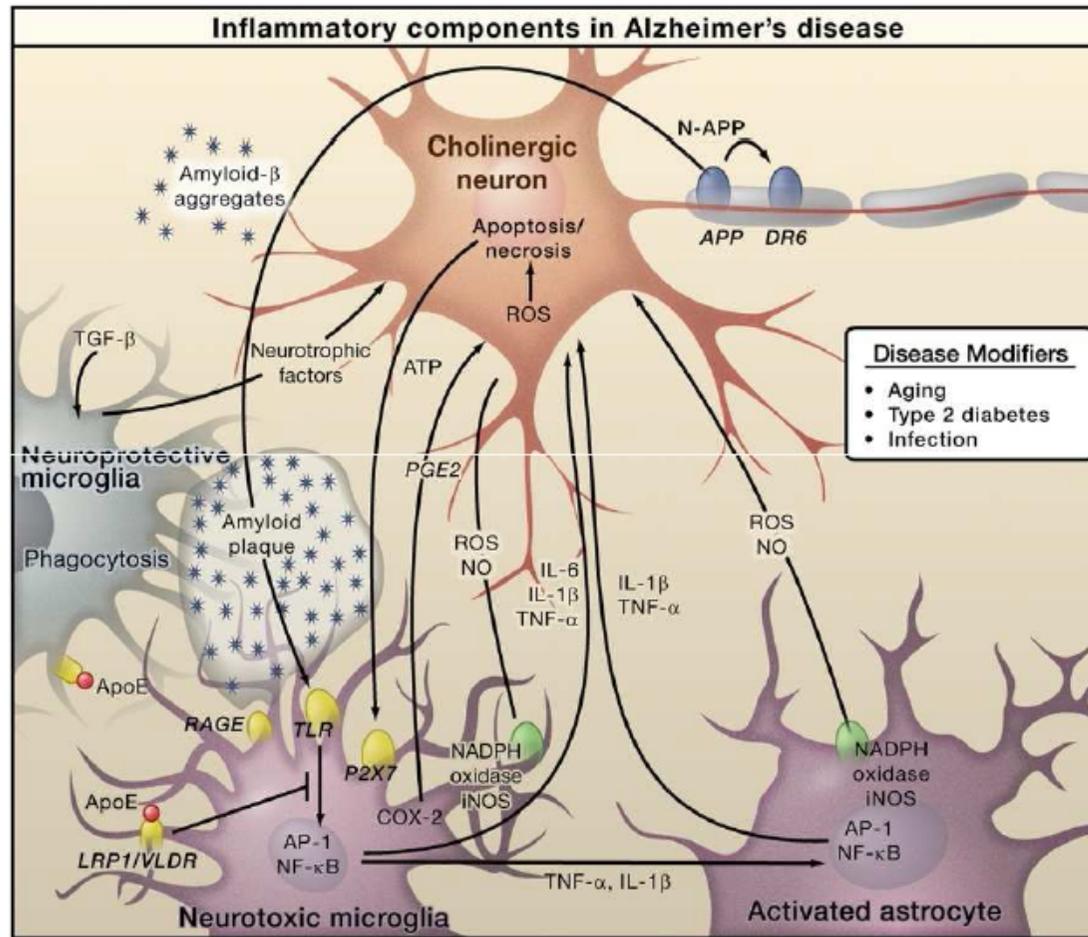
gina.devau@umontpellier.fr



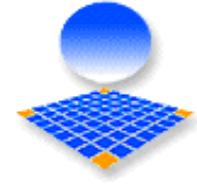
Maladies neurodégénératives : processus complexe



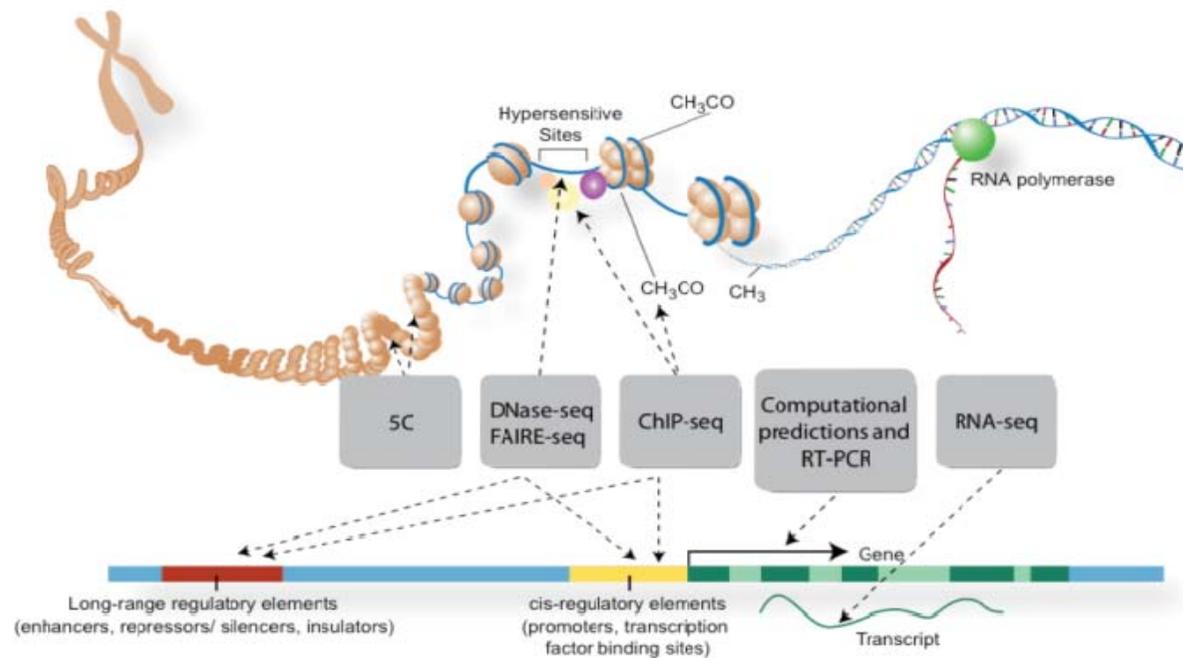
Comprendre les mécanismes moléculaires complexes



