



Relation entre l'expression des LAT et du gène RL2 pendant la latence du virus HSV-1

Nicolas Huot

► To cite this version:

Nicolas Huot. Relation entre l'expression des LAT et du gène RL2 pendant la latence du virus HSV-1. Sciences agricoles. Université Paris Sud - Paris XI, 2012. Français. NNT : 2012PA114860 . tel-00923161

HAL Id: tel-00923161

<https://theses.hal.science/tel-00923161>

Submitted on 2 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ECOLE DOCTORALE :

INNOVATION THÉRAPEUTIQUE : DU FONDAMENTAL A L'APPLIQUÉ

PÔLE : MICROBIOLOGIE / THERAPEUTIQUES ANTIINFECTIEUSES

DISCIPLINE : MICROBIOLOGIE

ANNÉE 2012 - 2013

SÉRIE DOCTORAT N° 1215

THÈSE DE DOCTORAT

Soutenue le 17/12/2012

par

Nicolas HUOT

Découverte d'un nouvel élément mobile dans le virus de
l'herpes simplex de type 1.

Directeur de thèse :
Co-directeur de thèse :

Mr Marc LABETOULLE
Mr Yves GAUDIN

Professeur (Paris XI, Kremlin-Bicêtre)
Directeur de Recherche (VMS, GIF sur Yvette)

Composition du jury :

Président du jury :

Mr Vincent MARECHAL

Professeur (UMPC, Paris)

Rapporteurs :

Mme Audrey ESCLATINE

Maître de conférences (ParisXI, Chatenay-Malabry)

Mr Nicolas GILBERT

Chargé de Recherche, (IGH, Montpellier)

Examinatrice :

Mme Joëlle MARIE

Directrice de Recherche, (CGM, GIF sur Yvette)

Au terme de ce travail, c'est avec émotion que je tiens à remercier tous ceux qui, de près ou de loin, ont contribué à la réalisation de ce projet.

Je tiens à remercier mes directeurs de thèse pour l'expérience inoubliable de ces dernières années.

Je tiens tout d'abord à adresser mes remerciements les plus sincères au Pr Marc LABETOULLE pour avoir co-dirigé cette thèse et m'avoir permis de la réaliser dans les meilleures conditions. Je tiens particulièrement à le remercier de la liberté d'action qu'il m'a donné à chaque étape de cette aventure. Merci Marc de t'être battu pour que je puisse faire cette thèse.

Merci à Yves GAUDIN directeur du laboratoire pour avoir co-dirigé cette thèse, pour son soutien sans failles et son accueil. Merci Yves d'avoir été là, pour ton écoute, tes conseils, et merci d'avoir partagé mes petites folies intellectuelles. Merci aussi pour tout ce que tu fais pour le laboratoire.

Merci à vous de m'avoir supporté et de m'avoir laissé vous contredire à quelques reprises...

J'espère avoir été digne de la confiance que vous m'avez accordé et que ce travail est finalement à la hauteur de vos espérances.

Quoi qu'il en soit, j'ai beaucoup appris à leurs côtés et je suis très honoré de les avoir eu pour encadrant.

Je tiens également à remercier les membres de mon comité de thèse pour l'intérêt qu'ils ont porté à ce travail et pour les remarques constructives qu'ils ont fait sur mon projet de thèse. J'adresse mes remerciements aux membres du jury d'avoir accepté de juger ce travail.

Merci également au personnel du laboratoire de virologie moléculaire et structurale qui à un moment ou un autre durant la durée de ce doctorat se sont dévoués pour le bon déroulement de mes études.

A mes parents, pour leur amour et leur soutien.

Mes remerciements vont aussi à ma famille et mes amis qui, avec cette question récurrente, « quand est-ce que tu la soutiens cette thèse ? », bien qu'angoissante en période fréquente de doutes, m'ont permis de ne jamais dévier de mon objectif final.

A mes grands-parents, pour tout ce qu'ils ont fait pour moi.

A mes frères et à mes sœurs, pour tout ce bonheur partagé lors de notre enfance.

A toute ma famille.

Merci à Frédérique et Benoit pour leurs encouragements.

A tous mes amis du passé et du présent, pour tous les merveilleux moments passés en votre compagnie et pour votre soutien le plus souvent sans faille.

Ces remerciements ne seraient pas complets sans une pensée pour mes amis de longue date, Arnaud, Sami, Yan, Alex, et Rachid. Merci de m'avoir aidé et encouragé, et pour m'avoir changé les idées quand j'en avais besoin.

Enfin, je remercie mon épouse Raphaëlle, pour son soutien quotidien indéfectible et son enthousiasme contagieux à l'égard de mes travaux comme de la vie en général. Notre couple a grandi en même temps que mon projet scientifique, le premier servant de socle solide à l'épanouissement du second.

Pour clore ce préambule, je souhaiterais préciser un dernier point. Dans une société où tout est, à l'image de la banque, rapidité, compétitivité, rentabilité, je revendique avoir fait, pour la première fois de ma vie, au travers de ce travail de thèse et dans l'esprit du philosophe et essayiste Pierre Sansot, « l'éloge de la lenteur » et son bon usage. En effet, il m'a fallu du temps pour dessiner les contours de mon sujet, comprendre ce que j'observais, avoir un certain recul, mettre en forme une thèse. ... prendre en fait le temps nécessaire à tout travail de recherche.

« On donne généralement le nom de découverte à la connaissance d'un fait nouveau ; mais je pense que c'est l'idée qui se rattache au fait découvert qui constitue en réalité la découverte. »
Claude Bernard

« Il y a plus de philosophie dans une bouteille de vin que dans tous les livres. »

Louis Pasteur

« La théorie de l'Évolution traite d'événements très anciens, détruit la position privilégiée de l'espèce humaine non seulement comme couronnement mais aussi comme finalité de la création, et foule aux pieds la moitié des religions de la planète. »

Ian Stewart

Résumé

Le virus de l'herpès simplex de type 1 (HSV1) établit une infection latente dans le système nerveux de l'homme, au cours de laquelle un type de transcrits, appelés LATs (pour *latency associated transcripts*), s'accumule dans les neurones infectés. Le rôle clef des LATs dans le contrôle de la latence virale est reconnu. Cependant, depuis leur découverte dans les années 80, leur mécanisme d'action reste non élucidé.

Le gène des LATs est transcrit en un LAT primaire de 8,3kb, qui est épissé, conduisant à la formation de deux LATs stables : le LAT_{2kb} et le LAT_{1.5kb}. De façon remarquable, le LAT_{2kb} et le LAT_{1.5kb} sont des introns. Leur stabilité est la conséquence d'un branchement non canonique qui se traduit par le maintien de la structure en lariat.

Par ailleurs, la région du génome codant les LATs contient également le gène *RL2* qui code ICP0, la protéine la plus en amont dans la cascade de réactivation du virus. Des études précédentes ont montré qu'au moment de la latence, des transcrits *RL2* non épissés, contenant encore l'intron 1, s'accumulaient au site principal de la latence (le ganglion de Gasser).

Nous avons caractérisé ces transcrits non épissés du gène *RL2* dans les tissus infectés de façon latente. Nous avons constaté avec surprise qu'ils contenaient des séquences du LAT_{2kb} insérées. Ces séquences insérées étaient aussi retrouvées dans le gène *RL2*. Ceci indiquait que les LATs étaient des éléments mobiles. En outre, une démarche visant à identifier l'ensemble des transcrits modifiés par insertion du LAT_{2kb} a révélé qu'un second gène viral, *US12*, codant pour la protéine ICP47, était également ciblé de manière similaire.

En plus de cette capacité du LAT_{2kb} de modifier des gènes viraux, nous avons profité de la fréquence élevée des éléments B1 dans le génome de la souris pour mettre en évidence par B1-PCR que le LAT_{2kb} pouvait également s'intégrer dans le génome de l'hôte.

Nous avons également vérifié si ces modifications pouvaient déjà être détectées lors de la phase aiguë de l'infection. Nous avons mis en évidence des insertions du LAT_{2kb} dans des transcrits viraux et cellulaires. Cependant, nous n'avons pas pu détecter ces insertions au niveau de l'ADN dans les gènes correspondants. Prises dans leur ensemble, nos données suggèrent qu'à ce stade de l'infection, l'insertion du LAT_{2kb} dans les transcrits est la conséquence d'un mécanisme d'épissage inverse. Certains des transcrits modifiés par cette insertion peuvent être ensuite soumis à une transcription inverse, ce qui conduit à l'intégration du LAT_{2kb} dans le gène correspondant. Ces génomes modifiés permettent ensuite la synthèse de transcrits modifiés.

Par conséquent, les LATs sont un nouveau type d'éléments mobiles dotés de propriétés uniques qui sont à la fois capables de modifier les génomes viraux et cellulaires. Ces propriétés des LATs ouvrent de nouveaux horizons sur la biologie du virus, en particulier sur la régulation de la latence.

Enfin, l'identification de cet élément mobile viral, qui partage des similitudes fonctionnelles et structurales avec des éléments mobiles procaryotes et eucaryotes, soulève des questions importantes dans le domaine de l'évolution.

Summary

The herpes simplex virus type 1 (HSV1) establishes a lifelong latent infection in the human nervous system, during which a subtype of transcripts, called *Latency Associated Transcripts* (LATs), accumulate in infected neurons. The key role of LATs in the control of HSV1 latency is well acknowledged. However, their mechanism of action remains unknown since their discovery in the 80's.

The *LAT* gene encodes an 8.3 kb primary transcript, which is spliced, leading to the formation of two stable LATs, LAT_{2kb} and LAT_{1.5kb} (2 kb and 1.5 kb in size, respectively). Remarkably, LAT_{2kb} and LAT_{1.5kb} are stable introns, as a consequence of a non-consensus branch-point that results in the maintenance of a lariat structure.

Interestingly, the region of the HSV1 genome that encodes LATs also contains the *RL2* gene, encoding ICP0, the most upstream factor in the reactivation cascade of HSV1. Previous RT-PCR assays showed that a sub-species of *RL2* transcripts, still containing intron 1, accumulate in the main site of HSV1 latency (i.e. the trigeminal ganglion).

In the course of identifying the nature of these unspliced *RL2* transcripts in latently infected tissues, we surprisingly found LAT_{2kb} sequences inserted in these transcripts and in the corresponding *RL2* gene. This revealed that LATs act as mobile elements. Furthermore, an approach aimed at identifying all the transcripts modified by LAT_{2kb} insertion revealed that a second viral gene, *US12* encoding the protein ICP47, was also targeted in a similar manner.

Besides this ability of LAT_{2kb} to modify viral genes, taking advantage of the high frequency of B1 elements in the mouse genome, we have performed B1-PCR assays and have demonstrated that LAT_{2kb} can also integrate the host genome.

We also checked whether these modifications could already be detected during the acute phase of infection. We detected LAT_{2kb} insertions in viral and host transcripts. However, we were unable to detect these insertions at the DNA level in the corresponding genes. Taken together, our data suggest that at this stage, LAT_{2kb} insertion into transcripts is a consequence of a reverse splicing mechanism. Some of the resulting LAT-containing transcripts can be then reverse-transcribed and the LAT insert is finally found integrated in the corresponding gene at the latent stage of infection. These modified genomes allow for the synthesis of modified transcripts.

Therefore, LATs are a new type of mobile elements with unique properties which are able to modify both the viral and the host genomes. This unique property of LATs opens new horizons on HSV1 biology, particularly on latency regulation.

Finally, the identification of this viral mobile element, that shares functional and structural similarities with both prokaryotic and eukaryotic mobile elements, raises major questions about evolution.

Table des matières

Chapitre I

Le virus Herpès Simplex de type 1	13
I.1 Le virus Herpes simplex-classification et généralités	14
I.2 La structure du HSV-1	16
I.2.1 Le tégment	17
I.2.2 L'enveloppe	18
I.3 Organisation du génome.....	18
I.4 La réplication du HSV-1	21
I.4.1 Attachement et entrée du virus	21
I.4.2 Le transport du virus.....	23
I.4.3 Expression des gènes	24
I.4.4 Réplication de l'ADN viral.....	26
I.4.5 Assemblage, encapsidation et bourgeonnement.....	29
I.5 Histopathologie de la cellule infectée.	31
I.6 Pathogenèse : neurovirulence et la latence.....	32
I.6.1 La neurovirulence	33
I.6.2 Établissement, maintien et réactivation de la latence	33
I.7 Les infections secondaires au HSV-1.....	39
I.7.1 Terminologie	37
I.7.2 Transmission du HSV-1	40
I.7.3 L'infection primaire et récurrente.....	41
I.7.4 Épidémiologie.....	42
I.8 Réponse immunitaire.....	45
I.9 La protéine virale ICP0	48

Chapitre II

Les transcrits associés a latence : LAT (<i>Latent Associated Transcript</i>)	52
II.1 Introduction.....	53
II.2 Organisation et rôle du locus des LAT.	54
II.3 Influence des LAT sur la réactivation.....	58

II.4 Influence des LAT sur l'établissement de la latence.	59
II.5 Influence des LAT sur la neuro-virulence.	60
II.6 Influence des LAT sur le lieu de la latence.....	60
II.7 Influence des LAT sur la protection des cellules neuronales et leur effet anti- et pro-apoptotique.	61
II.8 Influence des LAT sur la réponse immunitaire spécifique.	62
II.9 Action des LAT sur la transcription des gènes viraux.....	63
II.10 Action des LAT sur l'état de la chromatine	63
II.11 Par quel(s) mécanisme(s) les LAT influent-ils sur autant de phénotypes ?.....	65
II.11.1 Le LAT2KB : régulateur antisens de la protéine ICP0 ?	65
II.11.2 Le LAT2KB code-t-il pour une protéine ?	65
II.11.3 Le LAT 2kb peut-il affecter la biogenèse des ribosomes ?.....	66
II.11.4 Le LAT2KB comme régulateur de l'épissage des gènes viraux et cellulaires ?.....	67
II.11.5 Le LAT2KB comme source de micro-RNA (mi-RNA) ?	67
II.11.6 Le locus du LAT2KB peut-il jouer un rôle de « super enhancer » ?	68
II.12 Le LAT2KB peut-il agir à travers une combinaison de ses différents mécanismes ?	70
II.13 Comment un intron de grande taille sous forme circulaire peut-il agir sur autant de voies de régulations ?	75

Chapitre III

Les éléments mobiles	77
III.1 Généralités	78
III.2 La classification des éléments mobiles eucaryotes.....	79
III.4 Les éléments mobiles des organites cellulaires	80
III.4.1 Les introns du groupe II.....	80
III.4.1.1 Généralités.....	80
III.4.1.2 Structure des introns du groupe II.....	81
III.4.1.3 Mécanisme réactionnel.....	82
III.4.1.3.1 Auto-épissage et assemblage du complexe ribonucléoprotéique	82
III.4.1.3.2 Mobilité des introns du groupe II	82
III.5 Les éléments transposables de classe I	83
III.5.1 Les rétrotransposons à LTR.	83

III.5.2 Les rétrotransposons humains sans LTR	86
III.5.2.1 Les éléments L1	87
III.5.2.1.1 Généralités.....	87
III.5.2.1.2 Structure	87
III.5.2.1.3 Mécanisme de transposition	88
III.5.2.2 Les éléments Alu.....	90
III.5.2.3 Éléments non autonomes : Sines (short interspersed nuclear sequences).....	90
III.5.2.4 Les éléments SVA.....	91
III.6 Les transposons à ADN	92

Chapitre IV

Matériels et méthodes.....	94
IV.1 Rationnel de notre choix de modèle animal	95
IV.1.1 Les prérequis du modèle animal idéal	95
IV.1.2 Les différents modèles utilisés de façon historique.....	95
IV.1.3 Description précise du modèle murin oro-oculaire	98
IV.2 Type de souris utilisées	100
IV.3 Souche virale	101
IV.4 Préparation des concentrés viraux	102
IV.5 Titrage du virus	102
IV.6 Inoculation des animaux.....	103
IV.7 Sacrifice des animaux.....	103
IV.8 Extraction des ARN totaux à partir des biopsies de souris	103
IV.9 Extraction des ADN totaux à partir des biopsies de souris	104
IV.10 Extraction des ADN viraux à partir de cellules infectées en culture	104
IV.11 Les différentes techniques de Polymerase Chain Reaction (PCR) utilisées au cours de ce travail.....	105
IV.11.1 RT-PCR	105
IV.11.2 Touch-down-PCR.....	106
IV.11.3 i-PCR	107
IV.11.4 q-PCR	110
IV.11.4.1 Quantification des produits contenant les LAT	111

IV.11.4.2 Quantification des produits contenant la jonction 3' entre les LAT et le gène <i>RL2</i>	115
IV.11.4.3 Quantification des produits contenant le gène <i>RL2</i> du virus HSV-1	117
IV.12 Considérations générales sur le séquençage des produits de PCR.....	117
IV.13 Technique du RACE (Rapid Amplification of cDNA Ends).....	118
IV.14 Southern blot.	118
IV.4.1 Digestion de l'ADN génomique	118
IV.14.2 Migration de l'ADN	118
IV.14.3 Préparation du transfert de l'ADN.....	118
IV.14.4 Marquage des sondes radioactives.....	118
IV.14.5 Transfert de l'ADN sur la membrane	119
IV.14.6 Hybridation de la membrane	120
IV.14.7 Autoradiographie de la membrane	120

Chapitre V

Resultats 1^{er} partie "A new mobile element is encoded by HSV1"	121
--	-----

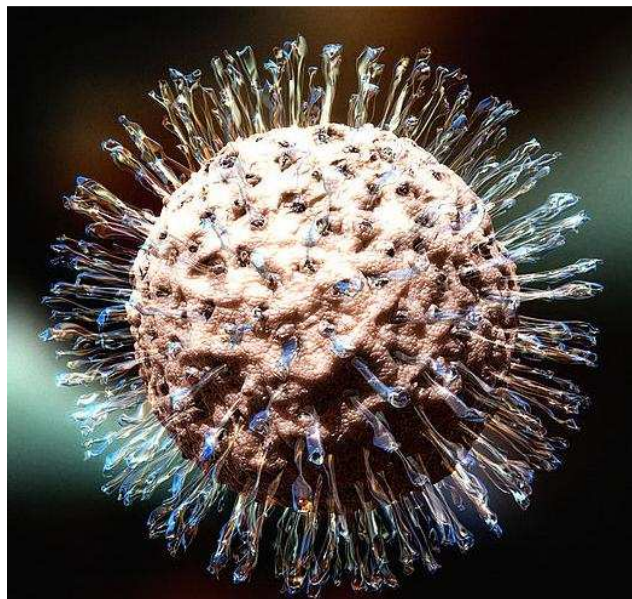
Chapitre VI

Resultats 2nd partie	164
VI.1 Introduction	165
VI.2 Recherche des jonctions par PCR	165
VI.3 Analyse des jonctions et des sites d'insertion du LAT dans les transcrits viraux et cellulaires au moment de la phase aigüe de l'infection virale.....	166
VI.3.1 Caractérisation des transcrits du gène <i>RL2</i> modifiés par insertion du LAT.	166
VI.3.2 Caractérisation des transcrits du gène <i>US12</i> modifié par insertion du LAT	168
VI.4 Recherche de l'ensemble des transcrits viraux et cellulaires susceptible d'être modifiés par le transcrits LAT.	170
VI.5 Recherche des jonctions 5' entre les gènes cellulaires, viraux et le LAT par PCR au moment de phase aigüe et latente de l'infection.....	172
VI.6 Conclusion : l'hypothèse du « reverse splicing ».....	173
VI.7 Données de séquençages	176

Bibliographie	201
Annexe 1	231
BIOLOGICAL FEATURES OF HERPES SIMPLEX VIRUS TYPE 1 LATENCY IN MICE ACCORDING TO THE VIRAL STRAIN, THE ROUTE OF INOCULATION AND THE TYPE OF NEURONES	
Annexe 2	270
Meganuclease-mediated Inhibition of HSV1 Infection in Cultured Cells	
Annexe 3	279
Structural Variability of the Herpes Simplex Virus 1 Genome <i>In Vitro</i> and <i>In Vivo</i>	

Chapitre I :

Le virus Herpes Simplex de type 1

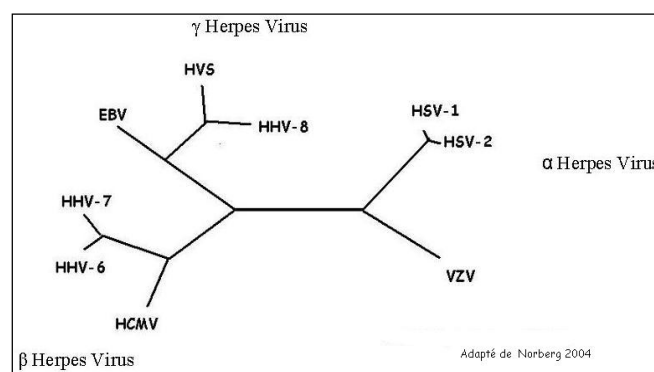


1.1 Le virus Herpes simplex-classification et généralités

Le virus de l'herpès simplex (HSV-1) fait partie de la grande famille des *Herpesviridae*. Ces virus sont grandement disséminés dans la nature et plus de 130 membres ont été identifiés jusqu'à maintenant. Les virus herpétiques partagent certaines caractéristiques physiologiques telles que leur structure, un large génome d'ADN double brin, des traits généraux de leur cycle de réplication, et notamment, leur capacité à établir la latence. Neuf virus herpès humains (HHV) ont été isolés jusqu'à maintenant (HHV-1 à -8; HHV-6A et -6B) et certaines particularités de leurs propriétés biologiques ont mené à leur subdivision en trois sous-familles; les *Alphaherpesvirinae*, les *Betaherpesvirinae*, et les *Gammapherpesvirinae* (Figure 1).

Les genres *Simplexvirus* (HSV-1 et -2), et *Varicellovirus* (HHV-3 ou varicella-zoster virus (VZV)) sont compris dans la sous-famille des *Alphaherpesvirinae*. (figure 2) Ces virus ont été initialement regroupés en raison du type de cellules qu'ils peuvent infecter, de leur rapidité de propagation en culture cellulaire, de leur efficacité à détruire les cellules qu'ils infectent (cycle lytique) et de leur capacité à établir une infection latente, principalement dans les neurones sensitifs (Roizman & Pellet, 2001). De façon typique, le HSV de type 1 (HSV-1) entraîne des infections orolabiales, alors que le HSV-2 infecte les muqueuses génitales. Le VZV est quant à lui l'agent étiologique de la varicelle et ses réactivations chez les individus plus âgés ou les patients immunodéficients sont responsables des épisodes de zona.

A



d'après Norberg 2004.

Figure 1 : Arbre phylogénétique des Herpesviridae humains et de l'herpesvirus Saimiri

Les *Betaherpesvirinae* comprennent les genres *Cytomegalovirus* (HHV-5 ou cytomégalovirus (CMV)) et *Roseolovirus* (HHV-6A, HHV-6B, et HHV-7) qui sont principalement caractérisés par le nombre restreint de cellules qu'ils peuvent infectées (figure 2). Leur cycle réplcatif est plus long, et ils se propagent plus lentement en culture cellulaire. Ces virus, et particulièrement le CMV, entraînent typiquement un élargissement des cellules infectées (cytomégalie). Les sites de latence de ces virus ne sont pas encore tous connus, mais sont en particulier impliquées les glandes sécrétrices (salivaires), les cellules lymphoréticulaires et les cellules rénales (Roizman & Pellet, 2001).

Le CMV est associé à plusieurs syndromes cliniques, notamment chez les individus immunodéprimés (spécialement les individus transplantés, qu'il s'agisse d'organes solides ou de moelle osseuse). A ce jour, le VHH-6A n'est pas associé à une maladie spécifique, mais il est souvent isolé chez les patients infectés par le virus de l'immunodéficience humaine (VIH) (Roizman & Whitley, 2001). En revanche, l'éventail de syndromes reliés au VHH-6B est en continuelle expansion, et il est entre autre retrouvé lors de roséoles chez l'enfant et d'encéphalites chez les transplantés (Mori Y *et al* .1988). Le VHH-7 est également un des agents causal de roséoles.

Herpesvirus	Nom commun (maladie associée)	sous unité	Taille du génome (Kpb)	siège de la latence
HHV-1	Herpes simplex type I (Herpes labial)	alpha	152	Neurones
HHV-2	Herpes simplex type II (Herpes génital)	alpha	152	Neurones
HHV-3	Varicella-Zoster virus (Varicelle, Zona)	alpha	125	Neurones
HHV-4	Epstein Barr virus (Mononucléose, lymphome de Burkitt, Carcinome du rhinopharynx)	gamma	172	Lymphocytes B
HHV-5	Cytomegalovirus (Mononucléose)	bêta	248	Cellules hematopoïétiques
HHV-6a	Roseolovirus	bêta	159	Monocytes
HHV-6b	Roseolovirus	bêta	162	Monocytes
HHV-7		bêta	145	CD4 +
HHV-8	KSHV (Sarcome de Kaposi)	gamma	170	Cellules épithéliales, Lymphocytes

Figure 2 : Principaux membres de la famille des Herpesviridae pathogènes pour l'homme.

Finalement, les *Gammaherpesvirinae* incluent deux genres : les *Lymphocryptovirus* (HHV-4, ou virus Epstein-Barr (EBV)) et les *Rhadinovirus* (HHV-8, ou virus associé au sarcome de Kaposi) (figure 2). Ceux-ci se distinguent par leur capacité à se répliquer *in vitro*

dans des cellules lymphoblastiques. Certains peuvent mener à une infection lytique dans des cellules de type épithélial et fibroblastique. Les virus de cette sous-famille infectent préférentiellement les lymphocytes B et les sites de latence semblent se situer majoritairement dans les tissus lymphoïdes (Roizman & Pellet, 2001). Ces virus possèdent également un pouvoir de transformation des cellules (oncogénicité). L'infection par EBV peut se manifester par une mononucléose infectieuse mais ce virus est également impliqué dans plusieurs néoplasies telles que le lymphome de Burkitt, le carcinome nasopharyngé, et le syndrome lymphoprolifératif en post-transplantation. L'EBV est également fréquemment isolé chez les patients atteints du SIDA (Cohen, 2000, Roizman & Whitley, 2001).

Le HHV-8 est le virus herpétique le plus récemment identifié et il est l'agent étiologique du sarcome de Kaposi. Il est également impliqué dans le syndrome de Castelman, ainsi que dans certains lymphomes, par exemple ceux des cavités corporelles (Cannon & Cesarman, 2000, Roizman & Whitley, 2001, Schulz, 2000).

1.2 La structure du HSV-1

Le HSV-1 a été le premier virus herpétique identifié et est l'un des virus les plus étudiés à travers le monde du fait de ses caractéristiques biologiques et des pathologies humaines qu'il induit. Le HSV-1, tout comme les autres membres de la famille des *Herpesviridae*, se caractérise par quatre éléments distincts ; I) un corps dense aux électrons comprenant l'ADN double brin; II) une capside icosaédrique; III) un tégument (une couche protéique amorphe entourant la capside) et; IV) une enveloppe (figure 3). L'ADN linéaire double brin loge dans la capside et comprend au minimum 152 kbp. La capside est constituée de 162 capsomères, 150 hexons et 12 pentons et mesure environ 100 nm de diamètre. Elle est entourée de façon très étroite par le tégument qui comprend des protéines impliquées dans l'initiation précoce de la réplication virale (α TIF ou VP16 chez HSV-1).

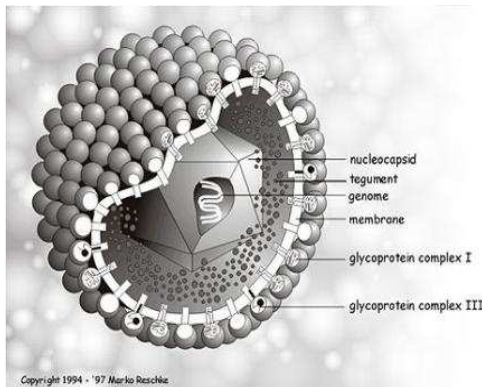
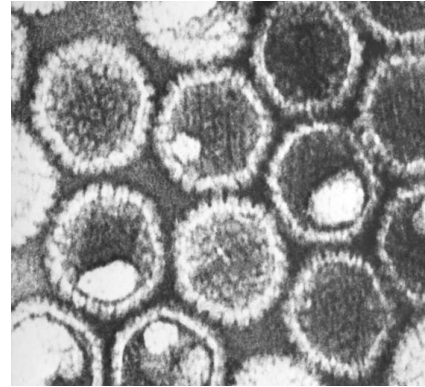
A**B**

Figure 3 : Représentation schématique de la particule virale (A) du virus HSV-1 et photo de microscopie électronique de particule virale (B).

A) Représentation de la structure de la particule virale de HSV-1, montrant les différents éléments composant celle-ci. Le virion est constitué d'une capsidie icosaédrique, d'un tégment (une couche protéique amorphe entourant la capsidie) et d'une enveloppe dans laquelle sont ancrées les glycoprotéines.

B) Photo de microscopie électronique de particule virale du virus HSV-1 la présence de l'ADN double brin génomique rend la partie centrale de la particule dense aux électrons.

I.2.1 Le tégment

Les protéines tégmentaires forment une couche de protéines diverses situées entre la nucléocapsidie (ADN viral et capsidie) et l'enveloppe lipidique du virus. Ces protéines sont acquises tout au long du périple du virus dans la cellule hôte et possèdent des fonctions indispensables à la maturation du virus. Il y aurait 24 protéines qui constituent le tégment chez HSV-1 et qui sont incorporées dans la particule virale (UL7; UL11; UL13; UL14; UL16; UL21; UL23; UL36; UL37; UL41; UL46; UL47; UL48; UL49; UL50; UL51; UL55; US2; US3; US10; US11; RL1; RL2; RS1)(Loret, S.,2008).

L'addition de tégments est séquentielle et peut se produire en différents sites dans la cellule hôte. Cette séquence d'ajout des tégments est peu détaillée et pourrait se révéler primordiale puisque les protéines tégmentaires jouent un rôle à toutes les étapes importantes du cycle viral de HSV-1. Les protéines qui constituent le tégment sont en effet impliquées dans : le ciblage des composants viraux vers les sites adéquats (site de réplication, site d'assemblage, site d'enveloppement secondaire, etc.); le transport du virus (de l'entrée du virus jusqu'à sa sortie de la cellule hôte) via l'interaction avec différentes composantes du cytosquelette et le recrutement de moteurs; la régulation de l'expression des protéines et des gènes cellulaires ou viraux; l'assemblage des virions; etc (Kelly, B.J., et al.,2009 ; Schlegel,

E.F. and J.A. Blaho, 2009 ; Yamauchi, Y., et al., 2008). Certaines protéines virales du tégument possèdent également une activité enzymatique fort utile au virus (protéine kinase, thymidine kinase, déubiquitinase, E3 ubiquitine ligase, deoxyuridine triphosphatase, etc.)(Morimoto, T., et al, 2009. Cook, W.J., et al, 1996 ; Kattenhorn, L.M., et al. 2005).

Parmi les protéines du tégument identifiées dans le tégument de HSV-1, neuf sont conservées parmi les différentes familles des virus herpès (UL7; UL11; UL13; UL14; UL16; UL21; UL36; UL37; UL51) d'où l'importance de bien caractériser leur(s) rôle(s). Les protéines tégumentaires forment des liens complexes entre elles, mais aussi avec les protéines de la capsid et les protéines intégrées dans l'enveloppe du virus.

I.2.2 L'enveloppe

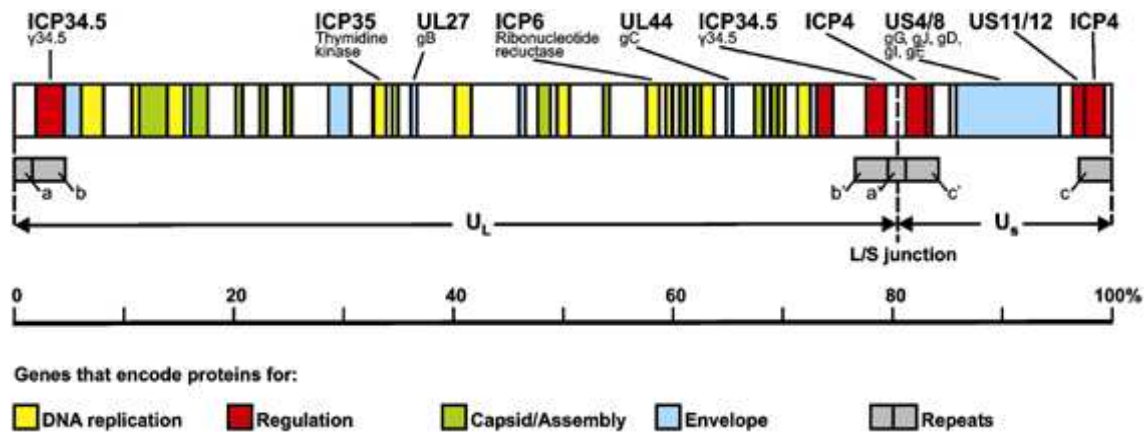
L'enveloppe entoure à son tour la capsid et le tégument, et consiste en une membrane lipidique variant entre 170 et 200 nm de diamètre. L'enveloppe virale comprend 12 types de glycoprotéines de surface différentes (gB à gM) dont plusieurs sont impliquées dans l'attachement et la pénétration du virus dans la cellule hôte. L'enveloppe contient aussi, d'autres protéines virales, comme les protéines, UL20; UL43; UL45; UL56; US9. Deux de ces glycoprotéines, gB et gD, sont particulièrement utilisées pour le développement de futurs vaccins contre l'infection au HSV-1, et la gG comprend suffisamment de différences en acides aminés (aa) entre le type 1 et le type de 2 de HSV-1 pour différencier les sérotypes (Figure 3).

I.3 Organisation du génome

Le génome du HSV-1 est constitué de deux segments désignés L (*Long*) et S (*Short* : court), liés de façon covalente. Chacun de ces segments est composé d'une séquence unique (UL ou US) flanquée par de larges séquences répétées inversées (SRI) (figure 4). Les SRI situées aux extrémités du segment UL sont désignés ab et b'a', alors que celles du segment US sont nommées a'c' et ca. Les deux segments du génome peuvent donc être inversés et combinés de diverses façons de manière à former jusqu'à quatre isomères qui diffèrent uniquement par l'orientation relative de leurs séquences (Roizman & Knipe, 2001). Le génome des souches de HSV-1 varie de 152 kbp à 155 kbp, cette variabilité étant principalement due aux variations du nombre de répétitions des SRI (Wagner *et al*, 1976). Le contenu en nucléotides G et C est de 68% pour le HSV-1 et de 69% pour le HSV-2, c'est à

dire plus élevé que celui des autres virus herpétiques humains (Roizman & Pellet, 2001). La conservation de séquence nucléotidique entre le génome complet des virus de type 1 et 2 est d'environ 85%. Les virus HSV-1 et -2 codent pour au moins 84 polypeptides différents (Ward & Roizman, 1994b).

A



Modifiée d'après R Argnani 2005

Figure 4 Représentation schématique du génome du virus HSV-1 sous sa forme linéaire.

La position des principaux gènes codant pour les protéines impliquées dans la réplication du virus (jaune), la régulation du cycle viral (rouge), la formation et l'assemblage de la capsid (vert), la formation de l'enveloppe (bleu) est indiquée. Les régions inversées répétées sont indiquées par un carré gris.

Les gènes viraux sont généralement nommés selon leur localisation dans chacun des segments UL et US, suivi d'un numéro correspondant à leur position relative de gauche à droite (ex : gène UL23=Thymidine kinase). Les protéines virales sont nommées basées sur leur propriété générale i.e. protéine infectieuse cellulaire (ICP) ou protéine du virion (VP), ou selon leur fonction respective (figure 5). Dans certains cas, elles peuvent être identifiées par le gène à partir duquel elles sont exprimées.

Les gènes compris dans les séquences uniques sont présents en une seule copie, alors que ceux encodés dans les séquences répétées sont présents en deux copies dans le génome. Les séquences répétées "a" sont importantes pour la circularisation de l'ADN dans le noyau, et dans l'emballage de l'ADN viral dans les capsides préformées. Les séquences *pac* 1 et *pac* 2 situées dans ces régions, assurent en partie cette fonction en permettant le clivage de l'ADN

en concatémères nouvellement formés et en les emballant dans les virions (Taylor *et al.*, 2002).

Les séquences "b" codent entre autre pour une protéine (ICP0) impliquée dans la régulation des gènes très précoces. Le segment UL code pour plus de 60 protéines distinctes qui sont entre autre impliquées dans la réplication de l'ADN viral et dans la formation de la capside. Les séquences "c" codent pour un activateur (ICP4) essentiel à la transcription des gènes très précoces. Finalement, le segment US code pour quatorze protéines dont les glycoprotéines de l'enveloppe virale (Roizman & Knipe, 2001).

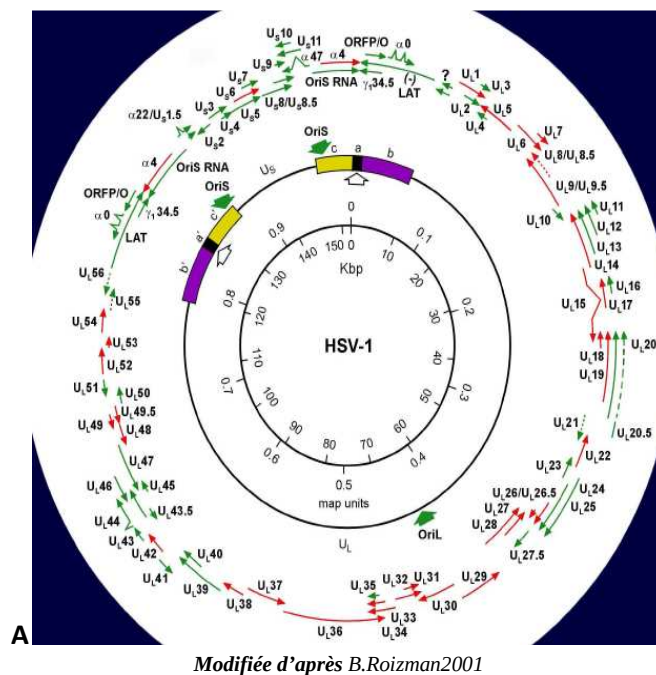


Figure 5 : Représentation schématique du génome du virus HSV-1 sous sa forme

circulaire. Le virus code près de 100 transcrits dont seuls quatre sont épissés. Les origines de réplication sont matérialisées par une flèche verte. Les transcrits des gènes indispensables sont représentés en rouge. Les autres en vert

Le génome du HSV-1 comprend trois origines de réplication nommées selon leur localisation (oriL ou oriS) (figure 5). OriL est retrouvée en une seule copie dans le segment UL, alors que oriS, située dans une séquence répétée du segment US, est présente en deux copies. OriL et oriS sont des séquences palindromiques constituées d'un centre riche en nucléotides A et T, flanquées par des séquences inversées répétées comprenant plusieurs sites

de liaison pour la protéine d'origine UL9. Une seule de ces oriS est nécessaire à l'initiation de la réplication de l'ADN viral (Taylor *et al* , 2002).

I.4 La réplication du HSV-1

L'infection productive d'une cellule par le HSV-1 requiert plusieurs étapes clés incluant l'entrée du virus, l'expression des gènes viraux, la synthèse de l'ADN viral et l'assemblage et le bourgeonnement des virions (Roizman & Knipe, 2001).

I.4.1 Attachement et entrée du virus

Pour initier l'infection, le HSV-1 doit s'attacher à des récepteurs présents à la surface des cellules et fusionner son enveloppe avec la membrane cellulaire. Les glycoprotéines de l'enveloppe du HSV-1 assurent l'attachement et la pénétration du virus dans les cellules. Parmi les douze glycoprotéines de l'enveloppe du HSV-1, gB, gC, gD, gH et gL semblent spécifiquement impliquées dans ce processus. Les glycoprotéines gE et gI forment une partie d'un complexe capable de lier le fragment Fc des anticorps. La fonction des autres glycoprotéines n'est pas encore tout à fait connue (Whitley, 2001). Les glycoprotéines induisent également une réponse immunitaire de l'hôte contre le virus.

L'attachement du virus est initié par l'interaction des glycoprotéines gC et gB avec les glycosaminoglycanes (héparane sulfates) de la surface cellulaire. Bien que ces interactions augmentent l'efficacité de l'infection, gC ne semble pas absolument indispensable à cette étape (Herold *et al* ., 1991) puisque des cellules n'exprimant pas les héparane sulfates demeurent permissives à l'infection par le virus HSV-1 (Shieh *et al.*, 1992, WuDunn & Spear, 1989). L'attachement viral est stabilisé par l'interaction de gD et l'un des nombreux co-récepteurs cellulaires (Geraghty RJ *et al* ., 1998). Ces récepteurs appartiennent à trois grandes familles et ceux-ci sont maintenant dénommés par l'acronyme Hve (pour herpes virus entry). HveA fait partie de la famille des récepteurs du facteur de nécrose tumorale alpha (TNF- α). HveB, HveC (également connu sous le nom de nectin-1 α) et HIgR (nectin-2 α) sont inclus dans la superfamille des immunoglobulines. Finalement, une troisième famille de récepteur comptant un seul membre a été identifiée et consiste en un protéoglycane d'héparane sulfate modifié (3-O héparanes sulfates) (Shukla *et al* ., 1999). HveC semble être le principal co-récepteur cellulaire nécessaire à l'infection des cellules humaines par le HSV-1.

L'attachement de la particule virale à la membrane cellulaire est suivie de l'étape de pénétration, c'est-à-dire de libération du génome dans la cellule hôte. Celle-ci passe par la fusion de la membrane virale avec une membrane de la cellule hôte (figure 6). Le lieu de la fusion fait encore débat : la plupart des images de microscopie électronique suggèrent qu'elle se déroule au niveau de la membrane plasmique mais il n'est pas complètement exclu que dans certains types cellulaires, elle aurait lieu après internalisation du virus dans un endosome. Les glycoprotéines, gB, gH-gL, forment le noyau de machinerie de fusion du virus de l'herpès. La résolution de la structure de la glycoprotéine gB de HSV-1 a montré qu'elle avait la même organisation structurale que d'autres glycoprotéines de fusion virale comme la glycoprotéine du virus de la stomatite vésiculaire (VSV) (Roche et al 2007) dont elle était donc homologue. Même si on ne connaît qu'une seule structure de

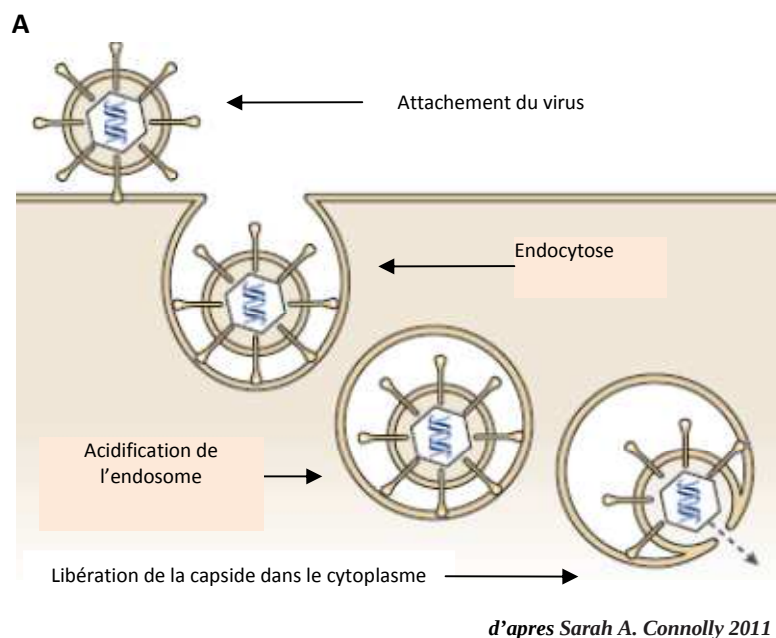


Figure 6 : Représentation des principales étapes de la pénétration du virus HSV-1 dans une cellule.

gB, par analogie avec la glycoprotéine du VSV, on pense que suite à un stimulus (probablement l'interaction de gD avec son récepteur), gB subit de grands changements conformationnels qui lui permettent de s'ancrer dans la membrane cellulaire, permettant ainsi

la formation d'un pore de fusion. Le rôle exact du couple gH-gL lors de la fusion virale est moins clair.

Une fois la fusion des membranes virale et cellulaire achevée, la nucléocapside et les protéines du tégment sont alors libérées dans le cytoplasme de la cellule hôte. La nucléocapside est transportée via les microtubules jusqu'au noyau, dans lequel l'ADN viral sera injecté à travers un pore nucléaire. Les protéines de capsid demeurent dans le cytoplasme de la cellule. Il en est de même pour la plupart des protéines du tégment à l'exception de VP16 et VP12 qui sont transportées dans le noyau indépendamment du génome (Roizman & Knipe, 2001).

I.4.2 Le transport du virus

Le transport des virus herpès demeure assez nébuleux et soulève plusieurs débats intéressants. À plusieurs étapes de déplacement du virus, différents modèles s'opposent pour tenter d'expliquer le processus de sortie. Après son entrée dans la cellule hôte, le virus HSV-1 doit se diriger vers le noyau qui constitue le site de transcription et de réplication de l'ADN viral de même que le site d'assemblage de nouvelles capsides et d'emballage du génome viral. Le virus doit aiguiller son transport vers le noyau et les protéines tégmentaires sont indispensables pour ce ciblage (Luxton, G.W., et al. 2005). Les capsides sont transportées notamment via l'interaction de la protéine virale tégmentaire UL36 avec les microtubules (Sodeik, B., et al, 1997 ; . Granzow, H., et al, 2005; Wolfstein, A., et al., 2006, Luxton, G.W., et al 2006, Mohl, B.S., et al., 2009). Il y a ensuite arrimage du virus avec le complexe du pore nucléaire (NPC) qui servira de passage pour l'insertion du matériel génétique viral dans le noyau (Batterson, W., et al, 1983. Ojala, P.M., et al. 2000. Copeland, A.M., et al, 2009. Padeloup, D., et al., 2009. Shahin, V., et al., 2006.). Les protéines virales UL6 (portal) et UL25 (bouchon du portal) ainsi que la protéine UL36 subissent des changements nécessaires à l'injection de l'ADN viral dans le noyau (Kuhn, J., et al., , 2008. Newcomb, W.W., et al., 2009.).

Les cellules humaines infectées par HSV-1 subissent d'importantes modifications morphologiques et moléculaires. HSV-1 étant un virus à réplication intranucléaire, l'infection des cellules hôte s'accompagne d'un remodelage de l'architecture du noyau, dont les nucléoles, ainsi que d'une forte inhibition de la synthèse de la majorité des protéines cellulaires. Plusieurs protéines virales se localisent au moins temporairement dans les

nucléoles, et certaines protéines nucléolaires sont délocalisées hors des nucléoles pendant l'infection. Dans le cadre des études sur les interactions virus-cellule (Greco *et al.* 2007 ; Diaz *et al.* 2002), il fut montré que la nucléoline, qui est la protéine nucléolaire majoritaire, se localise dans les compartiments de réplication virale, puis dans des aggrésomes et que l'infection par HSV-1 nécessite la présence de la nucléoline (Callé *et al.* 2008).

I.4.3 Expression des gènes

Les virus se servent de la machinerie de la cellule hôte pour exprimer leurs propres protéines, nécessaires à leur réplication et à leur prolifération. De façon générale, les virus expriment deux types de protéines : les protéines non-structurales impliquées dans la réplication de l'ADN viral et les protéines de structure qui forment le virion et une partie de l'enveloppe virale (Taylor *et al.*, 2002). Tous les virus herpétiques partagent la caractéristique d'une expression séquentielle hautement régulée de leurs gènes (Honess & Roizman, 1974). Les gènes du HSV-1 peuvent donc être classés en trois groupes selon leur apparition dans le temps: I) les gènes très précoces (dits IE ou α); II) les gènes précoces (dits E ou β) et; III) les gènes tardifs (dits L ou γ) (Figure 7). Les gènes IE sont transcrits immédiatement suite à l'infection et assurent la transcription des gènes E qui codent pour des protéines impliquées dans la réplication de l'ADN viral. Certaines protéines virales IE et E entraînent également l'expression des gènes L, qui codent majoritairement pour des protéines structurales (Whitley, 2001). Les événements clés de la réplication du HSV-1, i.e. la transcription de l'ADN, la synthèse, l'assemblage de la capside, l'emballage de l'ADN et l'acquisition de l'enveloppe, prennent place dans le noyau de la cellule infectée (Roizman & Knipe, 2001).

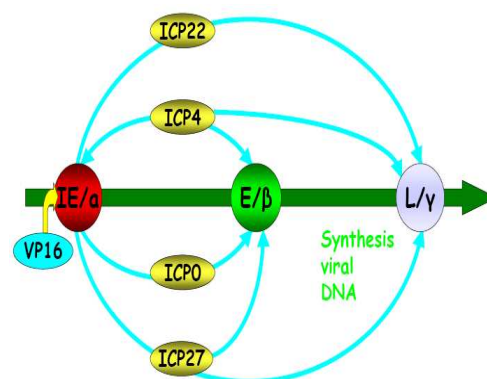


Figure 7 : Représentation de la cascade d'action des différentes classes de gènes avec leurs effecteurs viraux.

Très tôt suite à l'infection (2 à 4 h), les gènes IE sont exprimés. L'ARN polymérase II cellulaire est responsable de la synthèse de tous les ARNm viraux (Alwine *et al.*, 1974, Costanzo *et al.*, 1977). L'expression des gènes IE est indépendante des enzymes virales, mais est stimulée par un facteur régulateur de transcription appelé α TIF ou VP16 (Batterson & Roizman, 1983) (Figure 7).

Ce facteur est encodé par le gène viral UL48 et fait partie des protéines composant le tégment du virus. α TIF est donc apportée par le virus entrant et se combine à d'autres facteurs régulateurs cellulaires tels que Oct-1 et HCF (Marshall KR, *et al.*, 2000) et active la transcription des gènes IE en se liant à une séquence consensus, de type TAATGARAT, située en amont de ces gènes. Une autre protéine du tégment, le produit du gène UL41 (HSV-1), interagit avec une protéine cellulaire et dégrade rapidement les ARNm cellulaires, bloquant ainsi la traduction des protéines cellulaires. Des six protéines IE synthétisées, cinq sont impliquées dans le cycle réplcatif du virus (ICP0, ICP4, ICP27, ICP22, US1.5) (Figure 7), et par le fait même, modifient plusieurs fonctions clés de la cellule hôte. Une autre protéine (ICP47) bloque la présentation des antigènes à la surface des cellules infectées.

Les protéines IE sont souvent multifonctionnelles et elles stimulent l'expression subséquente des gènes E et L. Par exemple, en plus d'être un activateur indispensable à la transcription de tous les gènes précoces et tardifs (Dixon & Schaffer, 1980, Watson & Clements, 1980), ICP4 possède une fonction d'autorégulation en se liant avec des éléments répresseurs sur son propre promoteur (Faber & Wilcox, 1986, Kristie & Roizman, 1984). La diminution des niveaux d'ICP4 mène alors à la réduction de l'expression des gènes précoces et tardifs. La protéine ICP0 n'est pas absolument essentielle à la réplcation virale *in vitro* mais elle augmente significativement l'activité transcriptionnelle de la protéine ICP4. Pour sa part, ICP27 est essentielle à la transcription des gènes tardifs (Jean *et al.*, 2001). De plus, elle inhibe la maturation des ARN en empêchant leur épissage (Hardy & Sandri-Goldin, 1994). Contrairement aux gènes cellulaires, la plupart des transcrits viraux ne nécessitent pas d'épissage donc l'effet net d'une telle inhibition résulte en une diminution de la synthèse des protéines cellulaires. Les fonctions de ICP22 ne sont pas encore très bien comprises, mais elles semblent modifier l'ARN polymérase II cellulaire et optimiser l'expression de ICP0 (Carter & Roizman, 1996, Long *et al.*., 1999).

La deuxième classe de gènes exprimés est celle des gènes E. Leur expression nécessite la présence des protéines ICP4 et ICP27 (Figure 7). La synthèse des protéines IE atteint un niveau maximum de 5 à 7 h après le début de l'infection. Ces dernières sont majoritairement impliquées dans la réplication de l'ADN viral. Sept protéines virales issues de ces gènes sont strictement essentielles à la réplication de l'ADN viral (UL9, UL29, UL5/UL52, UL8, UL30, et UL42) dont on verra les détails dans la section suivante. Les autres, notamment la thymidine kinase virale (UL23), la ribonucléotide réductase (UL39/UL40), la dUTPase (UL50), et l'uracil ADN glucosylase (UL2), contrôlent le métabolisme des nucléotides. L'ADN polymérase (UL30) est la principale cible des thérapies anti-herpétiques courantes, qui bloquent aussi une partie de l'activité de la thymidine kinase (UL23).

On peut également subdiviser les protéines E en deux groupes selon l'initiation de leur expression dans le temps. Les gènes $\beta 1$ sont exprimés très précocement et de façon presque concomitante aux protéines IE. On y retrouve par exemple le gène UL29, codant pour la protéine de liaison à l'ADN simple brin (ICP8), et le gène UL39, codant pour la grosse sous-unité de la ribonucléotide réductase (ICP6). Les gènes $\beta 2$ pour leur part, sont exprimés dans un laps de temps un peu plus long suite à l'expression des protéines α . Le gène UL23 codant pour la thymidine kinase (TK) virale est représentatif de ce groupe de gènes (Roizman & Knipe, 2001).

Finalement, les gènes exprimés plus tardivement, i.e. les gènes L, sont principalement impliqués dans la formation de la structure du virion. Les protéines synthétisées sont alors assemblées pour former la capside et le tégument du virion. Le niveau d'accumulation de ces protéines atteint son maximum après le début de la synthèse de l'ADN viral et l'expression de ces gènes est stimulée par la synthèse de l'ADN. Cependant, les gènes précoces/tardifs ou $\gamma 1$ (par exemple : ceux codants pour ICP5, gG, gD, et ICP34.5) sont exprimés relativement tôt suite à l'infection et sont stimulés au départ par les gènes très précoces. Leur niveau se voit significativement augmenté par l'initiation de la réplication de l'ADN viral. Quant à eux, les gènes $\gamma 2$ typiques (i.e., UL44 : gC, UL41 : HSV-1, UL36 : ICP1-2, UL38 : ICP32/VP19C, et US11) sont exprimés plus tardivement dans l'infection et la synthèse de ces protéines est complètement freinée par des inhibiteurs de la réplication de l'ADN viral (Roizman & Knipe, 2001).

Sur les 84 polypeptides codés par HSV-1 (Ward & Roizman, 1994b), au moins 47 ne sont pas essentiels à la réplication virale en culture cellulaire (Figure 5). Cependant, ils peuvent être nécessaires pour pallier l'absence d'expression d'un gène cellulaire dans des cellules très différenciées ou pour altérer le métabolisme cellulaire et ainsi assurer l'efficacité de la réplication virale (Whitley & Roizman, 2001). De plus, plusieurs de ces gènes sont impliqués dans l'inhibition des mécanismes de défense de la cellule hôte. Le rôle de certaines protéines virales impliquées dans ce processus a été mentionné plus haut (ICP47, ICP27). L'apoptose est également un mécanisme de défense cellulaire important en réponse à l'infection virale, et trois protéines virales (US3, gJ, et gD) ont été identifiées comme ayant une action inhibitrice de l'apoptose (Whitley & Roizman, 2001). Par ailleurs, la présence d'ARN double brin (due à l'existence de transcrits antisens l'un de l'autre) stimule l'activation de la protéine kinase R cellulaire qui phosphoryle un facteur d'initiation de la traduction (eIF-2 α) menant à l'arrêt de la synthèse des protéines. Pour contrecarrer cet effet, la protéine virale ICP34.5 ou γ 134.5 se lie à une phosphatase cellulaire et désarme ce mécanisme de défense cellulaire (Ward & Roizman, 1994a).

I.4.4 Réplication de l'ADN viral

La réplication de l'ADN viral s'effectue dans des structures spécialisées formées dans le noyau de la cellule infectée, nommées compartiments de réplication (Quinlan *et al.*, 1984). L'ADN linéaire viral nouvellement arrivé dans la cellule hôte se circularise à son entrée dans le noyau (Uprichard & Knipe, 1996). La réplication de l'ADN est alors initiée à l'une des trois origines de réplication. Tel que discuté précédemment, la réplication de l'ADN viral nécessite absolument l'apport de sept protéines virales. La protéine de liaison d'origine UL9 se lie à des éléments spécifiques dans une des origines de réplication et a pour effet le dédoublement de l'ADN viral. UL9 recrute alors ICP8 (UL29), qui se lie à l'ADN simple brin. Ces deux protéines attirent les cinq autres protéines impliquées dans la réplication de l'ADN (UL5/UL52, UL8, UL30, UL42).

Le complexe hélicase primase (UL5/UL52/UL8) et l'ADN polymérase (pol) virale (UL30) s'associent et initient des cycles de réplication en mode θ (Figure 8). L'ADN pol est une holoenzyme comprenant deux sous-unités : une sous-unité catalytique (pol) (UL30) de 136 kDa et une sous-unité plus petite (UL42) de 65 kDa qui se lie en C terminal de l'enzyme et en stimule la processivité (Roizman & Knipe, 2001). L'interaction de ces deux unités est requise pour la réplication de l'ADN viral. L'ADN pol du HSV-1 possède une spécificité pour

un vaste éventail de substrats comparativement aux enzymes cellulaires, ce qui en fait une cible de choix pour des antiviraux. La réplication virale passe du mode thêta au mode du cercle roulant par un mécanisme encore inconnu. Ce dernier mode de réplication produit des ADN viraux concatémères qui seront ensuite clivés en monomères durant l’empaquetage de l’ADN viral (Roizman & Knipe, 2001, Taylor *et al.* , 2002).

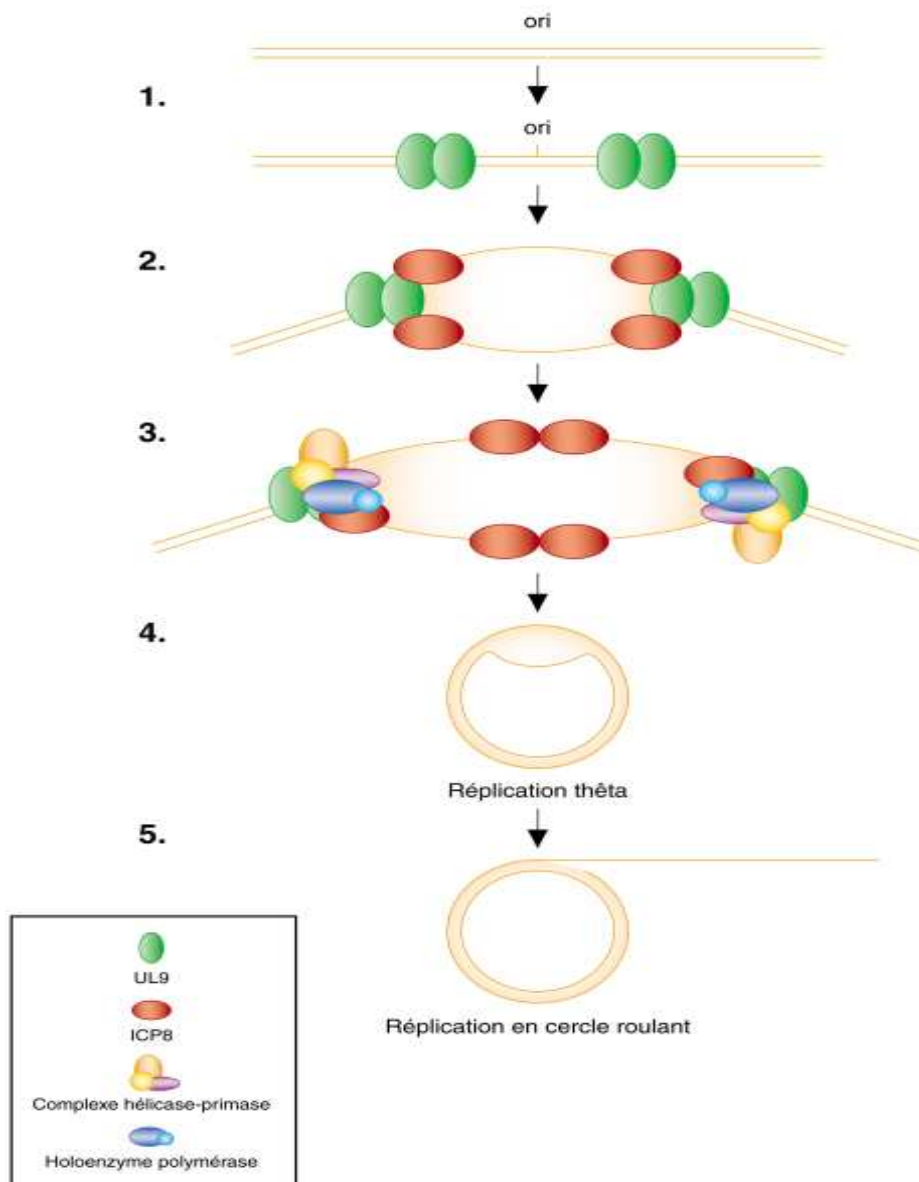


Figure 8: Schéma de la réplication de l’ADN du HSV-1

La protéine de liaison UL9 se lie à des sites spécifiques sur l’ADN viral (oriL ou oriS) et débute le désenroulement de l’ADN. 2) La seconde protéine de liaison, ICP8 est alors recrutée et se lie à l’ADN simple brin. 3) UL9 et ICP8 recrutent finalement les cinq autres protéines du complexe de réplication et 4) une réplication en mode thêta est alors initiée, pour ensuite passer au mode de réplication du cercle roulant (Modifiée de Roizman & Knipe, 2001, Taylor *et al.* 2002).

I.4.5 Assemblage, encapsidation et bourgeonnement

L'assemblage de la capside virale requiert la synthèse de plusieurs protéines virales tardives. Les quatre principales protéines de la capside (UL19:VP5, UL38:VP19c, UL18:VP23 et UL35:VP26), la protéine d'échafaudage codée par le gène UL26.5 et la protéase UL26 sont nécessaires à ce processus. D'autres protéines ont également été impliquées dans l'assemblage de la capside, mais ne sont pas indispensables, du moins dans des systèmes *in vitro* (Taylor *et al.*, 2002). Les capsides vides sont alors chargées de l'ADN viral par un processus qui implique simultanément le clivage de l'ADN concatémère en de simples unités monomériques et l'empaquetage de l'ADN, pour ainsi former la nucléocapside (Ladin *et al.*, 1980, Ladin *et al.*, 1982).

Ce mécanisme n'est pas encore très bien connu mais il a été observé que les coupures dans l'ADN s'effectuent à des sites spécifiquement déterminés par la distance entre les sites *pac* 1 et *pac* 2 (Deiss *et al.*, 1986, Varmuza & Smiley, 1985). La protéine codée par le gène UL15 en association avec celle codée par le gène UL26, semble impliquée dans le clivage de l'ADN en cours de réplication. L'empaquetage de l'ADN viral dans les capsides matures requiert plusieurs protéines virales incluant le produit des gènes UL6, 15, 17, 25, 28, 32, et 33. Le complexe UL15-UL26, semble initier l'entrée de l'ADN viral dans la capside. L'exonucléase codée par le gène UL12 joue également un rôle dans l'empaquetage de l'ADN viral (Roizman & Knipe, 2001). L'ADN alors encapsidé se trouve à une concentration de 300mg/mL dans un état liquide cristallin. (Booy *et al* 1991). Un travail, auquel notre équipe a collaboré au cours de ma thèse, a permis de montrer que les quatre isomères génomiques du virus n'étaient pas représentés de façon identique au sein de la population virale. La distribution dépend de la souche virale et du contexte dans lequel les particules virales ont été produites (Mahiet *et al* 2012, article en annexe). Une fois l'encapsidation réalisée les capsides résultantes ont des angles très marqués et sont donc facilement distinguables en microscopie électronique.

Les nucléocapsides bourgeonnent ensuite au niveau de la membrane nucléaire interne, où elles acquièrent une enveloppe et certaines protéines du tégument. La route empruntée par les virions pour sortir de la cellule est encore controversée. Une des hypothèses propose que les virus enveloppés fusionnent avec la membrane nucléaire externe, et que les nucléocapsides alors libérées dans le cytoplasme de la cellule pour être ré-enveloppées lors du bourgeonnement à partir des compartiments de l'appareil de Golgi (Roizman & Knipe, 2001)

(Figure 9). Ces virus ré-enveloppés sont alors conduits à la surface de la cellule par un transport vésiculaire.

L'autre mécanisme suggéré propose que les virions enveloppés soient transportés au sein de vésicules depuis la membrane externe du noyau, à travers l'appareil de Golgi et finalement sécrétés à l'extérieur de la cellule (Kim JS *et al.*, 1999) (Figure 9). Les évidences les plus récentes font maintenant pencher la balance en faveur du premier modèle de ré-enveloppement (Skepper *et al.*, 2001).

Plus récemment, un troisième modèle s'est ajouté pour tenter d'expliquer le transport des capsides fraîchement assemblées dans le noyau. Il s'agit du « modèle des pores élargis » qui suggère que les pores nucléaires dont le canal central qui mesure normalement moins de 25 nm, sont dilatés par un processus actif durant l'infection pour atteindre une taille capable d'accueillir le passage direct des capsides de plus de 100 nm (Figure 9). Ce modèle exclut le bourgeonnement entre les capsides intranucléaires et la membrane nucléaire interne (Wild, P., et al., 2005. Leuzinger, H., et al., 2005.).

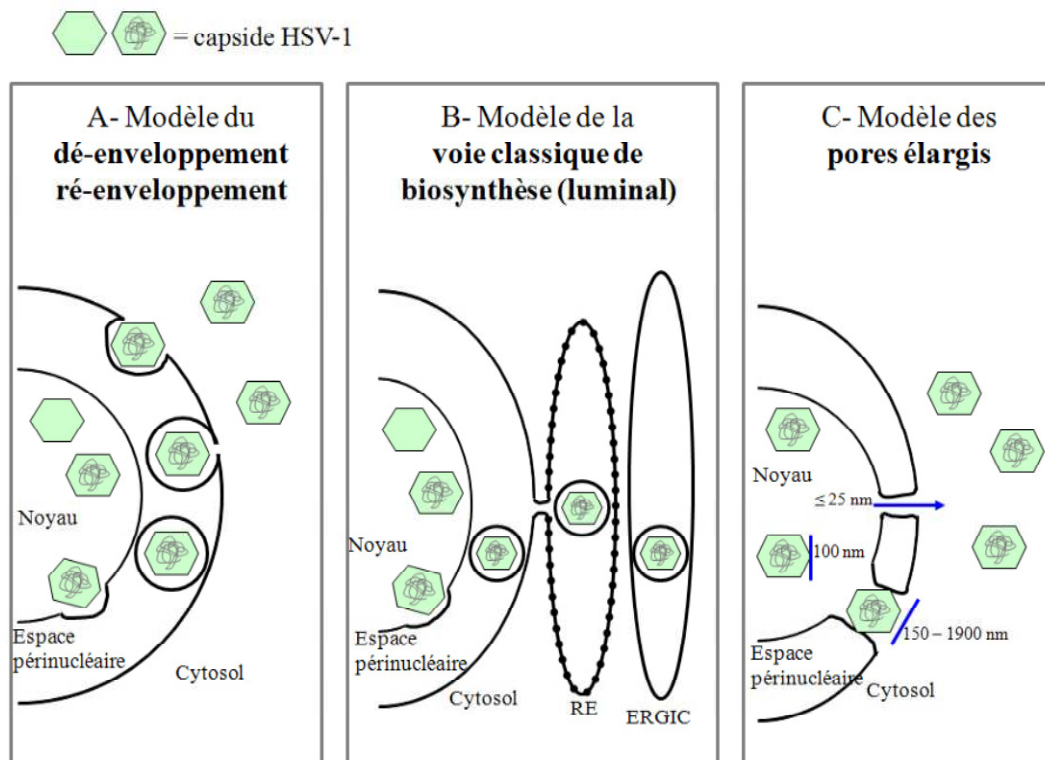


Figure 9: Schématisation des modèles de transport hors du noyau de la cellule hôte pour le virus HSV-1. A. Modèle de dé-enveloppement et ré-enveloppement développé par Stackpole (1969). B. Modèle luminal présenté par Darlington et Moss (1968) et C. Modèle des pores élargis amené par P. Wild (2005). ERGIC (compartiment intermédiaire entre le réticulum endoplasmique et l'appareil de Golgi). RE (réticulum endoplasmique)

Un autre aspect du transport du virus HSV-1 qui suscite des débats considérables concerne le transport du virus dans les cellules neuronales. Certaines équipes croient que les composantes virales (enveloppe et nucléocapside) sont transportées séparément dans l'axone et qu'elles sont rassemblées loin du corps neuronal près de la terminaison de l'axone. Il s'agit du « modèle séparé » de transport axonal (Enquist, et al 2002. LaVail, J.H., et al., 2003. Snyder, A., et al. 2006.). Un autre modèle « le modèle marié » suggère plutôt que le virus est complètement assemblé et enveloppé au niveau du corps du neurone. Les virions complets voyagent ensuite le long de l'axone dans des vésicules de transport (Del Rio, T., et al., 2005. Feierbach, B., et al 2007).

Comme les glycoprotéines virales sont disséminées dans toutes les membranes de la cellule suite à leur synthèse, elles sont acquises par le virion en même temps que l'enveloppe virale et se retrouvent donc à la surface de la particule. La réplication virale est approximativement achevée en l'espace de 18 h et ce processus entraîne des dommages cellulaires irréversibles (Roizman & Knipe, 2001). Ce schéma classique d'une réplication virale obligatoirement létale pour la cellule hôte est vérifié en culture cellulaire, mais il est possible que *in vivo*, les neurones, ou au moins certains sous types de neurones puissent survivre à un début de cycle de réplication virale, voire à la véritable production de particules infectieuses, grâce à l'invalidation des mécanismes apoptotiques et/ou autophagiques (Lussignol M ; 2012). Ce type de mécanisme pourrait rendre compte de certaines observations réalisées chez la souris.

I.5 Histopathologie de la cellule infectée.

L'aspect cytologique après infection par HSV-1 est le reflet des désordres métaboliques entraînés par l'infection virale et de la réponse inflammatoire associée. La première cellule cible du HSV-1 est la cellule muco-épithéliale sur laquelle les virus se fixent et se multiplient. L'infection virale entraîne le gonflement des cellules qui démontrent alors une chromatine condensée, suivi de la dégénérescence du noyau de celles-ci. Ces cellules perdent leur intégrité membranaire et forment des cellules géantes multinucléées pouvant contenir des inclusions intranucléaires de type Cowdry A. Lors de la lyse cellulaire, un liquide clair (liquide vésiculaire) renfermant une quantité importante de virions est libéré entre l'épiderme et le derme. Ce fluide contient des débris cellulaires, des cellules inflammatoires, et souvent des cellules multinucléées. La lésion caractéristique au niveau de la peau se

présente sous la forme d'une vésicule à liquide clair sur un fond érythémateux. Durant le processus de guérison, les vésicules évoluent en pustules, ulcérations et croûtes et ne laissent généralement pas de cicatrice. Sur une muqueuse, l'aspect de la lésion est celui d'une ulcération, du fait de l'érosion du toit de la vésicule (Roizman & Knipe, 2001)

1.6 Pathogenèse : neurovirulence et la latence

Le HSV-1 se caractérise spécialement par deux de ses propriétés biologiques : sa capacité à envahir le système nerveux central (SNC) et à s'y répliquer (i.e. la neurovirulence) et sa capacité à établir une infection latente. Cette dernière propriété est également partagée par les autres membres de la famille des *herpesviridae*, et le lieu de la latence, i.e. les ganglions sensitifs de la moelle épinière, est caractéristique de la sous-famille des *α -herpesvirinae*.

1.6.1 La neurovirulence

La neurovirulence du HSV-1 est habituellement reliée à sa virulence ou pathogénicité et celle-ci peut être évaluée dans des modèles animaux expérimentaux. La souris, le cochon d'Inde et les lapins sont les animaux les plus utilisés et ceux-ci possèdent des caractéristiques propres aux différents types d'infections étudiées. La souris, en raison de sa facilité de manipulation, demeure le modèle animal le plus utilisé pour l'étude de la pathogénicité du HSV-1. Suite à une inoculation intracérébrale ou intranasale du HSV-1, la réplication virale mène habituellement à une encéphalite, éventuellement mortelle selon les souches et le mode d'inoculation. La dose qui entraîne 50% de létalité (DL50) peut donc être mesurée pour évaluer la pathogénicité d'une souche de HSV-1, ou comparer plusieurs souches. Dans un autre type d'essais, la neurovirulence est évaluée par la capacité de la souche à envahir le SNC et/ou le système nerveux périphérique suite à son inoculation à un site périphérique (cornée, paume des pattes, oreille, racine du nez, lèvre). Pour un inoculum donné, on peut aussi mesurer le nombre de particules infectieuses au sein d'une structure anatomique précise dans le cerveau (e.g. ganglion trigéminal). Cette méthode, complétée par des approches de type FISH, permet également d'évaluer la capacité des virus à entrer en latence dans les neurones sensitifs de même que de suivre le maintien et la réactivation de la latence (discuté dans la section Matériel et Méthodes).

La neurovirulence du HSV-1 dépend de plusieurs gènes. En effet, la délétion de pratiquement tous les gènes dispensables pour la réplication du HSV-1 *in vitro* mène à une diminution de sa neurovirulence (Roizman & Knipe, 2001). Des mutations affectant la neurovirulence du HSV-1 ont également été identifiées dans les glycoprotéines de l'enveloppe virale (Whitley, 2001). Enfin, bien que des virus mutants portant une délétion du gène γ 34.5 se multiplient dans plusieurs types de cellules en culture, ces mutants sont parmi les virus les moins virulents identifiés jusqu'à maintenant (Chou *et al.* , 1990, Spector *et al.* , 1998,).

I.6.2 Établissement, maintien et réactivation de la latence

L'habileté du HSV-1 à entrer en latence et à maintenir cet état durant toute la vie d'un individu infecté est certainement l'une des propriétés les plus remarquables du HSV-1 dont on ne connaît malheureusement pas encore tous les détails. De façon théorique, la latence de HSV-1 peut se définir comme étant une infection durant laquelle le génome viral est présent sous un état non réplcatif dans une cellule infectée, à partir duquel il peut se réactiver de façon intermittente (Miller *et al.*, 1998).

En période de latence, le génome du HSV-1 est sous une forme circulaire épisomale dans le noyau des cellules nerveuses, et contrairement à l'infection lytique, aucune réplication virale ne peut être détectée. L'intégrité de la cellule hôte n'est donc pas compromise et les antigènes viraux caractéristiques d'une infection productive ne sont pas exprimés. Il est classiquement admis que les seuls transcrits produits en quantité facilement mesurables (northern-blot, hybridation *in situ*, et évidemment PCR) sont les transcrits associés à la latence (LAT pour *Latency Associated Transcript*). Il a été décrit dans certains modèles et dans certaines conditions d'inoculation que d'autres transcrits viraux peuvent aussi être décelés pendant cette période d'infection latente, sans pour autant que les protéines associées soient détectées.

En raison de cette extinction de l'expression des protéines virales, le virus est ainsi à l'abri du système immunitaire de l'hôte infecté qui ne semble pas reconnaître l'infection latente. Les sites habituels de latence du virus HSV-1 sont les ganglions sensitifs trigéminés alors que les ganglions sensitifs sacrés sont les sites les plus communs de la latence du HSV-2. De façon périodique, le virus se réactive et les capsides sont véhiculées par transport axonal vers les synapses situées aux extrémités des terminaisons neuronales (axone ou dendrite selon les cas), où les particules virales matures pourront bourgeonner pour se répliquer dans les

tissus alentours. En fonction de plusieurs facteurs, parmi eux l'état du système immunitaire, cette réactivation entraînera une infection symptomatique ou non. Il est classiquement admis que le site de primo-infection est le site privilégié des récurrences herpétiques, mais cela n'explique pas les cas d'herpès oculaire et mêmes les encéphalites, puisque la presque totalité des primo-infections à HSV-1 ont lieu dans les muqueuses de la sphère oro-pharyngée.

Dans les modèles d'infection animaux, les virus sont inoculés à des sites périphériques où ils se répliquent et la latence s'instaure dans les ganglions connectés au site de primo-infection. Cependant, contrairement à l'humain, la réactivation du virus chez l'animal s'effectue rarement de façon spontanée (sauf chez le lapin) et l'extrapolation des observations obtenues chez l'animal à l'infection chez l'humain doit être réalisée avec précaution. La latence du HSV-1 peut être divisée en trois phases : I) l'établissement, II) le maintien et III) la réactivation de la latence.

L'établissement de la latence est précédé par la réplication du HSV-1 dans des sites périphériques (muqueuses)(Figure 10). Dès lors, des virus pénètrent dans les terminaisons nerveuses des nerfs sensitifs innervant l'épithélium des muqueuses et sont transportés de façon rétrograde par les microtubules, jusqu'aux noyaux des cellules nerveuses (Lycke *et al* ., 1984). Les virus peuvent alors entrer dans un cycle de réplication productif ou lytique. En fait, il a été démontré que la réplication virale s'effectue dans une minorité de neurones et que finalement, après quelques semaines (4 semaines chez la souris), on ne détecte plus dans les ganglions que des génomes viraux et quasiment plus de transcrits (Miller *et al* ., 1998).

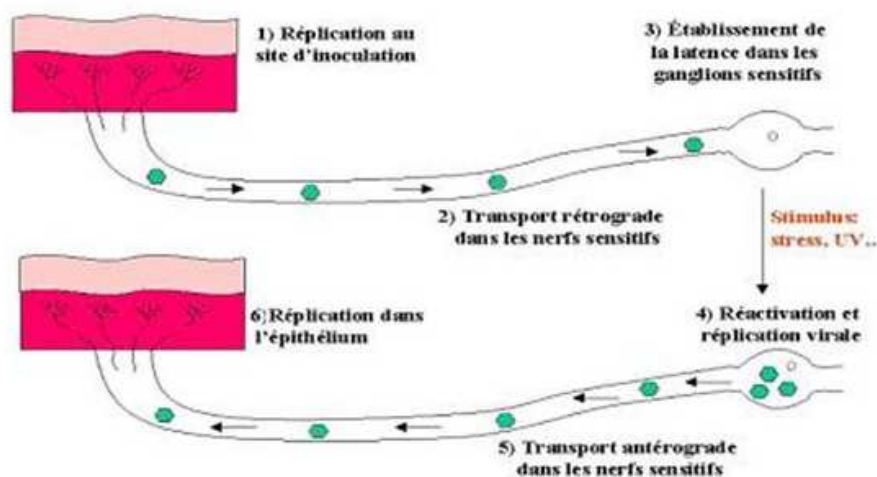


Figure 10: Représentation schématique des différentes étapes menant à la latence et à la réactivation de HSV-1

D'après les études *in vitro*, il est clair que la réplication virale et les protéines VP16, ICP4, et ICP27, essentielles à l'infection lytique, ne sont pas requises pour l'établissement de la latence (Sedarati *et al.*, 1993). D'ailleurs, les virus possédant des mutations dans les gènes codant pour ces protéines semblent tous capables d'entrer en latence, alors que la réactivation et le maintien de celle-ci paraissent en être affectés.

Une hypothèse plausible concernant l'établissement de la latence est un mécanisme dit par "défaut". L'expression des gènes viraux IE semble en effet être réprimée par certains facteurs cellulaires et viraux (Preston, 2000). Les facteurs de transcription cellulaire (*c-jun*, *c-fos*, et Oct-1) nécessaires à l'expression des gènes viraux (Sedarati *et al.*, 1993) sont absents ou présents en faible proportion dans les cellules neurales hautement différenciées, ce qui pourrait restreindre l'expression des gènes IE du HSV-1. De façon complémentaire, les cellules neurales expriment d'autres facteurs de transcription (Oct-2, et N-Oct-3) qui pourraient entrer en compétition avec VP16 pour les sites de liaison au facteur de transcription cellulaire Oct-1 (Kemp *et al.*, 1990) et mener à la répression de l'expression des gènes IE. Certains investigateurs ont suggéré que les niveaux réduits de VP16 dans les cellules neurales pourraient être en partie occasionnés par la perte de certaines protéines du tégment du HSV-1 en périphérie (Sears *et al.*, 1991). L'absence ou la diminution de

l'efficacité de VP16 entraînerait une diminution des niveaux d'expression de ICP0 et de ICP4, importantes pour l'initiation de l'infection lytique (Miller *et al.* , 1998).

Jusqu'à maintenant, aucun facteur cellulaire n'a été identifié dans le processus de maintien de la latence du HSV-1 *in vivo*. Toutefois, le facteur de croissance neurale (NGF) semble impliqué à certains égards puisque des cultures *in vitro* de ganglions infectés de façon latente en absence de NGF conduisent à une réactivation instantanée du virus alors qu'en sa présence, la production de particules virales est inhibée (Block *et al.*, 1994, Halford *et al.*, 1996, Wilcox & Johnson, 1988). Sa fonction n'a cependant pas encore été déterminée de façon très claire. Aucun gène viral n'a été impliqué de façon définitive dans le maintien de la latence. La TK virale a un rôle soupçonné mais cela est seulement dû au fait que des virus dépourvus de TK fonctionnelles ont une expression moindre de marqueurs de latence comme les LATs dans certains modèles d'infection (Tenser *et al.* , 1994). Plus récemment, des études ont démontré que des HSV-1 délestés des gènes UL5 et UL29 ont un défaut dans leur capacité à se maintenir de façon stable dans les ganglions sensitifs trigéminals (Da Costa *et al.*, 1999).

Plusieurs équipes ont clairement montré que la régulation de la latence peut se faire, au moins en partie, de manière épigénétique. Contrairement aux gamma- et beta-herpesvirus, la méthylation des nucléotides CpG ne joue pas de rôle dans la régulation des promoteurs des gènes lytiques (Dressler, Rock, Fraser, 1987). Cette différence peut refléter un mode de contrôle des mécanismes transcriptionnels différents dans les cellules où les virus établissent leurs latences. Pendant la latence de HSV, le génome viral est sous la forme d'un épisode circulaire chromatinisé (Deshmane, Fraser 1989). En fait, la totalité du génome est hétérochromatinisée, à l'exception du locus codant pour les LAT. Récemment, la protéine CTCF (CCCTC binding factor) connue pour reconnaître des éléments insulateurs chez les vertébrés, s'est avérée reconnaître plusieurs régions du génome de HSV (Amelio, 2006, Qi Chen 2007).

La protéine ubiquitaire CTCF est une protéine de 11 domaines en doigt de zinc, dont l'activité de régulation transcriptionnelle est dépendante de son état de phosphorylation. Le nombre important de domaines en doigt de zinc permet à la protéine d'augmenter son affinité pour l'ADN, là où un seul domaine a été montré comme nécessaire à son activité (Chao *et al.* 2002). L'attachement de CTCF à l'ADN permet la formation et la régulation des éléments barrières impliqués dans la chromatinisation d'un génome, grâce au recrutement de protéines

de modification de la chromatine. Il a ainsi été observé que lorsque la protéine CTCF était fixée sur le génome de HSV, le recrutement des protéines permettaient d'empêcher la diffusion de l'hétérochromatine. Deux régions majeures ont ainsi été mises en évidence. Une région dite B1 en aval du promoteur du LAT et une région dites B2 en amont de celui-ci. De plus 5 régions annexes ont été caractérisées. Pendant la latence, les 7 régions sont occupées par la protéine CTCF (Amelio 2006) permettant ainsi de moduler l'expression des gènes viraux.

La question la plus étudiée depuis une vingtaine d'années concerne l'action des LATs pendant la latence. Leurs différentes propriétés seront plus profondément débattues plus avant dans ce manuscrit. Brièvement, les LATs sont les seuls transcrits viraux détectés dans les neurones durant l'infection latente et, de ce fait, sont considérés comme les marqueurs de l'état de latence du HSV-1 (Cohrs *et al.*, 2000). Ceux-ci consistent en 4 types d'ARN hautement conservés, un long transcrit instable (8,3 kb) et trois autres transcrits très stables (2.0, 1.5, et 1.45 kb) qui s'accumulent abondamment dans les cellules infectées (Krause *et al.*, 1988, Rock *et al.*, 1987, Stevens *et al.*, 1987). Ces transcrits dérivent de séquences présentes dans les régions répétées "b" du segment UL du génome du HSV-1.

Le maintien du génome du HSV-1 dans un état de latence semble passif, il ne requiert pas l'expression de gènes viraux spécifiques ou d'un produit génétique quelconque. En fait, les facteurs impliqués dans le maintien de la latence virale ne sont pas encore bien compris. Des études récentes ont permis de démontrer que de 43 à 678 copies du génome viral sont présentes par 100 ng d'ADN de ganglions trigéminalisés d'individus infectés par le HSV-1 (Cohrs *et al.*, 2000). De plus, les LATs sont constamment détectés dans les ganglions comprenant des virus latents. Sur la base de la détection des LATs, on estime qu'au sein des ganglions, de 0,1 à 3% des neurones comprennent des génomes de HSV-1 maintenus en latence (Mehta *et al.*, 1995, Sawtell, 1997). Cependant, puisque toutes les cellules n'expriment pas des niveaux détectables de LATs, ces valeurs sont certainement très sous-estimées.

Finalement, il est admis, que le nombre de virus en latence dans une cellule infectée de même que leurs capacités à exprimer les LATs semble important dans le mécanisme d'établissement de la latence du virus HSV-1.

Chez l'homme, la réactivation du virus de son état de latence vers un mode de réplication lytique peut être observée de façon spontanée sans cause évidente ou suite à

différents stimuli. Par exemple, un traumatisme local dans les tissus innervés par des neurones infectés de façon latente avec le HSV-1 tel qu'une extraction dentaire, un stimulus systémique tel qu'un stress émotionnel, la fièvre et les menstruations, ou l'exposition aux ultraviolets, sont autant de situations qui peuvent mener à une réactivation du virus de son état de latence (Roizman & Knipe, 2001). Les virus sont alors transportés de façon antérograde le long des axones des nerfs sensitifs infectés, habituellement jusqu'au site périphérique de l'infection initiale ou à des sites voisins. À cet endroit, les virus se répliquent et entraînent une infection productive associée ou non à des symptômes cliniques.

Dans les modèles animaux, l'explantation des ganglions sensitifs, et leur mise en culture avec des cellules permissives entraînent la réactivation des virus latents (Coen *et al*, 1989b, Efsthathiou *et al*, 1989, Pelosi *et al.*, 1998a). Ce type d'étude a également démontré que la réactivation du HSV-1 se fait seulement à partir de quelques neurones infectés. Par contre, la nature du signal de transduction impliqué dans la réactivation du HSV-1 n'a pas encore été identifiée. Une induction de l'expression de certains facteurs cellulaires (*c-fos* , *c-jun* , *c-myc* et Oct-1), habituellement présents en faible quantité dans les neurones, pourrait être stimulée en raison d'un stress cellulaire quelconque et mener à l'activation de l'expression des gènes viraux (Miller *et al*, 1998). De plus, deux seconds messagers cellulaires (AMPc et protéine kinase C) semblent être impliqués dans le mécanisme de réactivation du HSV-1 *in vitro* et ceux-ci seraient alertés par une variété de stimuli. En effet, le promoteur des LATs est sensible à l'AMPc et la présence d'AMPc est corrélée avec une réactivation efficace du virus. Les événements de phosphorylation pourraient contribuer à l'activation de facteurs de transcription et d'expression de protéines virales (i.e., ICP0 et ICP4) impliquées dans l'expression des gènes viraux (Miller *et al* ., 1998).

Contrairement à ce qui se passe pour l'établissement de la latence, divers gènes viraux semblent impliqués dans la réactivation du HSV-1. En effet, plusieurs mutants démontrent une incapacité à se réactiver de l'état de latence. Parmi ceux-ci, on trouve les mutants portant des délétions dans des gènes essentiels à la réplication du virus en culture cellulaire ou dans les gènes de ICP0, de la ribonucléotide réductase et de γ 34.5, qui ont démontré une capacité de réactivation affaiblie ou nulle (Roizman & Knipe, 2001). Enfin, certaines mutations spécifiques dans le gène de la TK ont été associées, à une incapacité de HSV-1 à se réactiver bien que le virus soit resté capable d'établir et de maintenir la latence. Par ailleurs, la séquence d'activation temporelle des gènes IE, E et L dans les cellules supportant l'infection

lytique pourrait être modifiée dans les neurones et il est donc possible que l'expression des gènes viraux soit régulée de façon différente dans les cellules neuronales (Roizman & Knipe, 2001).

Une étude a ainsi démontré que les produits des gènes E (TK, ICP6, et VP5) sont détectables dans les premières heures suite à l'explantation de ganglions infectés (4 h), alors que les transcrits des gènes IE sont détectés seulement suite à une période de 8 à 24 h (Tal-Singer *et al.*, 1997). Ces observations suggèrent donc que les facteurs neuronaux permettant la réactivation du HSV-1 favoriseraient plutôt l'expression des gènes E que celle des gènes IE et que, lors de la réactivation, la réplication de l'ADN viral précéderait peut-être l'expression des gènes IE (Miller *et al.*, 1998). Finalement, le nombre de virus en latence dans une cellule infectée de même que le nombre total de copies du génome du HSV-1 présentes dans un ganglion donné semblent des facteurs prédictifs forts de la capacité de réactivation expérimentale du HSV-1 dans les modèles animaux (Miller *et al.*, 1998, Roizman & Knipe, 2001).

Le destin des neurones suite à la réactivation et à la réplication du HSV-1 est encore un sujet débattu. Si ceux-ci succombent à l'infection lytique, comment expliquerait-on que les sujets enclins à plusieurs épisodes d'infections récurrentes ne soient pas affectés par un déficit neurologique due à la destruction des sites de réactivation ? Comme on peut le constater, plusieurs questions concernant le phénomène de la latence du HSV-1 ne sont pas encore résolues et d'autres études sont nécessaires afin d'élucider plus clairement les différents mécanismes impliqués dans la latence du virus.

I.7 Les infections secondaires au HSV-1

I.7.1 Terminologie

Il importe tout d'abord de bien définir la terminologie employée pour identifier le type d'infection par HSV-1 chez l'homme, son seul hôte naturel. Les individus susceptibles à l'infection par le HSV-1, i.e. qui n'ont jamais été en contact avec le virus et qui ne possèdent donc pas d'anticorps contre celui-ci, développent une *infection primaire* suite à une première exposition au HSV-1. La grande majorité de ces primo-infections est asymptomatique (Liesegang 1989). Ces individus doivent alors être définitivement considérés comme infectés puisque HSV-1 a la propriété d'entrer et de maintenir un état de latence dans le système nerveux. Dans certains cas, le virus peut se réactiver et entraîner une *infection récurrente*,

symptomatique ou non. Une *primo-manifestation* concerne un individu qui a déjà rencontré le virus mais sans développer de signes cliniques (primo-infection asymptomatique puis entrée en latence) et qui développe pour la première des signes cliniques d'infection herpétique. Une *infection initiale* concerne un individu déjà immunisé avec un virus de type 1 ou 2, et qui développe une première infection avec le virus du type opposé.

Finalement, quoique très peu commun chez les individus immunocompétents, une *réinfection exogène* peut survenir. L'individu est donc infecté avec plus d'une souche de HSV-1 (Whitley *et al.*., 1998).

I.7.2 Transmission du HSV-1

Chez l'humain, la transmission du HSV-1 dépend de contacts plus ou moins directs durant lesquels il y a échange de salive (y compris dans les microgouttelettes de Pflügge émises lors de la respiration ou la toux) ou de fluide vésiculaire provenant de lésions infectées, entre un individu qui excrète du HSV-1 et une personne susceptible. Pour HSV-2, la contagion se fait par contact direct avec les sécrétions génitales. Dans tous les cas, les particules virales doivent entrer en contact avec une muqueuse ou de la peau abrasée pour qu'il y ait initiation de l'infection.

Aucun vecteur ni réservoir animal autre que l'homme n'a été décrit pour l'infection causée par HSV-1. L'espèce humaine demeure donc le seul réservoir pour la transmission du virus. Il ne semble pas non plus y avoir de variations saisonnières dans l'incidence de l'infection (Roizman & Knipe, 2001). Puisque l'infection au HSV-1 est rarement fatale, et que le virus demeure à l'état de latence, une personne infectée demeure une source d'infection constante tout au long de sa vie. Le HSV-1 est habituellement acquis tôt dans l'enfance, ou aux premières années de la vie d'adulte, alors que la contamination par HSV-2 intervient en général plus tard.

On sait que HSV-1 peut aussi entraîner des infections génitales, mais, de façon intéressante, elles sont habituellement à la fois moins sévères et moins susceptibles d'entraîner des récurrences que celles dues à HSV-2 (Lafferty, 2002, Lafferty *et al.* . 1987). Il y a également de plus en plus d'évidences qui suggèrent que la présence d'anticorps contre le HSV-1 protège quelque peu contre l'infection au HSV-2 (Bryson *et al.*, 1993, Mertz *et al.*, 1992), et presque totalement contre l'herpès génital causé par le HSV-1 (Lafferty, 2002). Les

femmes séronégatives pour le HSV-2 ont un risque de 80% de contracter une infection génitale (asymptomatique ou symptomatique) d'un partenaire infecté, alors que ce risque de 3 fois moindre pour un homme séronégatif en contact avec une partenaire infectée (Mertz *et al.*, 1992, Wald *et al.*, 2001, Whitley *et al.*, 1998).

Dans le cas de l'herpès néonatal, l'infection du nouveau-né se fait essentiellement au moment du passage dans la filière génitale au moment de l'accouchement. Les risques les plus élevés de la transmission du HSV-1 de la mère au fœtus surviennent lors d'une infection primaire ou initiale de la femme enceinte (30 à 50%) (Stanberry *et al.*, 2000). En revanche, les risques qu'un nouveau-né acquière l'infection herpétique suite à une exposition virale durant une infection récurrente asymptomatique sont beaucoup moins grands (moins de 4%) (Stanberry *et al.*, 2000). L'exposition du fœtus à des charges virales élevées augmenterait donc les risques de transmission du HSV-1, alors que le transfert passif d'anticorps de la mère au fœtus les diminuerait (Stanberry *et al.*, 2000). En effet, le statut immunitaire de la mère avant l'accouchement influence à la fois la sévérité de l'infection et les possibilités de transmission du virus au fœtus. Une rupture prolongée de la membrane placentaire (plus de 6 h) augmente les risques de l'infection du fœtus. De plus, l'utilisation de moniteurs intra-utérins peut également être un vecteur de contamination iatrogène (Whitley *et al.*, 1998).

I.7.3 L'infection primaire et récurrente

Dès l'infection primaire, des virus pénètrent dans les terminaisons nerveuses et sont transportés dans les ganglions sensitifs où ils demeurent à l'état de latence. Les ganglions trigéminals, cervicaux et lombo-sacrés sont les plus communément infectés. La réplication initiale du HSV-1 se fait habituellement dans les muqueuses oropharyngées et la latence dans les ganglions trigéminals et/ou les ganglions sensitifs rachidiens cervicaux. Le HSV-2 se réplique préférentiellement dans la muqueuse génitale et établit la latence dans les ganglions sensitifs sacrés et lombaires. Les infections primaires sont asymptomatiques dans plus de 90% des cas (Whitley *et al.*, 1998, Liesegang 1989). Cependant, dans le cas où celles-ci se manifestent, elles sont habituellement plus sévères que les infections récurrentes et un traitement antiviral adéquat peut dans certains cas limiter la douleur associée, la durée de l'infection ainsi que l'excrétion virale (Amir *et al.*, 1997). Exceptionnellement, le virus pourrait se disséminer par voie sanguine durant l'infection primaire et causer une infection systémique, avec méningite aseptique et/ou infection de plusieurs organes. Cette situation peut être rencontrée chez les grands immunodéprimés.

