

> L’empreinte parentale ou empreinte génomique impose un mode de reproduction sexué obligatoire chez les mammifères. Ce phénomène résulte de l’expression mono-allélique et monoparentale d’un groupe de gènes. Ce mécanisme de régulation génique concerne les mammifères vivipares comme les eutheriens et, dans une moindre mesure, les marsupiaux. Les mammifères ovipares, ou monotrèmes, ne semblent pas dotés d’empreinte parentale. Cette restriction phylogénétique conforte l’hypothèse selon laquelle l’émergence de l’organe placentaire aurait imposé une pression sélective pour l’acquisition de l’empreinte. Ces arguments physiologiques ont été plus récemment étayés par des arguments génomiques, grâce au séquençage du génome de l’étrange ornithorynque, un des rares monotrèmes encore existants. De nombreuses coïncidences temporelles et fonctionnelles existent entre l’apparition de l’empreinte et l’accumulation d’éléments transposables. Les analyses comparatives systématiques du génome de différentes espèces de mammifères, classiques et exotiques, sont essentielles à notre connaissance de l’émergence et de l’évolution de l’empreinte. <

Évolution de l’empreinte parentale chez les mammifères

Quelle ménagerie !

Charlotte Proudhon, Déborah Bourchis



UMR 3215/Inserm U934,
 Unité de génétique
 et biologie du développement,
 Institut Curie, 26 rue d’Ulm,
 75248 Paris Cedex 05, France.
Deborah.Bourchis@curie.fr
Charlotte.Proudhon@curie.fr

en captivité prolongée [1]. Chez les mammifères en revanche, la reproduction monoparentale est impossible. Des expériences de transfert nucléaire chez la souris ont démontré dans les années 1980 que le développement des mammifères requiert obligatoirement un génome maternel et un génome paternel [2, 3]. Nous avons célébré en 2009 les 25 ans de cette preuve formelle de la spécificité reproductive des mammifères liée à l’existence de l’empreinte parentale. Les mammifères ne sont cependant pas tous égaux face à l’empreinte parentale. Des contraintes physiologiques et génomiques ont ainsi contribué à l’émergence et à l’évolution de l’empreinte au cours de la spéciation des différents groupes de mammifères.

La majorité des espèces animales peuvent se reproduire sans partenaire sexuel, par activation d’un œuf non fécondé ou parthénogenèse. Les abeilles et les salamandres alternent naturellement entre reproduction sexué et parthénogenèse au cours de leur vie reproductive. La parthénogenèse peut aussi être déclenchée par une disponibilité limitée en partenaires sexuels : plusieurs cas de naissances parthénogénétiques ont été documentés chez des femelles requins et dragons de Komodo¹ maintenues

Définition biologique, génétique et épigénétique de l’empreinte parentale

La restriction à un mode de reproduction sexué chez les mammifères résulte d’une non-équivalence fonctionnelle des génomes parentaux, dictée par l’empreinte parentale ou empreinte génomique. L’empreinte parentale est un phénomène épigénétique qui affecte une centaine de gènes connus chez la souris et l’homme, leur imposant une expression monoallélique et dépendante de l’origine parentale. Le déterminisme parental résulte d’un marquage par méthylation de l’ADN, différenciellement apposé sur les allèles parentaux au moment de leur passage par la gamétogenèse.

¹ Le varan de Komodo (*Varanus komodoensis*) est le plus grand représentant actuel de la famille des Varanidae (dont font partie les lézards). Ces gigantesques varans (un mâle peut mesurer jusqu’à 3,10 m et peser 170 kg), capables d’engloutir un cerf en quelques minutes, nous font fantasmer. Komodo, l’île aux dragons, petite île volcanique au centre du triangle de Bornéo, est le dernier sanctuaire de ces grands reptiles.

