34e journéedu GAICRM
Symposium du 26 mars 2011
Suze la Rousse

Intégration des rétrovirus dans le génome humain
Fonctions et évolution

ProfesseurBernard David

Institut Pasteur

Très récemment, en 2001, la revue *Nature*a publié les résultats des travaux très attendus sur la séquence du génome humain, obtenue par le consortium public réunissant plusieurs laboratoires du monde entier. L'analyse de cette séquence révélait, entre autres, qu'au moins
8 % de l'ADN du génome humain était d'origine virale!Même s'il était connu que notre génome contenait, depuis plusieurs millions d'années, des traces de plusieurs séquences laissées par des virus infectieux (rétrovirus) dans certaines cellules, le fort pourcentage de cette intégration virale était surprenant. Certaines équipes de chercheurs ont donc poursuivi des travaux avec comme objectif d'analyser si ces séquences virales étaient susceptibles d'être fonctionnelles dans certaines cellules tant au plan physiologique que pathologique et aussi de connaître leur rôle éventuel dans l'évolution de la matière vivante

Historiquement, la nature et la structure des virus sont demeurées longtemps inconnues, du fait de l'absence de techniques permettant leur étude. C'est au début des années 40 que la présence d'acide nucléique a été mise en évidence chez de nombreux virus et en 1952 que se forge un concept de base de la biologie moléculaire, sur la base d'une démonstration particulièrement claire et exemplaire: le phage T4 d'Escherichia coli transmet son information héréditaire au moyen de son **ADN**tandis que ses protéines, constituant sa structure externe, demeurent à l'extérieur de la bactérie

Actuellement on définit les virus comme des particules contenant soit de l'**ADN** soit de l'**ARN,** constituant (l'un et l'autre) le matériel héréditaire transmissible de génération en génération. Un certain nombre de critères hiérarchisés permettent de classer les virus en très nombreuses familles, réparties en fonction du type de leur génome (ADN double ou simple brin, ARN double ou simple brin, ARN (+) double ou simple brin, ARN (-) double ou simple brin). Cette classification continue de s'enrichir de nouvelles familles, comme par exemple avec l'identification récente des **rétroviridés**(dont on connaît l'implication dramatique en santé publique).

En définitive, les virus ne peuvent se multiplier qu'à l'intérieur d'une cellule vivante. Cette multiplication, expression de la totalité du génome du virus (ADN ou ARN), a pour unique finalité la réplication du génome viral et son empaquetage dans une structure (capside) permettant de le véhiculer d'une cellule à l'autre.

Il est à noter que d'innombrables virus ne sont pas pathogènes, ou peuvent exister sous une «forme masquée» dans la cellule-hôte. L'exemple en est donné par les bactériophages (P1, P2, T4, λ…) qui sont des virus de bactéries, pouvant infecter ces cellules ou rester latents dans le génome bactérien. Il existe donc deux cycles, l'un ***productif*** ou ***lytique*** qui aboutit à la mort de la cellule infectée et l'autre en phase de ***dormance*** ou***lysogène***sans dommage pour la bactérie. En effet, on dit que la bactérie devient **lysogène,** en hébergeant un **prophage** (ou***provirus***) qui persiste de

génération en génération, sans cycle productif. De plus, ce cycle lysogèneest **inductible**suite à des stimuli tels que les UV qui excisent le génome du phage lequel est libéré dans la bactérie pour produire une lyse bactérienne. Le cycle lysogène dormant est devenu cycle lytique.C'est peut être le cas chez l'homme pour le virus de l'herpès et peut-être d'autres infections chroniques.

Pour s'intégrer au génome humain, le virus doit posséder un ADN, or tous les virus à ADN effectuent leur cycle productif en utilisant leur propre matériel génétique. Exceptionnellement, c'est un virus à ARN un **rétrovirus** de la famille des Rétroviridés qui va s'intégrer au génome humain en effectuant une synthèse d'ADN originale par rapport aux enzymes cellulaires, par le biais de la ***transcriptase inversevirale***mise en évidence par Temin en 1972.