Lors des réactivations périodiques, les virus sont transportés par voie antérograde le long des neurones sensitifs jusqu'à la peau ou aux muqueuses et entraînent un épisode d'infection récurrente, même en présence d'une immunité humorale et cellulaire intacte. L'excrétion muco-cutanée de HSV-1 lors des infections récurrentes peut aussi être asymptomatique ou se manifester sous plusieurs formes. Dans les deux cas cependant, une transmission virale à un nouvel hôte est possible (Mertz *et al.*, 1988, Wald *et al.*, 2000). Le site des infections récurrentes tend à se confiner au dermatome infecté lors de l'infection primaire, mais il peut également y avoir des lésions à des sites plus distants. Ce ci explique que l'herpès labial est la plus fréquente des récurrences herpétiques, touchant 15% de la population occidentale, alors que les atteintes oculaires ne touchent que 0,15% des individus, et que les encéphalites sont exceptionnelles. La fréquence des récurrences peut dépendre du type de HSV-1 et du site anatomique de l'infection (Lafferty *et al.*, 1987). Il semble également que la sévérité de l'infection primaire, caractérisée par le nombre, l'étendue, et la dimension des lésions, augmente les probabilités de récurrences futures.

I.7.4 Épidémiologie

L'infection au HSV-1 est omniprésente dans la population mondiale et elle atteint toutes les classes sociales. Les manifestations cutanées causées par l'infection au HSV-1 sont décrites depuis la Grèce ancienne (Roizman & Whitley, 2001), alors que l'infection génitale au HSV-2 a été considérée comme étant un sujet d'intérêt public seulement vers les années 1970. La manifestation la plus courante de l'infection au HSV-1 est encore la forme oro-labiale. Néanmoins, on dénote une augmentation marquée de l'incidence des infections génitales causées par HSV-1 (Cusini & Ghislanzoni, 2001, Kinghorn, 1999). Les données épidémiologiques sont généralement obtenues par la détection spécifique des anticorps contre le HSV-1 ou -2 dans le sang des individus (sérologie). Les méthodes sérologiques utilisées et les cohortes de patients évalués peuvent grandement influencer les résultats des études et il devient parfois difficile de comparer les données de celles-ci. Cependant, certaines tendances sont nettes, et seront principalement discutées dans cette section.

La prévalence de l'infection au HSV-1 est très élevée dans pratiquement tous les pays, et augmente avec l'âge pour atteindre plus de 90% chez les soixantenaires. De façon générale, la prévalence du HSV-1 est toujours plus élevée que celle du HSV-2, un fait probablement dû à leur mode de transmission différent. En effet, plus de 85% de la population mondiale est séropositive pour le HSV-1 (Spruance *et al.*, 1996). L'incidence de l'infection au HSV-1 varie

selon la localisation géographique, la classe socio-économique, et l'âge des individus. Dans les pays en voie de développement ainsi que dans les classes sociales économiquement plus défavorisées, les individus sont infectés primairement par HSV-1 plus tôt dans l'enfance comparativement aux individus des classes moyennes et/ou vivantes dans les pays développés. En effet, vers l'âge de 5 ans, plus du tiers des enfants issus de milieux sociaux moins favorisés sont séroconvertis comparativement à 20% des enfants de même âge, mais de statut plus privilégié (Nahmias *et al.*, 1990). À l'adolescence, 70 à 80% des individus défavorisés sont séropositifs tandis que ceux des classes moyennes ou élevées deviendront séropositifs dans 40 à 60% des cas durant la 2^{ème} et 3^{ème} décade de leur vie. Dans la plupart des études, la prévalence du HSV-1 augmente de façon constante avec l'âge et atteint un plateau après l'âge de 30 ans. Au-delà de la quarantaine, les différences de prévalence du HSV-1 entre les divers pays et classes sociales sont moins évidentes (Whitley *et al.*, 1998).

L'herpès génital est l'une des maladies sexuellement transmissibles (MST) les plus communes et est la première cause d'ulcérations génitales d'origine infectieuse à travers le monde (Fleming *et al.*, 1997). La prévalence du HSV-2 est également très variable et dépend de plusieurs facteurs incluant le pays et la région de résidence, les sous-groupes de population, le sexe et l'âge des individus. Puisque le HSV-2 est habituellement contracté durant des contacts sexuels, il est logique que les anticorps contre celui-ci soient rarement détectés avant l'adolescence. En effet, la prévalence du HSV-2 passe de niveaux négligeables chez les enfants en dessous de 12 ans à des taux de 80% chez des adultes à haut risque d'acquisition de l'infection. De façon générale, la prévalence du HSV-2 est plus élevée chez la femme que chez l'homme. Cela est probablement dû au fait que, comme pour les autres MST, la transmission du HSV-2 semble plus efficace de l'homme à la femme que l'inverse (Mertz *et al.*., 1992).

De façon générale, la prévalence du HSV-2 est constamment plus élevée chez les individus séropositifs pour le VIH, les prostitué(e)s et les individus présentant d'autres MST (Smith & Robinson, 2002). Aux États-Unis, des études ont révélé que parmi les individus séropositifs pour le VIH, 68% des homosexuels de sexe masculin (Stamm *et al.*., 1988), 63% des hommes hétérosexuels et 78% des femmes présentaient également des anticorps contre le HSV-2 (Hook *et al.*., 1992). En raison de sa nature ulcérate, l'infection génitale au HSV-2 semble faciliter la transmission du VIH (Telzak *et al.*, 1993, Wasserheit, 1992). De façon

intéressante, les homosexuels de sexe masculin semblent également plus infectés par le HSV-2 que les hétérosexuels du même sexe (Nahmias *et al.*, 1990).

La séroprévalence du HSV-2 varie énormément selon la localisation géographique. Certaines régions d'Afrique et d'Amérique comportent les séroprévalences les plus élevées, et la plus élevée au monde a été notée dans la région Sub-Saharienne où plus de 80% des individus testés étaient séropositifs pour le HSV-2 (Smith & Robinson, 2002). En Europe de l'Ouest et du Sud, la prévalence du HSV-2 est habituellement moins élevée qu'en Europe du Nord et qu'en Amérique du Nord. Finalement, l'Asie et l'Australie sont les régions les moins touchées et deux études menées au Japon ont révélé une prévalence inférieure à 7% dans tous les groupes d'âge observés (Hashido *et al.*, 1999, Hashido *et al.*, 1998). Ces données ont probablement varié depuis quelques années (ou varieront dans un avenir proche) en raison des changements économiques majeurs et la généralisation des échanges entre le monde occidental et le monde oriental.

Une étude nationale effectuée aux États-Unis (NHANES III) et couvrant les années 1988 à 1994, a permis de relever plusieurs facteurs prédictifs de l'infection au HSV-2 (Fleming *et al.*, 1997). Le sexe féminin, la race noire ou Mexicaine-Américaine, un âge élevé, un faible niveau d'éducation, la pauvreté, la consommation de cocaïne, et un nombre élevé de partenaires sexuels au cours d'une vie ont été les principaux facteurs dévoilés. En effet, une différence marquée entre la prévalence de l'infection chez les individus de race noire (45,5%) comparativement à ceux de race blanche (17,6%) a été observée (Fleming *et al.*, 1997). De plus, 25,6% des femmes comparativement à 17,8% des hommes étaient séropositifs pour le HSV-2 et la séroprévalence des individus en général, augmentait de façon significative avec l'âge pour se stabiliser vers l'âge de 40 ans où 24% à 28% des individus étaient infectés par le virus.

Le facteur le plus prédictif de l'acquisition de l'infection au HSV-2 demeure le nombre de partenaires sexuels au cours d'une vie (Fleming *et al.*, 1997). Prenons par exemple le cas d'une femme hétérosexuelle vivant aux États-Unis avec un seul partenaire ; la probabilité qu'elle contracte une infection au HSV-2 est inférieure à 10%. Avec le nombre de partenaires qui augmente de 2 à 10, de 11 à 50 ou à plus de 50, cette probabilité grimpe à 40%, 62% et plus de 80%, respectivement. L'étude nationale a également révélé que la prévalence du HSV-2 a subi une augmentation de 30% depuis les années 1976-1980

(NHANES II) (Fleming *et al.*, 1997, Johnson *et al.*., 1989). Plusieurs autres régions du monde semblent également suivre cette tendance, mais les données ne sont pas encore complètes (Lafferty, 2002). De plus, on constate que l'incidence de l'infection génitale causée par le HSV-1 chez la femme est en croissance dans plusieurs régions (Christie *et al.*., 1997, Edwards & White, 1994, Langenberg *et al.*., 1999, Carney *et al.*., 1993), contribuant jusqu'à 30-40% des nouveaux cas d'herpès génital. L'amélioration des conditions socio-économiques de ces pays entraînerait possiblement une séroconversion des individus pour le HSV-1 plus tardive, ce qui favoriserait la possibilité d'acquérir une infection génitale par ce type de virus. Les pratiques sexuelles orogénitales aujourd'hui plus fréquentes et mieux acceptées ont également été mises en causes.

Finalement, l'herpès néonatal affecte toutes les catégories de populations, mais le taux le plus élevé se situe aux États-Unis où l'incidence est de 1 cas par 2500-8800 naissances (Schomogyi *et al.*, 1998). Aux États-Unis, environ 75% des cas sont causés par le HSV-2, et 25% par le HSV-1. De façon contradictoire, au Royaume Uni, le HSV-1 est responsable de 50% des cas d'herpès néonatal (Tookey & Peckham, 1996) et compte pour plus de 60% des cas rapportés au Japon. Aux États-Unis, 5% des femmes en âge de procréer ont déjà présenté une histoire d'herpès génital (Prober *et al.*, 1992), et il est estimé qu'un million des grossesses annuelles se font chez des femmes séropositives pour le HSV-2.

1.8 Réponse immunitaire

Le cours naturel de l'infection au HSV-1 est influencé à la fois par la réponse immunitaire non spécifique (innée) et spécifique (acquise) de l'hôte infecté (Whitley, 2001). L'importance relative des diverses composantes de la réponse immune varie également selon que l'infection est initiale ou récurrente. Suite à une infection initiale, il y a rapidement apparition d'une réponse inflammatoire non spécifique incluant l'interféron (IFNs), les macrophages activés et les cellules tueuses (NK). Cette première ligne de défense limite l'extension de l'infection en inhibant significativement la réplication virale (Stanberry *et al.*, 2000). D'un autre côté, les anticorps neutralisants, les lymphocytes T spécifiques sécrétant les cytokines antivirales (CD4+), et/ou activées pour tuer les cellules infectées (CD8+) (réponse immunitaire spécifique) prennent plusieurs jours à se développer. D'un modèle animal à l'autre, cette réponse peut varier significativement (Whitley *et al.*, 1998). Par exemple, chez la souris, les lymphocytes T cytotoxiques et les immunoglobulines IgM et IgG spécifiques

apparaissent de 4 à 6 jours après le début de l'infection, alors que chez l'homme, l'immunité spécifique se développe 7 à 10 jours plus tard. De façon classique, les anticorps IgM apparaissent de façon transitoire, suivis des IgG et des IgA qui persistent sur une plus longue période (Whitley & Roizman, 2001). Les anticorps neutralisants sont dirigés contre les principales glycoprotéines de l'enveloppe (gB, gD, gH/L), alors que les lymphocytes T peuvent être dirigés contre une panoplie de protéines virales (Stanberry *et al.*, 2000). Les protéines et leurs épitopes antigéniques sont généralement bien conservés entre le HSV-1 et -2 mais des anticorps dirigés contre des épitopes type-spécifique ont été identifiés pour les glycoprotéines gB, gD, gC, et plus spécialement gG. Des épitopes T type-spécifiques ont également été identifiés chez ces glycoprotéines ainsi que chez la protéine virale VP16 (Koelle *et al.*, 1998a).

Face à un système de défense immunitaire efficace, le HSV-1 réussit tout de même à se réactiver de sa latence et à entraîner une infection récurrente. Ce phénomène demeure toujours aussi énigmatique malgré de nombreuses recherches effectuées sur ce sujet. Cependant, les infections récurrentes sont généralement moins sévères et plus brèves que les infections initiales. En effet, lors d'infections récurrentes, les mécanismes de l'immunité spécifique sont activés beaucoup plus rapidement (Cunningham & Mikloska, 2001). En fonction des individus, la sévérité de l'infection et la fréquence des récurrences peuvent être significativement différentes et reflètent peut-être la sélection préférentielle de l'une des composantes de la défense immunitaire. Les études effectuées chez les animaux suggèrent que l'immunité humorale et cellulaire sont toutes deux importantes dans le contrôle de l'infection initiale et des récurrences (Stanberry *et al.*, 2000).

La réponse humorale semble efficace pour réduire le titre viral au site d'infection et dans les tissus nerveux durant l'infection initiale. En effet, le transfert passif d'anticorps spécifiques du HSV-1 semble protéger l'animal contre une infection disséminée ou une infection du système nerveux (Bravo *et al.*, 1996, Mitchell & Stevens, 1996, Morrison & Knipe, 1997, Sanna *et al.*, 1996). Chez l'humain, le transfert d'anticorps de la mère au fœtus joue également un rôle important dans la protection contre l'herpès néonatal, mais ne semble pas totalement efficace (Kohl *et al.*, 1989, Sullender *et al.*, 1987). Par ailleurs, l'immunité humorale est moins efficace dans la protection contre les récurrences ou contre l'infection exogène (Whitley *et al.*, 1998). Le HSV-1 a développé plusieurs stratégies afin de déjouer cette composante de la réponse immune (Favoreel *et al.*, 2000). Par exemple, les

glycoprotéines gE et gI de l'enveloppe virale forment un complexe pouvant lier la fraction Fc des immunoglobulines, bloquant ainsi l'activation de la cascade du complément. De plus, la gC lie plusieurs composantes du complément. L'établissement de la latence virale dans les cellules du système nerveux, comprenant très peu de cellules impliquées dans la défense immunitaire, de même que l'absence d'expression de protéines virales pendant la phase d'infection latente demeure l'un des mécanismes les plus efficaces d'évasion du système immunitaire.

Plusieurs études suggèrent que la réponse à médiation cellulaire joue un rôle central dans le contrôle de la sévérité de l'infection et du nombre de récurrences (Cunningham & Merigan, 1983, Posavad *et al.*, 1996,). En effet, les infections mucocutanées sont beaucoup plus sévères chez les patients ayant une immunité à médiation cellulaire défective (Posavad *et al.*, 1996, Siegel & McCracken, 1981) comparativement aux patients agammaglobulinémiques. Plusieurs études effectuées chez l'humain incriminent les lymphocytes T CD8⁺ et T CD4⁺ dans la protection contre l'infection par HSV-1 (Jennings *et al.*, 1991, Manickan & Rouse, 1995, Posavad *et al.*, 1996). Lequel de ces deux types de cellules est le plus important pour contrôler les infections récurrentes demeure un sujet bien débattu surtout lorsqu'on considère le développement d'un vaccin efficace contre l'infection au HSV-1 (Cunningham & Mikloska, 2001). Une étude a rapporté que des individus séropositifs pour le VIH et souffrant d'infections génitales au HSV-1 sévères, présentaient des niveaux de lymphocytes T CD8⁺ précurseurs plus faibles que chez les patients VIH positifs avec des infections génitales herpétiques modérées (Posavad *et al.*, 1997). Une autre étude a démontré que la présence de lymphocytes T CD8⁺ locaux est associée à la clairance du HSV-1 des lésions herpétiques cutanées (Koelle *et al.*, 1998b). De façon intéressante, un des mécanismes d'évasion du HSV-1 (Favoreel *et al.*, 2000) contre les défenses immunitaires de l'hôte est en fait d'inhiber l'efficacité des lymphocytes T CD8⁺ cytotoxique en réponse à l'infection virale. La protéine virale ICP47 exprimée précocement dans l'infection, se lie à la protéine cellulaire TAP (*transporter associated with antigen processing*) qui permet la translocation de peptides dans le RE et leur association au complexe majeur d'histocompatibilité (CMH) de classe I. Ainsi ICP47 inhibe le transport et la présentation des antigènes associés au CMH de classe I. D'un autre côté, pour contrecarrer cet effet, les cellules T CD4⁺ sécrètent de l'INF- γ qui restaure partiellement l'expression du CMH de classe I par une stimulation de la production de la protéine TAP (Koelle *et al.*, 1993, Mikloska *et al.*, 1998, Posavad *et al.*, 1996). Des modèles animaux montrent que les

lymphocytes CD8 sont directement impliqués dans le contrôle des réactivations herpétiques, leur déplétion entraînant une sensibilité particulière des animaux aux stimuli de la réactivation du virus.

De façon générale, les lymphocytes T CD4⁺ peuvent sécréter des cytokines de la réponse de type Th1 (spécialement l'IFN- γ ainsi que le TNF- α et l'interleukine (IL)-2 et -12) ou de type Th2 (IL-4, IL-5, et IL-10). Au sein d'une population de lymphocytes, les deux types de réponse peuvent cohabiter. La réponse de type Th2 mène à la sécrétion d'anticorps neutralisants par les cellules B, alors que la réponse de type Th1 entraîne la production de cytokines antivirales et à l'activation des lymphocytes T CD8⁺ cytotoxiques. La réponse de type Th1 semble prédominante pour le contrôle de l'infection par HSV-1, alors que la réponse de type Th2 préviendrait la transmission du virus d'une cellule à l'autre (Cunningham & Mikloska, 2001).

Des études comparatives entre des individus ayant ou non des infections récurrentes ont démontré que des altérations dans la production des cytokines de la réponse de type Th1 sont corrélées avec le développement des infections récurrentes (Manickan & Rouse, 1995, Posavad *et al.*, 1998, Spruance *et al.*, 1995a). En effet, l'IFN- γ semble critique pour conférer une protection contre l'infection au HSV-1. Certains travaux ont permis de démontrer que les cellules mononucléées du sang périphérique (PBMC) prélevées chez des patients subissant des épisodes d'herpès labial fréquents ont une production d'IFN- γ inexistante en culture (Green *et al.*, 1985, Kuo & Lin, 1990, Sheridan *et al.*, 1987). De plus, ces PBMC produisent des titres élevés d'anticorps contre le HSV-1. D'un autre côté, les PBMC prélevés chez des patients séropositifs pour le HSV-1 mais sans histoires d'infections récurrentes, présentent des concentrations plus élevées d'IFN- γ et d'IL-2 et des titres plus faibles d'anticorps contre le HSV-1 (Spruance *et al.*, 1995b). Une meilleure compréhension de la réponse immunitaire spécifique dirigée contre l'infection causée par le HSV-1 permettra de prédire plus efficacement le type de réponse qui devrait être favorisé pour le développement de vaccins efficaces.

1.9 La protéine virale ICP0

La protéine ICP0 (infected cell protein 0) est une des protéines virales les plus étudiées. Initialement appelées Vmw110, ICP0 est une des cinq protéines très précoces (IE) synthétisées par le virus. (Everett 2000). ICP0 est le produit d'un gène de plus de 3Kb appelé

RL2 qui se situe dans les régions répétées "b" du segment UL du génome du HSV-1. La transcription du gène conduit à la formation d'un pré ARN messager composé de 3 exons et de 2 introns, qui est coiffé et poly-adenylé. *RL2* fait partie du petit groupe des 5 gènes épissés parmi les 84 gènes identifiés dans le génome viral. Quatre d'entre eux conduisent à la formation de protéines (ICP0, ICP22, ICP47, UL5) et le cinquième aboutit à la formation du LAT.

Après maturation, l'ARNm est exporté et traduit pour donner naissance à la protéine ICP0. Il s'agit d'un transactivateur ubiquitaire capable de stimuler l'expression des trois classes de gènes herpétiques (IE, E et L), ainsi que celle de nombreux autres gènes cellulaires. Toutefois, ICP0 ne se fixe pas à l'ADN et n'est donc pas un facteur transcriptionnel à proprement parler (Cheung et al 1997). La localisation de la protéine ICP0 évolue au cours de l'infection lytique du virus. C'est une protéine exclusivement nucléaire en début d'infection, qui devient cytoplasmique après 8h d'infection avant de retrouver une localisation nucléaire à des temps plus tardifs (Lopez et al., 2001 ; Bruni R et al., 1999). La protéine est composée de 775 acides aminés qui lui permettent de former plusieurs domaines fonctionnels (Figure 11).



D'après P .Lomonte 2005

Figure 11: Représentation schématique de la protéine ICP0 et de ses domaines d'activité ubiquitine ligase. Les rectangles blancs correspondent aux différentes régions de la protéine code par les 3 exons du gène *RL2*. Les deux régions minimales responsables de l'activité ubiquitine ligase de la protéine sont représentées par des rectangles orange. Le domaine RING Finger est représenté par un rectangle vert. la numérotation correspond à la position des différents acides aminés encadrant chaque domaine

La région N-terminale est responsable de l'activation de la transcription de plusieurs gènes lytiques, elle contient un domaine en doigt de zinc essentiel pour la majeure partie des activités biologiques de la protéine dont la plus importante est son activité E3 ubiquitine ligase. La région C-terminale contient un site de multimérisation important pour sa localisation dans les corps PML (*promyelocytic leukaemia protein*,).

L'activité E3 ubiquitine-ligase de la protéine ICP0 est la plus étudiée. L'ubiquitination désigne la fixation, covalente, d'une ou plusieurs ubiquitines sur une protéine cible. Cette modification post-traductionnelle nécessite l'intervention successive de 3 enzymes, E1, E2 et E3-ubiquitine ligases et dépend de l'ATP. Elle a pour principale fonction la reconnaissance puis la destruction de la protéine ainsi marquée, par le complexe protéolytique du protéasome. Il faut la liaison de 4 ubiquitines (polyubiquitination) pour qu'une protéine soit dégradée. Après dégradation de la protéine, l'ubiquitine est alors libérée de son substrat et peut être réutilisée.

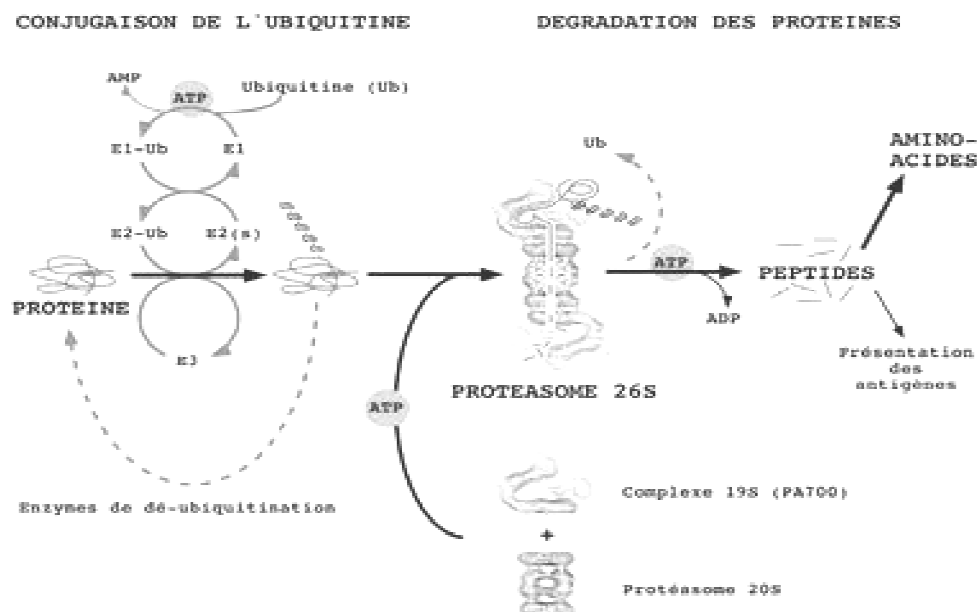


Figure 12: Représentation schématique des étapes de fonctionnement du système Ubiquitine-Protéasome. Ce système multi-enzymatique fonctionne en deux grandes étapes : marquage covalent (ubiquitylation) des protéines à dégrader par conjugaison d'une chaîne de poly-ubiquitine (polyUb), et dégradation du substrat ubiquitylé par le protéasome 26S. L'ubiquitination débute par la carboxylation d'une molécule d'ubiquitine sur l'enzyme E1, l'ubiquitine sera alors transférée sur l'enzyme de conjugaison E2 qui va interagir avec l'enzyme E3 (ICP0). Le complexe E2-E3 va alors transférer plusieurs molécules d'ubiquitine sur une protéine qui sera ensuite dégradée par le protéasome.

Comme les totalité des E3 ubiquitine ligase, ICP0 fonctionne de concert avec des protéines E2 ubiquitine conjugase, comme les protéines de la famille Ubch5 et Ubch6 (Boutell, 2002, Gu et Roizman, 2003). L'étude *in vitro* de l'activité enzymatique d'ubiquitination d'ICP0, a permis de mettre en avant plusieurs cibles dont ICP0 elle-même (Canning et al., 2004), USP7, protéine impliquée dans la coupure de la liaison entre un groupement ubiquitine et une protéine associée (Boutell et al., 2005), RNF8 qui est une ubiquitine ligase (Lilley et al., 2010), et enfin la protéine P53 qui est un facteur de

transcription régulant certaines fonctions cellulaires importantes comme la mitose ou la mort programmée (Boutell et al 2003).

Deux des cibles les plus remarquables visées par la protéine ICP0 sont les protéines PML (*promyelocytic leukaemia protein*) et SP100 qui sont des composants majeurs des corps PML. Les corps PML sont des structures sphériques de 0.1 à 1.0µm de diamètre, présents dans la plupart des lignées cellulaires et dans plusieurs tissus. Ils appartiennent à la matrice nucléaire, une superstructure du noyau mal définie dont on pense qu'elle régule plusieurs fonctions nucléaires incluant la réplication de l'ADN, la transcription ou l'« epigenetic silencing » . La protéine PML joue un rôle clef dans l'organisation de ces domaines qui recrutent un grand nombre de protéines, qui ont la plupart du temps une fonction de sumoylation. Les corps nucléaires PML sont régulés par le stress cellulaire comme, par exemple, les infections virales (Everett 2006), les dommages à l'ADN, la transformation (Koken et al. 1995) et le stress oxydatif (Yamada et al. 2001). De plus, la transcription des gènes codant pour PML et plusieurs de ses protéines partenaires est augmentée en présence d'interférons (Chelbi-alix, et al 1995). Ceci suggère que la dégradation des corps PML par la protéine ICP0 pourrait être une stratégie de défense mise en place par le virus afin de contrecarrer la réponse immunitaire de la cellule dans les premières étapes de l'infection.

L'autre cible majeure de l'activité d'ICP0 est la dégradation de certaines protéines centromériques (CENPS) CENP-C, CENP-B et CENP-A (Everett et al 1999, Lomonte et al 2001, Lomonte et al 2007). CENP-A est une protéine de 17 kDa localisée dans la région du kinétochore. Elle joue un rôle spécifique au niveau du centromère en favorisant la compaction de la chromatine. CENP-B est une protéine de 80 kDa qui joue aussi un rôle dans l'organisation de l'ADN centromérique. CENP-C est une protéine de 140 kDa qui intervient dans la relation entre le centromère et le fuseau mitotique.

En détruisant les protéines du centromère, ICP0 induit le blocage du cycle cellulaire. Les cellules seront alors incapables de finir la mitose (Lomonte et al 1999,). Le rôle exact de la dégradation des protéines centromériques par la protéine ICP0 n'est pas encore très bien défini, mais il semble que cette fonction soit utilisée au moment de la réactivation par le virus pour dégager son génome des régions centromériques, qui sont connues pour être associées à des mécanismes de répression de la transcription.

Chapitre II :

Les transcrits associés a latence : LAT

(Latent Associated Transcript)

II.1 Introduction

Depuis leur découverte à la fin des années 80 par hybridation *in situ*, le rôle exact des LAT n'a jamais pu être définitivement établi. Les LAT sont les seuls transcrits viraux abondamment détectés dans les neurones durant l'infection latente et, de ce fait, sont considérés comme les marqueurs de l'état de latence du HSV-1 (Cohrs *et al.*, 2000).

Les LAT sont transcrits à partir de séquences présentes dans les régions répétées "b" du segment UL du génome de HSV-1. Le groupe des LATs regroupe en réalité 4 types d'ARN hautement conservés; un long transcrit instable (8,3 kb) et trois transcrits très stables (2,0, 1,5, et 1,45 kb)(figure 13). Ces derniers s'accumulent abondamment uniquement dans les neurones infectés de façon latente (Krause *et al.* 1988, Rock *et al.* 1987, Stevens *et al.* 1987).

Le LAT de 8,3kb est le premier transcrit dans les neurones infectés. Il est coiffé et polyadénylé mais très instable, donc retrouvé en petites quantités. Pour cette raison, il est souvent appelé dans les publications « LAT mineur ». Ce transcrit est ensuite épissé pour donner naissance à un second transcrit de 2 kb, correspondant à un intron, qui s'accumule dans les neurones infectés de façon latente. Celui-ci est à son tour épissé pour former deux autres LAT de 1,5 kb et 1,45 kb, en fonction de sites donneurs et accepteurs d'épissage utilisés. Les LATs de 2 kb, 1,5 kb, et 1,45 kb sont retrouvés en quantité importante, jusqu'à 40000 copies par neurone infecté (Maillet *et al.* 2006). Pour cette raison, ils sont souvent dénommés « LAT majeurs ». Si le LAT de 2kb est déjà présent dès les stades précoces de l'infection en modèle animal et même dans les infections de cellules en culture, les LAT 1,5kb et 1,45 kb ne sont retrouvés que dans le cadre d'une réelle infection latente dans les modèles d'infection *in vivo*. La production des LAT 1,5kb et 1,45 kb semble donc dépendante de facteurs cellulaires propres aux neurones *in vivo*.

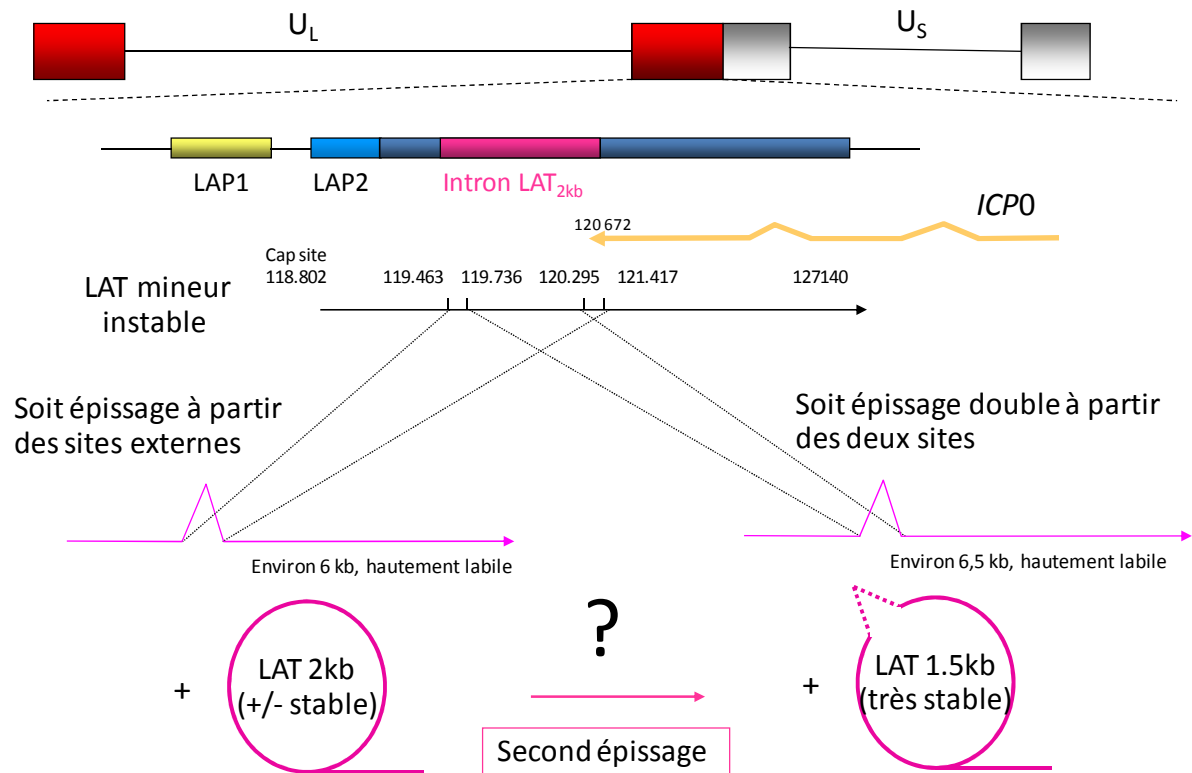


Figure 13: Représentation schématique de l'organisation du locus des LAT sur le génome de l'herpès simplex de type 1. Les LAT sont transcrits à partir de séquences présentes dans les régions répétées "b" du segment UL du génome de HSV-1. Le groupe des LATs regroupe en réalité 4 types d'ARN hautement conservés; le LAT mineur (8,3 kb) qui est épissé pour donner les LAT majeur (2,0, 1,5, et 1,45 kb). Le gène des LATs est anti-sens est partiellement chevauchant du gène RL2.

II.2 Organisation et rôle du locus des LAT.

Le promoteur des LAT est relativement plus complexe que la majorité des autres promoteurs de virus HSV. Les LAT sont sous la dépendance d'une région promotrice d'environ 200 pb appelée LAP1, indispensable à leur expression (figure 14). Il contient des sites de fixation pour plusieurs facteurs transcriptionnels cellulaires ou viraux parmi lesquels Sp1, YY1, Ap2, CRE, USF, ICP4. Cependant, la seule présence de cette région promotrice ne suffit pas à expliquer le maintien d'une expression à long terme des LATs pendant la latence. Il fut ainsi découvert une seconde région régulatrice de 650 pb, appelée LTE (*long term expression*), située en aval de LAP1. Son mécanisme d'action n'a pas encore été clairement défini, mais elle semble capable de maintenir l'activité du promoteur LAP1 au cours de l'infection latente, faisant de celle-ci un élément régulateur de type « enhancer » du promoteur LAP1 (Lokensgard 1994, Margolis1997).



Figure 14 : Représentation schématique de l'organisation du promoteur des LAT.

Le locus abritant les LAT est certainement la région transcriptionnelle la plus complexe du génome de HSV. Ainsi un autre promoteur, appelé LAP2, débutant juste après le nucléotide +1 de transcription, et finissant dans la partie 5' de l'exon 1 du LAT 2kb a pu être mis en évidence (Goins 1994, Maggioncalda 1994, Nicosia 1993). Tout comme LAP1, LAP2 contient des sites de fixation pour plusieurs facteurs transcriptionnels cellulaires dont Sp1, Ap2, E2F. De plus, il inclut aussi une région très riche en G+C, et une région riche en C+T, caractéristiques des promoteurs des gènes de ménage cellulaire. Contrairement à LAP1, LAP2 ne contient pas de TATA BOX. Il semble régulé négativement par la protéine ICP4. Le transcrit exact codé par ce promoteur n'a pas été clairement identifié et son rôle exact est encore à ce jour inconnu.

Dans le but d'étudier l'activité du promoteur des LAT et plus particulièrement de LAP1 en absence de facteurs viraux, une lignée de souris transgéniques a été établie (Gussow AM et al 2006). Celles-ci possèdent un génome dans lequel on a inséré une séquence correspondant à celle du LAT de 2kb sous la dépendance de sa région promotrice. Cette étude a montré que ce promoteur est potentiellement actif dans de nombreux types tissulaires, mais que seul les neurones sensitifs sont capables de maintenir de façon stable une expression des LAT 2kb au cours du temps. Cela conforte l'idée que l'activité promotrice de LAP1 dépend du type cellulaire et ne nécessite pas la présence d'autres facteurs viraux (Gussow AM et al 2006). Ainsi, la latence étant exclusivement détectée dans des cellules neuronales, on peut penser que des facteurs transcriptionnels neuronaux sont responsables de la surexpression du LAT.

La relation entre le virus et son hôte dépend de façon étroite du LAT, puisque l'échange des loci du LAT entre les types 1 et 2 de HSV inverse les sites de latence des virus ainsi génétiquement modifiés. De fait, un virus HSV-1 transformé avec la région LAT d'un virus HSV-2 établit sa latence dans les neurones de la chaîne ganglionnaire sacrée avec la même efficacité qu'un virus HSV-2 sauvage. Inversement, un virus HSV-2 modifié avec le locus LAT de type HSV-1 établit sa latence dans les neurones sensitifs du ganglion trijumeau

avec la même efficacité qu'un HSV-1 sauvage (Bertke AS et al. ? 2001). Il apparaît ainsi que les loci codant les transcrits LATs influencent nettement la capacité du virus à établir une infection latente efficacement dans certains types neuronaux sensitifs (trigémiques pour HSV-1 et sacrés pour HSV-2).

Au cours des dix dernières années, d'autres transcrits ont été attribués au locus codant pour les LAT. Un petit transcrit appelé UOL (upstream of LAT) est produit à partir d'une région située immédiatement en amont du promoteur LAP1, sur le même brin que le LAT mineur (Naito et al 2005). Dans la même région, un autre transcrit de 700 pb a été découvert. Il a été montré comme débutant en amont de LAP1 (Zhu et al 1999)(Ghiasi Het al 2002).

Il fut également découvert en 2002 par Perng (Perng et al 2002) des transcrits anti-sens du LAT mineur. Au nombre de 3, ceux-ci sont nommés AL-LAT (*antisens to LAT*). Le premier possède un promoteur chevauchant en partie le promoteur LAP2 et est transcrit à partir d'une région dans LAP1, les deux autres appelés AL2 et AL3 sont transcrits à partir d'une région interne au LAT de 2kb.

Il est important de noter que seuls les transcrits LAT (majeur et mineur) issus du promoteur LAP1 sont exprimés de façon abondante durant la latence du virus alors que tous les autres transcrits sont exprimés uniquement pendant la phase lytique de l'infection.

A la fin des années 90, deux transcrits correspondant à l'extrémité 3' du LAT de 8,3 kb contenant des ORF codant pour des protéines appelées P et O furent découverts. Le premier d'entre eux correspondant à l'ORF P, code pour une protéine de 233 aa dans le cas de la souche *syn17* et pour une protéine de 248 aa dans le cas de la souche *KOS*. Cette différence dans la taille des deux protéines est due à la présence d'une répétition de triplets de nucléotides codant pour trois acides aminés : alanine, tyrosine et proline dans la souche *KOS*. Le transcrit codant pour l'ORF P est presque complètement antisens du gène RL1 (codant pour ICP4) à l'exception de 24 nucléotides. La partie 5' du transcrit est chevauchante du second ORF codant pour la protéine O. ORF O code pour une protéine de 300 aa dans le cas de la souche *syn17* alors que celui-ci est interrompu dans le cas des souches *KOS* et *F*, ce qui donne naissance à des produits de 160 ou 173 aa.

Il a été montré qu'au cours de la phase répliquative de l'infection, le promoteur de ces deux transcrits était réprimé par la protéine ICP4. De plus, durant la phase lytique de

l'infection, la production du transcrit du gène RL1, antisens du transcrit contenant l'ORF P, semble interférer avec la synthèse des transcrits contenant l'ORF P (Randall et al 1997a ; Randall et al 1997b).

Des études menées avec des mutants, exprimant de manière constitutive chacune des protéines O et P, ont montré que, durant le cycle infectieux, l'expression des protéines ICP0 et ICP22 était retardée mais que le cycle viral n'était pas affecté à en juger par les titres viraux obtenus avec ces mutants. Des expériences de double hybride menées sur P ont montré que cette protéine interagissait avec des protéines du spliceosome. Ceci pourrait expliquer son action sur les protéines ICP0 et ICP22, qui sont produites à partir d'ARNm maturés par épissage (Randall et al 2000). Dans une étude *in-vivo*, il fut montré par les mêmes auteurs, que l'absence de ces deux ORF ne perturbait pas la mise en place et le maintien de la latence (Randall et al 2000). La protéine O a été moins étudiée en raison de son absence de conservation au sein des souches de HSV-1. Cependant, il est intéressant de noter qu'une étude a montré que le promoteur des LAT influence son expression.

Plusieurs mutants de la région LAT ont été construits dans le but d'identifier leur rôle exact dans le cycle biologique du virus. Une des premières difficultés, rencontrée par toutes les équipes, est la présence dans le locus du LAT de trois gènes codant pour des protéines indispensables pour la réplication du virus : ICP0, ICP4 et ICP34.5. Chacun de ses gènes est, au moins partiellement, antisens du gène des LAT, rendant impossible la construction d'un virus délesté de l'ensemble du locus LAT.

Une autre difficulté rendant l'étude des LAT fastidieuse est qu'il n'existe pas à ce jour de modèle de latence *in-vitro*. De plus, le choix du modèle animal ainsi que la souche virale utilisée modifie les observations phénotypiques pour un même type de modification génétique. Il est néanmoins à ce jour admis que la souris reste un modèle de choix pour l'étude de la latence, en raison des faibles fréquences de réactivation spontanée du virus dans ce modèle. Le lapin est quant à lui plutôt utilisé pour les études portant sur la réactivation et ses conséquences cliniques (évaluation des procédures thérapeutiques, par exemple).

A ce jour, pas moins de neuf phénotypes distincts sont attribués au LAT et plus particulièrement à la région comprenant le promoteur LAP1 et le LAT de 2kb. Ces deux éléments semblent intervenir dans (i) la neurovirulence, (ii) l'établissement de la latence, (iii) la nature des neurones infectés de façon latente, (iv) la neuro-protection, (v) les mécanismes

de réactivation et (vi) les interactions avec la réponse immunitaire. A l'échelle de la cellule, on leur attribue (vii) des fonctions anti-apoptotiques, (viii) une influence négative sur la transcription des gènes lytiques du virus et (ix) un rôle sur l'état de la chromatine du virus.

II.3 Influence des LAT sur la réactivation.

Historiquement, il s'agit du premier phénotype attribué au LAT par la communauté scientifique. Des études menées *in-vivo*, comparant des virus sauvages ou délétés pour différentes parties de la région 5' du gène (promoteur +/- début du gène) avaient montré un taux de réactivation diminué pour le virus mutant quelque soit le modèle animal utilisé (Bloom et al 1994 ; Hill et al, 1990 ; Leib et al 1989, Perng 1994). Cette réduction du taux de réactivation est particulièrement marquée dans un modèle murin après choc hyperthermique (Sawtell et al 1992) ou chez le lagomorphe après injection d'adrénaline (Bloom et al, 1994).

Depuis la découverte des LAT, plusieurs virus contenant des délétions de différentes parties des LAT ont été construits. Le type de délétion effectuée, le modèle animal utilisé et les principaux phénotypes observés sont résumés dans la figure 15.

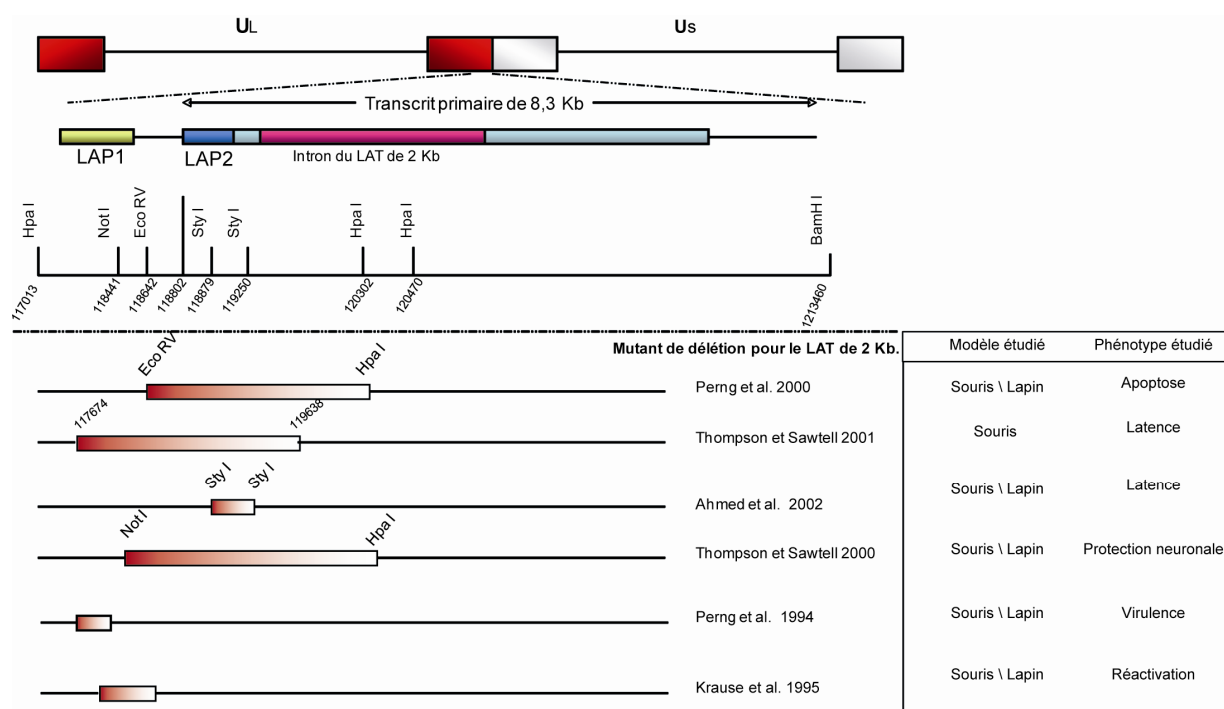


Figure 15: Représentation des différents mutants de construction pour l'étude du rôle des LAT. Représentation schématique de la région codant les LAT avec les principaux sites restriction utilisés pour la construction des différents mutants. Les zones en rouges représentent les différentes délétions de la région 5' du locus des LAT.

II.4 Influence des LAT sur l'établissement de la latence.

Les premières études menées par des techniques de Dot-Blot et de PCR semi quantitative n'avaient pas montré de différence de charge virale dans des ganglions infectés de façon latente entre des virus mutés dans le LAT et leur révertants (Sedarati *et al* 1989, Steiner *et al* ; 1989). Néanmoins, l'avancée des techniques a permis au cours du temps de mettre en évidence une réelle diminution de la capacité des virus mutants à établir de la latence. Avec sa méthode dite contextuelle qui permet d'analyser individuellement les neurones sensitifs du ganglion trigéminal, Sawtell a montré que l'inactivation des LATs réduit d'un facteur deux le nombre de neurones infectés de façon latente par rapport au virus sauvage (Sawtell *et al* 1997).

Ce premier résultat fut confirmé dans d'autres types d'études basés sur l'expression de gène rapporteur (Perng *et al* 2000, Thompson *et al*, 2001). Finalement le développement de la PCR quantitative a confirmé une réduction d'un facteur 4 du nombre de génomes présents dans des ganglions latents lorsque ceux-ci sont infectés par un virus délesté des LAT (Thompson *et al*, 2001). L'analyse globale des données sur le sujet a conduit à admettre que le LAT facilite l'établissement de la latence en agissant à la fois sur le nombre de particules virales présentes dans les sites de latence et le nombre de cellules infectées.

Ces résultats ont dans un premier temps poussé la communauté à s'interroger sur la capacité du virus à réactiver en absence de LAT. Si son absence diminue le nombre de particules virales présentes durant la latence, il semble alors logique d'observer moins de réactivation. La plupart des études utilisent des mutants dont la délétion inclut les régions promotrices et toute la partie 5' du gène jusqu'à la moitié de l'intron correspondant au LAT 2kb. Des études utilisant des virus porteurs de délétions plus petites n'ont pas retrouvé les phénotypes observés dans les études précédentes (O'Neil *et al* 2004). Cependant, l'analyse conjointe de toutes ces études montre qu'une séquence comportant les 1900 premières paires de base du LAT est essentielle pour obtenir des capacités normales d'entrée en latence et donc de réactivation.

II.5 Influence des LAT sur la neuro-virulence.

La plupart des délétions de tout ou partie de la région codant les LATs ont des effets sur la capacité du virus à infecter l'hôte (Perng *et al* 1999 & 2001). Certaines délétions procurent au virus une virulence accrue, conduisant à la mort de l'animal tandis que d'autres semblent diminuer le pouvoir pathogène du virus (Garber *et al* 1997). Il est important de noter qu'aucun des mutants de délétion obtenus à ce jour ne semble être affecté pour sa propagation en cellules en culture. Ceci suggère d'une part une absence de rôle des LAT durant le cycle lytique et d'autre part l'existence de facteurs présents exclusivement dans le neurone qui sont nécessaires à l'établissement de la latence.

II.6 Influence des LAT sur le lieu de la latence.

Comme mentionné plus haut, la séquence du locus LAT semble affecter le tropisme cellulaire dans lequel le virus peut entrer en latence. Il a été montré que lorsqu'on échange les séquences du LAT entre les types 1 et 2 de HSV, le virus mutant acquiert un tropisme correspondant au type de LAT acquis. Ainsi, une souche de type HSV-1 modifiée avec le LAT d'un HSV-2 entre plus efficacement en latence et se réactive plus volontiers dans les neurones de la chaîne ganglionnaire sacrée que dans les neurones du ganglion trigéminal, alors que c'est l'inverse pour un virus HSV-2 modifié avec le locus LAT de type 1. Plus récemment, en utilisant un modèle d'infection par scarification cornéenne (le détail des différents modes d'inoculation du virus sera donné dans le chapitre matériel et méthode), une équipe a montré que le locus LAT confère au virus la propriété d'établir une infection latente dans une sous population particulière de neurones sensitifs, appelée A5, dont la seule propriété identifiée à ce jour est d'exprimer à la surface l'épitope Galb1-4GlcNAc-R (Fenderson *et al* 1983 ; Yang *et al* ; 2000)(Imai *et al* 2009). La même expérience menée avec le virus HSV-2 a montré que les neurones du ganglion trigéminal infectés de façon latente expriment à leur surface l'épitope Galb α 1-Galb β 1-NAc-R, caractéristique d'une autre sous-population neuronale nommée KH10 (Margolis *et al* 2007). En culture cellulaire, des expériences d'échange des séquences LAT entre le type 1 et 2 de HSV ont montré que le tropisme pour les neurones A5 ou K10 était effectivement lié au type viral dont est issu le LAT. La nature du lien et les mécanismes biologiques en cause n'ont cependant pas été éclaircis à ce jour.

II.7 Influence des LAT sur la protection des cellules neuronales et leur effet anti- et pro-apoptotique.

Une des clefs de la mise en place de la latence est le maintien du réservoir de cellules infectées au cours de la phase aigüe du cycle lytique. Pour permettre au virus de basculer d'un cycle lytique vers un état de latence, il a été proposé que les LAT puissent moduler l'expression de la protéine ICP0, réduisant ainsi l'efficacité de l'ensemble de la cascade transcriptionnelle. Cette hypothèse est née de l'observation que le nombre de neurones infectés dans un ganglion trigéminal d'une souris infectée avec une souche virale déficiente pour l'expression des LAT (dite LAT-) est deux fois moindre que lors d'une infection par la souche virale sauvage correspondante (Thompson et Sawtell 2001). Ces résultats suggéraient que la présence de LAT bloquait le cycle lytique du virus et/ou facilitait l'entrée en latence grâce à un effet protecteur pour la survie des neurones infectés.

La publication de référence sur le rôle des LAT dans la survie neuronale repose sur l'analyse d'une souche LAT- obtenue par délétion des nucléotides -161 à +1667 du locus LAT, englobant ainsi le +1 du début site de transcription du LAT mineur (8,3kb) et une partie de la région correspondant au LAT 2kb. Les auteurs ont observé par TUNEL une quantité accrue de cellules apoptotiques 7 jours après l'infection par la souche mutante par rapport au virus sauvage et au révertant (Perng *et al* 2002). Ces résultats furent confirmés par quantification du nombre de neurones exprimant la protéine PARP clivée (caractéristique de l'apoptose), de l'ordre de 25% de la population neuronale infectée par la souche mutante contre 8% lors d'une infection par le virus sauvage (Perng *et al* 2000). Par ailleurs, il a été montré que la région du LAT est impliquée dans l'inhibition de la caspase 3. L'ensemble de ces données montrent que l'expression des LATs permet de maintenir un pool de neurones infectés, nécessaire à l'établissement de la latence (Thompson *et al*, 1997)(Perng *et al*, 2000)(Thompson *et al*, 2001).

Des études complémentaires menées sur des cellules transitoirement transfectées par un plasmide codant pour les nucléotides 301 à 2659 du LAT mineur ont montré que plus de 75 % des cellules transfectées étaient capables de résister à l'induction de l'apoptose par les ceramides (famille de lipides) ou les fumonisines (mycotoxines émises par certaines espèces de champignons du sol) (Perng *et al* 2000). Les niveaux de protection de cette construction codant pour le LAT sont comparables à ceux décrits pour des constructions comportant la protéine inhibitrice de l'apoptose du baculovirus (CpIAP). Comme les céramides et les

fumonisines n'induisent pas l'apoptose de la même façon, l'une agissant sur la voie des récepteurs de mort, l'autre sur la voie intrinsèque, il est probable que le LAT bloque l'apoptose en agissant sur une étape commune à ces deux voies probablement au niveau de la mitochondrie. De plus, une étude récente suggère que la présence de LAT semble activer, plus ou moins directement, un fort niveau de protéine AKT, protéine kinase B impliquée dans la voie de signalisation de la survie cellulaire, en inhibant l'apoptose (Li et al 2009).

II.8 Influence des LAT sur la réponse immunitaire spécifique.

Des travaux issus de différentes équipes ont montré que pendant la latence, une sous-population de neurones infectés dans les ganglions trigéminalés de souris sont entourés par des cellules lymphocytaires T CD8 (Knickelbein, J. E., *et al.* 2008.) (Branco, F. J., and N. W. Fraser. 2005.). Une observation similaire a par ailleurs été faite dans les ganglions trigéminalés humains. Il est maintenant admis que la réponse inflammatoire chronique, en particulier la présence de lymphocytes T CD8, favorise le maintien de la latence (Verjans *et al* 2007). Cette hypothèse est appuyée par des études montrant que le granzyme B a la capacité de cliver la protéine ICP4 (Knickelbein, J. E., *et al.* 2008). Le clivage de cette protéine virale essentielle mettrait ainsi fin à l'initiation du cycle répliatif, et donc au processus de réactivation.

Etonnamment, ces premières études n'avaient pas mis en évidence d'augmentation de l'apoptose consécutive au relargage de granzyme B par les cellules lymphocytaires. En effet, le granzyme B est une sérine-protéase présente dans les granules cytotoxiques des cellules NK et TCD8. Tout comme les caspases, le granzyme B coupe après un acide aspartique, ce qui lui permet d'activer les caspases, en particulier la caspase 3 (Chowdhury, D., and J. Lieberman. 2008. and Pardo, J., *et al.* 2009). Une fois libéré dans la synapse immunologique, il aurait donc dû activer l'apoptose, permettant ainsi la destruction de la cellule infectée par HSV-1. Ce n'est que deux ans plus tard que l'absence d'induction de l'apoptose dans ce contexte a été attribuée à la présence des LAT. Dans des cultures cellulaires de deux lignées neuronales de souris transfectées pour les LAT, les auteurs ont montré que le granzyme B n'active pas la caspase 3 (Jiang X *et al* 2011 ; Carpenter *et al* 2011). Les résultats de cette expérience ont été étendus à des cultures primaires de cellules neuronales, mais à ce jour il n'existe aucune preuve du blocage *in-vivo*, par les LATs, de l'activation de l'apoptose induite par le granzyme B .

Les LATs semblent aussi avoir un rôle sur le contrôle de l'extinction lymphocytaire CD8⁺ dirigée contre HSV-1. Une étude *in vivo* a montré que le nombre de lymphocytes activés ainsi que la sécrétion de TNF α et d'IFN- γ étaient moins importants après infection par une souche virale LAT- par rapport à la souche sauvage (Chentoufi *et al* ,2011). De plus, le nombre de lymphocytes CD8⁺ dirigés contre l'épitope gB₍₄₉₈₋₅₀₅₎ était moindre dans les TG infectés par le virus sauvage. Par ailleurs, il a été observé *in vitro* que des cellules neuronales de souris transfectées pour le LAT présentent une augmentation du niveau d'expression de PD-L1, ligand du récepteur PD-1 et du CMH1. Comme, en fixant son ligand, PD-1 favorise l'activation et l'expansion des cellules lymphocytaire (Yao, s 2006), il apparaît que la présence du LAT dans des neurones a une tendance à activer l'expansion lymphocytaire mais leur fait perdre leur spécificité antivirale.

II.9 Action des LAT sur la transcription des gènes viraux.

Puisque les LATs sont, de loin, les transcrits viraux les plus abondamment exprimés pendant l'infection latente, on peut penser que la production de LAT est incompatible avec la transcription des autres gènes viraux. En étudiant les profils transcriptionnels de deux souches virales LAT-, il est apparu deux phénotypes strictement opposés quant à la transcription des autres gènes viraux. Ainsi, en utilisant la souche KOS dans un modèle murin, la présence des LAT semble diminuer l'expression des gènes lytiques (Chen et al 1997), alors que l'opposé est observé avec un arrière-plan génétique de type syn17⁺ dans un modèle lapin (Giordani et al 2008). Il est difficile cependant dans ce type d'expériences de faire la part des phénomènes liés à la génétique du virus et de l'hôte (on connaît les différences importantes entre le lapin et la souris concernant les taux de réactivation spontanée du virus). Cependant, les avancées récentes sur les différences génétiques entre les deux souches virales et notamment sur la régulation épigénétique de l'expression de leurs génomes, il semblerait que le mode d'action des LATs sur l'expression des gènes lytiques soit beaucoup plus dépendant de la souche virale que du modèle animal concerné.

II.10 Action des LAT sur l'état de la chromatine

La régulation épigénétique de l'expression des gènes viraux est maintenant connue comme un phénomène central de la régulation entre période de réplication et période de latence du virus. En effet, il existe une corrélation forte entre l'état de la chromatine et le taux d'expression des différents gènes de HSV-1. Les études de chIP (*chromatin*

immunoprecipitation) réalisés sur des ganglions de souris infectées de façon latente ont montré qu'il s'agit plutôt d'une régulation passant par les maturations post-traductionnelles des histones et non le résultat de la méthylation du génome latent (Kubat et al 2004 a,b, Bloom et al 2010).

Pendant la latence, le locus LAT est enrichi en marqueurs d'euchromatine, état favorable à la transcription, à savoir une acétylation des lysines 9 et 14 de l'histone 3 et une diméthylation de la lysine 4 (Giordanni 2008, Kubat 2004b). Cependant, il existe, là encore, des différences entre les souches virales. Pour la souche KOS, l'analyse de mutants LAT- a montré que l'état de condensation de la chromatine ne semble pas lié à une transcription du LAT ni à la présence du promoteur du LAT (kubat et al. 2004a), alors que le même type de mutations introduites dans la souche *syn17+* entraîne une diminution de l'enrichissement en histone DiMe H3K4 au niveau du locus des LAT, synonyme d'une diminution de la perméabilité de la région pour la transcription. Dans ce cas le promoteur des LAT semble avoir un pouvoir suffisamment fort pour maintenir un état transcriptionnel actif. L'explication de ces différences entre les souches virales est actuellement inconnue.

Des études de cinétique de la répartition des marqueurs d'euchromatine sur l'ensemble du génome *in vivo* ont montré que les régions codant pour des gènes ICP4, TK, ICP8, ICP27 étaient enrichies en marqueur d'euchromatine au fur et à mesure de l'infection et de l'établissement de la latence (Wang et al., 2005), ce qui n'était pas le cas avec un virus LAT⁻. Cela suggérerait que la présence des LAT réprime les autres gènes en facilitant l'euchromatisation de leurs locus.

Toutefois, en 2009, d'autres études ont conclu à des résultats différents, rendant très débattu le rôle exact des LAT sur l'état de la chromatine virale (Cliffe et al., 2009, Kwiatkowski et al., 2009). S'il existe un consensus pour dire que la chromatine virale évolue au cours du cycle viral vers un état répressif favorable à la latence, le mécanisme exact par lequel le LAT influe sur cet état reste débattu. Pour certains, cette propriété serait plus liée au transcrit LAT lui-même qu'à son promoteur. En ce cas, le transcrit interagirait par un mécanisme qui reste à détailler, avec des protéines capables d'entraîner ces modifications, comme celles de la famille des PRCs (*polycomb repressive complex*), responsables de l'établissement et du maintien des modifications des histones H3 par méthylation de la lysine 27(triMe H3K27) (Clement C et 2011 ; Danaher RJ et al 2005).

II.11 Par quel(s) mécanisme(s) les LAT influent-ils sur autant de phénotypes ?

Plusieurs propriétés phénotypiques ont donc été attribuées au locus des LAT, mais à l'heure actuelle, aucun mécanisme commun ne permet d'expliquer l'ensemble des observations. Pour tenter de résoudre cette énigme, la grande majorité des études se sont concentrées sur l'étude du LAT2kb, car toutes les souches virales auxquelles les phénotypes sont associés sont génétiquement modifiées par délétion du promoteur et/ou du début du LAT 2kb, alors que les autres types de mutations n'entraînent pas les mêmes résultats. Cette idée est renforcée par l'observation qu'au cours de la latence herpétique, seul le LAT 2kb et ses dérivés (1,5 et 1,45 kb) s'accumulent dans les neurones infectés.

Dans les paragraphes qui suivent, nous allons donc passer en revue les différents modes d'action qui ont été proposés pour le LAT afin d'expliquer les divers phénotypes observés.

II.11.1 Le LAT_{2KB} : régulateur antisens de la protéine ICP0 ?

Le chevauchement partiel mais dans une orientation inverse entre le gène RL2, codant pour ICP0, et le gène codant le LAT 2KB, a conduit les premiers auteurs à suspecter un rôle régulateur des LATs sur l'expression de ICP0 par un mécanisme dit antisens (Stevens 1987). Cette hypothèse permettait d'expliquer les observations montrant que l'expression constitutive du LAT 2kb est associée à une diminution du nombre de transcrits du gène RL2 et de la charge virale pendant la phase aigüe de l'infection (Mador et al 1998). Cependant, la démonstration directe d'une régulation par mécanisme antisens du LAT2kb sur le transcrit RL2 n'a pas été concluante (Mador et al 2003) et cette hypothèse n'est plus actuellement retenue parmi les mécanismes d'action potentiels des LATs.

II.11.2 Le LAT_{2KB} code-t-il pour une protéine ?

Plusieurs protéines sont impliquées et exprimées pendant la latence de certains gamma-herpesviridae. Jusqu'à ce jour, aucune protéine traduite à partir du LAT n'a pu être mise en évidence de manière formelle même s'il existe de nombreux cadres ouverts de lecture (ORF, *open reading frame*) dans ce locus. Cependant, plusieurs arguments plaident en faveur d'une production peptidique à partir de ces ORF. Une mutation des divers ATG initiant les ORF putatives dans la partie du génome correspondant au LAT 1,5kb (sous-produit d'épissage du LAT 2kb) entraîne une moindre résistance des cellules infectées à l'apoptose

(Carpenter et al 2008). Une autre ORF putative, située sur le gène codant le LAT 1,5kb, mais antisens de ce dernier, pourrait conduire à la production d'une protéine nommée AL3 (Jaber et al 2009). La production d'anticorps monoclonaux à partir de la protéine synthétique a permis de retrouver un signal dans des expériences d'immunomarquage sur des coupes de ganglions trigémiques infectés de façon latente. Parmi les autres ORF capables d'être traduites dans des modèles *in vitro*, aucune ne semble réellement conduire à la production de protéines *in vivo*. Il n'est cependant pas exclu que les produits puissent avoir des niveaux d'expression trop faibles et/ou des demi-vies trop courtes pour que leur détection soit impossible avec les techniques actuelles. Cependant, la faible conservation de ces ORF au sein des différents *alpha-herpesviridae* et des différentes souches de HSV-1 soulève la question d'un rôle biologique réel.

II.11.3 Le LAT_{2kb} peut-il affecter la biogenèse des ribosomes ?

C'est au cours d'études visant à déterminer si le transcrit du locus LAT était capable d'être utilisé par la machinerie cellulaire de traduction pour donner naissance à des protéines que les premières observations de l'influence des LAT sur les ribosomes ont été effectuées (Nicosia et al., 1994).

En voulant vérifier que l'absence de protéine traduite à partir du LAT n'était pas la conséquence d'un défaut de reconnaissance de celui-ci par la machinerie de traduction, Nicosia et al. ont montré, au cours d'expérience de sédimentation, que le LAT 2kb migre avec la grande sous-unité ribosomique (Nicosia et al., 1994). Une analyse plus fine a ensuite montré qu'une quantité significative de LAT est associée aux polysomes *in-* et *ex- vivo* (Goldenberg et al 1997 a,b). La liaison entre les LAT et les ribosomes semblait même beaucoup plus forte que celle attendue entre un ARNm et son ribosome (Ahmed et al 2001). A l'aide de cellules transfectées pour le LAT 2kb puis traitées par des inhibiteurs de la biogenèse des ribosomes, les mêmes auteurs ont montré que l'interaction entre le LAT et la sous unité 60s du ribosome est irréversible, suggérant que le LAT 2kb pourrait se comporter comme un ARNr (Atanasiu et al 2007). Ces auteurs ont aussi montré que le LAT peut s'associer aux protéines de la famille des hnRNPs (*Heterogeneous nuclear ribonucleoproteins*) dont certaines sont impliquées dans la plupart des mécanismes d'épissage et de maturation des ARNm. Le rôle exact des LAT dans cette interaction reste à définir.

II.11.4 Le LAT_{2KB} comme régulateur de l'épissage des gènes viraux et cellulaires ?

Les expériences menées dans notre équipe ont montré, en comparant différents types de tissus neurologiques infectés de façon aigüe puis latente, que globalement plus le LAT est abondant moins on détecte la forme épissée du transcrit RL2, et plus on observe l'accumulation de formes non-épissées de ce transcrit (Maillet et al 2006). En culture cellulaire, une autre étude a montré que la présence du LAT influait l'épissage des transcrits du gène Bclx (Peng 2003). L'ensemble de ces données suggère que le LAT pourrait aussi réguler l'épissage de certains transcrits viraux et cellulaires.

II.11.5 Le LAT_{2KB} comme source de micro-RNA (mi-RNA) ?

Une des dernières hypothèses émises sur le rôle des LATs est qu'ils pourraient agir en tant que microARN (miARN).

Les miRNA sont des ARN non codants de 21 à 23 nucléotides de long, qui sont désormais considérés comme des acteurs majeurs dans la régulation de l'expression génique. Ils interagissent avec la région 3'UTR des ARN messagers par complémentarité de base entre les deux éléments. Chez les plantes, la complémentarité entre le miARN et l'ARNm est presque parfaite. Ceci engendre la dégradation de l'ARNm cible suivant un processus semblable au mécanisme d'ARN interférence, avec un clivage endonucléasique qui se produit au milieu du duplex miARN-ARNm (Jones-Rhoades et al. 2006).

Chez les métazoaires, les bases de l'appariement entre miARN et ARNm sont tout autres, puisque la complémentarité entre le miARN et l'ARNm est imparfaite. Cependant, cet appariement suit des règles précises, dont certaines sont indispensables à la répression exercée par le miARN sur l'ARNm. Dans ce cas de figure, les miARN contrôlent l'expression génique aux stades post-transcriptionnels en régulant (1) l'initiation de la traduction, (2) l'élongation de la traduction, (3) la protéolyse du peptide naissant, et enfin (4) la stabilité du messager en guidant sa déadénylation.

Des analyses bio-informatiques ont permis d'identifier des miRNA potentiellement codés par le transcrit mineur du locus LAT (Kramer et al 2011, Mulik S et al 2012). Leur surexpression *in vitro* a montré un effet inhibiteur sur l'expression des protéines ICP0, ICP4, ICP34.5 (seulement dans le cas de HSV-2 pour ce dernier) (Tang et al 2008a,b ; Umbach et al, 2008). Certains de ces miRNA pourraient aussi réguler des transcrits neuronaux, puisque

d'autres miRNA jouent un rôle dans le *silencing* de gènes cellulaires . Le *silencing* de gènes neuronaux comme de gènes pro-apoptotiques faciliterait l'établissement et le maintien de la latence. A ce jour aucune expérience n'a permis de confirmer cette hypothèse in-vitro ou in-vivo, en partie l'absence de souches génétiquement modifiées pour la seule expression de ces miRNA n'a pas permis de confirmer les prédictions bio-informatiques.

II.11.6 Le locus du LAT_{2KB} peut-il jouer un rôle de « super enhancer » ?

La partie 5' de la région codant pour les LAT, dont une partie du premier intron (LAT 2kb), n'est pas seulement importante pour l'expression du LAT lui-même, mais pourrait aussi agir en *trans* sur d'autres gènes en jouant le rôle d'*enhancer* (Berthomme et al 2000)(figure16). De plus, il a été démontré que cette région jouait un rôle important pour le maintien à long terme de l'expression des LAT pendant la latence c'est pourquoi elle est dénommée, région LTE (*Long Term Expression*) (Lokensgard et al 1994, Margolis et al 1993)

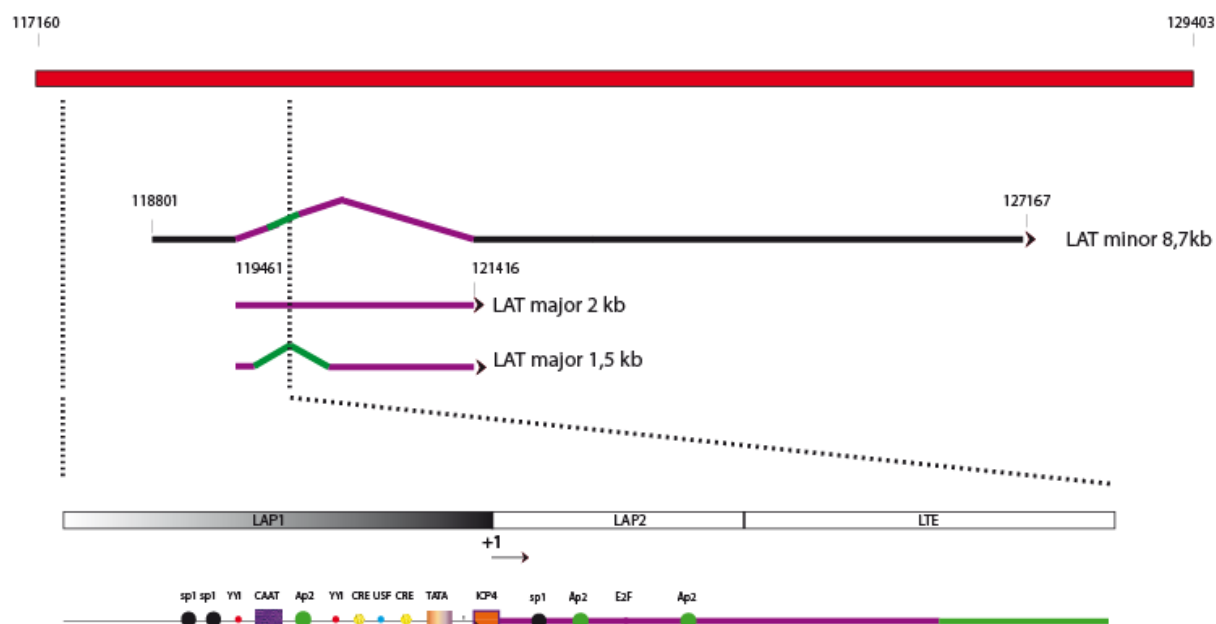


Figure 16: Représentation de la région du LAT impliquée dans le maintien à long terme de l'expression des LAT et pendant la latence. Le rôle d'*enhancer* du LAT est localisée dans la région 5' du gène, qui comprends, les promoteurs LAP1 et LAP2, la région LTE, et une partie du LAT 2kb.

Sa localisation distale par rapport aux promoteurs, potentiellement situés jusqu'à plusieurs milliers de pb, pourrait stimuler l'expression d'autres gènes viraux en recrutant des protéines activatrices. Cette hypothèse est compatible avec certains des phénotypes des souches LAT-, notamment ceux liées aux propriétés régulatrices de l'activité transcriptionnelle et anti-apoptotique.

Une des composantes majeures de ce système est la découverte récente de sites de fixations pour les protéines de la famille des CTCF (Amelio 2006b). La protéine CTCF comporte 11 domaines en doigt de zinc qui possède une activité de régulation transcriptionnelle dépendante de la phosphorylation. Bien qu'un seul site de fixation soit nécessaire à l'activité régulatrice de la protéine, ceux-ci sont souvent organisés en série, permettant d'accroître l'affinité de la protéine pour sa cible (Chao et al 2002 ; Weth et al 2011). La fixation de la protéine avec sa cible permet de bloquer l'interaction entre un *enhancer* et un promoteur, limitant ainsi l'action activatrice de l'*enhancer* (Cuddapah S,2009)(figure17). La protéine CTCF permet aussi le blocage de la diffusion de l'hétérochromatine pour un locus donné. Grâce à ce système, deux gènes situés sur le même chromosome et soumis aux mêmes séquences régulatrices peuvent s'exprimer indépendamment l'un de l'autre.

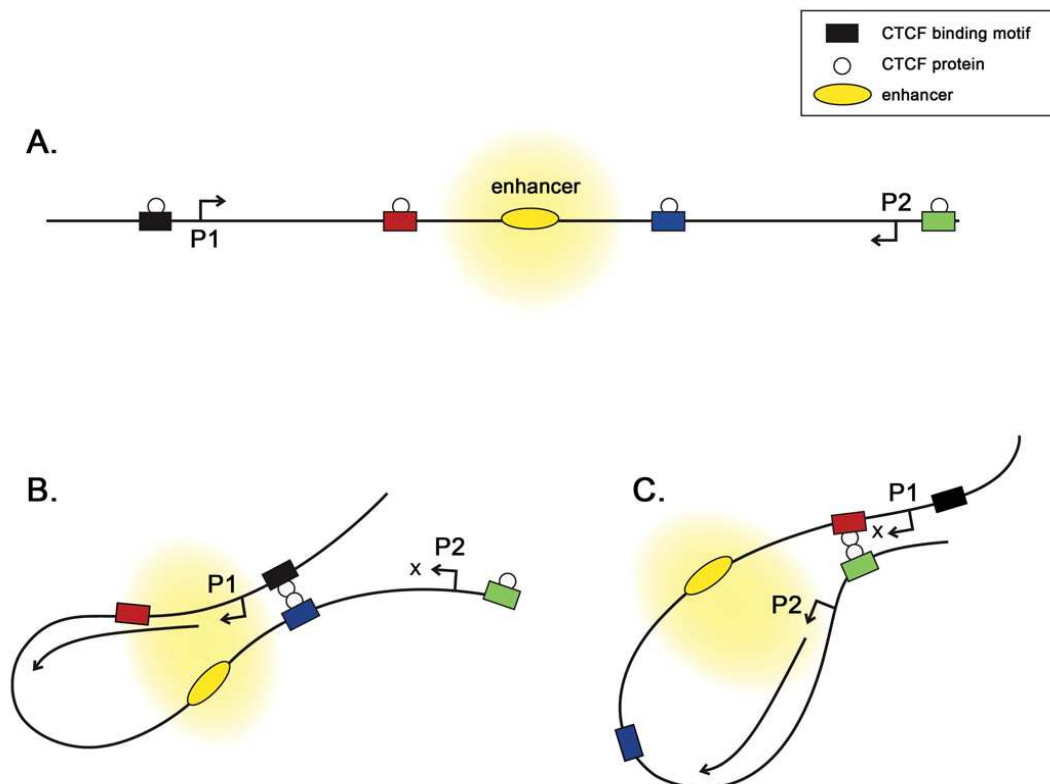


Figure 17: Représentation schématique du mode d'action de la protéine CTCF sur la régulation de la transcription

La recherche de domaines de fixation pour la protéine CTCF sur le génome de l'herpès a mis en évidence deux sites de fixation appelés domaines B1 et B2, situés de part et d'autre de LAP1, en plus de cinq autres domaines répartis sur le génome (Amelio, et al 2006 a

, b). La région B2, située dans la région codant le LAT 2kb, semble capable de contrôler l'expression de son propre locus et des gènes avoisinants.

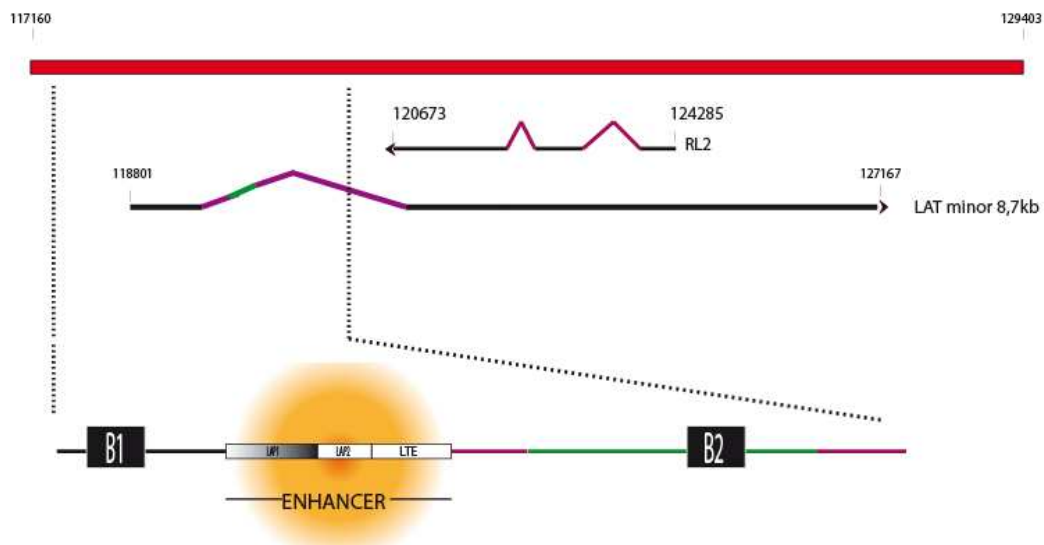


Figure 18: Schéma représentant la position des domaines de fixation pour la protéine CTCF sur le locus des LAT. La fixation de la protéine CTCF avec les régions B1 et /ou B2 permet de bloquer l'interaction entre l'*enhancer* et un promoteur, viral ou cellulaire, permettant ainsi au locus des LAT de réguler l'expression d'autres gènes.

Un mécanisme de régulation supposé se ferait *via* la formation de boucle ADN-ADN entre les régions B1 et B2 et d'autres régions de fixation aux protéines CTCF présentes sur les génomes viraux ou cellulaires, permettant ainsi l'activation ou la répression de promoteur (figure 18).

II.12 Le LAT_{2KB} peut-il agir à travers une combinaison de ses différents mécanismes ?

A ce jour, la multitude des phénotypes attribués aux LATs rend très difficile l'établissement d'un modèle unique permettant d'expliquer leurs modes d'action.

Au vu de la complexité du locus des LAT, il est possible d'imaginer que les produits de celui-ci puissent agir par différents mécanismes via plusieurs voies simultanément pour promouvoir la latence et la réactivation. Il est assez commun de penser que la diversité des transcrits codé par le locus lui-même puisse être la source des différents phénotypes à ce jour

observés. Néanmoins, si on se base sur l'ensemble des données disponibles, on peut réduire à une seule région l'attribution des phénotypes portés par des virus mutants pour les LAT. Cette région est définie par le promoteur LAP1 et le LAT majeur de 2kb.

Cette limitation conduit à proposer un modèle d'action sur la régulation et le maintien de la latence par les LAT à deux niveaux.

Dans un premier temps il semblerait que le promoteur LAP1 et sa région avoisinante n'aient pas qu'un rôle sur la transcription des LATs majeur et mineur. Ils pourraient aussi jouer le rôle d'enhancer capable d'activer la transcription de gènes codants pour des protéines impliquées dans le maintien et l'établissement de la latence (cellulaires et/ ou virales), en jouant sur l'état de la chromatine des gènes concernés (figure 19).

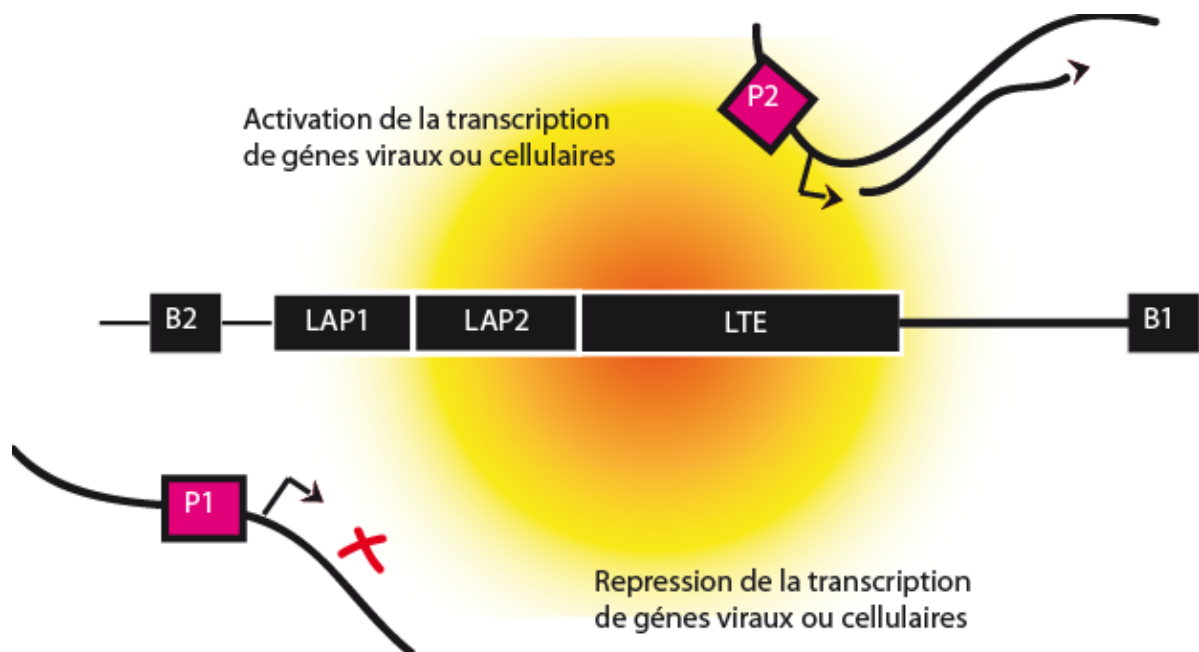


Figure 19 : Schématisation du mode d'action du locus des LAT sur l'expression de gènes viraux et cellulaires

Ce rôle de « super enhancer » sur l'expression de gènes viraux et cellulaires a en partie été démontré pour les gènes viraux avoisinants ou présents dans le locus des LAT (notamment ICP0) expliquant ainsi comment LAP1 peut favoriser l'établissement de la latence en bloquant l'expression de gènes qui sont défavorables à celle-ci et en activant d'autres qui lui sont propice. Ce modèle souffre néanmoins de lacunes. En effet, le blocage ou l'activation spécifique des gènes cellulaires, impliquerait la présence de génomes viraux au voisinage de

ceux-ci. Cela se heurte à deux problèmes : premièrement, il a été démontré que des neurones infectés de façon latente pouvaient héberger un nombre variable de génomes viraux (Lomonte et al, 2012). Cette variabilité implique que le nombre de locus des LAT pouvant agir à différents endroits du noyau est aléatoire et donc que le nombre de gènes potentiellement régulés l'est aussi. Deuxièmement, il est convenu de penser que si le LAT pouvait agir en trans sur d'autres gènes cellulaires, cela supposerait que les génomes viraux seraient distribués toujours de la même manière (au voisinage des gènes à réguler) dans des neurones latents, or à ce jour il n'existe aucune étude permettant de dire que la distribution des génomes latents n'est pas aléatoire.

La seconde partie du modèle concerne le mode d'action du LAT de 2kb qui à ce jour est vu comme un ARN non codant puisque aucune protéine fonctionnelle n'a jamais été obtenue à partir de celui-ci. Par ailleurs, le LAT_{2kb} n'est ni coiffé, ni polyadénylé. Ceci et son statut d'intron rendent difficile son export hors du noyau et expliquent pourquoi il n'est pas retrouvé libre (ou seulement en très faible quantité) dans le cytoplasme des cellules infectées de façon latente. Enfin, sa structure en lariat et l'absence de maturations propres aux ARN messagers rendent très peu probable la possibilité, qu'il puisse être pris en charge puis traduit par les ribosomes pour donner naissance à une protéine. C'est pourquoi, la possibilité que son action puisse se faire via la traduction d'une protéine « couteau suisse » qui serait capable de bloquer ou de modifier les différentes voies de signalisation est hautement improbable.

Un ARN non codant (ou ARNnm pour ARN non-messenger) est un ARN, issu de la transcription de l'ADN, qui ne sera pas traduit en protéine par les ribosomes. L'importance quantitative et fonctionnelle des ARN non codants a été très longtemps sous-estimée. On en connaît aujourd'hui une grande diversité. Par exemple, alors que seulement 1,2% du génome humain est traduit en protéine, il apparaît qu'une bien plus grande partie est transcrite en ARN. Les ARN non-codants assurent diverses fonctions pour la plupart déterminées par la structure spatiale qu'adopte la molécule d'ARN en se repliant sur elle-même (ou en interagissant avec sa cible - souvent un autre acide nucléique -).

Avec actuellement plus de 600 familles d'ARN non-codants connues et répertoriées dans des banques de données, il est difficile d'établir à quelle famille le LAT_{2kb} pourrait appartenir et donc d'en déduire un mécanisme d'action précis. Dans l'hypothèse où le LAT serait un ARN non codant il est essentiel que la fonction de celui-ci permette d'expliquer

l'ensemble des phénotypes observés. Certaines observations sur les caractéristiques du LAT peuvent néanmoins orienter notre réflexion.

Ainsi, une des particularités du LAT de 2kb est d'être un intron épissé lors de la maturation du LAT mineur. L'étude de la distribution des tailles des introns, au sein d'espèces allant de la plante à l'homme en passant par la mouche et la souris, a révélé que les introns de grande taille (supérieure à 1 kb) n'étaient pas les plus répandus au sein du monde eucaryote (Figure 20). Il semblerait donc que le virus de l'herpès ait gardé ou acquis au cours de son évolution un intron d'une taille peu fréquente. Cette observation est d'autant plus notable que le virus code pour plus de 100 transcrits parmi lesquels seuls cinq (LAT inclus) sont épissés.

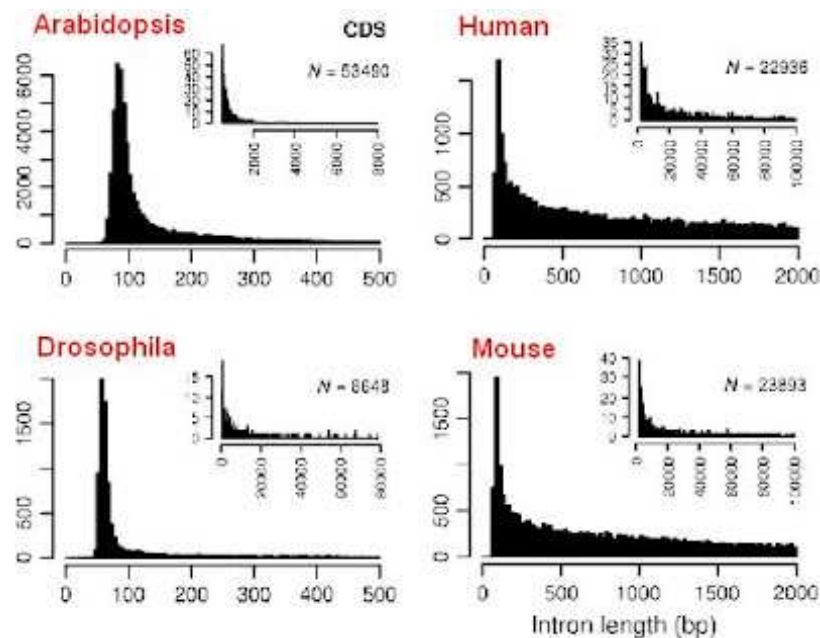


Figure 20: distribution de la fréquence des introns en fonction de leur tailles au sein de quatre espèce :*Arabidopsis thaliana*; *Drosophila melanogaster*; l'homme, et la souris. (Hong et al. 2006).

Une autre particularité du LAT de 2 kb est qu'il se maintient sous la forme d'un lariat stable. Le maintien de la structure en lariat est dû au fait que lors de l'épissage, le 2'-OH du ribose responsable de l'attaque sur le phosphate de la jonction exon-intron en 5' n'est pas porté par une adénosine mais par une guanosine. La liaison ainsi formée ne serait pas reconnue par l'enzyme impliquée dans l'étape de débranchement nécessaire à la libération de

l'intron (dans le cas présent le LAT de 2kb). L'attaque nucléophile du 2'-OH du ribose sur la guanosine est orientée grâce à la formation d'une tige boucle qui faciliterait ainsi la présentation du nucléotide (Krummenhacher et al 1997). Le maintien sous forme circulaire de l'ARN empêcherait alors l'action des Rnases cellulaires contribuant au moins partiellement à la stabilité du LAT 2kb dans la cellule infectée (Mukerjee 2004, Wu 1996, Wu 1998, Ng 2004, Krummenacher et al 1997).

De plus, la richesse en GC du LAT de 2kb environ 66% ainsi que le grand nombre de séquences répétées (stretch de T ; stretch de A ; stretch de G et de C) suggère aussi que le LAT est un ARN qui possède une structure secondaire complexe (Figure 21).

Nucleotides	Nombre	Pourcentage
A	320	16,37
C	735	37,60
G	555	28,39
T	345	17,65

Figure 21: fréquence de chaque nucléotide retrouvé dans la séquence du LAT_{2kb}.

L'idée que le LAT adopte une conformation secondaire hautement structurée, le rendant résistant à la dégradation, est supportée par sa capacité à s'accumuler dans les neurones infectés de façon latente ainsi que par sa grande demi-vie estimée à 24h (Thomas et al 2002). Comme pour les autres ARN, trois types d'interactions détermineraient la stabilité de la structure du LAT. La première est une contribution électrostatique liée à la formation des liaisons hydrogènes entre les bases. La seconde est l'empilement des plateaux des paires de bases (*stacking*) qui donne naissance à des interactions de van der Waals. La troisième composante est l'interaction avec le solvant aqueux et les ions de la solution dans laquelle se trouve le LAT, en particulier la présence de cations divalents comme Mg^{2+} , souvent nécessaires pour neutraliser la charge des groupements phosphate du squelette de l'ARN. Le repliement et la formation d'une structure tridimensionnelle par l'ARN implique en effet le rapprochement de ces groupements phosphate, ce qui est défavorable en l'absence de contrions. Le groupement 2'-OH porté spécifiquement par les riboses de l'ARN (mais absent dans l'ADN) joue également un rôle dans la stabilité et la formation de la structure, en formant des liaisons hydrogènes avec les bases.

Par ailleurs, certains ARN peuvent adopter plusieurs conformations alternatives en fonction de la liaison d'un ligand (protéine, petite molécule...) ou des conditions physico-chimiques (force ionique, pH). Cette constatation pourrait alors expliquer pourquoi le LAT de 2kb ne s'accumule que dans les neurones où les conditions physico-chimiques seraient favorables à la formation de sa structure secondaire puis tertiaire.

L'existence de structures secondaires et tertiaires bien définies dans les ARN est un des éléments importants de la fonction d'un certain nombre d'entre eux. Ces structures leur permettent de former des sites de liaison pour des ligands sélectifs, petites molécules ou protéines, et, pour les ribozymes, elles leur permettent d'assurer des fonctions catalytiques. Ainsi la supposition que le LAT_{2kb} possède une organisation structurale complexe suggère qu'il peut posséder des fonctions précises qui peuvent être modulées par son environnement physico-chimique ou par son association à d'autres molécules.

II.13 Comment un intron de grande taille sous forme circulaire peut-il agir sur autant de voies de régulations ?

En raison du grand nombre de phénotypes attribués au LAT de 2kb, ainsi que la multitude de processus cellulaires ou viraux qu'il touche, il est nécessaire d'imaginer un modèle dans lequel le LAT de 2kb agirait sur un mécanisme commun à l'ensemble de ses voies.

Lors du début de ce travail, deux phénotypes suggéraient que le LAT pouvait avoir comme cible différents ARN. Tout d'abord, il semblait être capable d'agir sur l'épissage de transcrits viraux (Maillet et al., 2006) et cellulaires (Peng et al, 2003). Ensuite, il était retrouvé associé aux ribosomes probablement parce qu'il pouvait se lier aux ARNr les constituant (Nicosia et al., 1994 ; Goldenberg et al., 1997).

Ces résultats étaient particulièrement intéressants, parce qu'en ciblant les ARN viraux et cellulaires, le LAT pouvait agir suffisamment en amont pour pouvoir modifier la régulation de plusieurs processus. On pouvait concevoir que le LAT de 2kb était un élément capable de reconnaître des ARN possédant une séquence nucléotidique particulière ou une structure secondaire (tige boucle), ou bien en complexe avec une protéine précise.

La capacité apparente du LAT à interagir avec des acides nucléiques et à influencer l'épissage, ainsi que sa structure en lariat et ses propriétés rappelant celles de certains

éléments mobiles d'origine procaryote, nous a suggéré que le LAT pouvait se comporter comme un élément mobile utile au virus pour contrôler l'établissement de la latence au cours d'une infection. C'est sur cette hypothèse que s'est construit le travail de cette thèse.

Avant de présenter les résultats, il est nécessaire de faire un état des lieux des propriétés des différents éléments mobiles eucaryotes décrits à ce jour.

Chapitre III

Les éléments mobiles

III.1 Généralités

C'est après plusieurs décennies d'études génétiques et moléculaires (travail commencé en 1940 et poursuivi jusqu'à sa mort en 1992) que Barbara Mc Clintock a mis en évidence chez le maïs des phénomènes d'instabilité génétique qu'elle a attribué à un élément mobile nommé *Ds* (pour dissociation) et à son compagnon l'élément *Ac* (pour activateur). Au cours de ses travaux, Barbara Mc Clintock a montré en détail comment fonctionnaient les éléments *Ds* et *Ac* en particulier comment ceux-ci font intervenir des mécanismes de transposition conduisant à l'altération de l'expression des gènes voisins. Ce travail a révolutionné les idées précédemment admises sur la stabilité du génome et ouvert la voie aux conceptions modernes de fluidité, de variabilité, et de constante réorganisation du génome.

Il est aujourd'hui connu que tous les organismes contiennent au sein de leurs génomes des éléments mobiles qui au cours de l'évolution ont envahi le génome. Ainsi, 45% du génome de l'homme est composé par des éléments transposables ou des séquences qui en dérivent, contre environ 15% chez *Drosophila melanogaster* et plus de 70% chez le maïs (*Zea mais*), comme chez la plupart des angiospermes. À elle seule, la présence de ces éléments peut expliquer d'importantes différences dans les tailles du génome d'organismes proches.

Chez les bactéries, les éléments transposables ont été décrits pour la première fois dans les années 60. Les premiers travaux concernaient la découverte du bactériophage Mu responsable de multiples mutations par insertion chez *Escherichia coli* K12 (Taylor et al , 1963). Ultérieurement, les équipes de Shapiro et de Starlinger ont rapporté l'existence de mutants d'*E. coli* dans l'opéron galactose et relié ces mutants spontanés à l'insertion de matériel génétique (Shapiro et al, 1969, Michaelis *et al.*, 1969). Finalement, chez les bactéries, certaines transmissions de résistances aux antibiotiques ont été reliées à la présence d'éléments transposables (Watanabe & Fukasawa, 1962). Avec l'évolution des connaissances, il est devenu de plus en plus évident que ces éléments génétiques mobiles faisaient aussi partie intégrante du paysage génomique des procaryotes.

III.2 La classification des éléments mobiles eucaryotes.

La multiplication des projets de séquençage de génomes a permis la découverte d'un nombre grandissant d'éléments transposables, rendant nécessaire un système de classification de plus en plus précis. À ce jour, les grandes familles décrites le sont sur la base de leurs mécanismes de transposition, sur la présence ou l'absence de certaines fonctions, sur le besoin de séquences-cibles pour leur insertion et sur la nature de ces séquences. Actuellement, les éléments transposables eucaryotes se répartissent en deux grandes classes.