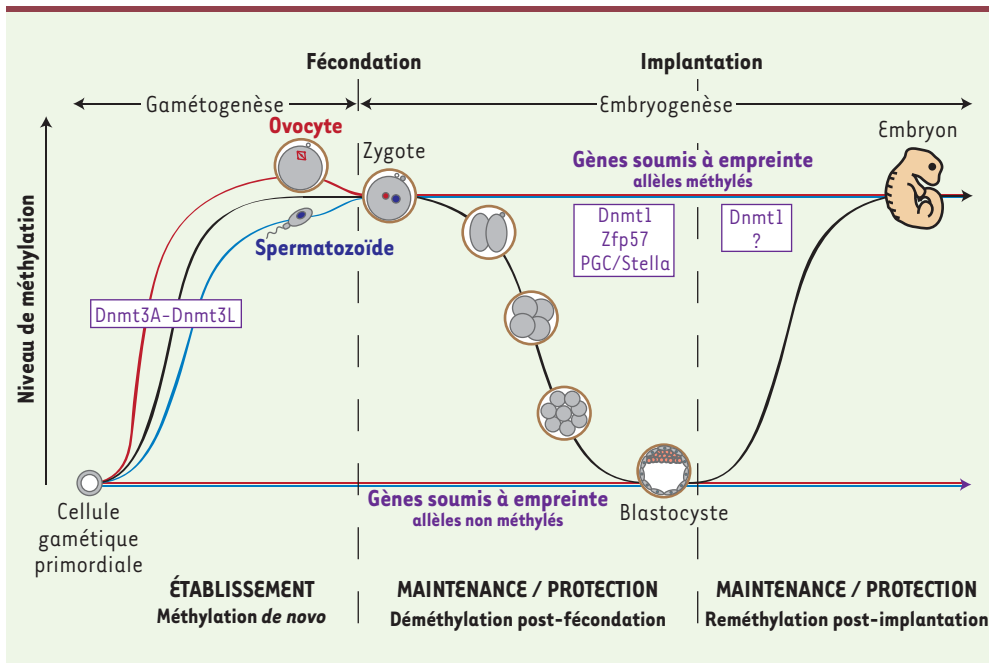


Figure 1. Établissement, maintenance et protection des marques de méthylation associées à l’empreinte parentale. Les courbes rouges et bleues représentent respectivement les allèles maternels et paternels des gènes soumis à empreinte et en noir le reste du génome. L’établissement de l’empreinte par apposition de marques de méthylation spécifiques du sexe se fait au cours de la gaméto-genèse. La maintenance de ces marques est assurée tout au long de la vie par Dnmt1. La protection de ces marques, par des facteurs identifiés ou non (?), est essentielle face aux vagues de déméthylation post-fécondation et de reméthylation post-implantatoire que subit le génome de l’embryon précoce.

La méthylation gamétique est spécifiquement ciblée sur des régions régulatrices appelées ICR (*imprinting control regions* ou centres d’empreinte) qui contrôlent en *cis* l’expression d’un ou plusieurs gènes. Les ICR maternelles sont méthylées spécifiquement dans l’ovocyte, tandis que les paternelles le sont dans le spermatozoïde. Ces marques d’empreinte ont la particularité d’être maintenues après la fécondation et interprétées par la répression d’un des deux allèles parentaux. Une question essentielle dans le champ de la biologie de l’empreinte est de comprendre l’identité unique des ICR. En particulier, quels sont les facteurs *trans* et *cis* responsables de l’établissement de l’empreinte et du maintien de la mémoire de l’origine parentale ?

La mise en place des marques de méthylation sur les centres d’empreinte est régulée lors de la gaméto-genèse par l’ADN méthyltransférase *de novo* Dnmt3A et son cofacteur Dnmt3L [4, 5] (Figure 1). Les marques d’empreinte sont conservées au cours de la vie par l’action de Dnmt1, l’ADN méthyltransférase de maintenance, qui reproduit les profils de méthylation au moment de la réplication [6]. Au cours du développement pré-implantatoire (de la fécondation à l’implantation dans l’utérus), les gènes soumis à empreinte doivent résister à des remaniements drastiques des profils de méthylation [7]. Les facteurs PGC/Stella et Zfp57 (*KRAB zinc finger protein*) jouent un rôle de protection contre la déméthylation qui suit la fécondation [8, 9]. Au moment de l’implantation, les allèles non méthylés des gènes soumis à empreinte doivent cette fois résister à la phase de reméthylation que subit l’embryon mais les facteurs responsables ne sont pas connus. Des caractéristiques nucléotidiques et épigénétiques particulières sont impliquées dans le recrutement des facteurs évoqués ci-dessus. Les gènes soumis à empreinte sont notamment préférentiellement associés à des régions riches en dinucléotides CpG, en éléments rétrotransposables de type LINE (*long interspersed elements*) et en motifs de liaison

