

Le gène flamenco de la drosophile, ou comment résister à un rétrovirus

Alain Bucheton

Pendant longtemps, on a pensé que les rétrovirus étaient spécifiques des vertébrés. Or, des résultats récents montrent que l'élément gypsy de la drosophile, considéré habituellement comme un élément transposable de la classe des rétrotransposons, est un véritable rétrovirus, le premier identifié chez les invertébrés. Un gène de l'hôte, appelé *flamenco*, maintient le virus à l'état réprimé dans la plupart des souches de drosophile. Lorsque le gène est muté, gypsy est transposé avec des fréquences très élevées et produit des particules infectieuses. De nombreux résultats indiquent que le rétrovirus gypsy a envahi l'espèce *Drosophila melanogaster* dans un passé relativement ancien, et que l'espèce n'a survécu à cette invasion que grâce à la sélection d'allèles du gène *flamenco* permettant de bloquer la propagation du virus. L'analyse des relations phylogénétiques entre rétrotransposons et rétrovirus indique que gypsy pourrait correspondre à la forme ancestrale des rétrovirus, à partir de laquelle les rétrovirus des vertébrés sont apparus.

Les rétrovirus ont la remarquable propriété d'intégrer leur information génétique, contenue dans la particule virale sous forme d'ARN, dans les chromosomes des cellules qu'ils infectent, grâce à la transcriptase inverse pour laquelle ils codent. Ce génome viral intégré est connu sous le nom de provirus. Certains provirus sont intégrés dans les chromosomes de la lignée germinale. Ils sont alors appelés rétrovirus endogènes, et ils sont transmis de génération en génération au même titre que les gènes de l'hôte, produisant

occasionnellement des particules virales infectieuses susceptibles d'infecter de nouveaux individus.

Jusqu'à une date récente, on pensait que les rétrovirus n'étaient présents que chez les vertébrés. Ils ont été notamment bien étudiés chez les oiseaux et les mammifères. Cependant, le génome de la plupart des groupes eucaryotes (par exemple les invertébrés, les plantes, les champignons, mais aussi certains vertébrés) contient des séquences dont la structure rappelle étrangement celle des provirus des rétrovirus, connues sous le nom de rétrotransposons à LTR

ADRESSE

A. Bucheton : directeur de recherche au Cnrs. Centre de génétique moléculaire, Cnrs UPR 2420, bâtiment 26, avenue de la Terrasse, 91198 Gif-sur-Yvette, France.

(*long terminal repeat*), ou rétrotransposons de type viral. Ces éléments représentent, en fait, une vaste classe d'éléments transposables, c'est-à-dire de parasites du génome ayant la propriété de se déplacer dans les chromosomes de leurs hôtes. Temin a avancé l'hypothèse selon laquelle les rétrovirus auraient pour origine les rétrotransposons à LTR [1]. Des résultats récents montrent que les rétrovirus se propagent, en fait, au-delà des vertébrés, ce qui apporte des informations intéressantes sur leur origine et sur leurs relations avec les rétrotransposons. La découverte d'un rétrovirus chez la drosophile, un des modèles d'étude préféré des généticiens, ouvre des perspectives très importantes pour l'étude du contrôle des rétrovirus par leur hôte.

Rétrovirus et rétrotransposons

Le rétrovirus le plus étudié est sans aucun doute le VIH, agent du SIDA. Le virus de la leucémie murine de Moloney (MoMLV) est plus simple et peut être pris comme modèle d'organisation générale des rétrovirus (figure 1). Ces derniers sont caractérisés par la présence de séquences de quelques centaines de nucléotides répétées en orientation directe aux extrémités du provirus, les LTR. Il existe trois gènes rétroviraux typiques, localisés dans la partie interne du génome viral comprise entre les LTR. Le gène *gag* code pour les constituants de la nucléocapside. Le gène *pol* code pour les enzymes nécessaires au cycle rétroviral: une protéase (PR), nécessaire au clivage des produits de traduction du rétrovirus, qui sont synthétisés sous la forme d'une polyprotéine; la transcriptase inverse (RT pour *reverse transcriptase*), requise pour produire l'ADNc à partir de l'ARN; une RNase H (RH) responsable de la dégradation de l'ARN dans la molécule hybride ADN/ARN résultant de la transcription inverse; enfin, l'intégrase (IN) nécessaire à l'intégration du provirus dans l'ADN de l'hôte. Le troisième gène, *env*, code pour les protéines de l'enveloppe du virus qui sont responsables de son pouvoir infectieux. Les rétrotransposons de type viral se caractérisent eux aussi par la présence de LTR typiques, mais ils ne contiennent que deux phases ouvertes de lecture (ORF, *open reading frame*) corres-

tésés par la présence de séquences de quelques centaines de nucléotides répétées en orientation directe aux extrémités du provirus, les LTR. Il existe trois gènes rétroviraux typiques, localisés dans la partie interne du génome viral comprise entre les LTR. Le gène *gag* code pour les constituants de la nucléocapside. Le gène *pol* code pour les enzymes nécessaires au cycle rétroviral: une protéase (PR), nécessaire au clivage des produits de traduction du rétrovirus, qui sont synthétisés sous la forme d'une polyprotéine; la transcriptase inverse (RT pour *reverse transcriptase*), requise pour produire l'ADNc à partir de l'ARN; une RNase H (RH) responsable de la dégradation de l'ARN dans la molécule hybride ADN/ARN résultant de la transcription inverse; enfin, l'intégrase (IN) nécessaire à l'intégration du provirus dans l'ADN de l'hôte. Le troisième gène, *env*, code pour les protéines de l'enveloppe du virus qui sont responsables de son pouvoir infectieux. Les rétrotransposons de type viral se caractérisent eux aussi par la présence de LTR typiques, mais ils ne contiennent que deux phases ouvertes de lecture (ORF, *open reading frame*) corres-

RÉFÉRENCES

1. Temin HM. Origin of retroviruses from cellular moveable genetic elements. *Cell* 1980; 21: 599-600.
2. Vaury C, Bucheton A, Pélisson A. The β -heterochromatic sequences flanking the I elements are themselves defective transposable elements. *Chromosoma* 1989; 98: 215-24.
3. Kim AI, Belyaeva ES. Transposition of mobile elements *gypsy* (*mdg4*) and *hobo* in germ-line and somatic cells of a genetically unstable mutator strain of *Drosophila melanogaster*. *Mol Gen Genet* 1991; 229: 437-44.
4. Mevel-Ninio M, Mariol MC, Gans M. Mobilization of the *gypsy* and *copia* retrotransposons in *Drosophila melanogaster* induces reversion of the ovoD dominant female-sterile mutations: molecular analysis of revertant alleles. *EMBO J* 1989; 8: 1549-58.
5. Kim AI, Lyubomirskaya V, Belyaeva ES, Shostack NG, Ilyin YV. The introduction of a transpositionally active copy of retrotransposon *gypsy* into the stable strain of *Drosophila melanogaster* causes genetic instability. *Mol Gen Genet* 1994; 242: 472-7.
6. Kim A, Terzian C, Santamaria P, Pélisson A, Prud'homme N, Bucheton A. Retroviruses in invertebrates: *gypsy* is an infectious retrovirus of *Drosophila*. *Proc Natl Acad Sci USA* 1994; 91: 1285-9.
7. Song SU, Gerasimova T, Kurkulos M, Boeke JD, Corces VG. An Env-like protein encoded by a *Drosophila* retroelement: evidence that *gypsy* is an infectious retrovirus. *Genes Dev* 1994; 8: 2046-57.