Les éléments de la classe I, aussi appelés rétro-éléments, se disséminent *via* un intermédiaire ARN grâce à une machinerie enzymatique codée par l'élément lui-même. Ce mécanisme est souvent résumé par l'expression « copier-coller ». Généralement de grande taille (quelques centaines à quelques milliers de paires de bases), ils peuvent représenter une part importante des génomes (voir par exemple SanMiguel et al,1998). Plusieurs sous-classes et sous-familles peuvent ensuite être distinguées selon les caractéristiques structurales (présence de LTR, ordre des gènes) et phylogénétiques. Ainsi, Tam et al en 2005 repartissent les éléments de classe I selon une hiérarchie définie par leur complexité. Les rétrovirus sont les éléments les plus complexes, capables de s'encapsider et de quitter le volume cellulaire. À l'opposé, les LINE (*Long Interspersed Nuclear Element*) sont des entités très simples, autonomes, incapables d'infecter d'autres cellules.

Les éléments de la classe II, appelés transposons à ADN, codent une enzyme, la «transposase», qui reconnaît les extrémités de l'élément, l'excise et l'insère ailleurs dans le génome. On parle dans ce cas de « couper-coller ». Plus courts que les éléments de classe I, leur mode de transposition limite leur importance en volume de séquence. Cependant, les éléments de classe II jouent un rôle évolutif important (Kazazian,et al 2004).

Enfin il existe dans les organites de certains eucaryotes, des éléments mobiles tels que les introns autocatalytiques qui ont une origine procaryote.

Aucun système de classification des éléments transposables n'étant complètement accepté, nous avons fait le choix d'opter pour une présentation simplifiée des différentes classes d'élément transposables nous permettant ainsi de décrire la majeure partie des mécanismes engendrant leurs mobilités.

III.3 Les éléments mobiles des organites cellulaires

III.3.1 Les introns du groupe II

III.3.1.1 Généralités

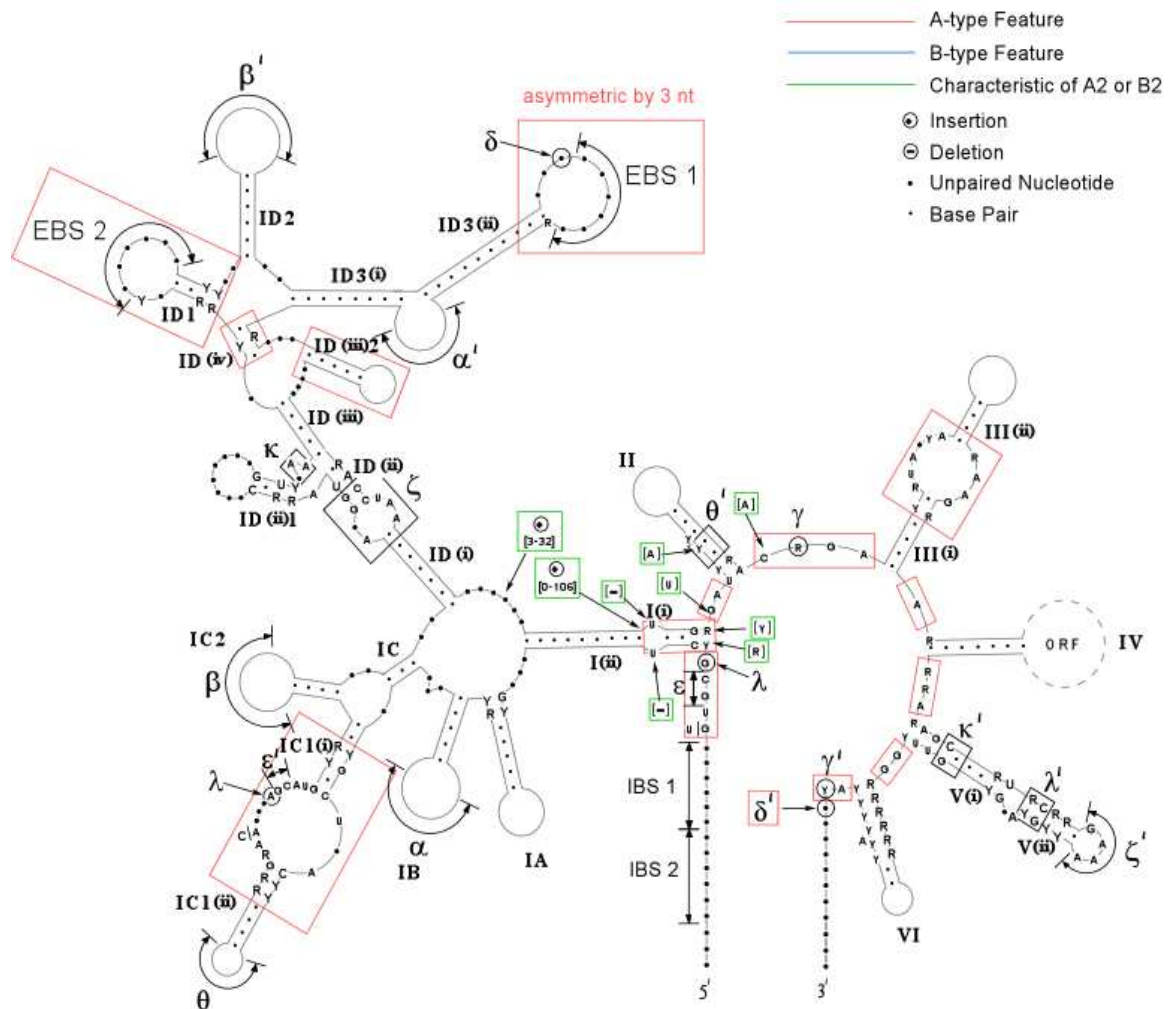
Initialement découverts dans les organites des protistes, des végétaux et des champignons, les introns de groupe II ont été décrits en 1993 chez les bactéries (Ferat & Michel, 1993). On pense d'ailleurs que ces éléments étaient initialement présents chez les bactéries dont certaines espèces seraient les ancêtres des mitochondries et des chloroplastes (Ferat & Michel, 1993). Grâce aux programmes de séquençage à grande échelle des génomes, on a découvert que les introns du groupe II étaient très largement répandus parmi les procaryotes et, en 2008, plus de 200 introns du groupe II avaient été identifiés dans une centaine d'espèces bactériennes.

Les introns du groupe II se replient en une structure secondaire conservée, constituée de six domaines arrangés autour d'une boucle centrale (Michel & Ferat, 1995). Le domaine V est le plus conservé et correspond au site actif du ribozyme. Un cadre de lecture est localisé dans la boucle du domaine IV et code pour une transcriptase inverse (RT). La protéine est constituée des sept domaines communs à toutes les RT ainsi que d'un domaine X et d'un domaine En. Le domaine X porte l'activité de maturase de la protéine et le domaine En a une activité de nucléase utilisée lors de la réaction de mobilité (Michel & Ferat, 1995).

Ils transposent via un intermédiaire ARN, appelé ribozyme, qui agit comme une enzyme. Le ribozyme catalyse *in vitro* sa propre excision à partir du transcrit primaire par un procédé d'auto-épissage. L'activité maturase de la protéine qu'il code coopère avec le ribozyme pour faciliter la mobilité de l'élément (Michel & Ferat, 1995, Ferat *et al.*, 1994). Les introns du groupe II sont localisés en majorité en dehors de cadres de lecture. On les trouve le plus souvent sur des plasmides ou localisés dans des îlots de pathogénicité. Une classification a été établie sur la base des phylogénies construites à partir des séquences des ORF. Huit classes ont été définies : CL, ML, A, B, C, D, E et F. Toutes ces classes sont retrouvées chez les bactéries. Seules les classes A, B et F sont retrouvées chez les mitochondries (Toor *et al.*, 2001, Zimmerly *et al.*, 2001).

III.3.1.2 Structure des introns du groupe II

Les introns du groupe II se replient en une structure secondaire conservée constituée de six domaines arrangés autour d'une boucle centrale (Figure 22)(Michel & Ferat, 1995).



d'après Michel & Ferat, 1995 et Zimmerly et al 2001

Figure 22 : Représentation schématique de la structure secondaire des introns de groupe II A.

Le domaine V est le plus conservé et correspond au site actif du ribozyme. L'ORF est localisé dans la boucle du domaine IV et code la RT. (Michel & Ferat, 1995).

III.3.1.3 Mécanisme réactionnel

III.3.1.3.1 Auto-épissage et assemblage du complexe ribonucléoprotéique

Le chemin réactionnel de l'auto-épissage est identique à celui des introns des ARN pré-messagers nucléaires et aboutit, après clivage et ligation des deux extrémités 5' et 3', à la formation d'un lariat (Lambowitz & Zimmerly, 2004) (Figure 24). *In vivo*, l'auto-épissage requiert l'assistance de l'activité maturase. La RT se lie au domaine IV de l'intron. Les interactions protéine-ARN, qui se créent, entraînent un changement de conformation de l'intron aboutissant à son auto-épissage. Après épissage, la RT reste liée fortement à l'intron et, par cela, augmente la fraction active des réactions de mobilité (Lambowitz & Zimmerly, 2004).

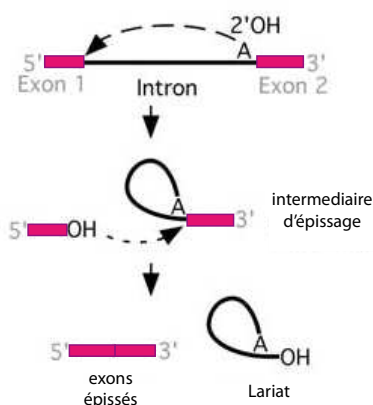


Figure 23: mécanisme d'épissage des introns de groupe II. Représentation simplifiée des différentes étapes d'épissage des introns de groupe II, au cours desquelles deux réactions séquentielles de trans-estérification ont lieu, aboutissant à la formation d'un lariat.

III.3.1.3.2 Mobilité des introns du groupe II

Le mécanisme de mobilité est caractérisé par les termes de *homing* ou *retrohoming* (Cousineau *et al.*, 1998). Le complexe ribonucléoprotéique (RNP) reconnaît un site spécifique sur l'ADN dépourvu de l'intron et l'ARN s'insère dans le brin sens de l'ADN au niveau de la jonction des exons (Michel & Ferat, 1995). La réaction est catalysée par le ribozyme mais

facilitée par la RT comme pour l'épissage. Le domaine En de la transposase clive le brin antisens et le domaine RT copie l'intron en ADNc. La réaction démarre à l'extrémité 3'OH de l'ADN clivé qui sert alors d'amorce (Lambowitz & Zimmerly, 2004). En général, la mobilité des introns du groupe II est très spécifique et cible une séquence spécifique d'environ 35 pb dépourvue d'intron. Les introns du groupe II sont aussi capables de s'insérer dans des sites ectopiques à un plus faible taux ; il s'agit alors du mécanisme de rétrotransposition (Cousineau *et al.*, 1998, Martinez-Abarca & Toro, 2000). Une troisième possibilité de mobilité existe chez les introns de la classe C où l'intron s'insère dans des motifs précis non linéaires impliquant éventuellement une structure locale de l'ADN (Dai & Zimmerly, 2002, Granlund *et al.*, 2001, Roberts *et al.*, 2008).

III.4 Les éléments transposables de classe I

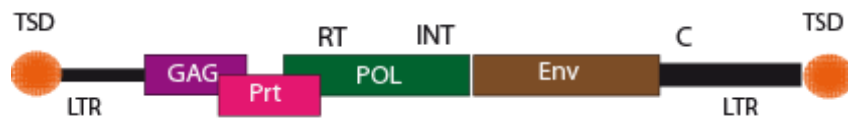
Les éléments transposables de classe I ou rétrotransposons ont la particularité d'utiliser un intermédiaire ARN qui sert à la fois de support pour la synthèse d'un brin d'ADN complémentaire par transcription inverse, et de support pour la traduction de l'ensemble des protéines nécessaires au cycle de rétrotransposition. Ainsi, une copie intégrée synthétise des ARN capable de s'intégrer à d'autres endroits, permettant ainsi la multiplication du nombre de copies de l'élément mobile. On distingue historiquement deux sous-classes de rétrotransposons.

III.4.1 Les rétrotransposons à LTR.

Les rétrotransposons à LTR sont des éléments génétiques ayant de nombreuses caractéristiques communes à celles observées chez les rétrovirus, exception faite de leur incapacité à se propager de cellule en cellule.

Les rétrotransposons ont été détectés dans des génomes aussi variés que ceux de la levure (Ty), des insectes (éléments de type copia), des plantes (Cin1 du maïs, Tnt1 du tabac) et de la souris (les séquences IAP (*intracisternal A particle*)).

A ce jour, aucun rétrotransposon à LTR actif n'a été identifié chez l'homme, en revanche les rétrovirus endogènes de la famille des HERVs sont les éléments génétiques les plus proches des rétrotransposons à LTR observés, et représentent à eux seuls environ 7% du génome humain (Figure 24).

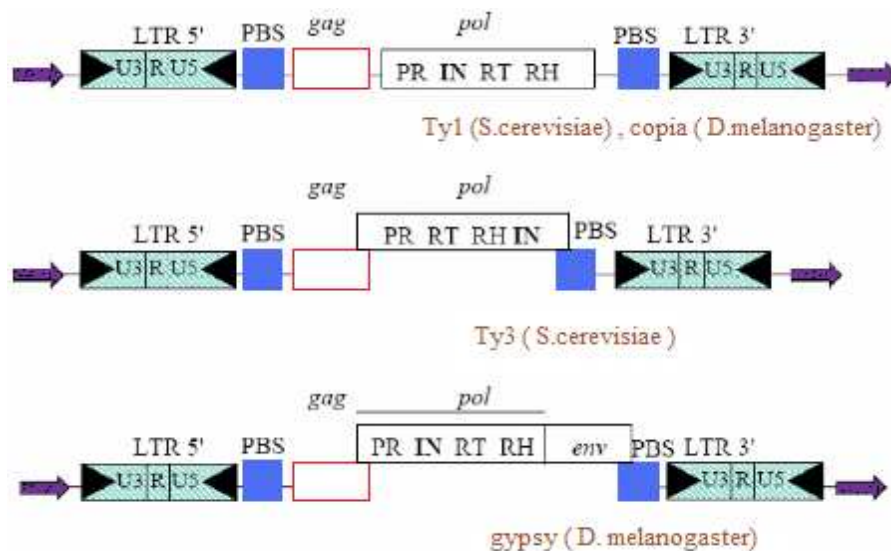


HERVs (7-9 kb)

Figure 24: représentation de la structure des rétrovirus endogènes humains. Sur la figure, sont représentés les sites dupliqués de l'insertion (TSD, *target site duplication*), le gène GAG, le gène Prt qui code pour la protéase virale, le gène POL qui code pour la polymérase virale qui possède l'activité RT et pour l'intégrase (INT) et le gène ENV. Les régions répétées terminales (LTR *long tandem repeat*) sont indiquées en noir.

Les génomes des rétrovirus et des rétrotransposons ont des organisations similaires. Ils présentent les caractéristiques suivantes :

- à chaque extrémité, des LTR identiques constitués des séquences U3, R, U5 sont retrouvés en orientation directe ;
- les séquences des rétrovirus comprennent en général quatre gènes dont l'ordre est *gag-pol-env*. Les protéines codées par la région Pol (RT et intégrase) sont nécessaires à la réplication des rétrovirus. Le gène *gag* code pour les protéines de capsid. Chez les rétrotransposons, le gène *env*, codant pour les protéines d'enveloppe, est absent, sauf dans le cas de *gypsy* de *Drosophila melanogaster*. (Figure 25)



Warmus et Brown 1989

Figure 25: Structure des différents types de rétrotransposons à LTR. Chez les rétrotransposons, l'utilité du gène gag n'est pas encore élucidée. Le gène pol code pour une polyprotéine qui, après protéolyse, forme les protéines requises pour la transposition, notamment une protéase nécessaire à la maturation protéolytique des diverses protéines impliquées dans cette étape, une transcriptase inverse, une Rnase H qui permet la dégradation de l'ARN lors de la formation de l'hétéroduplex ADNc-ARN et une intégrase, qui permet l'insertion de l'élément dans le génome. Les rétrotransposons à LTR sont subdivisés en deux groupes principaux suivant l'ordre des domaines protéiques du gène pol. Chez les éléments de type Ty1/copia, l'intégrase se trouve du côté 5' de la transcriptase inverse alors que chez les éléments de type Ty3, elle se place du côté 3'. Enfin, chez certains rétrotransposons à LTR du groupe gypsy/Ty3, on trouve aussi une troisième ORF, le gène env.

Les rétrotransposons, comme les rétrovirus, ont une copie d'ADN intégrée dans le génome de l'hôte (appelée provirus pour les rétrovirus). Cette copie sert de matrice à la synthèse d'un ARN polycistronique qui sera utilisé pour différentes fonctions. Il peut être traduit pour fournir les protéines nécessaires à la réplication du rétrotransposon. Il peut être inversement copié en ADN qui peut s'intégrer ailleurs dans le génome.

Le cycle de réplication des rétrovirus a été particulièrement étudié. Ce modèle de rétrotransposition est valable à la fois pour les rétrovirus et les rétrotransposons.

L'amorçage du cycle de la transcription inverse est assuré par la fixation d'un ARNt dans la partie U5 du LTR 5'. La synthèse par la transcriptase inverse s'effectue jusqu'à l'extrémité 5' du LTR, ce qui produit une molécule d'ADN d'environ 200 pb. La transcriptase inverse, par son activité ARNase H, dégrade alors l'ARN complémentaire jusqu'à la partie R. La partie 3' de la copie d'ADN ainsi libérée s'hybride alors sur la partie complémentaire du LTR 3'. La synthèse du premier brin d'ADN est ainsi complète. L'activité ADN polymérase de la transcriptase inverse permet la synthèse du deuxième brin d'ADN. On obtient une copie

d'ADN linéaire à deux LTR qui s'intégrera dans le génome. Dans certains cas, une copie circulaire à deux LTR est obtenue mais elle ne semble pas pouvoir se réintégrer. L'intégrase coupe de façon décalée au site d'insertion et assure la ligation des extrémités aux deux LTR, générant ainsi une duplication du site cible de part et d'autre des LTR.

Les études des rétrotransposons montrent qu'ils possèdent des caractéristiques générales semblables à celles décrites précédemment pour les rétrovirus. Les éléments Ty sont présents dans le génome de la levure à raison de 30 copies pour Ty1 et 6 copies pour Ty917. De plus, environ 100 copies de LTR indépendants (appelés « d solos ») sont détectées. Les Ty complets font environ 6 kb de long. Deux ARNm, dans deux phases de lecture différentes, sont produits. Un ARN de 5,0 kb code pour les protéines TyA (Gag) et TyB (Pol). L'élément Ty1 est détecté sous forme de VLP (particules de type virale) qui ne sont pas infectieuses mais qui comprennent des molécules d'ARN, d'ADN double brin et les protéines transcriptase inverse et intégrase.

Chez la drosophile, une vingtaine de familles différentes de rétrotransposons sont présentes dans le génome. Elles sont dites de type *copia* d'après le nom de la première famille décrite. Chaque famille d'éléments comprend 10 à 60 copies par génome, le nombre variant selon les espèces et les souches. L'élément *copia* a une longueur de 5 000 pb comprenant deux LTR de 276 pb et présente un cadre de lecture ouvert de 4 227 pb codant les protéines Gag et Env. Il ne présente pas de caractère infectieux. Par contre, chez *Drosophila melanogaster*, il a été démontré récemment que l'élément gypsy est un rétrotransposon infectieux; c'est le premier rétrovirus décrit chez les insectes (Malik et al 2005).

III.4.2 Les rétrotransposons humains sans LTR

Ces éléments ont également une phase ARN dans leur cycle de transposition. Ils diffèrent des rétrotransposons car ils ne possèdent pas de LTR. Les principaux représentants font partie des superfamilles «*Long Interspersed Nuclear Elements*» (LINE) et «*Short Interspersed Nuclear Elements*» (SINE). Les LINE sont autonomes et codent des protéines similaires à *gag* et *pol* alors que les SINE sont non-autonomes mais possèdent des promoteurs pour l'ARN polymérase III (Feschotte et al., 2002 a,b; Sabot et al.,2005; Wicker et al., 2007).

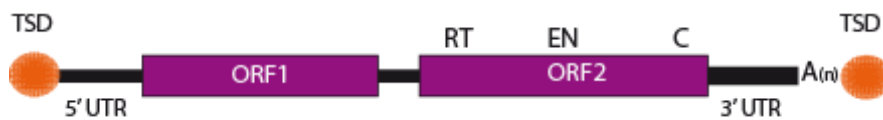
III.4.2.1 Les éléments L1

III.4.2.1.1 Généralités

Au sein du génome humain, il existe plus de 500 000 copies d'éléments mobiles de type LINE (L1), résultant de leur mobilisation continue au cours des 150 millions d'années d'évolution qui ont conduit à l'espèce humaine depuis son apparition. Ils représentent à eux seuls près de 17% du génome humain, ce qui fait d'eux les éléments transposables humains les plus représentés. Les éléments L1 complets ont une longueur d'environ 6 kb. Cependant, la majorité des éléments détectés dans les génomes ont subi une délétion du côté 5' et ne sont plus capables de transposer. On estime aujourd'hui qu'environ trois à quatre mille éléments L1 sont intégrés de façon complète et que seulement quarante à soixante sont toujours actifs. A l'inverse, on estime qu'il existe plus de trois mille éléments mobiles de type L1 actifs chez la souris (Deberardinis, et al 1998 ; Naas et al 1998). A titre d'exemple chez l'homme, l'insertion *de novo* d'un élément L1 dans le gène du facteur VIII a conduit à des mutations qui se sont traduites par une hémophilie A et une autre insertion dans le gène *myc* est associée à un type de cancer du sein.

III.4.2.1.2 Structure

Les éléments L1 ont une taille d'environ 6kb, et sont organisés de la manière suivante. Ils possèdent en 5' une région non traduite UTR (*untranslated region*) qui contient un promoteur nécessaire au recrutement de l'ARN polymérase de type II (Swergold et al 1990), deux ORF et une région 3'UTR contenant signal de poly-adenylation de taille variable (Babushok et al ,2007). Le premier cadre ouvert de lecture, ORF1, code pour une protéine de 40 kDa capable de se lier à l'ARN. Le second cadre ouvert de lecture code pour une protéine pouvant remplir deux rôles, une fonction de transcriptase inverse permettant la synthèse du brin complémentaire et une fonction d'endonucléase nécessaire à la coupure du brin d'ADN cible de la transposition. L'ensemble de ces éléments est impliqué dans le mécanisme de transposition de l'élément L1, faisant de celui-ci le seul élément mobile humain autonome (Figure 26).



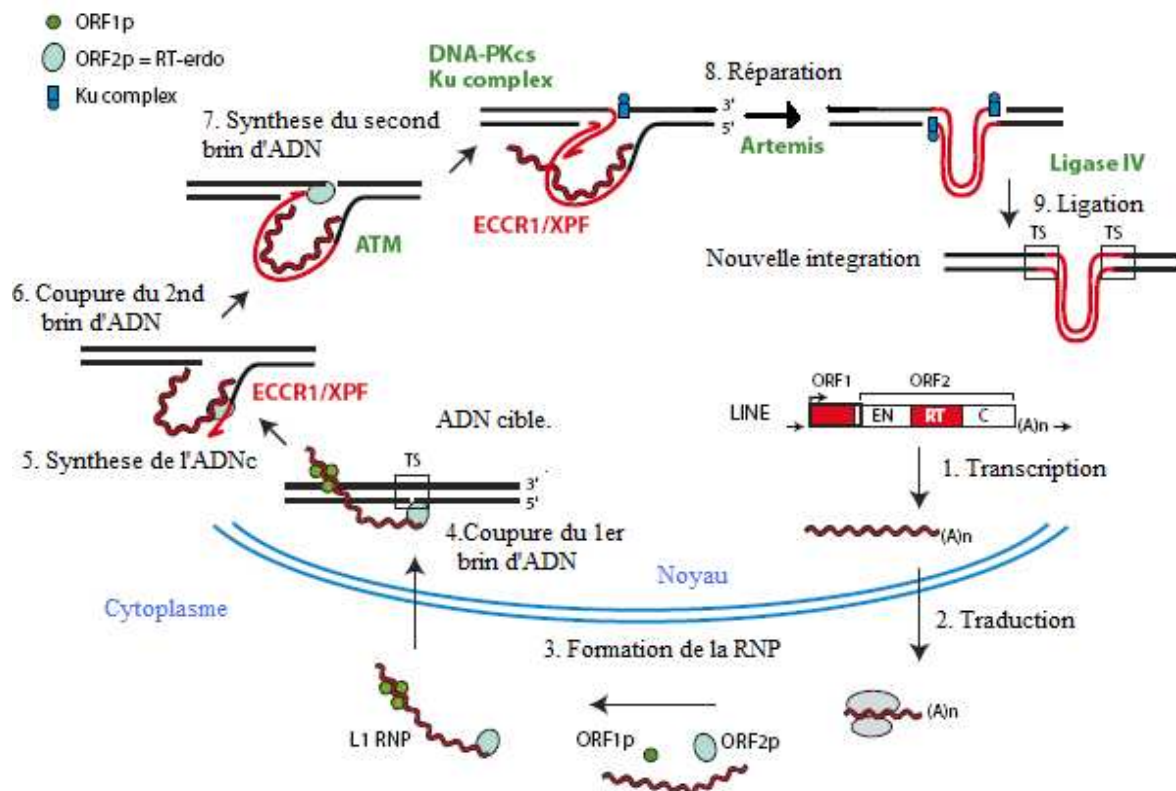
LINE 1 (6kb)

Figure 26: Représentation schématique de la structure des éléments Line humains. Sur la figure, sont représentés les sites dupliques de l'insertion (TSD *target site duplication*), les régions non traduites en 5' et 3' du transcrit (5'UTR et 3'UTR). Les deux ORF sont indiqués en magenta. L'ORF1, code pour une protéine de 40 kDa permettant la liaison à l'ARN. L'ORF2 code pour une protéine ayant une fonction de reverse transcriptase (RT) et une fonction d'endonucléase (EN) ainsi qu'un domaine en doigt de Zinc (C).

III.4.2.1.3 Mécanisme de transposition

L'importance du nombre de copies d'élément L1 hébergé par le génome humain est la résultante d'un mécanisme basé sur la duplication de l'élément L1 *via* un intermédiaire ARN appelé rétrotransposition (Figure 27).

La première étape de la rétrotransposition est la transcription de celui-ci par l'ARN polymérase de type II qui reconnaît un site de fixation situé sur le promoteur lui-même présent dans la région 5' UTR (Lavie et al 2004, swergold et al 1990). L'ARN est ensuite exporté dans le cytoplasme de la cellule où les ORF1 et 2 sont traduites. Chacune des protéines a une forte affinité pour l'ARN et peut s'y associer pour former une RNP (ribonucleoprotein particle) (Doucet et al 2001, Wei et al 2001). La RNP retourne ensuite dans le noyau par un mécanisme encore très mal compris.



d'après A.Beauregard et al 2008

Figure 27: Représentation schématique des différentes étapes de la rétrotransposition des éléments L1.

L'intégration à proprement parler de l'élément L1 se fait via un mécanisme appelé TPRT (*target primed reverse transcription*) (Moran et al 1996, , Cost et al 2002), qui fut pour la première fois décrit pour l'élément R2BM présent chez le vers *Bombyx Mori* (Luan et al 1993).

A ce jour, les étapes et le mécanisme exact du TPRT ne sont pas complètement élucidés. Il peut être néanmoins décrit de la manière suivante : il débute par le clivage du brin négatif d'ADN, grâce à l'activité endonucléase portée par l'ORF 2, le site de coupure ayant généralement lieu au niveau de la séquence consensus 5'TTTT /AAA3'. Le 3'OH ainsi libéré sert d'amorce à la transcription inverse initiée par cette même ORF. Le second brin d'ADN est alors clivé à son tour pour permettre la synthèse du brin complémentaire. La ligation des brins et la réparation du site de clivage restent les étapes les moins bien comprises, mais il est fortement supposé que cette étape fait appelle à la machinerie de réparation cellulaire (Gilbert et al 2005).

Après intégration, des marqueurs de transposition peuvent être observés. Il peut s'agir d'une troncation de la partie 5' de l'élément L1 attribué au décrochage de la RT au moment

de la transcription du 1^{er} brin. On observe aussi une queue poly A en 3' de l'élément intégré. Enfin, le clivage de l'ADN n'étant pas franc, il y a une duplication de la séquence d'insertion, cette duplication pouvant varier de 2 à 20 nucléotides.

III.4.2.2 Les éléments non autonomes : Sines. L'exemple des Alu

Il existe plus d'un million de copies d'élément Alu dans le génome humain. Ce nombre important est le résultat de leur mobilisation continue au travers de 65 millions d'année d'évolution. Les éléments Alu sont donc les éléments transposables les plus représentés dans le génome humain.

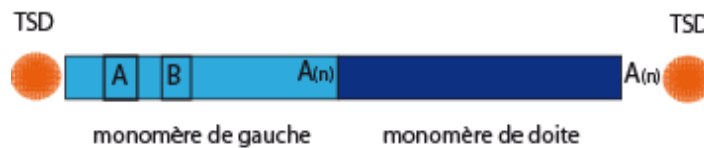


Figure 28: représentation schématique de la structure des éléments Alu humains

Un élément Alu est d'une longueur de 300 pb, sa structure est formée par deux régions répétées (monomères) d'environ 120 pb résultant de la fusion de deux séquences dérivées du gène codant pour l'ARN 7SL (figure 28). Chaque séquence est séparée par une région riche en adénines de 60 pb. La région 5' contient un promoteur pour l'ARN polymérase III (RNAPIII). L'élément se termine par une région de longueur variable, de type oligo(dA). Comme les éléments Alu ne possèdent pas des signaux de terminaison pour la RNAPIII, la transcription de l'élément Alu peut s'étendre jusqu'à un terminateur de la séquence en aval (généralement une série de quatre ou plus thymines consécutives). A l'inverse de l'élément LINE, Les éléments Alu ne codent pas pour une ORF et sont donc pas des rétrotransposons autonomes. Pour leur mobilité, les éléments Alu détournent la machinerie moléculaire de rétrotransposition codée par les éléments L1. Le mécanisme leur permettant d'utiliser la machinerie des elementL1 est à ce jour très mal compris(López-Flore et al 2010).

III.4.2.3 Les éléments SVA

Les éléments SVA (*SINE-VNTR-Alu*) ont été actifs tout au long des 25 millions d'année d'évolution des hominoïdes. A ce jour, on en compte environ 3000 exemplaires dans le génome humain dont 42% sont complets.

Un élément SVA complet est d'une longueur d'environ 2 kb et est composé par 5 régions successives (Figure 29).

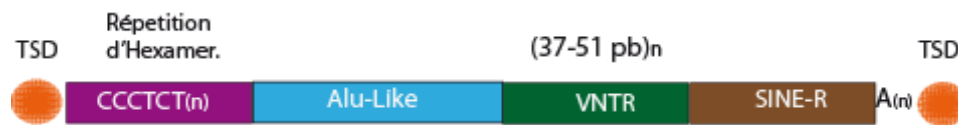


Figure 29:représentation schématique de la structure des éléments SVA humains. Sur la figure sont représentés les séquences hexamériques (CCCTCT) pouvant être répétées n fois, la séquence dite Alu-like, la région dite VNTR (*variable number of tandem repeat*) et la séquence appelée SINE-R.

On retrouve une séquence hexamérique (CCCTCT) pouvant être répétée n fois, une séquence dite Alu-like composée par deux fragment antisens, suivie par une petite séquence d'origine inconnue, une région dite VNTR (*variable number of tandem repeat*) composée de répétitions de séquences de 35 à 51 nucléotides, puis on retrouve une séquence appelé SINE-R d'environ 490bp, dérivée de la partie 3' du gène *env* des rétrovirus suivie d'une partie du LTR des rétrovirus endogène HERV-K10. Enfin on retrouve une région polyA. L'ensemble de cette séquence est flanquée par deux régions dupliquées (Wang et al 2005). Les éléments SVA ne contiennent apparemment pas de promoteur interne et pourraient s'appuyer, au moins en partie, sur l'activité de promoteur des régions flanquantes pour leur transcription. Comme les éléments Alu, les éléments SVA sont non-autonomes et sont sans doute mobilisés par la machinerie de rétrotransposition des éléments L1.

III.5 Les transposons à ADN

Les éléments de classe II se caractérisent par la nécessité de produire une copie d'ADN pour transposer. Ce sont des composants de génomes aussi variés que ceux des bactéries, plantes, insectes, poissons et mammifères. Ils codent pour une enzyme, la transposase, qui possède toutes les activités catalytiques nécessaires à la transposition de ces éléments. De plus, ils sont caractérisés par la présence d'ITR (*inverted terminal repeat*) en orientation inverse, à chaque extrémité du gène de la transposase.

Le mécanisme de transposition des transposons a été particulièrement étudié chez les bactéries. Les travaux réalisés sur les transposons eucaryotes en dérivent essentiellement.

Chez les bactéries, on détecte les unités transposables les plus simples d'organisation, les séquences d'insertion (IS). Chez *E. coli*, environ 10 copies de chaque famille d'IS sont présentes dans le génome. Les IS sont des unités autonomes codant pour les protéines nécessaires à leur propre transposition. Chaque IS ne possède qu'un cadre de lecture ouvert codant pour la transposase. Les familles d'IS sont différentes en séquence mais présentent les mêmes caractéristiques dans leur organisation. Les IS sont de courtes séquences d'environ 1000 pb présentant des répétitions inversées terminales (ITR) de 10 à 40 pb et, au point d'insertion, une duplication de la séquence cible. La séquence du site d'insertion varie pour une même IS, mais pas sa longueur. Cette séquence est dupliquée lors de l'intégration d'une IS, ce qui conduit à en trouver une copie à chaque extrémité de l'élément. Les répétitions inversées sont indispensables à la fixation de la transposase.

La transposition est déterminée par une ou plusieurs protéines spécifiques de chaque élément transposable. Tous les éléments transposables comprennent dans le segment d'ADN mobilisé un gène codant une protéine essentielle, la transposase. Les éléments transposables sont donc des éléments autonomes parasites des génomes dans lesquels ils se propagent. La transposase va reconnaître deux catégories distinctes de sites, d'une part les extrémités de l'élément et d'autre part un site cible (non spécifique), situé à un autre locus du génome. La recombinaison transpositionnelle repose donc sur un mécanisme complexe qui fait intervenir trois sites distincts séparés dans l'espace. De plus, l'événement de recombinaison n'est pas réciproque, et est en général suivi d'une étape de synthèse d'ADN consécutive aux étapes de coupure et ligation essentielles pour l'échange des brins d'ADN.

OBJECTIF DU TRAVAIL DE THESE

Ainsi que mentionné à la fin du chapitre 2, certaines propriétés du LAT_{2kb} et la multiplicité des phénotypes associés aux mutants viraux dans le locus du LAT nous ont fait penser que le LAT pouvait être un élément mobile.

Le but principal de cette étude a donc été de démontrer cette capacité du LAT_{2kb} à jouer le rôle d'élément mobile. Nous avons tout d'abord cherché à mettre en évidence sa capacité à transposer dans le gène RL2 au cours d'une infection latente. Le choix du gène *RL2* comme cible du LAT_{2kb}, faisait suite à deux observations réalisées dans le laboratoire. Premièrement, il avait été montré que les transcrits *RL2* ne sont pas maturés correctement dans le ganglion trigéminé. Deuxièmement, l'accumulation des LATs dans ces ganglions était corrélée avec l'accumulation de transcrits non épissés du gène *RL2*.

Une fois que nous avons montré l'insertion du LAT dans le gène RL2, la seconde partie de ce travail a été consacrée à l'utilisation d'approches globales afin d'identifier les diverses cibles du LAT_{2kb}.

Chapitre IV :

Matériels et méthodes

III.1 Rationnel de note choix de modèle animal

III.1.1 Les prérequis du modèle animal idéal

Si le virus HSV-1 est naturellement un agent pathogène de la seule espèce humaine, le large spectre d'hôte potentiel permet l'utilisation de plusieurs types de modèle animal pour l'étude de la pathogenèse virale, de la neuropathologie, de la neuroinvasion et de l'infection latente. Un modèle animal idéal serait en mesure de recréer tous les aspects de la maladie observée chez l'être humain, depuis la primo-infection jusqu'aux périodes de récurrences clinique en passant par les phases d'infection latente. Jusqu'à présent, ce modèle animal idéal n'a pas été caractérisé. Il existe en revanche plusieurs types de modélisation de l'infection herpétique chez l'animal (en faisant varier l'espèce hôte, la souche virale et la méthode d'inoculation), chacun d'entre eux permettant d'analyser une ou plusieurs aspects de la pathogénie de l'infection herpétique.

Toujours de manière idéale, le modèle animal le plus approprié pour l'étude de l'infection latente doit permettre d'observer une phase d'infection totalement quiescente à la fois cliniquement et biologiquement, puis d'obtenir des réactivations dans des conditions similaires à celles observées chez l'être humain. Chez un animal immunocompétent, ces réactivations devraient donc être à la fois spontanées mais rares, tout en étant inductibles par des stimuli tel que le stress. En outre, la primo-infection devrait être assez peu agressive pour limiter les séquelles tissulaires au niveau du site d'inoculation et dans les tissus neurologiques concernés par la période de dissémination initiale du virus, comme c'est le cas chez l'homme.

III.1.2 Les différents modèles utilisés de façon historique

Les deux espèces animales les plus fréquemment utilisées pour les expérimentations sur la pathogénie de HSV-1 sont le lapin et la souris. Le cochon d'Inde est quant à lui la référence dans les études sur HSV-2. Ces trois modèles animaux ont fourni à la communauté scientifique la grande majorité des données actuellement disponibles sur la pathogénie des virus HSV-1 et 2, notamment au regard des mécanismes impliqués dans la latence et la réactivation virale.

Le lapin présente l'avantage d'être un animal de taille suffisamment conséquente pour une observation clinique aisée et pour la réalisation de gestes chirurgicaux, le cas échéant. En revanche, l'infection herpétique induite chez le lapin est marquée par une tendance à la réactivation quasi-permanente (absence d'infection latente proprement dite) et une reproductibilité limitée des signes cliniques et biologiques. Ce dernier point est notamment dû à l'absence de souches consanguines permettant de limiter la variabilité interindividuelle. De même, les outils biologiques à disposition pour les études chez le lapin (anticorps monoclonaux par exemple) sont souvent moins fournis que pour d'autres espèces. Enfin, le lapin est un animal encombrant et couteux, ce qui limite rapidement le nombre d'individus incluables dans les expérimentations.

La souris présente des avantages et inconvénients en tous points opposés à ceux de lapin. Sa petite taille et le faible coût de son élevage permettent d'envisager des études avec un nombre d'individus satisfaisant pour les tests statistiques. De plus, les outils biologiques disponibles et l'existence de lignées consanguines facilite nettement les protocoles expérimentaux. Enfin, l'infection herpétique chez la souris entraîne rapidement une infection latente (mais potentiellement réactivable), à condition toutefois de faire le bon choix de la souche virale et des conditions d'inoculation (Crepin et al, Soumis). Pour toutes ces raisons, la souris est de loin l'hôte le plus communément utilisé dans les études sur les différentes phases de l'infection herpétique.

La première démonstration directe de la capacité de HSV-1 à établir et maintenir une infection latente dans les cellules neuronales a été démontrée chez des souris inoculées dans les coussinets plantaires. Après réplication virale au site d'inoculation, l'infection latente était observée dans les ganglions spinaux de la moelle épinière (Wagner EK, et Bloom DC. 1997). C'est aussi dans ce modèle qu'ont été effectuées les premières études sur le transport du virus le long des axones. Le modèle d'infection plantaire a donc permis de poser les bases de la compréhension des mécanismes impliqués dans la neurovirulence. Ce modèle est désormais plutôt utilisé pour l'étude de HSV-2, en raison du tropisme plus important de ce virus pour les ganglions spinaux que le virus HSV-1, dont le principal site de latence naturel est le ganglion trigéminal (TG), situé dans la base du crâne.

Le modèle murin le plus utilisé dans littérature est de loin est celui dit de scarification cornéenne (CS), dans lequel l'instillation du virus sur la cornée est réalisée après scarification

de l'épithélium. Grâce à cette manœuvre, les particules virales sont adsorbées par les cellules cornéennes et les terminaisons des fibres sensibles intracornéennes (d'origine trigéminal). Des antigènes viraux sont d'ailleurs détectables dans la cornée entre 2 et 8 jours post-inoculation (Crepin et al, Soumis). Après pénétration dans la terminaison des fibres nerveuses, la capside virale remonte le long de l'axone par transport rétrograde, à la vitesse d'au moins 1cm par jour, pour atteindre le noyau des neurones du TG homolatéral où débute un nouveau cycle viral. Des antigènes viraux sont détectables dans le TG dès le 4^{ème} jour post-inoculation (pi), plus en abondance à partir du 6^{ème} jour, puis cela décroît par la suite, preuve de l'entrée en latence du virus. Il existe aussi une réduction de la quantité d'ADN viral entre le 6^{ème} et le 28^{ème} jour, ce qui traduit aussi une réduction, voire une extinction du processus réplcatif. Cette période est aussi caractérisée par une augmentation de la quantité des LAT (Crepin et al, Soumis). Cependant, selon la souche virale utilisée, il peut persister au 28^{ème} jour post-infection certains marqueurs de la réplication virale, notamment des ARN messagers codant des protéines virales précoces, voire tardives. Ces données concourent d'ailleurs à remettre en question la réalité d'une extinction complète de la réplication virale, donc d'un réel état d'infection latente, dans le modèle CS au 28^{ème} jour post-inoculation, en particulier avec des souches virales comme KOS (Crepin et al, Soumis). Le modèle CS ne paraît donc pas le plus adapté pour avoir une idée précise de ce qu'il se passe réellement chez l'homme où les analyses de ganglion trigéminé post-mortem ont essentiellement retrouvé du l'ADN viral et des LAT, sans expression de transcrits codant pour des protéines virales (Mitchell et al, 2003). D'une façon plus générale d'ailleurs, une des principales limites des modèles animaux est la difficulté à reproduire ce qu'il se passe réellement chez l'homme. En particulier, de nombreux isolats cliniques de HSV-1 sont extrêmement neurovirulents et labiles lorsqu'ils sont inoculés chez l'animal, ou à l'inverse sont non pathogènes. Finalement, le nombre de souches de HSV-1 potentiellement à disposition en expérimentation animale courante est de l'ordre d'une quinzaine, ce qui est relativement faible si l'on compare à la diversité des génotypes réellement rencontrés chez l'homme.

Afin de disposer d'un modèle animal qui soit en mesure de reproduire un peu plus fidèlement la maladie herpétique telle qu'observée chez l'être humain, notre équipe a caractérisé dans les années 2000 un nouveau modèle murin dit « oro-oculaire » (Labetoulle et al, 2000, Son nom est lié au fait que la primo-infection est réalisée dans la lèvre (pour se rapprocher du principal site de primo-infection chez l'homme (Liesegang et al, 1989)), et qu'elle entraîne rapidement une infection herpétique aiguë de l'iris et de la cornée de l'œil

homolatéral, avant d'induire une infection latente dans les ganglions trigéminés, les ganglions cervicaux supérieurs, la moelle, le noyau moteur facial (Labetoulle et al, 2003), autant de localisations décrites chez l'homme comme sites de latence herpétique (Mitchell et al, 2003). Même si l'infection oculaire chez l'homme n'est quasiment jamais concomitante de la primo-infection, mais plutôt retardée de quelques années (Liesegang et al, 1989), ce modèle murin reproduit en de nombreux points chez la souris les voies de propagation du virus dans le système nerveux, depuis le site de primo-infection le plus fréquent (les voies oro-pharyngées) jusqu'aux principaux sites de latence observés chez l'homme. Pour toutes ces raisons, nous avons privilégié l'utilisation de ce modèle pour nos travaux plutôt que celui de la scarification cornéenne.

III.1.3 Description précise du modèle murin oro-oculaire

L'inoculation de la suspension virale est réalisée dans la lèvre gauche de la souris. Une lésion vésiculaire labiale, comparable à un bouton de fièvre, est observée au 2^{ème} jour pi en regard du site d'inoculation. Au 4^{ème} jour pi, une infection généralisée apparaît chez toutes les souris, caractérisée par un hérississement du poil et une perte de poids. Une infection oculaire, essentiellement de l'iris, du corps ciliaire, et de la cornée est observable chez les deux tiers des souris à 6 jours pi., dont la moitié (soit un tiers de l'effectif total) décède entre le 6^{ème} et le 10^{ème} jours pi, alors que les symptômes régressent chez les autres souris jusqu'à disparaître complètement. L'analyse systématique des tissus de souris euthanasiées pendant les 10 premiers jours de l'infection a montré que la propagation du virus HSV-1 se fait depuis la lèvre dans plusieurs voies neuronales de façon parallèle (Labetoulle et al, 2000).

La première chaîne neuronale positive pour l'expression d'antigènes viraux est la voie autonome sympathique. Depuis la lèvre, le virus emprunte les terminaisons nerveuses sympathiques jusqu'au ganglion cervical supérieur (SCG) gauche, premiers relais sympathique. Après réplication dans les neurones du SCG, qu'il infecte totalement, le virus remonte la chaîne sympathique pour gagner la colonne intermédiolatérale gauche (IML) de la moelle épinière, avant de passer dans l'IML droite entre le 8^{ème} et 10^{ème} jour chez un tiers des animaux.

En parallèle, le virus emprunte les terminaisons sensibles depuis la lèvre gauche jusqu'au ganglion trigéminal (TG) gauche, siège de l'innervation sensitive de la face, qu'il atteint entre le 4^{ème} et le 6^{ème} pi, pour se répliquer dans les corps neuronaux avant de se propager dans le noyau du tractus spinal trigéminal de l'hypothalamus, la zona incerta, le thalamus ventro-postérieur, qui sont donc les relais secondaires de l'infection du système sensitif. Enfin, le virus injecté dans la lèvre se propage aussi via les terminaisons nerveuses motrices situées dans la lèvre jusqu'au noyau moteur facial, responsable de l'innervation motrice d'une grande partie des muscles de la face.

Toutes ces données sur la propagation de HSV-1 depuis la lèvre sont parfaitement expliquées par les propriétés de transfert antérograde ou rétrograde du virus au sein d'un réseau de neurones connectés entre eux par des synapses, car on sait que les particules virales se déplacent depuis les noyaux des neurones jusqu'à leurs extrémités en empruntant les réseaux architectoniques des dendrites et des axones. Cependant, le virus HSV-1 a aussi une capacité de transfert local, c'est à dire de passer entre deux neurones adjacents situés dans une même structure anatomique mais non connectés entre eux par une synapse (Labetoulle et al, 2000).

Dans le modèle oro-oculaire, le virus peut ainsi se propager dans d'autres réseaux neuronaux que les seules voies sympathiques, sensibles et motrices décrites précédemment. Ainsi, le long des voies sensibles, les particules virales fabriquées dans les corps des neurones du TG dont les terminaisons innervent la lèvre peuvent aussi se propager à d'autres neurones sensitifs tels que ceux innervant le segment antérieur de l'œil. Cela explique l'infection irienne et cornéenne observée entre le 6^{ème} et le 8^{ème} jour pi. De même, les particules virales qui suivent les neurones depuis le TG jusqu'à l'œil via les voies sensibles rencontrent les neurones parasympathiques lors de leur passage dans le ganglion ciliaire, lui aussi infecté à 6 jours pi. Le transfert local dans cette structure explique les atteintes du noyau d'Edinger-Westphal puis du noyau paraventriculaire et du colliculus supérieur observées ce modèle. Un schéma simplifié de ces voies de propagation est proposé dans le schéma ci-dessous.

En complément des données de propagation virale obtenues par analyse immunohistologique de l'expression des antigènes viraux, le modèle oro-oculaire a aussi été largement utilisé pour analyser l'évolution des marqueurs biologiques moléculaires de l'infection virale. Ces données seront largement reprises dans les deux articles présentés dans

ce manuscrit. En bref, entre la phase aigu (6^{ème} jour pi) et latente (28^{ème} jour pi), il fut observé une extinction des transcrits précoces codant la thymidine kinase et des transcrits tardifs codant UL18, comme attendu lors de la mise en place de l'infection latente. Parallèlement était observé une réduction du nombre de copie d'ADN génomique dans toutes les structures et une augmentation des LAT dans les TG, alors que les quantités de LAT avaient plutôt tendance à diminuer dans les autres structures testées (moelle, SCG, hypothalamus) (Labetoulle et al, 2003, Maillet et al, 2006, Crepin et al, soumis).

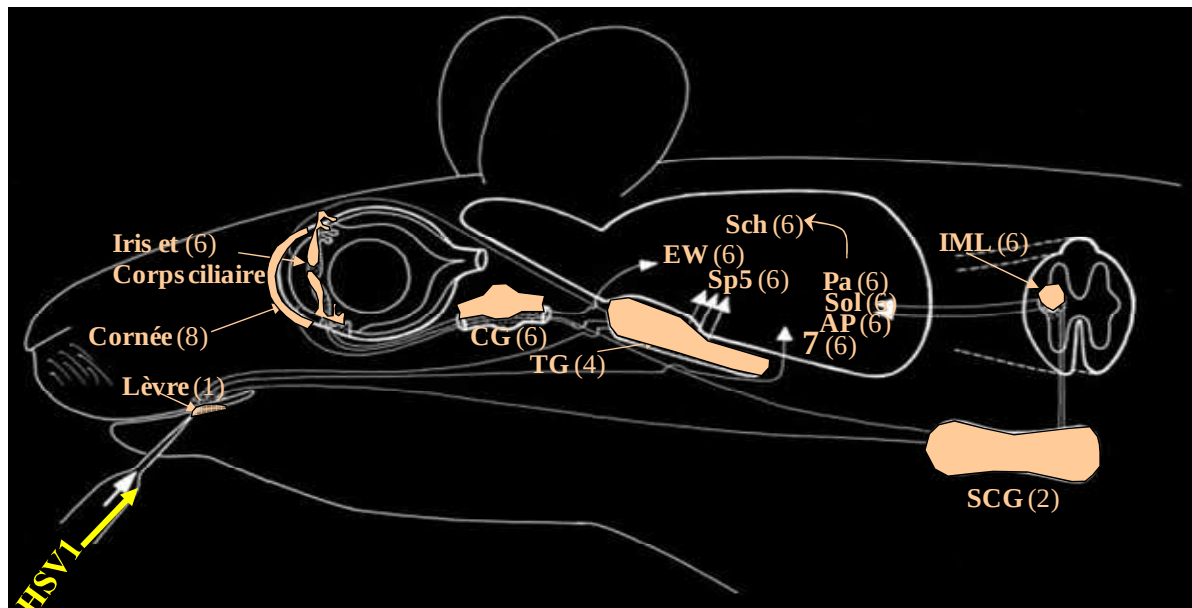


Figure 30: Représentation de la propagation du virus HSV-1 après injection dans la lèvre d'une suspension virale de la souche SC16 du virus HSV-1 au cours du temps. (adapté de Labetoulle et al., 2003). AP : area postrema ; CG : ganglion ciliaire ; EW : noyau d'Edinger-Westphal ; TG : ganglion Trigéminal ; IML : colonne intermédiolatérale de la moelle ; Pa : noyau para ventriculaire ; SCG : ganglion cervical supérieur ; Sch : noyau supra-chiasmatique ; Sol : noyau du tractus solitaire ; SP5 : noyau du tractus spinal trigéminal ; 7 : noyau du nerf facial.

III.2 Type de souris utilisées

Nous avons utilisé des souris consanguines BALB/c femelles de 6 semaines. L'âge de 6 semaines a été choisi en raison de l'existence d'un plateau de sensibilité à l'infection herpétique entre les âges de 4 et 12 à 14 semaines. Si les souris sont infectées trop jeunes, le risque est celui d'une trop grande sensibilité conduisant à la mort rapide des individus, ce qui est incompatible avec la compréhension des mécanismes de mise en place de l'infection latente. A l'inverse, l'utilisation d'animaux trop âgés est liée à une moindre reproductibilité de la sensibilité à l'infection, et donc à la nécessité d'utiliser plus d'individus pour obtenir le nombre suffisants d'animaux réellement infectés. La souche de souris Balb/c a été choisie

pour sa susceptibilité à l'infection par HSV1 (Lopez et al, 1975, Pepose et Whittum-Hudson 1987), ce qui permet de réduire le nombre d'individus utilisés. De même, la consanguinité des souris a permis de réduire les variations inter-individuelles (Lopez et al, 1975) et donc le nombre d'individus à utiliser pour obtenir des résultats fiables. Enfin, notre choix était guidé par le fait que le modèle d'infection oro-oculaire a été initialement caractérisé avec cette souche de souris, ce qui permettait, encore une fois, de réduire le nombre d'individus inclus dans les études au plus strict minimum pour obtenir des résultats analysables sur le nombre puisque les doses de virus pour induire tel ou tel type d'infection avaient été préalablement étalonnées (Labetoulle et al, 2003).

Toutes nos démarches expérimentales ont tenté de répondre à la consigne de « réduction » mise en valeur dans les règles d'éthique de l'expérimentation animale. De même, tous les animaux ont été nourris et abreuvés sans restriction tout au long de l'étude, conformément aux standards actuels des règles d'éthique, en particulier celles de l'*Association for Research in Vision and Ophthalmology (ARVO) Statement for the Use of Animals in Research*.

III.3 Souche virale

Les résultats présentés dans ce travail de thèse ont tous été réalisés avec la souche SC16 de HSV-1. Il s'agit d'une souche sauvage, isolée en Angleterre à partir de lésions herpétiques buccales d'un patient (Hill 1975). Cette souche est considérée comme hautement neurovirulente (Shimeld et al, 1985, Ugolini et al, 1987, Ugolini et al 1992). Son principal avantage est son caractère sauvage, c'est à dire l'absence de modifications génomiques majeures liées à des processus d'adaptation aux conditions de laboratoires. Elle est donc supposée refléter, au moins en partie, la virulence réelle des souches de HSV-1 humaines, à la différence d'autres souches utilisées pourtant de façon courante par un grand nombre d'équipes de recherche. C'est le cas par exemple de la souche KOS, très souvent utilisée dans le domaine de la recherche sur les mécanismes de régulation de l'infection herpétique latente. On sait en effet que la souche KOS est caractérisée par des mutations dans la glycoprotéine B (fortement impliqué dans les mécanismes de reconnaissance du récepteur) (Yuhasz et al, 1993, Mitchell et al, 1996), le gène gamma 34.5 (impliqué dans la neurovirulence), et le gène US9 (impliqué dans le transport des particules virales dans les neurones) (Negatsch et al, 2011). A titre d'exemple, dans un modèle d'infection rétinienne testé par notre équipe, $1,5 \times 10^4$ à 2×10^5 PFU de la souche KOS sont nécessaires pour induire une rétinite chez 19% à

57% des souris (Atherton et al, 1987, Pepose et al, 1987, Metzger et al, 1987) alors que 90 PFU de la souche SC16 sont suffisantes pour obtenir 100% d'infection rétinienne (Labetoulle et al, 2000). De la même façon, 70 PFU de la souche SC16 suffisent à induire le même taux d'atteinte oculaire dans le modèle de scarification cornéenne que 10^6 PFU de la souche KOS (Crépin et al, soumis).

III.4 Préparation des concentrés viraux

Des cellules BSR (clone de lignée cellulaire BHK 21, issues de rein de hamster nouveau-né) amenées en confluence en culture (boîtes Falcon, Becton-Dickinson, USA) ont été infectées à une multiplicité d'infection de 0,05 PFU (*Plaque Forming Unity*). Les cellules et le surnageant de culture ont été récoltés après 72 heures de culture, puis soumises à une centrifugation de 4500 tours/min (tpm, rotor JA 14, Beckman, USA) pendant 30 minutes à 4°C. Le culot a été soumis à des ultrasons (sonication) pendant 30 secondes puis à nouveau centrifugé dans les mêmes conditions. Le surnageant de la première centrifugation a été rassemblé avec le surnageant de la seconde, et le tout a été déposé dans un tube à centrifugation sur un coussin de TNE (Tris-HCL pH 7,5 :10mM, NaCl 50mM, EDTA : 1mM) + 25% glycérol. Après centrifugation à 27000 tpm pendant 1 h à 4°C (rotor sw28, Beckman), le culot a été suspendu dans du tampon de dilution (NaCl : 137mM, KCL : 5mM, Na₂HPO₄ : 0,7mM, Tris-HCL PH7,5 :25mM) puis conservé à -80°C.

III.5 Titrage du virus

Les virus purifiés ont été titrés selon la méthode des plages de lyse. Des dilutions croissantes de suspension virale (200µl/puits) ont été incubées 1 heure à 37°C sur des cellules Vero (cellules de rein de singe vert du Gabon) menées à confluence en culture monocouche dans des plaques de titrage à 6 puits (Multiwell, Falcon, Becton-Dickinson™). Les dilutions ont été calculées de manière à obtenir un nombre significatif de plage (20 à 200) dans plusieurs des 6 puits de la boîte de titrage. Après l'adsorption du virus sur les cellules puis rinçage, un gel de MEM + agarose 0,4% enrichi avec 2% de sérum de veau fœtal a été coulé dans chaque puits, puis les boîtes ont été mises à incuber à 37°C pendant 3 jours. Les tapis cellulaires ont été colorés au rouge neutre 0,1% (Gibco™) et les plages de lyse cellulaire ont été comptées. Des titres de 10^9 à 5×10^9 UFP/ml ont été obtenus selon les sessions de purification.

III.6 Inoculation des animaux

Toutes les inoculations ont été réalisées sous anesthésie générale par injection intrapéritonéale d'équithésine (pentobarbital sodique : 1 %, hydrate de chloral : 270mM, propylène glycol : 40%, MgSO₄ : 86 mM) à la dose de 40 µl/10g de poids de l'animal. L'injection a été réalisée dans la lèvre supérieure gauche, à la jonction cutanéomuqueuse, en utilisant une 1µl de suspension virale, à l'aide d'une micropipette en verre reliée à un système d'injection mécanisé (Infusionpump, Harvard™, Royaume-Uni).

III.7 Sacrifice des animaux

Le sacrifice des animaux a aussi été réalisé sous anesthésie générale à dose létale, par injection intrapéritonéale d'une solution de pentobarbital sodique à 54,7 mg/ml (dose injectée : 2mg/10g de poids). Après obtention de l'anesthésie complète, une sternotomie était pratiquée pour mise en place d'une perfusion intracardiaque (aiguille de 21 gauges, Terumo™, Belgique avec 20 mL de tampon PBS (phosphate buffer saline, NaCl : 137mM, KCl : 27mM ; Na₂HPO₄ : 101mM ; KH₂PO₄ : 17,6mM, PH 7,4), puis 150 ml de tampon PBS+ 4% de para formaldéhyde (Prolabo™, France) et enfin 60 ml de tampon PBS + 20% de saccharose (Prolabo™). Le débit de perfusion était réglé sur 10 ml/minute. Les souris ont ensuite été disséquées sous loupe binoculaire.

Pour les analyses en biologie moléculaire, les prélèvements (TG, SCG, moelle, hypothalamus) ont été placés pendant 24h à 4°C dans 10 volumes de RNA Later (Qiagen™), puis congelés à -80°C après retrait de la solution de RNA Later.

Pour les études histologiques, les crânes entiers et les colonnes vertébrales ont été placés pendant 10 jours à 4°C dans un tampon PBS + EDTA 0.1 M pour décalcification, puis dans un tampon PBS + 20% de saccharose pendant 2 jours pour protection des tissus au moment de la congélation. Les tissus ont ensuite été enrobés de milieu de congélation (Tissutek, Thermo Scientific™, USA) avant d'être stockés à -80°C.

III.8 Extraction des ARN totaux à partir des biopsies de souris

L'extraction des ADN totaux a été menée à partir de prélèvements tissulaires (TG, SCG, moelle, hypothalamus) à l'aide du RNeasy Plus Mini Kit (Qiagen™), conformément aux instructions fournies par le distributeur. Les ARN ainsi extraits ont ensuite été congelés à -80°C pour être stockés avant leur utilisation.

III.9 Extraction des ADN totaux à partir des biopsies de souris

L'extraction des ARN totaux a été menée à partir de prélèvements tissulaires (TG, SCG, moelle, hypothalamus) en plaçant ces derniers dans 170µl de tampon phosphate salin (PBS, NaCl 5g/l, KCl 0,2g/l, Na₂HPO₄ 1,44g/l, KH₂PO₄ 0,24g/l) et 10µL de Rnase A (Qiagen™) à la concentration de 7000 unités/ml. Après broyat de l'ensemble dans un pot, 20 µl de protéinase K (Qiagen™) à 600 mAU/ml était ajoutés avant incubation pendant 1h à 55°C. L'ADN total était alors extrait par la technique de phénol/chloroforme selon les étapes suivantes: i) ajout au produit de digestion de 1 volume de phénol/chloroforme/alcool isoamylique (25 :41 :1) (Fluka™) ; ii) agitation lente de haut en bas pour ne pas casser l'ADN puis centrifugation à 12000 tpm pendant 8 minutes ; iii) extraction de la phase aqueuse sur laquelle est répétée l'opération ; iv) extraction de la phase aqueuse et ajout de 1 volume de chloroforme, puis répétition de l'étape de centrifugation ; v) précipitation de l'ADN dans un mélange d'éthanol absolu et d'acétate de sodium 3M, pH = 5,7, pendant une nuit à -80°C ; vi) centrifugation pendant 20 min à 14000 tpm puis reprise du culot dans 1ml d'éthanol 70% pour éliminer les restes de sel ; vii) nouvelle centrifugation pendant 20 min à 14000 tpm puis reprise du culot dans 30µl d'eau microfiltrée avant stockage à -20°C.

III.10 Extraction des ADN viraux à partir de cellules infectées en culture

Des cellules BSR (clone de lignée cellulaire BHK 21, issues de rein de hamster nouveau-né) amenées en confluence en culture (boîtes Falcon, Becton-Dickinson, USA) ont été infectées à une multiplicité d'infection de 0,05 PFU (*Plaque Forming Unity*). Les cellules et le surnageant de culture ont été récoltés après 72 heures de culture, puis soumises à une centrifugation de 4500 tours/min (tpm, rotor JA 14, Beckman, USA) pendant 30 minutes à 4°C. Le culot a été soumis à des ultrasons (sonication) pendant 30 secondes puis à nouveau centrifugé dans les mêmes conditions. Le surnageant de la première centrifugation a été rassemblé avec le surnageant de la seconde, et le tout est re-suspendu dans un tampon TNE (50mM Tris (pH 7.5), 140mM NaCl, 5mM EDTA). Une précipitation au phénol/chloroforme a alors été réalisée selon les étapes i à vii décrites précédemment.

III.11 Les différentes techniques de *Polymerase Chain Reaction* (PCR) utilisées au cours de ce travail

III.11.1 RT-PCR

L'acronyme RT-PCR signifie *Reverse Transcriptase* PCR, c'est à dire PCR après transcription inverse d'un acide ribonucléique (ARN) en ADN complémentaire (ADNc). La RT-PCR est donc une technique qui associe une transcription inverse (RT) suivie d'une PCR. Elle permet de synthétiser le brin complémentaire d'un ARN avec des désoxyribonucléotides en utilisant une ADN polymérase ARN dépendante (transcriptase inverse).

Au cours de cette étude, 1 µl d'extrait d'ARN totaux ont été utilisés comme matrice et on a ajouté au mélange du *One Step RT PCR kit* (Qiagen™). Les réactions d'amplification ont été effectuées dans un volume final de 50µL suivant les conditions décrites ci-dessous :

Composant	Volume	Concentration finale
H2O	20µL	
5x Qiagen reaction mix	10µL	1x
dNTP (10mM par dNTP)	2µL	400 µM par dNTP
Amorce sens	2,5 µL	0,6 µM
Amorce reverse	2,5 µL	0,6 µM
Enzyme MIX	2µL	
Q-sol	10µl	1x
ARN	1µL	10 à 50 ng
Volume final	50µL	

Les réactions de PCR ont été réalisées dans un thermocycleur (T3000 Biometra™) suivant les conditions d'amplification suivantes :

Température	Temps	Nombre de cycles
50° C	45 min	1
95°C	15 min	1
95°C	30 sec	45
Variable °C	30 sec	
72 °C	1 min	
72 ° C	10 min	1

La température d'hybridation des amorces était définie pour chaque. Le contrôle de la qualité d'amplification a été réalisé par dépôt de 2 µl du produit final sur un gel d'agarose à 1,2 %. Après contrôle, l'intégralité des produits de PCR étaient déposée sur gel d'agarose à 1,2 % pour extraire les amplicons d'intérêt dans le but de les séquencer.

III.11.2 Touch-down-PCR

La *Touch-down PCR* est une technique développée pour amplifier une séquence d'ADN faiblement représentée et/ou pour laquelle il pourrait exister une compétition pour l'hybridation des amorces. Le principe est d'utiliser une température d'hybridation des amorces très haute lors des premiers cycles afin d'assurer une forte stringence et donc une amplification spécifique du produit souhaité. Une fois que la séquence d'intérêt devient majoritaire vis-à-vis de ses compétiteurs, la température d'hybridation est progressivement abaissée pour augmenter l'efficacité de la PCR proprement dite. L'efficacité de cette dernière n'est donc pas constante tout au long de la réaction, il est très difficile (sinon impossible) d'obtenir une quantification très précise.

Des quantités de 10 à 100 ng d'ADN extraits de TG de souris infectées ont été utilisées comme matrice dans un mélange réactionnel suivant les conditions décrites dans le tableau ci-dessous :

Composant	Volume	Concentration finale
H2O	21,75 – 26,75 μL	
10x PCR buffer	5μL	1x
10x Coraload	5μL	1x
dNTP(10mM par dNTP)	1μL	200μM par dNTP
Amorce sens	1 μL	0,2 μM
Amorce reverse	1 μL	0,2 μM
TopTaq DNA polymerase	0,25 μL	1,25 unité
Q-sol	10μl	1x
ADN	1-5 μL	10 à 100 ng
Volume final	50μL	

Les réactions de PCR ont été réalisées dans un thermocycleur (T3000 Biometra™) suivant les conditions d'amplification suivantes :

Température	Temps	Nombre de cycles
94° C	3 min	1
95°C	15 min	1
95°C	30 sec	10 avec diminution de 1°c du TM à chaque cycle
TM +10 °C	30 sec	
72 °C	1 min	
95°C	30 sec	35
Variable °C	30 sec	
72 °C	1 min	
72 ° C	10 min	1

La température d'hybridation des amorces était définie pour chaque expérience .Le contrôle de la qualité d'amplification a été réalisé par dépôt de 2 µl du produit final sur un gel d'agarose à 1,2%. Après contrôle, l'intégralité des produits de PCR étaient déposée sur gel d'agarose à 1,2% pour isoler les amplicons d'intérêt dans le but de les séquencer.

III.11.3 i-PCR

La PCR inverse (i-PCR) a été initialement décrite par Mr Ochman dès 1988 (Ochman et al, 1988) : elle permet l'amplification des régions flanquant un gène (ou une portion de gène) donné mais pour lesquelles aucune information n'est connue, sinon la séquence nucléotidique du gène (ou de la portion de gène) qu'elles entourent(Figure 31). De façon générale, l'intérêt de la i-PCR est donc d'identifier une ou des régions situées à proximité d'une séquence d'intérêt. Cette technique est basée sur une PCR conventionnelle à l'exception des amorces choisies pour reconnaître les deux extrémités de la séquence d'intérêt, qui sont orientées dans le sens inverse de l'orientation ordinaire (c'est à dire « vers » les régions flanquantes). Préalablement à leur hybridation, l'ADN total est digéré par une enzyme de restriction choisie de telle manière à ne pas couper dans la séquence d'intérêt. Les portions d'ADN ainsi coupées sont ensuite liguées pour circularisation, et donc, seuls quelques produits de ligation contiennent la séquence d'intérêt bordée par les régions flanquantes à déterminer. L'orientation particulière des amorces permet alors, en une seule réaction, de n'amplifier que ces produits de ligation particuliers.

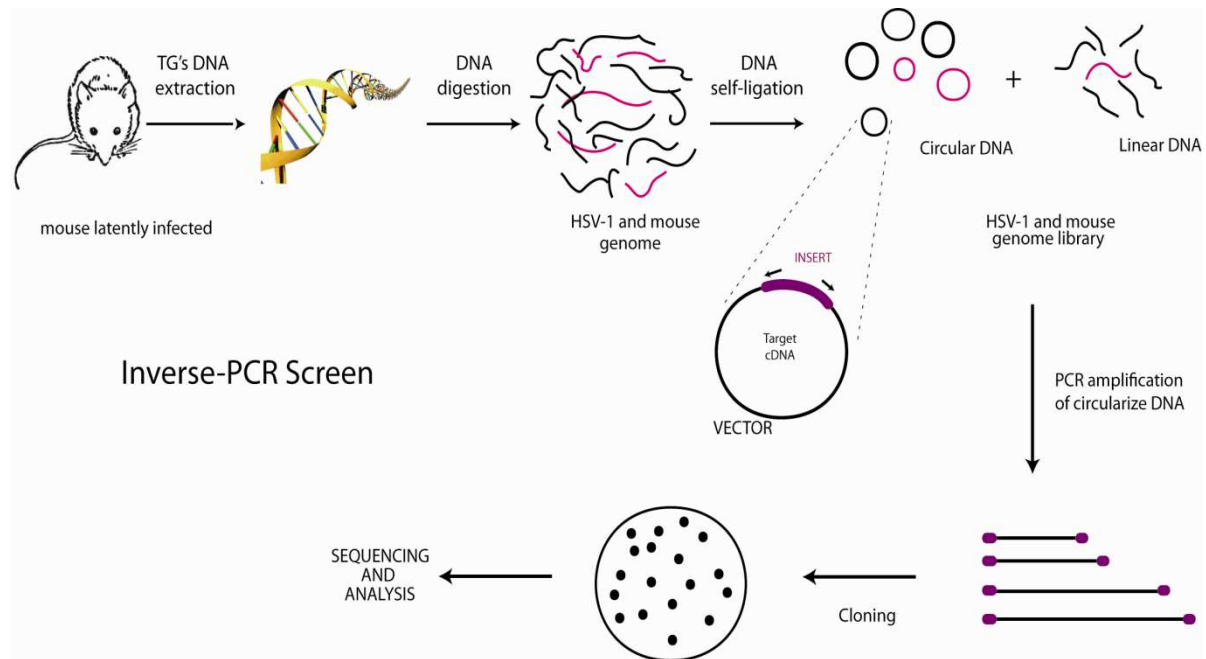


Figure 31: Principe de la PCR inverse appliqué à la recherche des sites d'insertions du LAT dans le génome murin. A partir de ganglion trigéminal de souris, l'ADN génomique est extrait, puis digéré, les fragments d'ADN sont ensuite ligués sur eux-mêmes. Une première PCR est effectuée l'aide d'un couple d'amorces spécifique de l'élément mobile intégré dans le génome de l'hôte. Une seconde PCR dite « nichée » est réalisée, avec un autre couple d'amorces. Les fragments obtenus sont ensuite purifiés sur gel et clonés. Les plasmides de chaque clones sont alors isolés et séquencés.

Pour réaliser nos expériences, 1µg d'ADN total extrait des biopsies de TG issus de souris infectées a été digéré par 20U d'enzyme de restriction pendant 3h à 37°C (5 types de digestions différentes ont été faites pour chaque extrait d'ADN, en utilisant différentes enzymes de restriction). Après purification des produits de digestion, la ligation a été réalisée par incubation pendant 18h à 16°C de 1ng/µl d'ADN digéré dans un mélange de T4 DNA ligase à 4U/µl (New England Biolabs™), Tris-HCl 50mM (pH 7.5), MgCl₂ 10mM, dithiothreitol 10mM, ATP 1.5mM, et BSA (bovine serum albumin) 25µg/ml. Pour éliminer les fragments d'ADN non circularisés, les produits de la réaction de ligation ont été soumis à l'action de l'Exonuclease V (New England Biolabs™) à 10U/µL avec 1mM/µL d'ATP. Les produits obtenus ont été purifiés avec le kit NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel™).

Par la suite, 5 µL des produits de ligation ont été ajoutés au mélange réactionnel décrit ci-dessous :

Composant	Volume	Concentration finale
H2O	29,5µL	
5x PCR buffer	10 µL	1x
dNTP(10mM par dNTP)	1 µL	200µM par dNTP
Amorce sens	2,5 µL	0, 5 µM
Amorce reverse	2,5 µL	0, 5 µM
PHUSION DNA polymerase	0,5 µL	1,25 U
ADN	5 µL	N .D
Volume final	50µL	

Les réactions de PCR ont été réalisées dans un thermocycleur (T3000 Biometra™) suivant les conditions d'amplification suivantes :

Température	Temps	Nombre de cycles
98°C	3 min	1
98°C	15 sec	34
Variable °C	15 sec	
72 °C	2 min	
72 ° C	5 min	1

Les produits de PCR obtenus ont alors été dilués au 1/25^{ème}, puis une aliquote de 5µL de chaque produit a été ajouté dans le même mélange réactionnel que précédemment mais en utilisant un nouveau couple d'amorces, située en aval des amorces précédentes (selon le principe de la PCR nichée).

Les secondes réactions de PCR ont été réalisées dans un thermocycleur (T3000 Biometra™) suivant les conditions d'amplification suivantes :

Température	Temps	Nombre de cycles
98°C	3 min	1
98°C	15 sec	29
Variable °C	15 sec	
72 °C	2 min	
72 ° C	5 min	1

Le contrôle de la qualité d'amplification a été réalisé par dépôt de 2 µl du produit final sur un gel d'agarose à 0,8%. Après contrôle, l'intégralité des produits de PCR était déposée sur gel d'agarose à 0,8 % pour isoler les amplicons d'intérêt dans le but de les séquencer.

III.11.4 q-PCR

La PCR quantitative (q-PCR), dite aussi « en temps réel », est une application particulière de la PCR conventionnelle. Elle permet de suivre en continu (en temps réel) le processus d'amplification en détectant la fluorescence émise par les produits de PCR néoformés. La q-PCR exploite une propriété inhérente à l'amplification en chaîne: plus il y a de molécules cibles à l'origine, plus le nombre de cycles d'amplification nécessaire pour atteindre un nombre déterminé de molécules amplifiées est réduit. Si on intègre dans la réaction de PCR un fluorophore dont l'énergie émise ne sera détectable que lorsqu'un nombre fixe et reproductible de copies de la matrice est atteint, on peut alors mesurer pour chaque échantillon le nombre de cycles de PCR à réaliser pour passer au dessus du seuil de détection. Ce nombre de cycle est communément nommé Ct (pour *cycle threshold*), et il est inversement corrélé au logarithme du nombre de copies de la matrice initialement présente (figure 32). Ainsi, il est possible de déterminer très précisément le nombre initial de copies par interpolation, en comparant le Ct observé par rapport à celui d'extraits du même produit mais dont les concentrations étaient connues.

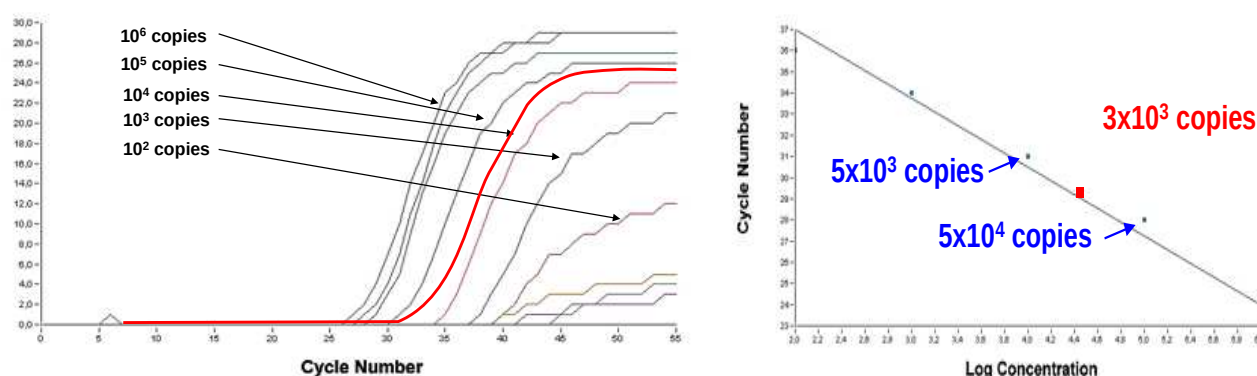


Figure 32: Exemple de q-PCR avec les courbes de fluorescence et la construction de la droite de régression nécessaire à la quantification du produit d'intérêt (en rouge)

La première partie de nos expériences en q-PCR a donc consistée à créer plusieurs gammes étalons des produits que nous désirions quantifier. De façon générale, la quantification des différents échantillons a été menée en triplicate pour chaque extrait.

III.11.4.1 Quantification des produits contenant les LAT

Deux constructions plasmidiques ont été réalisées à cet effet. Dans un premiers temps, les deux locus des LAT ont été clonés dans le vecteur pJET2.1. Pour cela, 2 µg d'ADN viral génomique ont été purifiés (voir paragraphe 10 ci-dessus), puis digérés par l'enzyme Nde1 (New England Biolabs™) avant séparation des fragments en gel d'agarose. Deux bandes principales étaient observées. Elles ont été extraites et purifiées séparément avec le kit NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel™).

Le produit correspondant à la 1^{ère} bande (taille de 12132 paires de base, pb) a été digéré par l'enzyme DraI (New England Biolabs™) avant séparation des fragments en gel d'agarose. Le produit de digestion d'une taille de 7985 bp était alors extrait puis purifié avec le kit *NucleoSpin® Gel and PCR Clean-up* (Macherey-Nagel™). Ce produit, correspondant au premier locus des LAT, ensuite été cloné dans le vecteur pJET2.1 à l'aide du kit *CloneJET PCR Cloning Kit* (Fermentas France™) et la séquence a été vérifiée.

La seconde bande résultant de la digestion de l'ADN génomique par Nde1, d'une taille de plus de 15kb et correspondant au reste du génome viral, a aussi été soumise, après purification, à l'action de l'enzyme DraI. Après séparation des produits de digestion en gel d'agarose, la bande correspondant à la taille de 7985 bp a été extraite et purifiée avec fut alors extrait avec le kit *NucleoSpin® Gel and PCR Clean-up* (Macherey-Nagel™). Ce produit, correspondant au second locus des LAT, a aussi été cloné dans le vecteur pJET2.1 à l'aide du kit *CloneJET PCR Cloning Kit* (Fermentas France™) et la séquence a été vérifiée. Finalement, aucune différence de séquence n'était observée entre les deux loci des LAT, ce qui a permis de réaliser une gamme unique pour la quantification des séquences d'ADN amplifié contenant tout ou partie des LAT.

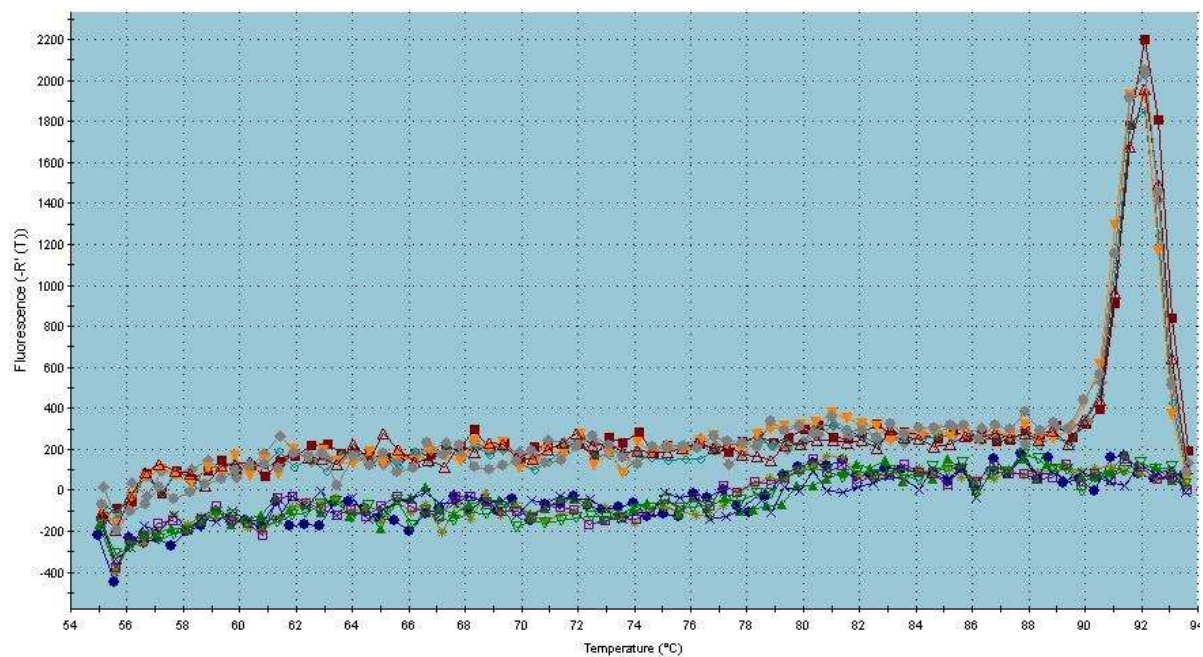
Pour obtenir cette gamme, l'une des deux constructions plasmidiques a ensuite été amplifiée, purifiée, et la concentration finale d'ADN a été mesurée à l'aide d'un Nanodrop 1000 (Thermo Fisher™), puis des séries de dilution par pas de 10 ont été réalisées pour obtenir la gamme requise pour la quantification. Chaque dilution a été réalisée dans une solution contenant 10 ng/µL d'ADN murin, pour placer les ADN étalons et les extraits à quantifier, issues de biopsies, dans des conditions identiques au moment de l'amplification. Pour la q-PCR proprement dite, 2 µl de chaque dilution ont été utilisés comme matrice dans le mélange *QuantiFast SYBR Green PCR master mix* (Qiagen™). Les réactions d'amplification ont été effectuées dans un volume final de 50µL suivant les conditions décrites ci-dessous :

Composant	Volume	Concentration finale
H2O	5,5 µL	
2x QuantiFast SYBR master mix	12,5µL	1x
Amorce sens	2,5 µL	1µM
Amorce reverse	2,5 µL	1 µM
ADN	2µL	variable
Volume finale	50µL	

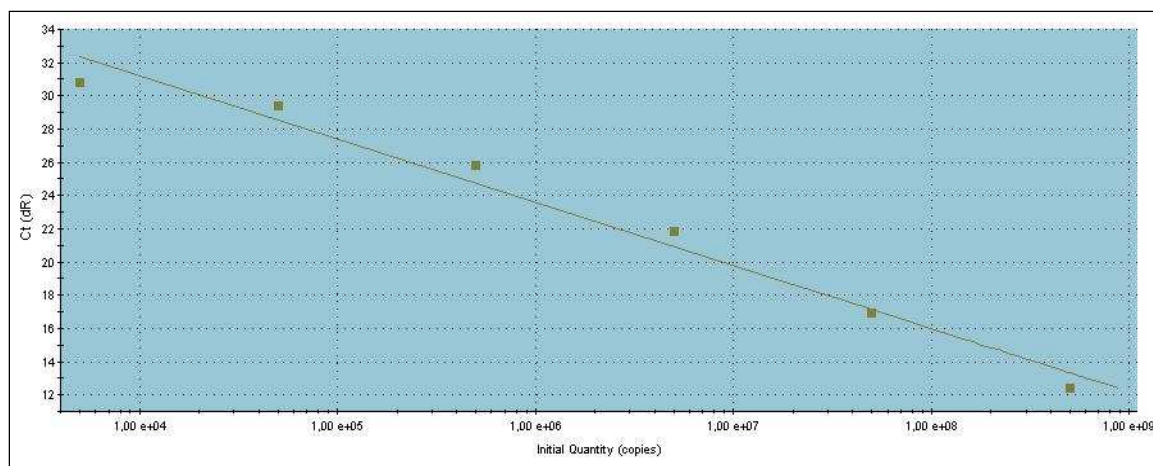
Les réactions de PCR ont été réalisées dans un thermocycleur (Applied Biosystems 7300) suivant les conditions d'amplification suivantes :

Température	Temps	Nombre de cycle
95° C	5 min	1 cycle
94° C	10 sec	40 cycles
60°C	30 sec	
65°c à 98°c	Analyse de la courbe de fusion des produits de PCR.	

L'étude de la courbe de fusion permet d'obtenir un graphe montrant la décroissance de fluorescence en fonction de la température. En réalisant la dérivée première de ces courbes, on obtient un chromatogramme permettant de suivre l'évolution de la croissance (ou de la décroissance) de la fluorescence en fonction de la température. On obtient ainsi des pics de modification de fluorescence qui correspondent à la formation de l'amplicon. Les courbes de fusion obtenues sont reportées dans le graphique suivant :



Pour les dilutions allant de 10^{-2} à 10^{-7} , la présence d'un seul pic à la même température montrait qu'une seule cible avait été amplifiée, ce qui prouvait la spécificité de la réaction et validait l'utilisation des ces dilutions pour la fabrication de la gamme. Pour la dilution de 10^{-1} , une inhibition de la PCR était observée en raison d'une trop grande quantité de matériel en début réaction. A l'inverse, pour les dilutions inférieure à 10^{-7} , des dimères d'amorces, (amorces qui se reconnaissent entre elles pour s'hybrider et réaliser des amplicons non spécifiques) se formaient en raison de la trop faible quantité de matrice (courbe non montrée). L'absence de dimère d'amorce ainsi que la spécificité de la réaction nous a permis de tracer la droite de régression de la gamme étalon nécessaire pour l'estimation, par interpolation, des quantités de produits recherchées au sein de nos échantillons. Les moyennes des Ct en fonction de la quantité d'ADN initiale de tous les échantillons de gamme ont été reportées dans un repère semi-logarithmique. La droite d'étalonnage moyenne (ci-dessous) a été modélisée à l'aide d'une régression linéaire, de type $Y = aX + B$.



Dans le cas présent, la droite de régression ne passait pas au centre de la dispersion des différentes mesures pour les concentrations allant de 10^4 à 10^5 copies, on constatait un « amortissement » de la pente (points non tracés sur le graphique). L'imprécision sur la quantification à ces concentrations était tellement importante qu'une modélisation de « l'amortissement » était peu pertinente.

De façon générale, cet « amortissement » correspond, lorsque des gammes moyennes sont réalisées, à un abaissement de la pente, donc à une augmentation de l'efficacité de PCR, alors que l'accroissement des effets stochastiques devrait l'abaisser. Il faut noter alors que cet « amortissement » est généré principalement par la concentration la plus faible (dans notre étude 10 copies, point non matérialisé). Or aucune amplification ne pourra se faire s'il n'y a pas au moins une molécule d'ADN initial. Le biais qui en résulte dans la distribution gaussienne de l'erreur est susceptible de provoquer cet « amortissement moyen ». Il faut avoir conscience que si la gamme d'étalonnage démontre l'aspect quantitatif du protocole expérimental, il est difficile d'éviter tous les biais potentiels, comme une différence de composition chimique entre le solvant des échantillons et le diluant des points de la gamme.

L'utilisation de cette gamme ne nous a donc pas permis de quantifier des échantillons ayant un nombre de copies initial inférieur à 10^4 unités, ce qui correspondait dans notre cas, à un nombre de copies dans le mélange réactionnel de 200 unités ($1/50^{\text{ème}}$ de l'extrait initial est utilisé pour la réaction.).

III.11.4.2 Quantification des produits contenant la jonction 3' entre les LAT et le gène RL2

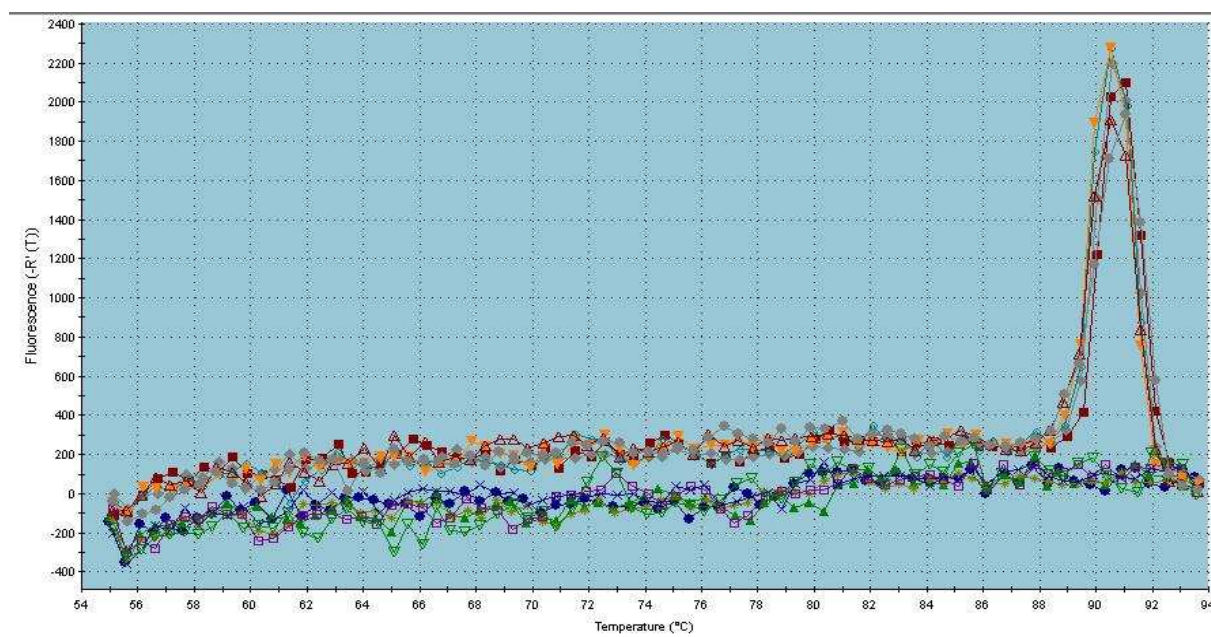
La gamme a été construite en clonant le produit de PCR de la jonction 3' entre le LAT de 2kb et le premier intron du gène RL2. Ce produit de PCR a été cloné dans le vecteur pJET2.1 à l'aide du kit *CloneJET PCR Cloning Kit* (Fermentas France™). Le plasmide correspondant a ensuite été amplifié, purifié, et la concentration finale a été mesurée à l'aide d'un Nanodrop 1000 (Thermo Fisher™). Une série de dilution au 1/10^{ème} dans une solution contenant 10 ng/μL d'ADN murin a alors été effectuée. Pour la q-PCR proprement dite, 2 μl de chaque dilution ont été utilisés comme matrice dans le mélange *QuantiFast SYBR Green PCR master mix* (Qiagen™). Les réactions d'amplification ont été effectuées suivant les conditions décrites ci-dessous :

Composant	Volume	Concentration finale
H2O	5,5 μL	
2x <i>QuantiFast SYBR master mix</i>	12,5 μL	1x
Amorce sens	2,5 μL	1 μM
Amorce reverse	2,5 μL	1 μM
ADN	2 μL	variable
Volume final	50 μL	

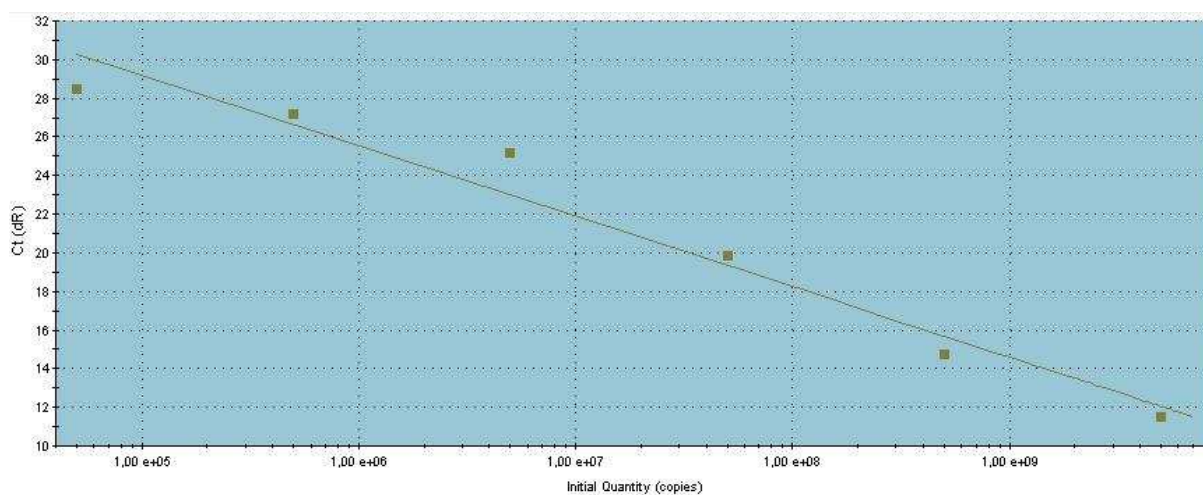
Les réactions de PCR ont été réalisées dans un thermocycleur (Applied Biosystems 7300) suivant les conditions d'amplification suivantes

Température	Temps	Nombre de cycle
95° C	5 min	1 cycle
94° C	10 sec	40 cycles
62° C	30sec	
65°c à 98°c	Analyse de la courbe de fusion des produits de PCR.	

Les courbes de fusion obtenues sont reportées dans le graphique suivant :



Pour les dilutions allant de 10^{-1} à 10^{-6} , la présence d'un seul pic à la même température montrait la spécificité de la réaction et validait son utilisation en tant que gamme. Pour les dilutions supérieures à 10^{-6} , des dimères d'amorces étaient observés. La droite de régression était donc la suivante :



Comme précédemment, la droite de régression ne passait pas au centre de la dispersion des différentes mesures, et on constatait un « amortissement » de la pente pour les concentrations allant de 10 à 10^5 copies (points non tracés sur le graphique), rendant non-fiable toute

extrapolation. L'utilisation de cette gamme ne nous a donc pas permis de quantifier des échantillons ayant un nombre de copies initial inférieur à 10^5 unités, ce qui correspondait dans notre cas à un nombre de copies dans l'extrait initial de 2000 unités ($1/50^{\text{ème}}$ de la solution initial est utilisé pour la réaction.).

III.11.4.3 Quantification des produits contenant le gene RL2 du virus HSV-1

Pour réaliser la quantification du nombre de copies de gène RL2, nous avons utilisé la même gamme étalon, que celle réalisée au cours des travaux précédents menés dans le laboratoire et publiés en 2006 (Maillet et al, 2006).

III.12 Considérations générales sur le séquençage des produits de PCR

Chaque produit de PCR obtenu au cours de cette thèse a été systématiquement analysé sur gel et séquencé. En bref, la totalité des produits amplifiés était déposée sur gel d'agarose, puis analysée sous lampe ultraviolette. Les bandes correspondant aux amplicons attendus (ou à ceux posant question) étaient découpées avant d'être purifiées à l'aide du kit NucleoSpin® Gel and PCR Clean-up (Macherey Nagel™) suivant les instructions fournies par le distributeur. Les produits purifiés étaient alors clonés dans le vecteur de séquence pJET2.1, à l'aide du kit CloneJET PCR Cloning Kit (Fermentas France™). Les produits de ligation ont été utilisés pour transformer des bactéries DH5α chimiquement compétentes (New England Biolabs™). Après une nuit de sélection sur boîtes de gélose LB contenant 100µg/ml ampicilline, chaque colonie a été amplifiée séparément dans 3ml de LB liquide contenant 100µg/ml ampicilline pendant 16h. Les plasmides ont alors été purifiés à l'aide du kit NucleoSpin Plasmid QuickPure (Macherey-Nagel™) suivant les instructions du distributeur. Les plasmides contenant l'insert attendu ont été adressés pour séquençage à la société GATC Biotech™ (Allemagne).