Glass et al., Cell, 2010

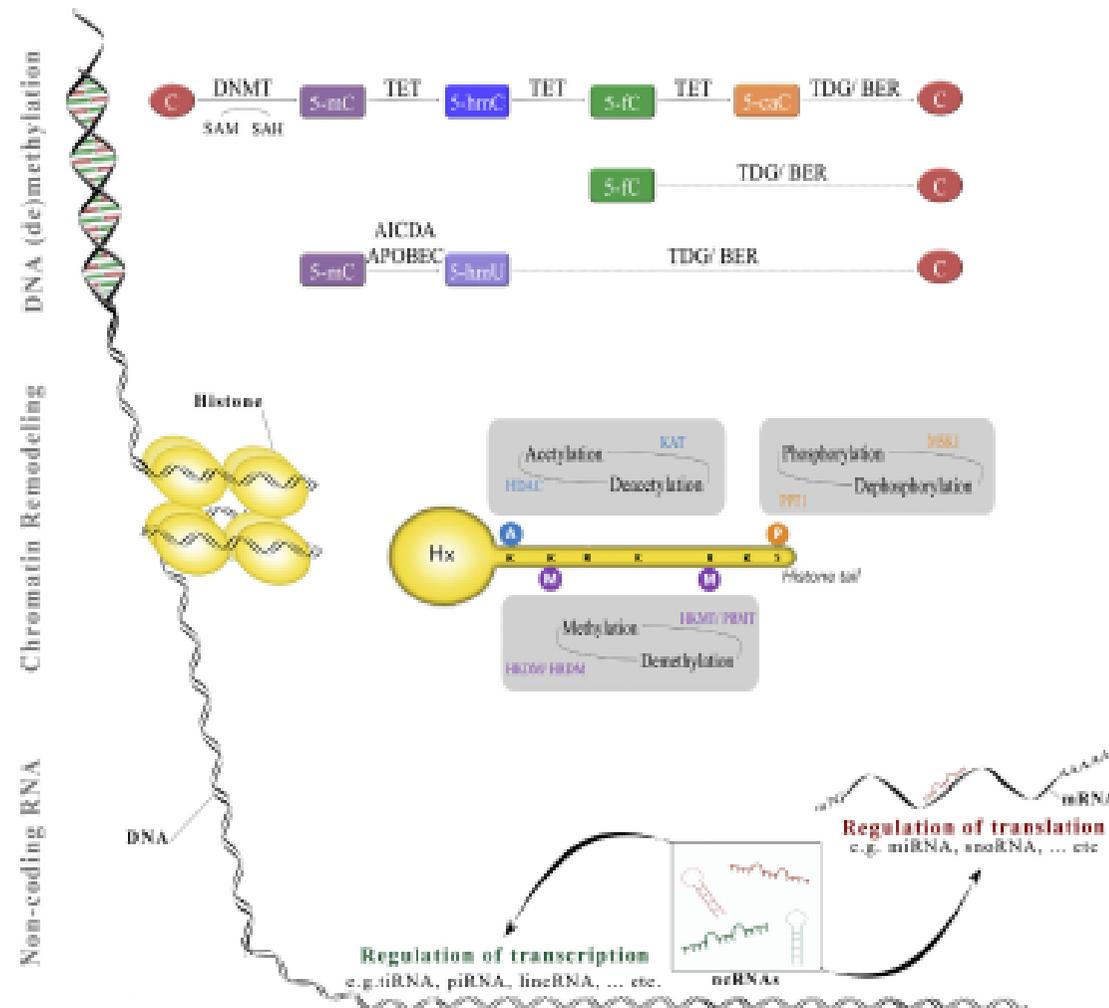


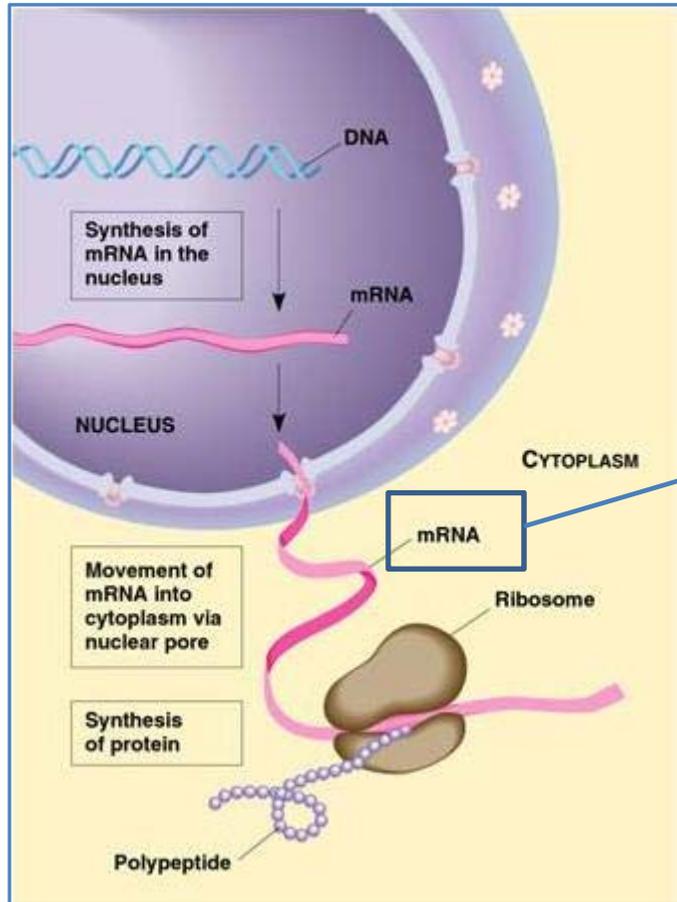
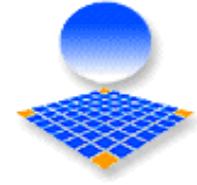
La même information génétique est présente quelque soit l'âge.



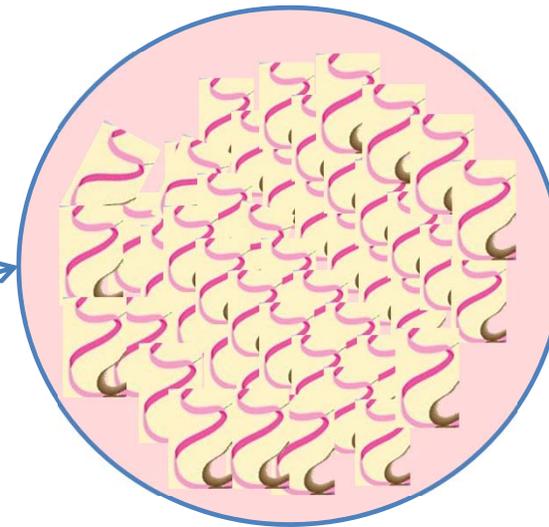
Modification de l'expression des transcrits

Contrôle de la transcription





ARNm : expression d'une séquence génique

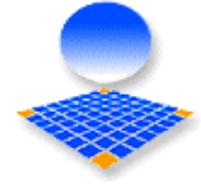


Le transcriptome

C'est l'ensemble de transcrits dans une cellule ou dans un ensemble de cellules à un moment donné dans une situation donnée

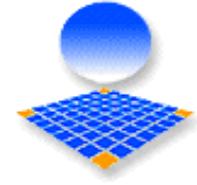


Questions ?



Quelles sont les modifications de l'expression des gènes au cours du vieillissement cérébral ?

Analyse du transcriptome permet d'étudier les modifications de l'expression de l'ensemble des transcrits (ARNm).



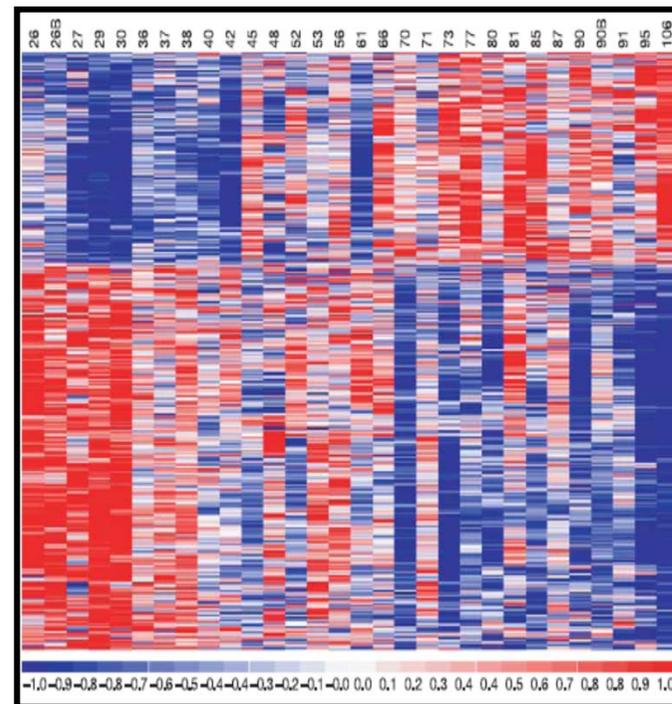
Puces à ADN Affymetrix

Profils d'expression de gènes du cortex frontal humain

Jeunes Adultes Agés

26 ans

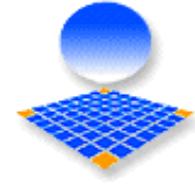
106 ans



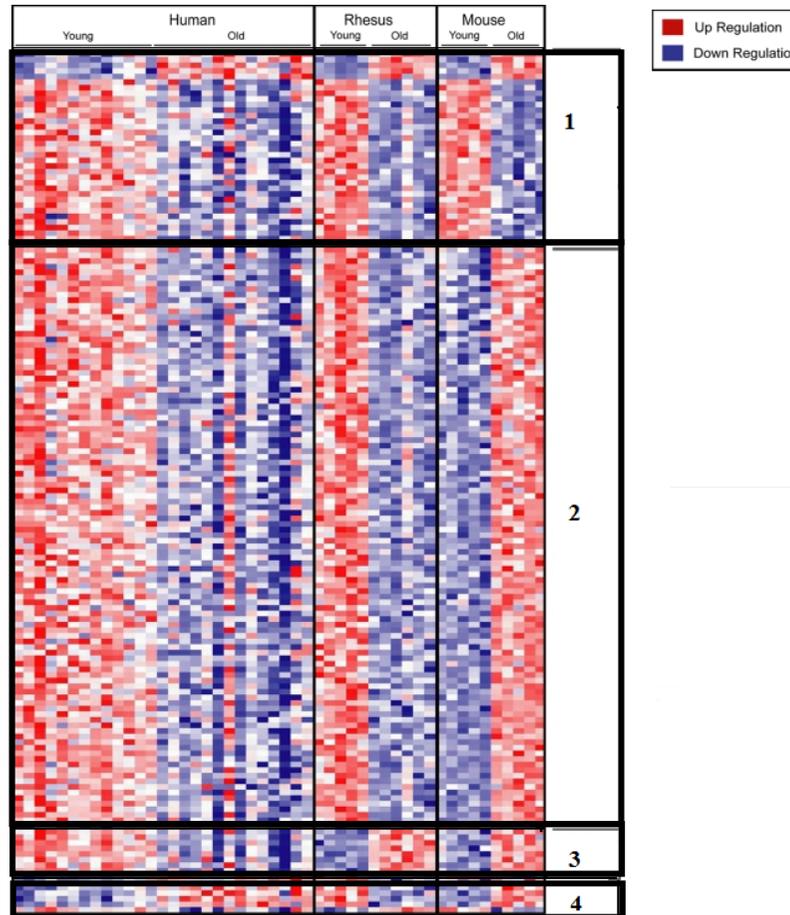
Augmentation
Diminution

Lu *et al.*, Nature 2004.

Modification d'expression de gènes au cours du vieillissement



- 3 espèces:
 - Humains
 - Macaques
 - Souris



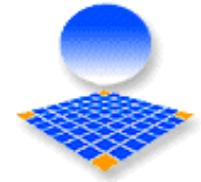
Cortex frontal

- Jeunes
- Agés

Loerch et al., Plos One, 2008

Conservation des profils d'expression chez les primates

Neural mechanisms of ageing and cognitive decline, Bishop, Lu & Yankner, Nature 2010 .



Primate lémurien prosimien



Ordre: Primate
Sous-Ordre: Lémuriformes
Famille: Cheirogaleidae
Origine: Madagascar

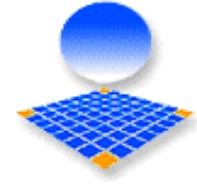
Taille: 20-30 cm
Poids: 80-100 g
Alimentation : fructivore,
insectivore

Animaux protégés par l'annexe 1 de la convention de Washington.

Colonies constituées dans les années 1960
Animalerie Primate de Montpellier:
250 individus,
Tous les âges sont présents.



Synonymes: microcèbes
grey mouse lemur.



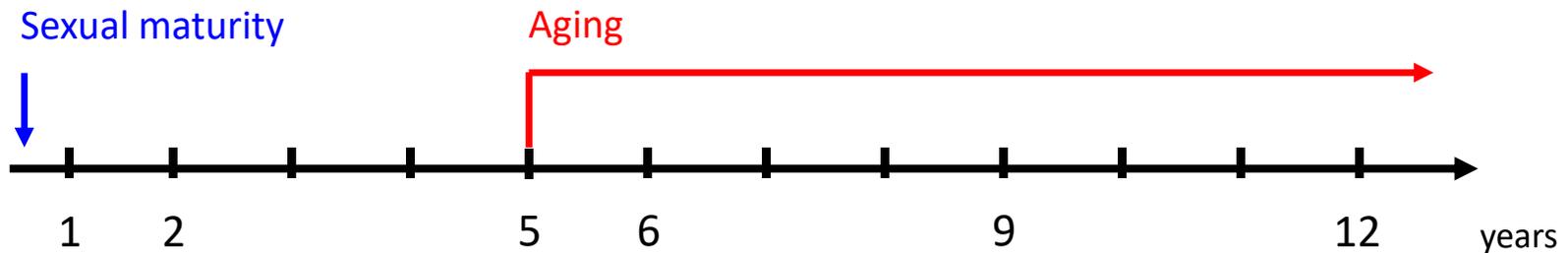
Primate lémurien prosimien



Microcebus murinus

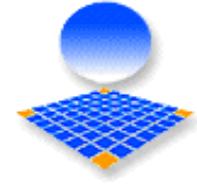
Perret & Aujard, 2006

La durée de vie en captivité 10-13 ans



91% d'identité des séquences entre les ARN Hs et Mm.