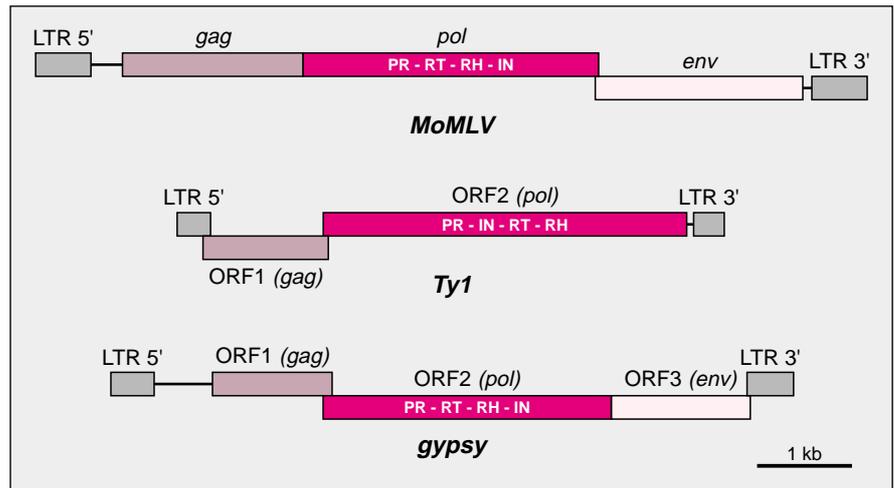


Figure 1. **Rétrovirus et rétrotransposons.** La figure représente l'organisation d'un rétrovirus de mammifère, le virus de la leucémie murine de Moloney (MoMLV), d'un rétrotransposon à LTR de la levure, l'élément Ty1, et d'un rétroélément de la drosophile, le rétrovirus *gypsy*. Les LTR sont représentés en grisé, et les phases ouvertes de lecture (ORF) sont schématisées par des rectangles. Les différents domaines du gène *pol* sont figurés dans l'ordre où ils sont localisés dans la phase de lecture pour chacun des éléments. Ces différents domaines dans l'élément *gypsy* sont organisés de la même manière que chez les rétrovirus des vertébrés. PR, protéase; RT, transcriptase inverse; RH, RNase H; IN, intégrase.

pendant aux gènes rétroviraux *gag* et *pol* (figure 1). L'ensemble des informations dont on dispose indique qu'ils transposent selon un mécanisme tout à fait semblable au processus d'intégration des provirus rétroviraux dans le génome de leur hôte. La principale différence est finalement que, étant dépourvus d'un gène d'enveloppe, leur cycle est entièrement intracellulaire. Ils ne sont donc pas infectieux et ne peuvent se maintenir que parce qu'ils sont présents dans la lignée germinale. Les éléments Ty1 chez la levure, copia chez la drosophile, Tnt1 chez le tabac, les IAP chez les rongeurs ou THE chez l'homme sont des exemples de ce type de rétrotransposons très largement répandus chez les eucaryotes.

La vision initiale d'une distribution des rétrovirus restreinte aux vertébrés posait cependant un certain nombre de questions. En effet, le génome des insectes possède des séquences qui sont beaucoup plus proches des rétrovirus que des éléments de type copia. C'est le cas par exemple de l'élément gypsy de *Drosophila melanogaster* (figure 1). Tout d'abord, cet élément possède trois ORF, au lieu de deux seulement pour les éléments de type Ty1 ou copia. D'autre part, la séquence et l'organisation du gène *pol* de gypsy sont beaucoup plus proches de celles des rétrovirus des vertébrés que de celles des éléments de type copia, dont un grand nombre cohabite pourtant avec lui dans la même espèce. Par exemple, les différents domaines du gène *pol* des rétrotransposons sont disposés dans l'ordre PR – IN – RT – RH, de l'extrémité 5' vers l'extrémité 3', alors que chez les rétrovirus, l'ordre est systématiquement PR – RT – RH – IN (figure 1). Le gène *pol* de gypsy possède cet arrangement typique des rétrovirus (figure 1). Des éléments ressemblant à gypsy ont été mis en évidence, parfois depuis longtemps, chez plusieurs espèces d'insectes. Il s'agit de 297 et 17.6 chez *Drosophila melanogaster*, tom chez *Drosophila ananassae* et TED chez *Trichoplusia ni*. Malgré leurs similitudes avec les provirus rétroviraux, ces éléments étaient considérés comme des éléments transposables, sans doute en vertu du dogme selon lequel les rétrovirus ne sont présents que chez les vertébrés.

La démonstration de propriétés infectieuses d'éléments de ce type chez les insectes n'est toutefois pas très facile. Cela implique qu'au moins trois conditions puissent être réunies. Tout d'abord, il est nécessaire de disposer de souches dans lesquelles le virus est actif et susceptible de se multiplier et de produire des particules infectieuses. De telles souches pourraient servir de donneuses dans des expériences destinées à mettre en évidence un pouvoir infectieux. Il faut également disposer de souches ne contenant pas le virus, pouvant être utilisées comme receveuses. Il est enfin nécessaire de disposer d'un test permettant de détecter facilement la présence du virus.

L'activité de gypsy conduit à des instabilités génétiques

Toutes les souches de *Drosophila melanogaster* contiennent des éléments gypsy défectifs localisés dans l'hétérochromatine péricentromérique. Des expériences de *Southern blot* montrent qu'ils sont localisés aux mêmes sites dans des souches d'origine indépendante, ce qui suggère qu'ils se sont intégrés dans ces régions chromosomiques très tôt au cours de l'histoire évolutive de l'espèce, peut-être même dans une espèce ancêtre [2]. Beaucoup d'autres souches possèdent, en plus, des éléments gypsy localisés dans l'euchromatine, à des sites différents selon les souches, ce qui indique qu'au moins certains d'entre eux sont capables de transposer et sont donc fonctionnels. Il n'existe qu'un très petit nombre de ces éléments dans la plupart des souches, souvent moins de trois ou quatre par génome. Bien que fonctionnels, ils ne transposent pas avec des fréquences appréciables car ils sont réprimés.