pour CTCF (*CCCTC-binding factor*)² [10]. Cependant, ces caractéristiques n’étant ni obligatoires ni exclusives, la méthylation parentale des ICR est plus vraisemblablement corrélée à un état chromatinien particulier. Il a ainsi été récemment montré que la mise en place de la méthylation sur les ICR maternelles dans l’ovocyte requiert le retrait des marques de méthylation sur la lysine 4 de l’histone H3 (H3K4me) par l’histone déméthylase Lsd2 [11], en relation avec le caractère répulsif de ces marques envers Dnmt3L [12]. Des événements de transcription qui traversent les ICR maternelles pourraient également participer à ce remodelage chromatinien [13]. En conclusion, le statut de l’empreinte d’un gène résulte de sa proximité avec un centre d’empreinte dont l’identité est dictée par une combinaison complexe de motifs génomiques et épigénétiques influençant l’établissement et le maintien des profils de méthylation parentaux.

Évolution de l’empreinte parentale chez les mammifères : pas de placenta... pas d’empreinte !

Le mécanisme de répression d’un des deux allèles parentaux soulève la question de comment et surtout pourquoi l’empreinte parentale a évolué, sachant qu’elle nie l’avantage de la diploïdie qui peut compenser l’apparition d’une mutation délétère. La distribution phylogénétique de l’empreinte fournit des arguments fonctionnels à cette ambiguïté.

² Selon le contexte, CTCF peut fonctionner comme un activateur ou un répresseur transcriptionnel.

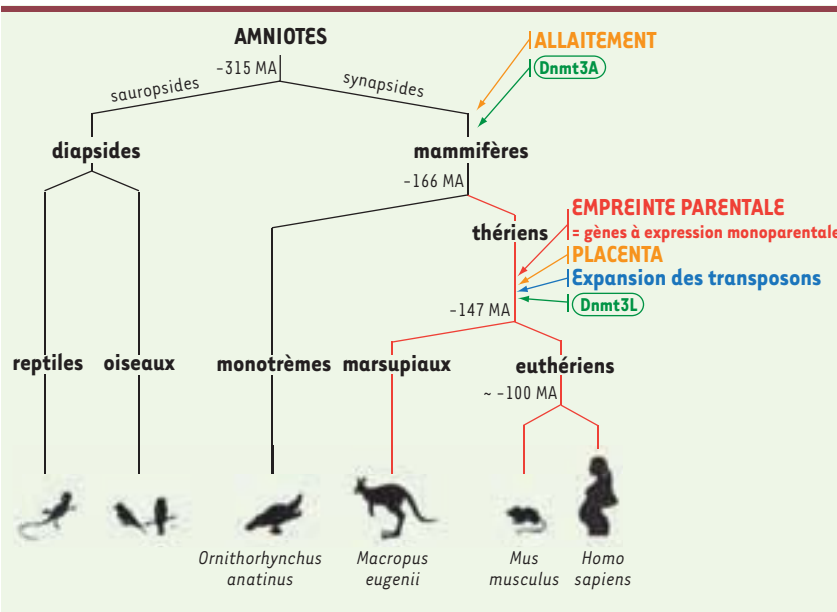


Figure 2. Apparition de l'empreinte parentale au cours de l'évolution. Trois groupes de mammifères existent : les monotrèmes et les thériens, composés eux-mêmes des marsupiaux et des euthériens. Tous ont recours à l'allaitement, tandis que le placenta est l'apanage des thériens. L'empreinte parentale est apparue avant la divergence des thériens, coïncidant avec l'acquisition du placenta, de Dnmt3L et de l'expansion des transposons. MA = millions d'années.