III.13 Technique du RACE (Rapid Amplification of cDNA Ends)

Cette technique est habituellement utilisée pour vérifier la nature du nucléotide +1 d'un ARN donné, afin d'être certain d'avoir isolé la totalité du bon produit. Pour cette raison, cette technique est souvent nommée 5'RACE. Dans nos travaux, nous avons utilisé le même principe mais pour analyser l'extrémité 5' des transcrits *RL2* non épissés. Cette technique, a permis de confirmer, par une autre technique que les PCR conventionnelles ou quantitatives, la présence d'un insert LAT dans l'intron 1 des transcrits *RL2* non-épissés ainsi que la présence du LAT dans des transcrits cellulaires.

III.14 Southern blot.

III.14.1 Digestion de l'ADN génomique

La réaction de digestion a été menée dans 50 µl contenant 10 µg d'ADN, 1,25 µl de spermidine, 5 µl du tampon adapté à l'enzyme et 1 µl d'enzyme. Le tout était incubé pendant 16 à 37°C. Les produits de restriction étaient précipités dans 100 µl d'éthanol à 95 % (v/v) et 5 µl d'acétate de sodium pendant une heure à -20°C. Après une heure de centrifugation à 13000 g (4°C), le culot était lavé deux fois par 500 µl d'éthanol 70 % (v/v), et séché.

III.14.2 Migration de l'ADN

Les échantillons de digestion ont été déposés à raison de 8 à 10µg par puits dans un gel d'agarose 0,8% ultra-pur, coulé dans du TAE 0,5X (1mg/ml de BET). La migration était réalisée sur 12h à un voltage constant de 20 volts.

III.14.3 Préparation du transfert de l'ADN

L'ADN présent dans le gel a tout d'abord été dénaturé dans du HCl 0,275M pendant 30 min, puis rincé à l'eau et neutralisé dans du NaOH 0,4M deux fois 10min, sous agitation constante. Pendant ce temps la membrane en nylon (Hybond N+, G.E Healthcare™) a été équilibrée dans un bain d'eau puis un bain contenant du NaOH 0,4M. Le transfert a duré 17h.

III.14.4 Marquage des sondes radioactives

Le marquage des sondes a été réalisé par amorçage aléatoire (*radon-priming ready to go DNA labelling beads*, G.E Healthcare™). L'ADNc, issu la PCR des LAT (50ng par

réaction), a été dénaturé pendant 10min à 95°C, avant que l'enzyme et le mélange de dNTPs soit ajouté, dont 50 à 100 microcuries de dCTP³². Après homogénéisation, la réaction de marquage était effectuée à 37°C pendant 20min. La sonde radiomarquée était purifiée des nucléotides marqués mais non incorporés à l'aide des colonnes *Illustra MicroSpin G-50* (G.E Healthcare™). Après agitation au vortex, la colonne était centrifugée pendant 1 min à 735G. L'éluât contenant la sonde marquée purifiée était finalement dénaturée par 5 µl de NaOH 3M pendant 5 min.

III.14.5 Transfert de l'ADN sur la membrane

Avant le transfert, l'ADN contenu dans le gel a été dépuriné par une incubation de ce dernier dans un bain d'HCl 0,25M pendant 15 min. Après rinçage, le gel était transféré pendant 30 min dans un bain de NaOH 0,4M pour dénaturer l'ADN, puis était réalisé le transfert sur une membrane de nylon chargée positivement (GeneScreen Plus, NEN™) grâce à l'appareil schématisé ci-dessous. Afin de conserver l'état de dénaturation de l'ADN, le transfert par capillarité se faisait en milieu alcalin (NaOH 0,4M). Pour pouvoir transférer l'ADN d'un même gel sur deux membranes à la fois, le gel était recouvert sur chacune de ses faces par, dans l'ordre (figure 33) :

- une membrane taillée un peu plus grand que le gel
- deux feuilles de papier de 3 mm d'épaisseur (*chromatography paper*, Whatman™) imbibées de soude 0,4M, également un peu plus grandes que le gel
- 6 cm de papier absorbant Cette superposition est recouverte d'une plaque surmontée d'un poids qui permettait un bon contact entre toutes les couches. Le phénomène de capillarité permettait à la soude de migrer du gel vers le papier absorbant, entraînant avec elle l'ADN jusqu'à la membrane. Une nuit suffit pour que tout l'ADN se soit fixé sur les membranes. Celles-ci sont alors rincées par deux bains successifs de SSC 1x.

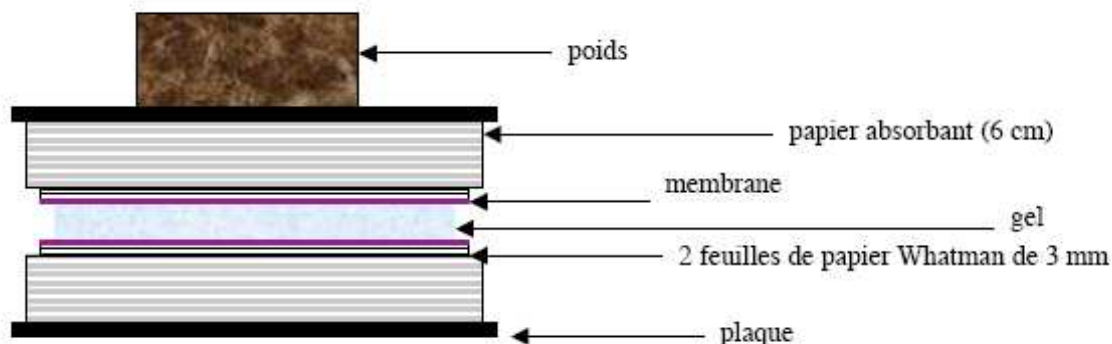


Figure 33 : Montage pour le transfert de l'ADN du gel sur les membranes.

III.14.6 Hybridation de la membrane

A l'aide de pinces plates, la membrane était placée dans un tube de verre rempli de 13 ml de tampon Church (1 % BSA; 1mM EDTA; 0,5 M NaHPO₄ pH 7,2 et 7 % SDS) et collée à la paroi. Le tube était mis au four à 62°C pendant au moins 2h. Le tampon Church était ensuite renouvelé et la sonde marquée et dénaturée était introduite dans le tube. L'hybridation était réalisée à 58°C pendant 24h. Les associations aspécifiques sonde-ADN étaient éliminées par deux lavages successifs, d'abord pendant 15 min avec le tampon LAV I (2x SSC, 0,1 % SDS (v/v)) puis pendant 5 min avec le tampon LAV II (0,2x SSC, 0,1 % SDS (v/v)). La membrane était ensuite séchée et emballée dans du film plastique.

III.14.7 Autoradiographie de la membrane

La membrane était placée à -80°C au contact d'un film photographique (Biomax MS film, Kodak™) entre deux écrans amplificateurs. Après 12 jours d'exposition, le film était révélé

Chapitre V :

Résultats 1^{er} partie

A new mobile element is encoded by HSV1

Nicolas HUOT^{1*}, David PASDELOUP¹, Thierry NAAS²,

Yves GAUDIN^{1*}, and Marc LABETOULLE.^{1,3*}

¹ *Laboratoire de Virologie Moléculaire et Structurale, Centre National de la Recherche Scientifique, 91198 Gif-sur-Yvette, France,*

² *Service de Bactériologie-Virologie, Hôpital de Bicêtre, Assistance Publique / Hôpitaux de Paris, Faculté de Médecine Paris-Sud, Université Paris XI, 94275 K.-Bicêtre, France,*

³ *Service d'Ophthalmologie, Hôpital de Bicêtre, Assistance Publique-Hôpitaux de Paris, 94275 Le Kremlin-³ Bicêtre, France*

Keywords: Mobile element, herpes simplex virus type 1, latency associated transcripts (LAT),

*To whom correspondence should be addressed. E-mails: Nicolas Huot (Nicolas.huot@vms.cnrs-gif.fr), Yves Gaudin (gaudin@vms.cnrs-gif.fr) and Marc Labetoulle (marc.labetoulle@bct.aphp.fr)

Introduction

Herpes simplex virus type 1 (HSV1) belongs to *Herpesviridae*, a family of viruses widely disseminated in animals that has the peculiar feature of being able to remain latent in their natural host (Roizman, 2001). HSV1 is a human neurotropic virus, responsible of life- and vision-threatening diseases due to episodes of viral reactivation.

HSV1 establishes lifelong latent infection in human neurons, especially in the trigeminal ganglia (TG), following a primary stage of propagation in the nervous system. After penetration into the cell, capsids migrate to the nuclear pore and DNA is ejected into the nucleus where the viral genome can either be replicated or persist in a latent state. By definition, no infectious particles can be detected in the host during latency and several studies failed to detect lytic viral gene expression during HSV1 latency (Roizman, 2001). During latent infection, the viral genome is circularized and maintained in a repressed state except for the region encoding the *Latency Associated Transcripts* (LATs) {Stevens et al,; Steiner et al, 1988 }, located in the inverted repeats of the viral genome (figure 1A). The *LAT* gene encodes an 8.3 kb primary transcript, which is spliced, leading to the formation of two stable LATs, LAT_{2kb} and LAT_{1.5kb} (2 kb and 1.5 kb in size, respectively, Figure 1B). Remarkably, LAT_{2kb} and LAT_{1.5kb} are stable introns {Farrell et al, 1991 }, as a consequence of a non-consensus branch-point {Mukerjee et al, 2004 } that results in the maintenance of a lariat structure {Wu et al, 1998; Krummenacher et al, 1997 a b }.

During latency, LAT_{2kb} and LAT_{1.5kb} are the most abundant viral transcripts and they accumulate with time {Roizman et al, 2011 }. Mutational analyses have shown the key role of LATs in the establishment of latent infection and in the reactivation process in a mouse model {Leib et al, 1989; Perng et al, 1994}. Specifically, (i) LATs expression is associated with a down regulation of the expression of immediate early genes {Garber et al, 1997; Mador et al, 1998}; (ii) LATs inhibit apoptosis, thus promoting neuronal survival {Perng et al, 2000; Ahmed et al, 2002; Branco et al, 2005}, and (iii) LATs interfere with both adaptive and innate immunity {Peng et al, 2005; Branco et al, 2005 ; Knickelbein et al, 2008 }. All these properties facilitate the implementation of a reservoir of latently infected neurons in the host {Perng et al, 2000; Thompson et al, 2001}. The molecular mechanisms behind these diverse properties are largely debated. Recently, it has been suggested that LATs encode miRNAs

{Umbach et al, 2008; Shen et al, 2009} and that they can also modify gene expression through chromatin remodeling {Wang et al, 2005; Bloom et al, 2010}.

Interestingly, the regions of the HSV1 genome that encode LATs also contain the *RL2* gene, encoding ICP0 (Figure 1B), a protein that is critical for initiating viral replication at low multiplicity of infection *in vitro* and for the experimental reactivation process *in vivo* (review in {Everett, 2000 }).

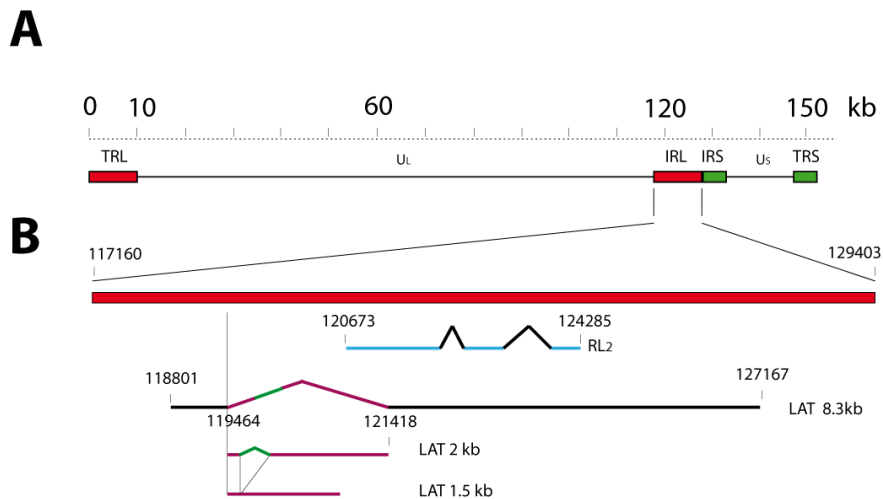


Figure 1. HSV-1 LATs region organization.

(A) HSV-1 genome is organized into two covalently-linked segments (a long and a short fragment). Each segment is composed of two unique sequences, U_L and U_S , flanked by two pairs of inverted repeats, referred as terminal repeats (TR_L and TR_S) or internal repeats (IR_L and IR_S). The final organization of the HSV1 genome can thus be summarized as $TR_L-U_L-IR_L-IR_S-U_S-TR_S$.

(B) Magnification of the IRL region, showing the position and orientation of *RL2* transcripts and of LATs. The LAT primary transcript (LAT_{8.3kb}) is spliced to generate the stable 2.0 and 1.5 kb LAT species. The *RL2* gene, encoding ICP0 protein of HSV1, is antisense to the LAT gene. Introns are indicated in magenta. The green segment in LAT_{2kb} is spliced to generate LAT_{1.5 kb}.

While it was initially assumed that viral proteins and viral transcripts (with the exception of LATs {Roizman, 2001 }) were absent during the latent phase of herpetic infection, RT-PCR assays showed that a sub-species of *RL2* transcripts, still containing intron 1, could be observed in the TG as late as 28 days post-inoculation (pi) {Chen et al, 2002 }.

We have previously developed a mouse oro-ocular model of HSV1 infection {Labetoulle et al, 2003} in which the inoculation of HSV-1 in the lip leads to an acute infection 6 days pi and to a latent infection in the TG 28 days pi. In this model, both intron 1-containing *RL2* transcripts (referred to as “unspliced” thereafter) and LATs accumulate in the TG during latency {Maillet et al, 2006 }.

In the course of identifying the nature of these unspliced *RL2* transcripts in infected tissues, we surprisingly found LAT_{2kb} sequences inserted in these transcripts and in the corresponding *RL2* gene. This revealed that LAT_{2kb} act as mobile elements. Furthermore, an approach aimed at identifying all the transcripts modified by LAT_{2kb} insertion revealed that a second viral gene, *US12* encoding the protein ICP47, was also targeted in a similar manner. Finally, taking advantage of the high frequency of B1 elements in the mouse genome, we performed B1-PCR and identified LAT_{2kb} insertion into the mouse genome. Therefore, LAT_{2kb} is a new type of mobile element which is able to modify both the viral and the host genomes. This unique property of LATs opens new horizons on HSV-1 biology particularly on latency regulation

Results

Part of LAT is found inserted in unspliced RL2 transcripts during latency

To determine the sequence of unspliced *RL2* transcripts from their 5' end, we used a primer extension assay with a primer located in intron 1 (n1RL2, corresponding to nucleotides 597 to 618 of the *RL2* gene, Figure 2A). Using RNA extracts from latently infected TGs (28 days pi), we found several products containing *RL2* gene sequences (Figure 2B and S1A). Among them, some had the expected exon1–intron1 junction and some others had a 5' end corresponding to LAT_{2kb} until nucleotide 1397 followed by the sequence of *RL2* gene starting at nucleotide 348 within the first intron. This unexpected sequence suggested that some *RL2* transcripts contained an insert corresponding to part of the LAT_{2kb}. An additional G was also observed downstream the junction, in a stretch of G nucleotides in the *RL2* gene (nucleotides 355 to 359) (Figure S1).

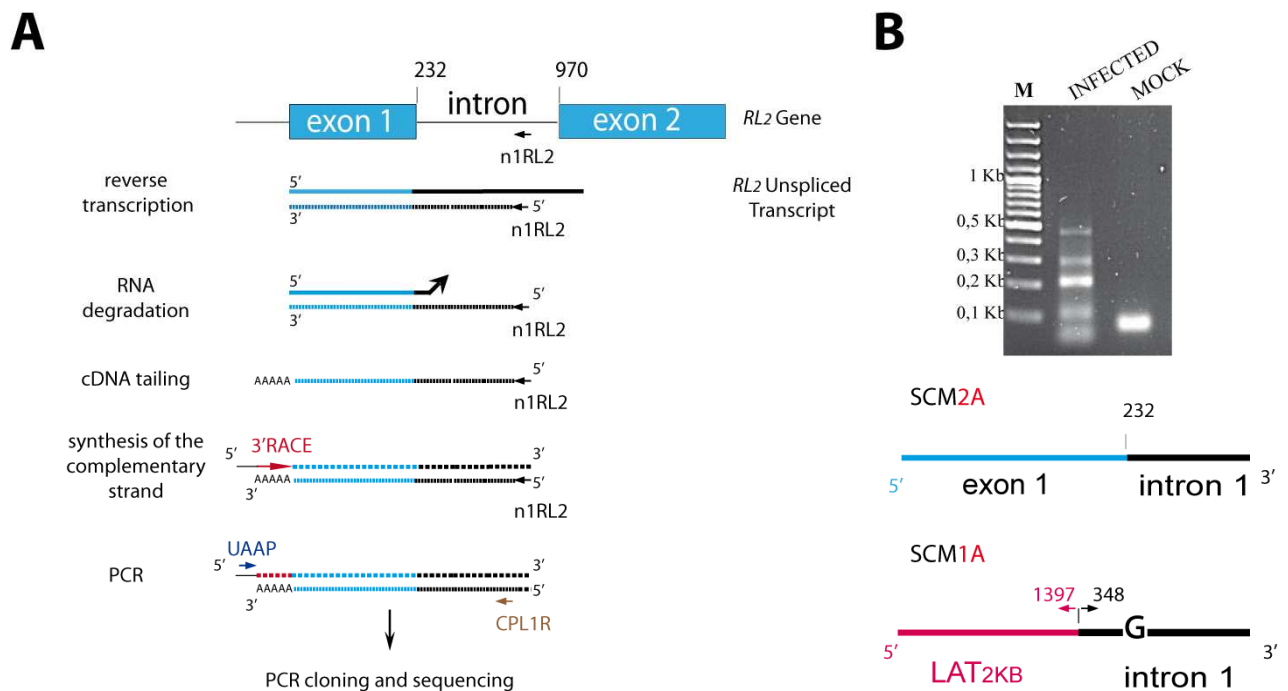


Figure 2: Analysis of the 5' end of unspliced *RL2* transcripts using primer extension assay

Inoculated mice were sacrificed at 28 days pi and total RNA was extracted from the left TGs. (A) The organization of the 5' end of the *RL2* gene of the SC16 strain is schematized. Below, are described the steps of the primer extension assay. After reverse transcription of the *RL2* transcript using the n1RL2 primer, the RNA template was degraded, the cDNA was 3' poly-adenylated and used as a template for synthesis of the complementary strand using primer 3'RACE. Finally, the cDNA was amplified by PCR using primers UAAP and CPL1R, before being cloned and sequenced.

(B) Electrophoresis gel of primer extension products, obtained from total RNA extracts of TG from infected or mock-infected mice. The organization of the longest primer extension products are indicated below the gel. The upper product corresponds to the expected unspliced *RL2* transcript, while the second reveals insertion of part of the LAT (in magenta) into *RL2* transcripts (sequences in figure S1). In the latter sequence, the additional G observed in intro 1 downstream the junction between *LAT* and *RL2* is indicated.

To check independently whether LAT_{2kb} could be inserted into *RL2* gene transcripts, different pairs of primers flanking the expected integration site (*i.e.* nt 348 of *RL2*) were used to amplify the complete LAT insert within intron 1 by RT-PCR. Although unmodified intron 1 could be amplified with these primers, we failed to amplify intron 1 containing the complete sequence of the LAT insert (data not shown). To circumvent this technical problem, we amplified each junction of the insert separately.

Two primers (LATF1 and CPL1R) were designed to amplify the 3' junction of the putative insertion (Figure 3A). LATF1 is a sense-strand primer located in the 3' end of LAT_{2kb} and CPL1R is an antisense-strand primer annealing to the 5' end of *RL2* intron 1. RT-PCR on RNA extracts from four TGs (three from latently infected mice and one mock-inoculated) revealed a unique product in all the 3 latently infected TGs, absent from the mock-infected TG (Figure 3A and S2). Upon cloning and sequencing (3 clones per mouse), it appeared that the 306bp band corresponded to a sequence joining the LAT_{2kb} and the intron 1 of *RL2*. The junction was reproducibly found between nucleotide 1397 of LAT_{2kb} and nucleotide 348 of *RL2* transcript, in agreement with the result of the primer extension assay. As already observed with the primer extension assay (Figure 2B and S1), an additional G was present downstream the junction, in a stretch of G nucleotides in the *RL2* gene (nucleotides 355 to 359) in all the analyzed clones (Figure 3A).

Two primers (O1 and PLAT2) were then designed to analyze the 5' end of the insert in RNA extracts from the 4 previously tested TGs. O1 is a sense-strand primer located in the 3' end of exon 1 of *RL2* gene and PLAT2 is an antisense-strand primer annealing to the 5' end of the LAT_{2kb}. RT-PCR assays revealed similar bands in the 3 latently infected TGs, while no product was observed from the mock-infected TG (Figure 3B). The PCR products were cloned and sequenced (total of 84 clones, giving 9 to 14 analyzable sequences per TG -for details, see the table inserted in figure 3B-). Four types of junction were found (Figure 3B and S3). The most frequently observed was (*RL2*-nt384)/(nt242-*LAT*), joining nucleotides 384 of *RL2* and 242 of *LAT*. According to the same nomenclature, the 3 others junctions were (*RL2*-nt362)/(nt181-*LAT*), (*RL2*-nt384)/(nt221-*LAT*) and (*RL2*-nt206)/(nt295-*LAT*).

Remarkably, none of the 3' ends of the *RL2* part of these junctions was positioned at nt 347 as expected from a simple insertion of *LAT* sequences into the *RL2* transcript. Instead, the sequence corresponding to junction (*RL2*-nt206)/(nt295-*LAT*) had 141 nucleotides missing (*i.e.* nt 207 to nt 347) compared to the expected sequence. Inversely, the sequences with junctions (*RL2*-nt362)/(nt181-*LAT*), (*RL2*-nt384)/(nt242-*LAT*) and (*RL2*-nt384)/(nt221-*LAT*)

had 15 (nt 348 to nt 362) or 37 (nt 348 to nt 384) additional nucleotides starting from nt 348 of *RL2*. These results suggested that part of the *RL2* sequence flanking the insertion site is either deleted or duplicated in these modified *RL2* transcripts.

LAT integration into the RL2 gene during latent infection

We next examined whether these LAT insertions were also present in the viral DNA. DNA was extracted from TGs from the same mice as above and the same primers were used (i.e. LATF1 and CPL1R). PCR assays on the putative 3' junction of LAT_{2kb} insert revealed two bands in the 3 latently infected TGs, but not in the mock-infected TG neither when performed on genomic HSV-1 (Figure 3C). This latter control excluded a template jump during the PCR. The PCR products were individually cloned and sequenced. The sequence of the PCR product corresponding to the upper band revealed an insertion of LAT_{2kb} in the *RL2* gene exactly at the same position as in the *RL2* transcripts, i.e. between the nucleotide 1397 of LAT_{2kb} and nucleotide 348 of *RL2* transcript (Figure 3C and S4). The G insertion in the stretch of G nucleotides of the *RL2* gene, downstream the inserted LAT, was found in 14 clones out of 15 that were sequenced. No *RL2* sequence was found in the PCR product corresponding to the faster migrating band. Therefore, except for one clone (for which the additional G was not detected), the 3' junction between the LAT insert and the *RL2* sequence observed at the DNA level was the same as that observed in the modified transcripts.

The 5' junction of the LAT_{2kb} insert was checked on the same DNA extracts using primers O1 and PLAT2. The PCR analysis revealed two major bands in all the 3 latently infected TGs, but not in the mock-infected TG, neither when performed on genomic HSV-1 (Figure 3D). The PCR products were individually cloned and sequenced (6 clones per band). Two types of junctions were found: (*RL2*-nt384)/(nt242-LAT) and (*RL2*-nt490)/(nt242-LAT) (Figure 3D and S5).

As in the transcripts, these data suggested a duplication of the *RL2* region flanking the insertion at the DNA level. The segments duplicated would then have respectively a length of 37 and 143 nucleotides for the junctions (*RL2*-nt384)/(nt242-LAT) and (*RL2*-nt490)/(nt242-LAT).

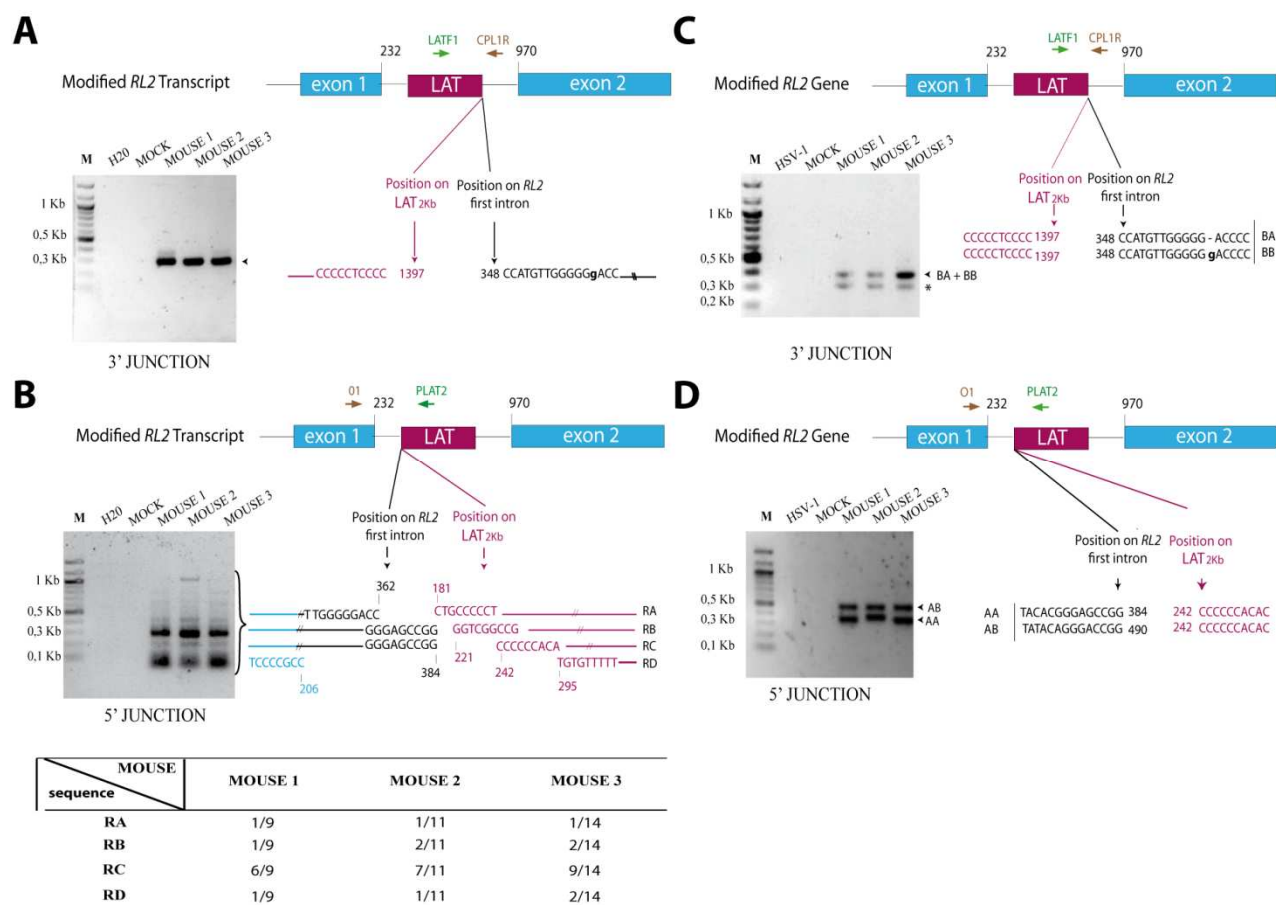


Figure 3: Part of LAT is found inserted in unspliced *RL2* transcripts and *RL2* gene during HSV1 latency

Inoculated mice were sacrificed at 28 days pi and DNA and total RNA were extracted from the left TGs. (A and B) RT-PCR analysis of the 3' (A) and 5' (B) junctions between LAT_{2kb} and intron 1 of modified *RL2* transcripts.

A schematic representation of the modified *RL2* transcript is drawn showing the *RL2* exons (cyan box), *RL2* intron 1 (black line) and the LAT insertion (magenta box). The primer pairs used for RT-PCR analysis are indicated by arrows. The electrophoresis gels of the RT-PCR products obtained from RNA extracts of TGs from mock or 3 different infected mice are shown. A control without RNA was also tested (H₂O).

The sequences of the junctions (as assessed by sequencing of the obtained RT-PCR products, see also figures S2 and S3) are indicated, showing the positions of nucleotides on *RL2* and LAT. In (A), the additional G in *RL2* sequence is indicated in low case and bold. It has been positioned arbitrarily at the end of the G stretch. In (B), the table indicates the relative frequencies of each 5' junction from positive clones (i.e. revealing an *RL2*-LAT junction) obtained from the 3 different mice tested.

(C and D) PCR analysis of the 3' (C) and 5' (D) junctions between LAT_{2kb} and intron 1 of modified *RL2* genes in DNA.

A schematic representation of the modified *RL2* gene is drawn (using the same color code as in A and B). The primer pairs used for PCR analysis are indicated by arrows. The electrophoresis gels of the PCR products obtained from DNA of TGs from mock or 3 different infected mice are shown. A control with HSV-1 DNA was also tested (HSV-1). In (C), no *RL2* sequence was found in the PCR product indicated by the asterisk. The sequences of the junctions (as assessed by sequencing of the obtained PCR products, see also figures S4 and S5) are indicated, showing the positions of nucleotides from *RL2* and LAT. In (C), the additional G in *RL2* sequence is indicated in low case. It has been positioned arbitrarily at the end of the G stretch.

Taken together, the sequence data indicate that the major species of modified transcript [the one with 5' junction (*RL2*-nt384)/(nt242-*LAT*) and 3'junction (*LAT*-nt1397)/(nt348-*RL2*)] corresponds to the expected transcript of the modified gene detected in our PCR assays. Two other modified transcripts [those with 5' junctions (*RL2*-nt362)/(nt181-*LAT*) or (*RL2*-nt384)/(nt221-*LAT*), and the unique 3'junction (*LAT*-nt1397)/(nt348-*RL2*)] contain duplications of the *RL2* regions [348-362] and [348-384], respectively, and thus are probably transcripts of modified genes that have not been detected in our PCR assay. The modified transcripts with 5' junction (*RL2*-nt206)/(nt295-*LAT*) and 3'junction (*LAT*-nt1397)/(nt348-*RL2*) could be the result of an alternative splicing in a transcript also expressed from a modified *RL2* gene. Therefore, all the modified transcripts observed at 28 days pi most probably derive from the transcription of modified *RL2* genes having an insertion of part of the *LAT*_{2kb} into intron 1.

Southern blot and inverse PCR confirm LAT insertion into RL2 gene.

To confirm *LAT* insertion into *RL2* gene, DNA extracted from latently infected TGs was submitted to an extensive digestion by restriction enzyme *Xho*II which does not cleave inside the *LAT* sequence. This enzyme was also chosen because the cleavage sites flanking the putative *LAT* insertion into *RL2* should result in products that could be easily distinguished from the expected products containing the *LAT* genes (Figure 4A). A southern blot was then performed on the digestion products and *LAT*-containing fragments were revealed using a single-stranded DNA probe corresponding to nt 971 to nt1310 of the *LAT* sequence (Figure 4B).

Two major bands (a and d, indicated by black arrows on figure 4B) were revealed. Their size was consistent with that of the two expected *LAT* containing fragments in HSV1 genome (Figure 4A). At least 4 other bands were also visible (indicated by orange arrows), two of which (labeled b+*LAT* and c+*LAT* on figure 4B) with a size potentially corresponding to a fragment in which *LAT* was inserted into *RL2* gene. These data confirmed the mobility of the *LAT* element and suggested that *LAT* insertion could occur in genes other than *RL2*.

In order to characterize *LAT* containing fragments, an inverse PCR (iPCR) strategy was used. Gel analysis (Figure 4C) revealed several PCR products which were individually cloned and sequenced. We obtained two clones corresponding to *LAT* insertion into *RL2* gene. This is another evidences that *LAT*_{2kb} is able to reintegrate the HSV1 genome in the *RL2* gene.

The sequence of the first clone (Figure 4D and S6) indicated a LAT insertion with a 5' junction corresponding to (RL2-nt384)/(nt242-LAT) and a 3' junction corresponding to (LAT-nt1397)/(nt348-RL2). This demonstrates that these previously identified junctions are indeed associated by a unique LAT insertion and that there is a duplication of the flanking *RL2* sequence (from nt 349 to 384) in such a modified gene.

The sequence of the second clone (Figure 4D and S7) revealed a 5' junction corresponding to (RL2-nt490)/(nt242-LAT) and a 3' junction corresponding to (LAT-nt1552)/(nt667-RL2). This revealed an unexpected new 3' junction between LAT insertion and *RL2* gene. In this case, LAT insertion leads to a deletion in the *RL2* sequence (from nt 491 to nt 666). As this junction could not be amplified with the primers previously used in figure 3C, we used a new pair of primers specifically designed to amplify this junction. Nevertheless, we were still unable to amplify this new junction using the same conditions as for the previous one [i.e. (LAT-nt1397)/(nt348-RL2)], which strongly suggests that this junction is much less represented.

Quantification of modified *RL2* genes and unspliced transcripts

To estimate the frequency of *RL2* modification by LAT insertion, quantitative PCR was performed on DNA extracted from TG of eight latently infected mice (3 independent quantifications for each extract). The absolute numbers of modified *RL2* genes varied from 5.68 to 7.21 Log₁₀ copies per TG, while the total numbers of *RL2* genes varied from 6.54 to 8.89 Log₁₀ copies per TG (Table 1). Depending on the mouse, the ratio of modified *RL2* genes varied from 0.15 % to 34 % with 5 mice out of 8 having an insertion ratio above 8 %.

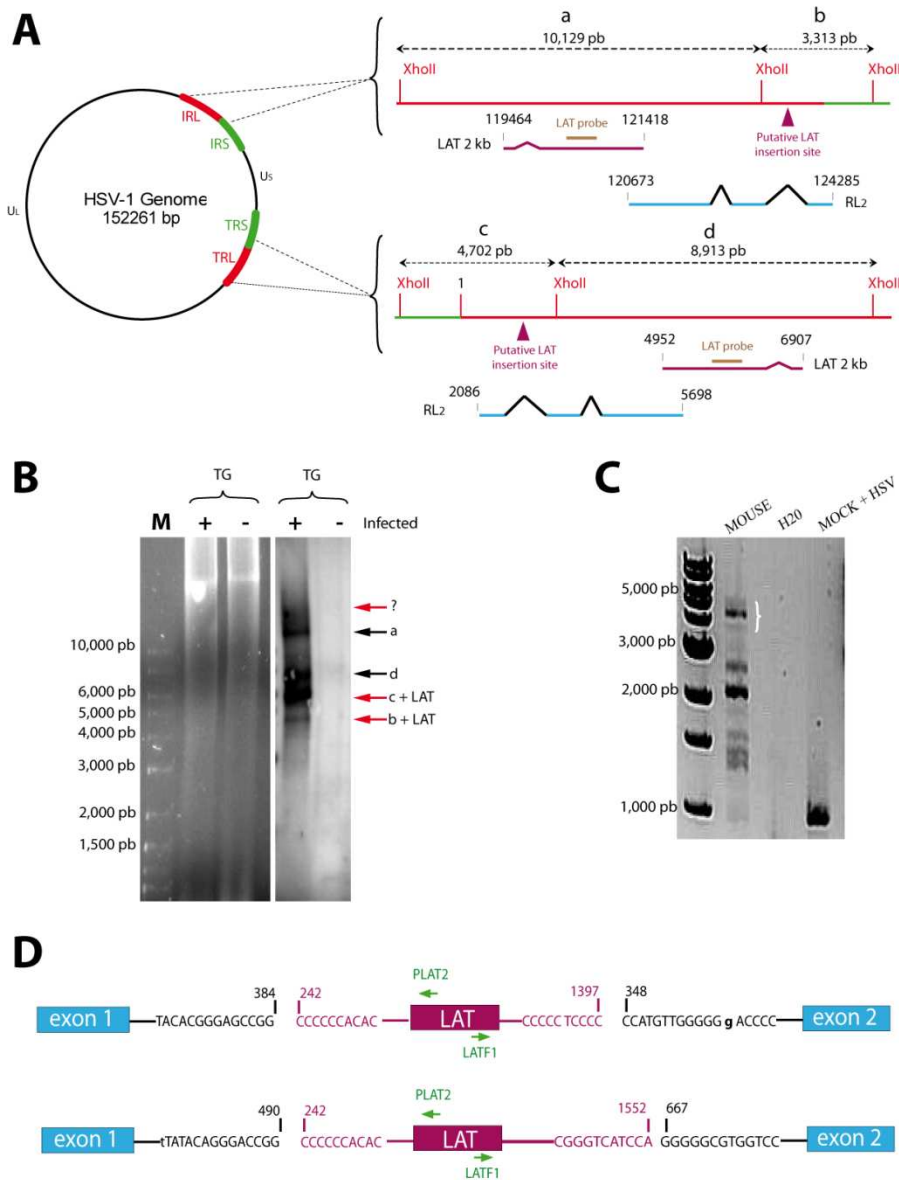


Figure 4: confirmation of LAT insertion into *RL2* gene by southern blot and iPCR.

(A) XhoII restriction map in the two genome regions containing *RL2* and *LAT* genes. Intraneuronal HSV1 genomes are expected to be exclusively in a circular form. The position of the probe (brown line), the putative insertion site of LAT in *RL2* gene (magenta triangle) and the XhoII restriction sites are indicated. The size of the four expected restriction fragments (a, b, c, d) is also given.

(B) Southern blot performed on total DNA extracted from the TG of an infected (+) or mock infected (-) mouse and digested by XhoII. On the left, is shown the electrophoresis gel of the digestion products (revealed by BET) before transfer and hybridization by LAT_{2kb} probe. On the right, is shown the autoradiography (14 days of exposure). The bands indicated by black arrows (a and d) could correspond to digestion fragments containing the original LAT gene. The bands indicated by red arrows could correspond to LAT insertion into cellular or viral DNA. Those labeled b+LAT and c+LAT could correspond to LAT integrated into *RL2* genes.

(C) Electrophoresis gel of iPCR products performed on total DNA extracted from the TG of an infected mouse, digested by XhoII and self-ligated. As controls, the same iPCR condition were made on H₂O or on a mixture of total DNA extracted from TG of a mock-infected mouse, digested by XhoII and self-ligated, and HSV-1 genome also digested and self-ligated (MOCK + HSV). PCR products were individually cloned and sequenced. The bracket indicates iPCR products corresponding to the insertion of LAT into *RL2*.

(D) Sequences of the junctions between LAT and *RL2* genes obtained from iPCR products (detailed sequences and chromatograms are shown in figures S6 and S7)

To estimate the amount of unspliced *RL2* transcripts with and without LAT insertion into intron 1, we used quantitative RT-PCR assay. The amounts of all unspliced *RL2* transcripts (whether containing LAT insertion or not) were very similar from one mouse to another and varied from 5.28 to 5.75 Log₁₀ copies per TG (average 5.58 +/- 0.25). It could not be distinguished from the amounts of LAT-containing *RL2* transcripts which varied from 5.09 to 5.93 Log₁₀ copies per TG (average 5.53 +/- 0.45) (Table S1). Therefore, it appeared that the major part of the unspliced transcripts contained LAT insertion. The amount of spliced *RL2* transcripts was below the threshold of detection (*i.e.* < 4,6 Log₁₀ copies per TG). Thus, more than 90 % of the *RL2* transcripts detected in the TG were unspliced (which was in agreement with previous results {Maillet et al., 2006 }), the vast majority of them containing LAT insertion.

Taken together, these data were consistent with the fact that only modified *RL2* genes were transcribed whereas unmodified genes are silenced.

Mouse	Day Post-infection	Log ₁₀ (total copies of <i>RL2</i> gene)	SD	Log ₁₀ (copies of modified <i>RL2</i> gene)	SD	Percentage of modified <i>RL2</i> genes
M1	28	7.06	 0.11	6,60	 0.66	34.5
M2	28	6.95	 0.30	6,10	 0.58	14
M3	28	8.89	 0.03	6,44	 0.09	3.5
M4	28	8.29	 1.03	7,21	 0.79	8.5
M5	28	6.54	 0.43	5,98	 0.70	27.5
M6	28	7.28	 0.38	6,25	 0.63	9.5
M7	28	8.51	 0.11	5,68	 0.68	0.15
M8	28	8.42	 0.31	6,71	 0.90	2

Table 1: Quantification of the amounts of total and modified genomes from DNA extracted from mice TG.

Quantification of all copies of *RL2* gene (*i.e.* modified and unmodified) and copies of *RL2* gene modified by LAT's insertion was performed by qPCR, as described in the methods section. TG were harvested from mice infected for the indicated time (days pi) and used for DNA extraction. Values are indicated as the mean of experiments conducted in triplicate. The percentage of modified versus total genome is deduced from the indicated experimental values. SD, standard deviation; NA, not applicable.

Part of LAT is also found inserted in US12 transcripts during latency

Although we only identified LAT insertion into RL2 genes, the Southern blot suggested that other genes were modified. We first tried to determine whether other transcripts were modified by LAT insertion, using a primer extension assay with RLAT3, a primer annealing from nt 1038 to nt 1058 of LAT_{2kb}. It was performed on poly-A transcripts purified from RNA extracts from latently infected TGs (28 days pi). We chose to work on poly-A transcripts in order to get rid of LAT_{2kb} and LAT_{1.5kb} (which are the most abundant RNA which are potentially amplified in such an assay).

Gel analysis revealed several bands among which 3 were absent in the mock-infected TG (Figure 5A). The corresponding products were individually cloned and sequenced. The longest one corresponded to LAT_{8.3kb} (which is polyadenylated). The second corresponded to a product which had a 5' end corresponding to RL2 gene until nucleotide 384 followed by the sequence of LAT_{2kb} starting at nucleotide 242 (*i.e.* corresponding to the major product observed previously –Figure 3B-) (Sequence in figure S8). The third product had a 5' end corresponding to HSV1 US12 gene until nucleotide 1218 followed by the sequence of LAT_{2kb} starting at nucleotide 224. This latter sequence (Figure S9) revealed that some US12 transcripts also contained an insert corresponding to part of the LAT_{2kb}.

In order to amplify the 3' junction of the putative insertion, primers LATF1 (previously used and located in the 3' end of LAT_{2kb}) and 47REV1, an antisense-strand primer annealing to the 3' end of US12 exon 2 were used for RT-PCR on RNA extracts from 4 TGs (3 from latently infected mice and one mock-inoculated). A single product was amplified in all the latently infected TGs, absent from the mock-infected TG (Figure 4B). Upon cloning and sequencing, it appeared to correspond to a sequence joining the LAT_{2kb} and the 3' end of exon 2 of US12. The junction was reproducibly found between nucleotide 1350 of LAT_{2kb} and nucleotide 1304 of US12 transcript (Figure 5B and S10).

We then used primers 47F1 (a sense-strand primer located in the 5' end of exon 2 of US12 gene) and PLAT2 (the previously used antisense-strand primer annealing to the 5' end of the LAT_{2kb}) to amplify and analyze the 5' end of the insert in RNA extracts from the 4 TGs that were used to analyze the 3' junction. RT-PCR assays revealed a single band in the 3 latently infected TGs, while no product was observed from the mock-infected one (Figure 5C). Upon cloning and sequencing, it appeared to correspond to a unique sequence revealing a junction between nucleotide 1218 of US12 transcript and nucleotide 224 of LAT_{2kb} (Figure

5C and S11), which was identical to that observed in the primer extension assay (Figure 5A and S9).

Therefore a single modified *US12* transcript was detected with the 5' junction of LAT insertion corresponding to (US12-nt1218)/(nt224-LAT) and the 3' junction corresponding to (LAT-nt1350)/(nt1304-US12). Here again, this was not a simple insertion of LAT sequences as 85 nucleotides were missing (i.e. nt 1219 to 1303) in these modified *US12* transcripts, compared to the expected sequence.

LAT integration into the US12 gene during latent infection

We next examined whether these LAT insertions were also present in the *US12* gene. Primers pairs (LATF1, 47 REV1) (Figure 5D) and (47F1, PLAT2) (Figure 5E) were respectively used in PCR assays to characterize the putative 3' and 5' junctions of LAT_{2kb} insert in the *US12* gene. The PCR assays were performed on DNA which had been extracted from TGs from the same 4 mice that were used to characterize the modified *US12* transcripts.

For both PCR assays, a single product was detected (Figure 5D, 5E). The first one revealed a 3' junction corresponding to (LAT-nt1350)/(nt1304-US12) (Figure 5D and S12). The second band revealed a 5' junction corresponding to (US12-nt1218)/(nt224-LAT) (Figure 5E and S13). The 3' and 5' junctions were thus identical to those found in the modified *US12* transcripts. Therefore, the modified *US12* transcripts observed at 28 days pi most probably derive from the transcription of the modified *US12* genes having an insertion of part of the LAT_{2kb} into exon 2.

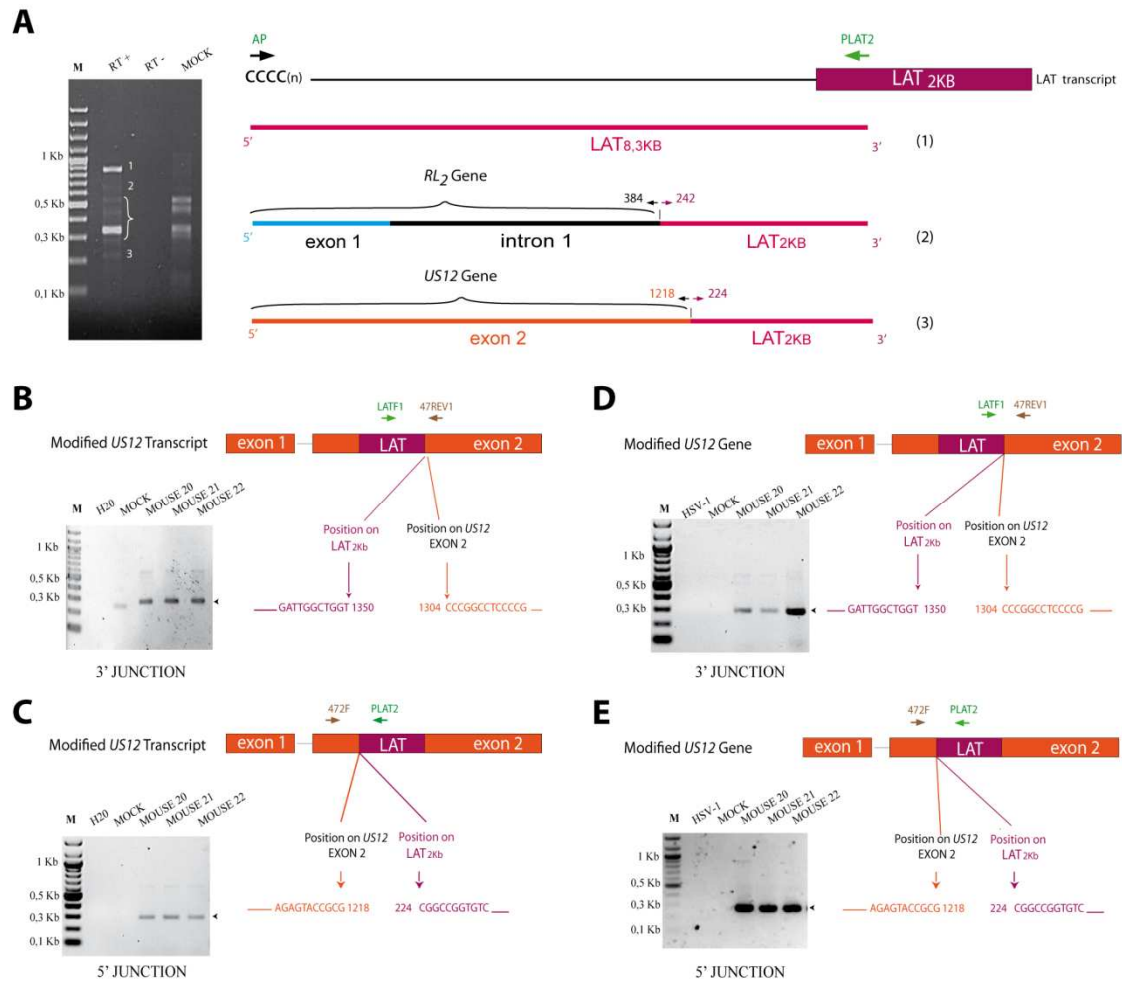


Figure 5: Part of LAT is found inserted in *US12* transcripts and gene during HSV1 latency

Inoculated mice were sacrificed at 28 days pi and total RNA (A, B, C) or DNA (D, E) was extracted from the left TG of infected or mock-infected mice.

(A) Polyadenylated RNA were purified and used as templates for primer PLAT2 extension assay. In a sample containing polyadenylated RNA from an infected mouse (RT-), the reverse transcription step was omitted. The electrophoresis gel of primer extension products is shown on the left. The PCR products labeled 1, 2 and 3 correspond respectively to LAT_{8.3kb}, *RL2* transcript containing LAT insert and *US12* transcript containing LAT insert. These product and the corresponding junctions are described (Sequences and chromatograms are in figure S8 and S9). The bracket on the gel indicates non-specific products also present in the mock lane.

(B and C) RT-PCR analysis of the 3' (B) and 5' (C) junctions between LAT_{2kb} and exon 2 of modified *US12* transcripts.

A schematic representation of the modified *US12* transcript is drawn showing the *US12* exons (orange box), *US12* intron 1 (black line) and the LAT insertion (magenta box). The primer pairs used for RT-PCR analysis are indicated by arrows. The electrophoresis gels of the RT-PCR products obtained from RNA extracts of TGs from mock or 3 different infected mice are shown. A control without RNA was also tested (H₂O).

The sequences of the junctions (as assessed by sequencing of the obtained RT-PCR products, see also figures S10 and S11) are indicated, showing the positions of nucleotides on *US12* and LAT.

(D and E) PCR analysis of the 3' (D) and 5' (E) junctions between LAT_{2kb} and exon 2 of modified *US12* genes in DNA.

A schematic representation of the modified *US12* gene is drawn (using the same color code as in B and C). The primer pairs used for PCR analysis are indicated by arrows. The electrophoresis gels of the PCR products obtained from DNA of TGs from mock or 3 different infected mice are shown. A control with HSV-1 DNA was also tested (HSV-1). The sequences of the junctions (as assessed by sequencing of the obtained PCR products, see also figures S12 and S13) are indicated, showing the positions of nucleotides from *US12* and LAT.

LAT is found integrated in the host genome during latent infection

As previous made clear that LATs are able to integrate into the HSV-1 genome, we wonder whether LAT was also able to integrate host DNA. With the aim to identify potential host targets, we performed B1-PCR experiments with a primer annealing to LAT and another annealing to conserved sequences of B1 elements which are highly present in mouse genome.

To amplify a potential junction between a host sequence and the 5' part of LAT_{2kb}, two pairs of primers were designed (Figure 6A). Amplification was performed on DNA extracts from the TG of infected mice for which the presence of a modified *RL2* gene had been previously checked. With both pairs, gel analysis revealed several bands (Figure 6A). The PCR products were cloned and sequenced. One of them revealed a LAT insertion in chromosome 9 (ref NCBI: AC174644.2) (Figure 6B and S14) which was found in the three mice tested. The 3' end of the LAT insertion corresponded to nt 254.

Using primers designed to amplify a junction between the 3' part of LAT_{2kb} and a host sequence, we amplified several products (Figure 6C). Sequence analysis revealed 4 LAT insertion in chromosomes 5 (ref NCBI: AC124754.5), 7 (ref NCBI: AC148980.6), 9 (ref NCBI: AC117679.8) and 12 (ref NCBI: AC132596.5) (Figure 6D and S15). These insertions were found in the three tested mice with the exception of the one in chromosome 5 which was found in two mice out of three. The 5' end of the LAT insertions in chromosome 5 and 7 corresponded respectively to nt 1718 and 1539 of LAT sequence. For the insertions in chromosomes 9 and 12, it was not possible to determine the exact 5' end of the insertion as the nucleotides located at the junction were common to the LAT and the host DNA sequence (Figure 6D and S15).

Interestingly, LAT insertion in chromosome 12 apparently led to the conversion of two nucleotides located respectively 8 and 10 nucleotides downstream the sequence common to the LAT and the host DNA at the level of the 3' junction (Figure 6D and S15).

All these insertions were independently and individually checked by PCR using the same primer for the LAT sequence and a host primer specifically designed for the junction. This definitively demonstrated that LATs are also able to integrate the host DNA.

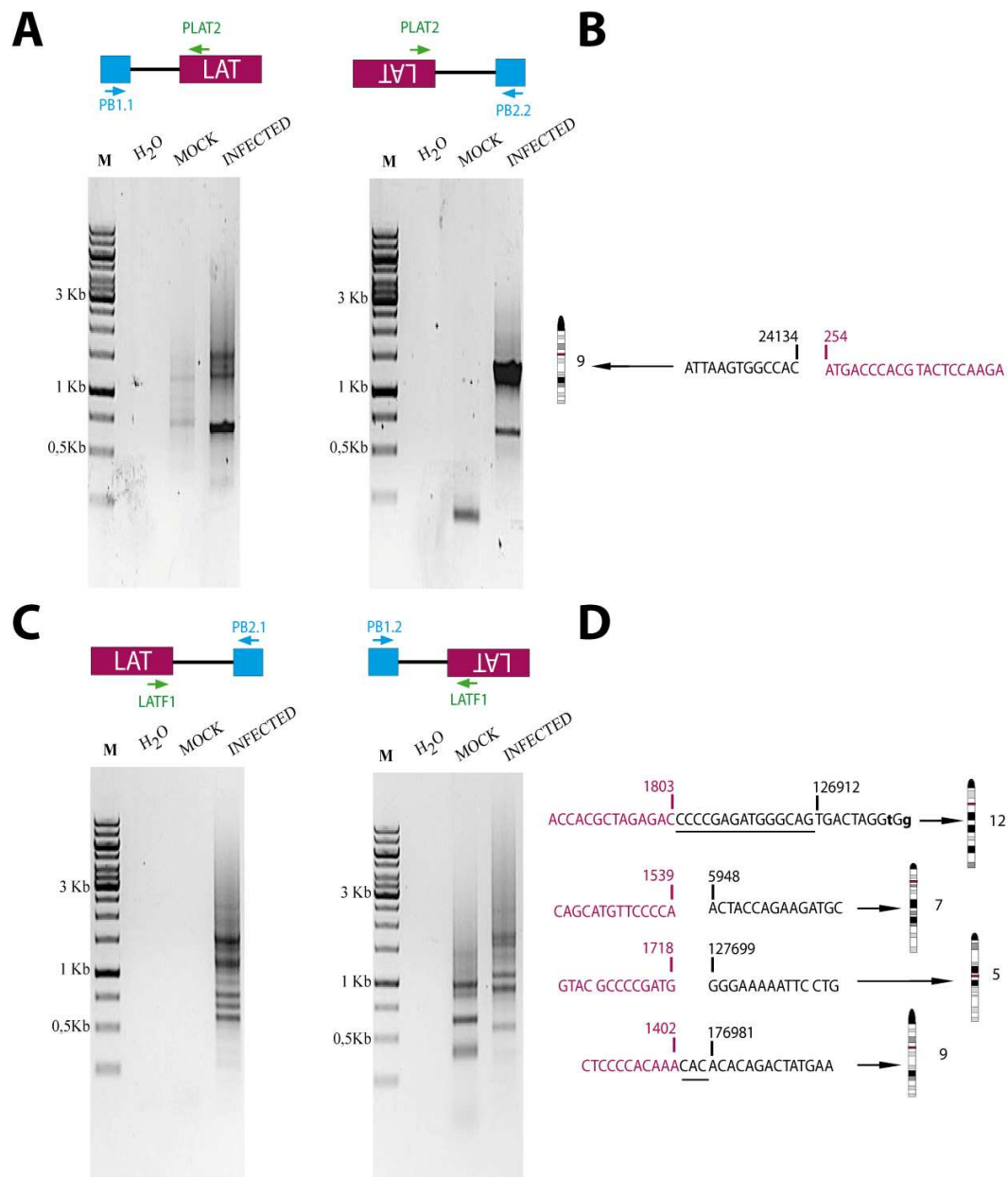


Figure 6: Part of the LAT_{2kb} is found integrated into different mouse loci.

B1-PCR were performed on total DNA extracted from the TG of an infected or mock infected mouse. Four primers pairs were used to take into account all the relative orientation between B1 elements and LAT integrated sequences. Primers positions are indicated.

(A) Electrophoresis gels showing products amplified using the two primers pairs allowing the amplification of a putative 5' junction between LAT_{2kb} and host DNA.

(B) The sequence of the junction of the unique clone revealing LAT_{2kb} integration is indicated. The positions of nucleotides from host DNA (chromosome 9) and LAT_{2kb} at the level of the 5' junction are mentioned (see also Figure S14 for detailed sequences).

(C) Electrophoresis gels showing products amplified using the two primers pairs allowing the amplification of a putative 3' junction between host DNA and LAT_{2kb}.

(D) The sequences of the junctions of the four clones revealing LAT_{2kb} integration are indicated. The positions of nucleotides from host DNA and LAT_{2kb} at the level of the 3' junction are mentioned (see also Figure S15 for detailed sequences).

The underlined nucleotides at the junction cannot be attributed due to the fact that the LAT and the chromosome sequences are the same. For chromosome 12, the nucleotides in low case and bold are different from those in the chromosome sequence (see figure S15A).

Discussion

In this study, we report that *in vivo* LATs sequences can be found inserted into viral and host genome sequences. Taken together, our data are consistent with LATs being mobile elements with unique properties. These properties are likely essential for their capacity to regulate latency and reactivation.

LATs have been described for years as the major latency regulators, apparently acting through different mechanisms (as detailed in the introduction). Our discovery that LATs are mobile elements explains some of the previous observations, particularly those showing that LATs are associated with a down-regulation of immediate early genes expression {Garber et al, 1997; Mador et al, 1998 }. Indeed, the majority of *RL2* genes are clearly unmodified at 28 days pi (table 1). Nevertheless, at this stage of infection, more than 90 % of *RL2* transcripts detected in the TG contain intron 1 and the LAT-insert, and are thus transcribed from modified genomes. This indicates that transcription of unmodified *RL2* genes is silenced. Furthermore, this also indicates that LAT insertion into the *RL2* transcript interferes with its splicing. This results in the inhibition of the production of ICP0, a key factor of the reactivation process {Everett, 2000 }. Therefore, LATs are latency keepers by impeding ICP0 expression from non-silenced genomes.

The physiological significance of LAT insertion in *US12* is more difficult to assess. The *US12* gene encodes the protein ICP47, a viral protein that binds the transporter associated with antigen processing, prevents peptide translocation into the endoplasmic reticulum and, as a consequence, inhibits antigen presentation to CD8⁺ T lymphocytes. This avoids the detection of infected cells by the host immune system {Hill et al, 1995; Fruh et al, 1995; York et al, 1994 }. This gene was hitherto not known to be transcribed during latency. It is worth to note that, as in the case of *RL2* gene, LAT insertion in *US12* occurs in a non-coding region of the transcript (part of exon 2 downstream the stop codon). One might speculate that stress conditions could induce the maturation of accumulated modified transcripts which would then be exported out of the nucleus and translated, which would lead to the synthesis of ICP0, which would trigger the reactivation cascade, and ICP47, which would immediately protect the cell from the immune system.

Besides this ability of LATs to modify viral genes, we have demonstrated that it can also integrate the host genome. As the PCR technique we used to identify LAT insertion into host DNA was biased (because it required the proximity of a B1 element), we have probably not identified all the LAT targets in the host genome. Nevertheless, LAT ability to insert into

the host genome is certainly a way for the virus to take the control of some cellular functions. This feature could also participate in maintaining latency and/or in the reactivation process.

A characteristic of this new mobile element is that it transposes through an RNA intermediate. In order to accomplish the complete integration process, as for other retromobile elements, at least three activities are required. LAT RNA needs to be reverse transcribed, target DNA needs to be cleaved and a recombination event must take place.

The characteristics of LAT_{2kb} provide some clues to the mechanisms used. First, similar to group II introns, LAT_{2kb} is a lariat. This suggests that, as for these introns, reintegration proceeds through direct attack of the target DNA by the lariat. Furthermore, there are several unusual features of the modified genes that may offer clues as to the mechanisms involved in the reverse transcription of the LAT_{2kb} containing transcripts and their integration, namely: (i) a partial duplication of the region flanking the insertion (for part of the modified RL2 genes) or (ii) a deletion of the target gene (for another part of the modified RL2 genes and most probably for the modified *US12* genes), (iii) the addition of an extra-nucleotide downstream LAT insertion in *RL2* gene and (iv) the conversion of nucleotides located downstream LAT insertion in a host DNA target. All these features are characteristic of “target-site primed reverse transcription” {Eskes et al, 1997 }, a process that is analogous to that described for mobile group II introns {Zimmerly et al, 1995 a b } and non-LTR retrotransposons {Gilbert et al, 2002 }.

Surprisingly, local alignment around the target sequences has not revealed a consensus. This raises the problem of target recognition by the LAT_{2kb}, a feature which is necessarily required at some stage of the process for integration.

The ability of a stable lariat to reintegrate genomic DNA opens a new field of investigation in eukaryotic cells. Furthermore, the fact that the HSV1 genome contains a mobile element that shares several properties with catalytic introns, which have so far been described only in procaryotes or in mitochondria (whereas LATs act in the nucleus of the infected cell), raises evolutionary questions. Whether this mobile element is shared with other herpesviruses, for which a common origin with some bacteriophages has been proposed {Baker et al, 2005; Cheng et al, 2004; Duda et al, 2006 }, or is a recent acquisition by HSV1 is now an open question.

Finally, as HSV1 establishes a long-life latent infection within the human nervous system, LAT_{2kb} insertion in the human genome could result after several years in events leading to chronic neuronal disorders or tumor development. This however remains to be

explored, on the basis of the epidemiological data. Our results thus open new fields of investigation for the long term consequences of infection by *Herpesviridae* through a direct genetic modification of the host, leading to chronic disorders.

Figure S1A

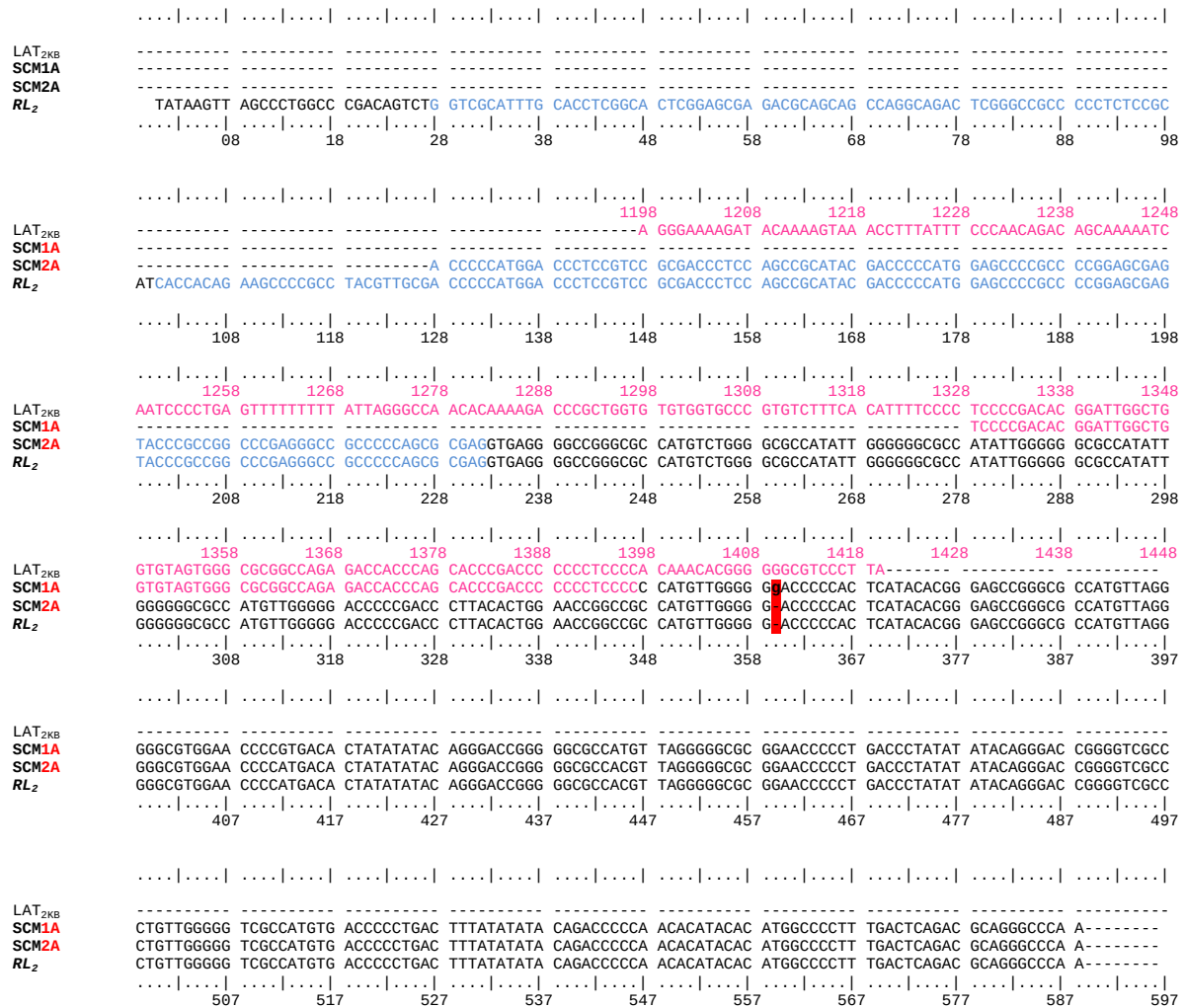


Figure S1A. Sequences of the products obtained by primer extension assay of unspliced *RL2* transcripts from 28 days infected mouse TG

The two sequences are aligned with those of *RL2* (blue and black) and *LAT*_{2kb} (magenta). The exon sequence of the *RL2* gene is indicated in blue. *LAT* sequence found in clone *SCM1A* is in magenta. Numbering in blue refers to *RL2* sequence, and numbering in magenta refers to *LAT*_{2kb} sequence. The red box on the alignment indicates an insertion of a G nucleotide (in lower case and bold) in *SCM1A* sequence in a G stretch of the *RL2* sequence. This additional nucleotide has been positioned arbitrarily at the end of the G stretch between positions 359 and 360 of *RL2* sequence.

Figure S1B

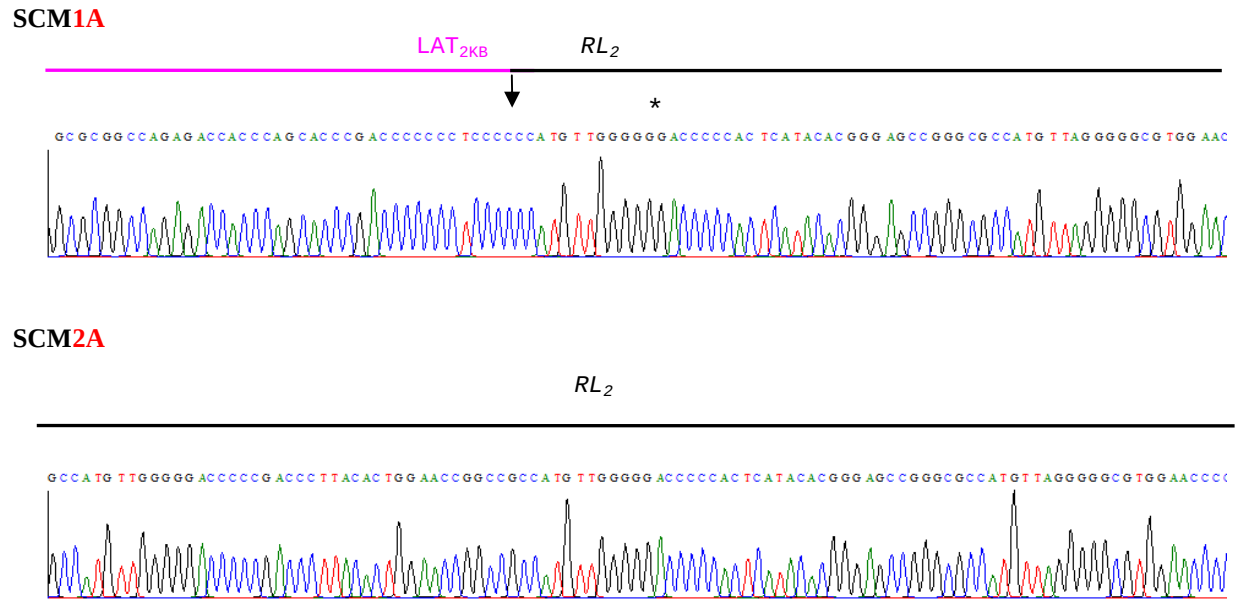


Figure S1B. Part of the histogram of SCM1A and SCM2A sequences obtained from primer extension assay corresponding to the junction between LAT_{2KB} and RL_2 sequences.

The RL_2 sequence is indicated by the black line above the histogram and LAT sequence is indicated by a magenta line. Black arrows above the histogram corresponding to SCM1A sequence indicate the position of the junction between LAT and RL_2 gene. The asterisk indicates the additional G inserted in the RL_2 sequence, which has been positioned arbitrarily at the end of the G stretch.

Figure S2A

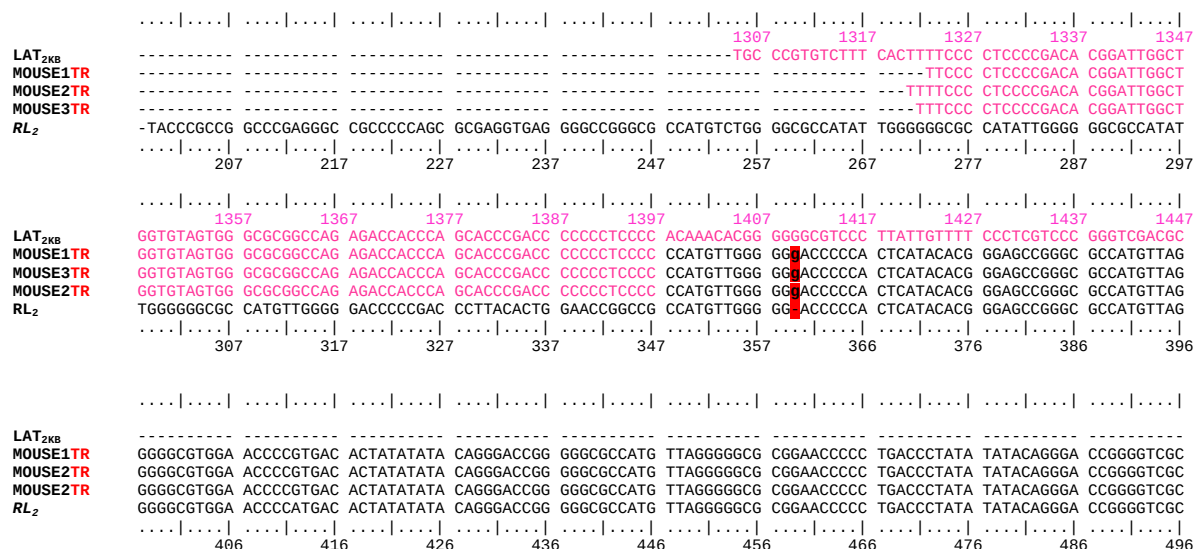
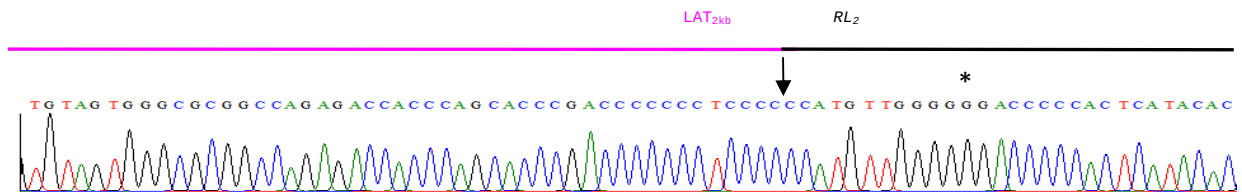


Figure S2A. Alignment of sequences of RT-PCR products, mapping the 3' junction between LAT_{2KB} insertion and intron 1 of *RL2* transcripts, with *RL2* and LAT sequences.

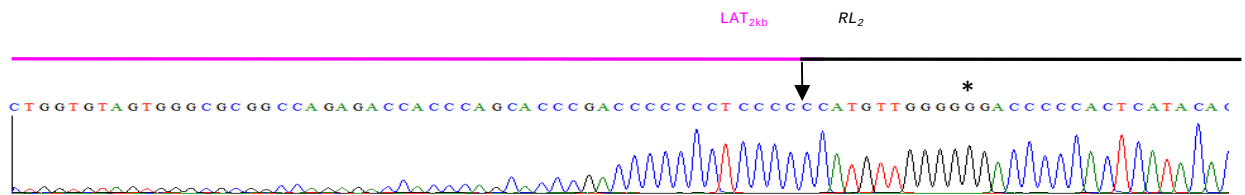
RT-PCR products were obtained from RNA isolated from three different 28 days infected mice TG. The three different sequences are aligned with the sequences of LAT_{2kb} and *RL2* to identify the 3' junction between LAT and *RL2* sequences. Names of the sequences are indicated on the left. Numbering and characters in black refer to the *RL2* sequence; numbering and characters in magenta refer to the LAT_{2kb} sequence. The red box on the alignment indicates an additional guanosine (in lower case and bold) in the *RL2* sequence of the modified transcript. This additional nucleotide has been positioned arbitrarily at the end of the G stretch.

Figure S2B

MOUSE1TR



MOUSE2TR



MOUSE3TR

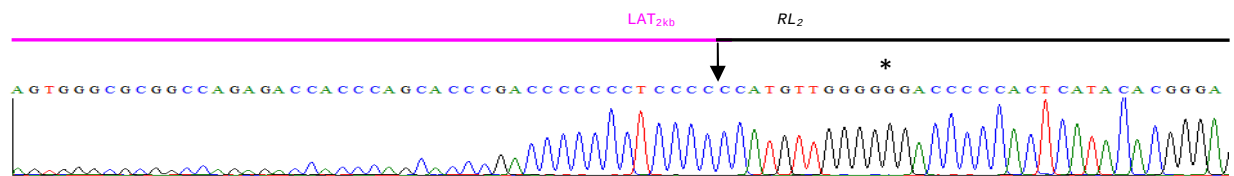


Figure S2B. Part of the histograms of the sequences of RT-PCR products from three different 28 days infected mice (named as in figure S2A) corresponding to the junction between LAT_{2KB} and RL₂ sequences.

The RL₂ sequence is indicated by the black line above the histogram and LAT sequence is indicated by a magenta line. Black arrows above the histograms indicate the position of the junction between LAT and RL₂ gene. The asterisk indicates the additional G inserted in the RL₂ sequence. This guanosine has been positioned arbitrarily at the end of the G stretch.

Figure S3A

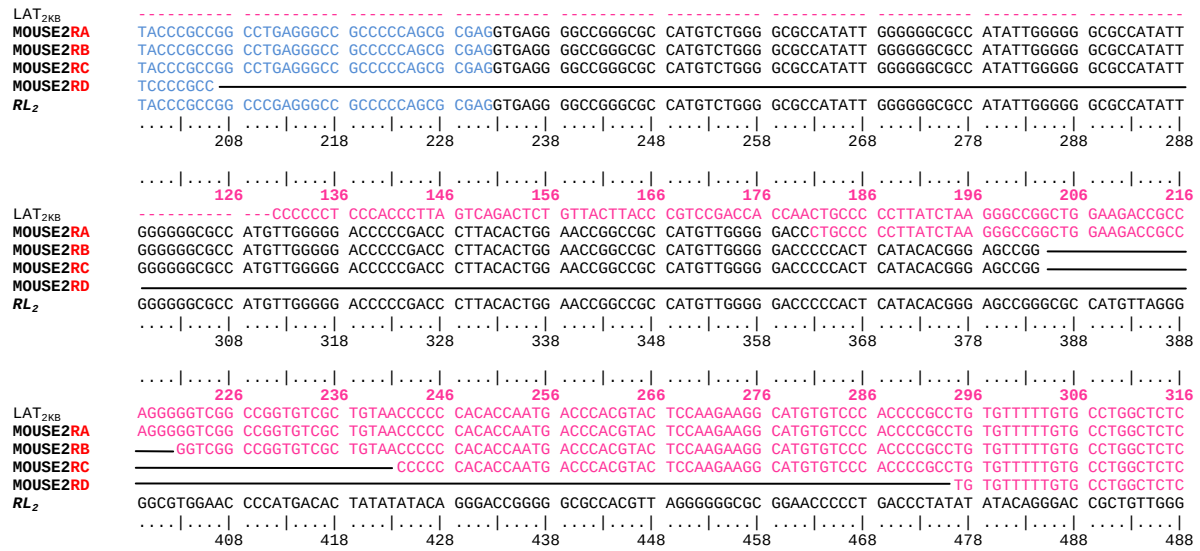


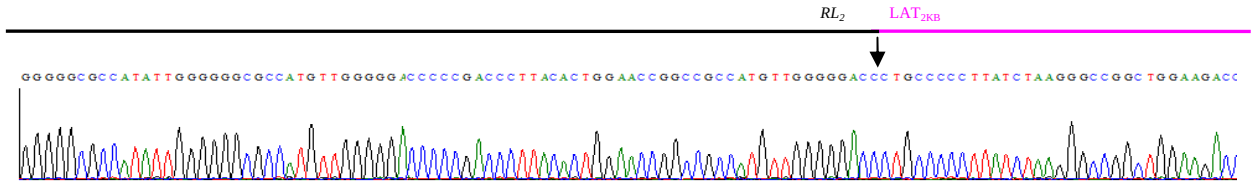
Figure S3A. Alignment of sequences of RT-PCR products, mapping the 5' junction between intron 1 of *RL2* and *LAT*_{2KB} insertion, with *RL2* and *LAT* sequences.

RT- PCR products were obtained from RNA isolated from 28 days infected mice TG.

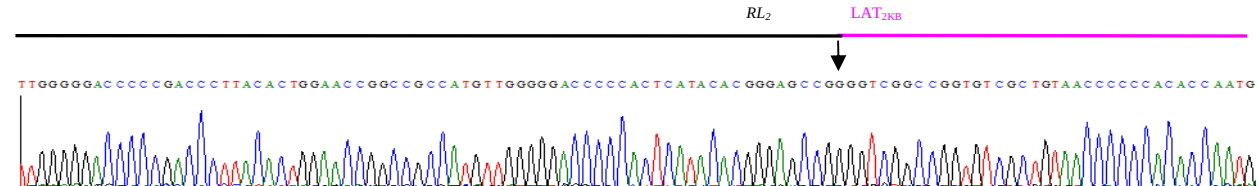
The sequences of 4 different cDNA clones are aligned with the sequences of *RL*₂ and *LAT*_{2kb} to identify the 5' junction between *RL*₂ and *LAT* sequences for each clone. Names of the sequences are indicated on the left. Characters in blue refer to the *RL*₂ first exon sequence, characters in black refer to the *RL*₂ first intron sequence and character in magenta refer to *LAT*_{2KB} sequence. Numbering in black refers to the *RL*₂ sequence and numbering in magenta refers to the *LAT*_{2kb} sequence.

Figure S3B

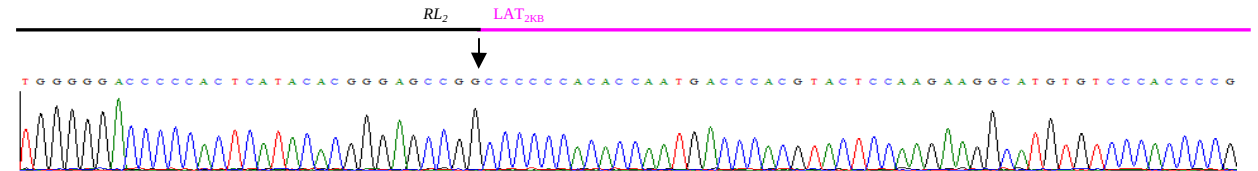
MOUSE2RA



MOUSE2RB



MOUSE2RC



MOUSE2RD

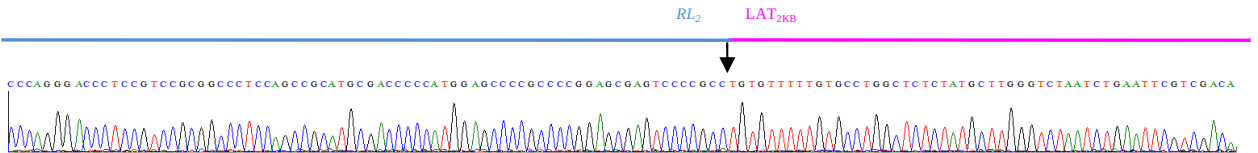


Figure S3B. Part of the histograms of the sequences of RT-PCR products from four different cDNA clones obtained from a 28 days infected mouse (named as in figure S3A) corresponding to the 5' junction between RL_2 and LAT sequences.

The intron 1 (resp. exon 1) RL_2 sequence is indicated by the black (resp. blue) line above the histogram and LAT sequence is indicated by a magenta line. Black arrows above the histograms indicate the position of the junction between RL_2 and LAT_{2kb} sequences.

Figure S4A

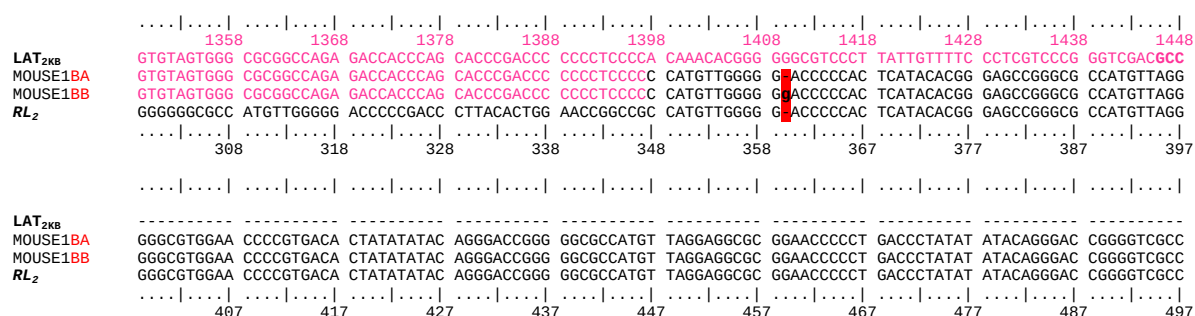
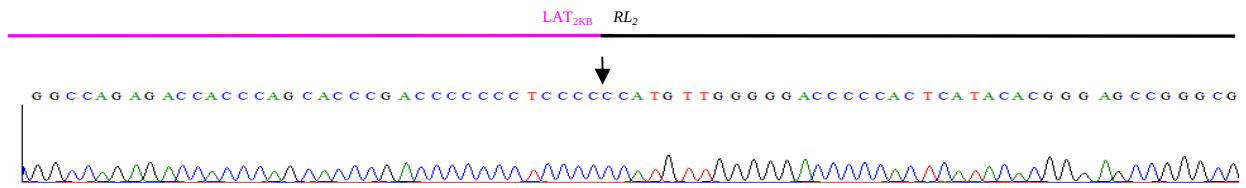


Figure S4A. Alignment of sequences of PCR products, mapping the 3' junction between LAT_{2KB} insertion and intron 1 of *RL2* gene, with *RL2* and LAT sequences.

PCR products were obtained from DNA isolated from a 28 days infected mouse TG. The sequences of 2 different clones are aligned with the sequences of LAT_{2kb} and *RL2* to identify the 3' junction between LAT and *RL2* sequences. Names of the sequences are indicated on the left. Numbering and characters in black refer to the *RL2* sequence; numbering and characters in magenta refer to the LAT_{2KB} sequence. The red box on the alignment indicates an additional G (in lower case and bold) in the *RL2* sequence of one of the modified gene. This additional nucleotide has been positioned arbitrarily at the end of the G stretch.

Figure S4B

MOUSE1BA



MOUSE1BB

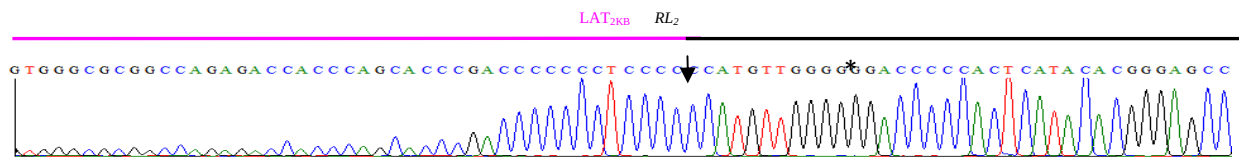


Figure S4B. Part of the histograms of the sequences of two different PCR products from a 28 days infected mouse (named as in figure S4A) corresponding to the junction between LAT_{2KB} and RL₂ sequences.

The RL₂ sequence is indicated by the black line above the histogram and LAT sequence is indicated by a magenta line. Black arrows above the histograms indicate the position of the junction between LAT and RL₂ gene. The asterisk indicates an additional G inserted in the RL₂ sequence of mouse1BB clone. This guanosine, which has been found in 14 clones out of 15, has been positioned arbitrarily at the end of the G stretch.

Figure S5A



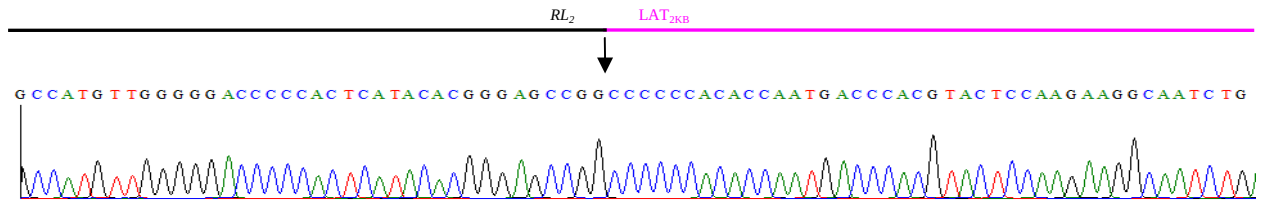
Figure S5A. Alignment of sequences of PCR products, mapping the 5' junction between intron 1 of *RL2* and *LAT*_{2KB} insertion, with *RL2* and *LAT* sequences.

PCR products were obtained from DNA isolated from 28 days infected mice TG.

The sequences of two different clones are aligned with the sequences of *RL2* and *LAT*_{2kb} to identify the 5' junction between *RL2* and *LAT* sequences for each clone. Names of the sequences are indicated on the left. Characters in blue refer to the *RL2* first exon sequence, characters in black refer to the *RL2* first intron sequence and character in magenta refer to *LAT*_{2KB} sequence. Numbering in black refers to the *RL2* sequence and numbering in magenta refers to the *LAT*_{2kb} sequence.

Figure S5B

MOUSE2AA



MOUSE2AB

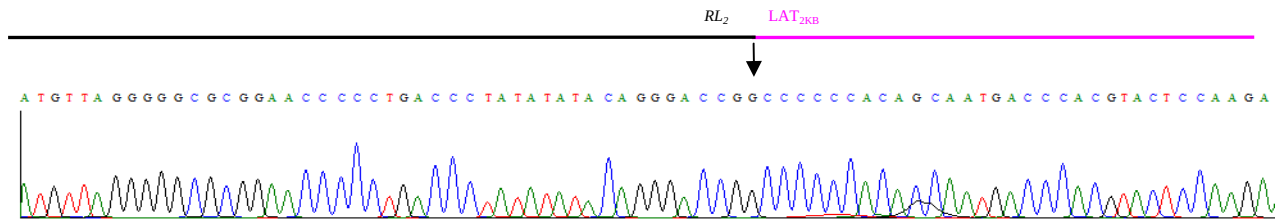
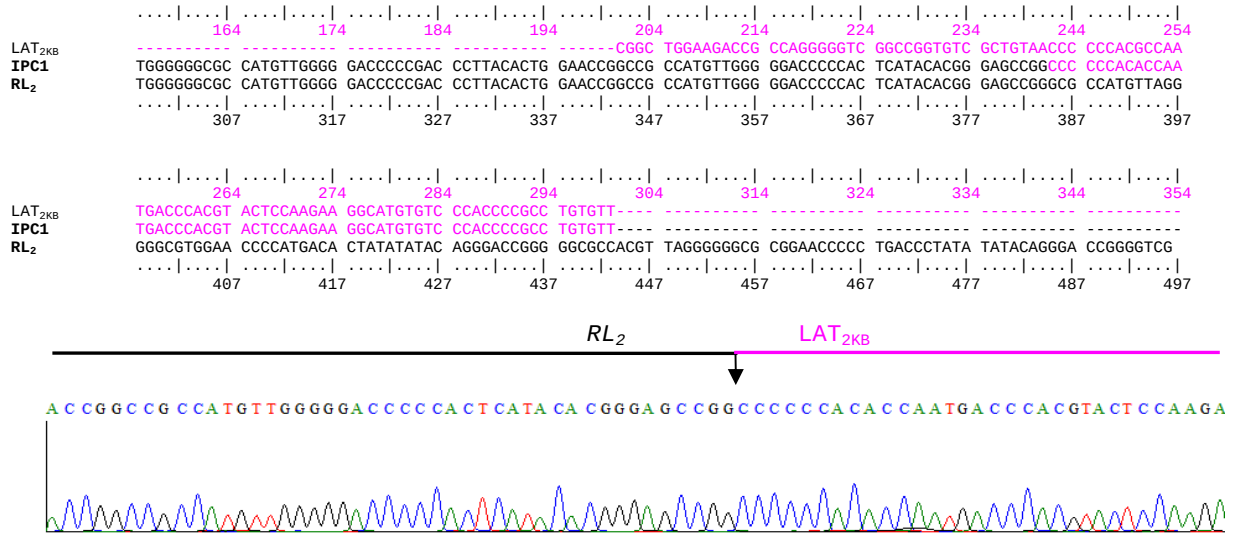


Figure S5B. Part of the histograms of the sequences of PCR products from two different clones obtained from a 28 days infected mouse (named as in figure S5A) corresponding to the 5' junction between RL_2 and LAT sequences.

The RL_2 sequence is indicated by the black line above the histogram and LAT sequence is indicated by a magenta line. Black arrows above the histograms indicate the position of the junction between RL_2 and LAT_{2kb} sequences.

Figure S6

A



B

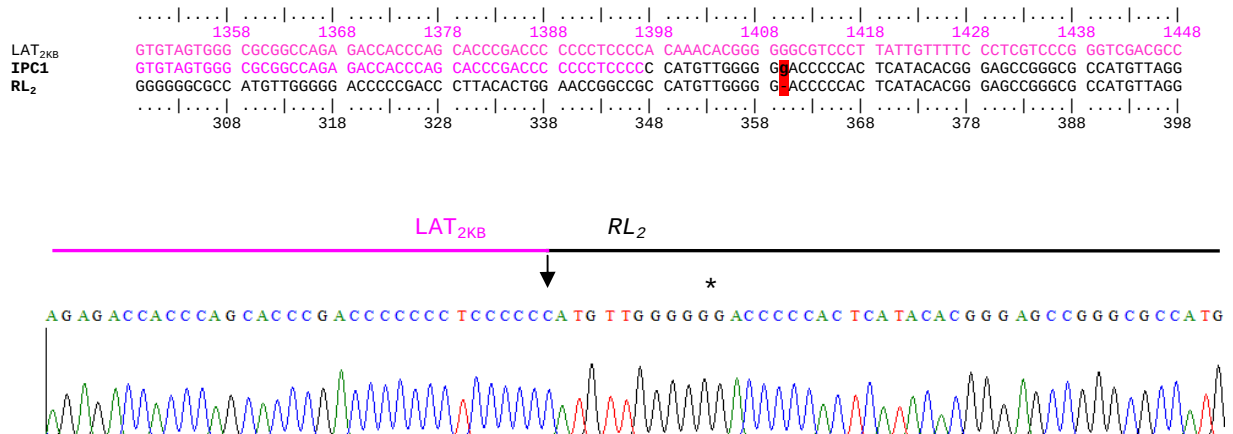
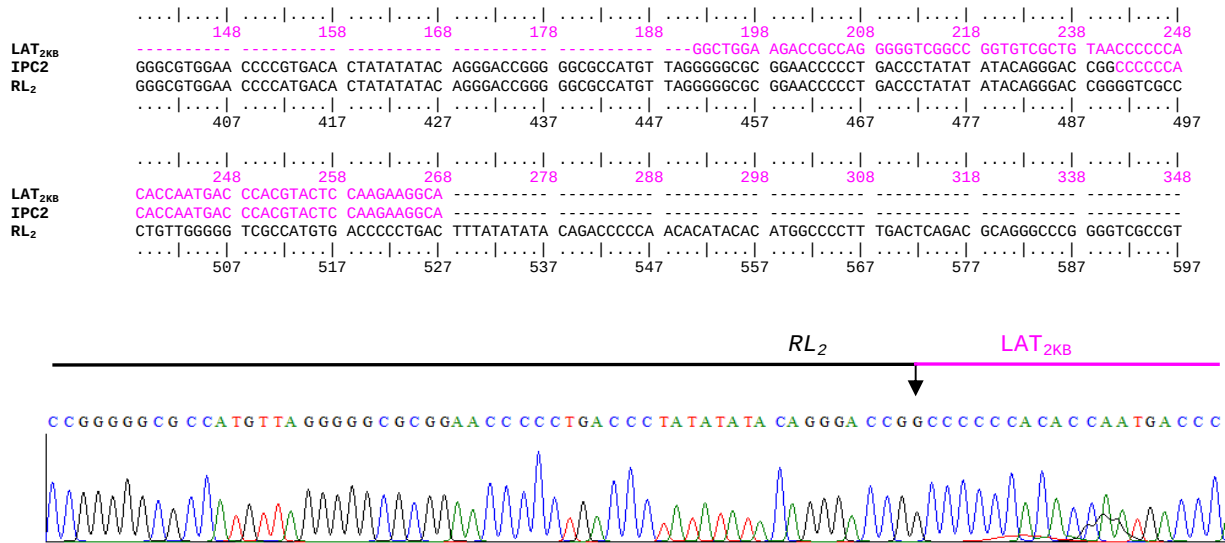


Figure S6. Chromatogram and alignment of sequences of an iPCR clone at the level of the 5' junction (A) and the 3' junction (B) between intron 1 of *RL2* and *LAT*_{2KB} insertion.

iPCR products were obtained from DNA isolated from 28 days infected mice TG, digested by XhoII and self-ligated. The two chromatograms were obtained from the same clone. The sequences are aligned with the sequences of *RL2* and *LAT*_{2kb} to identify the 5' (A) and 3' (B) junction between *RL2* and *LAT* sequences. Names of the sequences are indicated on the left. Characters in black refer to the *RL2* first intron sequence and characters in magenta refer to *LAT*_{2kb} sequence. Numbering in black refers to the *RL2* sequence and numbering in magenta refers to the *LAT*_{2kb} sequence.