Des souches contenant des éléments actifs en nombre bien supérieur ont toutefois été caractérisées. Gypsy transpose apparemment avec des fréquences très élevées dans ces souches, et certaines peuvent contenir jusqu'à 50 ou 60 copies de l'élément. Si gypsy produit des particules infectieuses, on peut penser que celles-ci sont particulièrement abondantes dans de telles souches.

Deux d'entre elles, appelées MG et MS, ont été spécialement étudiées. La fréquence élevée de transposition de gypsy se traduit dans ces deux cas par une forte mutabilité, les mutations produites résultant de l'insertion de l'élément dans des gènes [3, 4]. En particulier, de nombreuses mutations se produisent dans le gène *ovo* [4], ce qui permet, comme nous allons le voir, de détecter et de quantifier avec une relative facilité la transposition de gypsy.

Le gène *ovo*, localisé sur le chromosome X, est requis pour la différenciation de la lignée germinale femelle. Il existe une mutation dominante de ce gène, appelée *ovo^{DI}*, qui cause la stérilité des femelles qui le possèdent à l'état hétérozygote, leurs ovaires ne se développant pas. Cependant, lorsque des femelles des souches MG et MS sont croisées avec des mâles portant la mutation *ovo^{DI}*, elles produisent des filles fertiles avec une fréquence pouvant atteindre 15%. Ce phénomène résulte de la mutation de l'allèle dominant *ovo^{DI}* vers un allèle récessif nul, permettant ainsi le développement d'un ovaire. La mutation est due à l'insertion de l'élément gypsy dans le gène *ovo*, dans les cellules souches de la lignée germinale. Ce gène est donc un point très chaud pour l'insertion de l'élément gypsy [4]. La réversion phénotypique de la mutation *ovo^{DI}* représente un moyen simple de détecter l'activité de gypsy, puisqu'elle fournit un crible positif très puissant consistant à rechercher des femelles fertiles dans une population de femelles stériles.

Il existe également des souches de drosophile ne possédant aucun élément gypsy fonctionnel. C'est le cas notamment d'une souche, appelée SS (*stable strain*). Lorsque des femelles de cette souche sont croisées avec des mâles *ovo^{DI}*, elles ne produisent que des filles stériles. La souche SS est néanmoins permissive pour l'activité de gypsy. Cela a pu être démontré en introduisant dans son génome un élément gypsy fonctionnel, grâce aux techniques utilisées couramment pour produire des lignées transgéniques de drosophile. Gypsy se multiplie dans les lignées ainsi obtenues, conduisant à des souches possédant de nombreuses copies de l'élément, dites MSN. Gypsy est transposé à fré-

quence très élevée dans ces lignées et produit des mutations [5]. Lorsque des femelles MSN sont croisées avec des mâles portant la mutation *ovo^{D1}*, elles donnent des filles dont une certaine proportion sont fertiles. Les souches SS et MSN sont donc isogéniques et diffèrent uniquement par la présence dans ces dernières d'éléments gypsy fonctionnels transposant à fréquence élevée.

Un rétrovirus chez les invertébrés

Les trois conditions permettant de démontrer des propriétés infectieuses de gypsy sont donc remplies. Il existe en effet des souches, comme les souches MSN, dans lesquelles il est très actif, et dans lesquelles des particules infectieuses sont susceptibles d'être produites. Il existe également des souches, comme SS, dépourvues de copies actives de l'élément, mais permissives pour son activité. Enfin, le test de la réversion de la mutation *ovo^{D1}* permet de détecter aisément des mouches dans lesquelles gypsy est présent et actif. Si gypsy est infectieux, les particules virales éventuellement produites dans les souches MSN devraient pouvoir infecter des individus de la souche SS.

Les propriétés infectieuses de gypsy ont été étudiées dans deux séries d'expériences [6]. Dans la première, du cytoplasme prélevé dans des embryons d'une souche MSN a été micro-injecté dans des embryons de la souche SS. Dans la seconde, des individus de la souche SS ont été élevés en présence de broyats d'individus de la souche MSN. Les femelles provenant de ces deux séries d'expériences ont été croisées avec des mâles *ovo^{D1}*, et la fertilité de leurs filles a été étudiée. Dans les deux cas, des filles fertiles ont été obtenues. L'analyse moléculaire de leur descendance a montré que, dans tous les cas, la restauration de la fertilité était bien due à l'insertion d'un élément gypsy dans le gène *ovo*. Ces résultats montrent que gypsy a un pouvoir contaminant et peut infecter des mouches. C'est donc le premier rétrovirus identifié chez un invertébré.

Ces résultats ont en outre apporté d'autres informations intéressantes sur ce rétrovirus. En effet, dans les expé-

riences résumées ci-dessus, seuls les événements résultant de l'infection de la lignée germinale pouvaient être détectés. Le virus a donc un très net tropisme pour des cellules germinales. La seconde observation est que tous les «révertants» de la mutation *ovo^{D1}* correspondent à des insertions dans une même petite région du gène, ce qui révèle l'existence d'un point chaud d'intégration du provirus.

Les propriétés infectieuses de gypsy ont été confirmées par la suite grâce à une approche expérimentale similaire, mais en utilisant des particules purifiées sur gradient de sucrose à partir de la souche MG [7].

Tous ces résultats montrent que les rétrovirus se propagent au-delà des vertébrés, et que les séquences ressemblant à gypsy et présentes dans le génome des insectes, sont vraisemblablement aussi des rétrovirus. En fait, il ne fait guère de doute que si ces éléments avaient été identifiés chez les vertébrés, ils auraient été d'emblée qualifiés de rétrovirus endogènes.

Flamenco, un gène de l'hôte qui contrôle gypsy

Les éléments transposables se propagent dans les génomes grâce à leur pouvoir invasif. Toutefois, la transposition se manifeste par des mutations et diverses autres anomalies, et a le plus souvent des conséquences néfastes pour l'individu. Les éléments transposables ne se maintiennent donc dans leur hôte que parce qu'ils sont soumis à des régulations très strictes qui permettent leur transposition seulement à un très faible niveau. La même remarque s'applique aux rétrovirus endogènes. L'une des questions les plus intéressantes que l'on peut se poser au sujet de gypsy est donc celle de son contrôle. On peut en effet se demander pourquoi ce rétrovirus est habituellement stable, et pourquoi il est dérégulé et transpose à fréquence élevée dans certaines souches. L'existence de ces dernières offre une excellente opportunité d'étudier les modalités de sa régulation.