Hétérogénéité de l'empreinte : ornithorynque, marsupiaux, souris et homme

Les premiers signes d'empreinte parentale seraient apparus chez les mammifères il y a environ 150 millions d'années, chez l'ancêtre commun des thériens, après la divergence des monotrèmes (Figure 2). La branche des monotrèmes, dont le représentant actuel le plus connu est l'ornithorynque, est apparemment dépourvue de gènes à expression monoparentale. Les thériens comprennent les marsupiaux, dont le kangourou et l'opossum, et les euthériens, comme la souris et l'homme. Tous sont dotés de locus soumis à empreinte. Cependant, le petit nombre de locus observés chez les marsupiaux suggère que l'empreinte a été préférentiellement consolidée chez les euthériens, en fonction des besoins et de l'adaptation des espèces [14].

Les trois groupes de mammifères utilisent des modes de reproduction différents et il est possible de corréler l'évolution de l'empreinte parentale à des spécificités physiologiques. L'allaitement, associé à la présence de mamelles chez les femelles, est le dénominateur commun à tous les mammifères. Le développement *in utero*, ou viviparité, est une autre caractéristique des mammifères, mais n'est cependant pas universelle. Ainsi, les femelles monotrèmes allaitent leurs petits, mais sont ovipares. À la différence des oiseaux et des reptiles, la ponte de l'œuf est ici précédée d'une période relativement longue d'incubation dans le tractus maternel. Les marsupiaux, chez qui l'on retrouve une empreinte moins étendue que chez les euthériens, possèdent un placenta rudimentaire associé à un développement *in utero* très court. Chez le kangourou par exemple, la gestation dure 26 jours et s'achève par la migration du fœtus dans la poche, où la lactation prend le relais et peut s'étendre jusqu'à 10 mois. Par contraste, le groupe des euthériens arbore un placenta très développé, associé à une période de gestation longue et à un développement beaucoup plus abouti que chez les autres mammifères. Si l'on compare la masse relative des nouveau-nés par rapport à

leur mère au moment de la naissance, celle-ci est très inférieure chez les monotrèmes (0,08 %) et les marsupiaux (0,12 %) à ce qu'elle est chez les euthériens (15 %). Ces inégalités sont cependant compensées au moment du sevrage : la masse relative du nouveau-né y est équivalente dans les 3 groupes mammifères, ce qui illustre le rôle prédominant de la lactation pour le transfert de nutriments chez les monotrèmes et les marsupiaux (Figure 3) [15]. En conclusion, l'émergence de l'empreinte parentale chez les mammifères serait liée à l'apparition du placenta, et son amplification à la sophistication de cet organe.

Le « conflit des génomes parentaux »

D'un point de vue fonctionnel, le processus de lactation et les échanges placentaires génèrent chez les mammifères une asymétrie dans le transfert des ressources parentales à la progéniture, au profit d'une implication très largement maternelle. La théorie du « conflit des génomes parentaux », émise par Tom Moore et David Haig en 1991, tente d'expliquer d'un point de vue évolutif l'apparition de l'empreinte au sein du règne mammifère, en tenant compte de cet investissement parental inégal [16]. Cette théorie se base sur le profil d'expression des premiers gènes soumis à empreinte découverts au début des années 1990, le gène de croissance *Igf2* (*insulin growth factor 2*) à expression paternelle et son régulateur négatif *Igf2r* à expression maternelle [17, 18]. Par extrapolation, il a été postulé que les gènes à expression paternelle réguleraient positivement la croissance et l'apport de nutriments au fœtus, l'intérêt paternel étant de favoriser sa progéni-