Figure S7

A



B

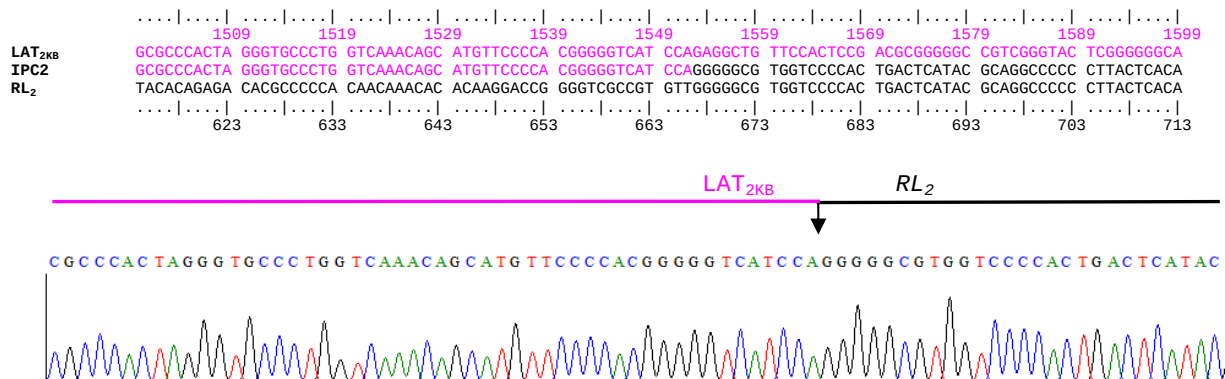


Figure S7. Chromatogram and alignment of sequences of an iPCR clone at the level of the 5' junction (A) and the 3' junction (B) between intron 1 of *RL2* and *LAT*_{2KB} insertion.

iPCR products were obtained from DNA isolated from 28 days infected mice TG, digested by XhoII and self-ligated. The two chromatograms were obtained from the same clone. The sequences are aligned with the sequences of *RL2* and *LAT*_{2kb} to identify the 5' (A) and 3' (B) junction between *RL2* and *LAT* sequences. Names of the sequences are indicated on the left. Characters in black refer to the *RL2* first intron sequence and characters in magenta refer to *LAT*_{2kb} sequence. Numbering in black refers to the *RL2* sequence and numbering in magenta refers to the *LAT*_{2kb} sequence.

Figure S8A



Figure S8A. Sequences of the products obtained by primer extension assay of total poly-A transcripts from 28 days infected mouse TG

The two sequences are aligned with those of RL₂ (blue and black) and LAT_{2kb} (magenta). The exon sequence of the RL₂ gene is indicated in blue. LAT sequence found in clone PELAT2 is in magenta. Numbering in blue refers to RL₂ sequence, and numbering in magenta refers to LAT_{2kb} sequence.

Figure S8B

PELAT2

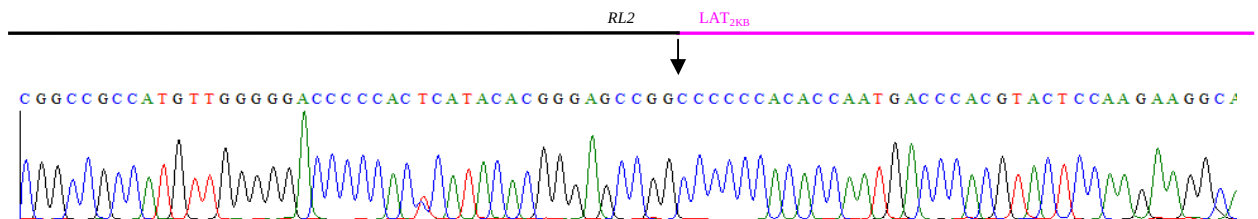


Figure S8B. Part of the histogram of PELAT2 sequences obtained from primer extension assay corresponding to the junction between RL₂ and LAT_{2kb} sequences.

The RL₂ sequence is indicated by the black line above the histogram and LAT sequence is indicated by a magenta line. Black arrows above the histograms indicate the position of the junction between RL₂ and LAT gene.

Figure S9A

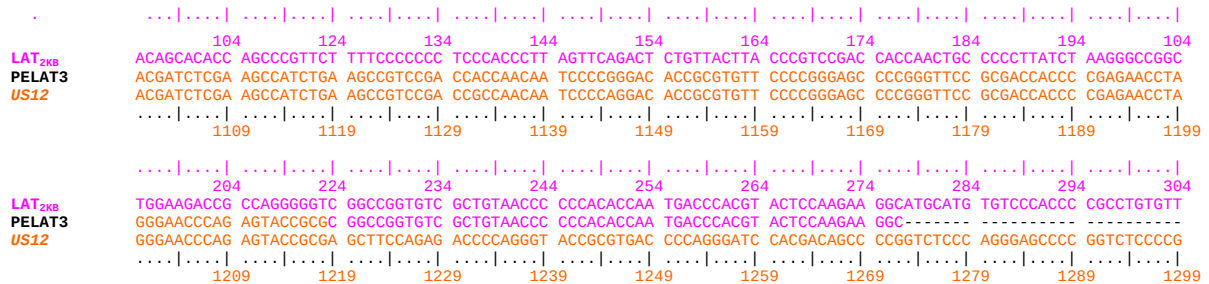


Figure S9A. Sequences of the products obtained by primer extension assay of total poly-A transcripts from 28 days infected mouse TG.

The two sequences are aligned with those of *US12* (orange) and *LAT*_{2kb} (magenta). The exon 2 sequence of the *US12* gene is indicated in orange. *LAT* sequence found in clone PELAT3 is in magenta. Numbering in orange refers to *US12* sequence, and numbering in magenta refers to *LAT*_{2kb} sequence

Figure S9B

PELAT3

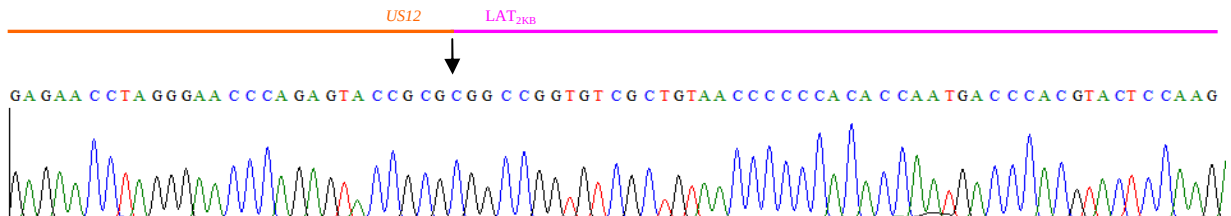


Figure S9B. Part of the histogram of PELAT3 sequence obtained from primer extension assay corresponding to the junction between *US12* and *LAT*_{2KB} sequences.

The *US12* sequence is indicated by the orange line above the histogram and *LAT* sequence is indicated by a magenta line. Black arrows above the histograms indicate the position of the junction between *US12* and *LAT* gene.

Figure S10A



Figure S10A. Alignment of sequences of RT PCR products, mapping the 3' junction between LAT_{2KB} and US12 insertion, with US12 and LAT sequences.

RT PCR products were obtained from DNA isolated from 28 days infected mice TG.

The sequence of the clone is aligned with the sequences of LAT_{2kb} and US12 to identify the 5' junction between LAT and US12 sequences. Names of the sequences are indicated on the left. Characters and numbering in orange refer to the US12 sequence; Characters and numbering in magenta refer to LAT_{2KB} sequence.

Figure S10B

MOUSE22RA

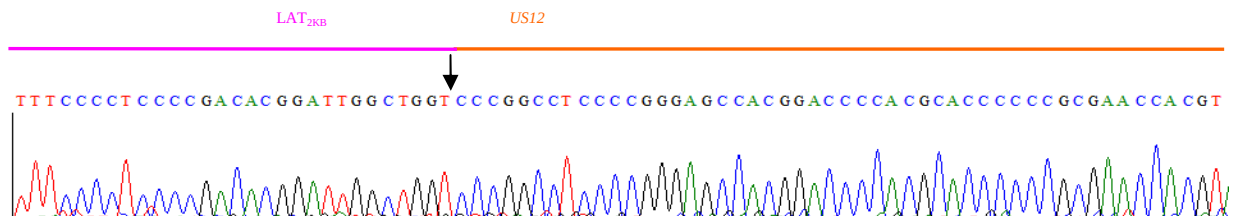


Figure S10B. Part of the histograms of the sequences of RT PCR products obtained from a 28 days infected mouse corresponding to the 3' junction between US12 and LAT sequences.

The US12 sequence is indicated by the orange line above the histogram and LAT sequence is indicated by a magenta line. Black arrows above the histograms indicate the position of the junction between LAT_{2kb} and US12 sequences.

Figure S11A

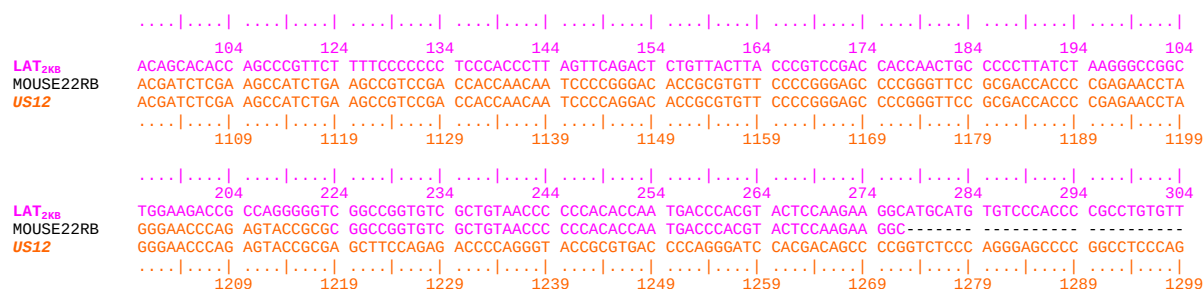


Figure S11A. Alignment of sequences of RT PCR products, mapping the 5' junction between *US12* and *LAT*_{2KB} insertion, with *US12* and *LAT* sequences.

RT PCR products were obtained from DNA isolated from 28 days infected mice TG.

The sequence of a clone is aligned with the sequences of *US12* and *LAT*_{2kb} to identify the 5' junction between *US12* and *LAT* sequences. Names of the sequences are indicated on the left. Characters and numbering in orange refer to the *US12* sequence; Characters and numbering in magenta refer to *LAT*_{2KB} sequence.

Figure S11B

MOUSE22RB

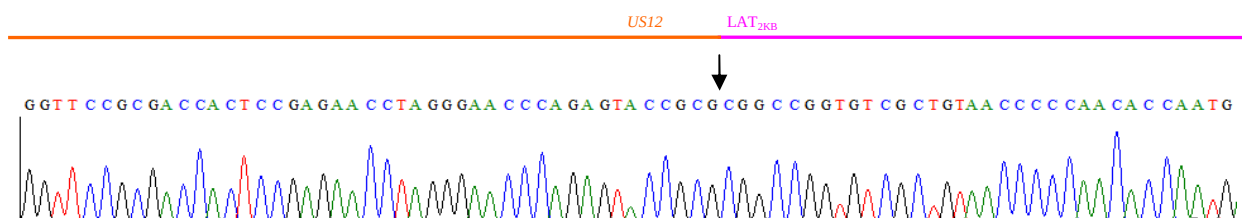


Figure S11B. Part of the histograms of the sequence of RT PCR product obtained from a 28 days infected mouse corresponding to the 5' junction between *US12* and *LAT* sequences.

The *US12* sequence is indicated by the orange line above the histogram and *LAT* sequence is indicated by a magenta line. Black arrows above the histograms indicate the position of the junction between *US12* and *LAT*_{2kb} sequences.

Figure S12A

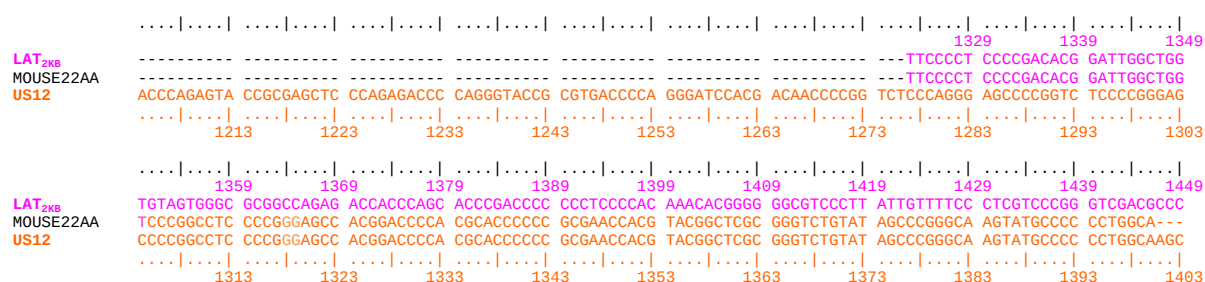


Figure S12A. Alignment of sequences of PCR products, mapping the 3' junction between LAT_{2KB} and US12 sequences.

PCR products were obtained from DNA isolated from 28 days infected mice TG.

The sequence of the clone is aligned with the sequences of LAT_{2kb} and US12 to identify the 3' junction between LAT_{2kb} and US12 sequences for each clone. Names of the sequences are indicated on the left. Characters and numbering in orange refer to the US12 sequence; characters and numbering in magenta refer to LAT_{2KB} sequence.

Figure S12B

MOUSE22AA

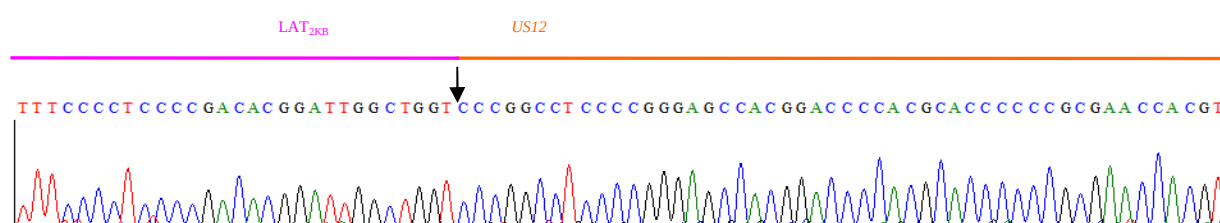


Figure S12B. Part of the histograms of the sequences of PCR products obtained from a 28 days infected mouse corresponding to the 3' junction between US12 and LAT sequences.

The US12 sequence is indicated by the orange line above the histogram and LAT sequence is indicated by a magenta line. Black arrows above the histograms indicate the position of the junction between LAT_{2kb} and US12 sequences.

Figure S13A

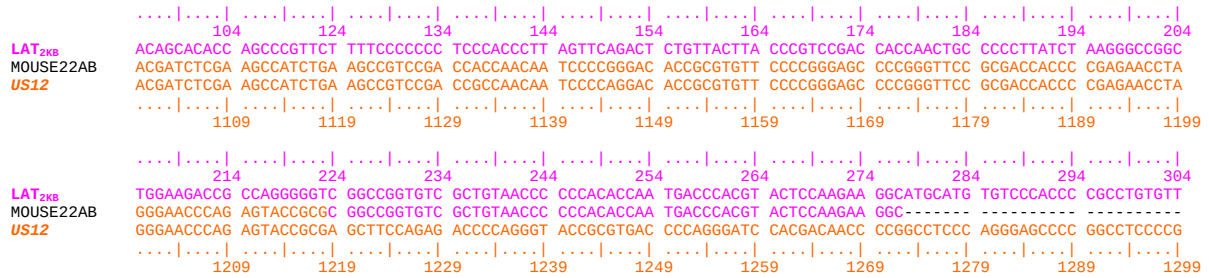


Figure S13A. Alignment of sequences of PCR products, mapping the 5' junction between US12 and LAT_{2KB} insertion, with US12 and LAT sequences.

PCR products were obtained from RNA isolated from 28 days infected mice TG.

The sequence of the clone is aligned with the sequences of US12 and LAT_{2kb} to identify the 5' junction between US12 and LAT sequences for each clone. Names of the sequences are indicated on the left. Characters and numbering in orange refer to the US12 sequence; Characters and numbering in magenta refer to LAT_{2KB} sequence.

Figure S13B

MOUSE22AB

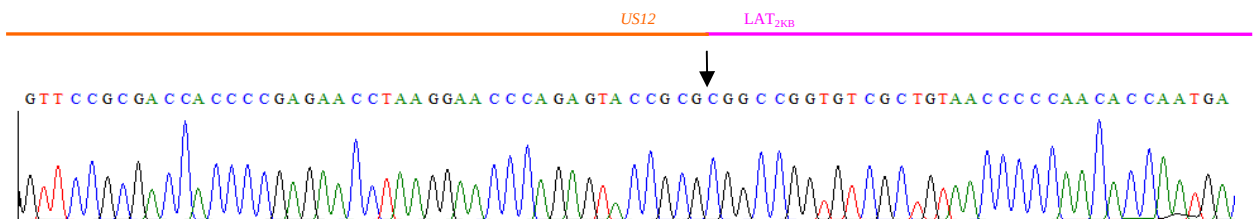
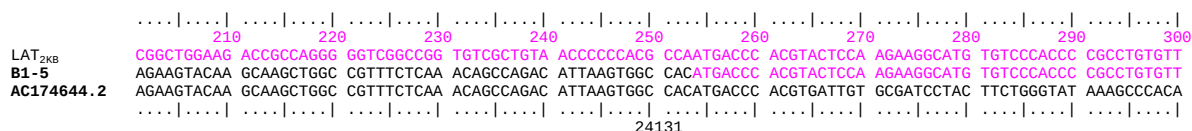


Figure S13B. Part of the histograms of the sequences of PCR products obtained from a 28 days infected mouse corresponding to the 5' junction between US12 and LAT sequences.

The US12 sequence is indicated by the orange line above the histogram and LAT sequence is indicated by a magenta line. Black arrows above the histograms indicate the position of the junction between US12 and LAT_{2kb} sequences.

A



B

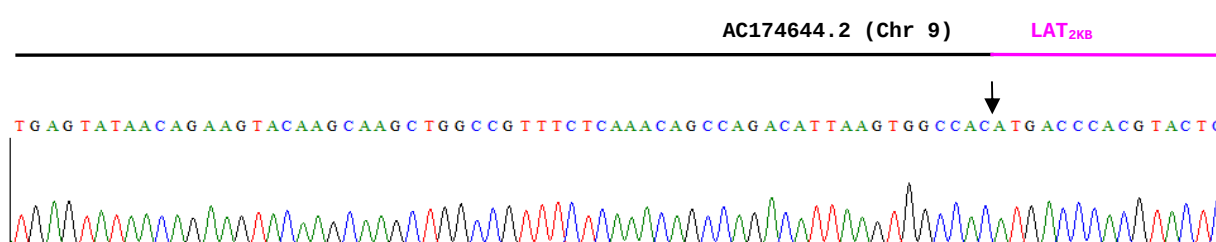


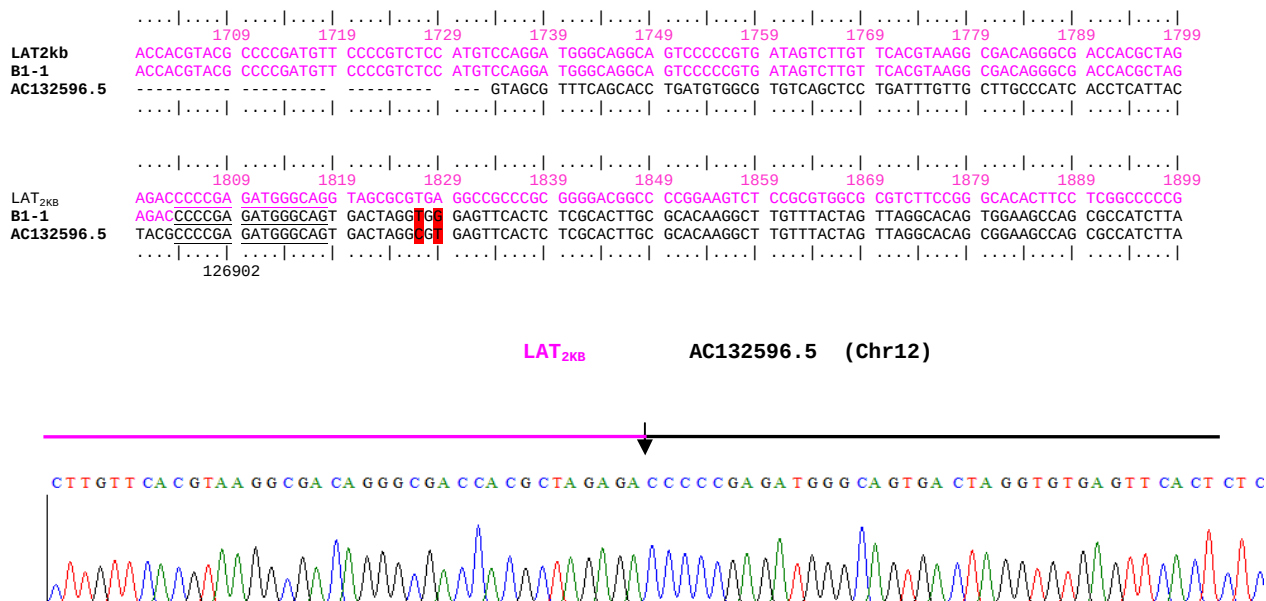
Figure S14. Sequence and histogram mapping the 5' junction of LAT insertion in chromosome 9 as identified by B1-PCR.

A) The sequence of the clone (B1-5) is aligned with the sequences of chromosome 9 (AC174644.2) and LAT_{2kb}. Names of the sequences are indicated on the left. Characters and numbering in black refer to the chromosome 9 sequence; Characters and numbering in magenta refer to LAT_{2KB} sequence.

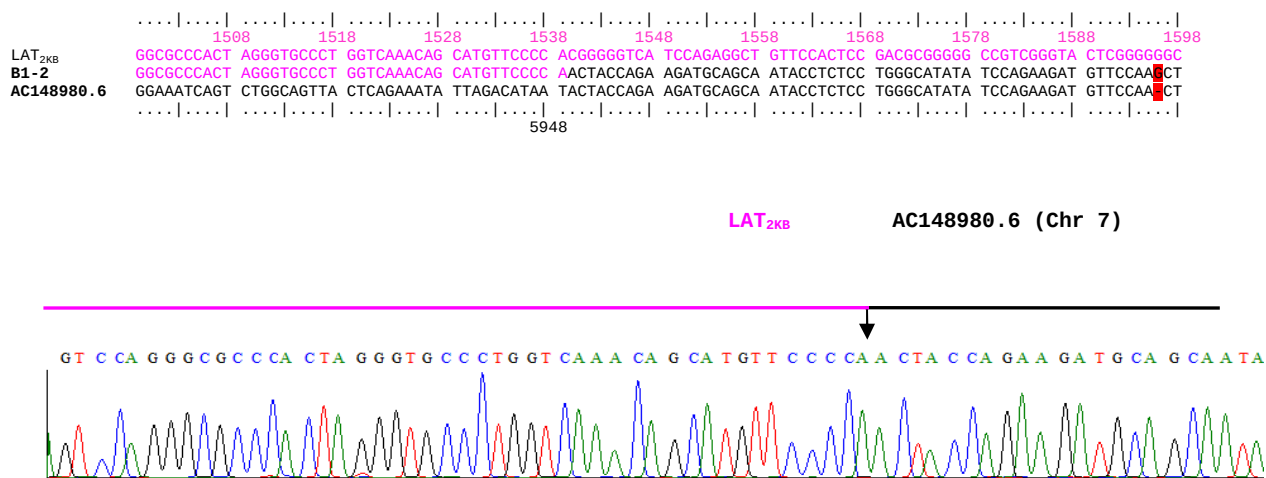
B) The host sequence is indicated by the black line above the histogram and LAT_{2kb} sequence is indicated by a magenta line. The black arrow above the histograms indicates the position of the junction between chromosome 9 and LAT_{2kb} sequences.

Figure S15

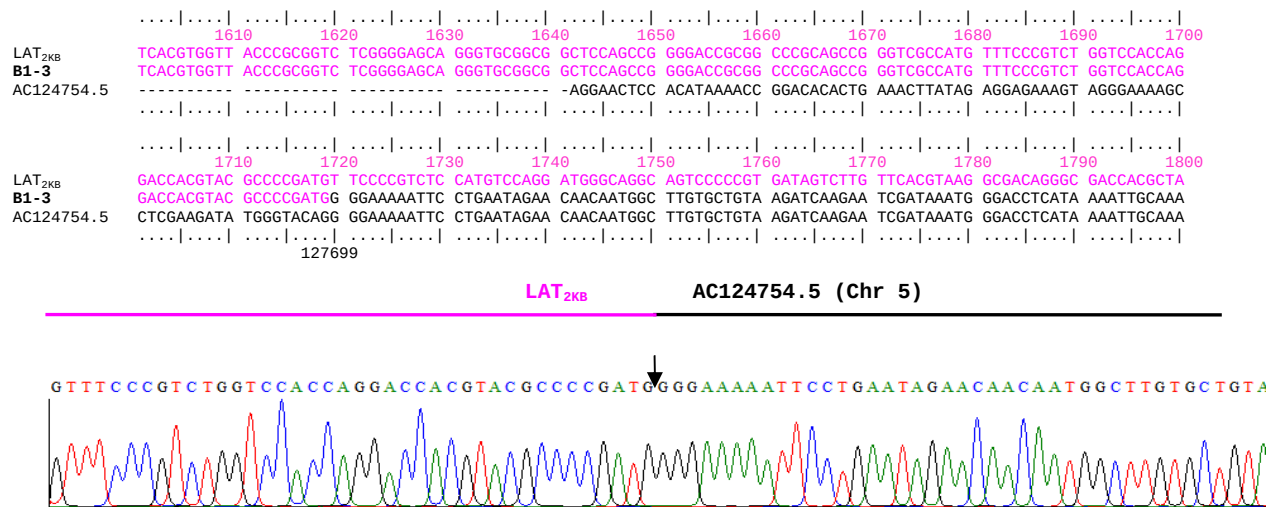
A



B



C



D

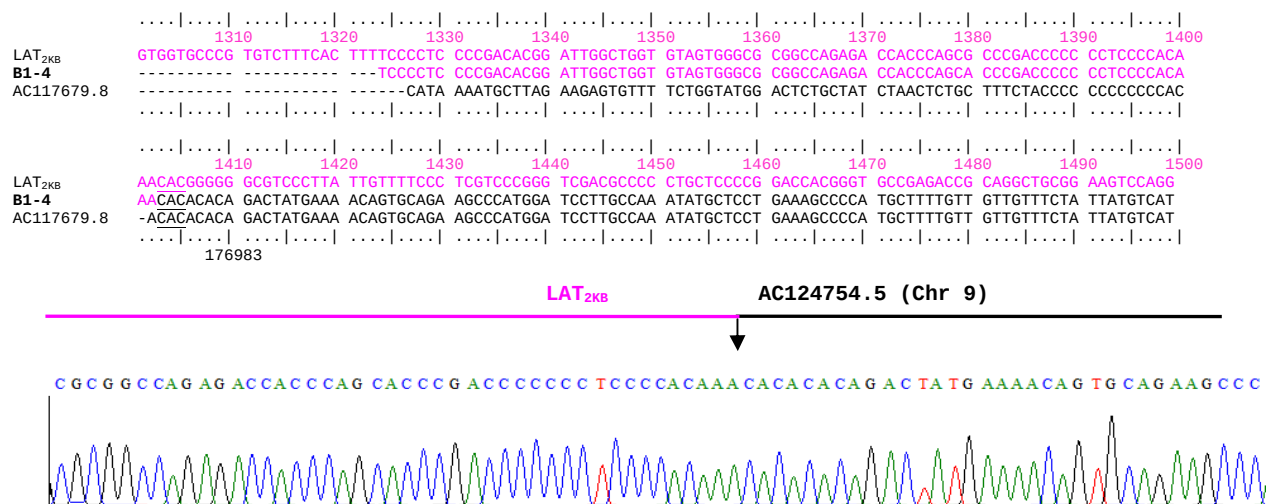


Figure S15. Sequences and histograms mapping the 3' junctions of LAT insertions in chromosome 12 (A), 7 (B), 5 (C) and 9 (D) as identified by B1-PCR.

For each panel, the top half shows the sequence of the clone which is aligned with the sequences of the corresponding chromosome and LAT_{2kb}. Names of the sequences are indicated on the left. Characters and numbering in black refer to the chromosome sequence; Characters and numbering in magenta refer to LAT_{2KB} sequence. The bottom half shows the corresponding histogram. The host sequence is indicated by the black line above the histogram and LAT_{2kb} sequence is indicated by the magenta line. The black arrow above the histograms indicates the position of the junction between host DNA and LAT_{2kb} sequences.

In panel A and D, the underlined characters at the junction cannot be attributed due to the fact that the LAT and the chromosome sequences are the same. Red boxes in A and B indicate differences between the chromosome and clone sequences.

Mouse	Day Post- infection	Log ₁₀ (RL ₂ unspliced transcripts)	SD	Log ₁₀ (Modified RL ₂ transcripts)	SD
M13	28	5.50	±0.37(5)	±5.86	±0.02(2)
M14	28	5.28	±0.19(5)	±5.93	±0.08(2)
M15	28	5.51	±0.27(5)	±5.66	±0.22(5)
M16	28	5.72	±0.05(5)	±5.09	±0.66(5)
M17	28	5.59	±0.18(5)	±5.61	±0.38(5)
M18	28	5.60	±0.18(5)	±5.37	±0.20(5)
M19	28	5.72	±0.3(5)	±5.67	±0.60(5)
M20	28	5.75	±0.17(5)	±5.47	±0.37(5)

Table S1: Quantification of the amount of unmodified and modified transcripts from RNA extracted from TG of mice infected 28 days.

Quantification of unmodified RL₂ transcripts and RL₂ transcripts modified by LAT's insertion was performed by real-time qPCR as described in experimental procedures section. Mice were euthanized on the indicated day. TG were harvested and used for RNA extraction. Values are indicated as the mean of experiments conducted in.

^bAbbreviations: SD, standard deviation; NA, not applicable.

Chapitre VI :

Résultats 2nd partie

VI.1 Introduction

Dans cette seconde partie, nous nous sommes intéressés au phénomène de transposition du LAT au moment de la phase aigüe de l'infection. En effet, bien qu'il ait été montré que le LAT n'était pas indispensable à l'établissement de la latence, nous avons souhaité savoir si les événements de transposition observés pendant la latence étaient déjà présents à des stades plus précoces de l'infection. Leur détection à ces stades précoces permettrait de spéculer sur la participation des LATs à l'établissement de la latence.

Afin de réaliser cette étude, un lot de souris a été infecté par la souche SC16 du virus HSV-1. Cinq souris ont été sélectionnées, 6 jours post-infection, sur des critères physiologiques témoignant de l'établissement de la maladie tels que l'apparition d'un pelage hérissé ou bien encore une diminution des capacités exploratrices de la souris.

Une fois les animaux sacrifiés, les ganglions trigémiques gauches ont été prélevés et coupés en deux. Le premier tiers a été utilisé pour extraire l'ADN. La seconde partie, correspondant aux deux autres tiers restant, a été utilisée pour en extraire les ARN totaux.

VI.2 Recherche des jonctions par PCR

L'étape suivante de ce travail a consisté en la recherche des insertions du LAT par PCR dans les deux gènes viraux identifiés comme étant modifiés au moment de la phase latente de l'infection (*RL2* et *US12*). Cette recherche a été faite dans des extraits d'ADN provenant des ganglions trigémiques de souris,

Malgré plusieurs tentatives et différentes approches de PCR, telles que l'augmentation du nombre de cycle ou l'emploi de PCR nichées, il n'a pas été possible d'obtenir le moindre produit d'amplification correspondant à l'une des jonctions observées au moment de l'infection latente et cela quel que soit la jonction recherchée.

L'absence de détection de gènes modifiés par le LAT peut être due au simple fait que ceux-ci ne sont pas assez nombreux pour être détectés par des méthodes de PCR conventionnelles.

Cependant, si, comme observé au moment de la latence, les génomes modifiés continuent à transcrire des ARNm contenant une insertion du LAT, les jonctions associées devraient être plus faciles à détecter dans ces ARN. Dans un premier temps, il a donc été décidé de rechercher des séquences de LAT insérées dans les ARNm des gènes *US12* et *RL2*.

VI.3 Analyse des jonctions et des sites d'insertion du LAT dans les transcrits viraux et cellulaires au moment de la phase aigüe de l'infection virale.

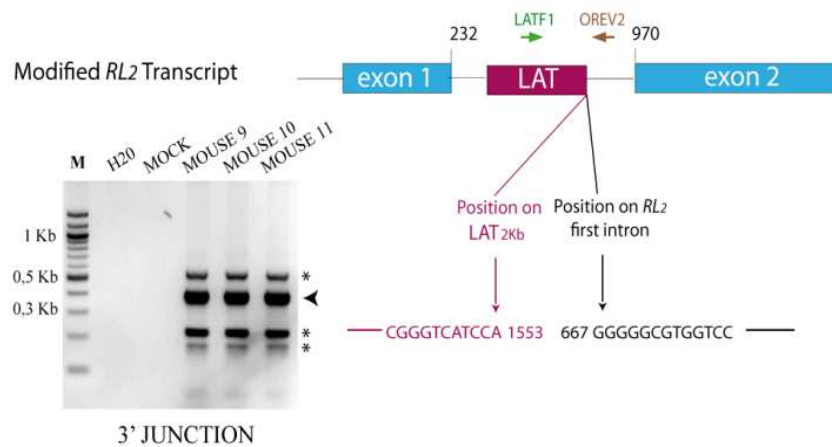
VI.3.1 Caractérisation des transcrits du gène RL2 modifiés par insertion du LAT.

A partir des ARN totaux extraits des ganglions, des RT-PCR utilisant le couple d'amorces (LATF1, CPL1R) permettant l'amplification de la jonction 3' la plus fréquente du LAT inséré dans le gène *RL2* [(LAT-nt1397)/(nt348-*RL2*)] ont été effectuées. Celles-ci n'ont néanmoins pas permis d'observer de produits d'amplification.

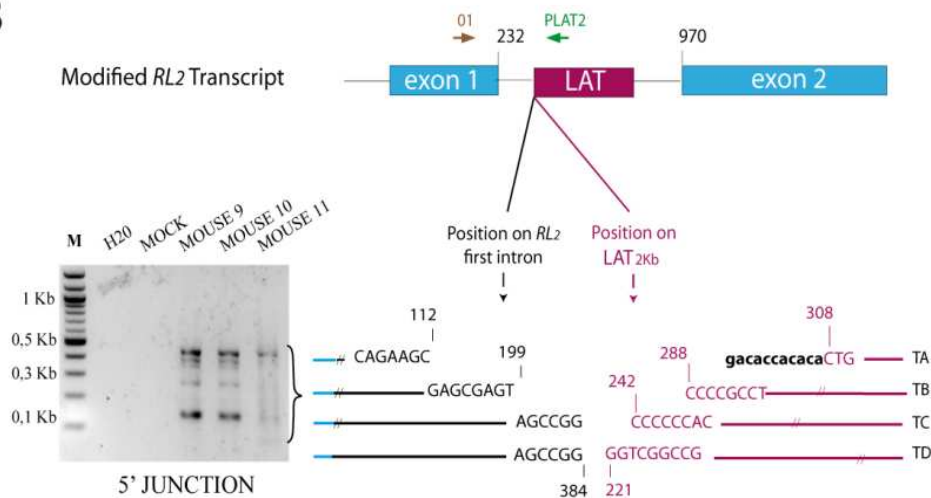
En revanche, l'utilisation du couple d'amorces (LATF1, OREV2) a permis l'amplification de quatre produits de RT-PCR. Chacun d'entre eux, a été cloné et séquencé séparément, mais un seul correspondait à une jonction valide (*RL2*-nt1553) / (nt667-LAT) (figure 1A). Les données de séquençage sont en annexe de ce chapitre (Figure S1A et S1B).

De façon surprenante, l'analyse des trois autres bandes, pourtant absentes des puits correspondant aux pistes H₂O et MOCK infectée, laissant supposer que celles-ci étaient spécifiques de produits d'amplifications correspondant à des jonctions entre le LAT et les transcrits *RL2*, a révélé, que l'une d'entre elles (la bande la plus haute), était exclusivement composée de nucléotides correspondant au gène *RL2*, alors que les deux autres correspondaient à des séquences virales sur lesquelles les amorces ne pouvaient en théorie pas s'hybrider (gène codant pour la protéine ICP34.5).

A



B



MOUSE	MOUSE 9	MOUSE 10	MOUSE 11
sequence			
TA	1/9	0/10	2/8
TB	1/9	3/10	2/8
TC	5/9	6/10	3/8
TD	2/9	1/10	1/8

Figure1 : Détection par RT-PCR des transcrits du gène *RL2* modifiés durant la phase aigüe de l'infection (6 jours p.i.).

Les souris infectées ont été sacrifiées 6 jours p.i. et les ganglions trigéminés gauches ont été extraits. L'ARN total a été isolé et utilisé comme matrice pour effectuer la recherche par RT-PCR des jonctions 3' (A) et 5' (B) entre le LAT_{2kb} et l'intron 1 des transcrits du gène *RL2*. Dans chaque cas, un schéma des transcrits modifiés du gène *RL2* est présenté. Les couples d'amorces (LATF1 et OREV2) utilisés pour amplifier la jonction 3' (A) ainsi que ceux utilisés pour amplifier la jonction 5' (O1 et PLAT2) sont indiqués par des flèches.

Les gels d'électrophorèse montrant les produits de RT-PCR obtenus, se trouvent sur la gauche de chaque panel. Un contrôle sans ARN a également été effectué (H20).

À droite des gels, sont représentées les séquences des jonctions obtenues après séquençage des produits amplifiés, le numéro des nucléotides au niveau de la jonction entre le gène *RL2* et le LAT est indiqué.

En (A), aucune séquence *RL2* n'a été retrouvée dans les produits de PCR indiqués par des astérisques.

En (B), les nucléotides en lettres minuscules grasses n'ont pu être attribués ni à *RL2*, ni au LAT. Le tableau indique les fréquences relatives de chaque jonction.

Nous avons ensuite cherché à amplifier la jonction 5' du LAT inséré dans le gène *RL2* par RT-PCR en utilisant les amorces O1, PLAT2. La migration des produits d'amplification sur gel a révélé plusieurs bandes dans les pistes correspondant aux extraits d'ARN provenant de souris infectées, alors qu'aucun produit n'a été détecté dans les pistes contrôles (RT-PCR réalisées sur de l'eau ou sur de l'ARN extrait de la souris MOCK infectée).

Les produits d'amplification ainsi obtenus ont été clonés. Le séquençage de 49 clones a permis la mise en évidence de quatre types de jonctions différentes (figure 1B ; figure S2A et S2B).

Deux d'entre elles, les jonctions (RL2-nt384) / (nt221-LAT), (RL2-nt384) / (nt242-LAT) ont été déjà observées dans des ARN modifiés au moment de la phase de latence du virus (28 jours p.i).

Les deux autres jonctions (RL2-nt199) / (nt288-LAT) et (RL2-nt112) / gacaccacaca / (nt 308-LAT), quant à elles, ont été mises en évidence pour la première fois à cette occasion. D'ailleurs, la dernière d'entre elles possède une particularité puisqu'elle comprend une insertion de 11 nucléotides entre le dernier nucléotide attribuable au gène *RL2* (nt112) et le premier nucléotide attribuable au LAT (nt308) (figure 1B. figure S2A et S2B).

VI.3.2 Caractérisation des transcrits du gène *US12* modifié par insertion du LAT

Nous avons réalisé le même type d'expériences pour les transcrits du gène *US12*. L'utilisation du couple d'amorces 472F, PLAT2, a permis l'amplification de deux produits de RT-PCR majoritaires qui correspondaient aux jonctions 5' du LAT inséré dans des transcrits du gène *US12* (figure 2B ; figure S3A et S3B).

La première jonction (US12-nt1218) / (nt224-LAT) a déjà été observée dans des ARN modifiés au moment de la phase de latence du virus, en revanche la seconde jonction, (US12-nt1345) / (nt129-LAT), a été mise en évidence pour la première fois.

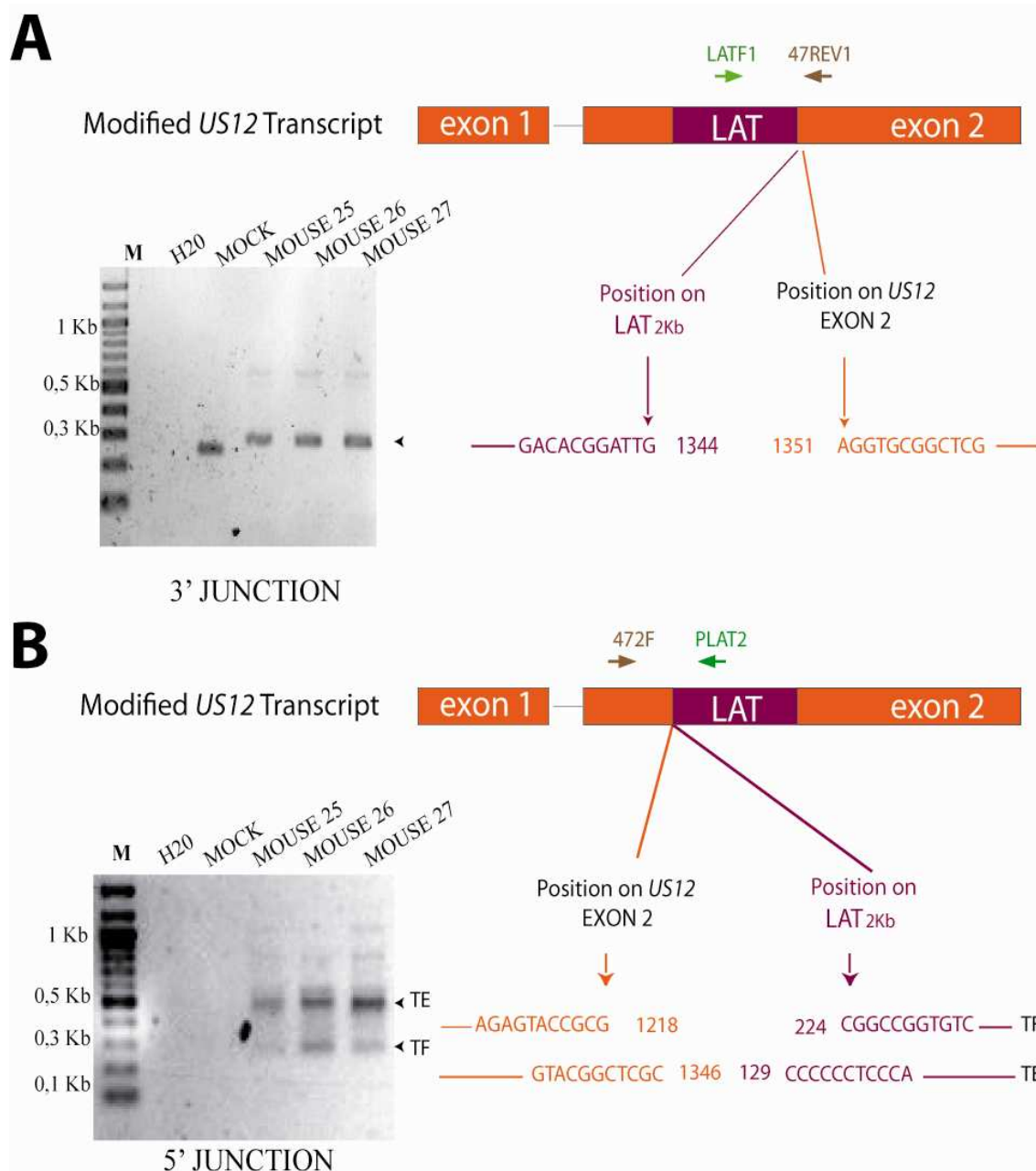


Figure 2: Détection des transcrits du gène *US12* modifiés par RT PCR au moment de la phase aigüe de l'infection 6 jours P.I

Les souris infectées ont été sacrifiées 6 jours p.i. et les ganglions trigéminés gauches ont été extraits. L'ARN total a été isolé et utilisé comme matrice pour effectuer la recherche par RT-PCR des jonctions 3' (A) et 5' (B) entre le LAT_{2kb} et les transcrits du gène *US12*.

Dans chaque cas de figure, une représentation schématisée des transcrits modifiés du gène *US12* a été réalisée. Les couples d'amorces (LATF1 et 47REV1) utilisés pour amplifier la jonction 3' (A) ainsi que ceux utilisés pour amplifier la jonction 5' (472F et PLAT2) sont indiqués par des flèches.

Les gels d'électrophorèse montrant les produits de RT-PCR obtenus sont sur la gauche de chaque panel. Un contrôle sans ARN a également été testé (H20).

À droite des gels, sont représentées, à la fois, les séquences des jonctions obtenues après séquençage des produits de RT-PCR, ainsi que le numéro des nucléotides au niveau de la jonction pour le gène correspondant. En A) les caractères en minuscule et gras correspondent à des conversions de nucléotides.

L'utilisation du couples d'amorces LATF1 ; 47REV1, a permis l'amplification d'un produit de RT-PCR unique correspondant à une jonction 3' entre le LAT et le transcrit du gène *US12*. Le clonage puis le séquençage de ce produit de PCR ont permis de révéler la jonction (LAT- nt1344) / (nt1351-Us12) (figure 2A ; figure S4A et S4B). En aval de la jonction deux nucléotides de la séquence d'*US12* étaient modifiés (la cytosine 1352 et l'adénine 1355 étaient toutes deux changées en guanine).

Parallèlement, a été amplifié un produit non spécifique à partir des ARN extraits des souris MOCK infectées. Le séquençage de ce produit a révélé une séquence correspondant à un transcrit murin sur lequel les amorces ne pouvaient en théorie pas s'hybrider.

VI.4 Recherche de l'ensemble des transcrits viraux et cellulaires susceptible d'être modifiés par le transcrits LAT.

Dans le but d'isoler l'ensemble des ARN viraux et cellulaires modifiés par le LAT, il a été décidé d'utiliser une approche plus globale similaire à celle utilisée dans le chapitre précédent.

Nous avons donc purifié les ARNm poly-adenylés à l'aide d'une chromatographie d'affinités sur une colonne dont la phase fixe est pourvue de fragments d'oligo (dT). Le choix d'isoler uniquement les ARNm avait deux principaux objectifs : caractériser l'ensemble des fonctions cellulaires et virales pouvant être affectées par la transposition du LAT, tout en se débarrassant des LAT_{2kb} ou LAT_{1,5kb} non insérés. Ces derniers n'étant pas polyadenylés.

Une fois les transcrits poly-adenylés purifiés, une extension d'une amorce spécifique du LAT a été effectuée par RT (figure 3A). Une PCR a été ensuite effectuée sur les ADNc ainsi obtenus. La migration des produits d'amplification sur gel a révélé plusieurs bandes dans les pistes correspondant aux extraits d'ARN provenant de souris infectées, alors qu'aucun produit n'a été détecté dans les pistes contrôles (RT-PCR effectuée sur H₂O ou sur de l'ARN extrait de la souris MOCK). (figure 3B).

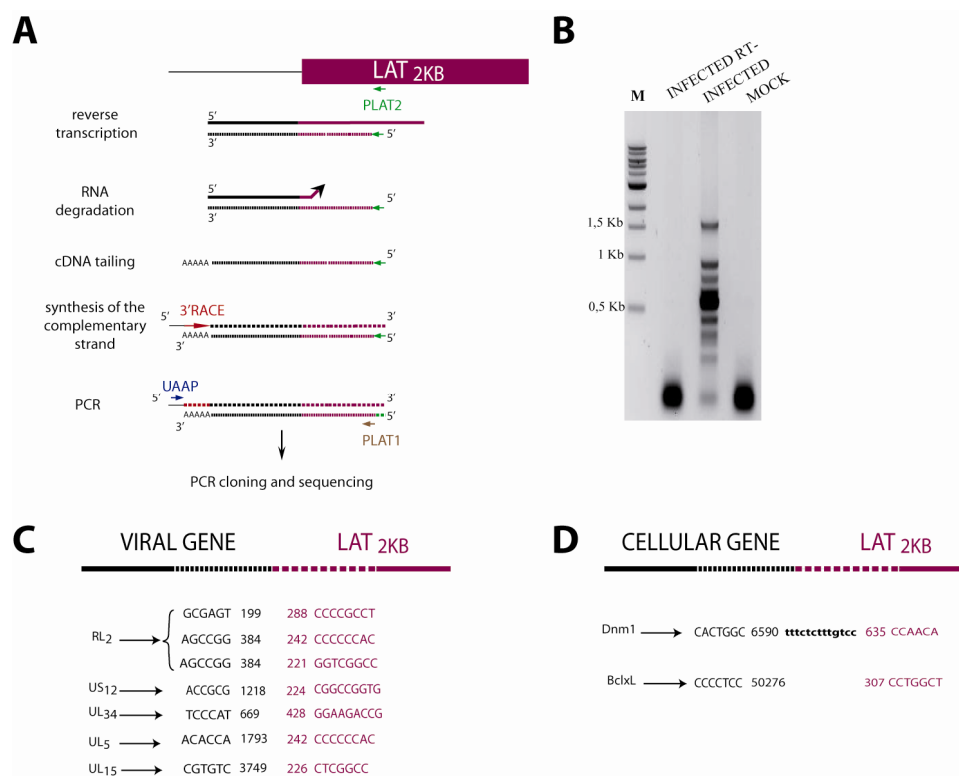


Figure 3: détection des transcrits viraux et cellulaires modifiés par 5'RACE au moment de la phase aigue de l'infection 6 jours P.I

En A, les principales étapes du 5'RACE utilisé pour amplifier les différents types d'ARN modifiés par insertion du LAT sont décrites. Le gel d'électrophorèse montrant les produits de PCR obtenus est présenté en B. Un contrôle sans transcriptase inverse a été effectué, en plus du contrôle utilisant des ARN extraits d'une souris MOCK infectée. En C et D sont représentées, à la fois, les séquences des jonctions obtenues après séquençage des produits de PCR, ainsi que le numéro des nucléotides engagés dans la jonction pour le gène correspondant. Pour le gène codant la dynamine, les caractères en minuscule et gras correspondent à une séquence nucléotidique qui n'a pu être attribuée ni au gène, ni au LAT

L'ensemble des produits obtenus a pu alors être cloné. Chacun des 69 clones obtenus s'est vu isolé puis séquencé. Malgré l'isolement des ARN poly-adenylés, la grande majorité des clones séquencés (soit un total de 38 clones) a révélé des séquences appartenant au LAT_{2kb}. Ces séquences peuvent provenir du LAT de 8,3kb qui est poly-adenylé ou bien de LAT_{2kb} qui n'auraient pas été complètement éliminés lors de la purification sur la colonne d'oligo dT.

Les autres séquences, obtenues à partir des clones isolés, ont permis de séparer les jonctions 5'entre le LAT et un transcrit donné, en deux groupes.

Le premier groupe (figure 3C.), est composé uniquement de séquences comportant des insertions du LAT dans des gènes viraux (21clones au total), quand le second comportait des insertions du LAT dans des gènes cellulaires (10clones au total) (figure 3D).

L'analyse des séquences du premier groupe, a permis de confirmer des résultats déjà obtenus. Ainsi parmi les 21 clones du 1^{er} groupes, 9 d'entre eux comportaient des séquences du LAT intégrées dans le gène *RL2*. Celles-ci correspondaient à trois types de jonctions distinctes déjà obtenues précédemment (figure S5A et S5B). Cinq clones comportaient des séquences du LAT intégrées dans le gène *US12* (figure S9A et S9B), 3 clones dans le gène *UL15* (figure S8A et S8B), 2 dans le gène *UL5* (figure S7A et S7B), et 2 dans le gène *UL34* (figure S6A et S6B).

Les séquences du second groupe ont permis de mettre en évidence des séquences du LAT dans des gènes cellulaires. Ainsi 6 clones comportaient des séquences du LAT intégrées dans les transcrits codant pour la protéine BCLxL du gène *BCL-X* (Figure 11A et 11B), et 4 clones comportaient des séquences du LAT intégrées dans les transcrits codant pour la dynamine de type 1 (figure s10A et s10B).

VI.5 Recherche des jonctions 5' entre les gènes cellulaires, viraux et le LAT par PCR au moment de phase aigüe et latente de l'infection.

La découverte de séquences du LAT insérées dans les transcrits des gènes codant pour les protéines cellulaires BCLxL et dynamine 1, ainsi que pour les protéines virales UL5, UL15 et UL34, nous a amené à rechercher la présence de séquences du LAT dans les gènes correspondants au niveau de l'ADN.

Des amorces ont été alors désignées pour chacun des gènes afin d'essayer de rechercher chacune des jonctions dans l'ADN extrait de souris infectées depuis 6 (phase aigüe de l'infection) ou 28 jours (phase latente de l'infection). Malgré plusieurs tentatives, il n'a pu être mis en évidence de séquences correspondant au LAT intégré dans les gènes identifiés par extension d'amorce (à l'exception des insertions déjà caractérisées de *RL2* et de *US12* à 28 jours).

VI.6 Conclusion : l'hypothèse du « *reverse splicing* »

Au cours de ce chapitre nous avons pu mettre en évidence la capacité du LAT à s'insérer dans les transcrits viraux et cellulaires aux phases précoces de l'infection (6jours p.i.).

Cependant, durant cette phase de l'infection, il n'a pas pu être mis en évidence des séquences du LAT insérées dans les gènes correspondant aux transcrits modifiés identifiés et cela même pour les gènes *RL2* et *US12*.

Il faut néanmoins être prudent dans l'interprétation de ce dernier résultat. En effet, il est possible que la transposition du LAT dans les gènes viraux comme *RL2* ou *US12* et *a fortiori* dans les gènes cellulaires, n'ait pas encore atteint un seuil raisonnable pour être détectable, contrairement aux transcrits modifiés, présents en plus forte quantité.

L'analyse des jonctions du côté 5' des transcrits *RL2* et *US12* modifiés pendant la phase aigüe de l'infection, révèle que certaines d'entre elles ne peuvent pas être obtenues à partir de la simple transcription des gènes modifiés tels que nous les avons décrits au chapitre précédent pendant la phase de latence de l'infection.

Ainsi si l'on considère la jonction (RL2-nt384) / (nt242-LAT) détectée au moment de la phase aigüe de l'infection, peut être le produit de la transcription d'un gène *RL2* modifié. L'hypothèse la plus simple est que des génomes modifiés existent au moment de la phase aigüe de l'infection, mais que ceux-ci ne sont pas en quantité suffisante pour être détectés.

La deuxième jonction (RL2-nt384) / (nt221-LAT) a été observée dans les transcrits modifiés obtenus durant l'infection latente. Dans le chapitre précédent, nous avons conclu que ces transcrits étaient probablement le résultat de gènes modifiés que nous n'avions pas été en mesure de mettre en évidence. Nous pouvons conserver cette hypothèse et considérer qu'au cours de la phase aigüe, des génomes avec ce type de modification sont aussi présents en petites quantités.

En revanche le cas des jonctions (RL2-nt199) / (nt288-LAT) et (RL2-nt112) /gacaccacaca/ (nt 308-LAT) est plus complexe. Nous n'avons pas détecté ces transcrits à 28 jours p.i. alors que nous les avons détectés de façon récurrente à 6 jours p.i. Par ailleurs, que ce soit durant la phase aigüe ou la phase latente, nous n'avons jamais mis en évidence de gènes modifiés permettant la synthèse de ces transcrits.

Les mêmes constatations peuvent être faites avec l'analyse des jonctions du gène *US12*. Ainsi la nouvelle jonction 5' détectée en phase aigüe, (US12-nt1345) / (nt129-LAT),

ne peut pas provenir des génomes modifiés détectés en phase de latence car les transcrits modifiés associés contiennent une insertion de LAT plus longue que celle observée au niveau de l'ADN.

De la même façon, la jonction 3' (LAT- nt1344) / (nt1351-US12) obtenue à partir de transcrits du gène *US12* modifiés, détectés lors de la phase aigüe, ne peut pas provenir des génomes modifiés détectés en phase de latence, car la séquence du LAT retrouvée dans les ARNm du transcrits du gène *US12* modifiés à J+6 est plus courte que celle retrouvée dans l'ADN à j+28 (LAT- nt1350) / (nt1304-US12).

Au final, nous pouvons dire qu'il semble exister deux types de transcrits modifiés. Les premiers d'entre eux sont clairement la résultante de la transcription de génomes modifiés. Pour les seconds, en absence de détection de génome modifié dont ils pourraient être les transcrits, il faut envisager un autre mécanisme.

Ils pourraient être la résultante d'une intégration directe du LAT dans les ARNm. Cette idée est séduisante : cette capacité du LAT à intégrer un ARNm pourrait être assimilée à du « *reverse splicing* », il s'agirait de la première fois que l'on observe un tel phénomène *in-vivo*. Cette hypothèse est en accord avec le fait que toutes les insertions du LAT sont sens. Si l'attaque se faisait directement au niveau de l'ADN, on s'attendrait aussi à rencontrer des insertions antisens (notons quand même qu'il y a un biais avec le protocole utilisé puisque nous avons fait appel à une extension d'amorce). Cette hypothèse est aussi en accord avec le fait que dans les 7 transcrits viraux et cellulaires modifiés, l'insertion se fait toujours dans une séquence non codante, le plus souvent un intron. Ce dernier point est difficile à expliquer si l'intégration se faisait d'abord dans l'ADN. Il signifie aussi que si l'hypothèse du « *reverse splicing* » est correcte alors celui-ci a lieu dans le noyau. Le dernier argument en faveur du « *reverse splicing* » réside dans l'analyse de la jonction 5' entre la dynamine et le LAT au sien du transcrit modifié : le nucléotide 6590 correspond au dernier nucléotide de l'exon 4 du gène !

Néanmoins, si « *reverse splicing* » il y a, il n'est pas canonique. En effet, (i) l'intégration du LAT est incomplète et est très certainement irréversible contrairement à ce qui a été montré dans des travaux effectués *in vitro*.

A ce stade, il est difficile de conclure sur la réalité du « *reverse splicing* » et si celui-ci est vrai, il est encore plus difficile de proposer un mécanisme. Néanmoins, si le mécanisme de « *reverse splicing* » était validé, la question se poserait de savoir si le LAT possède par lui-même toutes les activités nécessaires à sa réintégration ou s'il doit détourner une quelconque machinerie cellulaire.

VI.7 Données de séquençages

Figure S1A

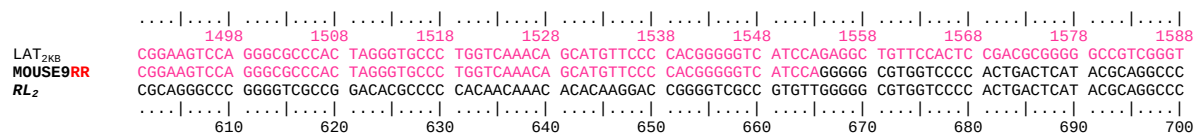


Figure S1A. Alignement de la séquence obtenue par RT-PCR représentant la jonction 3' entre le LAT_{2KB} et le transcrit du gène *RL2*. L'alignement a été effectué avec les séquences des gènes *RL2* et LAT_{2KB}.

Figure S1B

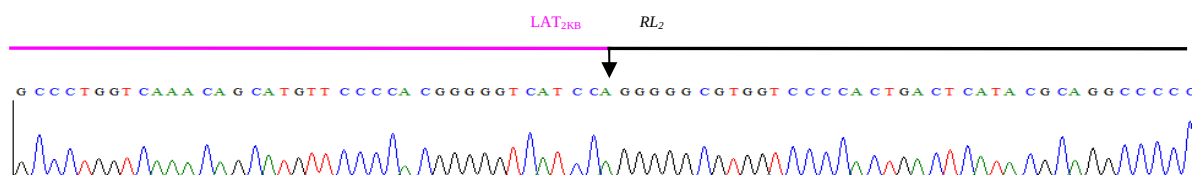


Figure S1B. Chromatogramme obtenue après le séquençage du clone « MOUSE 9RR », représentant la jonction 3' entre le LAT_{2kb} et l'ARNm du gène *RL2*. La numérotation fait référence aux positions des nucléotides à partir du 1^{er} nucléotide de l'amorce utilisée pour le séquençage du clone.

Figure S2A

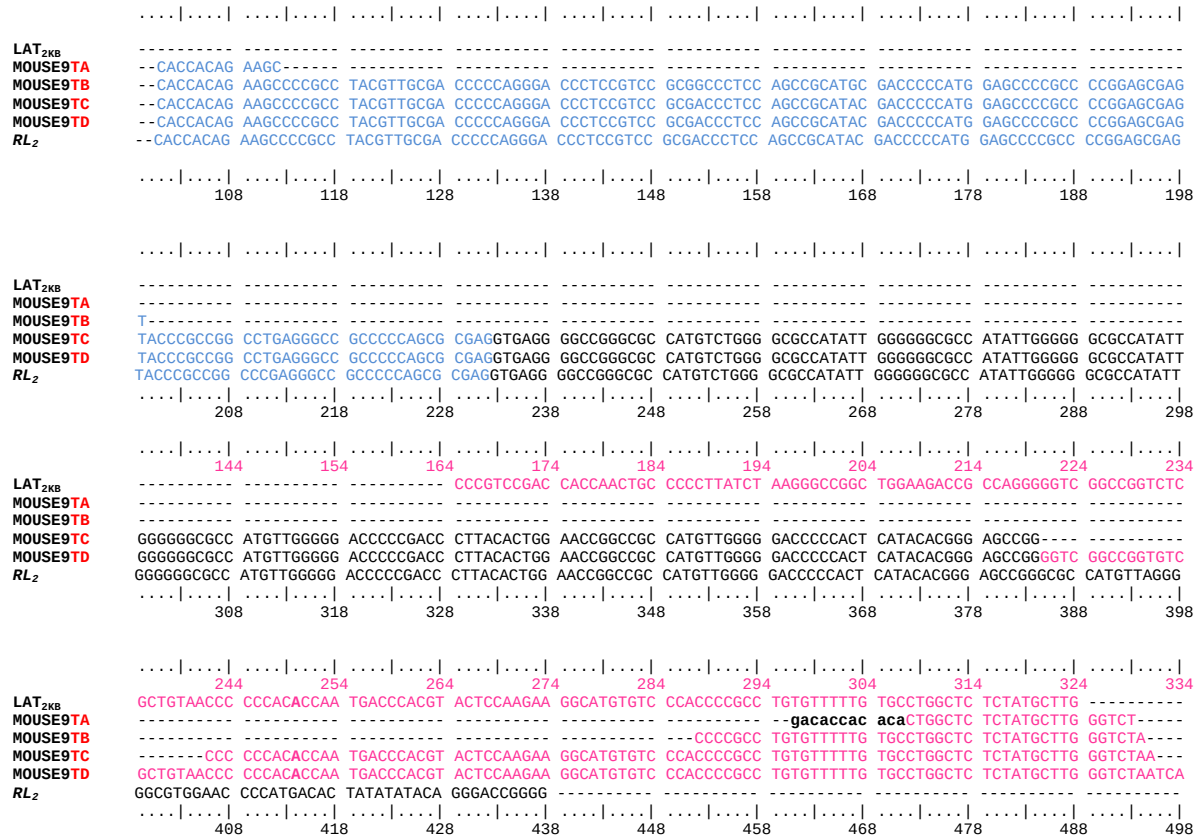
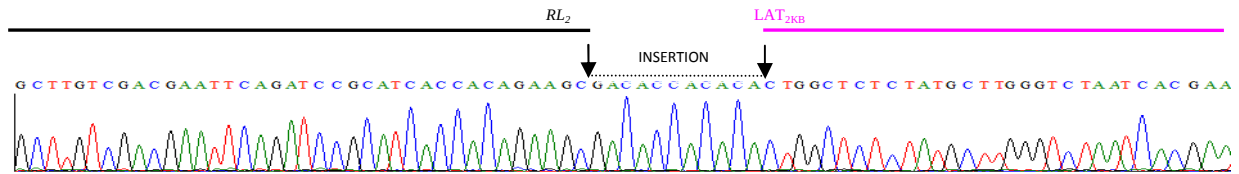


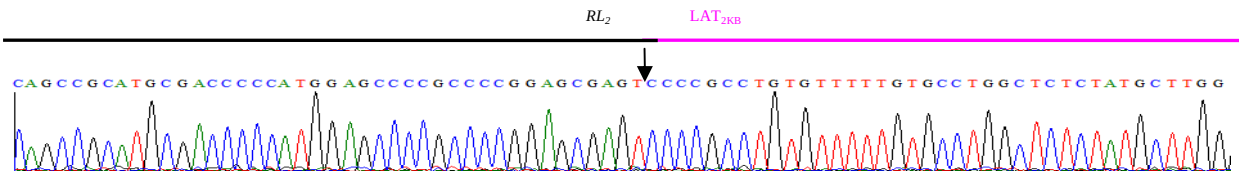
Figure S2A. Alignement des séquences obtenues par RT-PCR représentant la jonction 5' entre le transcrits du gène *RL2* et le LAT_{2KB}. L'alignement a été effectué avec les séquences des gènes *RL2* et LAT_{2KB}.

Figure S2B

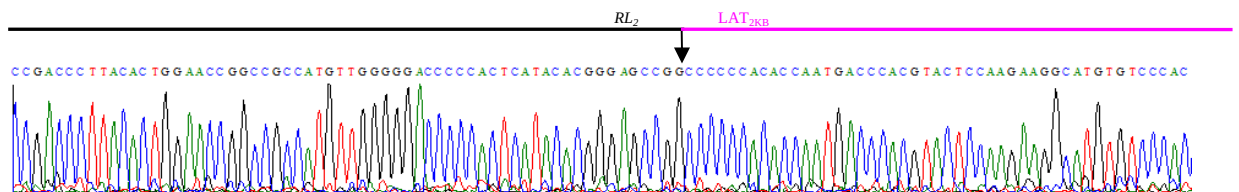
MOUSE9TA



MOUSE9TB



MOUSE9TC



MOUSE9TD

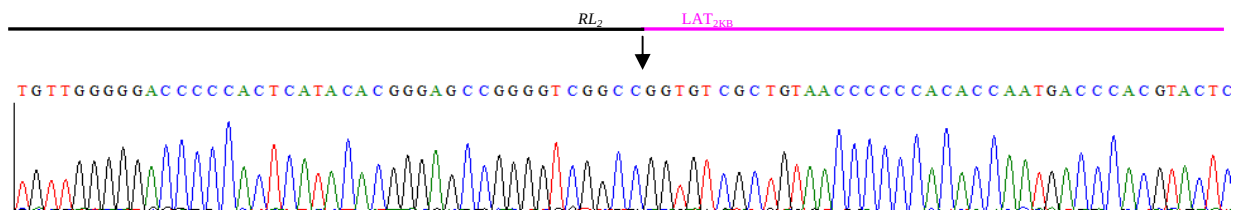


Figure S2B. Chromatogramme obtenue après le séquençage des clones « MOUSE 9TA ; MOUSE 9TB; MOUSE 9TC; MOUSE 9TD », représentant la jonction 5' entre l'ARNm du gène *RL2* et le *LAT*_{2kb}. La numérotation fait référence aux positions des nucléotides à partir du 1^{er} nucléotide de l'amorce utilisée pour le séquençage du clone.

Figure S3A

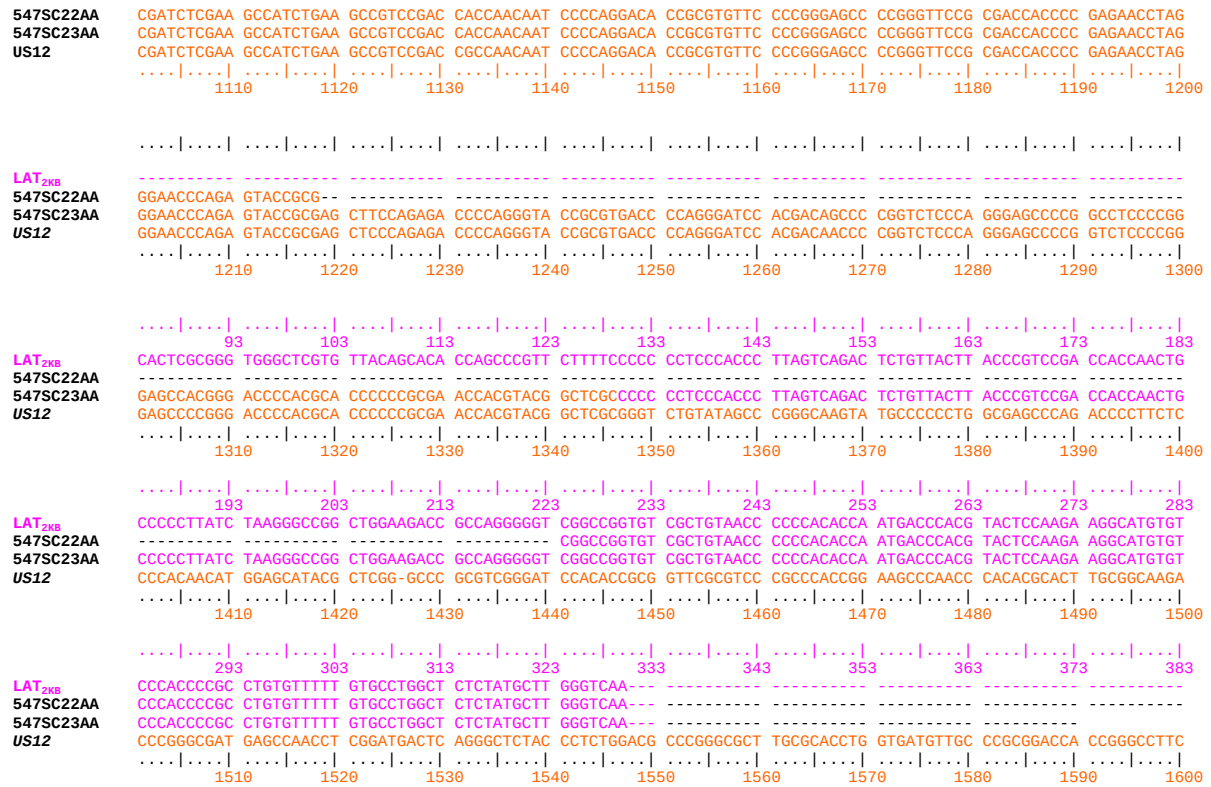
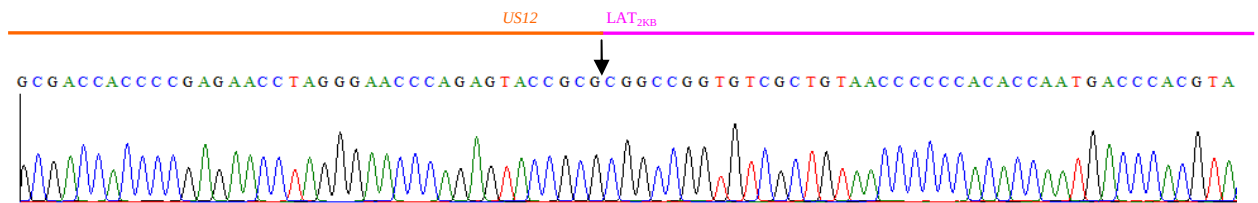


Figure S3A. Alignement des séquences obtenues par RT-PCR représentant la jonction 5' le transcrits du gène *Us12* et entre le LAT_{2KB}. L'alignement a été effectué avec les séquences des gènes *US12* et LAT_{2KB}.

Figure S3B

547SC22AA



547SC23AA

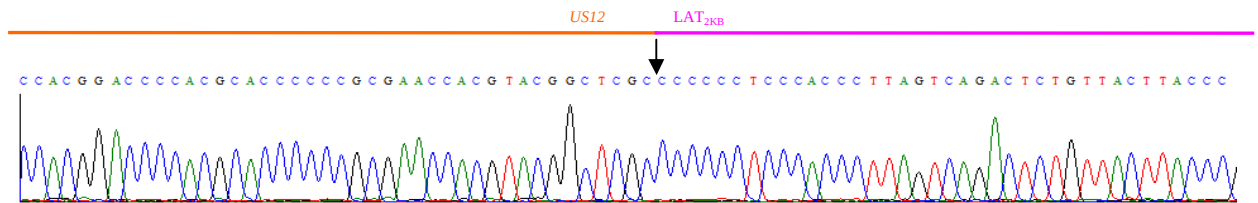


Figure S3B. Chromatogramme obtenue après le séquençage des clones « 547SC22AA ; 547SC23AA », représentant la jonction 5' entre l'ARNm du gène *US12* et le *LAT*_{2kb}. La numérotation fait référence aux positions des nucléotides à partir du 1^{er} nucléotide de l'amorce utilisée pour le séquençage du clone.

Figure S4A

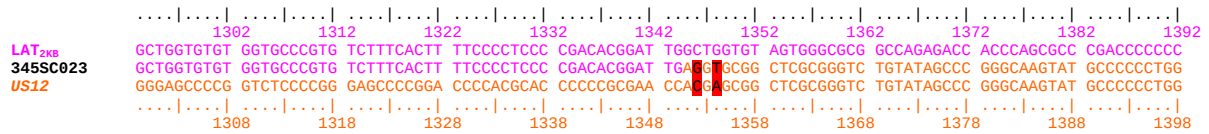


Figure S4A. Alignement de la séquence obtenue par RT-PCR représentant la jonction 3' entre le LAT_{2KB} et le transcrits du gène *Us12*. L'alignement a été effectué avec les séquences des gènes *US12* et LAT_{2KB}. Les nucléotides surlignés en rouge correspondent à des différences ponctuelles entre la séquence du clone et les séquences de références.

Figure S4B

345SC023

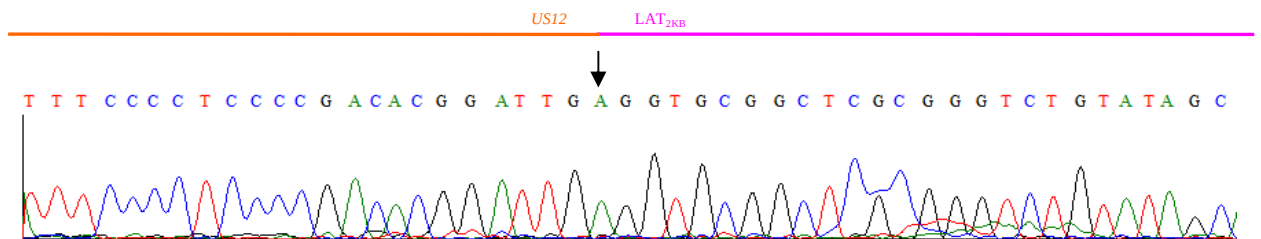


Figure S4B Chromatogramme obtenue après le séquençage du clone « 345SC023 », représentant la jonction 3' entre l'ARNm du gène *US12* et le LAT_{2kb}. La numérotation fait référence aux positions des nucléotides à partir du 1^{er} nucléotide de l'amorce utilisée pour le séquençage du clone.

Figure S5A

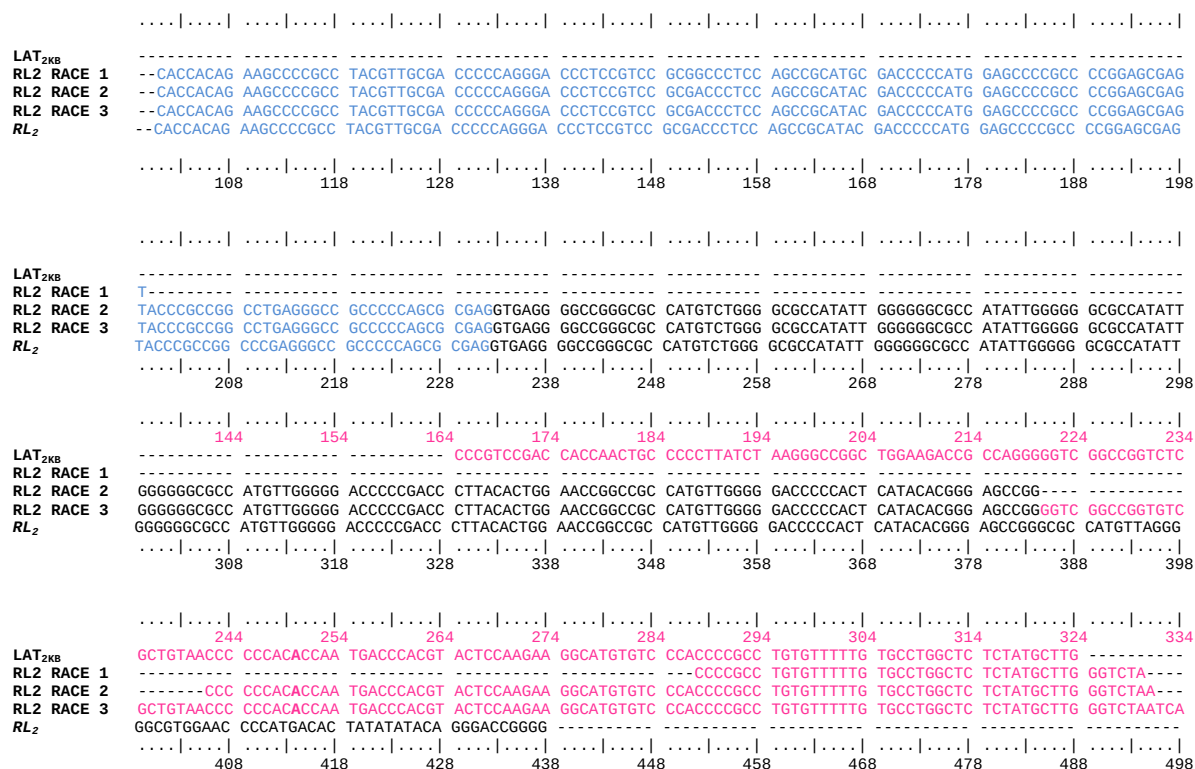
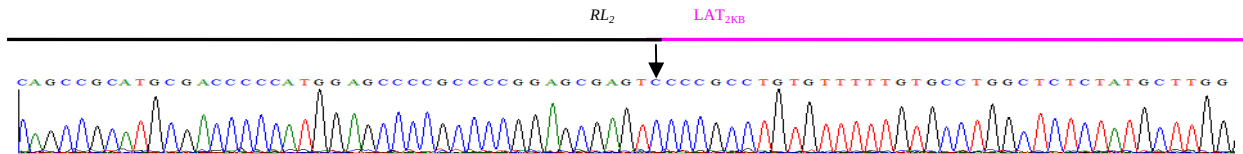


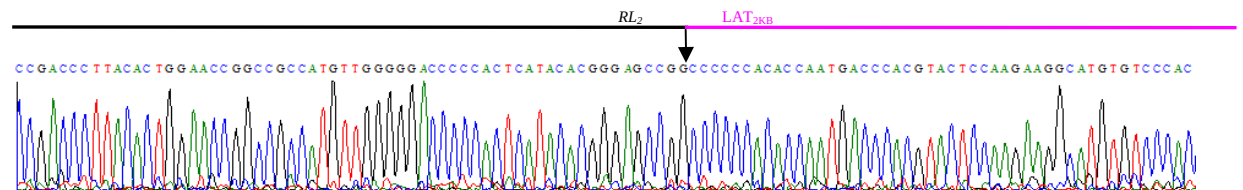
Figure S2A. Alignement des séquences obtenues par 5'RACE représentant la jonction 5' entre le transcrits du gène *RL2* et le *LAT*_{2KB}. L'alignement a été effectué avec les séquences des gènes *RL2* et *LAT*_{2KB}.

Figure S5B

RL2 RACE 1



RL2 RACE 2



RL2 RACE 3

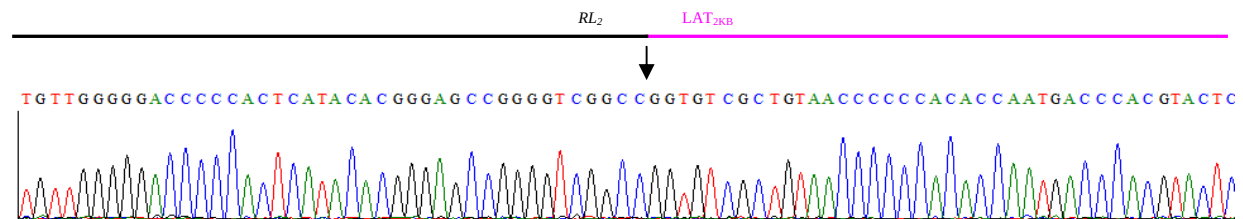


Figure S2B. Chromatogramme obtenue après le séquençage des clones « RL2 RACE 1; RL2 RACE 2 ; RL2 RACE 3 », représentant la jonction 5' entre l'ARNm du gène *RL2* et le LAT_{2kb}. La numérotation fait référence aux positions des nucléotides à partir du 1^{er} nucléotide de l'amorce utilisée pour le séquençage du clone.

Figure S6A

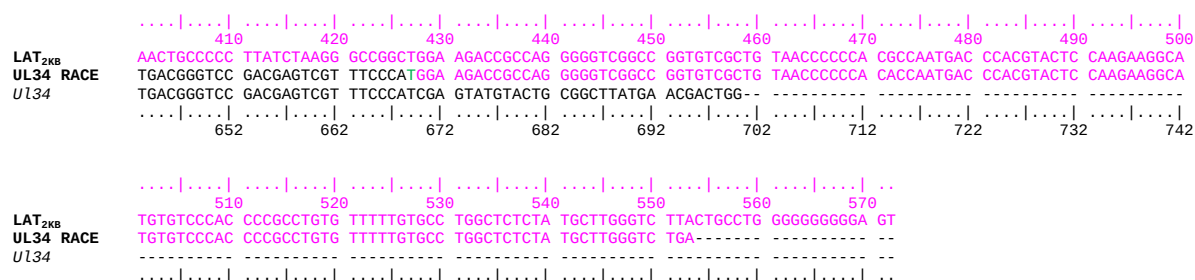


Figure S6A : Alignement d'une séquence obtenue par 5'Race représentant la jonction 5' entre le LAT_{2KB} et le transcrits du gène *UL34*. L'alignement a été effectué avec les séquences des gènes *UL34* et LAT_{2KB}.

Figure S6B

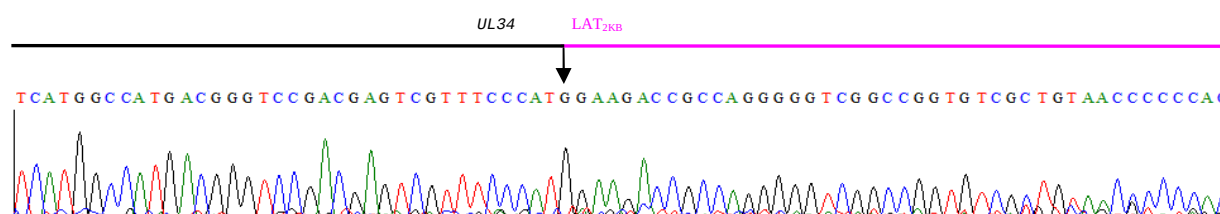


Figure S6B : Chromatogramme obtenue après le séquençage du clone « UL34 RACE », représentant la jonction 5' entre l'ARNm du gène *UL34* et le LAT_{2kb}. La numérotation fait référence aux positions des nucléotides à partir du 1^{er} nucléotide de l'amorce utilisée pour le séquençage du clone.

Figure S7A

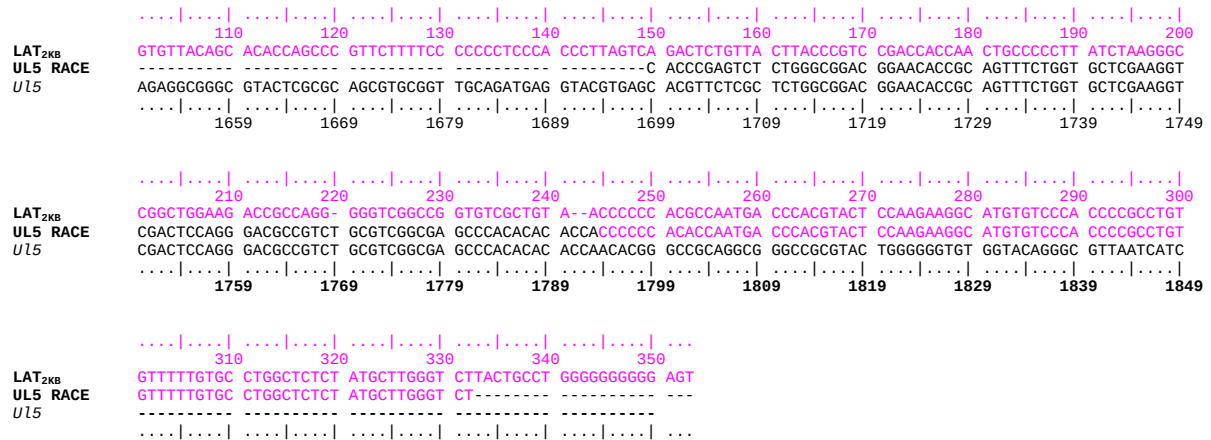


Figure S7A : Alignement d'une séquence obtenue par 5'Race représentant la jonction 5' entre le LAT_{2KB} et le transcrits du gène *UL5*. L'alignement a été effectué avec les séquences des gènes *UL5* et LAT_{2KB}.

Figure S7B

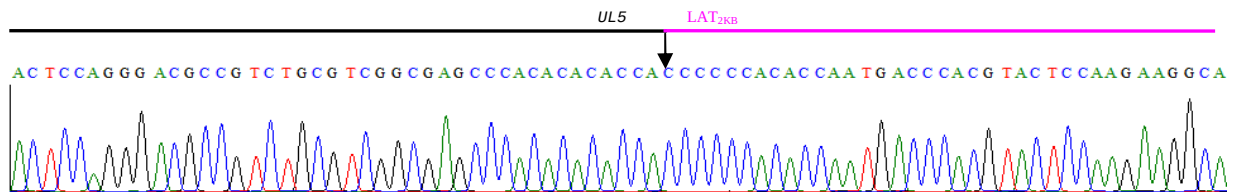


Figure S7B : Chromatogramme obtenue après le séquençage du clone « UL5 RACE », représentant la jonction 5' entre l'ARNm du gène *UL5* et le LAT_{2kb}. La numérotation fait référence aux positions des nucléotides à partir du 1^{er} nucléotide de l'amorce utilisée pour le séquençage du clone.

Figure S8A

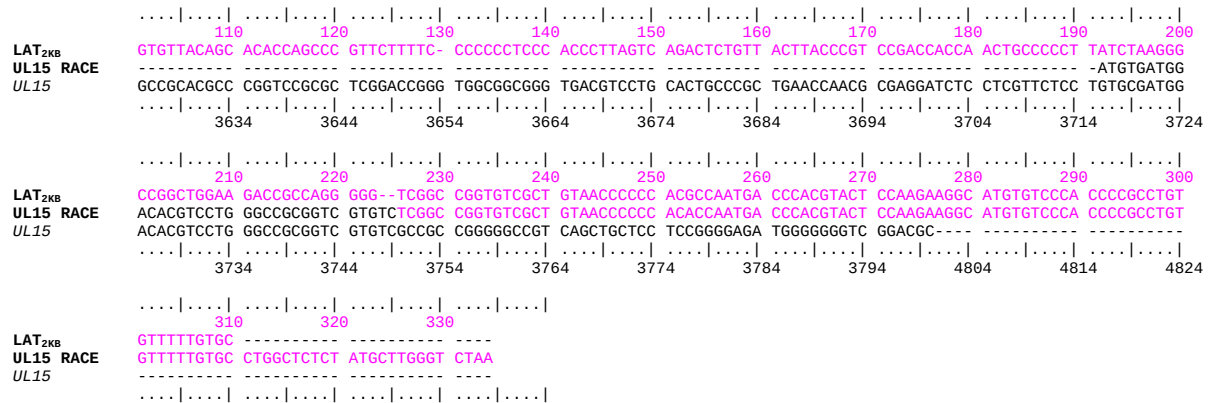


Figure S8A : Alignement d'une séquence obtenue par 5'Race représentant la jonction 5' entre le LAT_{2KB} et le transcrits du gène *UL15*. L'alignement a été effectué avec les séquences des gènes *UL15* et LAT_{2KB}.

Figure S8B

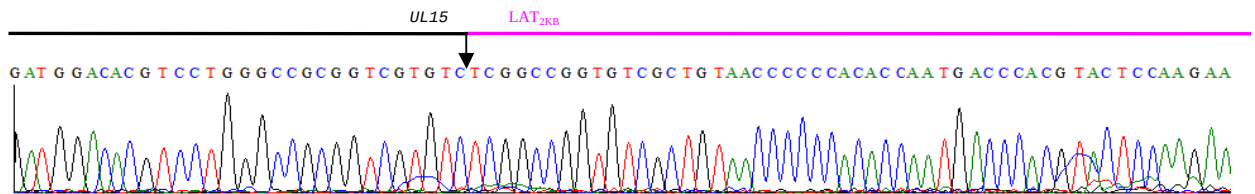


Figure S8B : Chromatogramme obtenue après le séquençage du clone « UL15 RACE », représentant la jonction 5' entre l'ARNm du gène *UL15* et le LAT_{2kb}. La numérotation fait référence aux positions des nucléotides à partir du 1^{er} nucléotide de l'amorce utilisée pour le séquençage du clone.

Figure S9A

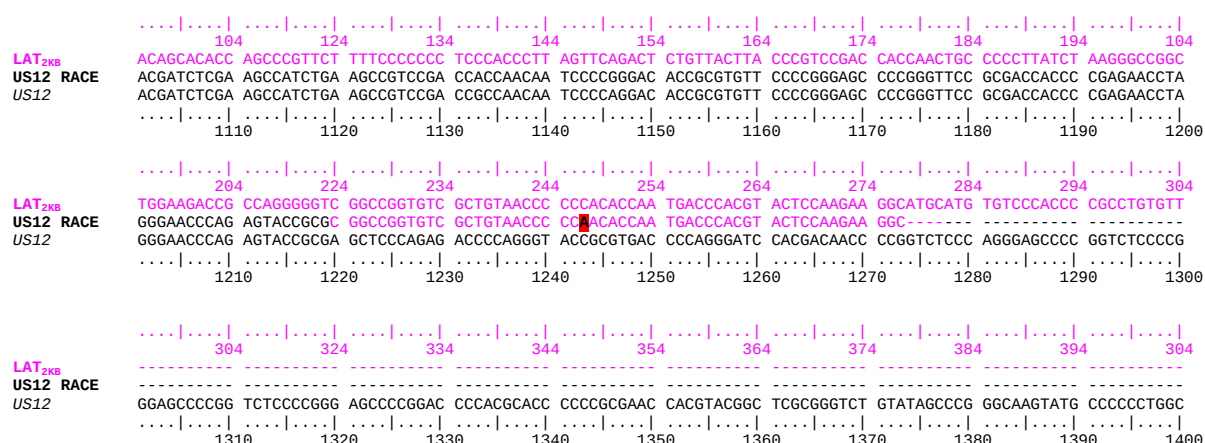


Figure S9A. Alignement d'une séquence obtenue par 5'RACE représentant la jonction 5' entre le LAT_{2KB} et le transcrits du gène *US12*. L'alignement a été effectué avec les séquences des gènes *US12* et LAT_{2KB}. L'alignement a été effectué avec les séquences des gènes *US12* et LAT_{2KB}. Les nucléotides surlignés en rouge correspondent à des différences ponctuelles entre la séquence du clone et les séquences de références.

Figure S9B

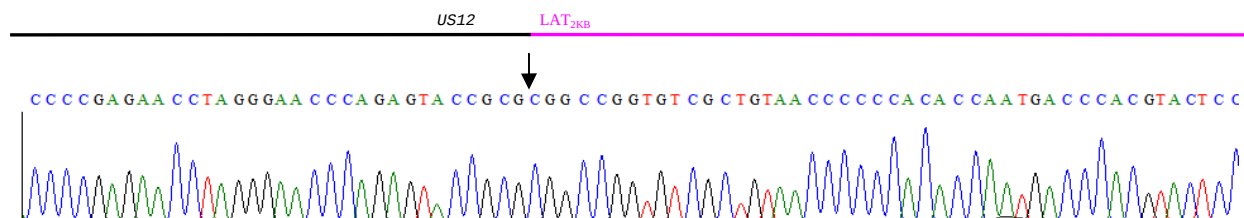
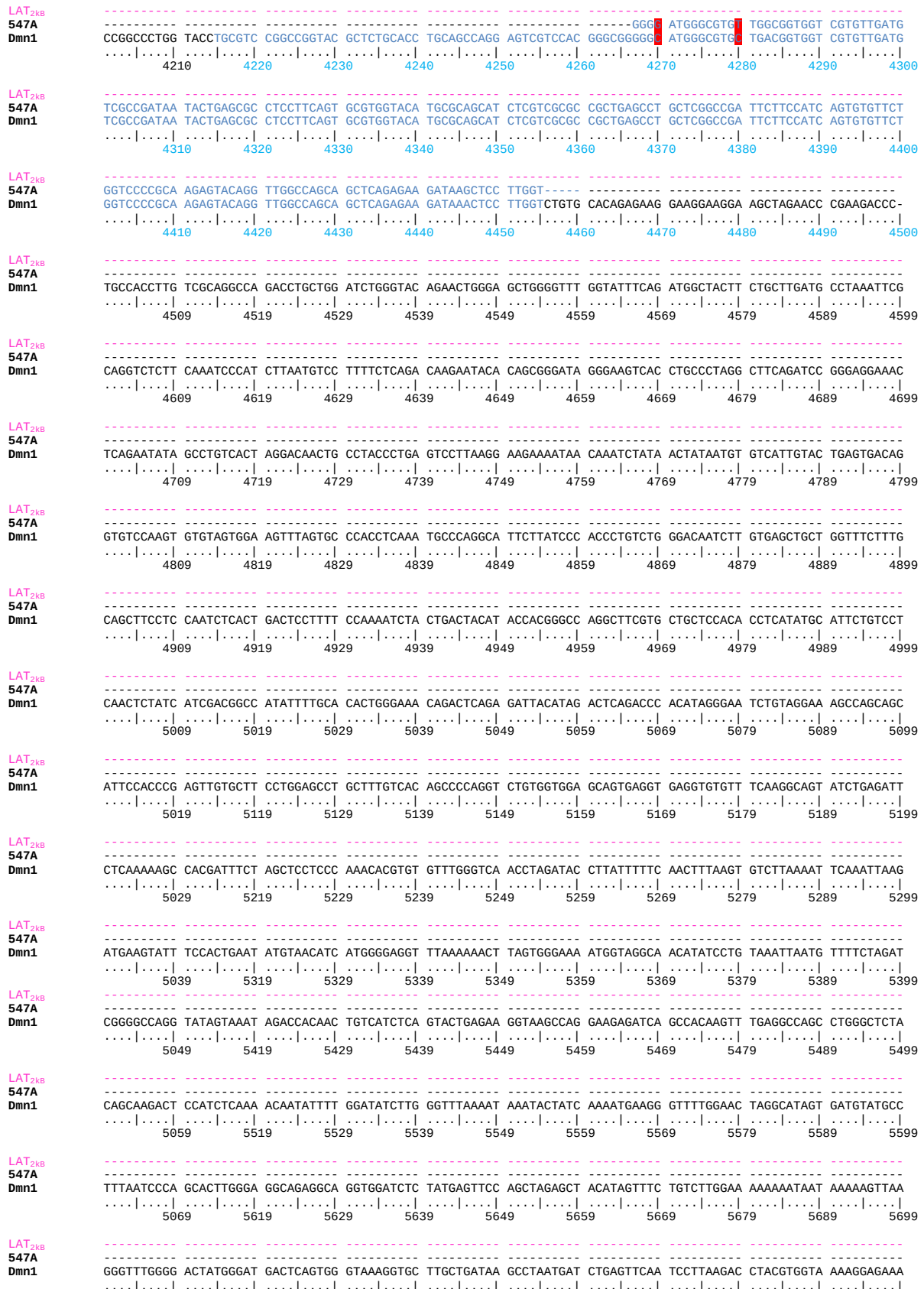


Figure S9B Chromatogramme obtenue après le séquençage du clone « US12 RACE », représentant la jonction 5' entre l'ARNm du gène *US12* et le LAT_{2kb}. La numérotation fait référence aux positions des nucléotides à partir du 1^{er} nucléotide de l'amorce utilisée pour le séquençage du clone.