L'étude génétique détaillée des phénomènes d'instabilité observés dans la souche MG a permis de montrer que gypsy est contrôlé par un gène de la drosophile, qui a été appelé *flamen-*

RÉFÉRENCES

8. Prud'homme N, Gans M, Masson M, Terzian C, Bucheton A. *Flamenco*, a gene controlling the *gypsy* retrovirus of *Drosophila melanogaster*. *Genetics* 1995 ; 139 : 697-711.
9. Pélisson A, Song SU, Prud'homme N, Smith PA, Bucheton A, Corces VG. Gypsy transposition correlates with the production of a retroviral-like protein under the tissue-specific control of the *Drosophila flamenco* gene. *EMBO J* 1994 ; 13 : 4401-11.
10. Arkhipova IR, Mazo AM, Cherkasova VA, Gorelova TV, Schuppe NG, Ilyin YV. The steps of reverse transcription of *Drosophila* mobile dispersed genetic elements and U3-R-U5 structure of their LTRs. *Cell* 1986 ; 44 : 555-63.
11. Bucheton A, Vaury C, Chaboissier MC, Abad P, Pélisson A, Simonelig M. I elements and the *Drosophila* genome. *Genetica* 1992 ; 86 : 175-90.
12. Busseau I, Chaboissier MC, Pélisson A, Bucheton A. I factors in *Drosophila melanogaster* : transposition under control. *Genetica* 1994 ; 93 : 101-16.
13. Jolicœur P. The Fv-1 gene of the mouse and its control of murine leukemia virus replication. *Curr Top Microbiol Immunol* 1979 ; 86 : 67-122.
14. Levy DE, Lerner RA, Wilson MC. The Gv-1 locus coordinately regulates the expression of multiple endogenous murine retroviruses. *Cell* 1985 ; 41 : 289-99.

co [8]. Ce gène est localisé sur le chromosome X, près de l'hétérochromatine péri-centromérique. L'allèle apparemment le plus répandu dans les souches de drosophile, *flam*⁺ (qualifié également de non permissif), réprime l'activité de gypsy. L'allèle présent dans la souche MG, *flam*¹, est permissif pour l'activité de gypsy. Le gène *flamenco* a un effet maternel sur la transposition (figure 2). En effet, celle-ci ne se produit que dans la lignée germinale de descendants de femelles homozygotes pour la mutation *flamenco*, quel que soit le génotype de ces descendants. Cette fréquence élevée de transposition se traduit par des taux de mutations élevés. Les mutations se produisent aussi bien dans la lignée germinale mâle que dans la lignée germinale femelle. Elles sont dues à des insertions de l'élément gypsy dans des gènes, et elles apparaissent souvent sous forme de clones, ce qui indique que la transposition est préméiotique. Elle ne se produit pas dans la descendance des femelles *flam*⁺/*flam*¹, ce qui montre qu'une dose de l'allèle *flam*⁺ est suffisante pour la réprimer (figure 2).

Le gène *flamenco* contrôle les transcrits de gypsy

Des expériences ont été réalisées dans le but de comprendre comment le gène *flamenco* contrôle l'activité de gypsy. Les transcrits du rétrovirus ont été comparés dans des femelles *flam*⁺/*flam*¹ (non permissives pour la transposition) et *flam*¹/*flam*¹ (permissives) [9]. Ces transcrits sont plus abondants chez les femelles permissives que chez les femelles non permissives. Deux transcrits dont l'abondance est augmentée chez les femelles permissives sont particulièrement intéressants. Leur structure est représentée sur la figure 3. L'un d'eux a une taille de 7 kilobases (kb) et avait déjà été caractérisé dans des cellules de drosophile en culture [10]. Il s'agit de l'ARN génomique du virus, dont la transcription débute dans le LTR 5' et se termine dans le LTR 3', qui possède la séquence complète de gypsy. La transcriptase inverse l'utilise comme matrice pour produire l'ADNc qui s'intégrera dans le génome de l'hôte. Cet ARN est également utilisé pour la synthèse des produits des gènes *gag* et *pol*, les seconds étant synthétisés grâce

à un déphasage traductionnel dans la région de chevauchement des deux gènes.

L'autre transcrit a une taille de 2,1 kb et n'avait jamais été décrit. Il s'agit d'un ARN épissé, dont la transcription débute et se termine de la même manière que celle de l'ARN génomique. L'épissage élimine les séquences correspondant aux gènes *gag* et *pol*, et crée un AUG qui se trouve dans un contexte favorable pour débiter la traduction de l'ORF3, dont le produit présente des similitudes avec les produits des gènes *env* rétroviraux [9]. Ce transcrit est tout à fait semblable à l'ARN sous-génomique utilisé par les rétrovirus pour produire les protéines de l'enveloppe.

Des produits de traduction de l'ORF3 de gypsy ont été détectés grâce à des anticorps monoclonaux sous forme de glycoprotéines dans des extraits de femelles *flam*¹/*flam*¹. Ces glycoprotéines sont associées à des particules ayant une activité transcriptase inverse [7].

La régulation de gypsy par le gène *flamenco* est spécifique des cellules folliculaires de l'ovaire

L'une des propriétés troublantes de la régulation de gypsy par le gène *flamenco* est son caractère maternel (voir ci-dessus et figure 2). Des expériences d'hybridation *in situ* ont été réalisées pour comparer la distribution des transcrits de gypsy dans les ovaires de femelles *flam*¹/*flam*¹ (donnant des descendants chez lesquels la transposition se produit) et *flam*⁺/*flam*⁺ (donnant des descendants chez lesquels la transposition ne se produit pas) [9]. Ces transcrits s'accumulent au stade 10 de l'ovogenèse dans les ovaires des femelles homozygotes pour la mutation *flam*¹. Curieusement, cette accumulation ne se produit pas dans les cellules de la lignée germinale, mais dans les cellules folliculaires qui entourent l'ovocyte et sont d'origine somatique (figure 4). Des expériences d'immunodétection avec des anticorps monoclonaux dirigés contre les protéines de l'enveloppe de gypsy indiquent également que ces produits sont localisés dans les mêmes cellules [9]. Les ARN et les protéines Env ne sont d'ailleurs pas distribués de manière uniforme dans les cel-

lules folliculaires, mais concentrés dans le cytoplasme de la partie apicale des cellules, proche de l'ovocyte.