Dès 2001, des groupes de chercheurs constatent que 8% du génome humain serait d'origine virale et plus précisément que les séquences analysées proviendraient des rétrovirus. Un certain nombre d'études approfondies ont tenté de décrypter ces séquences virales pour vérifier si elles pouvaient être potentiellement fonctionnelles

D'autre part, cette révélation va donner un éclairage nouveau sur le rôle important de ces vestiges ancestraux viraux au cours de l'évolution en référence à la ***théorie évolutionniste où le « monde » ARN aurait précédé l'ADN qui serait dès l'origine un ARN modifié (il y a de cela plus de 2 milliards d'années).***

Il est apparu que la plupart de ces séquences virales, appelées « rétrovirus endogènes » (ERVs), avaient au fil du temps subi des modifications les rendant incapables, pour la plupart, mais pas toutes, de coder la moindre protéine.

La séquence complète du génome humain a permis d'avoir une vue d'ensemble sur les ERVs humains **(HERVs**) qui ont été regroupées en fonction des homologies de séquences en ***80 familles distinctes***, chacune contenant une centaine d'éléments. Les fortes similarités entre

HERVs et les rétrovirus actuels peuvent être déduites des analyses phylogénétiques sur la transcriptase (gènes ***pol*** gène ***env***) qui révèlent un enchevêtrement des deux classes d'éléments et suggère une même histoire et des ancêtres communs. Des similarités sont aussi observées au niveau fonctionnel puisqu'en dépit du fait que la plupart des HERVs ont subi des mutations, délétions et troncatures qui se sont accumulées au cours de l'évolution, certains

éléments possèdent encore des fonctions rétrovirales avec formation de particules virales-like et l'apparition de protéines d'enveloppe dont certaines favorisent une fusion de cellule à cellule. **Deux exemples sont proposés sur la participation de ces HERVs dans un contexte physiologique (formation du placenta), l'autre pathologique (maladies auto-immunes).**

Au début des années 2000, plusieurs groupes ont identifié des protéines"rétrovirales" exprimées dans le placenta et en 2009 l'équipe de Thierry Heidman, a démontré, grâce à des souris génétiquement modifiées, que ces protéines étaient indispensables à la formation du placenta lors de la gestation, D'où l'hypothèse que les gènes qui les codent ont peut être contribué à l'émergence des mammifères placentaires, il y a 100 millions d'années,

Pour qu'une protéine virale soit impliquée dans la formation du placenta et soit transmise de génération en génération, il est nécessaire qu'elle provienne d'un ADN viral intégré dans le génome de la cellule. Or le seul virus qui répond à ces critères est le rétrovirus à condition qu'il infecte des cellules de la lignée germinale, car les virus ne s'introduisent que dans les cellules somatiques. La fréquence de ces évènements est donc rarissime, mais possible vu l'étendue de notre évolution.

On s'est aperçu qu'en général les rétrovirus endogènes étaient« dormants» au sein du génome, et donc incapables de former des particules infectieuses. De multiples mutations au cours du temps, ont abouti à l'inactivation de certains de leurs gènes. Cependant, des rétrovirus endogènes, incomplets donc non infectieux, ne sont pas totalement inactifs ; en effet, ils peuvent encore produire certaines protéines virales, en particulier les protéines d'enveloppe. Ces protéines d'enveloppe dites «fusiogènes » ont pour but de déclencher la fusion entre l'enveloppe des

rétrovirus et la membrane des cellules cibles avant leur pénétration. Mais, dans le cas des HERVs, elles ne peuvent être fonctionnelles que sur les cellules de l'hôte, donc avoir un rôle plus physiologique que pathologique. C'est ce qui se passe au niveau de la formation du placenta.

Au début des années 2000, l'équipe de T. Heidmann a repéré dans le **génome humain deux gènes de rétrovirus endogènes***codant des protéines ayant potentiellement des propriétés fusiogènes.* Ces deux protéines ont été baptisées **syncytine-l** et **syncytine-2.**

Repérées dans des échantillons de placenta, et uniquement là, ils ont émis l'hypothèse selon laquelle les **syncytines seraient impliquées dans la genèse du syncytium placentaire** et ont réalisé chez la souris des travaux expérimentaux

En 2005, cette même équipe a isolé deux gènes codant des protéines d'enveloppe fusiogènes s'exprimant spécifiquement au niveau du placenta, baptisées syncytine-A et syncytine-B gènes qui résultent de l'infection de rongeurs par deux rétrovirus distincts, il y a environ 20 millions d'années Les résultats obtenus à partir de souris génétiquement modifiées ont démontré que les gènes codant les syncytines seraient , sans aucun doute possible, impliqués dans la formation du placenta chez les mammifères actuels. Mieux cette équipe pense qu'ils ont peut-être joué un rôle plus important encore dans le passé et ***être à l'origine de l'apparition du placenta au cours de l'évolution, chez des mammifères archaïques de 70 millions d'années***.