Objectifs de notre étude



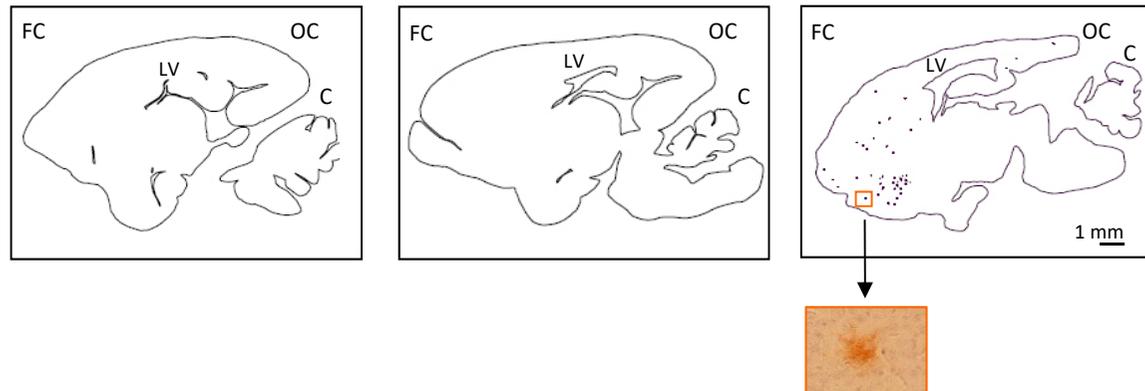
Quels sont les profils d'expression des gènes au cours du vieillissement chez *Microcebus murinus* ?



Microcebus murinus (M.m.)

Comparaison des transcriptomes :

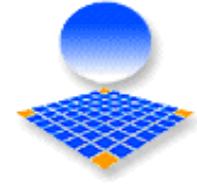
- animaux âgés *versus* jeunes adultes
- animaux AD-like *versus* âgés sains



Identification de gènes dont l'expression est modifiée par l'âge ou la présence de plaques.



Question ?

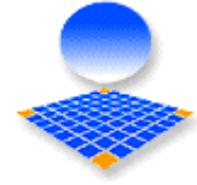


Comment différencier le vieillissement normal
des stades précoces de la maladie d'Alzheimer ?

Comparer le transcriptome de tissu cérébral de sujets
jeunes adultes, de sujets âgés et de sujets porteurs de
plaques amyloïdes de type maladie d'Alzheimer .



Les approches transcriptomiques



Objectif : Détection et quantification de tous les transcrits

Propriété : Hybridation complémentaire des bases nucléotidiques A-T ; C-G

Détection fluorescente de cette hybridation

Hybridation

Primers - Probesets

Microarrays, genechips

Biopuces, puces à ADN

ChIP on chip

Séquençage

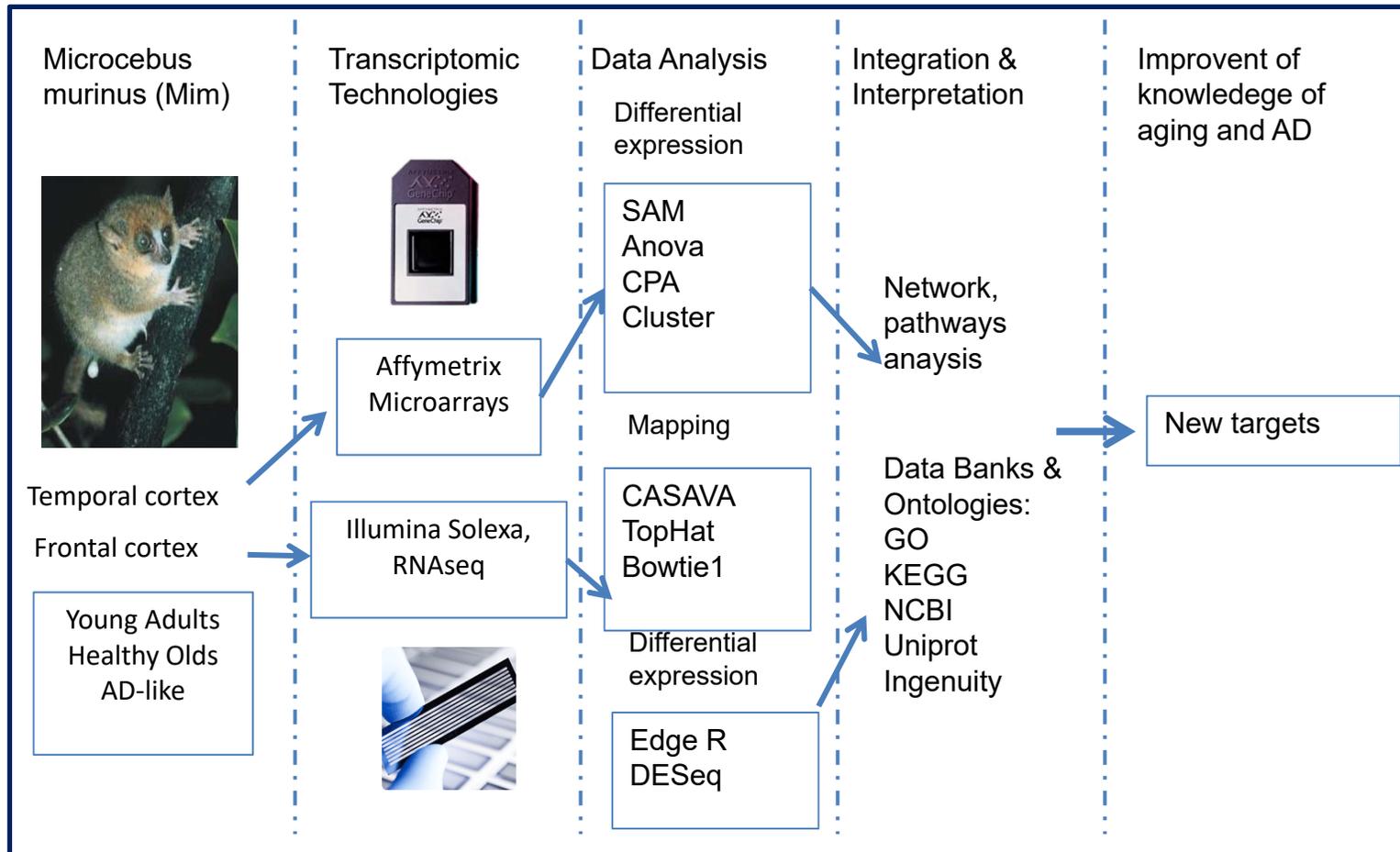
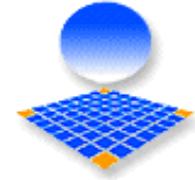
par synthèse inverse

NGS, Next Generation Sequencing

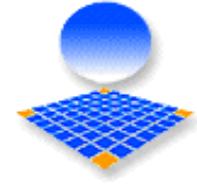
Séquençage à haut débit

NGS 3^{ème} génération

ChIP Seq



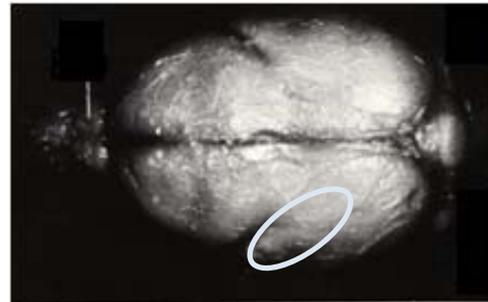
Comparaison des transcriptomes du cortex temporal



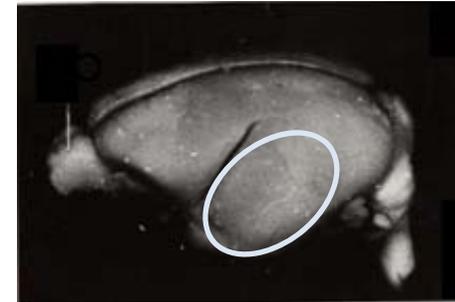
Le cortex temporal

Atlas: N. Bons et al., 1998.

Frontal cortex
Olfative Bulb



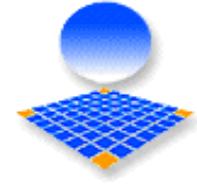
Vue de dessus



Vue latérale de
l'hémisphère gauche

Cerebral cortex
Sylvius scissure
Cerebellum
Brainstem
Spinal cord

About 91% of RNA sequence identity between
Homo sapiens and *Microcebus murinus*.