Il peut sembler contradictoire que l'accumulation des transcrits soit observée dans des cellules soma-

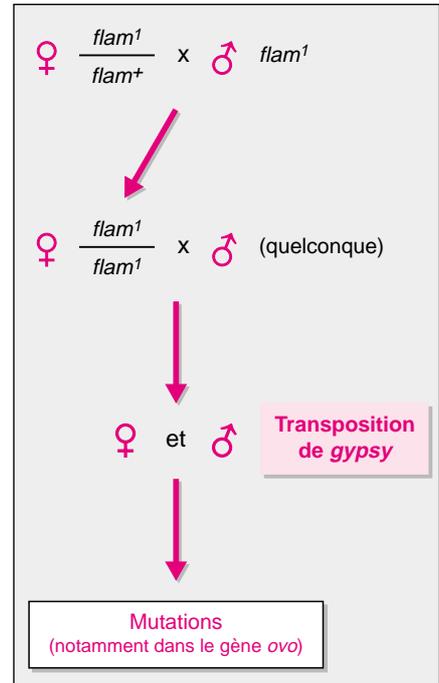


Figure 2. **Contrôle du rétrovirus gypsy par le gène *flamenco*.** Le gène *flamenco* (*flam*) est localisé sur le chromosome X de la drosophile. La plupart des souches possèdent des allèles non permissifs (*flam*⁺), qui répriment l'élément gypsy. Certaines portent des allèles permissifs, comme *flam*¹ qui permettent la transposition et l'expression des propriétés infectieuses de gypsy. Le gène *flamenco* a un effet maternel sur la transposition. En effet, celle-ci n'est observée que chez les descendants (mâles ou femelles) de femelles homozygotes pour des allèles mutants du gène *flam*, quel que soit leur génotype. La fréquence élevée de transposition se produisant dans la lignée germinale de ces individus se traduit par la production de mutations qui peuvent être observées à la génération suivante. Le gène *ovo* est un point très chaud d'intégration du provirus, et des mutations de ce gène sont très fréquemment observées. Une dose de l'allèle *flam*⁺ est suffisante pour réprimer la transposition. En effet, celle-ci ne se produit pas chez les descendants de femelles hétérozygotes *flam*¹/*flam*⁺.

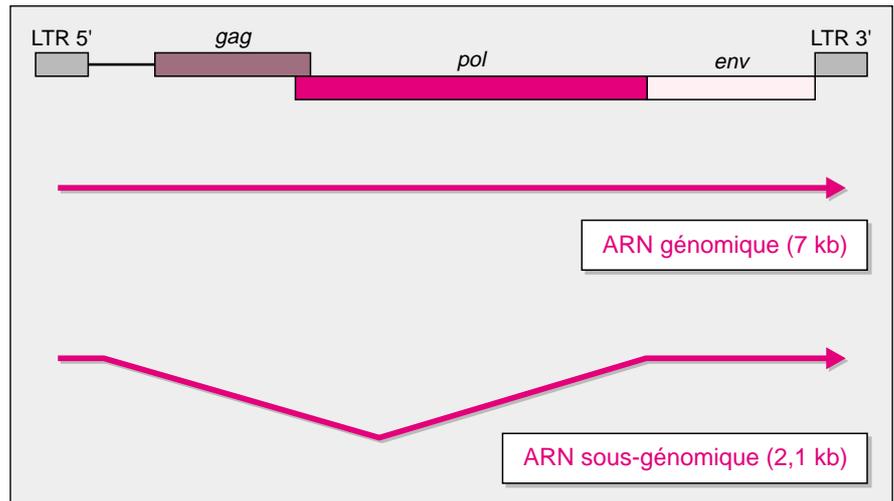


Figure 3. **Structure de l'ARN génomique et de l'ARN sous-génomique de gypsy.** La figure représente deux transcrits surexprimés chez les femelles *flam¹/flam¹* par rapport aux femelles *flam¹/flam⁺* ou *flam⁺/flam⁺*. L'un d'entre eux a une taille de 7 kb. C'est le transcrite génomique, dont la transcription commence dans le LTR 5' et se termine dans le LTR 3'. Il contient toute la séquence de gypsy, et est utilisé comme matrice pour la transcription inverse, ainsi que pour la synthèse des produits des gènes gag et pol. L'autre a une taille de 2,1 kb. C'est un ARN épissé, dont la transcription commence et se termine aux mêmes sites que l'ARN génomique. Il est utilisé pour la synthèse des produits du gène env. Cet ARN sous-génomique n'est détecté que chez les femelles portant à l'état homozygote des allèles mutants (permissifs) du gène flamenco.

RÉFÉRENCES

15. Gardner MB, Kozak CA, O'Brien SJ. The lake casitas wild mouse evolving genetic resistance to retroviral disease. *Trends Genet* 1991; 7: 22-7.
16. Gorska-Flipot I, Huang M, Cantin M, Rassart E, Massé G, Jolicœur P. U3 Long terminal repeat-mediated induction of intracellular immunity by a murine retrovirus: a novel model of latency for retroviruses. *J Virol* 1992; 66: 7201-10.
17. Kitagawa M, Aizawa S, Kamisaku H, Ikeda H, Hirokawa K, Sado T. Cell free transmission of Fv-4 resistance gene product controlling Friend leukemia virus-induced leukemogenesis: a unique mechanism for interference with viral infection. *Blood* 1995; 86: 1557-63.
18. Bucheton A. The relationship between the *flamenco* gene and *gypsy* in *Drosophila*: how to tame a retrovirus. *Trends Genet* 1994; 11: 349-53.
19. Doolittle RF, Feng DF, Johnson MS, McClure MA. Origins and evolutionary relationships of retroviruses. *Quart Rev Biol* 1989; 64: 1-30.
20. Xiong Y, Eickbush TH. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J* 1990; 9: 3353-62.
21. Britten RJ, McCormack TJ, Mears TL, Davidon EH. Gypsy/Ty3-class retrotransposons integrated in the DNA of herring, tunicate, and echinoderms. *J Mol Evol* 1995; 40: 13-24.

tiques de l'ovaire, alors que la transposition se produit à la génération suivante. L'hypothèse avancée pour expliquer cet apparent paradoxe est que, dans les cellules folliculaires des femelles homozygotes pour la mutation *flamenco*, des particules virales enveloppées, et donc infectieuses, seraient produites. Ces particules infecteraient l'ovocyte et seraient ainsi transmises à la descendance. La transposition observée à la génération suivante correspondrait à l'intégration de provirus dans les chromosomes des individus issus des ovocytes infectés [9]. La recherche de particules virales dans les ovaires des femelles *flamenco* devrait permettre d'apporter des précisions sur l'effet maternel de ce gène. La caractérisation moléculaire du gène *flamenco* devrait aussi apporter des informations importantes sur la façon dont il agit sur le rétrovirus, ainsi que sur sa fonction première chez la mouche.