Figure S10A



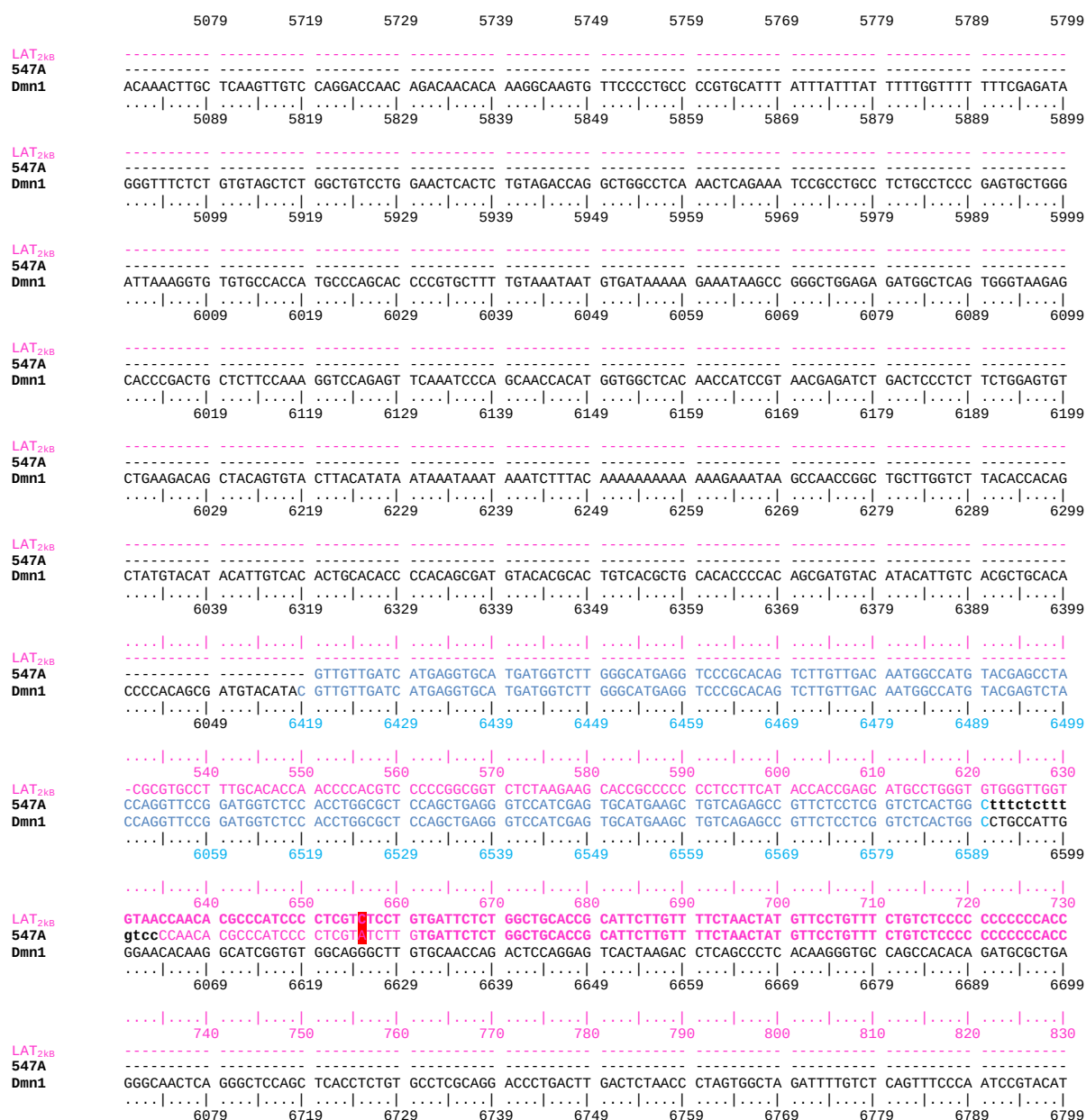


Figure S10A : Alignement d’une séquence obtenue par 5’Race représentant la jonction 5’ entre le LAT_{2KB} et les transcrits du gène codant pour la dynamine. L’alignement a été effectué avec les séquences des gènes de la dynamine et du LAT_{2KB}. L’alignement a été effectué avec les séquences des gènes *US12* et LAT_{2KB}. Les nucléotides surlignés en rouge correspondent à des différences ponctuelles entre la séquence du clone et les séquences de références.

Figure S10B

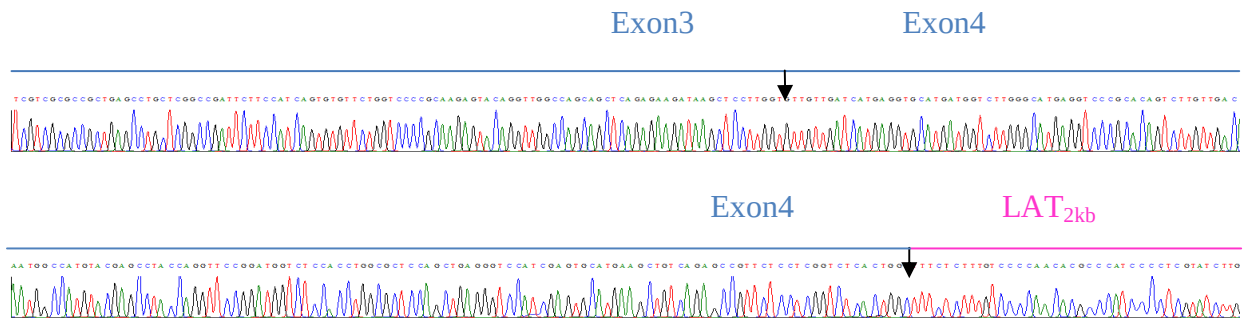


Figure S9B : Chromatogramme obtenue après le séquençage du clone « 547A », représentant la jonction 5' entre l'ARNm du gène codant pour la dynamine et le LAT_{2kb}. La numérotation fait référence aux positions des nucléotides à partir du 1^{er} nucléotide de l'amorce utilisée pour le séquençage du clone.

Figure S11A

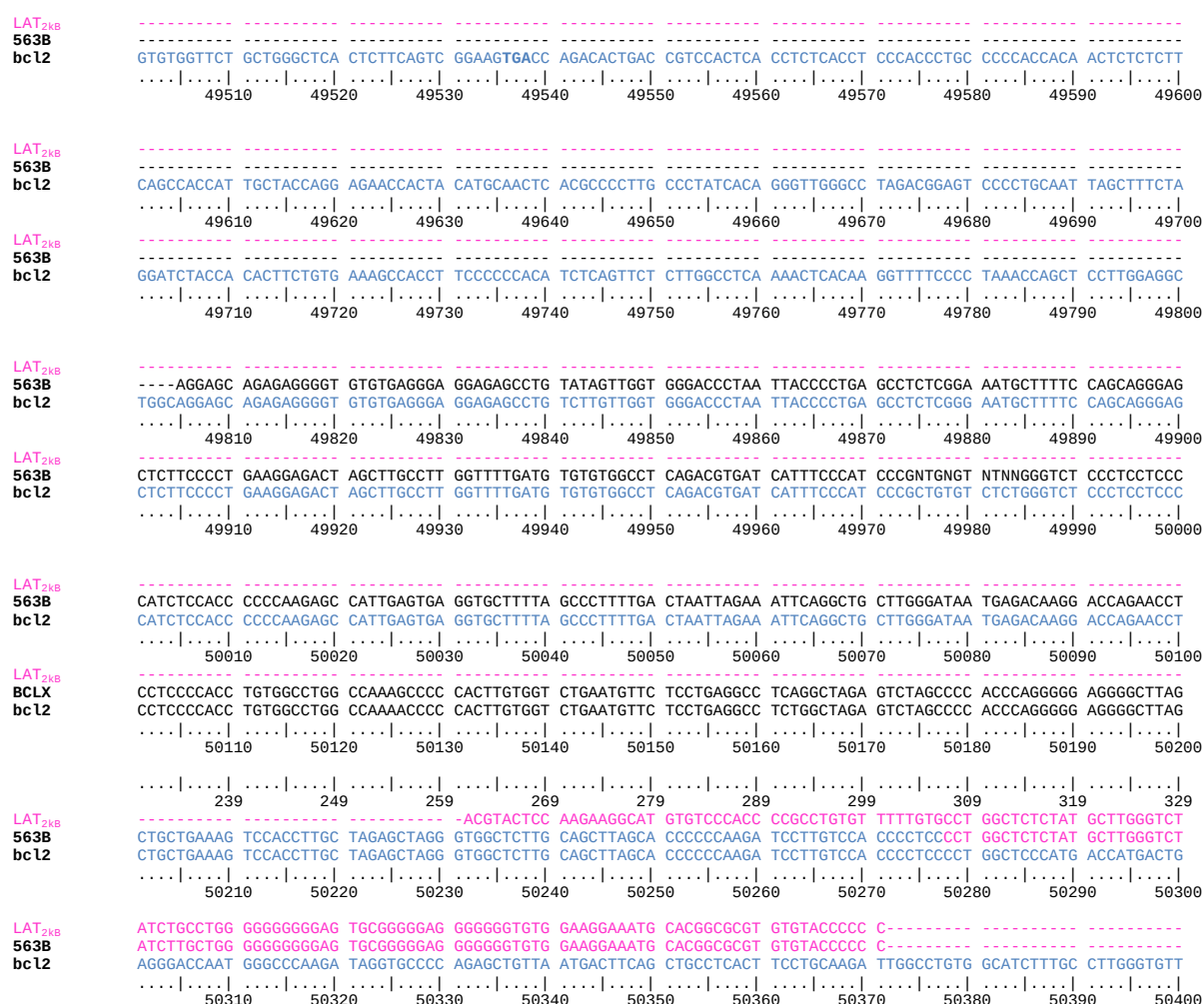


Figure S11A : Alignement d'une séquence obtenue par 5'Race représentant la jonction 5' entre le LAT_{2kb} et le transcrits du gène codant pour les protéines de la famille BcL-2. L'alignement a été effectué avec les séquences des gènes codant pour les protéines de la famille BcL-2 et du LAT_{2kb}.

Figure S11B

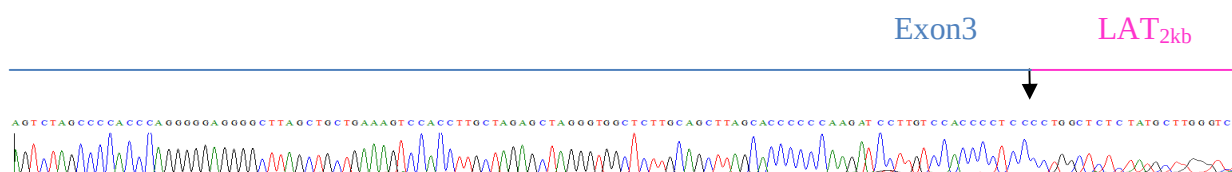


Figure S11B : Chromatogramme obtenue après le séquençage du clone « 563B », représentant la jonction 5' entre l'ARNm du gène codant pour les protéines de la famille BcL-2et le LAT_{2kb}. La numérotation fait référence aux positions des nucléotides à partir du 1^{er} nucléotide de l'amorce utilisée pour le séquençage du clone.

Conclusions et perspectives

Au cours de ce travail, nous avons pu mettre en évidence un phénomène de transposition du LAT dans des gènes viraux et au sein de loci génomiques de la cellule hôte.

Cette découverte apporte des réponses à plusieurs questions concernant la biologie du virus et plus particulièrement concernant les mécanismes impliqués dans la mise en place et le maintien de la latence ainsi que dans les épisodes de réactivation.

La transposition du LAT_{2kb} dans le gène *RL2* permet le maintien de la latence et suggère des pistes quant aux mécanismes de réactivation.

La transposition du LAT_{2kb} dans le gène *RL2* n'est pas anodine puisque celui-ci code pour la protéine ICP0. Comme nous l'avons vu dans l'introduction de ce manuscrit, ICP0 est considérée comme la protéine « *starter* » de la réactivation virale, ainsi en bloquant son expression le LAT pourrait contrôler la cascade transcriptionnelle virale et donc maintenir des génomes à l'état latent.

D'un point de vue mécanistique, la présence du LAT dans une séquence non codante d'un gène, peut engendrer plusieurs niveaux de régulation, comme un arrêt ou une stimulation de la transcription, un épissage alternatif des transcrits produits, une inhibition de l'export des ARNm hors du noyau ou même la synthèse d'une protéine de séquence distincte (conséquence d'un épissage alternatif).

La présence du LAT dans une séquence non codante du gène *RL2* pourrait aboutir au moment de la transcription à une atténuation voire un arrêt prématuré de celle-ci. Dans le cas de ce gène, l'intégration du LAT_{2kb} dans le 1^{er} intron altérerait alors l'expression de *RL2* conduisant à terme à un « *silencing* » du gène et donc au blocage du cycle viral dans un état latent. Cependant, nous avons pu montrer que plus de 90% des transcrits du gène *RL2* étaient non épissés et contenaient l'insertion du LAT. Par ailleurs, dans des expériences non décrites dans ce manuscrit, nous avons montré que les transcrits non épissés du gène *RL2* détectés au moment de la latence virale contenaient toujours l'intron 2. En conséquence, l'intégration du LAT_{2kb} dans le 1^{er} intron n'entraîne pas d'arrêt prématuré de la transcription. Cette hypothèse ne permet donc pas d'expliquer comment l'insertion du LAT dans le premier intron du gène *RL2* régule la production de la protéine ICP0 même si ce mécanisme de régulation reste potentiellement applicable à d'autres gènes viraux ou cellulaires.

En fait, nos résultats suggèrent plutôt que l'insertion du LAT_{2kb} dans le premier intron du gène RL2 inhibe l'épissage du transcrit tout en le protégeant de la dégradation. Il semble que cette action protectrice du LAT sur le transcrit dans lequel il s'intègre puisse être étendu au transcrit du gène US12 même si l'insertion a eu lieu dans un exon (mais toujours dans une région non codante).

De façon générale, l'épissage correct est une condition d'export de l'ARNm. La rétention dans le noyau d'ARN modifié dans des séquences non codantes permettrait alors au virus de bénéficier d'un pool de transcrits déjà présents qui n'auraient plus qu'à subir un épissage pour être exportés et ainsi permettre l'expression de la protéine. Ce mécanisme serait un avantage pour le virus car il lui permettrait de pouvoir exprimer très rapidement des protéines virales (notamment ICP0) en réponse à un stress. Des études restent à mener pour comprendre comment la présence du LAT dans les transcrits permet leur rétention dans le noyau et pourquoi ils ne subissent pas de contrôle qualité qui devrait conduire à leur dégradation. Il est de même intéressant de s'interroger sur les mécanismes moléculaires qui permettraient une maturation rapide des ARN modifiés lors d'un stress. Des changements de structure tertiaire ou secondaire du LAT intégré pourraient être envisagés.

La transposition du LAT_{2kb} dans le gène US12 pourrait permettre de réguler le contrôle de la réponse T contre les neurones infectés en début de réactivation.

Le second gène viral dans lequel le LAT_{2kb} est capable de transposer est le gène US12. Cette transposition conduit aussi à des transcrits modifiés, détectés pendant la phase de latence. Il faut d'ailleurs souligner que ce gène n'était pas connu pour être transcrit à ce stade de l'infection.

Le gène US12, qui est aussi un gène IE (*immediate early*), code la protéine ICP47, une protéine virale qui se lie au transporteur associé à la maturation des antigènes (en anglais TAP : *transporter associated with antigen processing*). Ce faisant, il empêche la translocation des peptides dans le réticulum endoplasmique et, par conséquent, empêche la présentation des antigènes aux lymphocytes T CD8 +. Cela permet d'éviter la détection des cellules infectées par le système immunitaire de l'hôte

Nous l'avons déjà mentionné : comme dans le cas du gène RL2, l'insertion du LAT dans US12 se produit dans une région non codante du transcrit (une partie de l'exon 2 en aval du codon d'arrêt). Comme pour les transcrits du gène RL2, on peut penser que les conditions

de stress peuvent induire la maturation des transcrits modifiés accumulés qui seront ensuite exportés hors du noyau et traduits. Dans ce cas, le phénomène qui conduirait à la synthèse d'ICP0 et qui donc déclencherait la cascade de réactivation, conduirait aussi à la synthèse rapide d'ICP47 qui protégerait la cellule du système immunitaire.

La transposition du LAT dans des loci génomiques de la cellule hôte permet d'expliquer les nombreux phénotypes associés aux mutants du LAT

Nous l'avons vu dans l'introduction, les mutations dans le LAT conduisent à des phénotypes très variés. La démonstration que le LAT est capable de transposer dans le génome de l'hôte permet d'apporter une explication simple à ces observations. Le ciblage de certains gènes spécifiques pourrait permettre au LAT de prendre le contrôle de voies de signalisation, de la réponse immune innée, de l'apoptose, etc.

Pour l'instant, les événements de transposition que nous avons identifiés révèlent des intégrations du LAT dans des loci génomiques qui ne semblent pas être associés à des gènes transcrits en ARNm. Quatre remarques s'imposent :

Premièrement, la stratégie utilisée pour identifier les séquences d'insertion est biaisée. L'insertion doit avoir eu lieu au voisinage d'un élément B1 pour être détectée. Il est probable que nous n'avons détecté qu'un petit nombre des insertions dans le génome, pas forcément les plus intéressantes du point de vue de la biologie du virus.

Deuxièmement, même si l'insertion a lieu dans des régions non transcrites en ARNm, on ne peut exclure que celle-ci ait un impact au niveau épigénétique. On rappellera ici que l'analyse de la séquence du LAT intégré révèle la présence de sites de fixation pour la protéine CTCF qui permet le blocage de la diffusion de l'hétérochromatine. Cela pourrait conduire à une régulation transcriptionnelle des gènes au voisinage de l'insertion. Un tel mécanisme serait avantageux pour le virus : un petit nombre de LAT intégrés dans le génome de l'hôte (petit nombre qui limiterait le risque d'insertion labile pour la cellule) pourrait être suffisant pour contrôler un grand nombre de gènes.

Troisièmement, lors de la phase aigüe de l'infection, nous avons quand même détectés la présence du LAT dans des transcrits associés à des gènes cellulaires. Nous n'avons pas pu mettre en évidence d'insertion dans les gènes correspondants (mais nous ne pouvons formellement exclure que nous n'avons pas été en mesure de détecter les insertions). Les deux transcrits identifiés ne sont pas anodins. La présence du LAT dans le transcrit du gène BCL-X

pourrait permettre un contrôle de l'apoptose. La présence du LAT dans le transcrit de la dynamine 1 engendrerait une régulation de l'endocytose. Dans ce dernier cas, cela pourrait permettre d'éviter une surinfection délétère par un autre virus (.

Enfin, l'ensemble de ce travail a été réalisé dans un modèle souris qui n'est pas l'hôte naturel du virus. Il est évident que les gènes de l'hôte modifiés qui sont pertinents quant à la biologie du virus seront trouvés chez l'homme.

Le mécanisme de transposition du LAT exploite probablement des machineries de la cellule hôte.

Une des premières questions qui fait suite à ce travail reste celle de comprendre le mécanisme de transposition du LAT. Nos données ne permettent absolument pas de conclure sur ce point. Les différentes étapes du processus qui conduisent à l'intégration du LAT_{2kb} dans l'ADN restent inconnues. Plusieurs modèles peuvent être proposés selon que la cible primaire est un ARN ou directement l'ADN.

Dans le **premier modèle**, le LAT_{2kb} attaque directement l'ADN. Dans ce cas, tous les ARN modifiés que nous avons détectés seraient le produit de la transcription de gènes modifiés par insertion du LAT_{2kb}. Cette hypothèse est la plus simple mais, comme nous l'avons mentionné à la fin du deuxième chapitre de résultats, un grand nombre de transcrits modifiés ont été identifiés sans que l'on puisse mettre en évidence le gène modifié dont ils pourraient être le produit.

C'est justement cette observation qui nous a conduits à proposer le **deuxième modèle** dans lequel le LAT_{2kb} a d'abord pour cible l'ARN. Le LAT_{2kb} serait donc capable de réintégrer un préARNm par une étape de « *reverse splicing* ». Nous sommes conscients que l'utilisation de ce terme est un peu abusive puisque nos données indiquent que la réintégration du LAT_{2kb} dans un ARN, si elle a vraiment lieu, n'est que partielle et irréversible (à la différence du « *reverse splicing* » observé *in vitro*). Outre le fait que, pour la plupart des transcrits détectés, nous avons été incapables de mettre en évidence le gène modifié dont ils pourraient être le produit, deux arguments viennent soutenir ce deuxième modèle. Tout d'abord, toutes les insertions mises en évidence dans des transcrits sont situées dans des séquences non codantes (généralement des introns). Ce point est d'autant plus remarquable que seuls 5 gènes de HSV-1 (sur 84 identifiés) sont épissés et que les transcrits de 3 d'entre eux (*RL2*, *US12*, *UL15*) sont modifiés (sur 5 gènes viraux dont les transcrits sont modifiés).

Le deuxième argument (très) fort en faveur du « *reverse splicing* » réside dans la nature de la jonction 5' entre la dynamine et le LAT : dans ce cas précis, le LAT s'est inséré juste après le dernier nucléotide de l'exon 4 du transcrit. Si la première étape est bien du « *reverse splicing* » alors on peut penser que seuls quelques transcrits modifiés vont par la suite permettre la réintégration du LAT au sein de l'ADN viral ou cellulaire.

Enfin le **troisième modèle** serait un hybride des deux précédents : le LAT pourrait à la fois attaquer des ARN et de l'ADN.

Si le « *reverse splicing* » est bien réel, il est possible que le LAT_{2kb} possède toutes les activités nécessaires à l'intégration d'un ARN. Notons que, dans ce cas, les séquences des transcrits modifiés suggèrent que cette intégration demande au minimum 4 étapes de clivage/ligation. D'une part, cela suppose une chorégraphie complexe lors de l'intégration et une structuration particulière du LAT_{2kb}.

Quelque soit le modèle retenu, l'événement de transposition implique plusieurs étapes successives : à savoir la reconnaissance de l'ADN cible, la coupure d'un des deux brins de l'ADN cible, la transcription inverse du LAT (ou du transcrit modifié), et enfin la ligation au site d'intégration. Chacune de ses étapes nécessite probablement l'action d'un ou plusieurs partenaires.

Considérons la première étape, c'est-à-dire la reconnaissance du site d'intégration. Au cours de ce travail, nous avons montré que le LAT transpose toujours dans les mêmes cibles et cela dans toutes les souris étudiées. De cela, nous pouvons conclure que les cibles de transpositions ne sont pas la résultante d'une transposition aléatoire. De plus, l'unicité des jonctions observée au sein des différentes souris pour un gène donné implique que la transposition du LAT est un phénomène hautement régulé nécessitant la reconnaissance précise du site d'intégration. Cependant, l'analyse de l'ensemble des séquences cibles d'intégration n'a pas permis de mettre en évidence un motif commun qui serait impliqué dans la reconnaissance du site d'intégration.

Cette absence de séquence spécifique semble exclure l'action d'une endonucléase qui cliverait l'ADN. Il est donc probable que cette activité catalytique soit portée par le LAT_{2kb} libre ou par le LAT intégré dans un transcrit. Ce type d'activité existe chez les introns autocatalytiques de groupe II avec qui le LAT_{2kb} partage la caractéristique d'être un lariat.

L'étape la plus critique dans le processus de transposition du LAT_{2kb} est la transcription inverse de celui-ci au site de coupure. Elle implique la présence d'une transcriptase inverse qui à ce jour reste à identifier. L'absence de cadre de lecture codant pour

une telle enzyme dans le génome de HSV-1 laisse supposer que celle-ci serait d'origine cellulaire. Dans ce cas, à l'instar des éléments ALU qui détournent la transcriptase inverse des éléments L1, le LAT détournerait la machinerie de transposition d'autres éléments mobiles.

Il se pose alors un problème : pour réussir son intégration, le LAT doit impérativement trouver une transcriptase inverse au moment où il transpose. Le détournement de la transcriptase inverse d'un autre élément mobile implique que cet élément mobile soit actif (*i.e.* en cours de transposition) au moment de l'infection virale. Pour réussir cette étape il faut donc qu'il y ait une co-régulation de la transposition du LAT et de l'élément mobile. Il faut donc imaginer un mécanisme dans lequel le stress induit par l'infection du virus HSV-1 entraînerait des événements de transposition.

De nombreux exemples ont démontré la mobilisation d'éléments transposables sous différentes conditions de stress, qu'ils soient biotiques ou abiotiques (Wessler, 1996; Capy et al., 2000; Ikeda et al., 2001; Grandbastien et al., 2005; Todeschini et al., 2005; Ebina et Levin, 2007). Ainsi, certains stress peuvent directement mobiliser les éléments transposables via leurs régions régulatrices. Par exemple, d'après Slotkin et Martienssen (2007), plusieurs éléments transposables chez *D. melanogaster* pourraient présenter des motifs semblables à ceux des promoteurs activables par des chocs thermiques (promoteurs Heat-Shock) au sein de leurs régions régulatrices. De même, chez l'amibe *Dictyostelium discoideum*, il a été démontré que les rétrotransposons DIRS1 présentent un promoteur Heat-Shock actif (Zuker et al., 1983 ; Chung et al 1983). Chez le tabac, l'élément Tnt1 peut être mobilisé en cas de blessures ou d'infections par des pathogènes, car il présente des séquences régulatrices similaires à celles impliquées dans l'activation de gènes de défense de la plante (Grandbastien et al., 1997; Melayah et al., 2001). Dans le cas qui nous intéresse, des éléments mobilisés par le stress induit par l'infection pourraient fournir la transcriptase inverse nécessaire à la transposition du LAT_{2kb}.

Le virus de l'herpes étant capable de faire de la latence au sein d'au moins deux espèces différentes (homme et souris), cela implique que la transcriptase inverse est portée par un élément mobile commun aux deux espèces dont la régulation a été conservée au cours de l'évolution. Dans ce contexte, les éléments mobiles qui retiennent particulièrement l'attention sont les éléments de type L1.

L'analyse des promoteurs du LAT et des éléments L1 révèle la présence de site de fixation pour une protéine commune le facteur transcriptionnel Yin Yang 1 (YY1) (athanikar et al 2004). YY1 est un facteur transcriptionnel ubiquitaire participant à la régulation de

l'expression de nombreux gènes et exerçant des effets divers sur la transcription. Il peut être activateur, répresseur ou participer à l'étape d'initiation de la transcription. Ces effets variés s'expliquent par sa capacité à interagir avec de nombreuses protéines clés de la transcription : des facteurs d'initiation comme TBP ou TFIIB, des complexes histones acétyltransférases pouvant jouer le rôle de coactivateurs transcriptionnels, ou des complexes histones déacétylases pouvant jouer le rôle de corépresseurs transcriptionnels. Enfin, il est très impliqué dans la biologie du neurone en agissant à différents moments de celle-ci (He et al 2008.) il est donc possible que le couplage de la transcription des LAT et des éléments L1 soit sous la dépendance d'un même facteur transcriptionnel. Le LAT aurait ainsi la capacité de recruter la transcriptase inverse codée par les éléments L1. Cette hypothèse est étayée par des travaux préliminaires réalisés en fin de thèse mais que je n'ai pu incorporer dans ce manuscrit.

La dernière étape nécessaire à l'intégration du LAT requiert la réparation du site de coupure et la ligation de l'insert dans son nouveaux locus. Cette étape reste la moins bien comprise pour l'ensemble des éléments mobiles. Plusieurs études laissent cependant supposer que celle-ci nécessite l'action de la machinerie de réparation de l'ADN cellulaire. En accord avec cette hypothèse, de nombreux travaux ont montré la capacité du virus à mobiliser les différentes voies de cette machinerie pour effectuer certaines étapes clés de son cycle (Schumacher et al. 2012 ; Balasubramanian et al. 2010 Sarisky et al. 1994). Par analogie avec les éléments L1, il est donc envisageable de penser que le LAT détourne l'une de ses voies pour finaliser son intégration.

La découverte des propriétés du LAT_{2kb} ouvre de nouveaux champs d'investigation

Le LAT_{2kb} est le premier élément mobile découvert chez un virus eucaryote. Sa structure de lariat et ses propriétés catalytiques le rapprochent des éléments mobiles de classe II. Néanmoins, à ce jour, ce type d'éléments n'a été mis en évidence que chez des procaryotes ou au sein des organelles eucaryotes (mitochondries et chloroplastes). Le LAT_{2kb} est le premier élément de ce type à être actif au sein du noyau de la cellule eucaryote.

La première question qu'on peut se poser porte sur l'existence d'autres éléments semblables au sein du monde eucaryote. En d'autres termes, existe-t-il des lariats stables au sein de la cellule eucaryote qui auraient des propriétés identiques à celles du LAT_{2kb} ? Peu d'exemple de lariats stables d'origine eucaryote ont été décrits dans la littérature à l'exception d'un intron de la région constante du gène codant la chaîne β du récepteur des cellules T

(Qian et al 1992). Néanmoins, la possibilité qu'il existe des éléments dotés de propriétés semblables chez les eucaryotes mérite d'être explorée.

Une autre possibilité est que cet élément soit spécifique aux herpesvirus. Il est important ici de se rappeler que tous les membres de cette famille virale ont une origine commune et appartiennent au même lignage que certains phages caudés. La présence de cet élément dans le génome de ce lignage viral remonterait donc à des temps ancestraux. Si tel est le cas, on devrait retrouver des éléments dotés de propriétés semblables chez de nombreux herpesvirus très éloignés dans l'arbre phylogénétique de la famille. Si cette hypothèse est correcte, les propriétés de ces éléments peuvent être très diverses. Celles du LAT sont le résultat de dizaine de millions d'années de co-évolution entre le virus et son hôte.

Dans le cas contraire (*i.e.* si un tel élément n'est retrouvé que dans une sous famille de la famille des herpes), il faudra supposer que le LAT a été acquis par transfert horizontal et cela implique que des éléments dotés de propriétés semblables existent chez les eucaryotes.

Implications en santé publique

Le fait que HSV-1 (et probablement d'autres herpesvirus, à commencer par HSV-2) soit capable de manipuler le génome de son hôte ouvre aussi de nouveaux champs d'investigation en santé publique. En effet, les études épidémiologiques montrent que plus de 90% de la population humaine est séropositive pour HSV-1. Le virus est donc latent dans la grande majorité de cette population. On peut presque considérer qu'il existe une symbiose entre le virus et son hôte, issue de millions d'années de coévolution. Il est probable même que l'infection par HSV-1 confère (ou plutôt conférerait) un avantage sélectif : on peut imaginer par exemple qu'elle protège d'infections par d'autres virus neurotropes autrement plus dangereux. L'espérance de vie de l'espèce humaine ayant considérablement augmenté en quelques générations, cette symbiose est peut-être maintenant déséquilibrée. Ainsi, il est possible que certaines maladies chroniques du système nerveux, certaines pathologies neurodégénératives soient directement ou indirectement causées par la latence herpétique et la capacité des LAT à transposer. Il n'est pas non plus impossible que certaines tumeurs cérébrales soient causées par des intégrations inappropriées du LAT dans certains gènes cellulaires.

Bibliographie

1. Ahmed, M., Lock, M., Miller, C. G. & Fraser, N. W. Regions of the herpes simplex virus type 1 latency-associated transcript that protect cells from apoptosis in vitro and protect neuronal cells in vivo. *J Virol* **76**, 717–29 (2002).
2. Alwine, J. C., Steinhart, W. L. & Hill, C. W. Transcription of herpes simplex type 1 DNA in nuclei isolated from infected HEp-2 and KB cells. *Virology* **60**, 302–7 (1974).
3. Amelio, A. L., Giordani, N. V., Kubat, N. J., O’neil, J. E. & Bloom, D. C. Deacetylation of the herpes simplex virus type 1 latency-associated transcript (LAT) enhancer and a decrease in LAT abundance precede an increase in ICP0 transcriptional permissiveness at early times postexplant. *J Virol* **80**, 2063–8 (2006).
4. Amelio, A. L., McAnany, P. K. & Bloom, D. C. A chromatin insulator-like element in the herpes simplex virus type 1 latency-associated transcript region binds CCCTC-binding factor and displays enhancer-blocking and silencing activities. *J Virol* **80**, 2358–68 (2006).
5. Amir, J., Harel, L., Smetana, Z. & Varsano, I. Treatment of herpes simplex gingivostomatitis with aciclovir in children: a randomised double blind placebo controlled study. *BMJ* **314**, 1800–3 (1997).
6. Amir, J., Nussinovitch, M., Kleper, R., Cohen, H. A. & Varsano, I. Primary herpes simplex virus type 1 gingivostomatitis in pediatric personnel. *Infection* **25**, 310–2 (1997).
7. Argnani, R., Lufino, M., Manservigi, M. & Manservigi, R. Replication-competent herpes simplex vectors: design and applications. *Gene Ther* **12 Suppl 1**, S170–7 (2005).
8. Atanasiu, D. & Fraser, N. W. The stable 2-kilobase latency-associated transcript of herpes simplex virus type 1 can alter the assembly of the 60S ribosomal subunit and is exported from nucleus to cytoplasm by a CRM1-dependent pathway. *J Virol* **81**, 7695–701 (2007).
9. Athanikar, J. N., Badge, R. M. & Moran, J. V. A YY1-binding site is required for accurate human LINE-1 transcription initiation. *Nucleic Acids Res* **32**, 3846–55 (2004).
10. Babushok, D. V. & Kazazian, H. H. J. Progress in understanding the biology of the human mutagen LINE-1. *Hum Mutat* **28**, 527–39 (2007).
11. Baker, M. L., Jiang, W., Rixon, F. J. & Chiu, W. Common ancestry of herpesviruses and tailed DNA bacteriophages. *J Virol* **79**, 14967–70 (2005).

12. Balasubramanian, N., Bai, P., Buchek, G., Korza, G. & Weller, S. K. Physical interaction between the herpes simplex virus type 1 exonuclease, UL12, and the DNA double-strand break-sensing MRN complex. *J Virol* **84**, 12504–14 (2010).
13. Batterson, W., Furlong, D. & Roizman, B. Molecular genetics of herpes simplex virus. VIII. further characterization of a temperature-sensitive mutant defective in release of viral DNA and in other stages of the viral reproductive cycle. *J Virol* **45**, 397–407 (1983).
14. Batterson, W. & Roizman, B. Characterization of the herpes simplex virion-associated factor responsible for the induction of alpha genes. *J Virol* **46**, 371–7 (1983).
15. Beauregard, A., Curcio, M. J. & Belfort, M. The take and give between retrotransposable elements and their hosts. *Annu Rev Genet* **42**, 587–617 (2008).
16. Benedetti, J., Corey, L. & Ashley, R. Recurrence rates in genital herpes after symptomatic first-episode infection. *Ann Intern Med* **121**, 847–54 (1994).
17. Bergefall, K. *et al.* Chondroitin sulfate characterized by the E-disaccharide unit is a potent inhibitor of herpes simplex virus infectivity and provides the virus binding sites on gro2C cells. *J Biol Chem* **280**, 32193–9 (2005).
18. Berthomme, H., Lokensgard, J., Yang, L., Margolis, T. & Feldman, L. T. Evidence for a bidirectional element located downstream from the herpes simplex virus type 1 latency-associated promoter that increases its activity during latency. *J Virol* **74**, 3613–22 (2000).
19. Bertke, A. S. *et al.* Latency-associated transcript (LAT) exon 1 controls herpes simplex virus species-specific phenotypes: reactivation in the guinea pig genital model and neuron subtype-specific latent expression of LAT. *J Virol* **83**, 10007–15 (2009).
20. Bloom, D. C., Devi-Rao, G. B., Hill, J. M., Stevens, J. G. & Wagner, E. K. Molecular analysis of herpes simplex virus type 1 during epinephrine-induced reactivation of latently infected rabbits in vivo. *J Virol* **68**, 1283–92 (1994).
21. Bloom, D. C., Giordani, N. V. & Kwiatkowski, D. L. Epigenetic regulation of latent HSV-1 gene expression. *Biochim Biophys Acta* **1799**, 246–56 (2010).
22. Booy, F. P. *et al.* Liquid-crystalline, phage-like packing of encapsidated DNA in herpes simplex virus. *Cell* **64**, 1007–15 (1991).
23. Boutell, C. *et al.* Herpes simplex virus type 1 ICP0 phosphorylation mutants impair the E3 ubiquitin ligase activity of ICP0 in a cell type-dependent manner. *J Virol* **82**, 10647–56 (2008).

24. Boutell, C., Orr, A. & Everett, R. D. PML residue lysine 160 is required for the degradation of PML induced by herpes simplex virus type 1 regulatory protein ICP0. *J Virol* **77**, 8686–94 (2003).
25. Branco, F. J. & Fraser, N. W. Herpes simplex virus type 1 latency-associated transcript expression protects trigeminal ganglion neurons from apoptosis. *J Virol* **79**, 9019–25 (2005).
26. Bravo, F. J. *et al.* Effect of antibody alone and combined with acyclovir on neonatal herpes simplex virus infection in guinea pigs. *J Infect Dis* **173**, 1–6 (1996).
27. Bruni, R., Fineschi, B., Ogle, W. O. & Roizman, B. A novel cellular protein, p60, interacting with both herpes simplex virus 1 regulatory proteins ICP22 and ICP0 is modified in a cell-type-specific manner and is recruited to the nucleus after infection. *J Virol* **73**, 3810–7 (1999).
28. Bryson, Y. *et al.* Risk of acquisition of genital herpes simplex virus type 2 in sex partners of persons with genital herpes: a prospective couple study. *J Infect Dis* **167**, 942–6 (1993).
29. Calle, A. *et al.* Nucleolin is required for an efficient herpes simplex virus type 1 infection. *J Virol* **82**, 4762–73 (2008).
30. Canning, M., Boutell, C., Parkinson, J. & Everett, R. D. A RING finger ubiquitin ligase is protected from autocatalyzed ubiquitination and degradation by binding to ubiquitin-specific protease USP7. *J Biol Chem* **279**, 38160–8 (2004).
31. Cannon, M. & Cesarman, E. Kaposi's sarcoma-associated herpes virus and acquired immunodeficiency syndrome-related malignancy. *Semin Oncol* **27**, 409–19 (2000).
32. Capy, P., Gasperi, G., Biemont, C. & Bazin, C. Stress and transposable elements: co-evolution or useful parasites? *Heredity (Edinb)* **85 (Pt 2)**, 101–6 (2000).
33. Carney, O., Ross, E., Ikkos, G. & Mindel, A. The effect of suppressive oral acyclovir on the psychological morbidity associated with recurrent genital herpes. *Genitourin Med* **69**, 457–9 (1993).
34. Carpenter, D. *et al.* A speculated ribozyme site in the herpes simplex virus type 1 latency-associated transcript gene is not essential for a wild-type reactivation phenotype. *J Neurovirol* **14**, 558–62 (2008).
35. Carter, K. L. & Roizman, B. Alternatively spliced mRNAs predicted to yield frame-shift proteins and stable intron 1 RNAs of the herpes simplex virus 1 regulatory gene alpha 0

- accumulate in the cytoplasm of infected cells. *Proc Natl Acad Sci U S A* **93**, 12535–40 (1996).
36. Carter, K. L., Ward, P. L. & Roizman, B. Characterization of the products of the U(L)43 gene of herpes simplex virus 1: potential implications for regulation of gene expression by antisense transcription. *J Virol* **70**, 7663–8 (1996).
 37. Catez, F. *et al.* HSV-1 genome subnuclear positioning and associations with host-cell PML-NBs and centromeres regulate LAT locus transcription during latency in neurons. *PLoS Pathog* **8**, e1002852 (2012).
 38. Chao, W., Huynh, K. D., Spencer, R. J., Davidow, L. S. & Lee, J. T. CTCF, a candidate trans-acting factor for X-inactivation choice. *Science* **295**, 345–7 (2002).
 39. Chelbi-Alix, M. K. *et al.* Induction of the PML protein by interferons in normal and APL cells. *Leukemia* **9**, 2027–33 (1995).
 40. Chen, Q. *et al.* CTCF-dependent chromatin boundary element between the latency-associated transcript and ICP0 promoters in the herpes simplex virus type 1 genome. *J Virol* **81**, 5192–201 (2007).
 41. Chen, S. H., Kramer, M. F., Schaffer, P. A. & Coen, D. M. A viral function represses accumulation of transcripts from productive-cycle genes in mouse ganglia latently infected with herpes simplex virus. *J Virol* **71**, 5878–84 (1997).
 42. Chen, S.-H. *et al.* Neither LAT nor open reading frame P mutations increase expression of spliced or intron-containing ICP0 transcripts in mouse ganglia latently infected with herpes simplex virus. *J Virol* **76**, 4764–72 (2002).
 43. Chen, X.-P., Mata, M., Kelley, M., Glorioso, J. C. & Fink, D. J. The relationship of herpes simplex virus latency associated transcript expression to genome copy number: a quantitative study using laser capture microdissection. *J Neurovirol* **8**, 204–10 (2002).
 44. Cheng, H., Shen, N., Pei, J. & Grishin, N. V. Double-stranded DNA bacteriophage prohead protease is homologous to herpesvirus protease. *Protein Sci* **13**, 2260–9 (2004).
 45. Chentoufi, A. A. *et al.* The herpes simplex virus 1 latency-associated transcript promotes functional exhaustion of virus-specific CD8⁺ T cells in latently infected trigeminal ganglia: a novel immune evasion mechanism. *J Virol* **85**, 9127–38 (2011).
 46. Cheung, P., Panning, B. & Smiley, J. R. Herpes simplex virus immediate-early proteins ICP0 and ICP4 activate the endogenous human alpha-globin gene in nonerythroid cells. *J Virol* **71**, 1784–93 (1997).

47. Chou, J., Kern, E. R., Whitley, R. J. & Roizman, B. Mapping of herpes simplex virus-1 neurovirulence to gamma 134.5, a gene nonessential for growth in culture. *Science* **250**, 1262–6 (1990).
48. Chowdhury, D. & Lieberman, J. Death by a thousand cuts: granzyme pathways of programmed cell death. *Annu Rev Immunol* **26**, 389–420 (2008).
49. Christie, S. N., McCaughey, C., McBride, M. & Coyle, P. V. Herpes simplex type 1 and genital herpes in Northern Ireland. *Int J STD AIDS* **8**, 68–9 (1997).
50. Chung, S., Zuker, C. & Lodish, H. F. A repetitive and apparently transposable DNA sequence in *Dictyostelium discoideum* associated with developmentally regulated RNAs. *Nucleic Acids Res* **11**, 4835–52 (1983).
51. Clement, C. *et al.* Upregulation of mouse genes in HSV-1 latent TG after butyrate treatment implicates the multiple roles of the LAT-ICP0 locus. *Invest Ophthalmol Vis Sci* **52**, 1770–9 (2011).
52. Cliffe, A. R., Garber, D. A. & Knipe, D. M. Transcription of the herpes simplex virus latency-associated transcript promotes the formation of facultative heterochromatin on lytic promoters. *J Virol* **83**, 8182–90 (2009).
53. Coen, D. M. *et al.* Thymidine kinase-negative herpes simplex virus mutants establish latency in mouse trigeminal ganglia but do not reactivate. *Proc Natl Acad Sci U S A* **86**, 4736–40 (1989).
54. Cohen, J. D. *et al.* [Association of Kaposi sarcoma--multiple myeloma. A new case]. *Rev Med Interne* **21**, 998–1000 (2000).
55. Cohrs, R. J. *et al.* Analysis of individual human trigeminal ganglia for latent herpes simplex virus type 1 and varicella-zoster virus nucleic acids using real-time PCR. *J Virol* **74**, 11464–71 (2000).
56. Connolly, S. A., Jackson, J. O., Jardetzky, T. S. & Longnecker, R. Fusing structure and function: a structural view of the herpesvirus entry machinery. *Nat Rev Microbiol* **9**, 369–81 (2011).
57. Cook, W. J., Wobbe, K. K., Boni, J. & Coen, D. M. Regulation of neighboring gene expression by the herpes simplex virus type 1 thymidine kinase gene. *Virology* **218**, 193–203 (1996).
58. Copeland, A. M., Newcomb, W. W. & Brown, J. C. Herpes simplex virus replication: roles of viral proteins and nucleoporins in capsid-nucleus attachment. *J Virol* **83**, 1660–8 (2009).

59. Cost, G. J., Feng, Q., Jacquier, A. & Boeke, J. D. Human L1 element target-primed reverse transcription in vitro. *EMBO J* **21**, 5899–910 (2002).
60. Costanzo, F., Campadelli-Fiume, G., Foa-Tomasi, L. & Cassai, E. Evidence that herpes simplex virus DNA is transcribed by cellular RNA polymerase B. *J Virol* **21**, 996–1001 (1977).
61. Cousineau, B. *et al.* Retrohoming of a bacterial group II intron: mobility via complete reverse splicing, independent of homologous DNA recombination. *Cell* **94**, 451–62 (1998).
62. Cuddapah, S. *et al.* Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res* **19**, 24–32 (2009).
63. Cunningham, A. L. & Merigan, T. C. gamma Interferon production appears to predict time of recurrence of herpes labialis. *J Immunol* **130**, 2397–400 (1983).
64. Cunningham, A. L. & Mikloska, Z. The Holy Grail: immune control of human herpes simplex virus infection and disease. *Herpes* **8 Suppl 1**, 6A–10A (2001).
65. Cusini, M. & Ghislanzoni, M. The importance of diagnosing genital herpes. *J Antimicrob Chemother* **47 Suppl T1**, 9–16 (2001).
66. Da Costa, X. J. *et al.* Humoral response to herpes simplex virus is complement-dependent. *Proc Natl Acad Sci U S A* **96**, 12708–12 (1999).
67. Dai, L. & Zimmerly, S. Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior. *Nucleic Acids Res* **30**, 1091–102 (2002).
68. Dai, L. & Zimmerly, S. The dispersal of five group II introns among natural populations of *Escherichia coli*. *RNA* **8**, 1294–307 (2002).
69. Danaher, R. J. *et al.* Histone deacetylase inhibitors induce reactivation of herpes simplex virus type 1 in a latency-associated transcript-independent manner in neuronal cells. *J Neurovirol* **11**, 306–17 (2005).
70. Darlington, R. W. & Moss, L. H. 3rd Herpesvirus envelopment. *J Virol* **2**, 48–55 (1968).
71. DeBerardinis, R. J., Goodier, J. L., Ostertag, E. M. & Kazazian, H. H. J. Rapid amplification of a retrotransposon subfamily is evolving the mouse genome. *Nat Genet* **20**, 288–90 (1998).
72. Deiss, L. P., Chou, J. & Frenkel, N. Functional domains within the a sequence involved in the cleavage-packaging of herpes simplex virus DNA. *J Virol* **59**, 605–18 (1986).

73. del Rio, T., Ch'ng, T. H., Flood, E. A., Gross, S. P. & Enquist, L. W. Heterogeneity of a fluorescent tegument component in single pseudorabies virus virions and enveloped axonal assemblies. *J Virol* **79**, 3903–19 (2005).
74. Deshmane, S. L. & Fraser, N. W. During latency, herpes simplex virus type 1 DNA is associated with nucleosomes in a chromatin structure. *J Virol* **63**, 943–7 (1989).
75. Diaz, J.-J., Giraud, S. & Greco, A. Alteration of ribosomal protein maps in herpes simplex virus type 1 infection. *J Chromatogr B Analyt Technol Biomed Life Sci* **771**, 237–49 (2002).
76. Dixon, R. A. & Schaffer, P. A. Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175. *J Virol* **36**, 189–203 (1980).
77. Doucet, A. J. *et al.* Characterization of LINE-1 ribonucleoprotein particles. *PLoS Genet* **6**, (2010).
78. Dressler, G. R., Rock, D. L. & Fraser, N. W. Latent herpes simplex virus type 1 DNA is not extensively methylated in vivo. *J Gen Virol* **68 (Pt 6)**, 1761–5 (1987).
79. Duda, R. L., Hendrix, R. W., Huang, W. M. & Conway, J. F. Shared architecture of bacteriophage SPO1 and herpesvirus capsids. *Curr Biol* **16**, R11–3 (2006).
80. Ebina, H. & Levin, H. L. Stress management: how cells take control of their transposons. *Mol Cell* **27**, 180–1 (2007).
81. Edwards, S. & White, C. Genital herpes simplex virus type 1 in women. *Genitourin Med* **70**, 426 (1994).
82. Efstathiou, S., Kemp, S., Darby, G. & Minson, A. C. The role of herpes simplex virus type 1 thymidine kinase in pathogenesis. *J Gen Virol* **70 (Pt 4)**, 869–79 (1989).
83. Enquist, L. W., Tomishima, M. J., Gross, S. & Smith, G. A. Directional spread of an alpha-herpesvirus in the nervous system. *Vet Microbiol* **86**, 5–16 (2002).
84. Everett, R. D. ICP0, a regulator of herpes simplex virus during lytic and latent infection. *Bioessays* **22**, 761–70 (2000).
85. Everett, R. D., Earnshaw, W. C., Findlay, J. & Lomonte, P. Specific destruction of kinetochore protein CENP-C and disruption of cell division by herpes simplex virus immediate-early protein Vmw110. *EMBO J* **18**, 1526–38 (1999).
86. Everett, R. D. Interactions between DNA viruses, ND10 and the DNA damage response. *Cell Microbiol* **8**, 365–74 (2006).

87. Faber, S. W. & Wilcox, K. W. Association of the herpes simplex virus regulatory protein ICP4 with specific nucleotide sequences in DNA. *Nucleic Acids Res* **14**, 6067–83 (1986).
88. Farrell, M. J., Dobson, A. T. & Feldman, L. T. Herpes simplex virus latency-associated transcript is a stable intron. *Proc Natl Acad Sci U S A* **88**, 790–4 (1991).
89. Favoreel, H. W., Nauwynck, H. J. & Pensaert, M. B. Immunological hiding of herpesvirus-infected cells. *Arch Virol* **145**, 1269–90 (2000).
90. Feierbach, B., Bisher, M., Goodhouse, J. & Enquist, L. W. In vitro analysis of transneuronal spread of an alphaherpesvirus infection in peripheral nervous system neurons. *J Virol* **81**, 6846–57 (2007).
91. Fenderson, B. A., Hahnel, A. C. & Eddy, E. M. Immunohistochemical localization of two monoclonal antibody-defined carbohydrate antigens during early murine embryogenesis. *Dev Biol* **100**, 318–27 (1983).
92. Ferat, J. L., Le Gouar, M. & Michel, F. Multiple group II self-splicing introns in mobile DNA from *Escherichia coli*. *C R Acad Sci III* **317**, 141–8 (1994).
93. Ferat, J. L. & Michel, F. Group II self-splicing introns in bacteria. *Nature* **364**, 358–61 (1993).
94. Feschotte, C., Jiang, N. & Wessler, S. R. Plant transposable elements: where genetics meets genomics. *Nat Rev Genet* **3**, 329–41 (2002).
95. Feschotte, C. & Wessler, S. R. Mariner-like transposases are widespread and diverse in flowering plants. *Proc Natl Acad Sci U S A* **99**, 280–5 (2002).
96. Fleming, D. T. *et al.* Herpes simplex virus type 2 in the United States, 1976 to 1994. *N Engl J Med* **337**, 1105–11 (1997).
97. Fruh, K. *et al.* A viral inhibitor of peptide transporters for antigen presentation. *Nature* **375**, 415–8 (1995).
98. Garber, D. A., Schaffer, P. A. & Knipe, D. M. A LAT-associated function reduces productive-cycle gene expression during acute infection of murine sensory neurons with herpes simplex virus type 1. *J Virol* **71**, 5885–93 (1997).
99. Gardner, D. P., Byrne, G. W., Ruddle, F. H. & Kappen, C. Spatial and temporal regulation of a lacZ reporter transgene in a binary transgenic mouse system. *Transgenic Res* **5**, 37–48 (1996).

100. Geraghty, R. J., Krummenacher, C., Cohen, G. H., Eisenberg, R. J. & Spear, P. G. Entry of alphaherpesviruses mediated by poliovirus receptor-related protein 1 and poliovirus receptor. *Science* **280**, 1618–20 (1998).
101. Ghiasi, H. *et al.* Infection of BALB/c mice with a herpes simplex virus type 1 recombinant virus expressing IFN-gamma driven by the LAT promoter. *Virology* **302**, 144–54 (2002).
102. Gilbert, N., Lutz, S., Morrish, T. A. & Moran, J. V. Multiple fates of L1 retrotransposition intermediates in cultured human cells. *Mol Cell Biol* **25**, 7780–95 (2005).
103. Gilbert, N., Lutz-Prigge, S. & Moran, J. V. Genomic deletions created upon LINE-1 retrotransposition. *Cell* **110**, 315–25 (2002).
104. Giordani, N. V. *et al.* During herpes simplex virus type 1 infection of rabbits, the ability to express the latency-associated transcript increases latent-phase transcription of lytic genes. *J Virol* **82**, 6056–60 (2008).
105. Goins, W. F. *et al.* A novel latency-active promoter is contained within the herpes simplex virus type 1 UL flanking repeats. *J Virol* **68**, 2239–52 (1994).
106. Goldenberg, D., Mador, N., Ball, M. J., Panet, A. & Steiner, I. The abundant latency-associated transcripts of herpes simplex virus type 1 are bound to polyribosomes in cultured neuronal cells and during latent infection in mouse trigeminal ganglia. *J Virol* **71**, 2897–904 (1997).
107. Goldenberg, D., Mador, N., Panet, A. & Steiner, I. Tissue specific distribution of the herpes simplex virus type 1 latency-associated transcripts on polyribosomes during latent infection. *J Neurovirol* **4**, 426–32 (1998).
108. Grandbastien, M. A. *et al.* The expression of the tobacco Tnt1 retrotransposon is linked to plant defense responses. *Genetica* **100**, 241–52 (1997).
109. Grandbastien, M.-A. *et al.* Stress activation and genomic impact of Tnt1 retrotransposons in Solanaceae. *Cytogenet Genome Res* **110**, 229–41 (2005).
110. Granlund, M., Michel, F. & Norgren, M. Mutually exclusive distribution of IS1548 and GBSi1, an active group II intron identified in human isolates of group B streptococci. *J Bacteriol* **183**, 2560–9 (2001).
111. Granzow, H., Klupp, B. G. & Mettenleiter, T. C. Entry of pseudorabies virus: an immunogold-labeling study. *J Virol* **79**, 3200–5 (2005).

112. Greco, A., Diaz, J.-J., Thouvenot, D. & Morfin, F. Novel targets for the development of anti-herpes compounds. *Infect Disord Drug Targets* **7**, 11–8 (2007).
113. Green, J. & Kocsis, A. Psychological factors in recurrent genital herpes. *Genitourin Med* **73**, 253–8 (1997).
114. Grosse, S. *et al.* Meganuclease-mediated Inhibition of HSV1 Infection in Cultured Cells. *Mol. Ther.* **19**, 694–702 (2011).
115. Gussow, A. M. *et al.* Tissue-specific splicing of the herpes simplex virus type 1 latency-associated transcript (LAT) intron in LAT transgenic mice. *J Virol* **80**, 9414–23 (2006).
116. Halford, W. P., Gebhardt, B. M. & Carr, D. J. Mechanisms of herpes simplex virus type 1 reactivation. *J Virol* **70**, 5051–60 (1996).
117. Hardy, W. R. & Sandri-Goldin, R. M. Herpes simplex virus inhibits host cell splicing, and regulatory protein ICP27 is required for this effect. *J Virol* **68**, 7790–9 (1994).
118. Hashido, M., Kawana, T., Matsunaga, Y. & Inouye, S. Changes in prevalence of herpes simplex virus type 1 and 2 antibodies from 1973 to 1993 in the rural districts of Japan. *Microbiol Immunol* **43**, 177–80 (1999).
119. Hashido, M. *et al.* Herpes simplex virus types 1 and 2, chlamydia, syphilis, and toxoplasma in pregnant Japanese women with HTLV-I. *J Acquir Immune Defic Syndr Hum Retrovirol* **17**, 95–7 (1998).
120. He, Y. & Casaccia-Bonnel, P. The Yin and Yang of YY1 in the nervous system. *J Neurochem* **106**, 1493–502 (2008).
121. Herold, B. C., WuDunn, D., Soltys, N. & Spear, P. G. Glycoprotein C of herpes simplex virus type 1 plays a principal role in the adsorption of virus to cells and in infectivity. *J Virol* **65**, 1090–8 (1991).
122. Hill, J. M., Sedarati, F., Javier, R. T., Wagner, E. K. & Stevens, J. G. Herpes simplex virus latent phase transcription facilitates in vivo reactivation. *Virology* **174**, 117–25 (1990).
123. Honess, R. W. & Roizman, B. Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. *J Virol* **14**, 8–19 (1974).
124. Hong, X., Scofield, D. G. & Lynch, M. Intron size, abundance, and distribution within untranslated regions of genes. *Mol Biol Evol* **23**, 2392–404 (2006).
125. Hook, E. W. 3rd *et al.* Herpes simplex virus infection as a risk factor for human immunodeficiency virus infection in heterosexuals. *J Infect Dis* **165**, 251–5 (1992).

126. Ikeda, K., Nakayashiki, H., Takagi, M., Tosa, Y. & Mayama, S. Heat shock, copper sulfate and oxidative stress activate the retrotransposon MAGGY resident in the plant pathogenic fungus *Magnaporthe grisea*. *Mol Genet Genomics* **266**, 318–25 (2001).
127. Imai, A. *et al.* [A case of successful treatment of disseminated herpes simplex virus infection with acyclovir on the basis of early diagnosis by esophageal biopsy]. *Nihon Shokakibyo Gakkai Zasshi* **102**, 1417–22 (2005).
128. Imai, Y., Apakupakul, K., Krause, P. R., Halford, W. P. & Margolis, T. P. Investigation of the mechanism by which herpes simplex virus type 1 LAT sequences modulate preferential establishment of latent infection in mouse trigeminal ganglia. *J Virol* **83**, 7873–82 (2009).
129. Jaber, T. *et al.* Identification of a novel herpes simplex virus type 1 transcript and protein (AL3) expressed during latency. *J Gen Virol* **90**, 2342–52 (2009).
130. Javier, R. T., Stevens, J. G., Dissette, V. B. & Wagner, E. K. A herpes simplex virus transcript abundant in latently infected neurons is dispensable for establishment of the latent state. *Virology* **166**, 254–7 (1988).
131. Jean, S., LeVan, K. M., Song, B., Levine, M. & Knipe, D. M. Herpes simplex virus 1 ICP27 is required for transcription of two viral late (gamma 2) genes in infected cells. *Virology* **283**, 273–84 (2001).
132. Jennings, S. R., Bonneau, R. H., Smith, P. M., Wolcott, R. M. & Chervenak, R. CD4-positive T lymphocytes are required for the generation of the primary but not the secondary CD8-positive cytolytic T lymphocyte response to herpes simplex virus in C57BL/6 mice. *Cell Immunol* **133**, 234–52 (1991).
133. Jiang, X. *et al.* The herpes simplex virus type 1 latency-associated transcript can protect neuron-derived C1300 and Neuro2A cells from granzyme B-induced apoptosis and CD8 T-cell killing. *J Virol* **85**, 2325–32 (2011).
134. Jin, L. *et al.* A herpes simplex virus type 1 mutant expressing a baculovirus inhibitor of apoptosis gene in place of latency-associated transcript has a wild-type reactivation phenotype in the mouse. *J Virol* **79**, 12286–95 (2005).
135. Jones-Rhoades, M. W., Bartel, D. P. & Bartel, B. MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol* **57**, 19–53 (2006).
136. Kattenhorn, L. M., Korbel, G. A., Kessler, B. M., Spooner, E. & Ploegh, H. L. A deubiquitinating enzyme encoded by HSV-1 belongs to a family of cysteine proteases that is conserved across the family Herpesviridae. *Mol Cell* **19**, 547–57 (2005).

137. Kazazian, H. H. J. Mobile elements: drivers of genome evolution. *Science* **303**, 1626–32 (2004).
138. Kelly, B. J., Fraefel, C., Cunningham, A. L. & Diefenbach, R. J. Functional roles of the tegument proteins of herpes simplex virus type 1. *Virus Res* **145**, 173–86 (2009).
139. Kemp, L. M., Gelman, I. H., Silverstein, S. J. & Latchman, D. S. Regulation of herpes simplex virus immediate-early gene promoters in mouse neuroblastoma cells. *Neurosci Lett* **118**, 185–8 (1990).
140. Kim, J. S., Enquist, L. W. & Card, J. P. Circuit-specific coinfection of neurons in the rat central nervous system with two pseudorabies virus recombinants. *J Virol* **73**, 9521–31 (1999).
141. Kinghorn, A. Interventions to reduce mother-to-child transmission in South Africa. *AIDS Anal Afr* **8**, 10–1 (1998).
142. Knickelbein, J. E. *et al.* Noncytotoxic lytic granule-mediated CD8⁺ T cell inhibition of HSV-1 reactivation from neuronal latency. *Science* **322**, 268–71 (2008).
143. *Fields Virology, 4th Edition.* (Lippincott Williams & Wilkins: 2001).
144. Koelle, D. M., Frank, J. M., Johnson, M. L. & Kwok, W. W. Recognition of herpes simplex virus type 2 tegument proteins by CD4 T cells infiltrating human genital herpes lesions. *J Virol* **72**, 7476–83 (1998).
145. Koelle, D. M. *et al.* Herpes simplex virus infection of human fibroblasts and keratinocytes inhibits recognition by cloned CD8⁺ cytotoxic T lymphocytes. *J Clin Invest* **91**, 961–8 (1993).
146. Kohl, S. *et al.* Effect of intravenously administered immune globulin on functional antibody to herpes simplex virus in low birth weight neonates. *J Pediatr* **115**, 135–9 (1989).
147. Koken, M. H. *et al.* The PML growth-suppressor has an altered expression in human oncogenesis. *Oncogene* **10**, 1315–24 (1995).
148. Kramer, M. F. *et al.* Herpes simplex virus 1 microRNAs expressed abundantly during latent infection are not essential for latency in mouse trigeminal ganglia. *Virology* **417**, 239–47 (2011).
149. Krause, P. R., Croen, K. D., Straus, S. E. & Ostrove, J. M. Detection and preliminary characterization of herpes simplex virus type 1 transcripts in latently infected human trigeminal ganglia. *J Virol* **62**, 4819–23 (1988).

150. Krause, P. R. *et al.* Expression of the herpes simplex virus type 2 latency-associated transcript enhances spontaneous reactivation of genital herpes in latently infected guinea pigs. *J Exp Med* **181**, 297–306 (1995).
151. Kristie, T. M. & Roizman, B. Separation of sequences defining basal expression from those conferring alpha gene recognition within the regulatory domains of herpes simplex virus 1 alpha genes. *Proc Natl Acad Sci U S A* **81**, 4065–9 (1984).
152. Krummenacher, C., Zabolotny, J. M. & Fraser, N. W. Selection of a nonconsensus branch point is influenced by an RNA stem-loop structure and is important to confer stability to the herpes simplex virus 2-kilobase latency-associated transcript. *J Virol* **71**, 5849–60 (1997).
153. Kubat, N. J., Amelio, A. L., Giordani, N. V. & Bloom, D. C. The herpes simplex virus type 1 latency-associated transcript (LAT) enhancer/rcr is hyperacetylated during latency independently of LAT transcription. *J Virol* **78**, 12508–18 (2004).
154. Kubat, N. J., Tran, R. K., McAnany, P. & Bloom, D. C. Specific histone tail modification and not DNA methylation is a determinant of herpes simplex virus type 1 latent gene expression. *J Virol* **78**, 1139–49 (2004).
155. Kuhn, J. *et al.* Partial functional complementation of a pseudorabies virus UL25 deletion mutant by herpes simplex virus type 1 pUL25 indicates overlapping functions of alphaherpesvirus pUL25 proteins. *J Virol* **82**, 5725–34 (2008).
156. Kuo, Y. C. & Lin, C. Y. Recurrent herpes simplex virus type 1 infection precipitated by the impaired production of interleukin-2, alpha-interferon, and cell-mediated cytotoxicity. *J Med Virol* **31**, 183–9 (1990).
157. Kwiatkowski, D. L., Thompson, H. W. & Bloom, D. C. The polycomb group protein Bmi1 binds to the herpes simplex virus 1 latent genome and maintains repressive histone marks during latency. *J Virol* **83**, 8173–81 (2009).
158. Labetoulle, M. *et al.* Neuronal pathways for the propagation of herpes simplex virus type 1 from one retina to the other in a murine model. *J Gen Virol* **81**, 1201–10 (2000).
159. Labetoulle, M. *et al.* Neuronal propagation of HSV1 from the oral mucosa to the eye. *Invest Ophthalmol Vis Sci* **41**, 2600–6 (2000).
160. Labetoulle, M. *et al.* HSV1 latency sites after inoculation in the lip: assessment of their localization and connections to the eye. *Invest Ophthalmol Vis Sci* **44**, 217–25 (2003).

161. Ladin, B. F., Blankenship, M. L. & Ben-Porat, T. Replication of herpesvirus DNA. V. Maturation of concatemeric DNA of pseudorabies virus to genome length is related to capsid formation. *J Virol* **33**, 1151–64 (1980).
162. Ladin, B. F., Ihara, S., Hampl, H. & Ben-Porat, T. Pathway of assembly of herpesvirus capsids: an analysis using DNA⁺ temperature-sensitive mutants of pseudorabies virus. *Virology* **116**, 544–61 (1982).
163. Lafferty, W. E., Coombs, R. W., Benedetti, J., Critchlow, C. & Corey, L. Recurrences after oral and genital herpes simplex virus infection. Influence of site of infection and viral type. *N Engl J Med* **316**, 1444–9 (1987).
164. Lafferty, W. E., Downey, L., Celum, C. & Wald, A. Herpes simplex virus type 1 as a cause of genital herpes: impact on surveillance and prevention. *J Infect Dis* **181**, 1454–7 (2000).
165. Lafferty, W. E. The changing epidemiology of HSV-1 and HSV-2 and implications for serological testing. *Herpes* **9**, 51–5 (2002).
166. Lambowitz, A. M. & Zimmerly, S. Mobile group II introns. *Annu Rev Genet* **38**, 1–35 (2004).
167. Langenberg, A. G., Corey, L., Ashley, R. L., Leong, W. P. & Straus, S. E. A prospective study of new infections with herpes simplex virus type 1 and type 2. Chiron HSV Vaccine Study Group. *N Engl J Med* **341**, 1432–8 (1999).
168. LaVail, J. H., Tauscher, A. N., Aghaian, E., Harrabi, O. & Sidhu, S. S. Axonal transport and sorting of herpes simplex virus components in a mature mouse visual system. *J Virol* **77**, 6117–26 (2003).
169. Lavie, L., Maldener, E., Brouha, B., Meese, E. U. & Mayer, J. The human L1 promoter: variable transcription initiation sites and a major impact of upstream flanking sequence on promoter activity. *Genome Res* **14**, 2253–60 (2004).
170. Leib, D. A. *et al.* A deletion mutant of the latency-associated transcript of herpes simplex virus type 1 reactivates from the latent state with reduced frequency. *J Virol* **63**, 2893–900 (1989).
171. Leuzinger, H. *et al.* Herpes simplex virus 1 envelopment follows two diverse pathways. *J Virol* **79**, 13047–59 (2005).
172. Li, S., Carpenter, D., Hsiang, C., Wechsler, S. L. & Jones, C. Herpes simplex virus type 1 latency-associated transcript inhibits apoptosis and promotes neurite sprouting in

- neuroblastoma cells following serum starvation by maintaining protein kinase B (AKT) levels. *J Gen Virol* **91**, 858–66 (2010).
173. Liesegang, T. J. Epidemiology of ocular herpes simplex. Natural history in Rochester, Minn, 1950 through 1982. *Arch Ophthalmol* **107**, 1160–5 (1989).
 174. Lilley, C. E. *et al.* A viral E3 ligase targets RNF8 and RNF168 to control histone ubiquitination and DNA damage responses. *EMBO J* **29**, 943–55 (2010).
 175. Lokensgard, J. R., Berthomme, H. & Feldman, L. T. The latency-associated promoter of herpes simplex virus type 1 requires a region downstream of the transcription start site for long-term expression during latency. *J Virol* **71**, 6714–9 (1997).
 176. Lokensgard, J. R., Bloom, D. C., Dobson, A. T. & Feldman, L. T. Long-term promoter activity during herpes simplex virus latency. *J Virol* **68**, 7148–58 (1994).
 177. Lomonte, P. & Everett, R. D. Herpes simplex virus type 1 immediate-early protein Vmw110 inhibits progression of cells through mitosis and from G(1) into S phase of the cell cycle. *J Virol* **73**, 9456–67 (1999).
 178. Lomonte, P., Sullivan, K. F. & Everett, R. D. Degradation of nucleosome-associated centromeric histone H3-like protein CENP-A induced by herpes simplex virus type 1 protein ICP0. *J Biol Chem* **276**, 5829–35 (2001).
 179. Lomonte, P. & Morency, E. Centromeric protein CENP-B proteasomal degradation induced by the viral protein ICP0. *FEBS Lett* **581**, 658–62 (2007).
 180. Lopez, C. Genetics of natural resistance to herpesvirus infections in mice. *Nature* **258**, 152–3 (1975).
 181. Lopez, P., Van Sant, C. & Roizman, B. Requirements for the nuclear-cytoplasmic translocation of infected-cell protein 0 of herpes simplex virus 1. *J Virol* **75**, 3832–40 (2001).
 182. Lopez, P., Jacob, R. J. & Roizman, B. Overexpression of promyelocytic leukemia protein precludes the dispersal of ND10 structures and has no effect on accumulation of infectious herpes simplex virus 1 or its proteins. *J Virol* **76**, 9355–67 (2002).
 183. Lopez-Flores, I. & Garrido-Ramos, M. A. The repetitive DNA content of eukaryotic genomes. *Genome Dyn* **7**, 1–28 (2010).
 184. Loret, S., Guay, G. & Lippe, R. Comprehensive characterization of extracellular herpes simplex virus type 1 virions. *J Virol* **82**, 8605–18 (2008).

185. Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72**, 595–605 (1993).
186. Lussignol, M. *et al.* The herpes simplex virus type 1 Us11 protein inhibits autophagy through its interaction with the protein kinase PKR. *J Virol* (2012).
187. Luxton, G. W. G. *et al.* Targeting of herpesvirus capsid transport in axons is coupled to association with specific sets of tegument proteins. *Proc Natl Acad Sci U S A* **102**, 5832–7 (2005).
188. Lycke, E., Kristensson, K., Svennerholm, B., Vahlne, A. & Ziegler, R. Uptake and transport of herpes simplex virus in neurites of rat dorsal root ganglia cells in culture. *J Gen Virol* **65** (Pt 1), 55–64 (1984).
189. Mador, N., Goldenberg, D., Cohen, O., Panet, A. & Steiner, I. Herpes simplex virus type 1 latency-associated transcripts suppress viral replication and reduce immediate-early gene mRNA levels in a neuronal cell line. *J Virol* **72**, 5067–75 (1998).
190. Mador, N. *et al.* Transgenic mouse with the herpes simplex virus type 1 latency-associated gene: expression and function of the transgene. *J Virol* **77**, 12421–9 (2003).
191. Mador, N., Panet, A. & Steiner, I. The latency-associated gene of herpes simplex virus type 1 (HSV-1) interferes with superinfection by HSV-1. *J Neurovirol* **8 Suppl 2**, 97–102 (2002).
192. Maggioncalda, J., Mehta, A., Fraser, N. W. & Block, T. M. Analysis of a herpes simplex virus type 1 LAT mutant with a deletion between the putative promoter and the 5' end of the 2.0-kilobase transcript. *J Virol* **68**, 7816–24 (1994).
193. Mahiet, C. *et al.* Structural variability of the herpes simplex virus 1 genome in vitro and in vivo. *J Virol* **86**, 8592–601 (2012).
194. Mahiet, C. *et al.* Structural variability of the herpes simplex virus 1 genome in vitro and in vivo. *J. Virol.* **86**, 8592–8601 (2012).
195. Maillet, S. *et al.* Herpes simplex virus type 1 latently infected neurons differentially express latency-associated and ICP0 transcripts. *J Virol* **80**, 9310–21 (2006).
196. Malik, H. S. & Henikoff, S. Positive selection of Iris, a retroviral envelope-derived host gene in *Drosophila melanogaster*. *PLoS Genet* **1**, e44 (2005).
197. Manickan, E. & Rouse, B. T. Roles of different T-cell subsets in control of herpes simplex virus infection determined by using T-cell-deficient mouse-models. *J Virol* **69**, 8178–9 (1995).

198. Margolis, T. P., Bloom, D. C., Dobson, A. T., Feldman, L. T. & Stevens, J. G. Decreased reporter gene expression during latent infection with HSV LAT promoter constructs. *Virology* **197**, 585–92 (1993).
199. Margolis, T. P., Imai, Y., Yang, L., Vallas, V. & Krause, P. R. Herpes simplex virus type 2 (HSV-2) establishes latent infection in a different population of ganglionic neurons than HSV-1: role of latency-associated transcripts. *J Virol* **81**, 1872–8 (2007).
200. Marshall, K. R., Lachmann, R. H., Efstathiou, S., Rinaldi, A. & Preston, C. M. Long-term transgene expression in mice infected with a herpes simplex virus type 1 mutant severely impaired for immediate-early gene expression. *J Virol* **74**, 956–64 (2000).
201. Mehta, A. *et al.* In situ DNA PCR and RNA hybridization detection of herpes simplex virus sequences in trigeminal ganglia of latently infected mice. *Virology* **206**, 633–40 (1995).
202. Melayah, D., Bonnivard, E., Chalhoub, B., Audeon, C. & Grandbastien, M. A. The mobility of the tobacco Tnt1 retrotransposon correlates with its transcriptional activation by fungal factors. *Plant J* **28**, 159–68 (2001).
203. Mertz, G. J., Benedetti, J., Ashley, R., Selke, S. A. & Corey, L. Risk factors for the sexual transmission of genital herpes. *Ann Intern Med* **116**, 197–202 (1992).
204. Metzger, E. E. & Whittum-Hudson, J. A. The dichotomy between herpes simplex virus type 1-induced ocular pathology and systemic immunity. *Invest Ophthalmol Vis Sci* **28**, 1533–40 (1987).
205. Mhiri, C. *et al.* The promoter of the tobacco Tnt1 retrotransposon is induced by wounding and by abiotic stress. *Plant Mol Biol* **33**, 257–66 (1997).
206. Michaelis, G., Saedler, H., Venkov, P. & Starlinger, P. Two insertions in the galactose operon having different sizes but homologous DNA sequences. *Mol Gen Genet* **104**, 371–7 (1969).
207. Michel, F. & Ferat, J. L. Structure and activities of group II introns. *Annu Rev Biochem* **64**, 435–61 (1995).
208. Mikloska, Z. & Cunningham, A. L. Herpes simplex virus type 1 glycoproteins gB, gC and gD are major targets for CD4 T-lymphocyte cytotoxicity in HLA-DR expressing human epidermal keratinocytes. *J Gen Virol* **79** (Pt 2), 353–61 (1998).
209. Mikloska, Z. & Cunningham, A. L. Alpha and gamma interferons inhibit herpes simplex virus type 1 infection and spread in epidermal cells after axonal transmission. *J Virol* **75**, 11821–6 (2001).

210. Miller, C. S., Danaher, R. J. & Jacob, R. J. Molecular aspects of herpes simplex virus I latency, reactivation, and recurrence. *Crit Rev Oral Biol Med* **9**, 541–62 (1998).
211. Mitchell, B. M. & Stevens, J. G. Neuroinvasive properties of herpes simplex virus type 1 glycoprotein variants are controlled by the immune response. *J Immunol* **156**, 246–55 (1996).
212. Mitchell, B. M., Bloom, D. C., Cohrs, R. J., Gilden, D. H. & Kennedy, P. G. E. Herpes simplex virus-1 and varicella-zoster virus latency in ganglia. *J Neurovirol* **9**, 194–204 (2003).
213. Mohl, B. S. *et al.* Intracellular localization of the pseudorabies virus large tegument protein pUL36. *J Virol* **83**, 9641–51 (2009).
214. Moran, J. V. *et al.* High frequency retrotransposition in cultured mammalian cells. *Cell* **87**, 917–27 (1996).
215. Morency, E., Coute, Y., Thomas, J., Texier, P. & Lomonte, P. The protein ICP0 of herpes simplex virus type 1 is targeted to nucleoli of infected cells. Brief report. *Arch Virol* **150**, 2387–95 (2005).
216. Mori, Y. & Yamanishi, K. HHV-6A, 6B, and 7: pathogenesis, host response, and clinical disease. (2007).
217. Morimoto, T. *et al.* Differences in the regulatory and functional effects of the Us3 protein kinase activities of herpes simplex virus 1 and 2. *J Virol* **83**, 11624–34 (2009).
218. Morrison, L. A. & Knipe, D. M. Contributions of antibody and T cell subsets to protection elicited by immunization with a replication-defective mutant of herpes simplex virus type 1. *Virology* **239**, 315–26 (1997).
219. Mukerjee, R., Kang, W., Suri, V. & Fraser, N. W. A non-consensus branch point plays an important role in determining the stability of the 2-kb LAT intron during acute and latent infections of herpes simplex virus type-1. *Virology* **324**, 340–9 (2004).
220. Mulik, S. *et al.* Role of miR-132 in angiogenesis after ocular infection with herpes simplex virus. *Am J Pathol* **181**, 525–34 (2012).
221. Naas, T. P. *et al.* An actively retrotransposing, novel subfamily of mouse L1 elements. *EMBO J* **17**, 590–7 (1998).
222. Nahmias, A. J., Lee, F. K. & Beckman-Nahmias, S. Sero-epidemiological and -sociological patterns of herpes simplex virus infection in the world. *Scand J Infect Dis Suppl* **69**, 19–36 (1990).

223. Naito, J., Mott, K. R., Osorio, N., Jin, L. & Perng, G.-C. Herpes simplex virus type 1 immediate-early protein ICP0 diffuses out of infected rabbit corneas. *J Gen Virol* **86**, 2979–88 (2005).
224. Naito, J. *et al.* Identification of a protein encoded in the herpes simplex virus type 1 latency associated transcript promoter region. *Virus Res* **108**, 101–10 (2005).
225. Negatsch, A., Mettenleiter, T. C. & Fuchs, W. Herpes simplex virus type 1 strain KOS carries a defective US9 and a mutated US8A gene. *J Gen Virol* **92**, 167–72 (2011).
226. Newcomb, W. W., Cockrell, S. K., Homa, F. L. & Brown, J. C. Polarized DNA ejection from the herpesvirus capsid. *J Mol Biol* **392**, 885–94 (2009).
227. Ng, A. K. *et al.* Construction of a herpes simplex virus type 1 mutant with only a three-nucleotide change in the branchpoint region of the latency-associated transcript (LAT) and the stability of its two-kilobase LAT intron. *J Virol* **78**, 12097–106 (2004).
228. Nicosia, M., Deshmane, S. L., Zabolotny, J. M., Valyi-Nagy, T. & Fraser, N. W. Herpes simplex virus type 1 latency-associated transcript (LAT) promoter deletion mutants can express a 2-kilobase transcript mapping to the LAT region. *J Virol* **67**, 7276–83 (1993).
229. Nicosia, M., Zabolotny, J. M., Lirette, R. P. & Fraser, N. W. The HSV-1 2-kb latency-associated transcript is found in the cytoplasm comigrating with ribosomal subunits during productive infection. *Virology* **204**, 717–28 (1994).
230. Norberg, P., Bergstrom, T., Rekabdar, E., Lindh, M. & Liljeqvist, J.-A. Phylogenetic analysis of clinical herpes simplex virus type 1 isolates identified three genetic groups and recombinant viruses. *J Virol* **78**, 10755–64 (2004).
231. Ochman, H., Gerber, A. S. & Hartl, D. L. Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**, 621–3 (1988).
232. Ohta, A., Yamauchi, Y., Muto, Y., Kimura, H. & Nishiyama, Y. Herpes simplex virus type 1 UL14 tegument protein regulates intracellular compartmentalization of major tegument protein VP16. *Virol J* **8**, 365 (2011).
233. Ojala, P. M., Sodeik, B., Ebersold, M. W., Kutay, U. & Helenius, A. Herpes simplex virus type 1 entry into host cells: reconstitution of capsid binding and uncoating at the nuclear pore complex in vitro. *Mol Cell Biol* **20**, 4922–31 (2000).
234. O’Neil, J. E. *et al.* Wide variations in herpes simplex virus type 1 inoculum dose and latency-associated transcript expression phenotype do not alter the establishment of latency in the rabbit eye model. *J Virol* **78**, 5038–44 (2004).

235. Pardo, J., Simon, M. M. & Froelich, C. J. Granzyme A is a proinflammatory protease. *Blood* **114**, 3968; author reply 3969–70 (2009).
236. Padeloup, D., Blondel, D., Isidro, A. L. & Rixon, F. J. Herpesvirus capsid association with the nuclear pore complex and viral DNA release involve the nucleoporin CAN/Nup214 and the capsid protein pUL25. *J Virol* **83**, 6610–23 (2009).
237. Pelosi, E., Mulamba, G. B. & Coen, D. M. Penciclovir and pathogenesis phenotypes of drug-resistant Herpes simplex virus mutants. *Antiviral Res* **37**, 17–28 (1998).
238. Peng, W. *et al.* The gene that encodes the herpes simplex virus type 1 latency-associated transcript influences the accumulation of transcripts (Bcl-x(L) and Bcl-x(S)) that encode apoptotic regulatory proteins. *J Virol* **77**, 10714–8 (2003).
239. Pepose, J. S. & Whittum-Hudson, J. A. An immunogenetic analysis of resistance to herpes simplex virus retinitis in inbred strains of mice. *Invest Ophthalmol Vis Sci* **28**, 1549–52 (1987).
240. Perng, G. C. *et al.* The latency-associated transcript gene of herpes simplex virus type 1 (HSV-1) is required for efficient in vivo spontaneous reactivation of HSV-1 from latency. *J Virol* **68**, 8045–55 (1994).
241. Perng, G. C. *et al.* Three herpes simplex virus type 1 latency-associated transcript mutants with distinct and asymmetric effects on virulence in mice compared with rabbits. *J Virol* **75**, 9018–28 (2001).
242. Perng, G. C., Ghiasi, H., Slanina, S. M., Nesburn, A. B. & Wechsler, S. L. The spontaneous reactivation function of the herpes simplex virus type 1 LAT gene resides completely within the first 1.5 kilobases of the 8.3-kilobase primary transcript. *J Virol* **70**, 976–84 (1996).
243. Perng, G. C. *et al.* Virus-induced neuronal apoptosis blocked by the herpes simplex virus latency-associated transcript. *Science* **287**, 1500–3 (2000).
244. Perng, G. C. *et al.* Herpes simplex virus type 1 serum neutralizing antibody titers increase during latency in rabbits latently infected with latency-associated transcript (LAT)-positive but not LAT-negative viruses. *J Virol* **73**, 9669–72 (1999).
245. Perng, G.-C. *et al.* A novel herpes simplex virus type 1 transcript (AL-RNA) antisense to the 5' end of the latency-associated transcript produces a protein in infected rabbits. *J Virol* **76**, 8003–10 (2002).

246. Perng, G.-C. *et al.* A gene capable of blocking apoptosis can substitute for the herpes simplex virus type 1 latency-associated transcript gene and restore wild-type reactivation levels. *J Virol* **76**, 1224–35 (2002).
247. Posavad, C. M., Koelle, D. M. & Corey, L. High frequency of CD8⁺ cytotoxic T-lymphocyte precursors specific for herpes simplex viruses in persons with genital herpes. *J Virol* **70**, 8165–8 (1996).
248. Posavad, C. M., Koelle, D. M. & Corey, L. Tipping the scales of herpes simplex virus reactivation: the important responses are local. *Nat Med* **4**, 381–2 (1998).
249. Posavad, C. M., Koelle, D. M., Shaughnessy, M. F. & Corey, L. Severe genital herpes infections in HIV-infected individuals with impaired herpes simplex virus-specific CD8⁺ cytotoxic T lymphocyte responses. *Proc Natl Acad Sci U S A* **94**, 10289–94 (1997).
250. Prober, C. G. *et al.* The management of pregnancies complicated by genital infections with herpes simplex virus. *Clin Infect Dis* **15**, 1031–8 (1992).
251. Qian, L., Vu, M. N., Carter, M. & Wilkinson, M. F. A spliced intron accumulates as a lariat in the nucleus of T cells. *Nucleic Acids Res* **20**, 5345–50 (1992).
252. Quinlan, M. P., Chen, L. B. & Knipe, D. M. The intranuclear location of a herpes simplex virus DNA-binding protein is determined by the status of viral DNA replication. *Cell* **36**, 857–68 (1984).
253. Randall, G., Lagunoff, M. & Roizman, B. The product of ORF O located within the domain of herpes simplex virus 1 genome transcribed during latent infection binds to and inhibits in vitro binding of infected cell protein 4 to its cognate DNA site. *Proc Natl Acad Sci U S A* **94**, 10379–84 (1997).
254. Randall, G., Lagunoff, M. & Roizman, B. Herpes simplex virus 1 open reading frames O and P are not necessary for establishment of latent infection in mice. *J Virol* **74**, 9019–27 (2000).
255. Randall, G. & Roizman, B. Transcription of the derepressed open reading frame P of herpes simplex virus 1 precludes the expression of the antisense gamma(1)34.5 gene and may account for the attenuation of the mutant virus. *J Virol* **71**, 7750–7 (1997).
256. Roberts, S. J., Owen, H. C. & Farquharson, C. Identification of a novel splice variant of the haloacid dehalogenase: PHOSPHO1. *Biochem Biophys Res Commun* **371**, 872–6 (2008).

257. Roche, S., Rey, F. A., Gaudin, Y. & Bressanelli, S. Structure of the prefusion form of the vesicular stomatitis virus glycoprotein G. *Science* **315**, 843–8 (2007).
258. Rock, D. L. *et al.* Detection of latency-related viral RNAs in trigeminal ganglia of rabbits latently infected with herpes simplex virus type 1. *J Virol* **61**, 3820–6 (1987).
259. Sabot, F. *et al.* Updating of transposable element annotations from large wheat genomic sequences reveals diverse activities and gene associations. *Mol Genet Genomics* **274**, 119–30 (2005).
260. SanMiguel, P., Gaut, B. S., Tikhonov, A., Nakajima, Y. & Bennetzen, J. L. The paleontology of intergene retrotransposons of maize. *Nat Genet* **20**, 43–5 (1998).
261. Sanna, P. P. *et al.* Protection of nude mice by passive immunization with a type-common human recombinant monoclonal antibody against HSV. *Virology* **215**, 101–6 (1996).
262. Sarisky, R. T. & Weber, P. C. Requirement for double-strand breaks but not for specific DNA sequences in herpes simplex virus type 1 genome isomerization events. *J Virol* **68**, 34–47 (1994).
263. Sawtell, N. M. Comprehensive quantification of herpes simplex virus latency at the single-cell level. *J Virol* **71**, 5423–31 (1997).
264. Sawtell, N. M. & Thompson, R. L. Herpes simplex virus type 1 latency-associated transcription unit promotes anatomical site-dependent establishment and reactivation from latency. *J Virol* **66**, 2157–69 (1992).
265. Schlegel, E. F. M. & Blaho, J. A. A conserved carboxy-terminal domain in the major tegument structural protein VP22 facilitates virion packaging of a chimeric protein during productive herpes simplex virus 1 infection. *Virology* **387**, 449–58 (2009).
266. Schlieker, C., Korbel, G. A., Kattenhorn, L. M. & Ploegh, H. L. A deubiquitinating activity is conserved in the large tegument protein of the herpesviridae. *J Virol* **79**, 15582–5 (2005).
267. Schomogyi, M., Wald, A. & Corey, L. Herpes simplex virus-2 infection. An emerging disease? *Infect Dis Clin North Am* **12**, 47–61 (1998).
268. Schulz, T. F. KSHV (HHV8) infection. *J Infect* **41**, 125–9 (2000).
269. Schumacher, A. J. *et al.* The HSV-1 exonuclease, UL12, stimulates recombination by a single strand annealing mechanism. *PLoS Pathog* **8**, e1002862 (2012).
270. Sears, A. E., Hukkanen, V., Labow, M. A., Levine, A. J. & Roizman, B. Expression of the herpes simplex virus 1 alpha transinducing factor (VP16) does not induce

- reactivation of latent virus or prevent the establishment of latency in mice. *J Virol* **65**, 2929–35 (1991).
271. Sedarati, F., Izumi, K. M., Wagner, E. K. & Stevens, J. G. Herpes simplex virus type 1 latency-associated transcription plays no role in establishment or maintenance of a latent infection in murine sensory neurons. *J Virol* **63**, 4455–8 (1989).
 272. Sedarati, F., Margolis, T. P. & Stevens, J. G. Latent infection can be established with drastically restricted transcription and replication of the HSV-1 genome. *Virology* **192**, 687–91 (1993).
 273. Shahin, V. *et al.* The genome of HSV-1 translocates through the nuclear pore as a condensed rod-like structure. *J Cell Sci* **119**, 23–30 (2006).
 274. Shapiro, J. *et al.* Isolation of pure lac operon DNA. *Nature* **224**, 768–74 (1969).
 275. Shen, W. *et al.* Two small RNAs encoded within the first 1.5 kilobases of the herpes simplex virus type 1 latency-associated transcript can inhibit productive infection and cooperate to inhibit apoptosis. *J Virol* **83**, 9131–9 (2009).
 276. Sheridan, J. F., Beck, M., Smith, C. C. & Aurelian, L. Reactivation of herpes simplex virus is associated with production of a low molecular weight factor that inhibits lymphokine activity in vitro. *J Immunol* **138**, 1234–9 (1987).
 277. Shieh, M. T., WuDunn, D., Montgomery, R. I., Esko, J. D. & Spear, P. G. Cell surface receptors for herpes simplex virus are heparan sulfate proteoglycans. *J Cell Biol* **116**, 1273–81 (1992).
 278. Shimeld, C., Tullo, A. B., Hill, T. J., Blyth, W. A. & Easty, D. L. Spread of herpes simplex virus and distribution of latent infection after intraocular infection of the mouse. *Arch Virol* **85**, 175–87 (1985).
 279. Shukla, D. *et al.* A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell* **99**, 13–22 (1999).
 280. Siegel, J. D. & McCracken, G. H. J. Sepsis neonatorum. *N Engl J Med* **304**, 642–7 (1981).
 281. Skepper, J. N., Whiteley, A., Browne, H. & Minson, A. Herpes simplex virus nucleocapsids mature to progeny virions by an envelopment --> deenvelopment --> reenvelopment pathway. *J Virol* **75**, 5697–702 (2001).
 282. Slotkin, R. K. & Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* **8**, 272–85 (2007).

283. Smith, J. S. & Robinson, N. J. Age-specific prevalence of infection with herpes simplex virus types 2 and 1: a global review. *J Infect Dis* **186 Suppl 1**, S3–28 (2002).
284. Snyder, A., Wisner, T. W. & Johnson, D. C. Herpes simplex virus capsids are transported in neuronal axons without an envelope containing the viral glycoproteins. *J Virol* **80**, 11165–77 (2006).
285. Sodeik, B., Ebersold, M. W. & Helenius, A. Microtubule-mediated transport of incoming herpes simplex virus 1 capsids to the nucleus. *J Cell Biol* **136**, 1007–21 (1997).
286. Spector, F. C. *et al.* Evaluation of a live attenuated recombinant virus RAV 9395 as a herpes simplex virus type 2 vaccine in guinea pigs. *J Infect Dis* **177**, 1143–54 (1998).
287. Spruance, S. L., Kriesel, J. D., Evans, T. G. & McKeough, M. B. Susceptibility to herpes labialis following multiple experimental exposures to ultraviolet radiation. *Antiviral Res* **28**, 57–67 (1995).
288. Spruance, S. L., Tying, S. K., DeGregorio, B., Miller, C. & Beutner, K. A large-scale, placebo-controlled, dose-ranging trial of peroral valaciclovir for episodic treatment of recurrent herpes genitalis. Valaciclovir HSV Study Group. *Arch Intern Med* **156**, 1729–35 (1996).
289. Stackpole, C. W. Herpes-type virus of the frog renal adenocarcinoma. I. Virus development in tumor transplants maintained at low temperature. *J Virol* **4**, 75–93 (1969).
290. Stamm, W. E. *et al.* The association between genital ulcer disease and acquisition of HIV infection in homosexual men. *JAMA* **260**, 1429–33 (1988).
291. Stanberry, L. R. Asymptomatic herpes simplex virus shedding and Russian roulette. *Clin Infect Dis* **30**, 268–9 (2000).
292. Stanberry, L. R. *et al.* Prospects for control of herpes simplex virus disease through immunization. *Clin Infect Dis* **30**, 549–66 (2000).
293. Steiner, I. *et al.* Herpes simplex virus type 1 latency-associated transcripts are evidently not essential for latent infection. *EMBO J* **8**, 505–11 (1989).
294. Steiner, I., Spivack, J. G., O’Boyle, D. R. 2nd, Lavi, E. & Fraser, N. W. Latent herpes simplex virus type 1 transcription in human trigeminal ganglia. *J Virol* **62**, 3493–6 (1988).
295. Stevens, J. G. Defining herpes simplex genes involved in neurovirulence and neuroinvasiveness. *Curr Eye Res* **6**, 63–7 (1987).