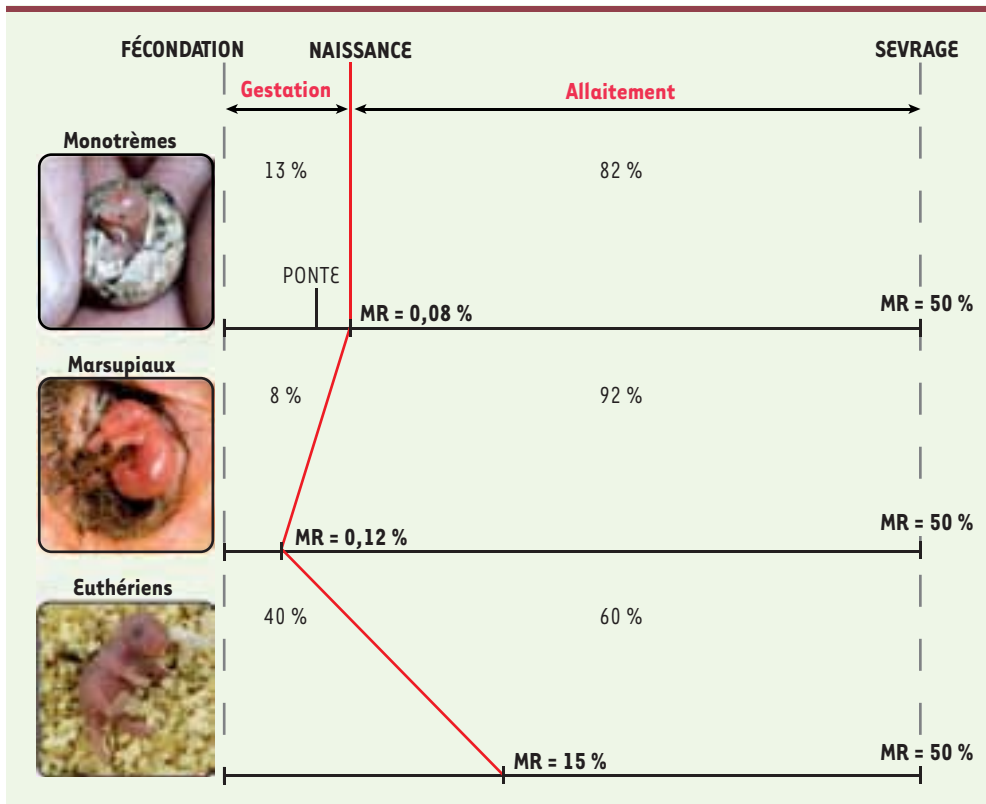


Figure 3. Différences physiologiques dans la reproduction des mammifères : proportion relative entre développement in utero et lactation. La lactation contribue de façon majeure au développement des petits chez les monotrèmes (82 %) et les marsupiaux (92 %) par rapport aux euthériens (60 %). Ces derniers ont en revanche recours à une phase intra-utérine nettement plus importante. MR = masse relative des nouveau-nés par rapport à celle de la mère.

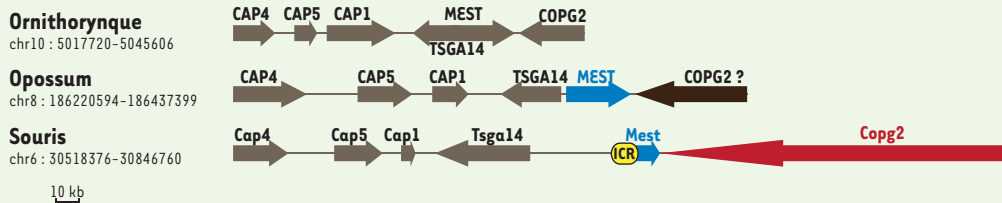
ture et de transmettre ses gènes. Par opposition, l'intérêt de la mère est de réduire le transfert de nutriments pour répartir ses ressources sur plusieurs portées, et les gènes à expression maternelle régulaient négativement la croissance et les transferts maternofœtaux. Les gènes soumis à empreinte actuellement décrits (une centaine) ont des fonctions souvent divergentes de la simple régulation de croissance ou du transfert nutritif. Plusieurs sont notamment exprimés dans le cerveau avec des fonctions essentielles au comportement reproductif des mammifères [19]. Une théorie actuelle moins « belliqueuse » prône ainsi un rôle coadaptateur de l'empreinte parentale, qui agirait pour la coordination de gènes impliqués dans les interactions entre la mère et ses petits, avant et après la naissance [20]. Le gène *Peg3* (*paternally expressed gene 3*), par exemple, améliore les capacités de reproduction par un double rôle. Son expression chez l'adulte contrôle la production de lait chez la mère, tandis qu'il influence la croissance placentaire et post-natale et le réflexe de succion chez le petit. Cependant, l'absence d'empreinte chez les monotrèmes suggère que le placenta reste la force évolutive dominante et/ou que le génome des monotrèmes n'est pas doté des régulateurs génétiques et épigénétiques nécessaires à l'évolution de l'empreinte.