L'acquisition d'une résistance de la drosophile à une invasion rétrovirale

Nous avons vu que toutes les souches de *Drosophila melanogaster* possèdent des éléments gypsy défectifs localisés dans l'hétérochromatine péricentromérique [2]. Ces éléments sont au nombre de 20 ou 30, et sont incapables de transposer. L'analyse de leur localisation génomique montre qu'ils occupent une position fixe dans des souches non apparentées. D'une manière générale, les copies défectueuses d'éléments transposables tendent à s'accumuler dans cette même région du génome. Par exemple, l'étude détaillée d'un élément transposable de la drosophile semblable aux LINE (*long interspersed nucleotidic element*) des mammifères, le facteur I, a montré que des copies défectueuses de cet élément se sont accumulées dans l'hétérochromati-

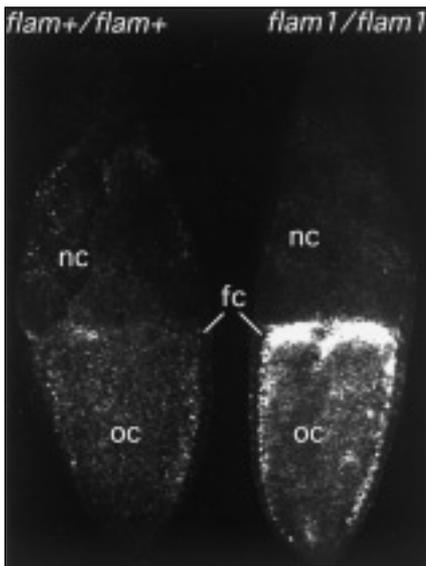


Figure 4. **Localisation des transcrits de gypsy dans les cellules folliculaires.** La figure représente les résultats d'expériences d'hybridation in situ sur les ARN dans des ovaires de femelles permissives ($flam^1/flam^1$) et non permissives ($flam^+/flam^+$) à l'aide d'une sonde de l'élément gypsy. Pratiquement aucune hybridation n'est détectée chez les femelles non permissives. Un très fort signal est observé dans les ovaires des femelles permissives. Celui-ci n'est pas détecté dans la lignée germinale, mais seulement dans le cytoplasme des cellules folliculaires qui entourent l'ovocyte, et sont d'origine somatique. oc, ovocyte; nc, cellules nourricières; fc, cellules folliculaires.

ne, du fait de l'inactivation des éléments fonctionnels très tôt au cours de l'histoire évolutive de l'espèce *Drosophila melanogaster* [11, 12]. Il est très vraisemblable qu'il en soit de même pour gypsy, et que les éléments défectueux localisés à des sites semblables dans l'hétérochromatine soient les vestiges d'une invasion ancienne par le rétrovirus gypsy. Il est ainsi probable qu'à cette période, l'espèce *Drosophila melanogaster*, ou son ancêtre, était permissive pour l'activité de gypsy, et possédait des allèles permissifs comme *flam*¹. Nous savons que les souches de drosophile possédant à la fois des allèles permissifs de *flamenco* et des éléments gypsy fonctionnels sont caractérisées

par des fréquences élevées de mutations, et présentent des anomalies telles qu'une fertilité ou une viabilité réduites. Il est donc tout à fait concevable que la propagation de gypsy dans les populations ait été accompagnée de l'apparition d'anomalies diverses. L'espèce n'aurait survécu que grâce à la sélection de formes alléliques de certains gènes ayant la propriété de stopper la propagation du rétrovirus. On peut imaginer que les allèles non permissifs de *flamenco* ont été sélectionnés de cette façon. Cela expliquerait pourquoi gypsy est maintenant stable dans la plupart des souches. La mutation des allèles non permissifs vers des allèles permissifs restaurerait ainsi la situation qui prévalait dans les populations avant l'invasion du virus. La sélection d'allèles non permissifs dans une souche instable qui possédait au départ l'allèle *flam*¹ et des éléments gypsy fonctionnels a déjà été observée. Ce fait vient à l'appui de l'hypothèse selon laquelle les différentes formes alléliques de *flamenco* pourraient résulter de la pression de sélection exercée par le rétrovirus.

Le gène *flamenco* et le rétrovirus gypsy représentent ainsi une situation exceptionnelle dans le domaine de la rétrovirologie pour étudier les relations entre un rétrovirus et son hôte. Celle-ci est particulièrement intéressante pour étudier les conséquences de la présence d'un rétrovirus dans le génome de son hôte, et pour déterminer les mécanismes qui peuvent être sélectionnés dans une espèce pour répondre à des invasions rétrovirales. Il est presque impossible d'aborder expérimentalement ce genre de questions chez les vertébrés. Seuls quelques cas de régulation de rétrovirus de mammifères par des séquences de l'hôte, qui restent d'ailleurs pour l'essentiel très mal compris, ont été rapportés [13-17]. La drosophile est sans aucun doute le modèle le plus approprié au développement d'une étude génétique détaillée des relations rétrovirus-hôte.

■ L'ancêtre des rétrovirus ?

La différence fondamentale entre rétrovirus et rétrotransposons réside dans leur mode de propagation [18]. Les rétrovirus typiques, comme le VIH, ont un mode de transmission

strictement horizontal, car ils n'infectent que des cellules somatiques. Les rétrotransposons, comme les éléments copia de la drosophile ou les éléments Ty1 de la levure, sont transmis uniquement verticalement car ils ne sont pas infectieux. Ces éléments ne se propagent donc que parce qu'ils se multiplient dans les cellules de la lignée germinale. Les rétrovirus endogènes comme gypsy possèdent les deux propriétés. Étant installés dans la lignée germinale, ils sont transmis verticalement comme les rétrotransposons, mais la présence d'un gène *env* permet la production de particules infectieuses assurant une dissémination horizontale. Gypsy possède une particularité supplémentaire qui est, comme nous l'avons vu plus haut, son très fort tropisme pour la lignée germinale.

Des arbres phylogénétiques des rétroéléments ont été établis par comparaison de certains domaines de la transcriptase inverse [19, 20]. La figure 5 donne une représentation schématique de ces arbres. Les rétrovirus des vertébrés forment une classe particulière de rétroéléments, et les rétrotransposons appartenant au groupe nommé groupe Ty1/copia forment une autre classe bien distincte de la précédente. Ce groupe Ty1/copia comprend des rétrotransposons d'animaux (principalement des invertébrés), de plantes et de champignons.