Enfin leurs derniers travaux ont montré que les syncytines avaient une seconde propriété: comme les protéines d'enveloppe des rétrovirus infectieux, elles peuvent *inhiber les mécanismes de défense du système immunitaire.*

Concernant les syncytines du placenta, cette propriété immunosuppressive jouerait un rôle central dans l'instauration du mode de vie interactif entre le fœtus et la mère pendant la gestation. Elle expliquerait pourquoi le *fœtus,* dont les antigènes sont pour moitié paternels, n'est pas rejeté par le système immunitaire de la mère. Il a été suggéré que cette propriété immunosuppressive aurait entraîné l'apparition du placenta au cours de l'évolution. Les échanges entre la mère et le *fœtus* in utero auraient été rendus possibles par la protéine d'enveloppe d'un rétrovirus «fondateur", qui se serait intégré il *y* a environ 100 millions d'années dans le génome d'animaux pondant des œufs.

Finalement, il reste à confirmer définitivement si la fonction immunosuppressive des syncytines participe à l'établissement de la tolérance entre la mère et le fœtus.

Le deuxième exemple s'inscrit dans un contexte pathologique où il est suggéré que les **HERV**s pourraient être impliqués dans ***certaines maladies chroniques*** (cancers, maladies affectant le système nerveux et le tissu conjonctif, maladies rhumatoïdes et auto-immunes)**.** L'hypothèse émise est que l'insertion des HERVs proches de certains gènes de l'ADN de l'hôte pourrait altérer ceux-ci et créer une situation propice à amplifier les effets de certaines maladies. Une autre voie par laquelle ces maladies seraient en progression porte sur la sécrétion de protéines codées par les HERVs. Ces protéines seraient considérées comme étrangères par notre organisme, stimuler les lymphocytes B et produire des anticorps non seulement contre elles, mais aussi contre les protéines de notre corps par le biais de réactions croisées..

Ces mécanismes seraient susceptibles d'augmenter la sévérité, la fréquence et la pérennité des maladies auto-immunes (lupus érythémateux disséminé (LED), arthrite rhumatoïde…). D'autre part ces protéines (auto anticorps) pourraient agir comme des super antigènes. Enfin, certains auteurs évoquent l'action possible de certains agents de l'environnement (UV, produits chimiques) sur l'expression des gènes de HERVs en réactivant des séquences rétrovirales et en altérant leur structure sans modifier la composition nucléotidique de l'ADN. Ceci expliquerait les changements épigénétiques observées dans le LED. D'autres études doivent être poursuivies pour infirmer ou confirmer le rôle direct ou indirect des HERVs sur l'initiation et l'évolution de certaines maladies chroniques.

En conclusion, le génome comprend des ***séquences d'ADN*** d'origine **rétrovirale** (8%), vestiges de l'infection de cellules de nos ancêtres primates il y a plusieurs millions d'années et la plupart de

ces séquences semblent inactives ; cependant certaines d'entre elles codent encore des protéines. Il existe une forte présomption pour qu'elles puissent jouer chez l'homme un rôle en physiologie (placenta) ou en pathologie (maladies auto-immunes). Enfin ces séquences émanent des rétrovirus qui ont la capacité de transformer leur ARN en ADN (transcriptase reverse), ce qui les place dans une position favorable pour être à l'origine de la vie et de son évolution sur la terre**.L'ARN et /ou le virus à ARN serait-il à l'origine de la vie ?** La question est actuellement débattue par la communauté scientifique.

Mes remerciements à mes collègues de l'Institut Pasteur pour leurs précieux conseils.