Coévolution de l'empreinte parentale avec l'expansion des transposons

L'analyse génomique des espèces dotées et dépourvues d'empreinte a fourni des informations clefs quant à l'évolution du processus d'empreinte. Une centaine de gènes soumis à empreinte sont connus chez les euthériens, et seulement 5 sont répertoriés chez les marsupiaux. Seulement 2 de ces gènes sont de plus associés à des ICR porteuses de marques de méthylation gamétique. Ceci suggère un déploiement spécifique des ICR chez les euthériens ou, alternativement, l'utilisation par les marsupiaux d'un mode de régulation indépendant de la méthylation. Le consortium SAVOIR (*sequence analysis of vertebrate orthologous imprinted regions*) a entrepris depuis plusieurs années l'analyse comparative systématique chez le poulet, le wallaby/kangourou et l'ornithorynque des régions orthologues des locus soumis à empreinte connus [21]. La publication récente de la séquence du génome de l'ornithorynque a facilité l'identification des régions synténiques et des éléments fonctionnels importants, en relation avec l'évolution des génomes [22]. L'organisation des locus soumis à empreinte est généralement conservée chez tous les mammifères, avec une expansion des régions chez les euthériens, due à des phénomènes de duplication et de transposition. Cette observation, liée au fait que le génome de l'ornithorynque est relativement pauvre en éléments transposables, suggère une évolution coïncidente de l'empreinte avec la nécessité de contrôler une masse critique d'éléments génomiques parasites chez les mammifères supérieurs [23]. Le degré de conservation des facteurs impliqués dans la mise en place dévelop-

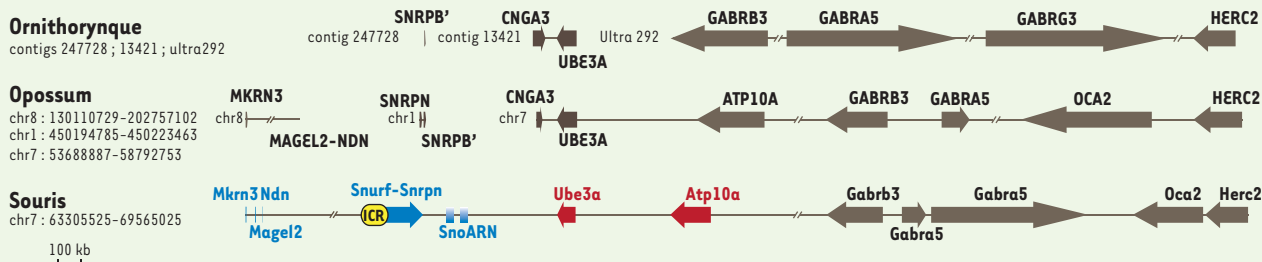
A Acquisition progressive à partir d'un locus conservé chez les mammifères