Comme nous l'avons vu précédemment, Temin a avancé l'hypothèse, communément admise, selon laquelle les rétrovirus ont évolué à partir des rétrotransposons à LTR, par acquisition du gène *env* [1]. Cette hypothèse se heurte à une difficulté majeure, qui est la séparation complète du groupe des rétrovirus des vertébrés de celui des rétrotransposons de la classe Ty1/copia. En effet, aucun rétrovirus ne possède un gène *pol* du type Ty1/copia. Inversement, aucune transcriptase inverse semblable à celle des rétrovirus de vertébrés n'a été décrite dans la classe Ty1/copia. Gypsy appartient à un autre groupe de rétroéléments connu sous le nom de groupe Ty3/gypsy. Ce groupe, jusqu'à une date récente, était considéré comme homogène, puisqu'il n'était censé contenir que des éléments transposables, les propriétés infectieuses de

gypsy étant inconnues. On sait maintenant qu'il est en fait hétérogène et comprend d'une part de vrais rétrotransposons, comme l'élément Ty3 de la levure, qui ne possèdent que deux ORF correspondant aux gènes *gag* et *pol*, et d'autre part tous les éléments identifiés jusqu'à présent chez les insectes qui, comme gypsy, possèdent trois ORF et sont vraisemblablement des rétrovirus endogènes. Ainsi, les éléments 17.6, 297, tom et TED appartiennent à cette classe. La transcriptase inverse des éléments de cette catégorie est plus proche de celle des rétrovirus de vertébrés que de celle des rétrotransposons du groupe Ty1/copia. Originellement, on pensait que les éléments du groupe Ty3/gypsy étaient restreints aux champignons et aux invertébrés, mais des résultats récents ont permis de mettre en évidence des éléments de cette classe chez les poissons [21].

Il est donc tout à fait clair maintenant que le groupe Ty3/gypsy comprend à la fois des rétrotransposons et des rétrovirus, appartenant à des espèces très diverses, allant des champignons aux invertébrés et vertébrés. Cela laisse penser que les rétrovirus se sont formés à partir de rétrotransposons du type Ty3, par acquisition d'un gène d'enveloppe, comme cela est illustré sur la figure 5, et que cet événement a pu se produire très tôt au cours de l'évolution, avant la divergence des vertébrés et des invertébrés. En ce sens, gypsy représenterait une classe d'éléments correspondant à une forme ancestrale de rétrovirus, dont les rétrovirus des vertébrés seraient issus.

Des perspectives dans le domaine médical

L'existence de rétrovirus chez les insectes, tout particulièrement chez la drosophile, ouvre de nombreuses perspectives. Les travaux réalisés sur cet organisme modèle ont généralement eu des retombées importantes pour l'étude des autres organismes. Il suffit pour s'en convaincre de songer par exemple aux gènes homéotiques, initialement découverts chez la drosophile, dont l'étude a permis d'identifier les gènes *Hox* chez les mammifères (notamment chez l'homme) et de mettre en évidence la remarquable conservation de ces

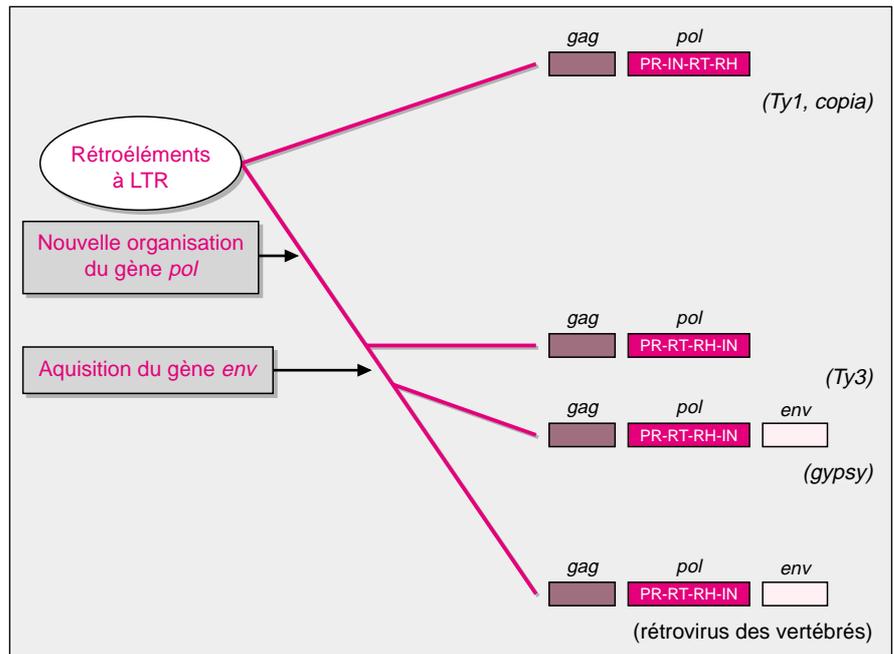


Figure 5. **Relations phylogénétiques entre rétrovirus et rétrotransposons à LTR.** Cet arbre schématique est fondé sur la comparaison des séquences des domaines transcriptase inverse du gène *pol* [19, 20]. Les événements dont on peut penser qu'ils ont conduit aux rétrovirus des vertébrés sont indiqués sur la figure. Il s'agit tout d'abord d'une réorganisation du gène *pol* d'éléments transposables semblables à Ty1 ou copia, caractérisée par une nouvelle localisation terminale du domaine intégrase, aboutissant aux éléments transposables du groupe Ty3. Le second événement serait l'acquisition du gène *env* par un élément du groupe Ty3, et serait à l'origine des rétrovirus endogènes du type gypsy. Dans cette hypothèse, gypsy appartiendrait à un groupe représentant la forme ancestrale des rétrovirus, à partir duquel les rétrovirus des vertébrés auraient évolué. Les différents domaines du gène *pol* sont représentés : PR, protéase; RT, transcriptase inverse; RH, RNase H; IN, intégrase.

gènes et de leur rôle au cours de l'évolution. La drosophile peut se révéler aussi un modèle de choix pour l'analyse des relations, sur le plan génétique, entre les rétrovirus et leurs hôtes. L'intérêt de comprendre comment cette espèce a probablement survécu à une invasion rétrovirale est évident. De ce point de vue, l'étude du gène *flamenco* est prometteuse.