Françoise Barré-Sinoussi: Chef de l'unité des Régulations des Infections Rétrovirales (Prix Nobel 2008)

Pierre Tiollais: Biologie moléculaire des infections virales

Fredj Tekaia: Unité de Génétique Moléculaire des Levures

Simon Wain-Hobson: Chef de l'unité de Rétrovirologie Moléculaire

Edouard Yeramian : Unité de Bio-Informatique Structurale

Bibliographie

Balada E, Ordi-Ros J, and Vilardell-Tarrés M. Molecular mechanisms mediated by Human Endogenous Retroviruses (HERVs) in autoimmunity Rev Med Virol 19: 273-286 (2009)

Blond JL Lavillette D, Cheynet V, and all. An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. J. Virol 74:321-339 (2000)

de Parseval N, Heidman T, Human endogenous retroviruses: from infectious elements to human genes. Cytogenet Genome Res 110:318 (2005)

de Parseval N, Casella JF, Gressin L, Heidman T. Characterisation of the three HERV-H proviruses with an open envelop reading frame encompassing the immunosuppressive domain and evolutionary history in primates. Virology 279: 558-569 (2001)

Forterre P. Des virus ont-ils inventé l'ADN? Pour la science, n° 55: 26-31(2007)

Forterre P, Grimaldo S, et Brochier C. Luca: à la recherche du plus proche ancêtre commun universel. MEDECINE/SCIENCE21:860-5 (2005)

Forterre P. L'origine du génome. La Recherche, 336, 34-39 (2000)

……

Place des rétrovirus à l'origine de la vie et de son évolution sur la planète Terre

- la vie est-elle d'origine virale?

- la nouvelle genèse version 2011

- Au commencement était ……..l'ARN

……

……..

LE MONDE ARN

- **ARN précurseur**: - information génétique et catalyse chimique
- **Ribozymes:**- ARN enzymes et auto réplication
- **Coexistence:** - ARN primitif et structures moléculaires (espace clos)
- **ARN ribosomique:** - ARN-enzymefossile
- **Protéines:** - « inventées » par l'ARN ribosomique
  (assemblage des Acides Aminés)

- **ADN:** - apparu à partir de l'ARN avec l'aide des protéines
- **Rétrovirus:** - ARN transcriptase inverse → ADN

Coexistence entre les cellules et les virus ARN et ADN
Sans capside, le virus est incapable de voyager

Automates moléculaires (ARN sans cellules) → virus → cellules

……

……...

L'EVOLUTION EN MARCHE

De la terre primitive à …….

….à LUCA (Last Universal common ancestor)……..

…...…de LUCA à l'embranchement des espèces……

……. des espèces……..à nos jours.

……

…… …….

Calendrier de la Terre

- **Il y a près de 4,6 milliards d'années: la Terre se forme.**
- **Il faudra attendre plusieurs centaines de millions d'années pour qu'elle se refroidisse et que la vie puisse débuter.**

……..