Microarrays



Affymetrix genechips
HG A133 plus2.00
54 675 probesets

Coll. J. De Vos, IRB

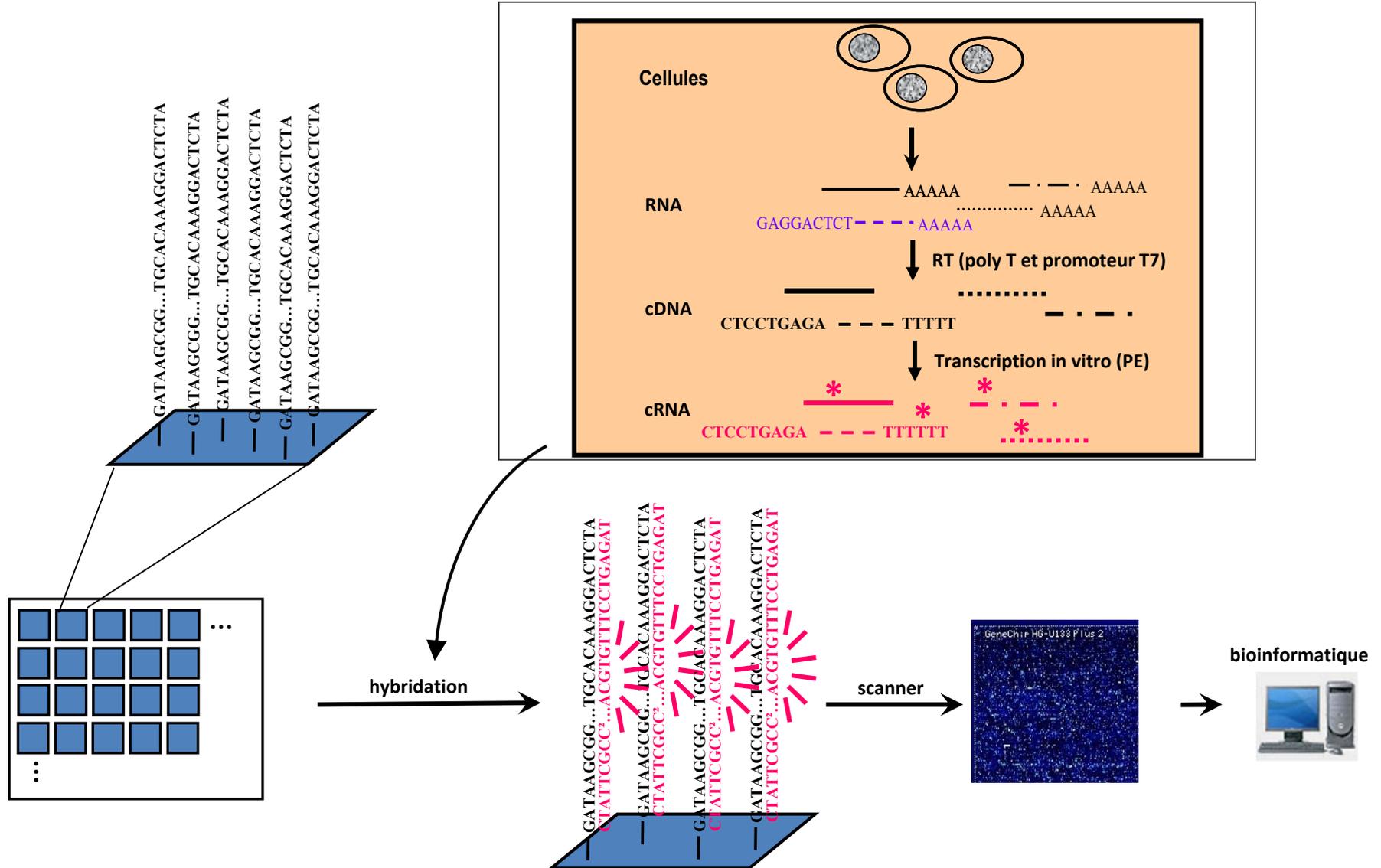
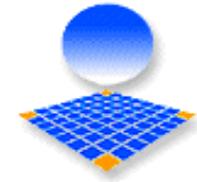
18 Microcebes:

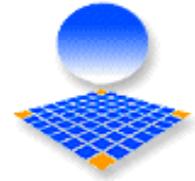
- 6 young adults
- 10 healthy old
- 2 AD-like

Microarrays: to compare the expression of thousands of genes in different tissues, cells or conditions

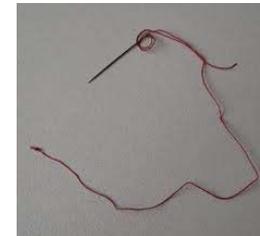
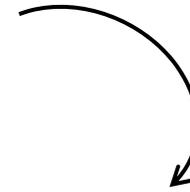
Processing microarray analysis for making biomedical sense is a big challenge because of the large amounts of data

All steps will be conducted according to the MIAME checklists

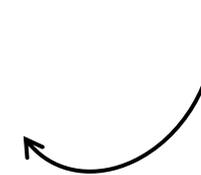




Fouille de données

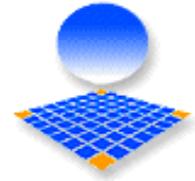


Interprétation des données





Transcriptomic data



Affymetrix Data

Probe Set ID	139 CF2 D_Signal	139 CF2 D_Detection	139 CF2 D_p-value
AFFX-BioB-5_at	1188.18	P	0.0004459014
AFFX-BioB-M_at	2142.838	P	4.428729E-05
AFFX-BioB-3_at	1354.257	P	4.428729E-05
AFFX-BioC-5_at	3856.865	P	5.167325E-05
1007_s_at	382.0397	P	0.02186644
1053_at	7.341878	A	0.8494729
117_at	85.2521	A	0.1276453
121_at	489.4013	P	0.02186644

data Format : (.cel) or numerisation

Normalization by GCOS-MAS5

Expression signal

Call: Present or Absent

P-value at 0.005

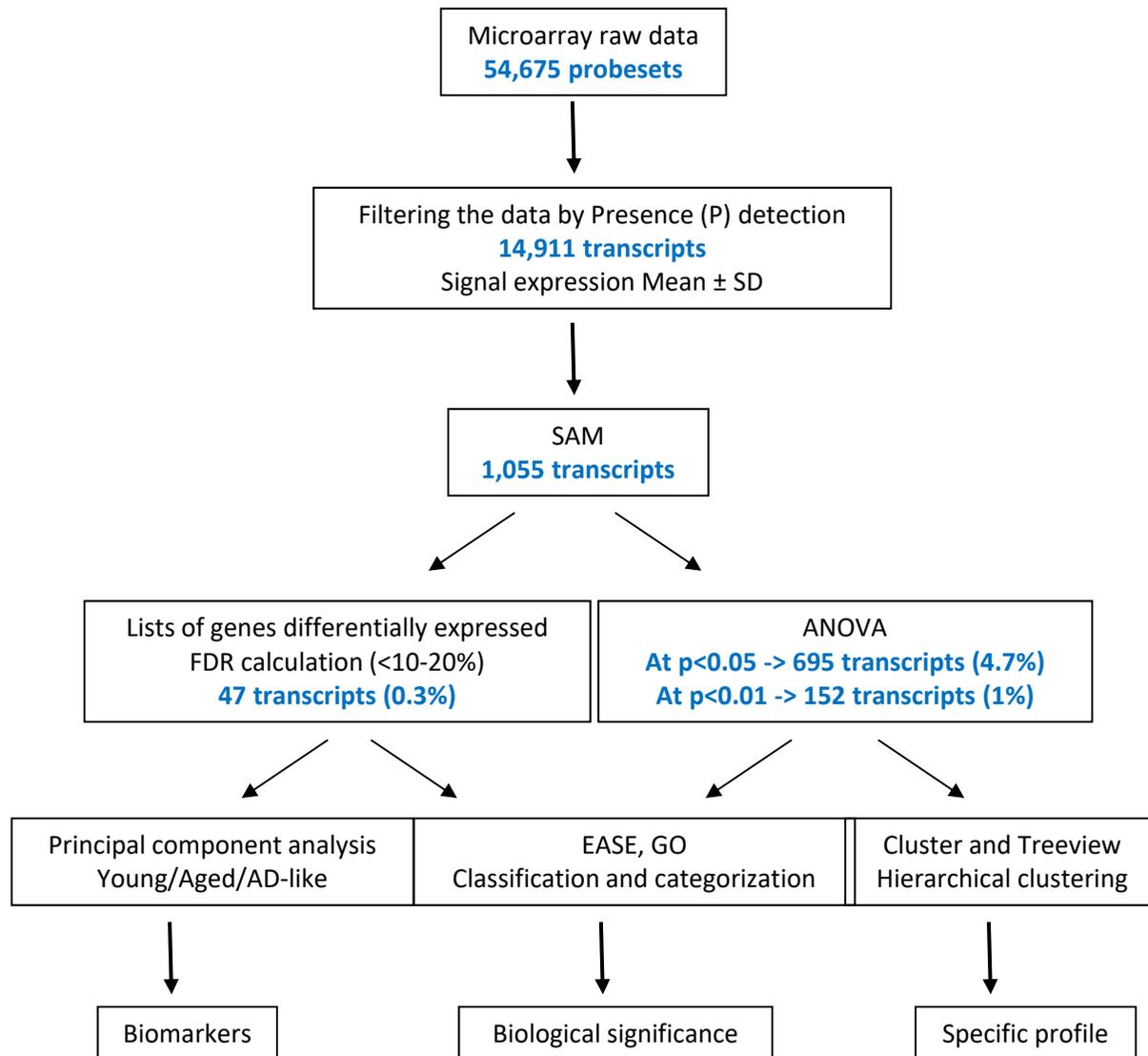
Gènes exprimés dans le cortex temporal

Nombre de gènes détectés présents (P) par hybridation avec la puce HGA-U133 Plus 2.0

Temporal cortex	Gene Number (P)	SD
Young adult	8914.67	1482.68
Healthy old	9456.20	1378.93
AD-like	8259.00	45.25
Over all	9142.67	1349.59

20% de transcrits

Analyse des données transcriptomiques



Analyses des données : SAM analysis

SAM: Significance analysis of microarray

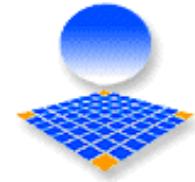
Identification des gènes différemment exprimés en fonction de l'âge