ex : locus *MEST/PEG1*



B Acquisition par remaniement chromosomique

ex : locus *SNRPN-UBE3A*



C Acquisition par rétrotransposition

ex : locus *PEG10*

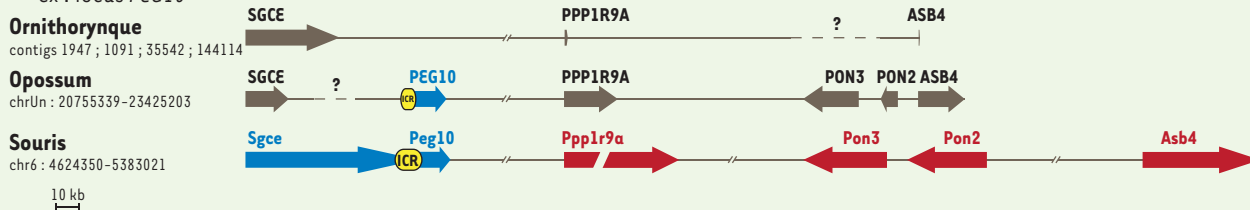


Figure 4. Mécanismes évolutifs d'acquisition d'un locus soumis à empreinte, illustration par analyse comparative chez les trois groupes de mammifères. Rouge : gène à expression maternelle, bleu : gène à expression paternelle, gris : gène à expression biallélique, jaune : ICR. **A.** Acquisition progressive des caractéristiques au sein d'un locus préexistant chez l'ancêtre commun de tous les mammifères. Exemple du gène *PEG1*, qui existe chez les monotrèmes, mais n'est exprimé paternellement que chez les thériens. L'ICR n'est présente que chez les eutheriens. **B.** Acquisition de l'empreinte par réarrangement chromosomique. Exemple du locus *UBE3A-SNRPN*. Chez la souris ce locus est situé sur le chromosome 7 et correspond à des blocs séparés sur les chromosomes 1, 7 et 8 chez l'opossum. **C.** Acquisition par rétrotransposition. Exemple du locus *PEG10*. Le rétrotransposon nouvellement inséré chez les marsupiaux est devenu le gène *Peg10*. L'ICR associée au promoteur du gène est présente chez l'opossum suggérant son apparition au moment de l'insertion du gène. La régulation monoparentale s'est ensuite étendue aux gènes voisins chez les eutheriens.

parentale des marques de méthylation est aussi instructif. Dnmt3A est présent chez tous les mammifères, tandis que Dnmt3L est spécifique des thériens [24]. Or Dnmt3L est aussi requis pour le contrôle par méthylation des éléments transposables dans les gamètes mâles chez la souris [25]. Le placenta, l'expansion génomique des transposons, Dnmt3L et l'empreinte parentale ont donc émergé de façon concomitante au cours de l'évolution des mammifères (Figure 2), sans que l'on puisse désigner les relations de cause à effet entre ces événements. Le scénario le plus accepté est : (1) l'acquisition du gène *Dnmt3L* comme élément de défense gamétique contre les transposons colonisant le génome des thériens, et (2) la cooptation de ce facteur pour la régulation de l'expression monoparentale d'une poignée de gènes, en relation avec la

fonction d'un organe placentaire complexe. Le plus grand nombre de gènes soumis à empreinte dans le placenta murin par rapport à l'homme suggère que la monogamie et la production moyenne d'un seul fœtus par gestation auraient induit une pression de sélection moindre dans l'espèce humaine [26].

Modus operandi de l'acquisition de nouveaux gènes soumis à empreinte

L'empreinte parentale est donc, comme nous l'avons décrit ci-dessus, hétérogène en quantité (nombre de

gènes) mais aussi en qualité (mode de régulation) selon les espèces. Elle représente un phénomène en marche, qui évolue par vagues de sélection au cours de la spéciation des mammifères. On peut définir deux modes de recrutement de nouveaux gènes soumis à empreinte. Les gènes de la première catégorie existent chez les monotrèmes, mais n'ont été sélectionnés pour une expression mono-allélique que chez les thériens. L'ICR fonctionnelle riche en motifs CpG et portant les marques de méthylation gamétique n'est souvent acquise que chez les euthériens, comme dans le cas du gène *PEG1* (Figure 4A). Des réarrangements chromosomiques ont aussi participé à la construction de larges locus soumis à empreinte, réunissant plusieurs blocs chromosomiques autour d'une même ICR apparue au même moment. L'exemple le plus connu est celui du locus associé aux syndromes de Prader-Willi et d'Angelman³ (PWS-AS), qui regroupe les gènes *SNRPN* (*small nuclear ribonucleoprotein N*) et *UBE3A* (*ubiquitin protein ligase E3A*) chez les euthériens, alors qu'ils sont situés sur des chromosomes différents chez les autres mammifères (Figure 4B) [27].

La seconde catégorie correspond à des gènes qui sont apparus spécifiquement chez les thériens et ont immédiatement acquis un statut soumis à empreinte. Plusieurs de ces gènes dérivent de rétrotransposons, dont la séquence a été domestiquée et utilisée pour développer de nouvelles fonctions cruciales pour la fonction placentaire. C'est notamment le cas des gènes *Peg10* (*paternally expressed gene 10*) et *Rtl1*, tous deux dérivés de l'insertion d'un rétrotransposon sushi-ichi⁴ de type LTR (*long terminal repeat*) et indispensables au développement du placenta murin (Figure 4C) [28, 29]. On voit encore ici la coïncidence évolutive entre l'empreinte parentale, l'organe placentaire et la défense de l'organisme contre les éléments transposables qui utilise la méthylation de l'ADN pour réprimer la mobilisation de séquences parasites endogènes et maintenir l'intégrité génomique. Dans cette catégorie de « néogènes soumis à empreinte », on trouve également un nombre important de gènes qui ont détourné la machinerie de rétrotransposition pour se mouvoir