296. Stevens, J. G., Wagner, E. K., Devi-Rao, G. B., Cook, M. L. & Feldman, L. T. RNA complementary to a herpesvirus alpha gene mRNA is prominent in latently infected neurons. *Science* **235**, 1056–9 (1987).
297. Sullender, W. M. *et al.* Humoral and cell-mediated immunity in neonates with herpes simplex virus infection. *J Infect Dis* **155**, 28–37 (1987).
298. Swergold, G. D. Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol Cell Biol* **10**, 6718–29 (1990).
299. Tal-Singer, R. *et al.* Gene expression during reactivation of herpes simplex virus type 1 from latency in the peripheral nervous system is different from that during lytic infection of tissue cultures. *J Virol* **71**, 5268–76 (1997).
300. Tam, S. M. *et al.* Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR. *Theor Appl Genet* **110**, 819–31 (2005).
301. Tanaka, M., Nishiyama, Y., Sata, T. & Kawaguchi, Y. The role of protein kinase activity expressed by the UL13 gene of herpes simplex virus 1: the activity is not essential for optimal expression of UL41 and ICP0. *Virology* **341**, 301–12 (2005).
302. Tang, S. *et al.* An acutely and latently expressed herpes simplex virus 2 viral microRNA inhibits expression of ICP34.5, a viral neurovirulence factor. *Proc Natl Acad Sci U S A* **105**, 10931–6 (2008).
303. Tang, S., Patel, A. & Krause, P. R. Novel less-abundant viral microRNAs encoded by herpes simplex virus 2 latency-associated transcript and their roles in regulating ICP34.5 and ICP0 mRNAs. *J Virol* **83**, 1433–42 (2009).
304. Taylor, T. J., Brockman, M. A., McNamee, E. E. & Knipe, D. M. Herpes simplex virus. *Front Biosci* **7**, d752–64 (2002).
305. TAYLOR, A. L. BACTERIOPHAGE-INDUCED MUTATION IN ESCHERICHIA COLI. *Proc Natl Acad Sci U S A* **50**, 1043–51 (1963).
306. Telzak, E. E. *et al.* HIV-1 seroconversion in patients with and without genital ulcer disease. A prospective study. *Ann Intern Med* **119**, 1181–6 (1993).
307. Tenser, R. B., Edris, W. A., Gaydos, A. & Hay, K. A. Secondary herpes simplex virus latent infection in transplanted ganglia. *J Virol* **68**, 7212–20 (1994).
308. Thomas, D. L., Lock, M., Zabolotny, J. M., Mohan, B. R. & Fraser, N. W. The 2-kilobase intron of the herpes simplex virus type 1 latency-associated transcript has a

- half-life of approximately 24 hours in SY5Y and COS-1 cells. *J Virol* **76**, 532–40 (2002).
309. Thompson, R. L. & Sawtell, N. M. The herpes simplex virus type 1 latency-associated transcript gene regulates the establishment of latency. *J Virol* **71**, 5432–40 (1997).
 310. Thompson, R. L. & Sawtell, N. M. Herpes simplex virus type 1 latency-associated transcript gene promotes neuronal survival. *J Virol* **75**, 6660–75 (2001).
 311. Todeschini, A.-L., Morillon, A., Springer, M. & Lesage, P. Severe adenine starvation activates Ty1 transcription and retrotransposition in *Saccharomyces cerevisiae*. *Mol Cell Biol* **25**, 7459–72 (2005).
 312. Tookey, P. & Peckham, C. S. Neonatal herpes simplex virus infection in the British Isles. *Paediatr Perinat Epidemiol* **10**, 432–42 (1996).
 313. Toor, N., Hausner, G. & Zimmerly, S. Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA* **7**, 1142–52 (2001).
 314. Ugolini, G. Transneuronal transfer of herpes simplex virus type 1 (HSV 1) from mixed limb nerves to the CNS. I. Sequence of transfer from sensory, motor, and sympathetic nerve fibres to the spinal cord. *J Comp Neurol* **326**, 527–48 (1992).
 315. Ugolini, G., Kuypers, H. G. & Simmons, A. Retrograde transneuronal transfer of herpes simplex virus type 1 (HSV 1) from motoneurons. *Brain Res* **422**, 242–56 (1987).
 316. Umbach, J. L. *et al.* MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. *Nature* **454**, 780–3 (2008).
 317. Uprichard, S. L. & Knipe, D. M. Assembly of herpes simplex virus replication proteins at two distinct intranuclear sites. *Virology* **229**, 113–25 (1997).
 318. Varmuza, S. L. & Smiley, J. R. Signals for site-specific cleavage of HSV DNA: maturation involves two separate cleavage events at sites distal to the recognition sequences. *Cell* **41**, 793–802 (1985).
 319. Verjans, G. M. G. M. *et al.* Selective retention of herpes simplex virus-specific T cells in latently infected human trigeminal ganglia. *Proc Natl Acad Sci U S A* **104**, 3496–501 (2007).
 320. Wagner, E. K. & Bloom, D. C. Experimental investigation of herpes simplex virus latency. *Clin Microbiol Rev* **10**, 419–43 (1997).
 321. Wagner, E. K., Swanstrom, R. I., Rice, M., Howell, L. & Lane, J. Variation in the molecular size of the DNA from closely related strains of type I herpes simplex virus. *Biochim Biophys Acta* **435**, 192–205 (1976).

322. Wald, A. *et al.* Reactivation of genital herpes simplex virus type 2 infection in asymptomatic seropositive persons. *N Engl J Med* **342**, 844–50 (2000).
323. Wald, A. & Link, K. Risk of human immunodeficiency virus infection in herpes simplex virus type 2-seropositive persons: a meta-analysis. *J Infect Dis* **185**, 45–52 (2002).
324. Wang, H. *et al.* SVA elements: a hominid-specific retroposon family. *J Mol Biol* **354**, 994–1007 (2005).
325. Wang, K., Lau, T. Y., Morales, M., Mont, E. K. & Straus, S. E. Laser-capture microdissection: refining estimates of the quantity and distribution of latent herpes simplex virus 1 and varicella-zoster virus DNA in human trigeminal Ganglia at the single-cell level. *J Virol* **79**, 14079–87 (2005).
326. Wang, Q.-Y. *et al.* Herpesviral latency-associated transcript gene promotes assembly of heterochromatin on viral lytic-gene promoters in latent infection. *Proc Natl Acad Sci U S A* **102**, 16055–9 (2005).
327. Wang, Y., Yu, L. & Geller, A. I. Diverse stabilities of expression in the rat brain from different cellular promoters in a helper virus-free herpes simplex virus type 1 vector system. *Hum Gene Ther* **10**, 1763–71 (1999).
328. Ward, P. L., Campadelli-Fiume, G., Avitabile, E. & Roizman, B. Localization and putative function of the UL20 membrane protein in cells infected with herpes simplex virus 1. *J Virol* **68**, 7406–17 (1994).
329. Ward, P. L. & Roizman, B. Herpes simplex genes: the blueprint of a successful human pathogen. *Trends Genet* **10**, 267–74 (1994).
330. Wasserheit, J. N. Epidemiological synergy. Interrelationships between human immunodeficiency virus infection and other sexually transmitted diseases. *Sex Transm Dis* **19**, 61–77 (1992).
331. WATANABE, T. & FUKASAWA, T. Episome-mediated transfer of drug resistance in Enterobacteriaceae IV. Interactions between resistance transfer factor and F-factor in *Escherichia coli* K-12. *J Bacteriol* **83**, 727–35 (1962).
332. Watson, R. J. & Clements, J. B. A herpes simplex virus type 1 function continuously required for early and late virus RNA synthesis. *Nature* **285**, 329–30 (1980).
333. Wei, W. *et al.* Human L1 retrotransposition: cis preference versus trans complementation. *Mol Cell Biol* **21**, 1429–39 (2001).
334. Wessler, S. R. Turned on by stress. Plant retrotransposons. *Curr Biol* **6**, 959–61 (1996).

335. Weth, O. & Renkawitz, R. CTCF function is modulated by neighboring DNA binding factors. *Biochem Cell Biol* **89**, 459–68 (2011).
336. Whitley, R. J. & Roizman, B. Herpes simplex virus infections. *Lancet* **357**, 1513–8 (2001).
337. Wicker, T. & Keller, B. Genome-wide comparative analysis of copia retrotransposons in Triticeae, rice, and Arabidopsis reveals conserved ancient evolutionary lineages and distinct dynamics of individual copia families. *Genome Res* **17**, 1072–81 (2007).
338. Wilcox, C. L. & Johnson, E. M. J. Characterization of nerve growth factor-dependent herpes simplex virus latency in neurons in vitro. *J Virol* **62**, 393–9 (1988).
339. Wild, P. *et al.* Impairment of nuclear pores in bovine herpesvirus 1-infected MDBK cells. *J Virol* **79**, 1071–83 (2005).
340. Wolfstein, A. *et al.* The inner tegument promotes herpes simplex virus capsid motility along microtubules in vitro. *Traffic* **7**, 227–37 (2006).
341. Wu, T. T., Su, Y. H., Block, T. M. & Taylor, J. M. Evidence that two latency-associated transcripts of herpes simplex virus type 1 are nonlinear. *J Virol* **70**, 5962–7 (1996).
342. Wu, T. T., Su, Y. H., Block, T. M. & Taylor, J. M. Atypical splicing of the latency-associated transcripts of herpes simplex type 1. *Virology* **243**, 140–9 (1998).
343. WuDunn, D. & Spear, P. G. Initial interaction of herpes simplex virus with cells is binding to heparan sulfate. *J Virol* **63**, 52–8 (1989).
344. Yamada, M. *et al.* Interaction between neuronal intranuclear inclusions and promyelocytic leukemia protein nuclear and coiled bodies in CAG repeat diseases. *Am J Pathol* **159**, 1785–95 (2001).
345. Yang, L., Voytek, C. C. & Margolis, T. P. Immunohistochemical analysis of primary sensory neurons latently infected with herpes simplex virus type 1. *J Virol* **74**, 209–17 (2000).
346. Yao, S. & Chen, L. Reviving exhausted T lymphocytes during chronic virus infection by B7-H1 blockade. *Trends Mol Med* **12**, 244–6 (2006).
347. Yuhasz, S. A. & Stevens, J. G. Glycoprotein B is a specific determinant of herpes simplex virus type 1 neuroinvasiveness. *J Virol* **67**, 5948–54 (1993).
348. Zabolotny, J. M., Krummenacher, C. & Fraser, N. W. The herpes simplex virus type 1 2.0-kilobase latency-associated transcript is a stable intron which branches at a guanosine. *J Virol* **71**, 4199–208 (1997).

349. Zimmerly, S. *et al.* A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell* **83**, 529–38 (1995).
350. Zimmerly, S., Guo, H., Perlman, P. S. & Lambowitz, A. M. Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell* **82**, 545–54 (1995).
351. Zimmerly, S., Hausner, G. & Xc, W. Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res* **29**, 1238–50 (2001).
352. Zuker, C., Cappello, J., Chisholm, R. L. & Lodish, H. F. A repetitive Dictyostelium gene family that is induced during differentiation and by heat shock. *Cell* **34**, 997–1005 (1983).

Annexe 1

BIOLOGICAL FEATURES OF HERPES SIMPLEX VIRUS TYPE 1 LATENCY IN MICE ACCORDING TO THE VIRAL STRAIN, THE ROUTE OF INOCULATION AND THE TYPE OF NEURONES

Sophie CREPIN-CAVALLERO¹, Nicolas HUOT¹, Laetitia FRANCELLE¹, Frédéric CATEZ³, Patrice NORDMANN², Patrick LOMONTE³, Thierry NAAS², Marc LABETOULLE^{1,4*}

(1) Laboratoire de Virologie Moléculaire et Structurale, Centre National de la Recherche Scientifique, 91198 Gif sur Yvette, France ; (2) Service de Bactériologie-Virologie-Hygiène, INSERM U914 "Emerging Resistance to Antibiotics", Hôpital de Bicêtre, Assistance Publique / Hôpitaux de Paris, Faculté de Médecine and Université Paris Sud, Kremlin-Bicêtre, France ; (3) Centre de Génétique et Physiologie Moléculaire et Cellulaire CNRS, UMR5534, Université de Lyon 1, F-69000 Lyon, France, Laboratoire d'excellence, Labex DEVweCAN, F-69000 Lyon, France ; (4) Service d'Ophtalmologie, Hôpital de Bicêtre, Assistance Publique-Hôpitaux de Paris, Assistance Publique / Hôpitaux de Paris, Faculté de Médecine and Université Paris Sud, Kremlin-Bicêtre, France

* Corresponding author : Marc Labetoulle & Thierry Naas

Mailing address : Service d'Ophtalmologie, Hôpital de Bicêtre, Assistance Publique-Hôpitaux de Paris, Université Paris Sud, 94275 Le Kremlin Bicêtre, France

Phone : 33 1 69 82 38 40

Fax : 33 1 69 82 43 08

E-mail : marc.labetoulle@bct.aphp.fr

KeyWords: HSV1, Latency, Virulence, LAT, ICP4

Running title : Biological patterns of HSV1 infection

Figures: 10

Appendixes: 7 tables

ABSTRACT.

Following acute infection, the herpes simplex virus type 1 (HSV1) becomes latent in neuronal tissues through mechanisms that are not yet completely understood. Using the oro-ocular murine model of herpetic infection with the wild type strain SC16 of HSV1, we previously showed that both HSV1 *Latency-Associated Transcripts* (LATs) and non-spliced transcripts of the RL2 immediate-early (IE) gene encoding the ICP0 protein accumulate in the trigeminal ganglia (TG) during the onset of latent infection while they decrease in several other neurological tissues. In the present study we show that, 1 month post-inoculation (early stage of latent infection), some IE gene transcripts encoding the ICP4 protein were still present, especially in the TG, while early and late viral genes transcripts were not detected, as expected. ICP4-encoding transcripts were however in much lower amounts as compared to acute infection, and were not detected anymore at 3 months post-inoculation (late stage of latent infection). In contrast, such progress in the viral gene expression was only partially observed when inoculation was performed using a corneal scarification model. In particular, the classical pattern of HSV1 latency, i.e. no expression of early and late viral genes, was not fully obtained, especially with the KOS strain of HSV1. These results highlight that the behaviour of the virus in the nervous system depends on the method and site of inoculation, the type of infected neurons and the viral strain. All these parameters are likely to influence the establishment of latent infection.

INTRODUCTION

The Herpes Simplex Virus type 1 (HSV1) is a common human pathogen that affects the face and the upper body with recurrent infections, mostly in tissues innervated by sensitive neurons (i.e. lips and cornea) and more rarely in the brain, the iris or the retina. It is consequently a major cause of blindness and severe neurological disorders in human {Labetoulle, 2005 #3245}. After a primary infection that is almost always located into the oral mucosa {Liesegang, 1989 #1649}, the virus moves by axonal transport to neuronal cell bodies {Smith, 2001 #4722}. Within the neuron nuclei, HSV1 establishes a latent infection. Spontaneously or stimulated by various triggering factors (stress, ultraviolet radiations, immuno-depression), the virus can reactivate and migrate to peripheral tissues by anterograde axonal transport, leading to clinical or subclinical herpetic recurrences {Roizman, 2001 #673}{Curanovic, 2009 #5523}{Negatsch, 2010 #5563}.

During the lytic phase of infection, the synthesis of viral proteins progresses in a regulated cascade, successively immediate-early (IE), early (E) and late (L) proteins, resulting in the formation of new infectious particles {Roizman, 2001 #673}. In contrast, during latency, no viral protein is produced. Latent viral DNA is circularized, and tightly associated with histones {Knipe, 2008 #5938}{Bloom, 2010 #5264}. Moreover, latent genomes are not randomly distributed in the nucleus of infected neurons, and show a preferential association with promyelocytic leukemia nuclear domains (PML-NBs) and centromeres {Catez, 2012 #5937}. Latent HSV1 is essentially transcriptionnaly repressed with the exception of the repeat domains that flanks the unique long region of the HSV1 genome. These domains contains a locus allowing the synthesis of *Latency-Associated transcripts* (LATs), which classically are considered as the only viral transcript that strongly accumulate during HSV1 latency {Rock, 1983 #1751;Stevens, 1987 #2154;Steiner, 1988 #1827;Bloom, 2010 #5264}{Deshmane, 1989 #1757}. Major LAT species are 1.5-kilobases (kb) and 2-kb stable non polyadenylated introns, spliced from a primary 8.3-kb transcript {Farrell, 1991 #2158;Devi-Rao, 1991 #3360}{Wu, 1996 #2132;Wu, 1998 #2133;Ng, 2004

#3335;Krummenacher, 1997 #5888}. The exact biological pathways in which LATs are involved are not fully understood. Several experiments suggest that they promote the establishment of latent viral infection and regulate viral reactivation **{Steiner, 1989 #1825;Leib, 1989 #4766;Sawtell, 1992 #2167;Peng, 1994 #1799}**, the repression of viral gene expression in sensory neurons during latent infection **{Garber, 1997 #2207;Mador, 1998 #2094;Giordani, 2008 #5087}{Chen, 1997 #2359}**, the protection of infected neurons from death by apoptosis inhibition **{Peng, 2000 #2155;Ahmed, 2002 #2515;Branco, 2005 #3326}**, and the interference with immunity **{Peng, 2005 #3329;Branco, 2005 #3326;Verjans, 2007 #5757;Knickelbein, 2008 #5160;Jiang, 2011 #5596}**. Some of these properties seem to involve LATs-encoded miRNAs **{Umbach, 2008 #5804;Shen, 2009 #5074;Peng, 2008 #5078}** and/or chromatin remodeling **{Kubat, 2004 #3296;Wang, 2005 #4363;Amelio, 2006 #5699;Bloom, 2010 #5264}**.

A mouse model of oro-ocular (OO) HSV1 infection that mimics the natural history of HSV1 infection in human was previously developed **{Labetoulle, 2000 #2047}{Labetoulle, 2003 #2689}**. In this model, mice develop an acute infection in the iris and the cornea (6 to 8 days post-inoculation), following viral inoculation in the lip that spontaneously resolves and which is followed by a latent infection in the nervous system. Latency was demonstrated by the expression of LATs in absence of production of infectious particles at 28 days post-inoculation (dpi) in the trigeminal ganglia (TG), the superior cervical ganglia (SCG), the spinal cord and the hypothalamus (respectively peripheral sensory or sympathetic, and central neurons) **{Labetoulle, 2000 #2047}{Labetoulle, 2003 #2689}**. Using the same OO model, we showed that transcripts of the ICP0 gene of HSV1, which plays a key role in viral reactivation **{Cai, 1992 #3290}{Everett, 2000 #2315}{Halford, 2001 #2350}**, were also expressed at 28 dpi, with however different levels of expression according to the site of HSV1 latency. Particularly, the TG was characterized by a significant accumulation of both LATs and non-spliced ICP0 transcripts, suggesting that the balance between LATs and ICP0 transcripts may be involved in the maintenance of HSV1 latency in TG

neurons. In the same tissues, neither transcripts for thymidine kinase (TK, an early gene) nor UL18 transcripts (a late gene, encoding the capsid protein VP23) were detected, supporting that HSV1 productive cycle was blocked {Maillet, 2006 #3281}.

The presence of both ICP0 transcripts had also been observed at 28 dpi in the corneal scarification (CS) murine model, which is used in the vast majority of studies on HSV1 latency. However, in these studies, transcripts encoding ICP4 (IE viral protein), TK (E protein) and even UL15 (L protein) were also evidenced {Chen, 2002 #2811}{Kramer, 1995 #2157} {Thompson, 2006 #4364}, which raised the question of the complete extinction of HSV1 replication cycle in tested tissues. Besides the differences in the method of inoculation (corneal scarification *versus* lip injection), these experiments also differed by the HSV1 strain used (KOS *versus* SC16, respectively).

To assess the differences in herpetic pathogenesis induced by various conditions of infection, we compared both routes of inoculation (i.e. the mouse OO and CS models) using the two strains of HSV1 strains (SC16 and KOS). We observed significant differences in the amounts of viral genomes and the expression of viral genes according to the viral strain, the route of inoculation and the type of neurological tissues tested at the late stages of infection. These results suggest that investigational conditions have significant implications for experiments aiming at studying HSV1 latency and reactivation.

MATERIAL AND METHODS

Virus strains,, viral inoculation and euthanasia. Wild-type HSV1 strain SC16 {Hill, 1975 #54} and KOS {Smith, 1964 #4744} were used. Virus were propagated and titrated as previously described {Labetoulle, 2000 #1297}.

Six-week-old female BALB/c mice (Janvier Breeding, Le Genest Saint Ile, France) were used for the experiments. Animals received unrestricted access to food and water. All experimental procedures involving animals were performed under general anesthesia, conformed to ethical standards such as Association for Research in Vision and Ophthalmology (ARVO) Statement for the Use of Animals in Research, and were approved by a local ethical committee.

For the oro-ocular model, 1µl suspension of either SC16 or KOS strain was inoculated into the left upper lip. For the corneal scarification model, the viral suspension was inoculated on the scarified cornea of the left eye. Mice were observed daily for clinical signs of ocular infection from days 0 to 100 dpi, as previously described {Labetoulle, 2000 #2047}{Labetoulle, 2003 #2689}. Mice that showed clinical signs of ocular disease in the days following inoculation were randomly sacrificed at 6 dpi (during the acute phase of the ocular disease), at 28 dpi (during the expected latent phase of infection), and at 100 dpi (late stages of latent infection). As controls, 5 mock- infected mice in each model were sacrificed at 6, 28 and 100 dpi.

Before tissue dissection, intra cardiac perfusion of the mice was performed with phosphate-buffered saline (PBS) for RNA or DNA extraction {Maillet, 2006 #3281} or successively with PBS, 4% paraformaldehyde in PBS, and 20% saccharose in PBS for tissue cryosection {Labetoulle, 2000 #1297}.

Cryostat sections. The spinal cord and the whole head (containing the SCG and the TG) were prepared as previously described, and 10 µm frontal cryosections were collected in three parallel series {Labetoulle, 2003 #2689}.

In situ hybridization (ISH). A digoxigenin-labeled riboprobe was produced after HindIII

digestion of the pSLAT2 plasmid, provided by S. Efstathiou {Arthur, 1993 #1301} and *in vitro* transcription with T7 RNA polymerase, using the RNA Dig transcription kit (Roche, Meylan, FRANCE), according to manufacturer's instructions. ISH was performed as previously described {Labetoulle, 2003 #2689}.

Detection of HSV1 antigens. For the study on the propagation of KOS strain in the corneal scarification model, 12 mice were scarified at 2, 4 and 6 dpi. After dissection and preparation of the tissues, HSV1 positive cells were detected as previously described {Labetoulle, 2000 #1297}.

DNA and RNA extraction. For RT-PCR and PCR studies, 39 mice that demonstrated clinical eye infection were used. For each of the 3 combinations of viral strain and method of inoculation, 5 animals were sacrificed at 6 dpi, 5 animals at 28 dpi, and 3 animals at 100 dpi. 15 additional mice were used as control, sacrificed at 6, 28 and 100 days following mock-inoculation.

Spinal cord, TG, SCG and hypothalamus were quickly dissected, incubated overnight at 4°C in RNA later (QIAGEN, Courtabeuf, France) and stored at -80°C. Frozen tissues were homogenized using the QIAshredder columns (QIAGEN) and 1/5 of each homogenized tissue was used for DNA extraction by phenol-chloroform protocol, whereas the remaining 4/5 were used for total RNA extraction using the Rneasy Mini extraction kit (Qiagen) as previously described {Maillet, 2006 #3281}.

Production of DNA and RNA standards for quantification. The gG gene of HSV1 and the cellular GADPH (Glyceraldehyde 3-phosphate dehydrogenase) gene were amplified by PCR, and the 2kb LATs, ICP4, TK, UL18 and GADPH transcripts were amplified by RT-PCR. Products of amplification were then cloned into pCRII-TOPO vector (Invitrogen, Cergy Pontoise, France). The RNA standards were produced following digestion of the vector with EcoRV (Ozyme, Saint-Quentin-en-Yvelines, France) and *in vitro* transcription using SP6 polymerase (Roche). The amounts of DNA and RNA standards were evaluated by optical density at 260 nm.

Quantitative PCR and RT-PCR. Quantification of DNA copies and transcripts was performed using a Light Cycler 2.0 machine (Roche) with Taqman technology, with the exception of GADPH gene and GADPH transcripts that were quantified using SYBR Green technology, as previously described {Maillet, 2006 #3281}. Respective primers are given in Table 1.

For 2-kb LATs, 2.5 µl of extracted RNA were amplified with the Quantitect Probe RT-PCR kit (Qiagen), using 10µM of each primers and 4µM of the LAT probe. Reverse transcription was performed at 60°C for 20 min. After denaturation (95°C, 15 min), amplification was performed using 45 cycles of PCR (20 s at 95°C, 1 min at 60°C).

For GADPH gene and transcripts, 1 µl of extracted DNA or RNA was amplified with the Light Cycler DNA Master SYBR Green I kit or with the Light Cycler RNA Master SYBR Green I kit (Roche) at an annealing temperature of 50°C, using 6µM of primers.

For gG gene, 2µl of extracted DNA were amplified with the Quantitect Probe PCR kit (Qiagen) using 10µM of primers and 4µM of gG probe. After denaturation (95°C, 15 min), amplification was performed using 40 cycles of PCR (20s at 95°C, 1 min at 60°C).

For quantification of ICP4, TK and UL18 transcripts, we performed quantitative nested RT-qPCR. One µl of RNA was reverse transcribed at 45°C for 45 min and subsequently amplified using 10µM of the external pair of primers using the AccessQuick RT-PCR system (Promega, Paris, France) and the following conditions: denaturation (95°C, 2 min), 20 cycles of PCR (45s at 95°C, 1 min at 55°C, 1 min at 72°C) before final extension (5 min at 72°C). Finally, 1µl of the RT-PCR was further amplified using Quantitect Probe PCR kit (Qiagen) using 10µM of the internal pair of primers and 4µl of probe. After denaturation (95°C, 15 min) amplification was performed using 40 cycles of PCR (20s at 95°C, 1 min at 60°C).

Flow cytometry analysis (FACS). After the sacrifice and exsanguination of animals with intracardiac perfusion of PBS {Maillet, 2006 #3281}, TGs from mice at 28 dpi were rapidly dissected and incubated in PBS containing 40 units/ml papain (Worthington Biochemical

Corporation), 4 mg/ml collagenase type II (CLS2) (Worthington Biochemical Corporation) and 4,6 mg/ml dispase type II (MB, cat.no. 16589), at 37°C for 20 min with agitation every 5 min. Dissociated cells were pelleted (400 x G for 4 min), washed twice in PBS, resuspended in 500µl of ice-cold PBS and fixed in 4% PFA. Three aliquots of 2.5×10^5 cells were incubated with antibodies specific for either HSV1 antigens (polyclonal rabbit Anti-HSV1, Dako®, Trappes, France), macrophages (mouse anti CD11b, eBioscience®, Paris, France) or lymphocytes (mouse anti-CD8a, eBioscience), diluted at 1/250 in FACS buffer (PBS 1x, bovine serum albumin 1%). After three washes in PBS, cells were incubated for 1 hour with an anti-rat IgG, texas red-conjugated antibody (Santa Cruz Biotechnology Inc., Heidelberg, Germany) for anti-lymphocytes and anti-macrophages antibodies, or anti-rabbit IgG, fluorescein conjugated (Dako®). Cells were analyzed by flow cytometry (Cytometer ELITE ESP, Beckman-Coulter, Paris, France). Thresholds were set-up to include viable cells only, as assessed by forward scatter data, and to include a range of fluorescence representing less than 1% of the fluorescence of control cells.

RESULTS

Morbidity and mortality induced by SC16 and KOS strains according to the inoculation site. In the oro ocular (OO) model, the injection of 10^6 PFU of SC16 strain of HSV1 into the lip resulted in a 60% rate of morbidity and a 45% of mortality within the first 15 dpi, similarly to previously published results with the same animal model {Labetoulle, 2000 #2047}{Labetoulle, 2003 #2689} (**Figure 1C, 2C**). In contrast, even when using the maximum load of infectious particles (i.e. 10^6 PFU in 1 μ l), no death or clinical signs of infection were observed with the KOS strain (**Figure 1D, 2D**).

In the corneal scarification (CS) model, inoculum higher or equal to 10^3 PFU of the SC16 strain of HSV1 resulted in a 100% rate of mortality at 10 dpi, while inoculum of 10^1 and 10^2 PFU provided survival rates of 80% and 40%, respectively (**Figure 1A**). To allow a survival rate that was compatible with longer follow-up studies (i.e. 28 dpi to compare with the OO model), we estimated 70 PFU as the optimal SC16 inoculum. To validate this inoculum, 30 mice were inoculated in the cornea with 70 PFU of SC16 strain, which resulted in 60% of morbidity and 45% of mortality (i.e. similar rates to those observed with the inoculation of 10^6 PFU into the OO model, data not shown). In contrast, with the KOS strain of HSV1, no mortality was observed whatever the size of inoculum (**Figure 1B**), and 10^6 PFU were required to induce ocular morbidity similar to that observed with 70 PFU of the SC16 strain (**Figure 2B**).

In addition to the differences in the size of inoculum required to induce a similar rate of ocular disease, clinical differences in affected eyes were also observed. Persistent inflammation and numerous neo-vessels in the cornea were still observed at 28 dpi in the CS model, especially with the KOS strain. In contrary, in the OO model, dilation of vessels and corneal clouding dramatically reduced from 14 dpi, leading to almost normal cornea at 28 dpi (**Figure 3**).

In order to compare viral strains and routes of inoculation with similar rates of morbidity and mortality, the following inoculums were chosen for the molecular analysis of the viral gene

expression: 10^6 PFU of KOS or 70 PFU of SC16 in the CS model, and 10^6 PFU of SC16 in the OO model. Since no comparison was possible using 10^6 PFU of KOS strain in the OO model (no clinical signs of ocular infection), this combination was not further analysed.

Viral propagation pathways in the oro-ocular (OO) model and the corneal scarification (CS) models. The propagation pathways of the SC16 strain of HSV1 in the OO model have been previously described {Labetoulle, 2000 #2047}. Briefly, from the left upper lip, the virus propagated simultaneously via the sympathetic neurons to the left SCG (immunostained at 2 dpi) and via the sensitive neurons to the left TG (immunostained at 4 dpi). From these first order neuronal connections, the virus propagated to the brainstem and the spinal cord, and finally to the right SCG and TG (immunostained at 6 dpi, but to a much fewer extent) (**Figure 4A**).

Similarly, we traced the propagation of the KOS strain after corneal scarification (CS) using immunochemical staining on tissues of mice inoculated for 2, 4 and 6 dpi {Labetoulle, 2000 #2047}. Corneal and iris infection were detected at 2 and 4 days, respectively (data not shown). The infection of the iris was probably the consequence of local transfer of virus into the ciliary ganglion, from sensitive fibers connected to the cornea to autonomic fibers connected to the iris, both of them being present within this ganglion. A few infected neurons were observed in the left TG at 2 dpi, with an increasing number at 4 dpi and then a decrease at 6 dpi. In the SCG and spinal cord, only rare infected cells were detected at 6 dpi, while no immunochemical staining was observed in the brain (schematically represented in **Figure 4B**).

Detection of LATs on histological sections using in situ hybridization (ISH) at 28 dpi. 17 mice that had survived inoculation into the lip (10^6 PFU of SC16 strain) or into the cornea (10^6 PFU of KOS strain or 70 PFU of SC16 strain) were sacrificed at 28 dpi to detect LAT-positive TG neurons using ISH assays {Labetoulle, 2003 #2689}. As a consequence of the side of inoculation, LAT-positive neurons were more frequent in the left TG than in the right one (**Table 2**). They were also more abundant in the OO model (SC16 strain) than in the CS model with the KOS strain (38.8 versus

9.2 in the left TG respectively), despite a similar size of inoculum (10^6 PFU) (**Table 2**). Using the same ISH method, no LAT-positive neuron was detected in TGs following SC16 inoculation in the CS model (**Figure 1**), but few positive neurons were detected using an ISH with tyramide-biotin amplification system (method described in {Maillet, 2006 #3281}). This suggested that, at the single neuron level, the expression of LATs may differ according to the animal model and the viral strain. This also could be the consequence of the very low inoculum (70 PFU) used for SC16 in the OO model.

As controls, no positive hybridization signal was observed on sections from mock-infected mice, and ISH signal persisted on sections of infected animals after DNase treatment. In addition, no antigen-positive neurons, as assessed by immunoassay using polyclonal anti-HSV1 antibodies (mainly anti-glycoproteins), was observed in histological sections adjacent to those positive for LATs at 28 dpi, suggesting that HSV1 did not reactivate in the ISH-positive tissues.

Quantification of viral transcripts and genomes at 6, 28 and 100 dpi. All results of quantitative PCR (or RT-PCR) were normalized with the amounts of GAPDH gene copies (or transcripts) in the corresponding tissue extracts. To reduce the intra-experimental variability, all samples have been tested twice. A variability between mice was however observed, inherent to the experimental nature of the work. The average number of genome or transcripts copies, and the rate of positive structures, are represented in figures 5 to 9 and raw data as supplemental tables A1-A5.

Viral transcripts and genomes in the oro-ocular (OO) model with SC16 strain. The amounts of viral products diverged markedly between neurological structures at 6 dpi (**Tables A1 to A5**). For example, viral genome varied from 10^9 (SCGs and spinal cord) to 10^7 copies (left TG and hypothalamus), and 2-kb LATs varied from 10^8 (left SCG and spinal cord) to 10^5 copies (right SCG). A similar order of magnitude was also found for ICP4, TK and UL18 transcripts, with generally the highest number of copies in the spinal cord and the lowest in the TGs (**Figure 5A to 9A**).

From 6 to 28 dpi, the 2kb LATs amount diminished in a large proportion in all tissues (10 to 1000-fold reduction), with the exception of the TGs where they were stable (left TG) or slightly increased (right TG). In comparison, reduction of viral genome copies and ICP4 transcripts was globally more important (10^2 to 10^7 -fold), with TGs being the sites of minimal reduction and SCGs the site of maximal reduction. Indeed, only one fifth of SCGs was finally positive at 28 dpi for ICP4 transcripts (**Figures 5A and 7A**). At this time of infection, TK and UL18 transcripts were not detected in all tissues, suggesting a complete extinction of the transcription of these early and late genes of HSV1 (**Figure 8A and 9A**).

From 28 to 100 dpi, the amounts of 2kb LATs were almost stable in all structures (less than 10-fold variation), whereas a reduction was clearly observed for genome copies, especially in the TGs (10^2 to 10^3 -fold reduction). Interestingly, no ICP4 transcripts were detected in all tested tissues at 100 dpi, neither were TK and UL18 transcripts (**Figures 5A to 9A**).

Viral transcripts and genomes at 6, 28 and 100 dpi in the corneal scarification (CS) model with the SC16 strain. For viral transcripts and genome copies, a striking difference between tissues was also found at the three periods of examination (**Figures 5B to 9B, Tables A1 to A5**). Whereas only 70 PFU were inoculated in the cornea, the amounts of viral DNA and transcripts in the TG at 6 dpi were similar to those observed with the OO model (using 10^6 PFU of SC16 strain), while SCG (especially right side), hypothalamus and spinal cord were less infected. This could reflect a very effective infection of primary neurons in the CS model, due to the very high number of nerve endings into the cornea and/or a more efficient replication of the virus at the inoculation site in the cornea as compared to the lip. In contrast, the later infection of the SCG in this model probably leads to a delay in the infection of second and third order neurons, located in the spinal cord, hypothalamus and right SCG (**Figure 5B to 9B, Tables A1 to A5**).

The difference between the OO and CS model (both using the SC16 strain of HSV1) was however more important concerning the evolution of viral products from 6 to 28 dpi. In the SCG,

hypothalamus and spinal cord, the decrease in the number of viral genome copies and 2-kb LATs was much less important in the CS model, with the exception of the TGs (no dramatic reduction, as in the OO model). In 3 out of 6 tissue types, the average number of ICP4 transcripts was only slightly reduced and about one half of tissue samples were still positive for TK transcripts, leading to an average number of 10^3 to 10^5 copies per structure. However, UL18 was no more detected at 28 dpi **(Figures 5D to 9D)**.

From 28 to 100 dpi, the average number of 2-kb LATs increased in the TGs (10^2 -fold), and the number of viral genome copies decreased, leading to a final viral load less important than in the OO model. This type of evolution was thus quite similar to the one observed between 6 and 28 dpi in the OO model. However, 3 out of 6 TGs and 2 out of 6 SCGs were still positive for TK and ICP4 transcripts at 100 dpi, respectively, but these viral transcripts were not found in the other tissues, and UL18 was not detected **(Figure 5 to 9, Tables A1 to A5)**.

Quantification of viral products at 6, 28 and 100 dpi in the corneal scarification model (CS) with the KOS strain. As with the two previous combinations of viral strain and route of inoculation, a difference of 2 to 3 orders of magnitude in the amounts of viral transcripts and genome copies (after standardization with GAPDH copies) was found in the various neurological tissues tested at the three times of infection **(Figure 5C to 9C, Tables A1 to A5)**.

However, the inoculation of the KOS strain following corneal scarification revealed important singularities. We first observed that the levels of genome copies, and ICP4, TK or UL18 transcripts were globally less important (about 10^2 -fold) at 6 dpi than with the OO model with SC16 strain, despite a similar size of inoculums (10^6 PFU). This also was the case when comparing with the inoculation of only 70 PFU of the SC16 strain in the CS model. Concerning the 2-kb LATs, the more important finding was their increase in all structures from 6 to 28 dpi and 100 dpi, whereas a decrease was observed in the other models (with the exception of the TGs where they were stable). Finally, corneal scarification with KOS strain was characterized by a delay in the extinction of productive

infection (and/or the onset of latent infection), as proved by the presence of UL18 transcripts in 3 out of 30 tissue samples at 28 dpi, and the presence of TK transcripts in 6 out of 18 tissue samples at 100 dpi (**Figure 5 to 9, Tables A1 to A5**).

Flow cytometry analysis of TG dissociated cells. The observation of clinical differences between animal models (i.e. a prolonged inflammation of the corneas in the CS model, especially with the KOS strain), led us to estimate the level of inflammatory response at a time that was supposed to correspond to the latent phase of HSV1 infection (28 days), using FACS analysis on dissociated cells from TGs at 28 dpi. We observed less positive cells for HSV1-infection (polyclonal anti-HSV1 antibodies) in the OO model than with both of CS models (with either KOS or SC16 strain, $p=0.006$ and 0.002 , respectively), confirming that the extinction of viral productive infection was much more efficient in the OO model. We also observed that the KOS strain induced more lymphocytes and less macrophages in TGs at 28 dpi than the SC16 strain with either the CS or the OO model ($p= 0.04$ and 0.004 , respectively) (**Figure 10**).

DISCUSSION

The expression of HSV1 genes during the latent phase of infection has been studied for years, but several questions remain unanswered, including the real state of gene expression. Early works showed that LATs were the only viral product abundantly expressed during HSV1 latency {**Rock, 1983 #1751;Stevens, 1987 #2154;Steiner, 1988 #1827**}{**Preston, 2000 #2186**}{**Roizman, 2001 #673**}{**Kent, 2003 #3101**} which correlates with the fact that the LAT regions are globally enriched in histone modification marks of open, transcriptionally active chromatin during latency unlike the rest of the viral genome {**Bloom, 2010 #5264**}{**Deshmane, 1989 #1757**}{**Catez, 2012 #5937**}. Interestingly, genes encoding lytic infection-associated ICP0 and ICP4 proteins (RL2 and RS1, respectively) are also localized in the LAT locus which could explain some basal level of transcription during the latent stage of infection. Indeed, a weak expression of transcripts encoding

ICP0 and ICP4 was reported in the TG at 28 days after the inoculation of the KOS or the 17+ strains of HSV1 by the corneal scarification method {Chen, 2002 #2811}{Kramer, 1995 #2157}{Thompson, 2006 #4364}. In these studies, TK transcripts were also found, which was rather unexpected since the corresponding gene is located in a theoretically highly repressed part of the HSV1 genome during latency {McGeoch, 1998 #4839}{Deshmane, 1989 #1757}{Bloom, 2010 #5264}. It is not clear whether expression of these transcripts was due to an occasional neuron that had recently reactivated from latency (spontaneous reactivation), to basal level of transcription leaking from the latent genome, or to continuous low level of HSV1 genome transcription.

In previous experiments based on the inoculation of the HSV1 SC16 strain in the lip, we observed the presence of ICP0-encoding transcripts in the TGs at 28 dpi, which was consistent with the results described above. However, in the same experiments, we did not detect TK or UL18 transcripts (respectively early and late transcripts of HSV1) {Maillet, 2006 #3281}, which was suggestive of a strong repression of the replication process in the tested tissues. To further understand these differences concerning the expression of the HSV1 lytic genes at times that were expected to be relevant for studying HSV1 latency, we analyzed the contribution of both, the route of viral inoculation (i.e. lip or cornea), and the viral strain for the discrepancies of gene expression during latency in different types of neurological tissues.

The expression of viral transcripts at 6 dpi varies with the viral strain and the route of inoculation. A striking difference in susceptibility to viral infection was observed when comparing the CS and the OO model. To induce similar rates of ocular morbidity and mortality, 10^6 PFU of the HSV1 SC16 strain were required in the OO model versus only 70 PFU in the CS model (**Figure 1 and 2**). Similarly, 10^6 PFU of the KOS strain induced an ocular disease in the CS model but not in the OO model. This is likely related to the very high numbers of nerve endings into the cornea, a tissue that is known in human as one of the most innervated of the whole body {Muller, 2003 #5425}.

As a consequence, the low virulence of the KOS strain combined with the reduced accessibility to nerve endings following inoculation into the lip likely explains the lack of morbidity of this viral strain in the OO model.

The low virulence of the KOS strain was also highlighted when comparing the amounts of viral transcripts and viral genome copies in the neurological tissues at 6 dpi following a corneal inoculation of 10^6 PFU (**Figure 5 to 9, Tables A1 to A5**), or the immunochemical staining of infected cells. Only few neurons were infected in the TG from 2 to 6 dpi following corneal scarification with the KOS strain, and other structures (SCG, hypothalamus and spinal cord) were very weakly infected, compared to the strong immuno-staining of infected cells already reported in these tissues at 6 dpi with the SC16 strain in the OO model **{Labetoulle, 2000 #2047}**.

The lower virulence of the KOS, compared to the SC16 strain, had already been shown in other experiments, such as in the rabbit ocular infection model **{Hill, 1987 #4813}** or in the mouse after an inoculation into the vitreous body. In the latter, where only 90 PFU of SC16 strain were sufficient to induce a bilateral retinitis in 100% of animals at 6 dpi after unilateral injection **{Labetoulle, 2000 #1297}**, whereas $1,5 \times 10^4$ to 2×10^5 PFU of the KOS strain induced bilateral retinitis in only 19% to 57% of mice **{Atherton, 1987 #20}{Pepose, 1987 #21}{Metzger, 1987 #22}**. The mutation within the glycoprotein B (gB) that is present in the KOS strain **{Yuhasz, 1993 #4814}{Mitchell, 1996 #2183}**, could explain the reduced virulence since gB is involved in viral adsorption and penetration, and in the cell-to-cell viral spreading **{Cai, 1988 #4815}**. The KOS strain is also characterized by two mutations in the ICP34.5 gene, which is known to have a major impact on neurovirulence **{Perng, 2002 #3082}{Mao, 2003 #4811}**, and also mutations in US9 gene, also involved in transport in neurons **{Negatsch, 2011 #5934}**. In contrast, the SC16 strain is considered as a wild type and neurovirulent strain **Shimeld, 1985 #16}{Ugolini, 1987 #59}{Ugolini, 1992 #9}**, isolated in 1976 from a patient with labial herpes **{Hill, 1975 #54}**.

ICP4 transcripts are expressed during the early stages of HSV1 latency in the oro-ocular (OO) model. We found ICP4-encoding transcripts in the OO model with SC16 strain at 28 dpi in all the tested neurological tissues, however in smaller amounts than at 6 dpi (10^3 to 10^6 -fold lesser, depending on the type of tissues). The number of ICP4 transcripts copies was not directly related to the amounts of the viral genomic matrix, as proved by the decrease in the ratio between ICP4 transcripts and viral genome copies (from 5×10^1 to 3×10^3 at 6 dpi to 2×10^{-2} to 8×10^{-4} at 28 dpi). Since the half-life of IE transcripts of HSV1 has been reported to be 5 to 7 hours {Henley, 1991 #5426}, the presence of ICP4 transcripts at 28 dpi likely reflects some low level persisting transcription of IE genes, rather than remnants from the first days of infection. Interestingly, UL18 and TK transcripts (respectively encoding L and E viral proteins) were produced in at least similar amounts than ICP4 at 6 dpi but were not detectable anymore at 28 dpi. At 100 dpi, none of the ICP4, TK and UL18 transcripts were found in any of the tested tissues. The decrease in the expression of late and early genes of HSV1 (namely UL18 and TK), observed between the 6th and the 28th dpi, seems thus to precede the silencing of the ICP4 gene during the onset of HSV1 latency.

In the CS model, ICP4-encoding transcripts were also detectable at 28 dpi when using the KOS strain, in accordance with previous studies {Chen, 2002 #2811}{Kramer, 1995 #2157}, or the SC16 strain. In both cases, however, TK and UL18 transcripts were also found, raising the question of the efficacy of the latency establishment in this model.

Expression of viral transcripts at late stages of infection: is viral latency achieved? While no TK or UL18 transcript were detected in the OO model at 28 dpi (even with using nested-PCR amplification method), TK transcripts were found in a significant proportion in tissue extracts at 28 dpi in the CS model with both viral strains, and even at 100 dpi with the KOS strain (1/6 of samples). Similarly, 10% of the tissue extracts at 28 dpi were also positive for UL18 transcripts. Flow cytometry showed a higher proportion of HSV1 antigen positive cells from dissociated TGs at 28 dpi with the CS model than with the OO model (**Figure 10**). Furthermore, *in situ* hybridization (ISH)

(Table 2) and quantitative RT-PCR (Figure 6) also found more LAT-expressing neurons in the TG and more amounts of LATs on whole tissue extracts in the OO model with the SC16 strain than in the CS model (with KOS or SC16 strain), suggesting that the accumulation of LATs in the TG neurons at 28 dpi, described as a marker of the balance from productive latent HSV1 infection {Preston, 2000 #2186}{Maillet, 2006 #3281} does not occur as efficiently in the CS model than in the OO model with the SC16 strain.

Taken together, these results suggest that the complete extinction of HSV1 replication cycle (as proven by the lack of TK and UL18 transcripts) is not fully obtained in the CS model, especially with the KOS strain. Noteworthy, the two important markers of the onset of HSV1 latency observed in the OO model from 6 to 28 dpi (i.e. the accumulation of LATs in the TGs and the low reduction of viral genome copies) were less obvious in the CS model with SC16 strain, and not observed with the KOS strain. These markers however were achieved between 28 and 100 dpi with the SC16 strain in the CS model, together with an almost disappearance of TK and UL18 transcripts, suggesting that extinction of HSV1 replication takes place later in the CS model, as compared to the OO model.

Role of inflammation induced by corneal scarification. The lack in the reduction of viral genome copies and the presence of TK or UL18 transcripts in the CS model at 28 and 100 dpi suggest that the inoculation by corneal scarification may induce events that lead to the persistence of some viral replication several weeks after the inoculation. Indeed, corneal scarification is a major trauma that induces a persistent inflammatory response in the TG {Shimeld, 1997 #1217} and an up-regulation of transcriptional factors such *oct-1*, *c-jun*, and *c-fos* {Valyi-Nagy, 1991 #1308}, which have binding sites in the promoter regions of several IE and E genes of HSV1 {Jones, 1985 #5265}{Stern, 1989 #3301}{Zabierowski, 2008 #4787}. In the CS model, all animals still exhibited ocular inflammation at 28 dpi (which persisted over three months), with active corneal neo vessels and clouding of the cornea, whereas eyes were not clinically inflamed in the OO model at 28 dpi and later (figure 3). Flow cytometry on dissociated TG cells at 28 dpi showed more lymphocytes in the CS

model, especially with the KOS strain, than in the OO model. Indeed, unrelenting infiltration of inflammatory cells has been previously described in the TG following corneal scarification {Shimeld, 1995 #1223}{Liu, 1996 #5266}, and the modifications of gB in the KOS strain {Yuhasz, 1993 #4814} {Mitchell, 1996 #2183}, which is one of the targets of both protective antibodies and cytotoxic T cells {Mester, 1990 #4816} may enhance this phenomenon. Since corneal nerve injury is a well-known triggering factor for HSV1 reactivation {Liesegang, 2001 #2726}{Shimeld, 1989 #1237}, the up-regulation of the viral gene expression in the CS model could be due to injury-like stimulation of the nerve ending by the persisting inflammation in the cornea.

HSV1 behavior and type of neurons. The OO model clearly showed a distinct evolution of HSV1 infection according to the type of neurological structures. From 6 to 28 dpi, there was a less important decrease of ICP4 transcripts and viral genome copies in the TG neurons than in the SCG, the hypothalamus or the spinal cord, and LATs accumulated instead of decreasing. While the differences were not in the same order of magnitude in the CS model (especially with the KOS strain), the TGs also exhibited a peculiar pattern in the evolution of viral infection from the acute to the latent phase of infection. Together with our previous results showing an accumulation of non-spliced ICP0 transcripts in the TG neurons only {Maillet, 2006 #3281}, these results strengthen the observation that HSV1 exhibits a special behavior in the TG during the onset of latent infection, which could be defined as follows: few reduction (10^1 to 10^3) of both viral genome copies and ICP4 transcripts, and accumulation of LATs and of non-spliced ICP0 transcripts.

The typical pattern of HSV1 latency observed in the OO model was less clearly observed in the CS model, suggesting that viral replication was not completely repressed at 28 dpi nor at 100 dpi, especially with the KOS strain. These findings suggest that the OO model using SC16 strain may be particularly relevant for studying both the cellular and viral events that occur during the latent stage of infection. It also appears as an excellent basic system for studying HSV1 reactivation. Since the latent state of infection is characterized by an almost complete extinction of the viral gene

expression, with the exception of genes located in the LAT region, any modification induced by a reactivation triggering factor should be more easily detected than in an animal model where the expression of some lytic viral genes is not completely repressed. Consequently, this animal model may also qualifies as a well-adapted system for assessing emerging preventive antiviral strategies in the future, as the basal level of viral gene expression during latent infection is clearly defined.

Finally, our results show how the viral gene expression pattern may also differs upon the type of neurons, especially during the establishment of HSV1 latency. Other studies had already suggested that viral behavior may vary between neuron types, as for example the speed of viral propagation during neuro-invasion **{Ugolini, 1992 #9}{Labetoulle, 2000 #2047}**. Henceforth, HSV1 expression should thus be analyzed in the light of the neurons used for the study. Trigeminal neurons are more susceptible than other neurons to experimental viral reactivation **{Shimeld, 1990 #1235}{Bertke, 2011 #5941}**. This is also supported by the clinical observation that the most frequent sites of HSV1 recurrence in human are tissues innervated by TG neurons (lips and cornea) whereas upper body tissues that are not directly connected to the TG (e.g. the iris, the retina or the brain) are rarely injured **{Liesegang, 2001 #2726}**. The unique pattern of HSV1 gene expression observed in the TG neurons suggests that some special events occur in these cells, but the question on whether these peculiar features are the consequence or the source remains to be addressed.

Acknowledgements

Many thanks to Noura AHMAR-ERRAS and Christel PICARD for technical assistance, and to Magali BRECKLER for careful reading of the manuscript.

This work was partially supported by a grant (ANR-005-MIIM-008-01) from the ANR (*Agence Nationale de la Recherche*)

The authors have no conflicting financial interests.

TABLES

Table 1 : Primers and probes

Target	Name of primers ^a	Sequence of primers
Viral LATs	LAT For	5' CCC ACG TAC TCC AAG AAG GC 3'
	LAT Rev	5' AGA CCC AAG CAT AGA GAG CCA G 3'
	LAT Taq	5' FAM CCC ACC CCG CCT GTG TTT TTG TG Tamra3'
Viral ICP4 transcripts	ICP4 Nfor	5' TTT TCC CAC CCA AGC ATC G 3'
	ICP4 Nrev	5' TCG TCG TCG GCT CGA AAG 3'
	ICP4 For	5' AGG TGA CCT ACC GTG CTA C 3'
	ICP4 Rev	5' CTT GTT CTC CGA CGC CAT C 3'
	ICP4 Taq	5' FAM TCC GCC GTC GCA GCC GTA TCC Tamra 3'
Viral TK transcripts	TK Nfor	5' TGA AAC TCC CGC ACC TCT TCG 3'
	TK Nrev	5' ACA CCC GCC AGT AAG TCA TCG 3'
	TK For	5' TTC TCG CGG CCA TAA CAA C 3'
	TK Rev	5' ACC GTC TAT ATA AAC CCG CAG 3'
	TK Taq	5' FAM AGC AAG AAG CCA CGG AAG TCC GCC CG Tamra3'
Cellular GADPH	GADPH For	5' CAA GGT CAT CCA TGA CAA CTT TG 3'
	GADPH Rev	5' GGC CAT CCA CAG TCT TCT GG 3'
Viral DNA (gG gene)	gG For	5' CTG TTC TCG TTC CTC ACT GCC T 3'
	gG Rev	5' CAA AAA CGA TAA GGT GTG GAT GAC 3'
	gG Taq	5' FAM CCC TGG ACA CCC TCT TCG TCG TCA G Tamra3'
Viral UL18 transcripts	UL18 Nfor	5' TGGATTCGCTGGACCTCACG 3'
	UL18 Nrev	5' ATTTAGGTATGCTGGGGTTTCTCC 3'
	UL18 For	5' AGACCTCAACGTGCTGTACTAC 3'
	UL18 Rev	5' CGAGTAAACCATGTTAAGGACCAG 3'
	UL18 Taq	5' FAM CGGGAGTCGCCTCTCGCTGCT Tamra 3'

^a Primer names containing a capital N, correspond to the primers used in the nested PCR

Table 2 : Numbers of cells positive for LAT detected by ISH in trigeminal ganglion depending on viral strain and inoculation site

Strain, viral dose, inoculation site	(Mouse)	Number of positive neurons		Number of positive neurons	
		Right TG	Left TG	Right TG	Left TG
SC16 10 ⁶ PFU Oro-ocular	(125)	24	44	1.26	2.32
	(126)	42	43	2.62	2.69
	(127)	10	25	0.62	1.56
	(128)	6	15	0.27	0.68
	(129)	29	68	1.53	3.58
	(130)	20	38	1.05	2.00
	Average	21.8	38.8	1.47	2.57
SC16 70 PFU Corneal scarification	(150)	UTT ^a	UTT	UTT	UTT
	(151)	UTT	UTT	UTT	UTT
	(152)	UTT	UTT	UTT	UTT
	(153)	UTT	UTT	UTT	UTT
	(154)	UTT	UTT	UTT	UTT
	Average	UTT	UTT	UTT	UTT
KOS 10 ⁶ PFU Corneal scarification	(162)	7	17	0,64	1.54
	(163)	15	21	1,67	2.33
	(180)	2	5	0,14	0.36
	(181)	0	6	0,00	0.43
	(182)	2	1	0,14	0.07
	(183)	0	5	0,00	0.36
	Average	4.3	9.2	0.52	1.0

(a) UTT : under the threshold with the usual ISH method (but some positive neurons when the signal was amplified using the tyramide-biotin system)

Figure 1 Rates of survival of mice inoculated with either the SC16 or the KOS strains of HSV1, with respect to the route of inoculation

A and B: Survival rates of mice infected via corneal scarification with different sizes of inoculums of the SC16 strain (A) or the KOS strain (B) of HSV1.

C and D: Survival rates of mice infected via lip inoculation of different sizes of inoculums of the SC16 strain (C) or the KOS strain (D) of HSV1

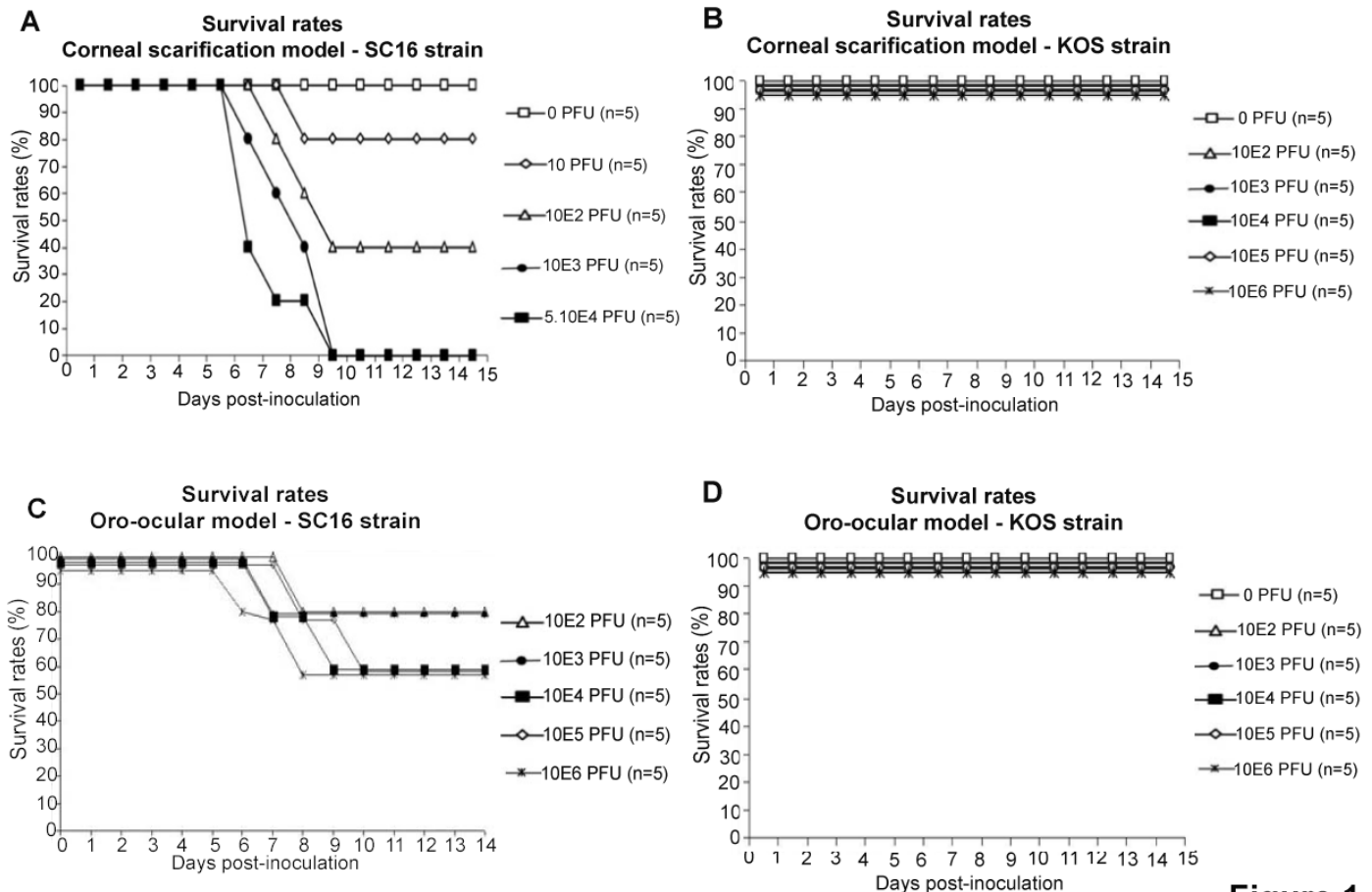


Figure 1

Figure 2 Rates of ocular disease in mice inoculated with either the SC16 or the KOS strains of HSV1, with respect to the route of inoculation

A and B: Rates of ocular disease in mice infected via corneal scarification with different sizes of inoculums of the SC16 strain (A) or the KOS strain (B) of HSV1.

C and D: Survival rates of mice infected via lip inoculation of different sizes of inoculums of the SC16 strain (C) or the KOS strain (D) of HSV1

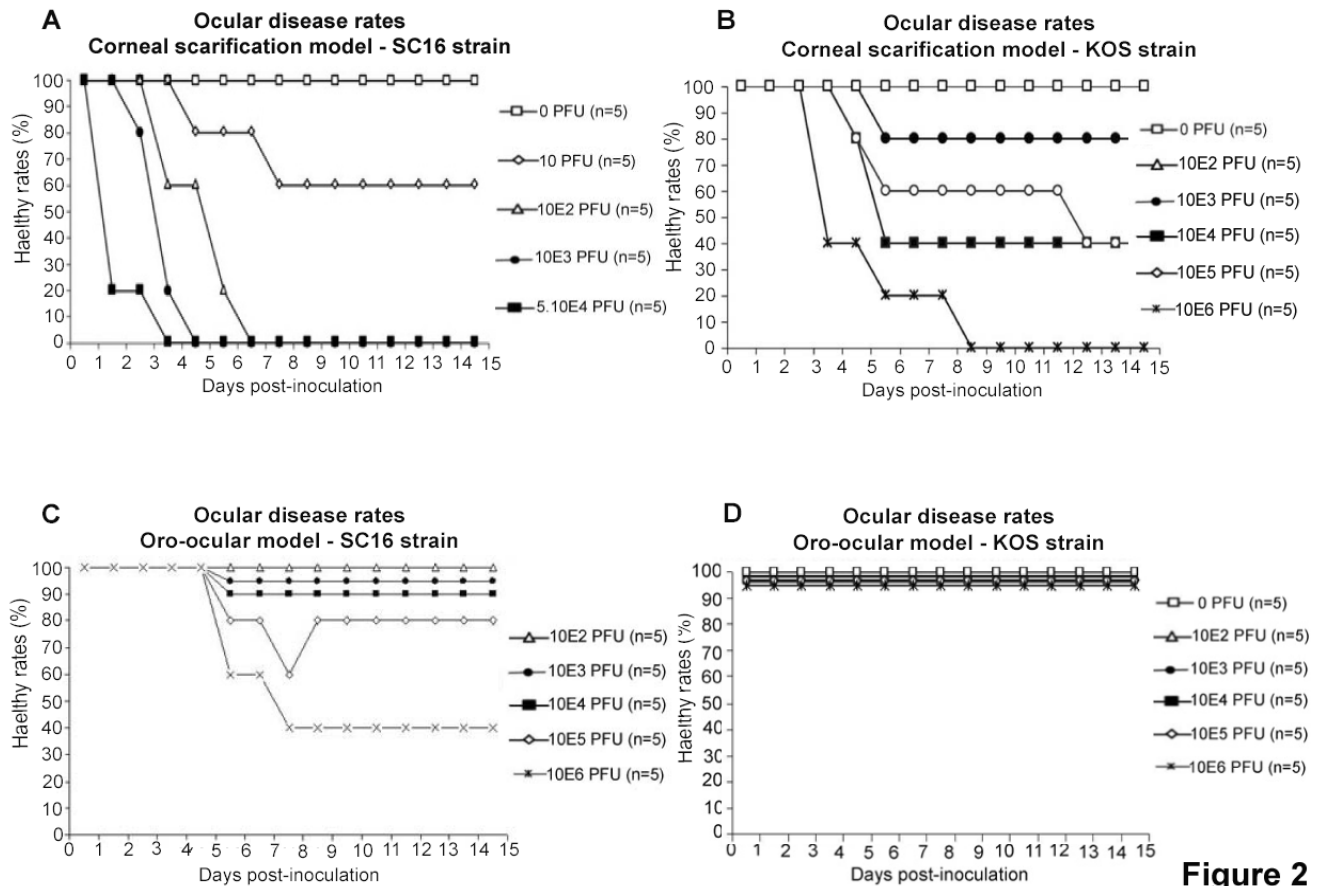


Figure 2

Figure 3 Clinical patterns of the ocular disease

Photographs of eyes of mice inoculated with the SC16 strain of HSV1 into the lip: the corneal inflammation had vanished at 28 days post-inoculation (A), whereas persistent corneal inflammation was found following inoculation with the KOS strain of HSV1 using corneal scarification (B and C).

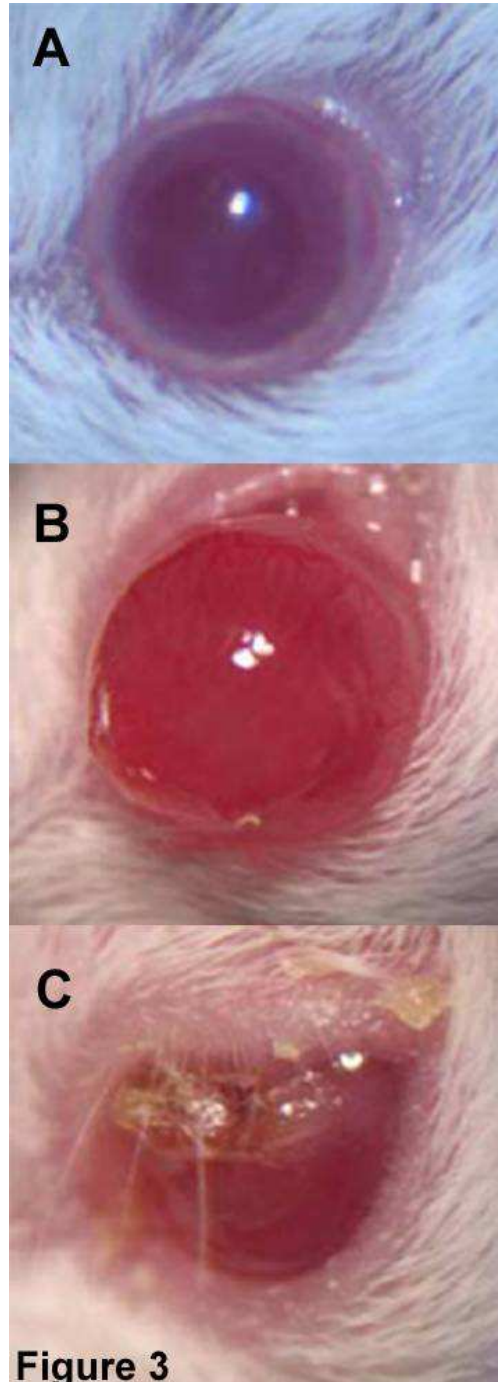


Figure 3

Figure 4 Pathways of viral propagation regardless inoculation site

(A) Spread of viral strain SC16 in oro-ocular model (B) Spread of viral strain in corneal scarification model. The day of infection of each tissue is noted in parentheses. TG, trigeminal ganglion; SCG, cervical superior ganglion; ciliary ganglion.

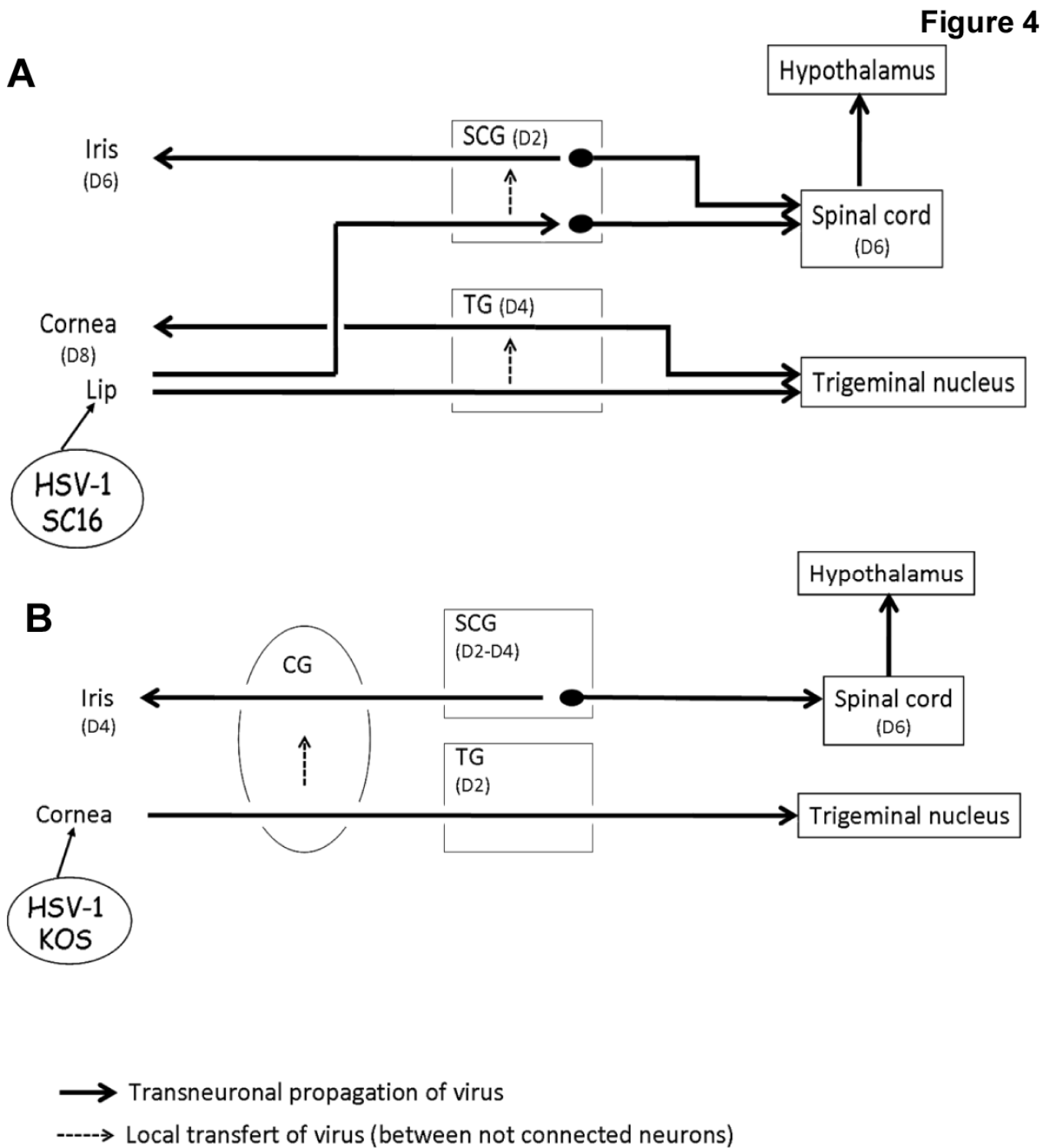


Figure 5 Bar graphs representing mean HSV1 DNA copies at 6, 28 and 100 days post-inoculation, depending on the viral strain and the route of inoculation

Amounts of HSV1 DNA copies were standardized with the number of GADPH DNA copies in the same tissue extracts. Bars represent the average amounts at respectively 6, 28 and 100 days post-inoculation, in either the oro-ocular model with the SC16 strain (A), or the corneal scarification model with the SC16 strain (B) or the KOS strain (C).

The evolution of the amounts of HSV1 DNA copies (after standardization) is represented as the ratio between the mean rates observed at 6 and 28 days post-inoculation (D) and between 28 and 100 days post-inoculation (E).

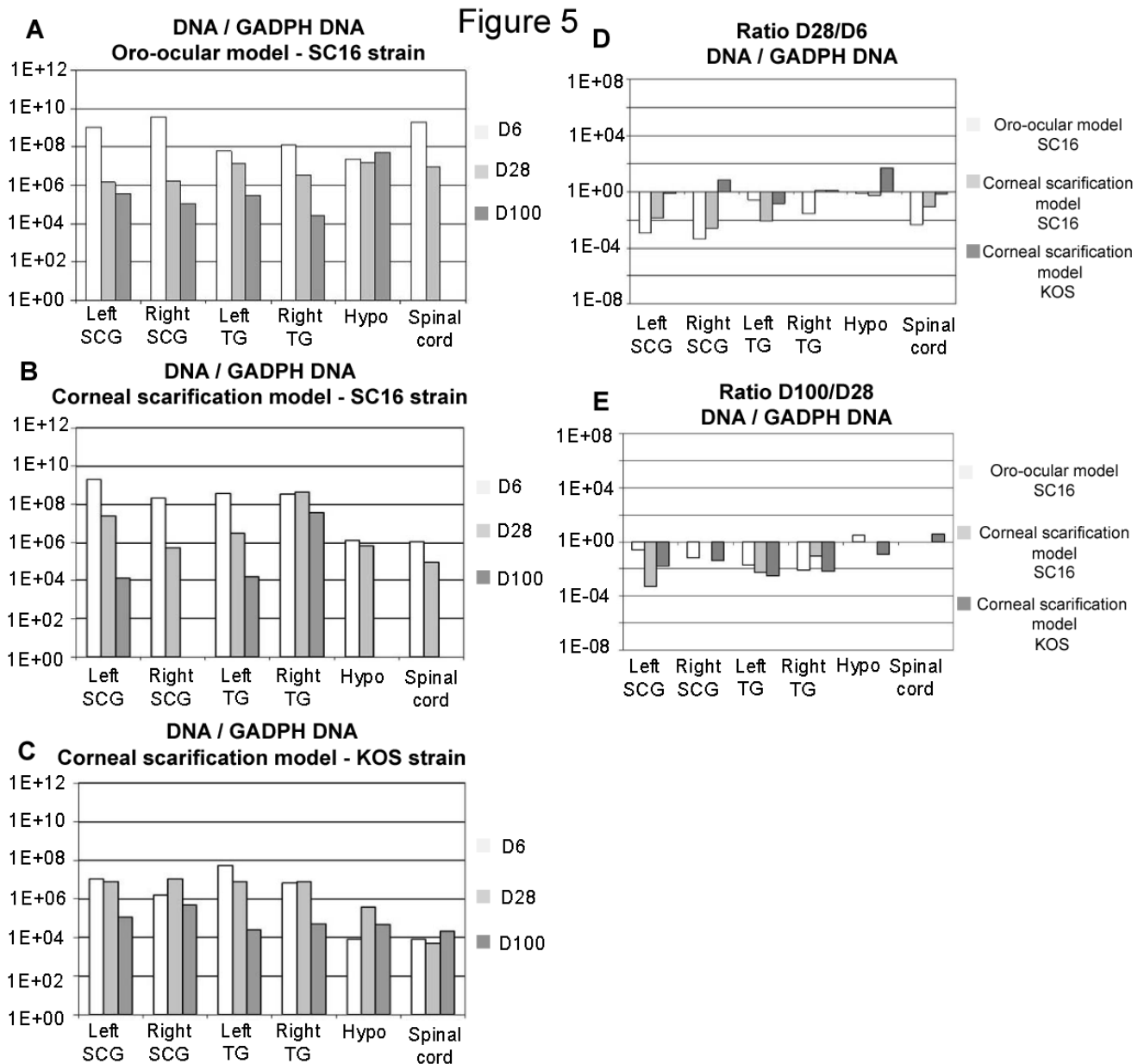


Figure 6 Bar graphs representing mean amounts of LAT copies at 6, 28 and 100 days post-inoculation, depending on the viral strain and the route of inoculation

Amounts of LAT copies were standardized with the number of GADPH transcripts in the same tissue extracts. Bars represent the average amounts at respectively 6, 28 and 100 days post-inoculation, in either the oro-ocular model with the SC16 strain (A), or the corneal scarification model with the SC16 strain (B) or the KOS strain (C).

The evolution of the amounts of LATs (after standardization) is represented as the ratio between the mean rates observed at 6 and 28 days post-inoculation (D) and between 28 and 100 days post-inoculation (E).

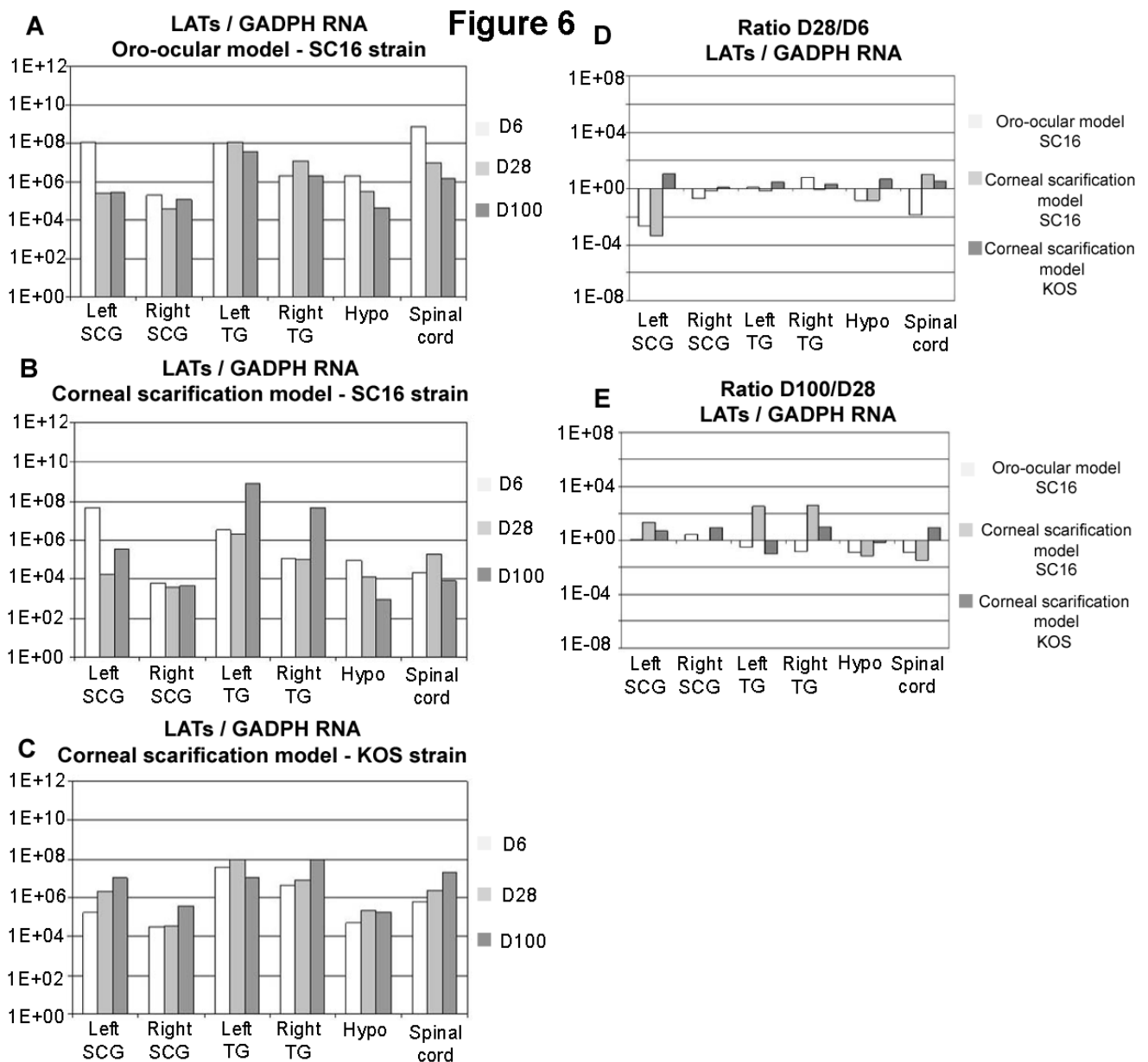


Figure 7 Bar graphs representing mean amounts of ICP4 transcripts at 6, 28 and 100 days post-inoculation, depending on the viral strain and the route of inoculation

Amounts of ICP4 transcripts were standardized with the number of GADPH transcripts in the same tissue extracts. Bars represent the average amounts at respectively 6, 28 and 100 days post-inoculation, in either the oro-ocular model with the SC16 strain (A), or the corneal scarification model with the SC16 strain (B) or the KOS strain (C).

The evolution of the amounts of ICP4 transcripts (after standardization) is represented as the ratio between the mean rates observed at 6 and 28 days post-inoculation (D) and between 28 and 100 days post-inoculation (E).

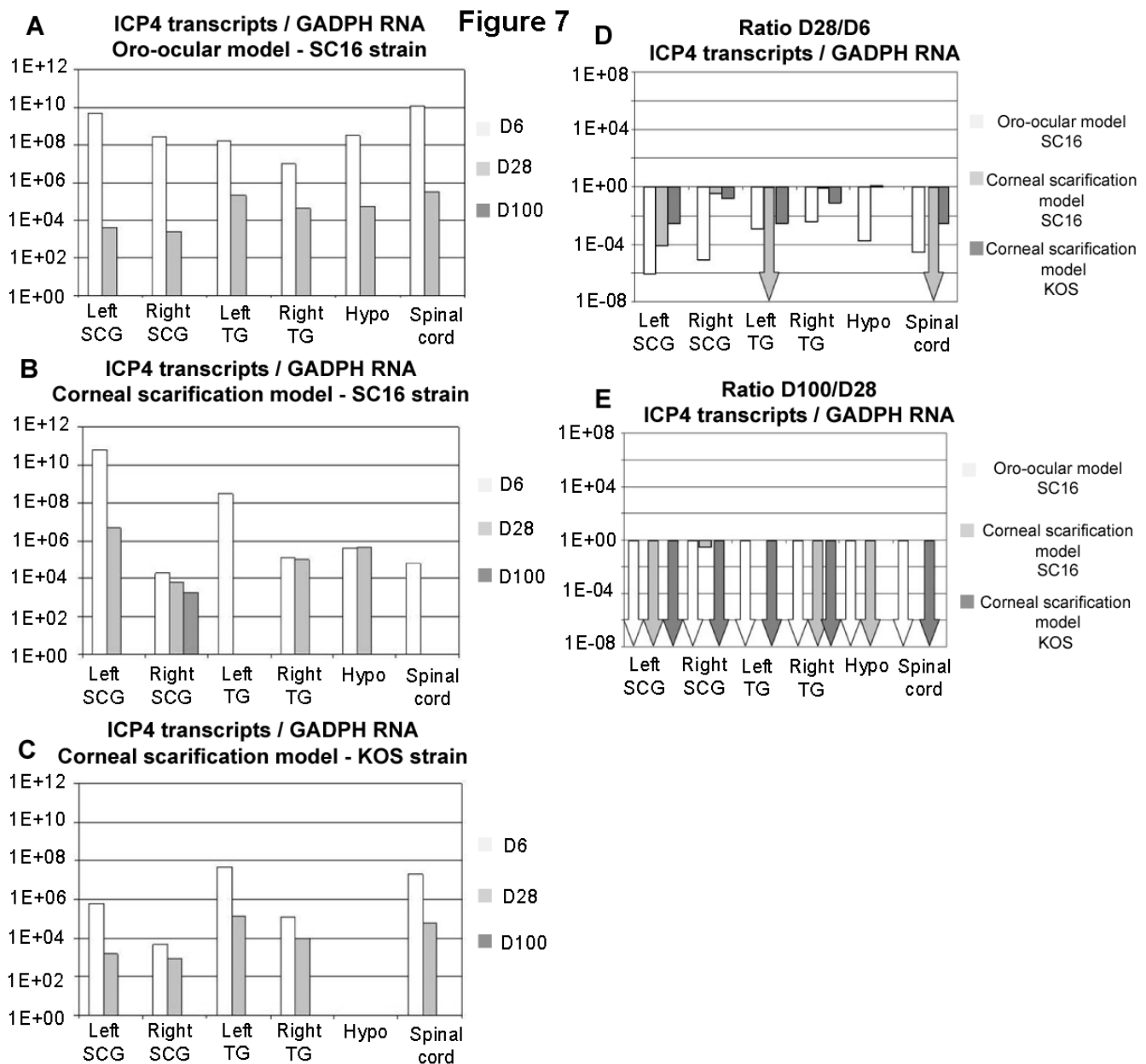


Figure 8 Bar graphs representing mean amounts of TK transcripts at 6, 28 and 100 days post-inoculation, depending on the viral strain and the route of inoculation

Amounts of TK transcripts were standardized with the number of GADPH transcripts in the same tissue extracts. Bars represent the average amounts at respectively 6, 28 and 100 days post-inoculation, in either the oro-ocular model with the SC16 strain (A), or the corneal scarification model with the SC16 strain (B) or the KOS strain (C).

The evolution of the amounts of TK transcripts (after standardization) is represented as the ratio between the mean rates observed at 6 and 28 days post-inoculation (D) and between 28 and 100 days post-inoculation (E).

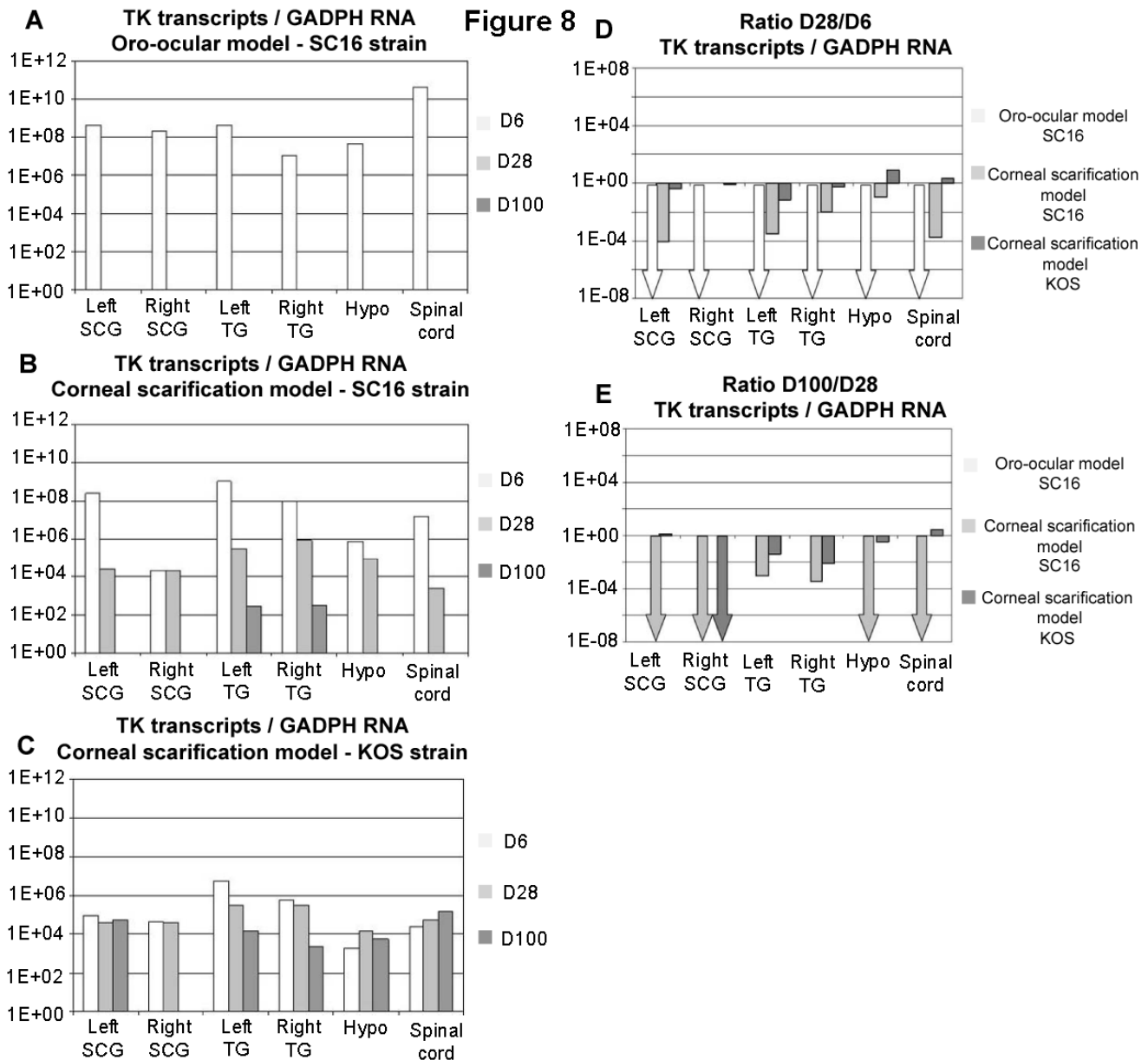


Figure 9 Bar graphs representing mean amounts of UL18 transcripts at 6, 28 and 100 days post-inoculation, depending on the viral strain and the route of inoculation

Amounts of UL18 transcripts were standardized with the number of GADPH transcripts in the same tissue extracts. Bars represent the average amounts at respectively 6, 28 and 100 days post-inoculation, in either the oro-ocular model with the SC16 strain (A), or the corneal scarification model with the SC16 strain (B) or the KOS strain (C).

The evolution of the amounts of UL18 transcripts (after standardization) is represented as the ratio between the mean rates observed at 6 and 28 days post-inoculation (D) and between 28 and 100 days post-inoculation (E).

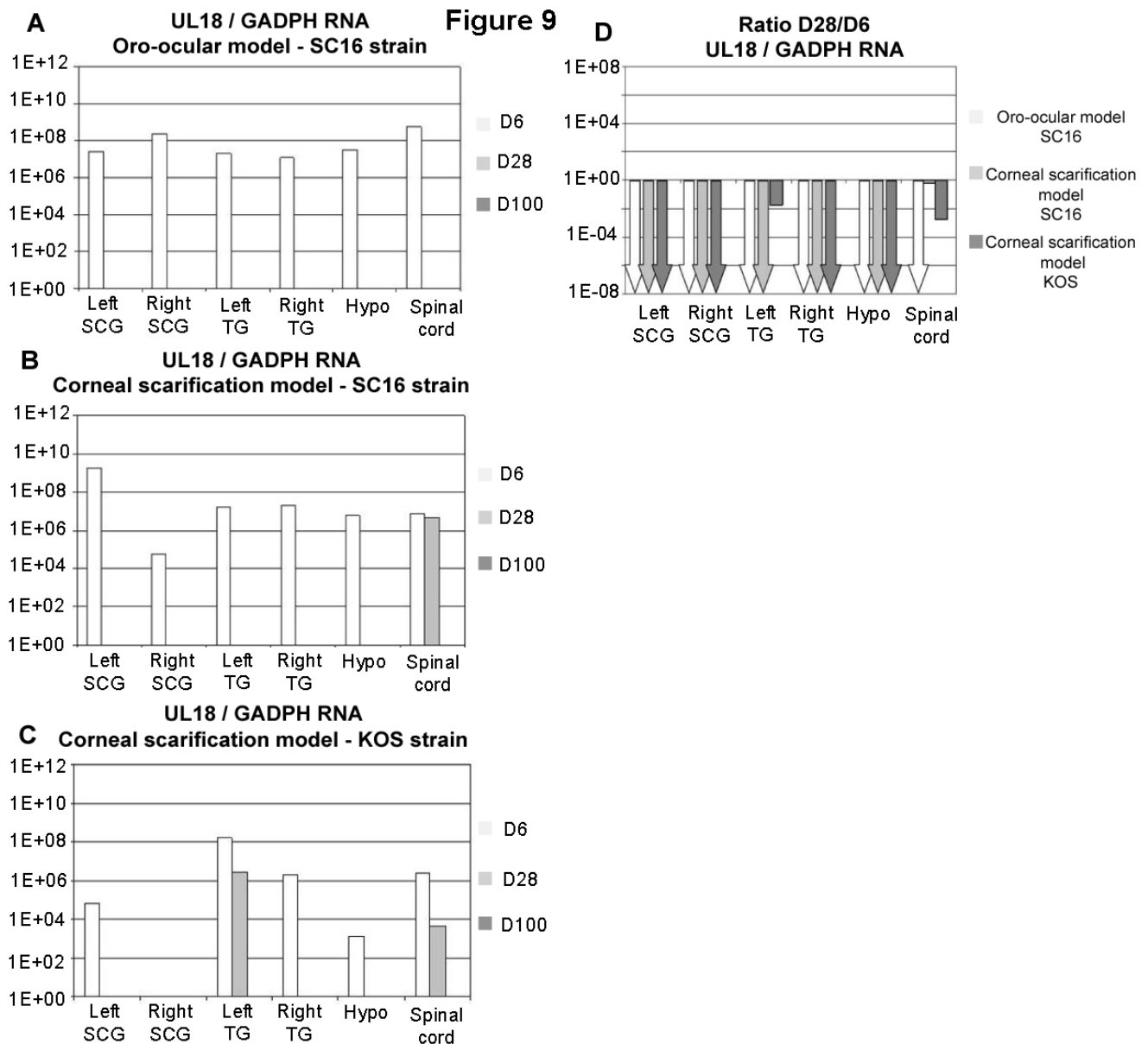
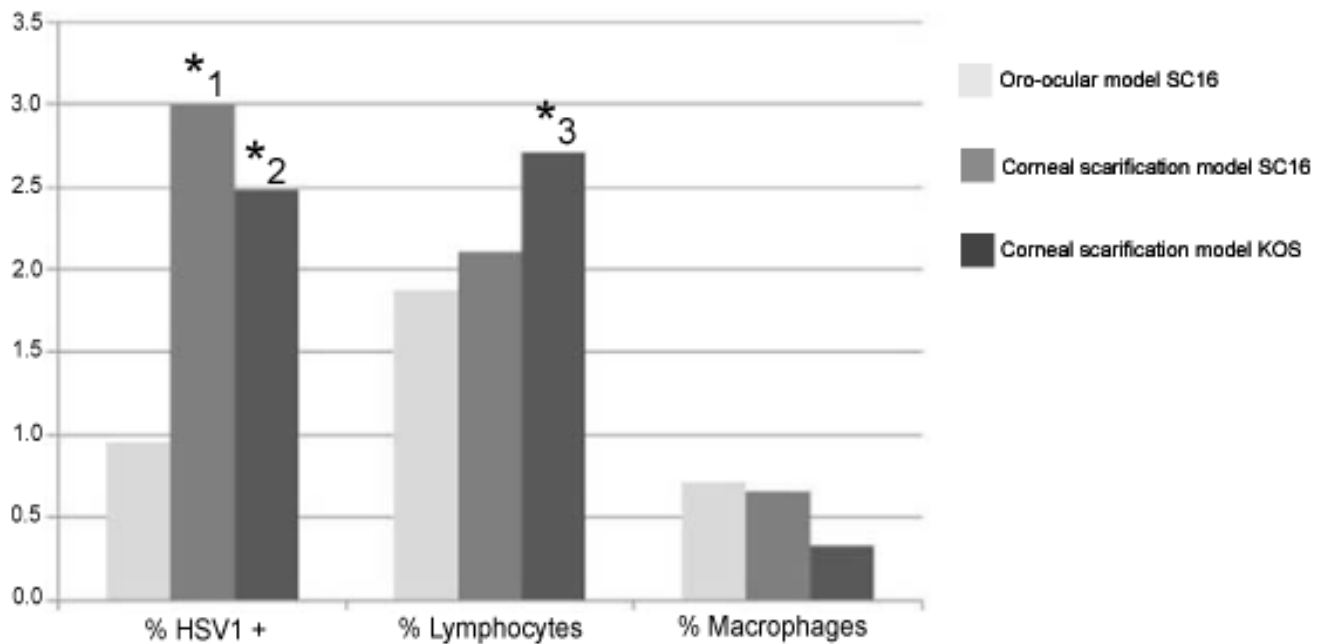


Figure 10 Quantification by flow cytometry of TG dissociated cells at 28 days post-inoculation, with respect to the viral strain and the route of inoculation

Aliquots of cells were immunostained for makers of either HSV1 infection (polyclonal anti-HSV1 antibodies), macrophages (anti-CD11b antibodies) and lymphocytes (anti-CD8a antibodies). The results are given as a percentage of total dissociated cells.

The percentage of HSV1 cells was significantly lower following inoculation of SC16 in the oro-ocular model than in the corneal scarification model, with either the SC16 (*1) or the KOS (*2) strain of HSV1. The percentage of lymphocyte was higher following the inoculation of KOS compared to SC16 in the corneal scarification model (*3), whereas the opposite was observed for macrophages (*4).

Figure 10



Annexe 2

Meganuclease-mediated Inhibition of HSV1 Infection in Cultured Cells

Stéphanie Grosse¹, Nicolas Huot², Charlotte Mahiet^{3,4}, Sylvain Arnould¹, Sébastien Barradeau³, Diane Le Clerre¹, Isabelle Chion-Sotinel¹, Cécile Jacqmarcq¹, Benoît Chapellier⁵, Ayla Ergani², Carole Desseaux¹, Frédéric Cédrone⁶, Emmanuel Conseiller³, Frédéric Pâques^{1,6}, Marc Labetoulle^{2,7} and Julianne Smith¹

¹Collectis Genome Surgery, Romainville, France; ²Laboratoire de Virologie Moléculaire et Structurale, Centre National de la Recherche Scientifique, Gif-sur-Yvette, France; ³Genomic Vision, Pépinière Paris Santé Cochin, Paris, France; ⁴Molecular Retrovirology Unit, Virology Department, Institut Pasteur, Paris, France; ⁵Institut de la Vision, Paris, France; ⁶Collectis, Romainville, France; ⁷Ophthalmology Department, Hôpital Bicêtre, APHP, Université Paris Sud, Kremlin-Bicêtre, France

Herpes simplex virus type 1 (HSV1) is a major health problem. As for most viral diseases, current antiviral treatments are based on the inhibition of viral replication once it has already started. As a consequence, they impair neither the viral cycle at its early stages nor the latent form of the virus, and thus cannot be considered as real preventive treatments. Latent HSV1 virus could be addressed by rare cutting endonucleases, such as meganucleases. With the aim of a proof of concept study, we generated several meganucleases recognizing HSV1 sequences, and assessed their antiviral activity in cultured cells. We demonstrate that expression of these proteins in African green monkey kidney fibroblast (COS-7) and BSR cells inhibits infection by HSV1, at low and moderate multiplicities of infection (MOIs), inducing a significant reduction of the viral load. Furthermore, the remaining viral genomes display a high rate of mutation (up to 16%) at the meganuclease cleavage site, consistent with a mechanism of action based on the cleavage of the viral genome. This specific mechanism of action qualifies meganucleases as an alternative class of antiviral agent, with the potential to address replicative as well as latent DNA viral forms.