Delta table

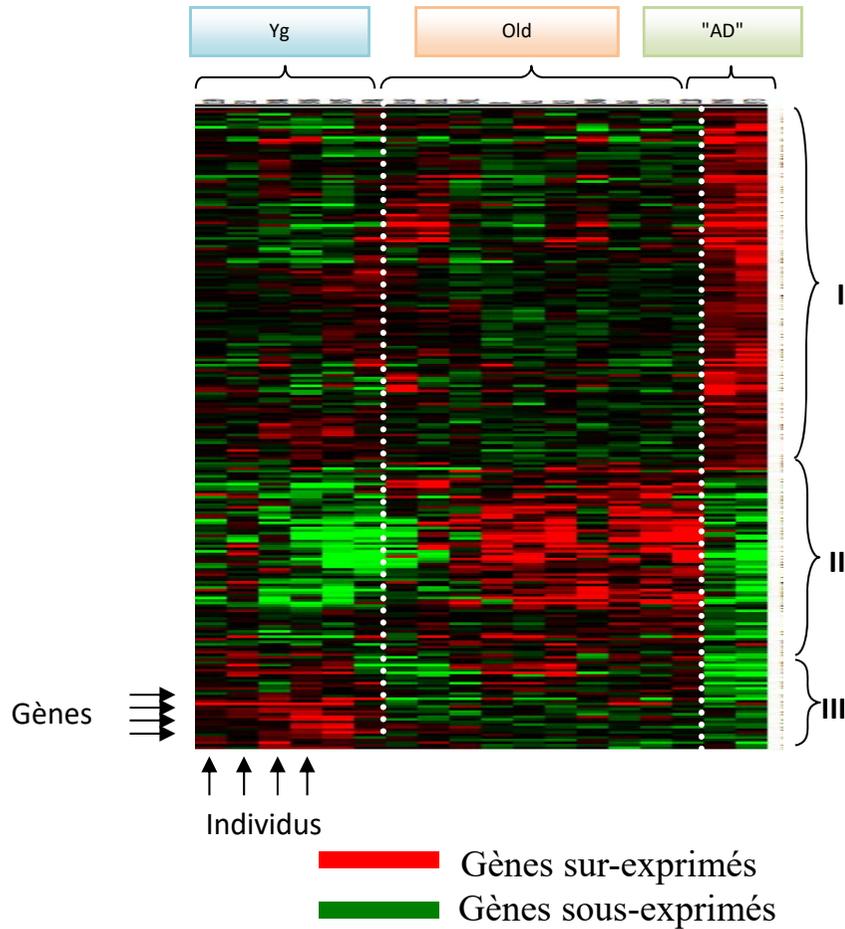
Delta	Significant genes	FDR%
0.76	6	0.00
0.74	14	6.84
0.64	30	9.58
0.64	44	10.88
0.56	53	12.65
0.56	70	13.68
0.49	83	17.31
0.43	105	20.06
0.43	116	20.64
0.42	148	21.35
0.41	177	22.18
0.40	213	23.83
0.39	233	24.25
0.36	334	27.24
0.33	399	31.08
0.30	532	34.74
0.28	605	37.67
0.25	710	41.41

FDR: False discovery rate

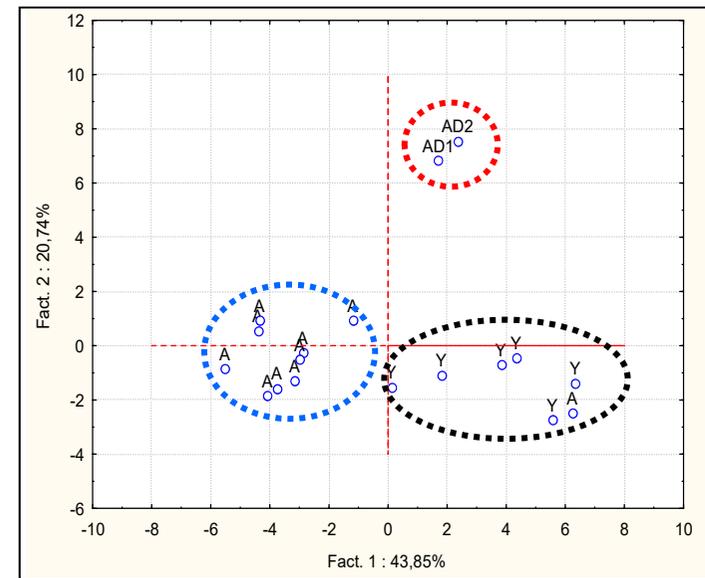
Abdel rassoul et al., PlosOne, 5(9), e12770, 2010



Hierarchical clustering



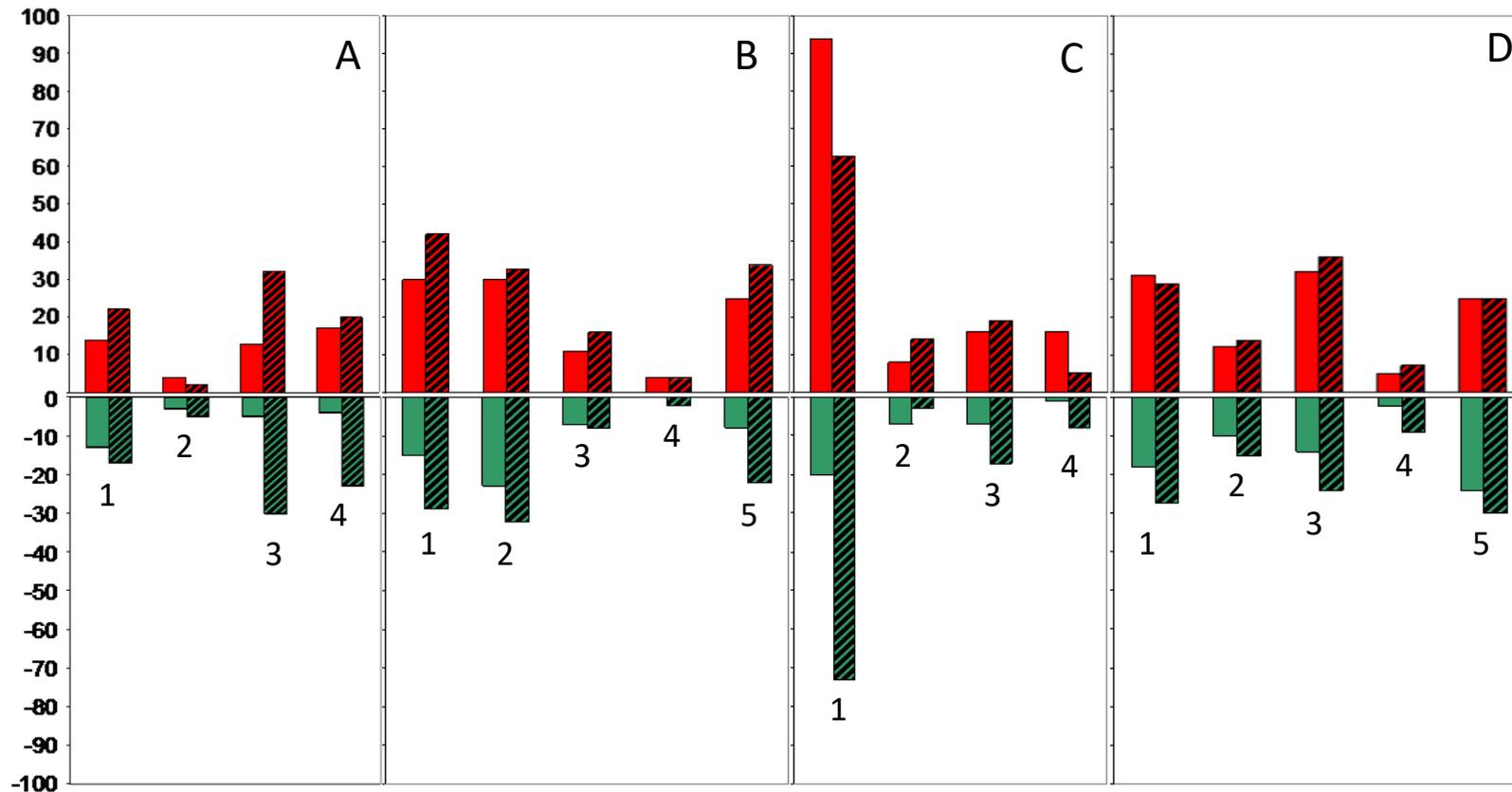
Analyse en composante principale



Projection des individus

Expression de 3 Profils différents
 Identification de signatures moléculaires.

Classification des gènes selon leur fonction biologique



A: Synaptic plasticity

- A1: Synaptic transmission and plasticity
- A2: Neurogenesis and extracellular matrix
- A3: Growth factors
- A4: Cytoskelet