et se dupliquer à une autre position génomique. Ce mécanisme est notamment à l'origine de l'acquisition des gènes *Nap1l5*, *U2af1-rs1*, *Inpp5f-v2* et *Mcts2*, qui, pour une raison encore inconnue, proviennent systématiquement d'un orthologue originaire du chromosome X [30]. Dans tous les cas de gènes dérivés directement ou indirectement de rétrotransposons, l'apparition d'une ICR fonctionnelle est simultanée ou postérieure à l'événement d'insertion. D'une manière générale, on observe aussi que toute nouvelle ICR tend à étendre son influence dans la région (Figure 4), induisant en *cis* l'expression mono-allélique d'un nombre croissant de gènes chez les euthériens, comme un recrutement fortuit à la cause de l'empreinte parentale.

Conclusion

L'empreinte parentale reflète un mécanisme d'expression singulier, appliqué à une centaine de gènes mammifères et dépendant de déterminants génétiques et épigénétiques. L'empreinte est phylogénétiquement liée à la viviparité et s'est répandue avec la complexité de la placentation. Les gènes soumis à empreinte sont généralement exprimés au cours du développement embryonnaire chez les mammifères placentaires, en accord avec une physiologie reproductive centrée sur la gestation *in utero*. Notre vision souffre cependant de plusieurs biais qui influencent notre interprétation évolutive de ce phénomène : (1) la majorité des études sont réalisées chez la souris et l'Homme, (2) l'ornithorynque est une espèce protégée et seuls des tissus adultes sont disponibles pour étude, (3) l'allaitement joue un rôle majeur dans l'asymétrie de transfert des ressources parentales chez les marsupiaux et les monotrèmes. Il est concevable de proposer qu'une empreinte associée à la lactation puisse exister chez ces espèces. Les organes liés à la fonction de l'allaitement, comme la glande mammaire et l'hypophyse, devraient être étudiés en priorité chez les marsupiaux et les monotrèmes pour l'identification de gènes à expression monoparentale. ♦

SUMMARY

Evolution of genomic imprinting in mammals: what a zoo !

Genomic imprinting imposes an obligate mode of biparental reproduction in mammals. This phenomenon results from the monoparental expression of a subset of genes. This specific gene regulation mechanism affects viviparous mammals, especially eutherians, but also marsupials to a lesser extent. Oviparous mammals, or monotremes, do not seem to demonstrate monoparental allele expression. This phylogenetic confinement suggests that the evolution of the placenta imposed a selective pressure for the emergence of genomic imprinting. This physiological argument is now complemented by recent genomic evidence facilitated by the sequencing of the platypus genome, a rare modern day case of a monotreme. Analysis of the platypus genome in comparison to eutherian genomes shows a chronological and functional coincidence between the appearance of genomic imprinting and transposable element accumulation. The systematic comparative analyses of genomic sequences in different species is essential for the further understanding of genomic imprinting emergence and divergent evolution along mammalian speciation. ♦

³ Le syndrome de Prader-Willi (1/15 000 naissances) associe une hypotonie sévère, une hyperphagie entraînant une obésité, une petite taille, un hypogonadisme hypogonadotrophique, des mains et des pieds courts, un visage caractéristique, un retard mental modéré. Il résulte de l'absence ou de l'inactivation d'un ensemble de gènes de la région 15q12 du chromosome 15 paternel. Le syndrome d'Angelman (1/20 000 naissances) associe un retard mental sévère, une absence de langage, une microcéphalie, une ataxie, des convulsions, un EEG caractéristique et parfois une hypopigmentation. L'origine est l'absence de contribution maternelle du gène *UBE3A*. Seule la copie maternelle de ce gène est active dans le cerveau, alors que l'expression du gène est biallélique dans tous les autres tissus.

⁴ Ces rétrotransposons sushi-ichi sont représentatifs d'une famille de rétrotransposons détectés en de multiples copies dans le génome de poissons (*fugu fish Fugu rubripes*).