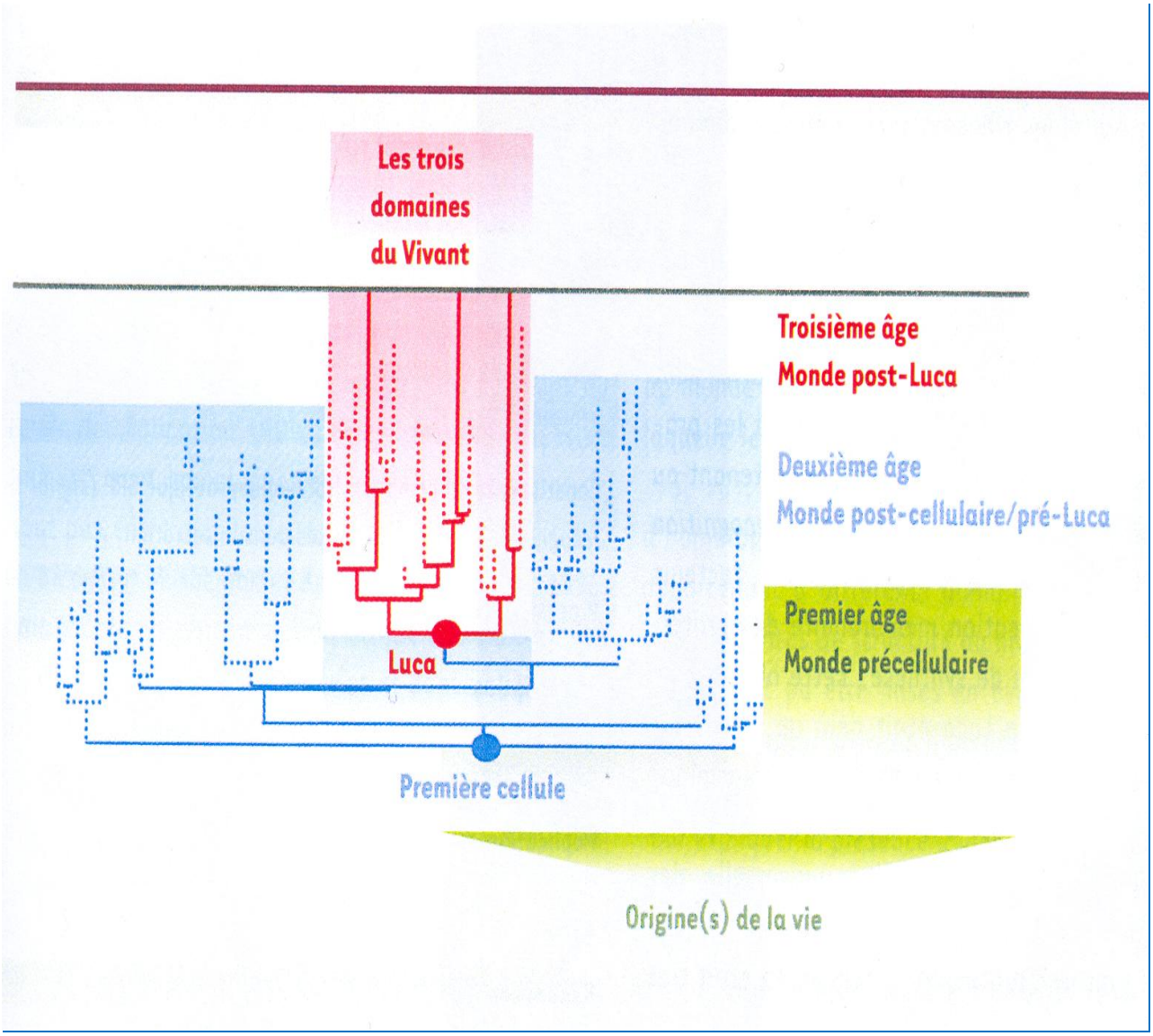
Evolution

4,5 Milliards (MM) d'années
Automates moléculaires (ARN sans cellules) → virus → cellules
↓
Première cellule: 3,5MM
↓
LUCA Last Universal common ancestor

- **3 milliards d'années: la vie est déjà florissante dans les océans?**
- **540-370 millions d'années: apparition des animaux à squelette externe**
- **apparition des insectes et des crustacés**
- **apparition des poissons**
- **premiers végétaux à sortir de l'eau: mousses**
- **arthropodes foulent la Terre**
- **apparition des premiers vertébrés à pattes**
  ce sont les tétrapodes
-
- **320 millions d'années: apparition des premiers reptiles**
- **230 millions d'années: apparition des dinosaures**
- **120 millions d'années: apparition des plantes à fleurs**
- **100 millions d'années: les mammifères: ancêtre commun des primates et des**
- **hominidés**

……

Evolution de la vie sur terre depuis son (ou ses) origine(s)



Luca : À la recherche du plus proche ancêtre commun universel

Patrick Forterre, Simonetta Gribaldo et Céline Brochier***M/S: Médecine sciences 2005 vol 21860-65***