B: Transduction and signaling

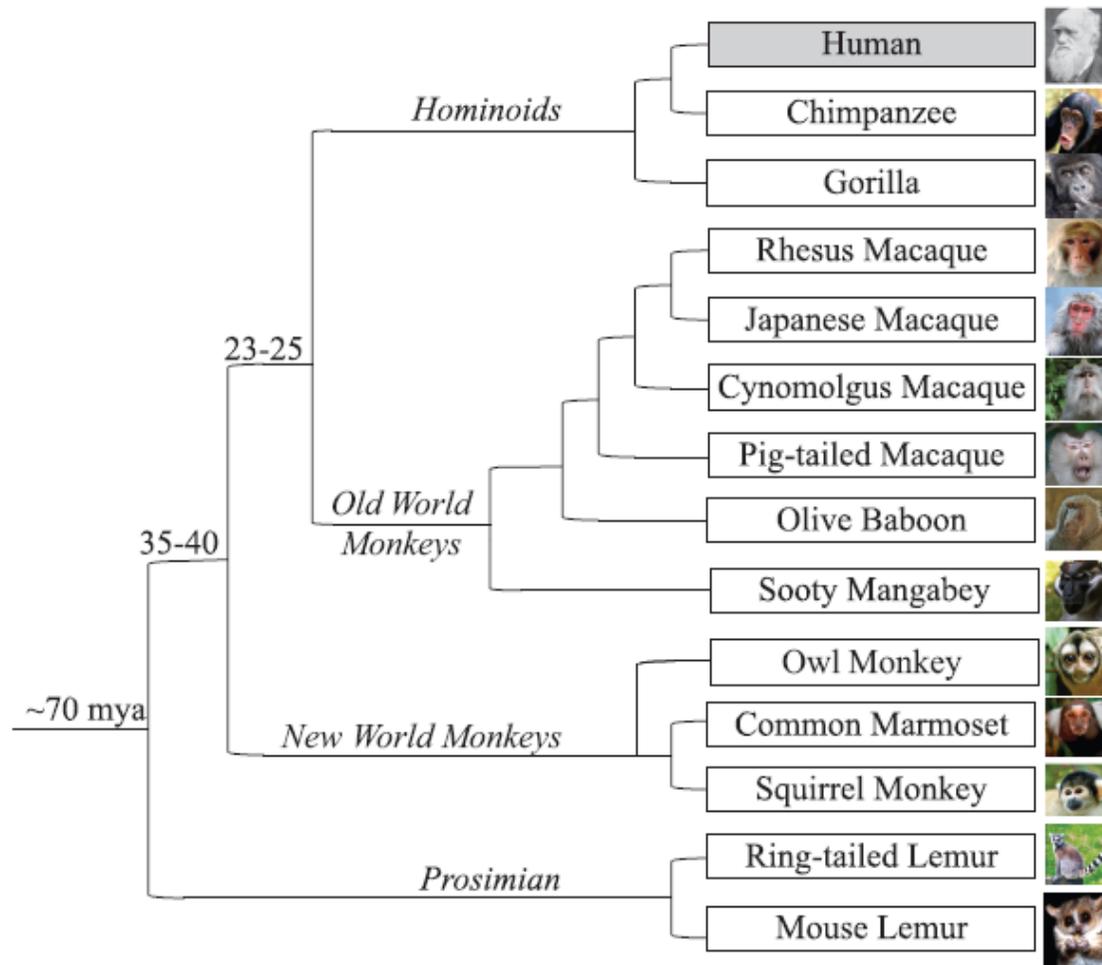
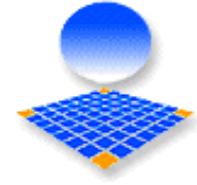
- B1: Calcium signaling; Ion channels, homeostasis
- B2: Transduction signal and second messenger
- B3: Kinases
- B4: Phosphatases
- B5: Transferases

C: Metabolism

- C1: Proteique maturation and synthesis
- C2: Protein degradation
- C3: General metabolism
- C4: Mitochondrial metabolism

D: Nuclear regulation

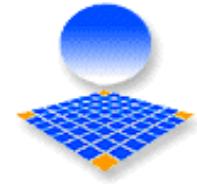
- D1: Nuclear factors
- D2: Transcriptional factors
- D3: Epigenetics
- D4: Cell death
- D5: Cell cycle



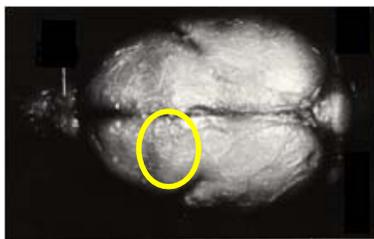
91% d'identité de séquence entre *Homo sapiens* et *Microcebus murinus*.

Détection des transcrits conservés avec les puces Affymetrix HG A133 Plus 2
Données incomplètes

2- Données issues de NGS

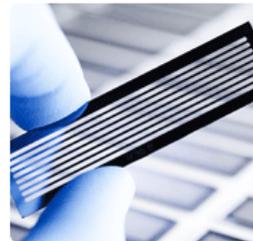


↓
Cerveau
Cortex frontal



Extraction des ARN

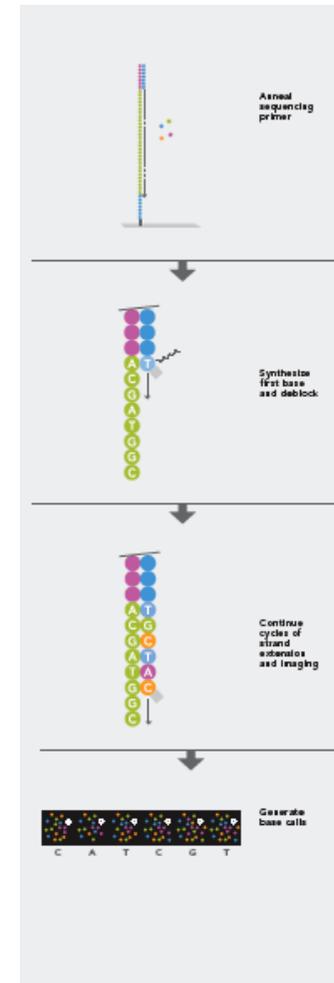
Séquençage à Haut débit
Illumina Solexa HiSeq 2000



Flow-cell

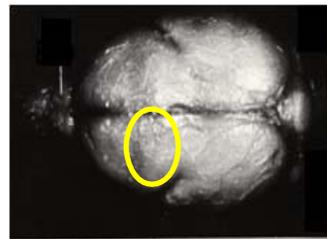
Synthèse de Sanger
4 Fluorochromes

↓
Identification de séquences
"Reads" de 50 nt ou 100 nt



Construction d'une banque par échantillon

Cerveau
Cortex frontal moteur



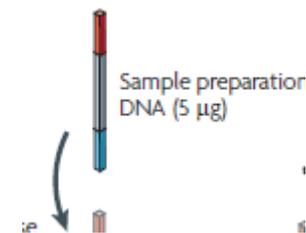
Extraction des ARNm



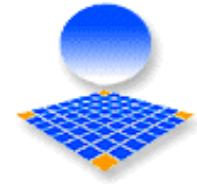
Préparation de l'échantillon



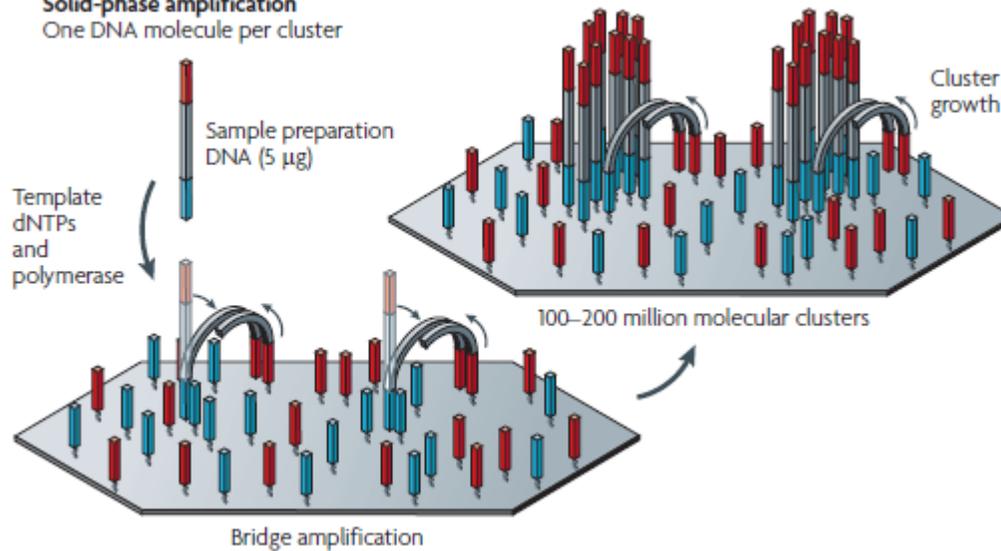
Fragmentation de 50 à 100 nucléotides,
amplification,
Fixation d' adaptateurs spécifiques