La présence de rétrovirus chez les invertébrés est également intéressante sur un plan plus pratique. La performance du matériel biologique qu'est la drosophile repose très largement sur l'utilisation d'un élément transposable (l'élément P), fondée sur des connaissances précises du mécanisme de sa transposition et de son contrôle. Les outils développés à partir de cet élément permettent en

particulier de réaliser très facilement des mutagenèses par insertion en vue de cloner des gènes, et des expériences de transgénèse. La drosophile est cependant le seul insecte chez lequel la génétique moléculaire est aussi performante. Cela est dû, notamment, au fait que les outils dérivés de cet élément transposable ne fonctionnent pas chez des espèces plus éloignées. Ainsi, la constitution de lignées transgéniques, indispensable à un grand nombre d'études, n'est pour l'instant pas possible chez les autres insectes, y compris ceux d'intérêt médical ou agronomique. De nombreux domaines de recherche pourraient progresser plus rapidement si l'étude des autres insectes n'était pas ainsi limitée. Citons le cas du paludisme dans lequel les difficultés à mettre au point un vaccin ou à

proposer des traitements efficaces contre la maladie amènent nombre de chercheurs à penser qu'il serait peut-être judicieux de lutter contre le parasite responsable de la maladie non pas chez l'homme, mais chez le moustique. Il est possible d'envisager de construire, à partir de gypsy, des vecteurs rétroviraux du même type que ceux développés chez les vertébrés, qui pourraient être utilisés pour la transgénèse chez les insectes. Ainsi, la découverte d'un rétrovirus chez les invertébrés pourrait bien ne pas avoir que des retombées sur le plan fondamental, mais aussi sur le plan agronomique et médical ■

Remerciements

Je remercie mes collègues du groupe « rétrotransposition » du Centre de Génétique Moléculaire du Cnrs pour leurs innombrables et fructueuses discussions, et Alain Pélisson, Nicole Prud'homme et Christophe Terzian pour leurs commentaires sur ce manuscrit. Notre travail sur le rétrovirus gypsy et le gène *flamenco* est financé notamment par des subventions du Cnrs, de l'ARC, et du Groupement de Recherches et d'Études sur les Génomes (GREG).

TIRÉS À PART

A. Bucheton.

Summary

The flamenco gene of *Drosophila*: how to resist a retrovirus

For a long time, retroviruses, that are responsible for serious diseases, have been thought to be restricted to vertebrates. They are actually more widely dispersed. Recent results indicate that the gypsy element of *Drosophila melanogaster*, usually considered as a transposable element of the class of retrotransposons, has infective properties and is therefore the first retrovirus identified in invertebrates. The genome of insects contain other retroelements which, like gypsy, are strikingly similar to vertebrates proviruses of retroviruses. It is likely that they are also endogenous retroviruses. Gypsy is controlled by a *Drosophila* gene called *flamenco*. Many strains contain non permissive alleles of this gene that maintain the retrovirus in a repressed state. When the *flamenco* gene is mutated gypsy transposes at high frequency and produces infectious particles. Many results show that gypsy invaded the *Drosophila melanogaster* species, or an ancestor of this species, a long time ago, indicating that the species at that time was permissive for propagation of the retrovirus and presumably contained permissive alleles of the *flamenco* gene. The simultaneous presence in flies of permissive alleles of the gene and of func-

tional gypsy elements results in the occurrence of many abnormalities, and one can imagine that the species survived the retroviral invasion because non permissive alleles of *flamenco* were selected. The characterization of a retrovirus in *Drosophila*, one of the most advanced model organisms for molecular genetics, provides us with an exceptional clue to study how a species can resist a retroviral invasion. Similar investigations are presently not possible in vertebrates. Analysis of the phylogenetic relationships between retroviruses and retrotransposons indicate that gypsy might be a member of an ancestral class of retroviruses from which vertebrate retroviruses evolved. Therefore the identification of a retrovirus in *Drosophila* provides a very powerful experimental system to investigate both the genetic relationships between retroviruses and their hosts and the evolutionary relationships between retroviruses and retrotransposons. In addition, this endogenous retrovirus might be used to construct retroviral vectors. Therefore the characterization of a retrovirus in *Drosophila* might be extremely useful to develop tools which could be used also to study insects of medical and agricultural interest.

DIPLÔME D'UNIVERSITÉ DE CHRONOBIOLOGIE, Année Universitaire 1996-1997

- Un enseignement de **Chronobiologie** est organisé à la Faculté de Médecine Pitié-Salpêtrière, sous la direction du Professeur Yvan Touitou. Il a pour but de donner une formation théorique et pratique aux étudiants pour leur permettre l'utilisation des méthodes chronobiologiques. Le diplôme est ouvert aux médecins, pharmaciens, chirurgiens-dentistes, internes des hôpitaux, maîtres ès sciences et, sur proposition du directeur d'enseignement, aux candidats intéressés par la **Chronobiologie** ayant tous autres titres et travaux. L'enseignement se déroule sous la forme de 5 séminaires de 2 jours chacun, en novembre, décembre, janvier, février et mars. Il est dispensé à la Faculté de Médecine Pitié-Salpêtrière à Paris. Les étudiants salariés peuvent s'inscrire dans le cadre de la formation permanente (prise en charge de l'inscription par l'employeur). La date limite des inscriptions est fixée au 12 novembre 1996.
- L'enseignement porte sur les aspects fondamentaux et appliqués des rythmes biologiques, de la cellule à l'homme. Il est sanctionné par un examen écrit et oral permettant l'obtention du Diplôme d'Université.

Le programme des cours est le suivant :

Mardi 12 et mercredi 13 novembre 1996 : Propriétés fondamentales et méthodes d'étude des rythmes biologiques.

Lundi 9 et mardi 10 décembre 1996 : Rythmes à l'échelon cellulaire et moléculaire : mécanismes. Rythmes endocriniens et neuroendocriniens.

Lundi 13 et mardi 14 janvier 1997 : Rythmes en pharmacologie et toxicologie.

Lundi 10 et mardi 11 février 1997 : Pathologie et chronothérapeutique en endocrinologie, en cancérologie, en psychiatrie, etc.

Lundi 10 et mardi 11 mars 1997 : Développement, vieillissement et adaptation. Photopériodisme et régulation des rythmes biologiques.

Les candidats intéressés doivent faire une demande écrite précisant leur formation universitaire au Professeur Yvan Touitou, DU de Chronobiologie, Faculté de Médecine Pitié-Salpêtrière, 91, boulevard de l'Hôpital, 75013 Paris, France. Tél. (33-1) 40.77.96.63 - Fax (33-1) 40.77.96.65.