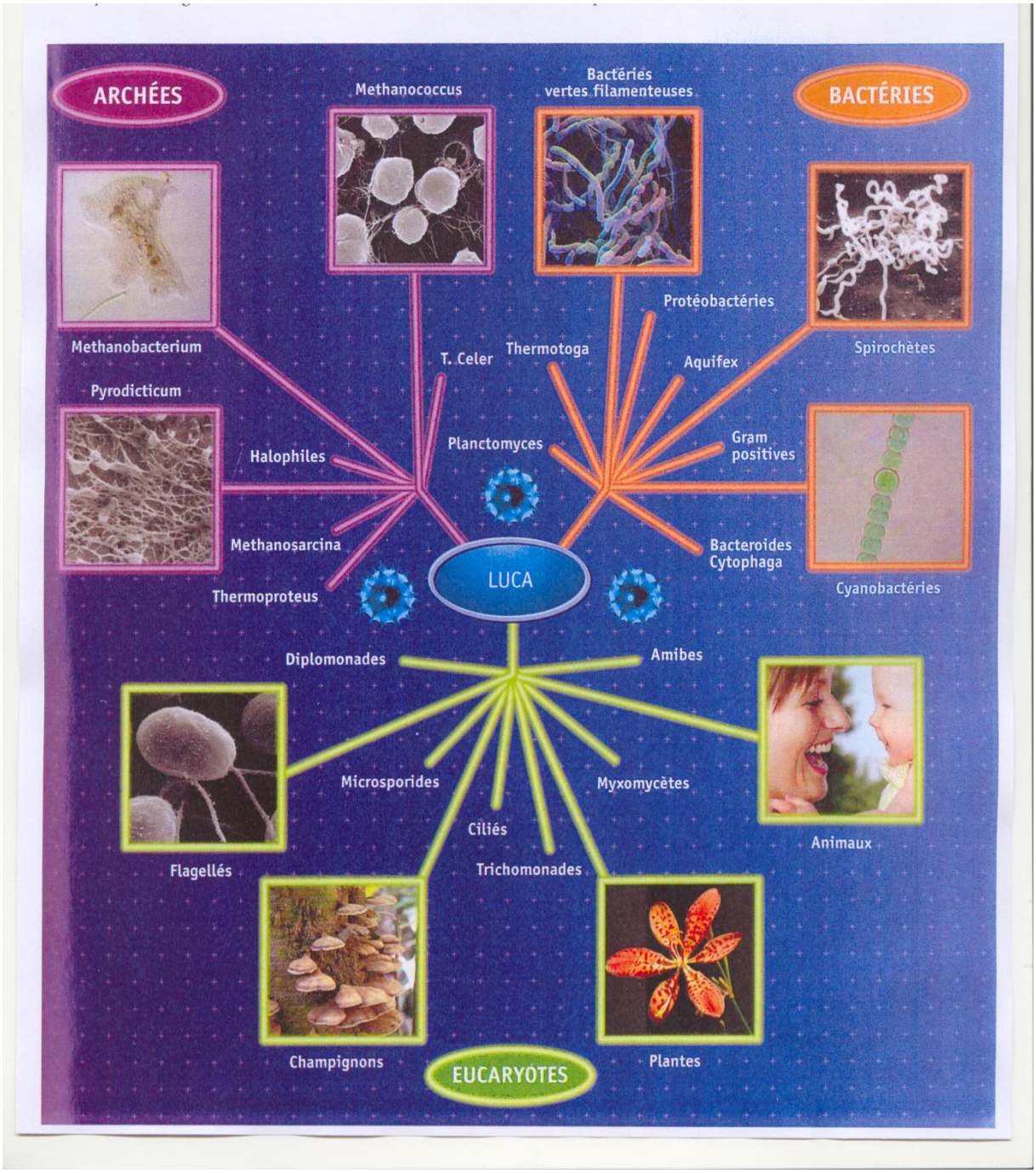
LUCA Last Universal common ancestor

↓

Eucaryotes Procaryotes Archées

↓ ↓

HOMMEBACTERIE



Dossier n° 55 Pour la Science avril - juin 2007

. Des virus ont-ils inventé l'ADN ? Patrick Forterre

L'Homme et les mammifères

100 millions d'années :les mammifères « **ancêtre commun des primates et des hominidés** »

70 millions d'années:les prosimiens

« le **rétrovirus** à l'origine des syncytines des simiens actuels a été capturé par un ancêtre primate entre **70 et 45 millions** d'années et le gène codant sa protéine d'enveloppe a été conservé jusqu'à nos jours sous forme codante et fonctionnelle dans toutes les branches primates dérivées ». (T. Heidman)

14 millions années: Les **grands singes** dont la branche a divergé des **hominidés**

7 millions années : Hominidé Toumaï(sahelanthropus tchadensis)

2 millions année :Homo ergaster(pierre taillée)

200.000 ans: Homo sapiens (150 milliards de neurones)

……

HOMO SAPIENS

« affligeante banalité biologique »

Sur le plan génétique, notre proximité est:

- avec le **chimpanzé**, de **98,7%**
- avec la **souris**, de **80%**
- avec la **levure**, de **50%**

D'où l'on vient????.....probabilités……. Où l'on va **????incertitude**



***Sahelanthropus tchadensis*, Toumaï, est le plus ancien hominidé connu avec ses 7 millions d'années. Il était bipède, comme l'indique la position du trou par lequel la colonne vertébrale pénétrait dans le crâne.**

La RechercheMensuel - n°438 février 2010 pp 38-54
La nouvelle histoire de l'Homme