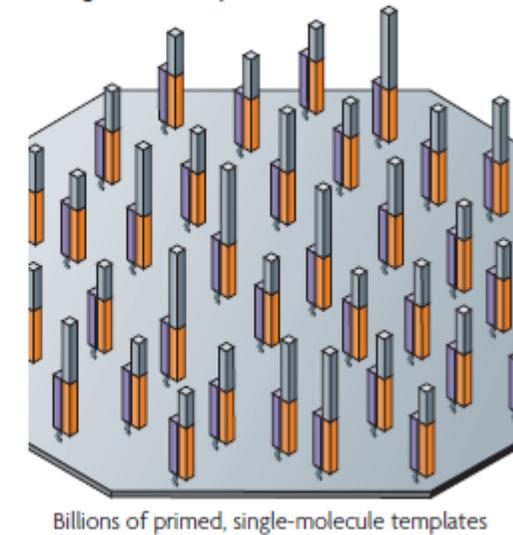
Procédure expérimentale

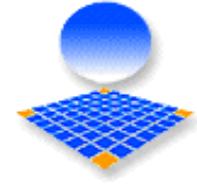


b Illumina/Solexa
Solid-phase amplification
One DNA molecule per cluster



c Helicos BioSciences: one-pass sequencing
Single molecule: primer immobilized



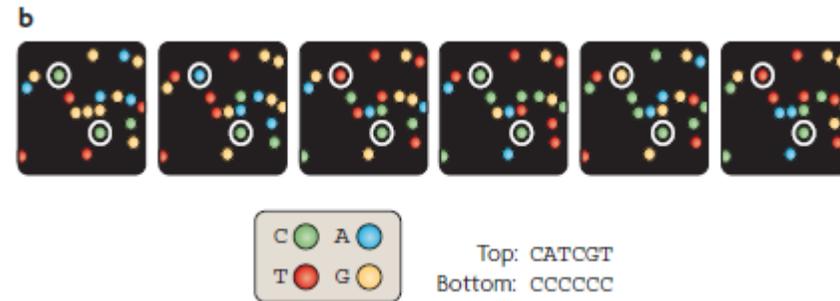
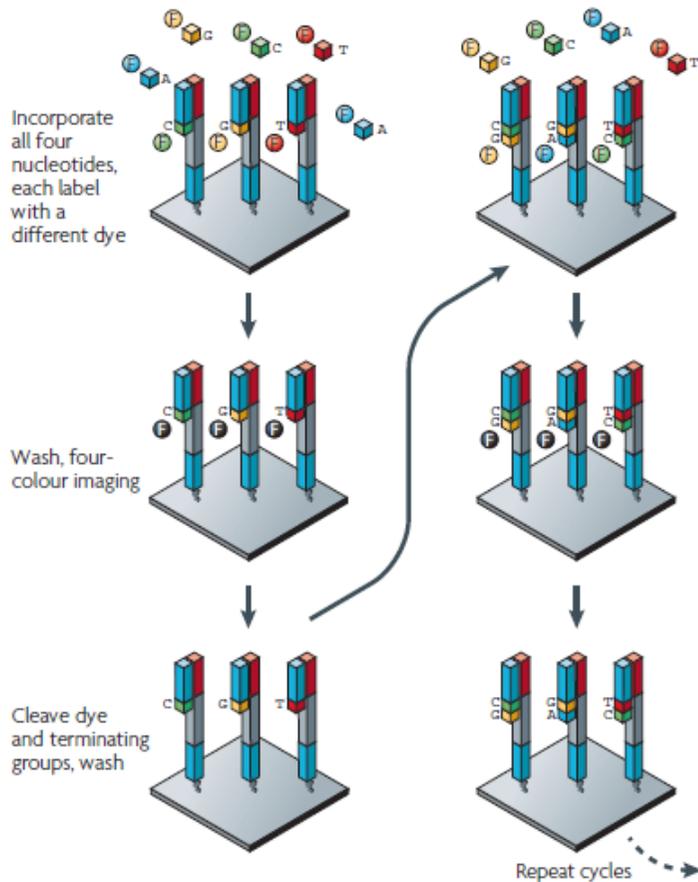


Séquençage : Identification de la succession des bases
Méthode de Sanger

Utilisation de 4 chromophores différents

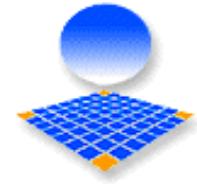
Lecture par scan à chaque étape de la synthèse

a Illumina/Solexa — Reversible terminators





Détection des "reads"



Condition	Echantillon	Nombre de reads
Jeune	CF52	12 199 550
Jeune	CF274	9 939 226
Jeune	CF986	9 035 638
Agée	CF232	9 861 433
Agée	CF268	10 257 894
Agée	CF269	9 218 643
ADlike	CF896	7 816 464
ADlike	CF973	7 182 893

Nombre moyen des lectures = 10 363 121 reads

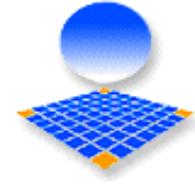
Vérification de la Qualité des séquences

Stockage des données

Analyse des données



Génome de référence: Mmur 2.0 (2015)



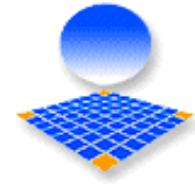
The screenshot shows a web browser window displaying the NCBI website. The address bar shows the URL: https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Microcebus_murinus/100/#BuildInfo. The page title is "NCBI Microcebus murinus Annotation Release 100".

The page content includes:

- NCBI Microcebus murinus Annotation Release 100**
- A paragraph explaining that the RefSeq genome records for *Microcebus murinus* were annotated by the NCBI Eukaryotic Genome Annotation Pipeline.
- A list of links for more information: [Annotation Release information](#), [Assemblies](#), [Gene and feature statistics](#), [Alignment of the annotated proteins to a set of high-quality proteins](#), [Masking of genomic sequence](#), and [Transcript and protein alignments](#).
- A section titled "Annotation Release information" with details: Annotation release ID: 100, Date of Entrez queries for transcripts and proteins: May 29 2015, Date of submission of annotation to the public databases: Jun 10 2015, Software version: 6.3.
- A section titled "Assemblies" with a table listing the assembly used.
- A section titled "Gene and feature statistics" with a note that counts and lengths of annotated features are provided below for each assembly.

Assembly name	Assembly accession	Submitter	Assembly date	Reference/Alternate	Assembly content
Mmur_2.0	GCF_000165445.1	Baylor College of Medicine and Broad Institute	05-01-2015	Reference	unplaced scaffolds

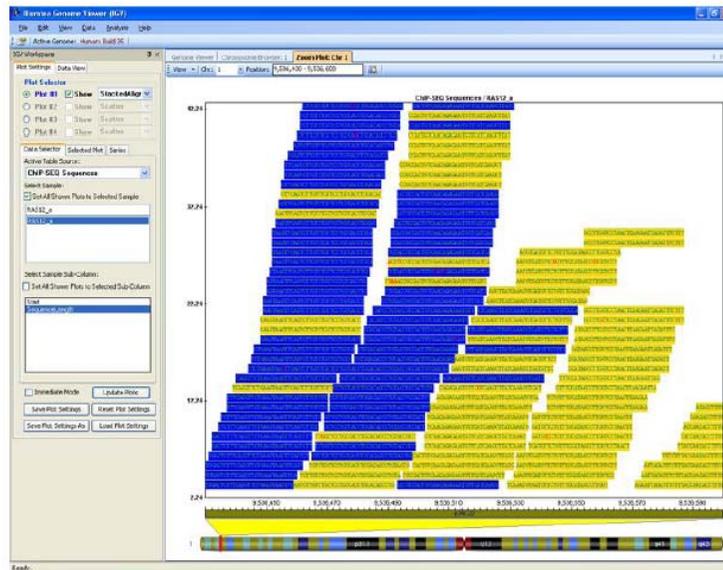
The Windows taskbar at the bottom shows the date and time as 13:36 on 12/10/2016, along with various application icons.



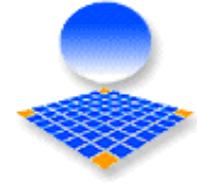
1- Alignement avec le logiciel CASAVA_Illumina

```

GTGCCCGATCTCGATCAGGT
GTGCCCGATCTCGATCAGGT
...GTGCCCGATCTCGATCAGGTAATCGTATATTAAGTGCCCGATCTCGATCAGGT...
CTCGATCAGGTAATCGT
GTAATCGTATATTA
TATATTAAGTGCCCGAT
TATATTAAGTGCCCGAT
TATATTA AGTGCCCGAT
    
```



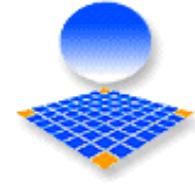
- 2 – Assemblage des fragments
- 3- Normalisation des données



Animaux	Nombre de transcrits
Jeunes	13 585
Agées	12 827
AD-like	13 677

16 244 transcrits détectés avec 0 ou 1 mismatch.

Poursuite de l'analyse pour découvrir de nouveaux transcrits.
Identification de site de régulation de la transcription,
Identification de site d'épissage alternatif



Fichier d'entrée Ensembl

ENSMICT00000000075	KIAA0494
ENSMICT00000000080	SLC29A3
ENSMICT00000015496	ATP2B4

Clé	Valeur
ENSMICT00000000075	KIAA0494
ENSMICT00000000080	SLC29A3
ENSMICT00000015496	ATP2B4

Fichier de résultats NGS

ENSMICT00000012857	503	229
ENSMICT00000007908	2495	1636
ENSMICT00000015496	2378	2706
ENSMICT00000002864	3745	2522

Fichier de sortie

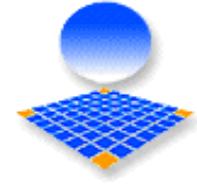
Identifiant <u>Ensembl</u>	Symbole du gène	Individu 1	Individu 2
ENSMICT00000015496	ATP2B4	2378	2706

Associer les noms des gènes
aux identifiants Ensembl

Nombre de reads par individu
pour ce transcrit

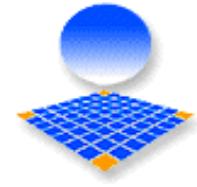


Comparaison des données NGS et biopuces



	Données NGS	Données Affymetrix HG
Détection Présent	16 000 Transcrits	14 000 Transcrits
Expression différente	2 258 gènes	1 050 gènes
Anova 1%		152 gènes

80 gènes



Comparaisons: AD-like versus Agées

Agées versus Jeunes adultes

AD-like versus Jeunes adultes

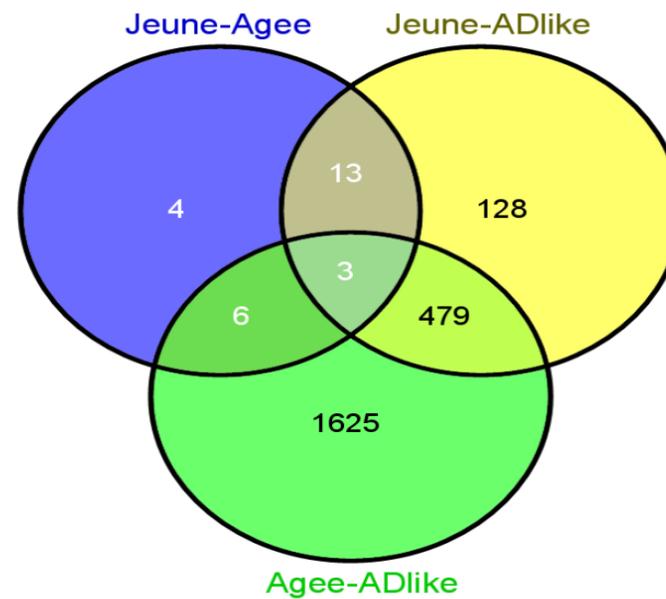
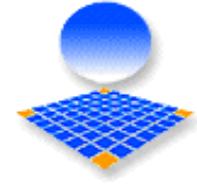
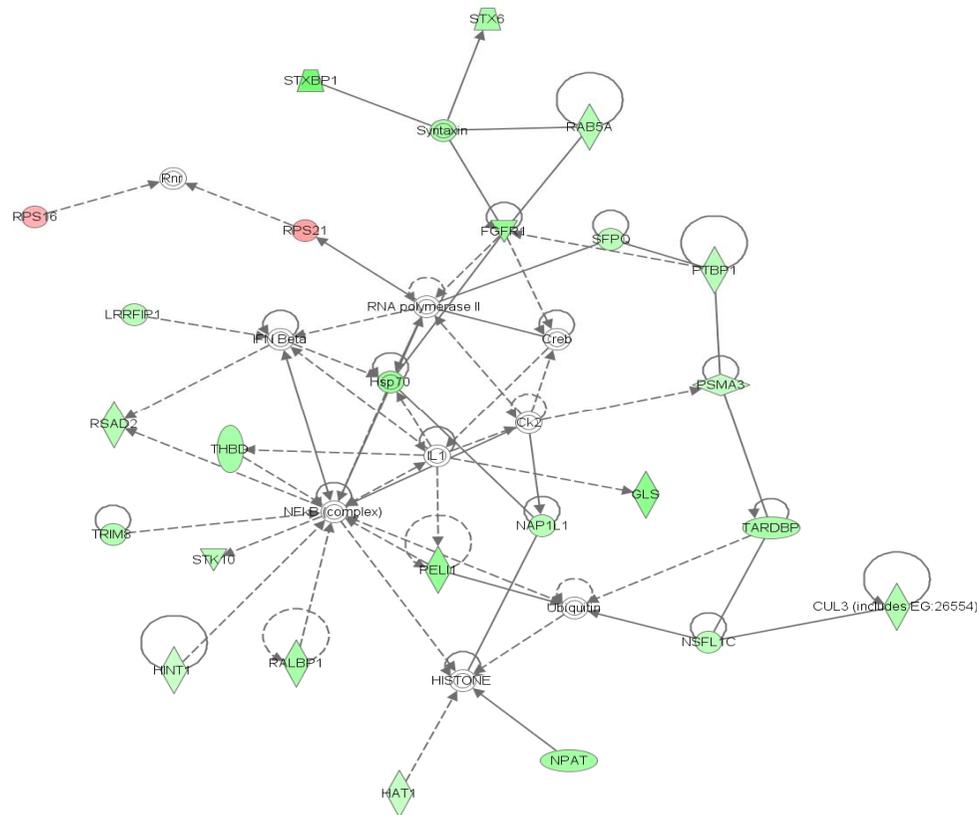


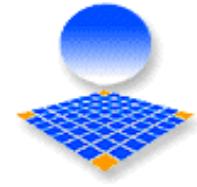
Diagramme de Venn

2258 gènes différentiellement exprimés



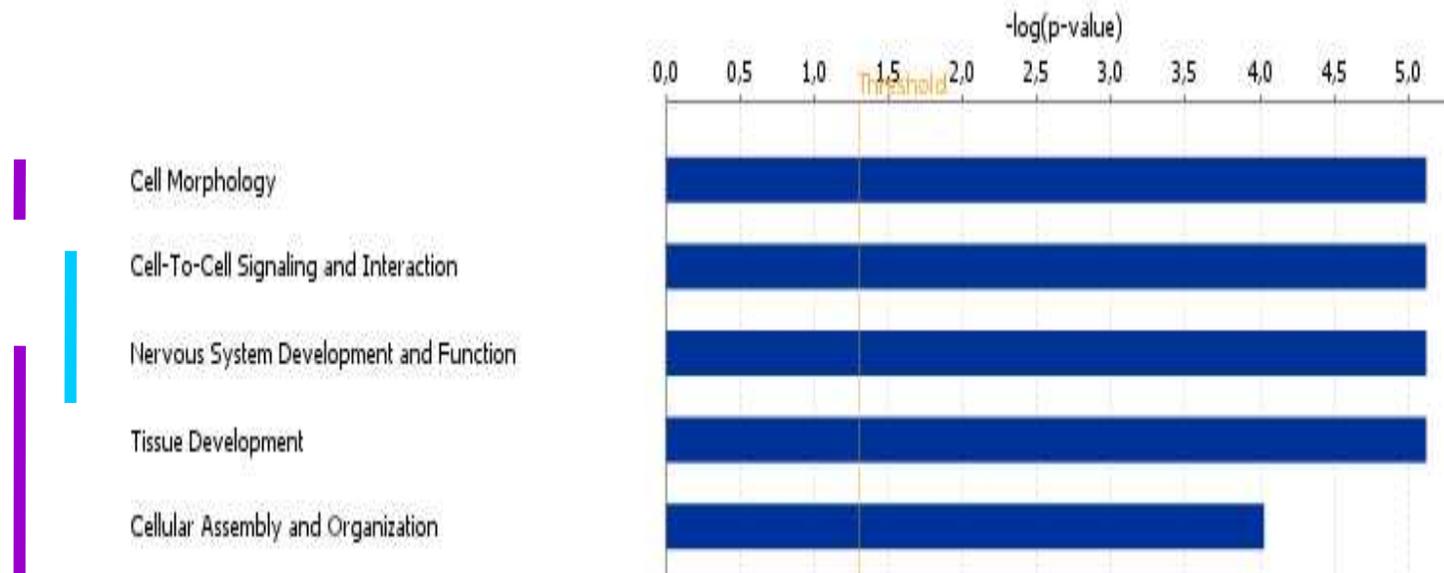
Construction du network d'interactions





Analyse des fonctions avec le logiciel Ingenuity

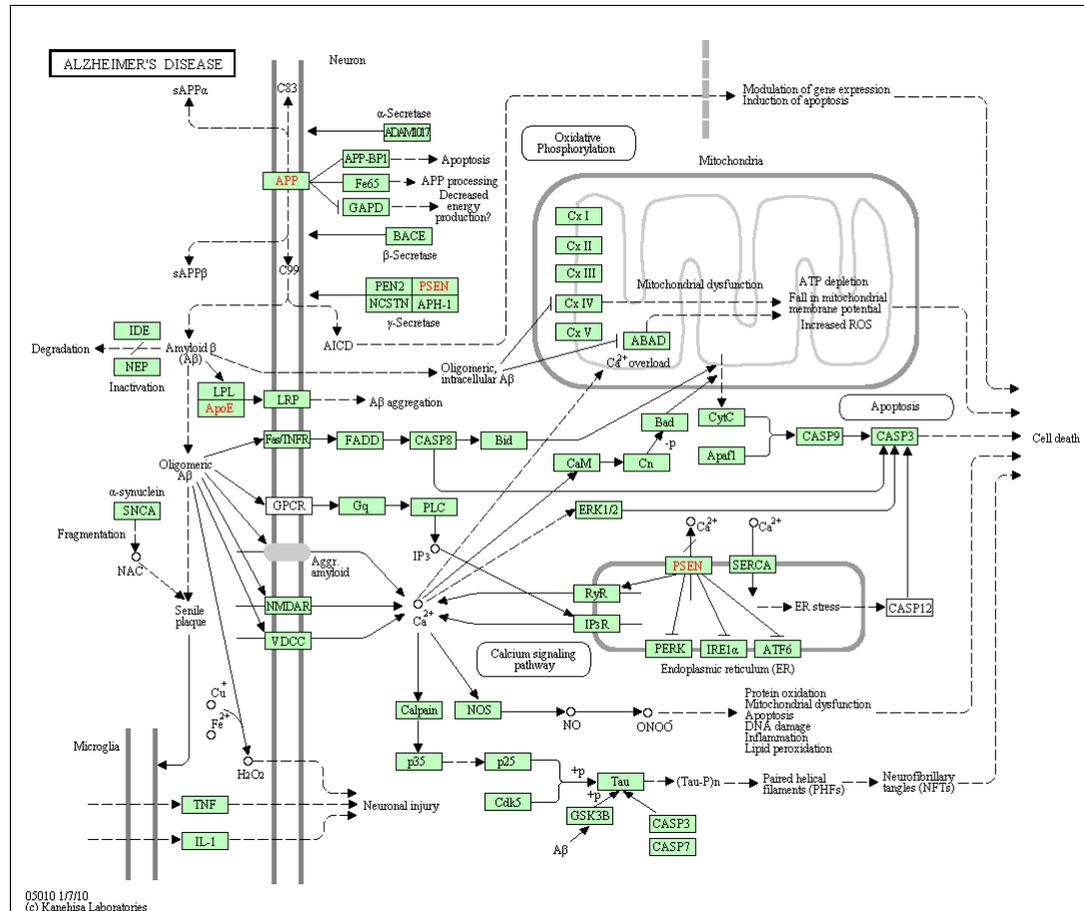
Data AD vs Aging NGS 2012 - 2012-06-21 03:02 PM



- Gènes impliqués dans une réorganisation cellulaire
- Gènes impliqués dans la communication cellulaire

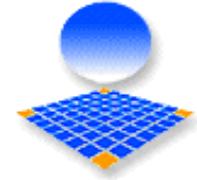
Signifiant biologique : construction de "pathways"

Alzheimer pathways construit par KEGG





Régulation de la transcription



Dans le vieillissement physiologique

Processus de Compensation:

Augmentation des synthèses protéiques (*RPS* et *RPL*)

Régulation de la transcription (*HAT*, *SERPIN*...)

Dans la pathologie

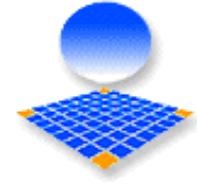
Perte des processus de compensation

Diminution des synthèses de protéines

Modifications de facteurs épigénétiques (*HAT1*, *PCAF*,...)



Potentiel des approches transcriptomiques



Les possibilités

- Détection des transcrits
- Quantification de l'expression des transcrits
- Comparaison de l'expression des transcrits
- Profils et signatures transcriptomiques

Les applications

- Augmentation de la compréhension des processus biologiques
- Diagnostic dans une approche médecine personnalisée
- Suivi du patient sur l'évolution de la maladie ou de l'efficacité du traitement.

Les procédures techniques

- Biopuces- Microarrays
- Séquençage à haut débit (NGS -Illumina)
- Très haut débit –séquençage de molécules uniques
- Développement des outils d'analyse des données