Received 4 June 2010; accepted 12 December 2010; published online 11 January 2011. doi:10.1038/mt.2010.302

INTRODUCTION

Herpes simplex virus type 1 (HSV1) is a major health problem with substantial impact on quality of life and disability-related costs.¹ As with two other human α -herpesvirinae (namely HSV type 2 and varicella-zoster virus), HSV1 is able to become latent in neurons before inducing recurrent infections in peripheral tissues. The primary infection, asymptomatic in >90 % of cases,² takes place in the oral mucosae, as a consequence of contact with infected particles in saliva. After replication in epithelial tissues, viruses propagate in neurons before becoming latent. Following various triggering factors, HSV1 may reactivate and thus invade the peripheral tissues

connected to the reactivated neuron. Since the principal location of latent HSV1 is the trigeminal ganglion (TG), responsible for sensory innervation of the face, most recurrences are located in the eyes or the lips. The seroprevalence of HSV1 in the general population ranges from 24.5% to 67%, with 15 to 45% of positive subjects experiencing recurrent herpes labialis.¹ The eye, and particularly the cornea, is the second most frequent location of HSV1 infection. The prevalence of herpes simplex keratitis is 149/100,000,³ with >30 new events per 100,000 inhabitants annually.⁴ As a consequence, HSV1 is a leading cause of blindness throughout the world. In addition, despite the use of topical or oral antiviral agents, herpes simplex keratitis remains the most frequent cause of infectious corneal opacities in the most developed areas of the world,⁵ accounting for about 10% of patients undergoing corneal transplantation.⁶ Moreover, the natural risk of HSV1 reactivation in recently grafted cornea is about 25% in the first year following surgery.⁷ Today, HSV1 remains a major cause of corneal graft failure, accounting for about 22% of all cases of regrafting.⁸

To date, commercially available anti-HSV1 agents are able to inhibit viral replication through an inhibition of the DNA-polymerase,⁹ meaning that any reduction in the bioavailability of the drug and/or the sensitivity of the virus may result in treatment failure. Moreover, current treatments do not reduce the load of the DNA matrix, and thus are unable to reduce the risk of further viral reactivation. Ideally, an ultimate weapon against HSV1 infection should be durably present in the cells, avoid questions of sensitivity and, if possible, reduce the load of viral genomes. Very specific endonucleases such as zinc finger nucleases^{10,11} or meganucleases^{12,13} could be used to address these goals, especially if they are delivered by a gene transfer process.

With the aim of a proof of concept study, we assessed anti-HSV1 activity in cells transfected with meganuclease-encoding plasmids. Meganucleases are endonucleases that recognize large (>12 bp) DNA sequences. In nature, they induce the spreading of mobile genetic elements by a process called homing,¹⁴ and are therefore also called homing endonucleases. Homing endonucleases have also been used to stimulate targeted recombination

in immortalized mammalian cells^{15,16} and in various cell types and organisms (for review see ref. 13). Moreover, homing endonucleases with tailored specificities can be engineered^{17–23} and redesigned homing endonucleases targeting the human *XPC* and *RAG1* genes have been described previously.^{24–26}

In this report, we used both an I-SceI natural meganuclease and a set of redesigned meganucleases derived from I-CreI, to cleave a recombinant rHSV1 virus and/or a wild-type strain of HSV1. Our results demonstrate that HSV1-specific meganucleases can have a significant inhibitory effect on HSV1 infection at low and moderate multiplicities of infection (MOIs) (*i.e.*, 10^{−3} to 1). Furthermore, this inhibition is associated with a high rate of mutation in the viral genome, at the meganuclease cleavage site, consistent with a mechanism of action based on virus clipping.

RESULTS

Engineering of meganucleases recognizing sequences from the HSV1 virus

We have described in a series of previous reports the engineering of redesigned meganucleases targeting human genes.^{17,24,26,27} Briefly, we have gathered a proprietary collection (called omegabase) of >30,000 engineered endonucleases, derived mostly from the I-CreI meganuclease, and used these proteins as “building blocks” in a combinatorial process to create novel endonucleases with chosen specificities. This process, allows us to create endonucleases that cleave sequences from virtually any gene. Furthermore, although I-CreI derivatives are based on a dimeric scaffold, the engineered endonucleases can be transformed into single chain molecules, by fusing the two monomers through a peptidic linker. Single chain design can improve specificity, and enhance efficacy over a broad dose range.²⁸ All the meganucleases used in this study are single chain versions.

In order to design meganucleases cleaving HSV1 sequences, we first identified a series of “hits” or potential targets, *i.e.*, patchworks of sequences cleaved by proteins from Omegabase, in the HSV1 genome. Altogether, 546 such “hits” were found along the entire genomic sequence (NC_001806 from <http://www.ncbi.nlm.nih.gov>). Six of them were selected as final targets, including four in the 9 kb latency region. Functional meganucleases could be obtained for four of these six targets. These proteins, described in **Table 1**, are called HSV1m1, HSV1m2, HSV1m4,

and HSV1m12 and the position of their target sites HSV1t1, HSV1t2, HSV1t4, and HSV1t12 is shown in **Figure 1a**. The HSV1m4 and HSV1m12 meganucleases recognize two distinct sequences in the *ICP₀* gene, which encodes a protein (infected cell protein 0) that plays a central role in viral reactivation and/or viral growth at low MOIs.⁹ Since this gene is duplicated, both proteins have two recognition sites in the HSV1 genome. In contrast, HSV1m1 and HSV1m2 each have a single target site, in *US₂*, an envelope associated protein, and in *UL₁₉*, a major capsid component, respectively.

Characterization of anti-HSV1 meganuclease activity and toxicity

The activity of the anti-HSV1 endonucleases was characterized with a previously described cell based assay,²⁷ which directly monitors meganuclease-induced tandem repeat recombination in living cells (**Figure 1b**). We used as positive controls the I-SceI meganuclease, which has been used to induce highly efficient genome engineering by a large number of groups,¹³ and RAG1m, a previously described meganuclease shown to induce up to 6% of recombination at the *RAG1* locus in cultured cells.²⁸ HSV1m4 and HSV1m12 displayed very similar levels of activity, matching the activity of I-SceI and RAG1m, and HSV1m1 proved slightly less active (**Figure 1c**). However, HSV1m2 displayed a markedly different profile, with maximal activity being observed at a very low dose (0.39 ng), indicative of an extremely active protein. For comparison, the I-SceI and RAG1m proteins reached approximately the same maximal activity at a 16 times higher dose (6.25 ng) of plasmid.

Although rare cutting endonucleases are powerful tools for genome engineering, these proteins can cleave promiscuous sequences, resulting in off-site cleavage and a toxicity that can be limiting for therapeutic applications. In order to evaluate potential meganuclease toxicity, we used a previously described cell survival assay that uses green fluorescent protein (GFP) as a marker for cell viability.^{28–30} The I-SceI and RAG1m, shown to display little toxicity, if any, in a former study^{25,28} were used as controls as well as the wild-type I-CreI meganuclease (**Figure 1d**). Among the four anti-HSV1 meganucleases, HSV1m1 proved quite innocuous, with its profile mimicking the I-SceI profile. HSV1m2 behaved very similar to I-CreI, while the two other proteins displayed intermediate

Table 1 Anti-HSV1 target sequences and meganucleases

Target	Meganuclease
HSV1t1 ATGGGACGTCGTAAGGGGGCCTGG	HSV1m1 (a) I-CreI 7E30R33G38T96E106P132V (b) I-CreI 8K19S30G38R44K57E61R70E 75N108V
HSV1t2 ATAAACTCACACACGGCGTCTCGG	HSV1m2 (a) I-CreI 7E44D68T70S75R77R80K96E (b) I-CreI 8K19S28E38R40K44K54I61R70S75N 80K105A132V
HSV1t4 CCAAGCTGGTGTACCTGATAGTGG	HSV1m4 (a) I-CreI 7E44M70A75N80K96E105A132V 146K156G (b) I-CreI 8K19S32E38Y44A61R68Y70S75Y77K 80K105A132V
HSV1t12 CCTGGACATGGAGACGGGGAACAT	HSV1m12 (a) I-CreI 7E24V33C38S44I50R70S75N77R 96E132V (b) I-CreI 8K19S30R33S44K61R66H68Y70S77T 87I132V139R163S

For each meganuclease, (a) and (b) are the sequences of the two I-CreI monomers fused in the final single chain molecule (see Material and Methods).

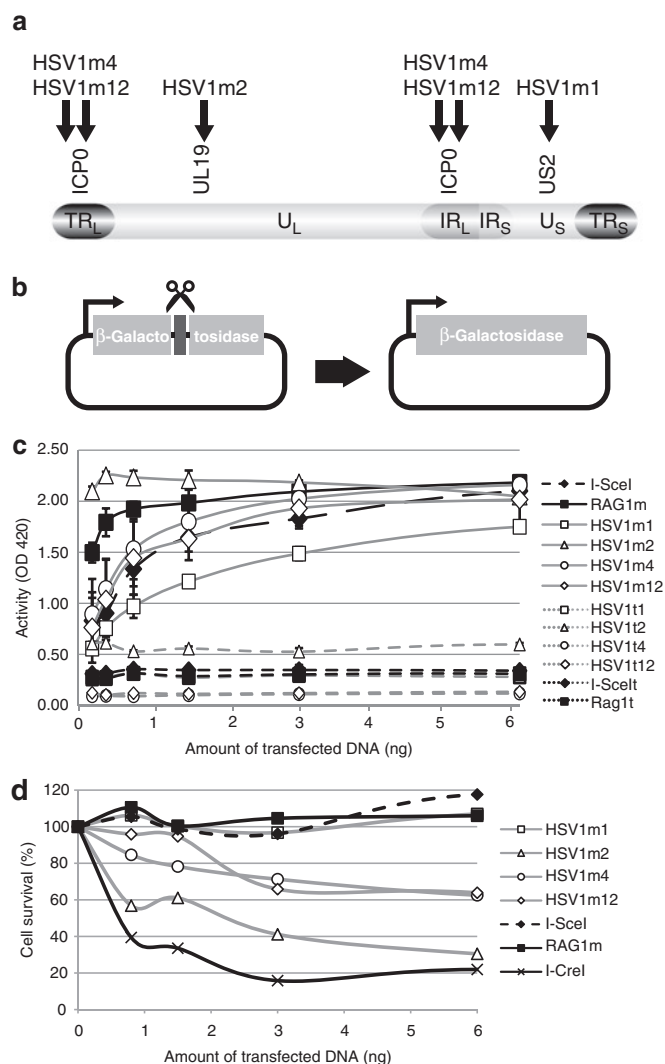


Figure 1 Characterization of anti-herpes simplex virus type 1 (HSV1) meganeucleases. **(a)** Schematic representation of the HSV1 genome with the position of recognition sites for meganeucleases HSV1m1, m2, m4, and m12 indicated. The overall structure of the HSV1 genome is shown with the unique long (U_L) and unique short (U_S) regions, the TR_L and TR_S terminal repeats regions, and the IR_L and IR_S inverted internal repeat regions. **(b)** and **(c)** Cleavage activity in an extra-chromosomal functional assay in CHO-K1 cells for HSV1m1, HSV1m2, HSV1m4, and HSV1m12 compared to I-SceI and RAG1m meganeucleases as positive controls. HSV1t1, HSV1t2, HSV1t4, HSV1t12, I-SceI, and Rag1t represent cleavage activity in the absence of meganeucleases. Experiments were performed in triplicate and the error bars represent standard deviation. **(d)** Evaluation of the toxicity of meganeucleases by a cell survival assay in CHO-K1 cells. Various amounts of plasmid expressing HSV1m1, HSV1m2, HSV1m4, or HSV1m12 and a constant amount of plasmid encoding green fluorescent protein (GFP) were used to cotransfect CHO-K1 cells. Cell survival is expressed as the percentage of cells expressing GFP 6 days after transfection, as described in Materials and Methods. I-SceI and RAG1m meganeucleases are shown as a control for nontoxicity and I-CreI is shown as a control for toxicity.

patterns. However, it is unclear which protein displays the best activity/toxicity ratio, as HSV1m2 is also significantly more active than I-SceI, RAG1m, and the other anti-HSV1 meganeucleases (see above).

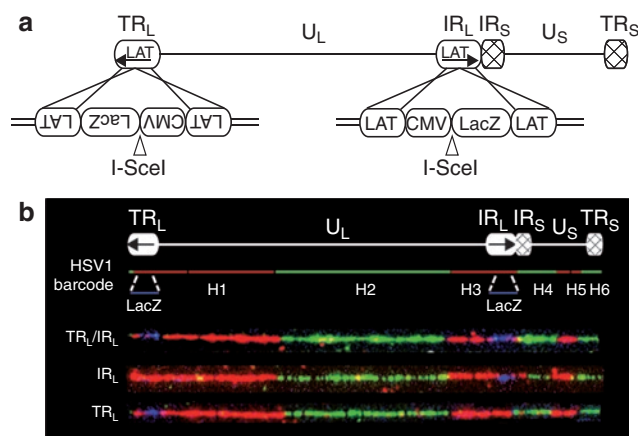


Figure 2 Construction and characterization of the recombinant herpes simplex virus type 1 virus (rHSV1). **(a)** Genomic structure of rHSV1. The overall structure of the HSV1 genome is shown with unique long (U_L) and unique short (U_S) regions and the TR_L/IR_L and TR_S/IR_S repeats. An expression cassette containing the cytomegalovirus (CMV) promoter and the LacZ coding sequence was inserted in the major latency-associated transcript (LAT) genes by homologous recombination. The I-SceI target site was cloned between the CMV promoter and the LacZ gene. **(b)** Schematic representation of the rHSV1-specific barcode. Alexa 594-fluorescence H1, H3, and H5 probes are indicated by the red bars, Alexa 488-fluorescence H2, H4, and H6 probes are depicted by the green bars and the LacZ probe by a blue bar. Underneath, hybridization signals that are representative of the different observed hybridization patterns are shown. IR, inverted internal repeat region; TR, terminal repeat region.

Generation of a marked recombinant HSV1 virus containing an I-SceI cleavage site

Among rare cutting endonucleases, the I-SceI meganeuclease remains today the gold standard in terms of activity and specificity.¹³ Therefore, in order to compare the antiviral efficacy of our engineered meganeucleases with this widely recognized benchmark, we built a recombinant HSV1 virus (rHSV1), containing a *LacZ* reporter gene with an I-SceI site between the *LacZ* ORF and its promoter (Figure 2a). The I-SceI cleavage site does not interfere with β-galactosidase expression. The recombinant virus was built by homologous gene targeting into the latency-associated transcript region. However, since this sequence is duplicated, two kinds of insertions can be expected (Figure 2a). Molecular combing, a powerful technique enabling the direct visualization of individual molecules, was used to determine the structure of the rHSV1. Basically, individual molecules of high molecular weight DNA are “combed” on a silanized glass coverslip, hybridized with different DNA probes³¹ and examined by microscopy. We applied molecular combing to uniformly stretch rHSV1 DNA extracted from viral particles and hybridized the resulting combed rHSV1 DNA with labeled adjacent and overlapping HSV1-specific DNA probes and a *LacZ* probe. Immunofluorescence microscopy exhibited 955 multicolor linear patterns that fulfilled the criteria for evaluation (Figure 2b). 94.6 % of these patterns ($n = 903$) showed two *LacZ* signals within the H1 and H3 regions, showing that insertions are present in both the terminal repeat region (TR_L) and inverted internal repeat region (IR_L) of the rHSV1 genome. In contrast, only a few signal patterns exhibited a single *LacZ* signal: 4.3% ($n = 41$) and 1.1% ($n = 11$) in the IR_L region and the TR_L region, respectively. Thus,

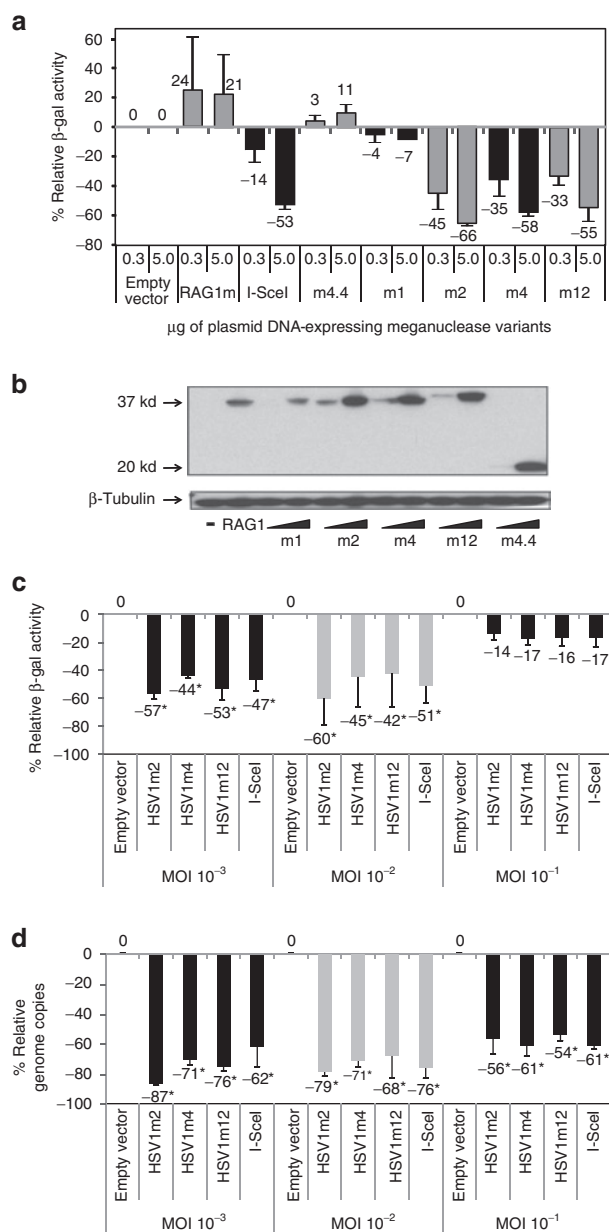


Figure 3 Meganuclease-mediated inhibition of infection by recombinant herpes simplex virus type 1 virus (rHSV1). **(a)** COS-7 cells were transfected with various amounts (0.3 and 5 μ g) of empty vector or meganuclease expressing plasmids, and infected 24 hours later with rHSV1 at a multiplicity of infection (MOI) of 10^{-3} . The level of infection was assessed 24 hours later with a β -galactosidase activity test, as described in Materials and Methods. The relative β -galactosidase activity (in %) compared to samples transfected with the empty vector is indicated. Four anti-HSV1 meganucleases were tested with several controls. HSV1m4.4 (20 kDa) is a homodimeric variant of HSV1m4. **(b)** Meganuclease (37 kDa) expression levels were analyzed in COS-7 cells by western blotting at 24 hours after transfection with various amounts of plasmid (0.3 and 5 μ g). Antibody against β -tubulin (50 kDa) was used for the loading control. **(c and d)** HSV1m2, HSV1m4, and HSV1m12 meganucleases were tested at increased MOIs. A single concentration (5 μ g) of meganuclease expression plasmid was introduced in COS-7 cells and infected 24 hours later with rHSV1 at a MOI of 10^{-3} , 10^{-2} , or 10^{-1} . Viral load was monitored by a β -galactosidase activity test **(c)** and by quantitative PCR **(d)**. Results are the average of three independent experiments. Error bars represent standard deviation. *Statistical significance, $P < 0.05$.

the vast majority of recombinant HSV1 molecules in our viral preparation display two inserts.

Meganucleases can prevent the infection of cultured cells by a recombinant HSV1 virus

The replication competent rHSV1 virus was used to test the activity of four anti-HSV1 meganucleases in African green monkey kidney fibroblast (COS-7) cells. We first tried to identify inhibition at a low MOI, mimicking the initial steps of infection. Briefly, COS-7 cells were transfected with meganuclease-encoding plasmids at two different doses. One day later, these cells were infected at a MOI of 10^{-3} , with the rHSV1, and viral infection was estimated 1 day postinfection using a standard β -galactosidase test. In an initial set of experiments (**Figure 3a**), we evaluated the effect of HSV1m1, HSV1m2, HSV1m4, HSV1m12, I-SceI, RAG1m, and HSV1m4.4, a meganuclease related to HSV1m4, but recognizing only one half of the HSV1t4 target. As expected, no infection inhibition was observed with the RAG1m and HSV1m4.4 meganucleases. In contrast, an inhibitory effect was observed with HSV1m2, HSV1m4, HSV1m12, and I-SceI. HSV1m1 had virtually no impact on infection. Western Blot performed with extracts of the transfected COS-7 cells indicates that this protein was expressed at a lower level than the three other anti-HSV1 meganucleases (**Figure 3b**), and thus may partially explain the lack of activity observed. The reduced levels of HSV1m1 protein may result from altered transcription, mRNA stability, mRNA translation, or protein stability due to differences in the coding sequence as has been previously shown with GFP variants.³²

To further characterize the antiviral potential of the HSV1m2, HSV1m4, and HSV1m12 meganucleases under more stringent conditions, cells were infected at MOIs of 10^{-3} , 10^{-2} , and 10^{-1} . In order to directly evaluate the effect of the meganuclease on viral DNA production, a quantitative PCR assay (**Figure 3d**) was performed in addition to a β -galactosidase activity test (**Figure 3c**). LacZ monitoring detected a significant inhibition of viral infection with HSV1m2, HSV1m4, and HSV1m12 at a MOI of 10^{-3} and 10^{-2} (**Figure 3c**). However, this effect was no longer visible at a MOI of 10^{-1} , likely due to saturating levels of LacZ activity at elevated levels of infection. In contrast, a significant reduction in viral DNA was detected for the three anti-HSV1 meganucleases by quantitative PCR in all conditions, with up to 61% of reduction of viral load at a MOI of 10^{-1} (**Figure 3d**). In addition, viral inhibition is not specific to rHSV1 as similar experiments performed with wild-type HSV1 resulted in an equivalent reduction in viral DNA as determined by quantitative PCR (data not shown).

To further evaluate the ability of meganucleases to inhibit viral replication a plaque assay was performed. In this assay COS-7 cells were transfected with a control plasmid or a meganuclease-encoding plasmid, 1 day later, cells were infected with rHSV1 at a MOI of 10^{-3} and 2 days postinfection the number of plaques formed was determined after staining with X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside). A dramatic reduction in the formation of plaques was observed with both HSV1m2 and HSV1m4 (**Figure 4a,b**), indicating that both of these meganucleases are able to effectively inhibit viral replication. In contrast, no reduction in plaque formation was observed with two nonrelevant meganucleases, RAG1m, and wild-type I-CreI. These results indicate that inhibition of viral replication is specific to the anti-HSV1 meganucleases.

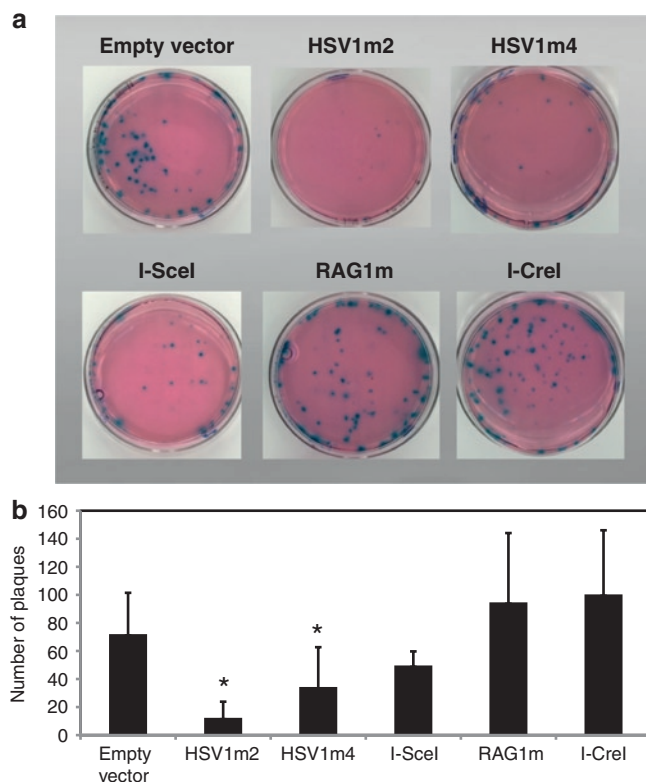


Figure 4 Prevention of recombinant herpes simplex virus type 1 virus (rHSV1) infection by HSV1m2 and HSV1m4. Cos-7 cells were transfected with 8 μ g of empty vector or meganuclease expressing plasmids, and infected 24 hours later with rHSV1 at a MOI of 10^{-3} , as described in Materials and Methods. **(a)** The β -galactosidase positive "blue" plaques were assessed 2 days later as described in Materials and Methods. **(b)** The average number of "blue" plaques detected in three independent experiments is indicated. Error bars represent standard deviation. *Statistical significance, $P < 0.05$.

In addition, the absence of inhibition with wild-type I-CreI, which displays a toxicity profile similar to that of HSV1m2 (**Figure 1d**), indicates the antiviral activity observed with the anti-HSV1 meganucleases is not due to a general toxicity effect.

Meganucleases prevent infection of cultured cells by a wild-type HSV1 virus

To determine if anti-HSV1 meganucleases were also active on a wild-type HSV1 virus, their antiviral activity was assessed in BSR cells with the wild-type SC16 strain of HSV1. Since the results obtained above suggested that anti-HSV1 meganucleases had a strong antiviral effect, we raised the virus dose in this experiment, using several MOIs ranging from 0.1 to 8. Briefly, BSR cells were cotransfected with GFP and meganuclease-encoding plasmids. Two days later, cells were infected at different MOIs and 8 hours postinfection the ratio of infected cells among transfected (GFP^+) and nontransfected (GFP^-) cells was determined (**Figure 5a,b**). Infected cells were identified using an antibody against the viral envelope glycoprotein C of HSV1.

As shown in **Figure 5b**, a two to fourfold inhibitory effect was observed at MOIs of 0.1 and 0.5. However, this effect disappeared (compared to negative control) by a MOI of 1. The HSV1m1 meganuclease, which was found to be inefficient in a previous

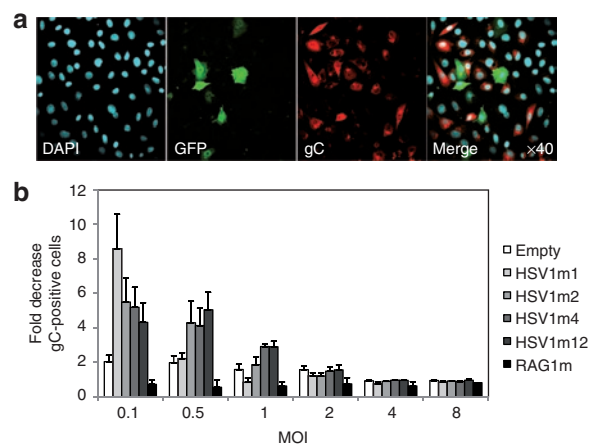


Figure 5 Meganuclease-mediated inhibition of infection by a wild-type herpes simplex virus type 1 virus (HSV1) virus in BSR and COS-7 cells. **(a)** BSR cells were cotransfected with green fluorescent protein (GFP) and meganuclease expression plasmids. Two days later, cells were infected with wtHSV1 at an MOI of 1 and 8 hours postinfection cells were fixed and analyzed by immunocytochemistry. Infected cells were detected with an antibody recognizing the gC viral glycoprotein while GFP was used as a marker for meganuclease transfection. **(b)** Fluorescence-activated cell sorting analysis was used to monitor the number of $GFP^+ HSV1^+$, $GFP^+ HSV1^-$, $GFP^- HSV1^+$, and $GFP^- HSV1^-$ cells. The level of infection inhibition was determined as the following ratio: (number of $HSV1^+ GFP^-$ cells/number of GFP^- cells)/ (number of $HSV1^+ GFP^+$ cells/number of GFP^+ cells). Results are the average of six (HSV1m2 and control) or three (others) independent experiments. DAPI, 4',6-diamidino-2-phenylindole.

assay (**Figure 3a**), displayed a high efficacy in this assay at a MOI of 0.1. This discrepancy can be attributed to a difference in the sensitivity of the assays: indeed, the β -galactosidase assay used in **Figure 3a,c** seems to be less sensitive than the quantitative PCR assay (compare **Figure 3c,d**, at high MOIs), and the immunodetection assay (compare **Figures 3c,5b** at MOI = 0.1). It is unclear why, with the immunodetection assay, HSV1m1 has the highest impact at a MOI of 0.1. However, in accordance with previous results, this meganuclease appears to be less potent as its activity is no longer detected at a MOI of 0.5 (**Figure 5b**).

In addition, the effects observed appear to be specific to anti-HSV1 meganucleases since the introduction of an unrelated meganuclease, RAG1m, resulted in no detectable antiviral activity. One should also note that at a low MOI, a small but reproducible effect was observed even with the negative control, suggesting that transfection of a GFP plasmid might have an effect by itself, maybe by transiently disrupting cellular membrane metabolism.

Anti-HSV1 meganucleases induce high rates of mutations at their target site

The inhibition of HSV1 infection by the anti-HSV1 meganucleases is thought to be due to cleavage of viral DNA. Cleavage of an episomal sequence can result in its degradation and loss, but also in its repair by the endogenous maintenance systems of the cell. DNA double strand breaks can be repaired by homologous recombination or by nonhomologous end joining, two alternative mechanisms.^{33–35} Furthermore, there is an error prone nonhomologous end joining pathway that results mostly in small insertions or deletions (indels) at the cleavage site.^{34,36–38} Thus, upon

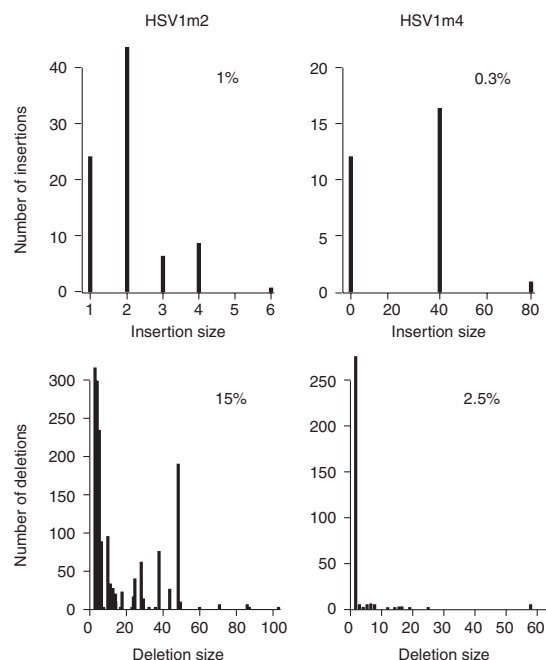


Figure 6 Distribution and frequency of meganuclease-induced deletions and insertions (indels) in the recombinant herpes simplex virus type 1 virus (rHSV1) genome after treatment with HSV1m2 and HSV1m4. 10,356 PCR products were sequenced for HSV1t2 and, 12,228 for HSV1t4. The distribution and frequency of observed deletions or insertions is indicated. We also sequenced 23,527 PCR products for HSV1t2 and 16,961 for HSV1t4, in the absence of meganuclease treatment and found 12 events for HSV1t2 (0.05%) and, no indel for HSV1t4.

meganuclease inhibition of infection, the remaining viral genomes should display a detectable level of indels at the cleavage site.

We amplified by PCR the HSV1t2 and HSV1t4 DNA regions from cells treated with the HSV1m2 and HSV1m4 meganucleases or empty vector, and used deep sequencing to characterize individual PCR products (**Figure 6**). In the absence of meganuclease, indels were absent or barely detectable, with no observed events for HSV1t4, and 0.05% for HSV1t2. However, mutation frequencies increased up to 2.8% in samples transfected with 5 μ g of the HSV1m4 expression vector, and 16% in samples treated with 5 μ g of the HSV1m2 expression vector. In both cases, deletions largely outnumbered insertions (2.5% of deletions versus 0.3% of insertions with HSV1m4, 15% of deletions and 1% of insertions with HSV1m2). As shown in **Figure 6**, there was also a strong bias in favor of small deletions. These results confirm that a very high rate of cleavage occurs at the meganuclease cleavage sites, and are consistent with a mechanism of inhibition based on viral DNA cleavage.

DISCUSSION

In this report, we show that the expression of meganucleases cleaving sequences from an HSV1 virus can significantly prevent HSV1 infection in cultured cells. Three different and nonmutually exclusive mechanisms of action could explain these results. First, cleavage followed by double strand break repair by the error prone nonhomologous end joining pathway could result in mutations that would alter essential functions of the virus and prevent

its efficient replication. The presence of meganuclease-induced mutations is indeed demonstrated by our results (**Figure 6**). As the meganuclease sites are present in coding sequences, the majority of the resulting deletions/insertions result in either frameshifts or inactivating mutations. However, it cannot be excluded that in certain cases mutations will occur that render the virus resistant to meganuclease cleavage without affecting viral fitness. Further studies examining the effect of individual mutations will be necessary to address this question. Nonetheless, it is unlikely that meganuclease-induced mutations are responsible for the largest part of infection inhibition as the I-SceI meganuclease, which is one of the most potent infection inhibitors, cleaves the LacZ reporter cassette, which is not involved in viral functions.

Second, virus cleavage could be followed by DNA degradation and clearance of the cleaved molecules. Inhibition of infection would then result from a loss of viral sequences. An unrepaired double strand break is a deadly event for a cell, for the only DNA ends that can be stably maintained in the cell, without the loss or degradation of genetic material, are the telomeres.³⁹ Cells have evolved a series of maintenance pathways responsible for the repair of such damages.^{33,35} However, these repair pathways could be overwhelmed, being unable to respond to strong cleavage of the target sequence, and opening the way to the extensive resection of double strand breaks by exonucleases, as was observed for yeast chromosomes.⁴⁰ The degradation or disappearance of non-replicative episomal sequences upon cleavage by an endonuclease has been previously documented,⁴¹ and it is extremely likely that a similar effect would be observed with a latent, nonreplicative HSV1 form. Third, cleavage could interfere with viral replication, which is disconnected from the host cell's cycle. Upon infection, the viral genome adopts a circular conformation, and is replicated by a rolling circle mechanism, that is not compatible with a linearized viral form.⁴² Inhibition of infection would then result from a lack of amplification, or at least a decrease in the level of amplification of the viral particles initially entering the cell.

Indeed, we found that anti-HSV1 meganucleases could reduce the number of viral genome copies in infected-cells (**Figure 3d**) as well as block viral replication as determined by a plaque reduction assay (**Figure 4a,b**). Thus our findings are consistent with the inhibition of viral replication either by the linearization of replication intermediates or the direct elimination of incoming viral genomes before replication is initiated. These results illustrate a new mechanism of action, based on virus clipping, to inhibit viral infection. Recently, zinc finger nucleases, another class of very specific endonucleases, have been shown to cleave hepatitis B virus sequences present on a plasmid substrate.⁴³ Our study expands these initial results by demonstrating the potential of meganucleases to inhibit an intact, replicating virus, and thus provide an essential proof of concept toward the use of endonucleases as antiviral agents.

In our study, efficient inhibition could be observed with MOIs ranging from 10^{-3} to 10^{-1} , but not at higher doses. The first steps of viral reactivation within the TG likely involve only a very small number of neurons, as shown by experimental studies.⁴⁴ However, HSV1 reactivation is a rare and self-limited event. In mice, spontaneous reactivation of the virus occurs at any given moment in about one neuron per 10 TG⁴⁴ leading to an estimation of one reactivation for about 50,000 latently infected neurons. Furthermore,

most of the patients with recurrent corneal herpetic disease display only dendritic epithelial keratitis,^{1,4} *i.e.*, a self-limited infection of the superficial layer of the cornea, while major necrosis of the entire cornea is very rare. Taken together, these data suggest that the early steps of clinical herpetic relapse are probably the consequence of only a few infectious particles emerging from TG nerve endings into corneal tissues. As a consequence, a sentinel antiviral process localized in the corneal cells, could block the viral infection in the early stages, and thus prevent the extension of a huge and devastating productive cycle of viral replication in the ocular tissues.

However, quantitative studies in relevant animal models will be necessary to determine if the levels of inhibition obtained with anti-HSV1 meganucleases are sufficient to have a significant impact on the clinical course of infection.

One of the limits to this type of strategy is the toxicity that could result from long-term expression of the meganucleases in the target cells. Sustained expression can result in off-site cleavage potentially resulting in mutations or gross chromosomal rearrangements.⁴⁵ One way to circumvent this kind of effect would be to control the expression of the meganuclease, and limit it to a critical period of time. Inducible meganucleases, resulting from fusions with the ligand binding domain of nuclear receptors have already been described^{46–48} and could be used for this purpose.

The use of an efficient delivery method is another prerequisite for the effective use of meganucleases as an antiviral agent. Vectorization of internal structures such as the TG might prove more difficult than delivery into external tissues or organs, such as lips and superficial ocular tissues. However, even in easily accessible tissues, the ability of meganucleases to be effectively used as a stand-alone treatment will need to be evaluated since current vectorisation techniques do not permit *in vivo* vectorisation of a large majority of cells. Alternatively, meganucleases could be combined with classic antiviral treatments to give a potentially synergistic effect that would provide more potent protection than current approaches. This strategy is particularly attractive for the prevention of recurrent infections during corneal transplantation. The introduction of meganucleases into the graft before transplantation may provide the additional protection necessary to successfully address the low amplitude of the initial reinfection steps. An *ex vivo* animal model is currently being established in order to test the effectiveness of these approaches to inhibit infection directly in corneal tissue.

MATERIALS AND METHODS

Meganuclease engineering. The engineering methods used have been described before.^{17,26–28} To create single chain molecules, the two monomers forming the most active heterodimer were selected, we added mutations K7E, K96E and E8K, E61R to the first and second monomer respectively, and the two monomers were fused as described in Grizot *et al.*²⁸

Evaluation of meganuclease activity in an extrachromosomal assay in mammalian cells. CHO-K1 cells were transfected with Polyfect transfection reagent (QIAGEN, Courtabouef, France) in 96-well plates. 150 ng of target vector was cotransfected either with empty vector or with increasing quantities of a meganuclease expression plasmid (0.75 to 6.25 ng). Total amount of transfected DNA was completed to 175 ng with empty vector. Seventy-two hours after transfection, cells were treated with 150 μ l of β -galactosidase lysis/revelation buffer (typically for 1l of buffer we use 100 ml of lysis buffer (10 mmol/l Tris-HCl pH 7.5, 150 mmol/l NaCl, 0.1 %

Triton X-100, 0.1 mg/ml bovine serum albumin, protease inhibitors), 10 ml of Mg 100 $\times$ buffer (100 mmol/l MgCl₂, 35% β -mercaptoethanol), 110 ml ortho-nitrophenyl- β -galactopyranoside 8 mg/ml and 780 ml of sodium phosphate 0.1 mol/l pH 7.5). After incubation at 37°C, OD was measured at 420 nm. The entire process is performed on an automated Velocity11 BioCel platform (Velocity11, Menlo Park, CA).

Cell survival assay. CHO-K1 cells were transfected with various amounts of meganuclease expression vectors and a constant amount of GFP-encoding plasmid in 96-well plates using Polyfect transfection reagent (QIAGEN). GFP levels were monitored on day 1 and 6 after transfection, by flow cytometry. Cell survival is expressed as a percentage and was calculated as a ratio: meganuclease-transfected cell expressing GFP on day 6/control transfected cell expressing GFP on day 6, corrected for the transfection efficiency determined on day 1.

Construction of rHSV1 recombinant virus. The HSV1 strain (F1) was purchased from American Type Culture Collection (ATCC, Teddington, UK). Viruses were propagated on immortalized COS-7 cells, obtained from ATCC. Recombinant virus (rHSV1) was generated in a manner similar to that previously described.⁴⁹ An ~4.6 kb PstI-BamHI viral genomic fragment was cloned into pUC19. Based on HSV1 sequence from the database (GenBank NC_001806) this represents nucleotides 118869–123461 and 7502–2910 in the TR_L and IR_L repeats of the HSV1 genome. A cassette containing the cytomegalovirus promoter driving LacZ expression was introduced into a 19 bp SmaI/HpaI deletion. The I-SceI cleavage site (TAGGGATAACAGGGTAAT) was inserted after the cytomegalovirus promoter and before the ATG of the LacZ gene. This construct was linearized by XmnI digestion and 2 μ g was cotransfected with 10 μ g of HSV1 genomic DNA into COS-7 cells using Lipofectamine 2000 (Invitrogen, Pontoise, France). After 3 days, infected cells were harvested, sonicated and used to infect a COS-7 monolayer. Cells were overlaid with 1% agarose medium and 3 days later, 300 μ g/ml of X-gal was added to the overlay. β -galactosidase positive “blue” plaques were picked and subjected to three rounds of plaque purification.

Molecular combing, analysis of rHSV1 genome. 4 $\times$ 10⁵ rHSV1 viral particles were embedded in 1% low melting point agarose and lysed in 1% sodium dodecyl sulfate-0.5 mol/l EDTA (pH 8.0) solution at 50°C for 30 minutes. After three washing steps in 0.5 mol/l EDTA (pH 8.0) buffer, agarose plug-embedded viral particles were digested by overnight incubation at 50°C with 2 mg/ml Proteinase K (Eurobio, Les Ulis, France) in digestion buffer (0.5 mol/l EDTA (pH 8.0)). Final DNA extraction was achieved by β -agarase I digestion (1.5 units; New England Biolabs, Ipswich, MA) in 0.5 mol/l of 2-(N-morpholino)ethanesulfonic acid (pH 5.5) at 42°C for up to 16 hours. The DNA solution was then poured into a Teflon reservoir and combed on silanized glass cover slips as described previously.³¹ Probes were labeled with digoxigenin-11-dUTP (LacZ), biotin-11-dCTP (H1, H3, and H5) or Alexa488-7-OBEA-dCTP (H2, H4, and H6) by random priming with BioPrime DNA kit (Invitrogen) according to the manufacturer's instructions. Hybridization and multilayer antibodies revelation was performed as described previously.³¹ For data evaluation, only fluorescent *in situ* hybridization signal arrays composed of a continuous signal chain of Alexa 594-fluorescence for H1 and H3 and Alexa 488-fluorescence for H2 were considered.

Transfection of COS-7 cells and infection by rHSV1. COS-7 cells were seeded in six well plates at 2 $\times$ 10⁵ cells per well and incubated overnight at 37°C in complete growth medium. The next day, cells were transfected with 0.3 or 5.0 μ g of meganuclease expression plasmid or empty vector using Lipofectamine 2000 (Invitrogen). In these conditions, the transfection efficiency was greater the 90% with 5.0 μ g and 55% with 0.3 μ g of plasmid DNA. Twenty-four hours post-transfection, cells were infected with rHSV1 at a MOI from 10⁻³ to 1. Cells were harvested 24 hours after infection and β -galactosidase activity was assayed on a total of 1.0 $\times$ 10³ or

2.5×10^4 rHSV1 infected-cells using a luminescent β -galactosidase assay (Beta-Glo assay, Promega, Charbonnières-les-Bains, France).

Western blot analysis. COS-7 cells were harvested 24 hours after transfection and lysed using RIPA buffer (Santa Cruz Biotechnology, CA). Twenty micrograms of protein were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transferred to a nitrocellulose membrane. Membranes were incubated with a primary rabbit polyclonal antibody against I-CreI (1:20,000 dilution) followed by a secondary goat anti-rabbit antibody conjugated with horseradish peroxidase (1:5,000 dilution). Specific protein bands were visualized with a chemiluminescence reagent (Santa Cruz Biotechnology). A rabbit antibody against β -tubulin (Cell Signaling Technology, MA) was used for the loading control.

Quantification of viral DNA. Total DNA (COS-7 and HSV1 genomes) from transfected and infected COS-7 cells was extracted and purified using DNeasy Blood and Tissue Kit (QIAGEN). Real-time PCR was performed using primers specific to the gB gene in the HSV1 genome (forward primer: 5'-AGAAAGCCCCATTGGCCAGGTAGT and reverse primer: 5'-ATTCTCTCCGACGCCATATCCACCAC) normalized to COS-7 DNA using primers specific to the glyceraldehyde-3-phosphate dehydrogenase gene (forward primer: 5'-GGCAGAACCCGGGTTTATACTGTC and reverse primer: 5'-CCAGTCCTGGATGAGAAAGG). The PCR was carried out using SYBR Premix Ex taq (TaKaRa, Saint-Beauzire, France) and PCR amplification included initial denaturation at 95°C for 5 minutes, followed by 40 cycles of 95°C for 15 seconds, 60°C for 15 seconds, and 72°C for 30 seconds.

Plaque assay. To visualize the infection by rHSV1, transfected COS-7 cells (4×10^5) were seeded in 60-mm dishes. Twenty-four hours post-transfection, cells were incubated with rHSV1 at a MOI of 10^{-3} . One hour after infection, cells were washed once with medium without serum to remove input virus particles that did not enter cells, overlaid with 1% agarose growth medium and incubated at 37°C. At 2 days postinfection, cells were overlaid with 1% agarose medium containing 0.5% of X-gal to visualize the β -galactosidase positive “blue” plaques.

Transfection of BSR cells and infection by wild-type SC16 strain of HSV1. The wild-type SC16 strain of HSV1 was propagated on BSR cells, as previously described.⁵⁰ BSR cells were seeded in 6-well plates at 2×10^5 cells per well and incubated overnight at 37°C in complete growth medium. The next day, cotransfections with 1.5 μ g of plasmid expressing an anti-HSV1 meganuclease and 1.5 μ g of plasmid expressing GFP were performed using Lipofectamine 2000 (Invitrogen). In these conditions, the transfection efficiency was 60%. After 48 hours, cells were infected with the wild-type SC16 strain of HSV1 at a MOI varying from 0.1 to 8. Eight hours postinfection cells were fixed with 2% paraformaldehyde at RT for 15 minutes. For analysis by microscopy, cells were permeabilized with 0.2% Triton X-100 and treated with 4',6-diamidino-2-phenylindole (DAPI). Immunostaining was performed with (i) 1:500 dilution of rabbit antibodies against glycoprotein C (gC) of HSV1 (Clinisciences) and (ii) 1:500 dilution of phycoerythrin conjugated goat anti-rabbit IgG secondary antibodies (Santa Cruz Biotechnology). Cells were analyzed either with an EPICS ELITE ESP (Beckman-Coulter, Villepinte, France) Cytometer or a Leica DMR confocal microscope with a $\times 40$ objective. The analysis by cytometer was performed on a total of 1×10^4 cells.

PCR and sequencing analysis of rHSV1 genome. COS-7 cells were transfected with 5 μ g of meganuclease expressing plasmids, and infected 24 hours later with rHSV1 at a MOI of 10^{-3} . Twenty-four hours after infection total DNA (COS-7 and rHSV1 genomes) was recovered and analyzed using a pair of tag-forward and biotin-labeled reverse PCR primers recognizing a part of the rHSV1 genome corresponding to sequences recognized by meganucleases HSV1m2 and m4. PCR reactions were run for 30 cycles

with an annealing temperature of 67°C. Gel purified PCR products were sent to GATC Biotech (Mullhouse, France) for sequence analysis.

Statistical analysis. A Kruskal–Wallis rank test was used for the immunodetection assay. Student's *t*-test was used for all other analyses.

ACKNOWLEDGMENTS

This work is part of the project ACTIVE that is supported by the Strategic Industrial Innovation program of OSEO (OSEO-ISI). S.G., S.A., D.L., I.C.-S., C.J., C.D., F.C., F.P., and J.S. are full time employees of Collectis or Collectis Genome Surgery. C.M., S.B., and E.C. are full time employees of Genomic Vision.

REFERENCES

- Liesegang, TJ (2001). Herpes simplex virus epidemiology and ocular importance. *Cornea* **20**: 1–13.
- Liesegang, TJ (1989). Epidemiology of ocular herpes simplex. Natural history in Rochester, Minn, 1950 through 1982. *Arch Ophthalmol* **107**: 1160–1165.
- Liesegang, TJ, Melton, LJ 3rd, Daly, PJ and Ilstrup, DM (1989). Epidemiology of ocular herpes simplex. Incidence in Rochester, Minn, 1950 through 1982. *Arch Ophthalmol* **107**: 1155–1159.
- Labetoulle, M, Auquier, P, Conrad, H, Crochard, A, Daniloski, M, Boué, S *et al.* (2005). Incidence of herpes simplex virus keratitis in France. *Ophthalmology* **112**: 888–895.
- Group THEDS (1998). Acyclovir for the prevention of recurrent herpes simplex virus eye disease. Herpetic Eye Disease Study Group. *N Engl J Med* **339**: 300–306.
- Léger, F, Vital, C, Négrier, ML and Bloch, B (2001). Histologic findings in a series of 1,540 corneal allografts. *Ann Pathol* **21**: 6–14.
- Lomholt, JA, Baggesen, K and Ehlers, N (1995). Recurrence and rejection rates following corneal transplantation for herpes simplex keratitis. *Acta Ophthalmol Scand* **73**: 29–32.
- Cockerham, GC, Bijwaard, K, Sheng, ZM, Hidayat, AA, Font, RL and McLean, IW (2000). Primary graft failure: a clinicopathologic and molecular analysis. *Ophthalmology* **107**: 2083–90;discussion 2090.
- Knickerbein, JE, Hendricks, RL and Charukamnoetkanok, P (2009). Management of herpes simplex virus stromal keratitis: an evidence-based review. *Surv Ophthalmol* **54**: 226–234.
- Porteus, MH and Carroll, D (2005). Gene targeting using zinc finger nucleases. *Nat Biotechnol* **23**: 967–973.
- Carroll, D (2008). Progress and prospects: zinc-finger nucleases as gene therapy agents. *Gene Ther* **15**: 1463–1468.
- Galettto, R, Duchateau, P and Pâques, F (2009). Targeted approaches for gene therapy and the emergence of engineered meganucleases. *Expert Opin Biol Ther* **9**: 1289–1303.
- Pâques, F and Duchateau, P (2007). Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. *Curr Gene Ther* **7**: 49–66.
- Chevalier, BS and Stoddard, BL (2001). Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res* **29**: 3757–3774.
- Choulitka, A, Perrin, A, Dujon, B and Nicolas, JF (1995). Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol Cell Biol* **15**: 1968–1973.
- Rouet, P, Smih, F and Jasin, M (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol* **14**: 8096–8106.
- Smith, J, Grizot, S, Arnould, S, Duclert, A, Epinat, JC, Chames, P *et al.* (2006). A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res* **34**: e149.
- Steuer, S, Pingoud, V, Pingoud, A and Wende, W (2004). Chimeras of the homing endonuclease PI-SceI and the homologous *Candida tropicalis* intein: a study to explore the possibility of exchanging DNA-binding modules to obtain highly specific endonucleases with altered specificity. *ChemBiochem* **5**: 206–213.
- Ashworth, J, Havranek, JJ, Duarte, CM, Sussman, D, Monnat, RJ Jr, Stoddard, BL *et al.* (2006). Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature* **441**: 656–659.
- Chen, Z and Zhao, H (2005). A highly sensitive selection method for directed evolution of homing endonucleases. *Nucleic Acids Res* **33**: e154.
- Eastberg, JH, McConnell Smith, A, Zhao, L, Ashworth, J, Shen, BW and Stoddard, BL (2007). Thermodynamics of DNA target site recognition by homing endonucleases. *Nucleic Acids Res* **35**: 7209–7221.
- Rosen, LE, Morrison, HA, Masri, S, Brown, MJ, Springstubb, B, Sussman, D *et al.* (2006). Homing endonuclease I-CreI derivatives with novel DNA target specificities. *Nucleic Acids Res* **34**: 4791–4800.
- Silva, GH, Belfort, M, Wende, W and Pingoud, A (2006). From monomeric to homodimeric endonucleases and back: engineering novel specificity of LAGLIDADG enzymes. *J Mol Biol* **361**: 744–754.
- Grizot, S, Epinat, JC, Thomas, S, Duclert, A, Rolland, S, Pâques, F *et al.* (2010). Generation of redesigned homing endonucleases comprising DNA-binding domains derived from two different scaffolds. *Nucleic Acids Res* **38**: 2006–2018.
- Redondo, P, Prieto, J, Muñoz, IG, Alibés, A, Stricher, F, Serrano, L *et al.* (2008). Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. *Nature* **456**: 107–111.
- Arnould, S, Perez, C, Cabaniols, JP, Smith, J, Gouble, A, Grizot, S *et al.* (2007). Engineered I-CreI derivatives cleaving sequences from the human XPC gene

- can induce highly efficient gene correction in mammalian cells. *J Mol Biol* **371**: 49–65.
27. Arnould, S, Chames, P, Perez, C, Lacroix, E, Duclert, A, Epinat, JC *et al.* (2006). Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. *J Mol Biol* **355**: 443–458.
 28. Grizot, S, Smith, J, Daboussi, F, Prieto, J, Redondo, P, Merino, N *et al.* (2009). Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease. *Nucleic Acids Res* **37**: 5405–5419.
 29. Maeder, ML, Thibodeau-Beganny, S, Osiaik, A, Wright, DA, Anthony, RM, Eichinger, M *et al.* (2008). Rapid “open-source” engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell* **31**: 294–301.
 30. Pruett-Miller, SM, Connelly, JP, Maeder, ML, Joung, JK and Porteus, MH (2008). Comparison of zinc finger nucleases for use in gene targeting in mammalian cells. *Mol Ther* **16**: 707–717.
 31. Michalet, X, Ekong, R, Fougereux, F, Rousseaux, S, Schurra, C, Hornigold, N *et al.* (1997). Dynamic molecular combing: stretching the whole human genome for high-resolution studies. *Science* **277**: 1518–1523.
 32. Sacchetti, A, El Sewedy, T, Nasr, AF and Alberti, S (2001). Efficient GFP mutations profoundly affect mRNA transcription and translation rates. *FEBS Lett* **492**: 151–155.
 33. Hartlerode, AJ and Scully, R (2009). Mechanisms of double-strand break repair in somatic mammalian cells. *Biochem J* **423**: 157–168.
 34. Guirouilh-Barbat, J, Rass, E, Plo, I, Bertrand, P and Lopez, BS (2007). Defects in XRCC4 and KU80 differentially affect the joining of distal nonhomologous ends. *Proc Natl Acad Sci USA* **104**: 20902–20907.
 35. Pâques, F and Haber, JE (1999). Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* **63**: 349–404.
 36. Perez, EE, Wang, J, Miller, JC, Jouvenot, Y, Kim, KA, Liu, O *et al.* (2008). Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* **26**: 808–816.
 37. Doyon, JB, Pattanayak, V, Meyer, CB and Liu, DR (2006). Directed evolution and substrate specificity profile of homing endonuclease I-SceI. *J Am Chem Soc* **128**: 2477–2484.
 38. Liang, F, Han, M, Romanienko, PJ and Jasin, M (1998). Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proc Natl Acad Sci USA* **95**: 5172–5177.
 39. O’Sullivan, RJ and Karlseder, J (2010). Telomeres: protecting chromosomes against genome instability. *Nat Rev Mol Cell Biol* **11**: 171–181.
 40. Lee, SE, Moore, JK, Holmes, A, Umez, K, Kolodner, RD and Haber, JE (1998). *Saccharomyces* Ku70, mre11/rad50 and RPA proteins regulate adaptation to G2/M arrest after DNA damage. *Cell* **94**: 399–409.
 41. Liang, F and Jasin, M (1996). Ku80-deficient cells exhibit excess degradation of extrachromosomal DNA. *J Biol Chem* **271**: 14405–14411.
 42. Roizman, B and Knipe, D (2001). Herpes simplex viruses and their replication. In: Knipe, D, Howley, P, Griffin, D. *et al.* (eds) *Fields Virology*, 4th edn., vol. 72. Raven Publishers: Philadelphia, pp. 2399–2459.
 43. Cradick, TJ, Keck, K, Bradshaw, S, Jamieson, AC and McCaffrey, AP (2010). Zinc-finger nucleases as a novel therapeutic strategy for targeting hepatitis B virus DNAs. *Mol Ther* **18**: 947–954.
 44. Feldman, LT, Ellison, AR, Voytek, CC, Yang, L, Krause, P and Margolis, TP (2002). Spontaneous molecular reactivation of herpes simplex virus type 1 latency in mice. *Proc Natl Acad Sci USA* **99**: 978–983.
 45. Brunet, E, Simsek, D, Tomishima, M, DeKelver, R, Choi, VM, Gregory, P *et al.* (2009). Chromosomal translocations induced at specified loci in human stem cells. *Proc Natl Acad Sci USA* **106**: 10620–10625.
 46. Wehrkamp-Richter, S, Degroote, F, Laffaire, JB, Paul, W, Perez, P and Picard, G (2009). Characterisation of a new reporter system allowing high throughput in planta screening for recombination events before and after controlled DNA double strand break induction. *Plant Physiol Biochem* **47**: 248–255.
 47. Soutoglou, E, Dorn, JF, Sengupta, K, Jasin, M, Nussenzweig, A, Ried, T *et al.* (2007). Positional stability of single double-strand breaks in mammalian cells. *Nat Cell Biol* **9**: 675–682.
 48. Bennardo, N, Cheng, A, Huang, N and Stark, JM (2008). Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet* **4**: e1000110.
 49. Lachmann, RH and Efsthathiou, S (1997). Utilization of the herpes simplex virus type 1 latency-associated regulatory region to drive stable reporter gene expression in the nervous system. *J Virol* **71**: 3197–3207.
 50. Labetoulle, M, Kucera, P, Ugolini, G, Lafay, F, Frau, E, Offret, H *et al.* (2000). Neuronal propagation of HSV1 from the oral mucosa to the eye. *Invest Ophthalmol Vis Sci* **41**: 2600–2606.

Annexe 3

Structural Variability of the Herpes Simplex Virus 1 Genome *In Vitro* and *In Vivo*

Charlotte Mahiet, Ayla Ergani, Nicolas Huot, Nicolas Alende, Ahmed Azough, Fabrice Salvaire, Aaron Bensimon, Emmanuel Conseiller, Simon Wain-Hobson, Marc Labetoulle and Sébastien Barradeau
J. Virol. 2012, 86(16):8592. DOI: 10.1128/JVI.00223-12.
Published Ahead of Print 6 June 2012.

Updated information and services can be found at:
<http://jvi.asm.org/content/86/16/8592>

These include:

REFERENCES

This article cites 69 articles, 39 of which can be accessed free at: <http://jvi.asm.org/content/86/16/8592#ref-list-1>

CONTENT ALERTS

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), [more»](#)

Information about commercial reprint orders: <http://journals.asm.org/site/misc/reprints.xhtml>
To subscribe to to another ASM Journal go to: <http://journals.asm.org/site/subscriptions/>

Structural Variability of the Herpes Simplex Virus 1 Genome *In Vitro* and *In Vivo*

Charlotte Mahiet,^{a,b} Ayla Ergani,^c Nicolas Huot,^c Nicolas Alende,^a Ahmed Azough,^a Fabrice Salvaire,^a Aaron Bensimon,^a Emmanuel Conseiller,^a Simon Wain-Hobson,^b Marc Labetoulle,^{c,d} and Sébastien Barradeau^a

Genomic Vision, Bagneux, France^a; Molecular Retrovirology Unit, Institut Pasteur, Paris, France^b; Laboratoire de Virologie Moléculaire et Structurale, Centre National de la Recherche Scientifique, Gif-sur-Yvette, France^c; and Ophthalmology Department, Hôpital Bicêtre, APHP, Université Paris-Sud, Le Kremlin-Bicêtre, France^d

Herpes simplex virus 1 (HSV-1) is a human pathogen that leads to recurrent facial-oral lesions. Its 152-kb genome is organized in two covalently linked segments, each composed of a unique sequence flanked by inverted repeats. Replication of the HSV-1 genome produces concatemeric molecules in which homologous recombination events occur between the inverted repeats. This mechanism leads to four genome isomers (termed P, IS, IL, and ILS) that differ in the relative orientations of their unique fragments. Molecular combing analysis was performed on DNA extracted from viral particles and BSR, Vero, COS-7, and Neuro-2a cells infected with either strain SC16 or KOS of HSV-1, as well as from tissues of experimentally infected mice. Using fluorescence hybridization, isomers were repeatedly detected and distinguished and were accompanied by a large proportion of noncanonical forms (40%). In both cell and viral-particle extracts, the distributions of the four isomers were statistically equivalent, except for strain KOS grown in Vero and Neuro-2a cells, in which P and IS isomers were significantly overrepresented. In infected cell extracts, concatemeric molecules as long as 10 genome equivalents were detected, among which, strikingly, the isomer distributions were equivalent, suggesting that any such imbalance may occur during encapsidation. *In vivo*, for strain KOS-infected trigeminal ganglia, an unbalanced distribution distinct from the one *in vitro* was observed, along with a considerable proportion of noncanonical assortment.

Herpes simplex virus 1 (HSV-1) is a pathogen that commonly leads to primary and recurrent orofacial and ocular lesions (32). Epithelial lesions result from massive production of virions with attendant immune responses. Primary infection is followed by the spread of viral particles to sensory neurons, especially those of the trigeminal ganglia, and establishment of a lifelong latent infection. Reactivation can take place sporadically, causing a new productive infection in tissues innervated by the neurons that harbored the reactivation process.

Although HSV-1 seroprevalence in the general population ranges from 60 to 90% depending on the country and age (9), HSV-1 infections are mostly asymptomatic. However, HSV-1 may lead to recurrent labial lesions in 15 to 45% of the adult population (18), encephalitis in one per 500,000 people per year (65), and corneal keratitis with an annual incidence of 30 cases per 100,000 people (30). The genome is a linear ~152-kb DNA molecule that is organized into two covalently linked segments: a long (L; 82% of the genome) and a short (S; 18% of the genome) fragment. Each segment is composed of a unique sequence, U_L and U_S , bounded by inverted repeats referred to as b and c , respectively. They are both flanked by the a inverted repeats, so that the final canonical structure is $ab-U_L-b'a'c'-U_S-ca$, also designated $a-TR_L-U_L-IR_L-a'-IR_S-U_S-TR_S-a$, with TR and IR standing for terminal repeat and internal repeat, respectively (51, 63). During productive infection, HSV-1 DNA replication generates concatemeric DNA more than one unit in length that includes complex branched structures (26, 34, 50, 69). It has long been thought that the linear viral genome circularized and was replicated through a rolling-circle mechanism (7). However, this model is being challenged, as analysis of wild-type HSV-1 DNA on Gardella gels, which allow the differentiation of latent episomal from lytic linear viral DNA on the basis of electrophoretic mobility, has shown that its circularization may not occur after its entry into cells (25).

Furthermore, it was suggested that synthesis of HSV-1 DNA proceeded through recombination-dependent replication (67).

There is clear evidence that recombination is intrinsically linked to replication (16, 37, 45, 64). Recombination events are frequent, as implied by the study of the products of cells coinfecting with two types of HSV-1 genome (8, 21). During replication, homologous recombination may occur between the inverted repeats of the HSV-1 genome (54, 55, 64), resulting in the generation of four isomers differing in the relative orientations of their L and S genomic fragments. They are designated P, for prototype; IL, for inversion of the L fragment; IS, for inversion of the S fragment; and ILS, for inversion of the L and S fragments (19, 44, 51). Homologous recombination is not sequence specific (10, 11, 42, 54–56, 64). Although the a sequence is dispensable for isomerization (33), it seems to play an important role, as it supports recombination with high efficiency despite its short length (11, 54, 61, 62, 64). Moreover, it contains the packaging signal and the cleavage site of concatemeric DNA (35, 38, 56). The alternative cleavage events account for the generation of the four genome isomers (56, 62).

Recombination events seem to occur early during HSV-1 DNA replication, as genome isomers have been observed in concatemeric DNA in the earliest detectable products (69). They are probably initiated by double-strand breaks (45). Moreover, it has been shown that the cellular endonuclease G cleaves in the a sequences

Received 26 January 2012 Accepted 21 May 2012

Published ahead of print 6 June 2012

Address correspondence to Marc Labetoulle, marc.labetoulle@bct.aphp.fr, or Sébastien Barradeau, s.barradeau@genomicvision.com.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.00223-12

(23, 24, 68). Lastly, virion and cellular HSV-1 DNAs present nicks and gaps (3, 17, 26, 43, 52, 66) that could lead to double-strand breaks during the passage of the replication forks, favoring recombination (29).

Recombinant mutants of HSV-1 that are not able to undergo homologous recombination due to the deletion of part or all of the inverted repeats at the junction between the L and S fragments (*a'b'/a'c'*) showed lower efficiency of replication *in vitro* than the corresponding wild-type virus and an absence of acute infection *in vivo* (27, 28, 41, 46). These observations suggest that either the doubled rate of synthesis of proteins encoded by the *b'a'/a'c'* sequences and/or the ability to undergo homologous recombination is required for efficient *in vivo* infection. It has been previously shown that the four genome isomers of strains MP and 17 of HSV-1 are randomly distributed in virion preparations obtained from infected Hep-2 and BHK cells, respectively (12, 13, 19). These data were obtained either by the use of endonucleases that cleave within a ³²P-labeled HSV-1 DNA or by the analysis of partially denatured HSV-1 DNA by electron microscopy. The genome isomers were also detected in concatemeric DNA using techniques such as pulsed-field gel electrophoresis/field inversion gel electrophoresis coupled with HSV-1 DNA endonuclease restriction and Southern blotting (1, 2, 34, 49, 53, 69).

In the present study, we tested the hypothesis that the process of isomerization systematically results in a random distribution no matter what cells or strain is being considered. We applied molecular combing for the direct visualization and analysis of a large number of single HSV-1 DNA molecules through hybridization of uniformly and irreversibly stretched DNA on silanized glass surfaces (36, 48). Molecular combing analysis was performed on stretched DNA extracted from either Vero, COS-7, BSR, or Neuro-2a cell lines infected with the SC16 or KOS strain of HSV-1 or on corneas and trigeminal ganglia from mice experimentally infected with the same viral strains. Using fluorescence hybridization with a set of digoxigenin- or biotin-labeled probes covering the entire HSV-1 genome, surprisingly, we observed that the distribution of the genome isomers in both cell and viral-particle extracts varied significantly from a random distribution depending on the strain and of the cell lines used for production. A similar observation was made *in vivo*. As molecular combing allows the visualization of individual DNA molecules within a heterogeneous population, we also identified concatemers of up to 10 genome equivalents, along with a large proportion of noncanonical forms, many of them showing signs of recombination. Among the large numbers of HSV-1 genomes generated, some probably are noninfectious, indicating that although HSV-1 replication is a costly process, it is not incompatible with efficient transmission.

MATERIALS AND METHODS

Virus strains and cells. Wild-type HSV-1 strains SC16 (20) and KOS (M. Levine, University of Michigan, Ann Arbor, MI) were used. Stocks of HSV-1 SC16 and KOS were prepared and titrated in BSR cells as previously described (31). African green monkey kidney epithelial cell (Vero), African green monkey kidney fibroblast (COS-7), and subclones of baby hamster kidney cell (BSR) and albino mouse neuroblast (Neuro-2a) cell lines were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin and grown at 37°C in a humidified atmosphere with 5% CO₂.

Preparation of viral particles and infected cells. Cells were seeded in 6-well plates at 2×10^5 cells per well and incubated overnight at 37°C in

Dulbecco's modified Eagle's medium. After 24 h, the cells were infected with HSV-1 strain SC16 or KOS at a multiplicity of infection of 0.5. Up to 16 h postinfection, culture supernatants were recovered and centrifuged for 5 min at $5,000 \times g$ (Heraeus 7593 rotor) at room temperature. The viral particles and infected cells embedded in agarose plugs were prepared as previously described (48), except for the viral-particle preparation, which required a 1% (wt/vol) low-melting-point agarose (Nusieve GTG; Cambrex) solution in $1 \times$ phosphate-buffered saline (PBS).

Preparation of mouse trigeminal ganglia and corneas. All procedures were performed in compliance with the Association for Research in Vision and Ophthalmology guidelines for the use of animals in ophthalmic and visual research and the tenets of the European guidelines for protection and welfare of animals. The mice were infected with either HSV-1 strain SC16 (100 PFU per eye) or KOS (10^6 PFU per eye) by corneal scarification, based on preliminary experiments on the optimal inoculum to induce reproducible and nonlethal infection of the cornea and connected neurons. The mice were sacrificed on the sixth day postinoculation. The corneas and trigeminal ganglia were recovered, rinsed three times in $1 \times$ PBS for 15 min at room temperature, and then cut into small pieces. Dissociation of cornea and trigeminal ganglion tissues was performed for 3 h at 37°C in a 0.3-mg/ml collagenase type A (Roche) and 0.8-mg/ml Gibco dispase (Invitrogen) solution and in a 1-mg/ml collagenase type A and 0.25-mg/ml Gibco trypsin (Invitrogen) solution prepared in Gibco $1 \times$ Hanks' balanced salt solution buffer (Invitrogen), respectively. The lysates were pelleted by centrifugation (Heraeus 7593 rotor) at $5,000 \times g$ for 10 min, resuspended in a $1 \times$ PBS solution, and mixed thoroughly at a 1:1 ratio with a 1.2% (wt/vol) solution of low-melting-point agarose prepared in $1 \times$ PBS at 50°C. Ninety microliters of the tissue lysate-agarose mixture was poured into a Chef plug-forming well (Bio-Rad) and left to cool for 30 min at 4°C.

DNA extraction and molecular combing. Extraction of genomic DNA for all samples except viral particles was performed by digestion of the agarose plugs according to the procedure described by Schurra and Bensimon (48). Genomic DNA from HSV-1 particles was extracted either by a phenol-chloroform method, as previously described (4), or in agarose plugs with a slightly modified procedure. Lysis of viral particles was performed by dipping agarose plugs in 1% SDS-0.5 M EDTA, pH 8.0, solution at 50°C for 30 min. After three washing steps in digestion buffer (0.5 M EDTA, pH 8.0) for 10 min at room temperature, the plugs were digested by overnight incubation at 50°C with 2 mg/ml proteinase K (Eurobio) in 250 µl digestion buffer.

Molecular combing was performed using the Molecular Combing System (Genomic Vision S.A.) and Molecular Combing coverslips (20 by 20 mm; Genomic Vision S.A.) as previously described (48). After the combing process, the coverslips were dried for 4 h at 60°C. For direct visualization of combed HSV-1 DNA fibers, the coverslips were mounted with 20 µl of 1/1,000 (vol/vol) Yoyo-1 iodide (Molecular Probes)-Prolong Gold antifade reagent (Invitrogen).

Cloning and labeling of HSV-1 probes. HSV-1 genomic fragments ranging from 1,110 to 9,325 bp were obtained by either SacI or BspEI (New England BioLabs) enzymatic digestion of HSV-1 strain SC16 or by long-range PCR using LR *Taq* DNA polymerase (Roche), a set of HSV-1-specific primers (Table 1), and the DNA from HSV-1 strain SC16 as the template DNA. SacI and BspEI HSV-1 fragments were ligated in the SacI- and XmaI-digested pNEB193 plasmid (New England BioLabs). The PCR products were ligated in the pCR2.1 vector using the Topo TA cloning kit (Invitrogen). Both extremities of each cloned fragment were sequenced with the universal primer M13 forward primer —48, 5'-CGCCAGGGTTTCCCCAGTCACGAC-3', and the M13 reverse primer —48, 5'-AGCGGATAACAATTTTCACACAGGA-3', for verification and compared to the reference HSV-1 genome sequence (GenBank accession number NC_001806). DNA labeling was performed using the conventional random-priming protocol with either biotin-14-dCTP or digoxigenin-11-dUTP and 200 ng of DNA as a template. For biotin-14-dCTP labeling, the BioPrime DNA kit (Invitrogen) was used according to the manufacturer's

TABLE 1 Primer pairs used for cloning of HSV-1^a

Probe	Coordinates	Primer (5'→3')	
		Forward	Reverse
P4	9273–11276	TGGTTGTGTTACTGGGCAAA	TCGATCGACGACACCATAAA
P5	11090–13245	CAGATACGACTCCCGCAGAT	CGACGACCTCGACGTTATTT
P8	73229–77332	AGATGTCCACGAGCACCAG	CCTGACTTTGTGGGGCTAAA

^a The sequences and coordinates of the primers are based on the sequence of the genome of HSV-1 strain 17 (GenBank accession no. NC_001806).

instructions, except that the labeling reaction proceeded overnight. For digoxigenin-11-dUTP, a deoxynucleoside triphosphate (dNTP) mixture composed of 2 mM dATP, 2 mM dCTP, 2 mM dGTP, 1 mM dTTP, and 1 mM digoxigenin-11-dUTP (Roche) was used in place of the dNTP mixture furnished with the BioPrime DNA kit.

Hybridization and detection of HSV-1 probes. Hybridization steps were performed as described by Schurra and Bensimon (48). Detection of labeled probes was carried out using three layers of antibodies in a 1:25 dilution. Biotin-14-dCTP-labeled probes were revealed with Alexa Fluor 594-conjugated streptavidin (Invitrogen) as the first step, followed by incubation with a biotinylated goat anti-streptavidin antibody (Vector Laboratories) and then with Alexa Fluor 594-coupled streptavidin. Digoxigenin-11-dUTP-labeled probes were consecutively revealed with an aminomethylcoumarin acetate (AMCA)-conjugated monoclonal mouse anti-digoxigenin antibody (Jackson ImmunoResearch) and then AMCA-coupled rat anti-mouse (Jackson ImmunoResearch) and Alexa Fluor 350-conjugated goat anti-rat IgG (Invitrogen) as a final step. After the last washing steps, the hybridized coverslips were gradually dehydrated in 70%, 90%, and 100% ethanol solutions and air dried.

Statistical analysis of the distribution of the HSV-1 genome isomers. Yoyo-1 iodide-stained or hybridized-combed DNA coverslips were scanned with an inverted automated epifluorescence microscope equipped with a 40× objective (ImageXpress Micro; Molecular Devices). All images were uniformly contrasted using ImageJ (<http://rsbweb.nih.gov/ij/>) with identical parameters to allow comparison. The lengths of Yoyo-1 iodide-stained DNA fibers and HSV-1-specific hybridization sig-

nals were measured using a custom image analysis system (5) and converted to kilobases using an elongation factor of 2 kb/μm. The minimum number of fluorescent signals analyzed was determined using the central-limit theorem that leads to the following equation: $n = [\hat{e}(1-\hat{e})Zc^2]/m^2$, with Zc as the quantile of the standard normal distribution in a 95% confidence interval, $\hat{e}$ as the estimated proportion of a given isomer, m as the margin of error between the estimated proportion $\hat{e}$ and the actual proportion of a given isomer, and n as the size of the sample. With a tolerated margin of error of 5% and the proportion $\hat{e}$ estimated at 50%, 405 signals were randomly selected and analyzed. The distribution was tested by a chi-square test and accepted when the P value observed was above 0.05. The binomial error was calculated according to the equation $100N_{\text{total}} \times \sqrt{N[1 - (NN_{\text{total}})]}$, where N_{total} is the total number of signals analyzed and N is the number of a given isomer.

RESULTS

Improvement of the DNA extraction method for the detection of HSV-1 genomes in viral particles. In order to characterize the structure and the organization of the HSV-1 genome, we applied molecular combing, a technology that allows the direct visualization and analysis of single DNA molecules at the kilobase level of resolution. Since the HSV-1 DNA is a 152-kb-long genome, high-resolution mapping of the genome requires very high-molecular-weight DNA. First, HSV-1 strain SC16 DNA was extracted from viral particles by the standard phenol-chloroform method and

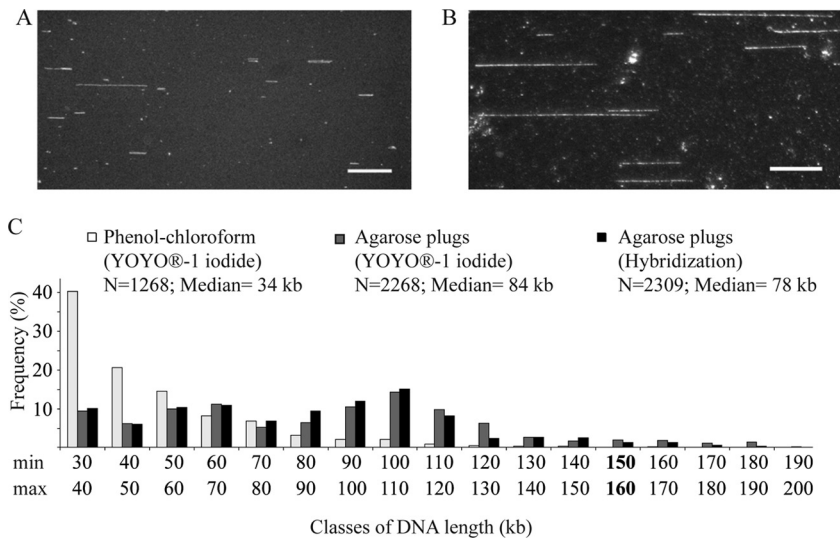


FIG 1 Combed HSV-1 DNA from viral particles. (A) Detection of Yoyo-1 iodide-stained combed DNA of HSV-1 strain SC16 particles extracted by a phenol-chloroform method. The scale bar corresponds to 25 μm (50 kb). (B) Alexa Fluor 594 fluorescence hybridization HSV-1 DNA arrays corresponding to HSV-1 strain SC16 DNA extracted from agarose plugs with virions embedded. Scale bar, 25 μm (50 kb). (C) Size distribution of combed DNA extracted from HSV-1 strain SC16 particles. The number of DNA fibers was classified per bin of 10-kb width, for which the minimum (min) and maximum (max) lengths are indicated. The fiber lengths of Yoyo-1 iodide-stained viral DNA extracted by the phenol-chloroform method ($n = 1,268$; median = 34 kb) and using agarose plugs ($n = 2,268$; median = 84 kb) are depicted by light and dark gray bars, respectively. Black bars correspond to the measurement of the Alexa Fluor 594-fluorescent HSV-1 signals ($n = 2,309$; median = 78 kb). The boldface numbers indicate the bin containing full-length HSV-1 genomes.

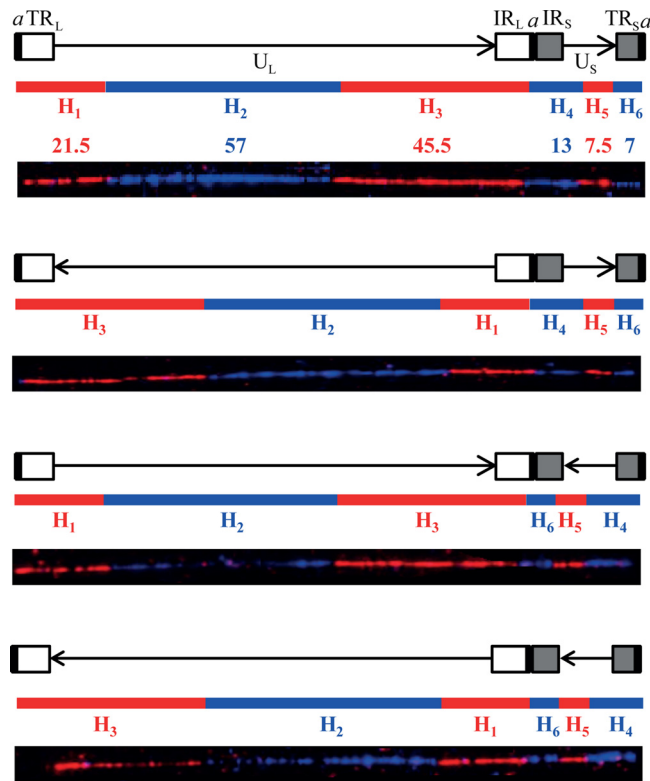


FIG 2 Genomic Morse code for the detection of HSV-1 isomers. Fluorescence hybridization arrays illustrating the HSV isomers (from the top down) P, IL, IS, and ILS are shown below the schemas. The structure of the HSV-1 genome is shown, with the unique long (U_L) and short (U_S) regions (arrows), the TR_L/IR_L (empty boxes) and TR_S/IR_S (gray boxes) repeats, and the a sequences (black boxes). After detection, the apparent HSV-1 genomic Morse code consists of continuous arrays of Alexa Fluor 594 fluorescent H_1 , H_3 , and H_5 probes (red bars) and Alexa Fluor 350/AMCA fluorescent H_2 , H_4 , and H_6 probes (blue bars). The sizes of the fragments shown at the top are given in kilobases according to the reference HSV-1 genome sequence (GenBank accession number NC_001806).

combed. Individual DNA fibers were revealed with the intercalating agent Yoyo-1 iodide (Fig. 1A), and 1,268 stained fibers were then measured. The majority of the fibers had lengths between 15 and 20 μm (30 and 40 kb, using the elongation factor of 2 kb/ μm) (Fig. 1C). With a median size of 17 μm (34 kb), only 0.9% of the measured DNA molecules were over 70 μm (140 kb), indicating that full-length HSV-1 genomes were rare. These results suggest that the majority of DNA molecules are partly fragmented by several factors, for example, as a consequence of uncontrolled forces during phenol-chloroform extraction or the activity of nucleases. DNA extraction from viral particles embedded in agarose plugs, as described by Schurra and Bensimon (48), did not lead to detection of Yoyo-1 iodide-stained DNA molecules (data not shown), suggesting that the HSV-1 particles are not lysed within the agarose plug. We slightly modified the extraction method by decreasing the concentration of low-melting-point agarose to 1% and by performing lysis of viral particles in two consecutive lysis buffers containing 1% SDS (instead of 1% sodium *N*-lauroylsarcosinate) and 2 mg/ml of proteinase K, respectively. As shown in Fig. 1C, the analysis of 2,268 individual DNA molecules revealed a different type of DNA fiber length distribution than did phenol-chloroform extraction—the median size was 42 μm (84 kb), with 4% of

DNA fibers having a length over 70 μm (140 kb). As 2.5% of the DNA fibers had a size greater than 76.5 μm (153 kb), this suggests that the viral-particle preparations were contaminated with cell genomic DNA.

The combed DNA fibers were then cohybridized with a set of 41 digoxigenin-11-dUTP-labeled cloned fragments that covered the entire HSV-1 genome. Continuous Alexa Fluor 594-fluorescent hybridization signals were obtained, showing that the observed signals corresponded to HSV-1 DNA (Fig. 1B). Measurement of 2,309 linear fluorescent signals yielded a median size of 39 μm (78 kb) (Fig. 1C). This time, 3% of the individual fibers were over 70 μm (140 kb). It was noted that 1.5% of the HSV-1 DNA fibers had lengths greater than 76.5 μm (153 kb), suggesting that the viral-particle preparations are partially contaminated with HSV-1 concatemeric DNA from infected cells. In conclusion, the data show that the method of extraction was improved in terms of DNA length, allowing analysis of the HSV-1 genome structure in viral particles by molecular combing.

Detection of the four HSV-1 genome isomers. In order to distinguish the four isomers of the HSV-1 genome, the 41 HSV-1 probes were labeled with either digoxigenin-11-dUTP or biotin-14-dCTP and detected with Alexa Fluor 350/AMCA- and Alexa Fluor 594-coupled antibodies, respectively, to create six groups of probes, referred as H_1 to H_6 , that differ in size and color (Fig. 2). This constitutes the HSV-1-specific genomic Morse code. The order of the 6 probes within the genomic Morse code differs for each isomer; for example, the expected $H_1H_2H_3H_4H_5H_6$ fluorescent array corresponds to a P isomer, while an IL isomer is revealed by the theoretical $H_3H_2H_1H_4H_5H_6$ fluorescent array. Immunofluorescence microscopy of hybridized combed DNA extracted from particles of HSV-1 strain SC16 displayed a large diversity of multicolor linear patterns. Figure 2 shows an example of each full HSV-1 genome isomer that was successfully detected. Nevertheless, most of the canonical signals detected (95%) were found to be fragmented. However, they can be considered for further analysis, since we determined the minimal genomic Morse code pattern necessary to identify an isomer (Table 2). For example, a continuous signal composed of a 10-kb-long AMCA/Alexa Fluor 350 signal followed by a 21.5-kb-long Alexa Fluor 594 fluorescent signal, a 13-kb-long AMCA/Alexa Fluor 350 signal, and a 7.5-kb-long Alexa Fluor 594 fluorescent signal corresponds to a fragment of H_2 followed by $H_1H_4H_5$ and thus can only be interpreted as an IL isomer. On 1,082 measurements, the mean sizes of the H_1 , H_3 , H_4 , and H_6 probes were 21.3 ± 1.46 kb (theoretical 21.5 kb), 45.5 ± 2 kb (theoretical 45.5 kb), 12.7 ± 0.7 kb (theoretical 13 kb), and 7.8 ± 0.5 kb (theoretical 7 kb), respectively, indicating that the designed genomic Morse code detects HSV-1 genomes that structurally correspond to what was theoretically expected. All these observations indicate that the designed genomic Morse code can discriminate the four HSV-1 genome isomers at the single-

TABLE 2 Minimum fluorescence hybridization patterns for the identification of HSV-1 isomers

Isomer	Minimum pattern (size [kb])				
P	H_3 (36)	H_4 (13)	H_5 (3)		
IL	H_2 (10)	H_1 (21.5)	H_4 (13)	H_5 (7.5)	
IS	H_3 (37.5)	H_6 (7)	H_5 (7.5)		
ILS	H_2 (16)	H_1 (21.5)	H_6 (7)	H_5 (7.5)	

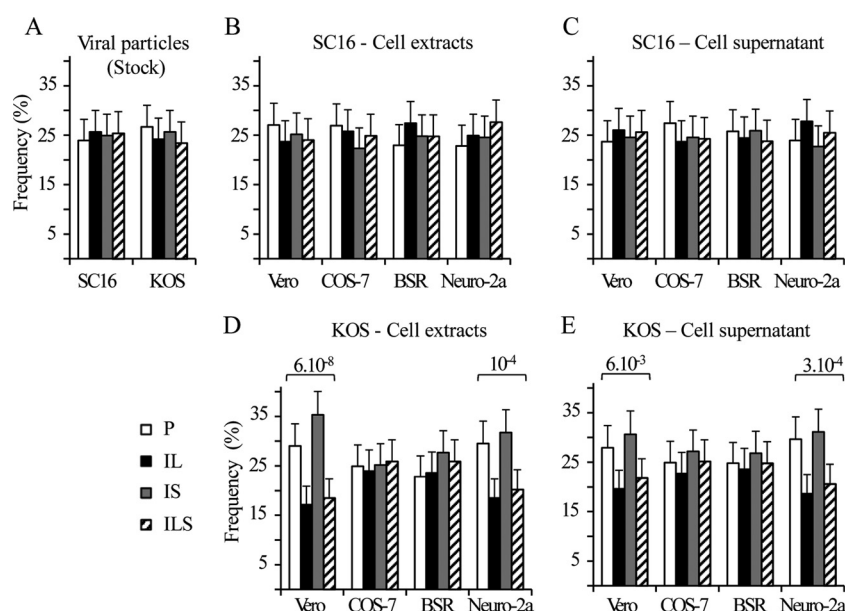


FIG 3 Distribution of the HSV-1 isomers in viral particles and in infected cells. The histograms show the average frequencies of the four HSV-1 genome isomers in SC16 and KOS viral-particle extracts produced in BSR cells (Stock) (A); Vero, Cos-7, BSR, and Neuro-2a cell extracts infected with strain SC16 (B) and strain KOS (D); and the corresponding cell culture supernatants (C and E). The frequencies of the P, IL, IS, and ILS isomers are depicted. *P* values of <0.05 (χ^2 test) are indicated by the brackets. The error bars indicate binomial errors.

molecule level. To obtain the distribution of the isomers from viral-particle extracts of the HSV-1 strains SC16 and KOS produced in BSR cells, 405 fluorescent arrays that fulfilled the criteria for attribution of an isomer were counted. In both cases, the distributions were found to be equivalent (Fig. 3A), with chi-square *P* values of 0.96 and 0.80, respectively. Those viral-particle preparations were designated stocks. They were subsequently used to infect various cell lines and to analyze the HSV-1 isomer distribution during productive infection.

Cell- and strain-dependent distributions of the HSV-1 genome isomers. In order to perform a systematic analysis of the distributions of the four isomers of the HSV-1 genome, stocks of strains SC16 and KOS were used to infect Vero, COS-7, BSR, and Neuro-2a cells at a multiplicity of infection of 0.5. The distributions of the four genome isomers were determined for one genome equivalent from cell extracts and the corresponding supernatants. In the cell extracts, the distribution was analyzed globally, as HSV-1 genomes may have various origins: HSV-1 DNA that has just entered cells, fragmented concatemeric DNA, residues of A capsids that lack the CCSC complex (UL25/UL17 proteins) that failed to retain DNA (60), and encapsidated HSV-1 DNA. Cell supernatants contained the viral particles that had been released from infected cells. As shown in Fig. 1C, the preparations of cell supernatants are slightly contaminated with HSV-1 DNA longer than one genome equivalent. Here again, the analysis of the distribution was performed on the overall population. For strain SC16, both particles and infected cell samples showed that the frequencies of the four HSV-1 isomers were about 25% (Fig. 3B and C), indicating that the distributions were statistically equivalent (chi-square *P* value, between 0.6 and 0.9). For the KOS strain, the frequencies of the four isomers were also about 25% for cell and cell supernatant DNA extracts of the COS-7 and BSR cells. On the other

hand, the P and IS isomers were overrepresented in Vero and Neuro-2a cell extracts, with proportions between 30% and 35%. These differences in the distributions of the four HSV-1 isomers were statistically confirmed, as the *P* values of the chi-square tests were 6×10^{-8} and 10^{-4} with Vero and Neuro-2a cells, respectively (Fig. 3D). A similar statistically unbalanced distribution was noticed in the corresponding viral particles recovered from the culture supernatant. P and IS were again overrepresented, with proportions between 28% and 31%, and the chi-square *P* values were 6×10^{-3} and 3×10^{-4} with Vero and Neuro-2a cells, respectively (Fig. 3E). This means that both P and IS viral genome isomers from HSV-1 genomes in infected cells and from released viral particles are more highly represented than the IL and ILS genome isomers. Thus, the process of isomerization seems not to systematically result in a random distribution.

Detection of intermediates of replication in infected cells. In infected Vero, COS-7, BSR, and Neuro-2a cells, fluorescent signal arrays longer than 152 kb (likely corresponding to intermediates of replication) were detected, in addition to the canonical HSV-1 fluorescent signals (one-genome-equivalent elements with the usual *a*-TR_L-UL_L-IR_L-a'-IR_S-US_S-TR_S-a organization). Fluorescence hybridization was performed using the genomic Morse code labeled with digoxigenin-11-dUTP on combed DNA from noninfected Vero cells and SC16-infected Vero cell DNA extracts. As expected, in noninfected Vero cells, used as a control, no fluorescence hybridization signals were detected, indicating that the probes do not cross-hybridize with host genome DNA and are specific to the HSV-1 genome (data not shown). Continuous Alexa Fluor 594-fluorescence hybridization signals were detected in infected-cell DNA extracts. The measurement of 1,031 signals revealed the presence of HSV-1-positive DNA fibers as long as 1,393 kb. Fibers were grouped into bins of one-genome-equivalent elements

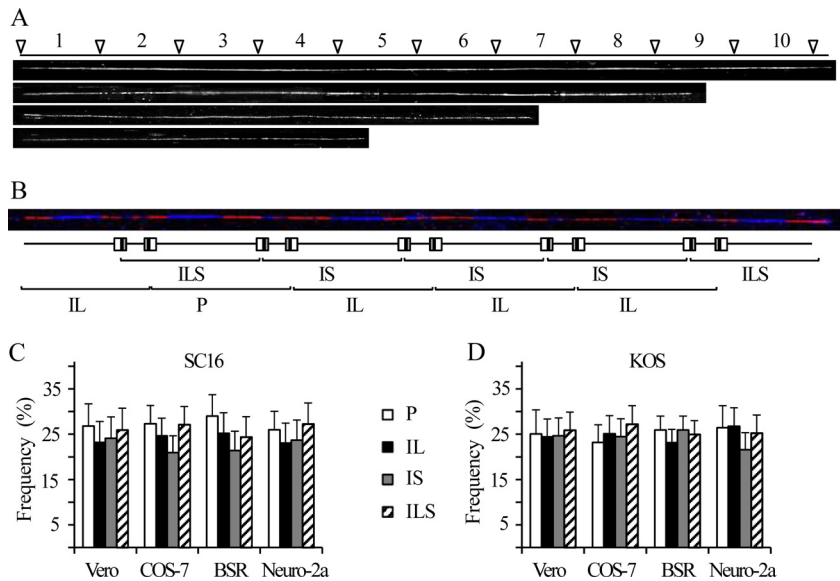


FIG 4 Detection of HSV-1 genome concatemers. (A) Alexa Fluor 594-fluorescence hybridization HSV-1 DNA arrays corresponding to HSV-1 strain SC16 concatemerized DNA molecules extracted from infected Vero cell extracts. The arrowheads delimit genome equivalents (1 genome equivalent is ~152 kb), and the genome equivalent numbers are given between them. (B) Example of a hybridization array corresponding to an intermediate of replication from strain SC16-infected BSR cell extracts. The red Alexa Fluor 594-fluorescence signal corresponds to the H₁, H₃, and H₅ probes, while the blue AMCA/Alexa Fluor 350-fluorescence corresponds to the H₂, H₄, and H₆ probes. Below is a schematic representation of the concatemeric DNA chain. The empty boxes correspond to the *b/b'* inverted repeats, and the gray boxes correspond to the *c/c'* inverted repeats. The *a* sequences are depicted as black boxes. U_L and U_S are represented as black straight lines. The two possible genome isomers that can be generated are indicated below the lines. (C and D) Histograms showing the frequencies of the four HSV-1 isomers in concatemeric HSV-1 of strains SC16 (C) and KOS (D) produced in infected Vero, COS-7, BSR, and Neuro-2a cells. The error bars indicate binomial errors.

(76.5 μm or 153 kb) (Fig. 4A and Table 3). The majority of fibers (62%) was less than 153 kb and 29% were between 153 and 306 kb, corresponding to two genome equivalents, while 3.6% were between 306 and 459 kb long, as expected for three genome equivalents. The remaining 5.4% were between 459 and 1,530 kb, i.e., 3 to 10 genome equivalents. It is notable that fibers of three genome equivalents or more were detectable and were possibly underestimated, considering the breakage during the DNA extraction or the molecular combing processes, probably due to the presence of nicks and gaps reported in replicative HSV-1 DNA (17).

Equivalent distributions of the four HSV-1 isomers in the intermediates of replication. Using the validated genomic Morse code, the isomer distribution was analyzed within the concatemers that were produced in Vero, COS-7, BSR, and Neuro-2a cells

infected with either strain SC16 or KOS of HSV-1. As during the encapsidation process unit-length genomes are cleaved in the *a* sequence, and as it is not possible to orient the DNA molecules, there are two possibilities for cleavage per molecule, generating two possible sets of isomers per molecule (Fig. 4B). The frequencies of the four isomers were about 25%, even for concatemers from strain KOS-infected Vero and Neuro-2a cell DNA extracts (Fig. 4C and D). These observations suggested that the nonequivalent distributions of the one genome equivalents produced in cells or cell supernatants from KOS-infected Vero or Neuro-2a cells likely occur throughout encapsidation rather than during homologous recombination between the inverted repeats surrounding the U_L and the U_S fragments.

Noncanonical HSV-1 genomes are frequent in viral particles. During the analysis of isomer distribution, we noticed several types of unconventional DNA structures. First, there were individual molecules for which the sequence of the H₁ to H₆ probes corresponded to one isomer but with unexpected length or order of at least one of the probes. For example, the 104-kb AMCA/Alexa Fluor 350 array for the first molecule in Fig. 5A is 47 kb longer than the longest expected blue H₂ signal. A second example, typified by the second molecule in Fig. 5A, corresponds to a 75-kb AMCA/Alexa Fluor 350 array disconnected by a 5-kb Alexa Fluor 594 signal. These examples suggest events like internal duplications, deletions, or rearrangements. Third, several fibers showed regular repeats (Fig. 5A) for which a 17.8-kb repetition composed of a 5.8-kb AMCA/Alexa Fluor 350 fluorescence signal and a 12-kb Alexa Fluor 594 signal are observed. Their repetitive structure is similar to those described by Schroder et al. (47) as parts of viral genomes built up from repeats of restricted regions of

TABLE 3 Numbers of HSV-1 genome equivalents in replication intermediates

No. of genome equivalents	Size class (kb)		No. of fibers	Frequency (%)
	Minimum	Maximum		
1	>0	<153	635	61.6
2	>153.2	<306	296	28.7
3	>306.2	<459	37	3.6
4	>459.2	<612	17	1.6
5	>612.2	<765	20	1.9
6	>765.2	<918	6	0.6
7	>918.2	<1071	8	0.8
8	>1071.2	<1224	4	0.4
9	>1224.2	<1377	7	0.7
10	>1377.2	<1530	1	0.1

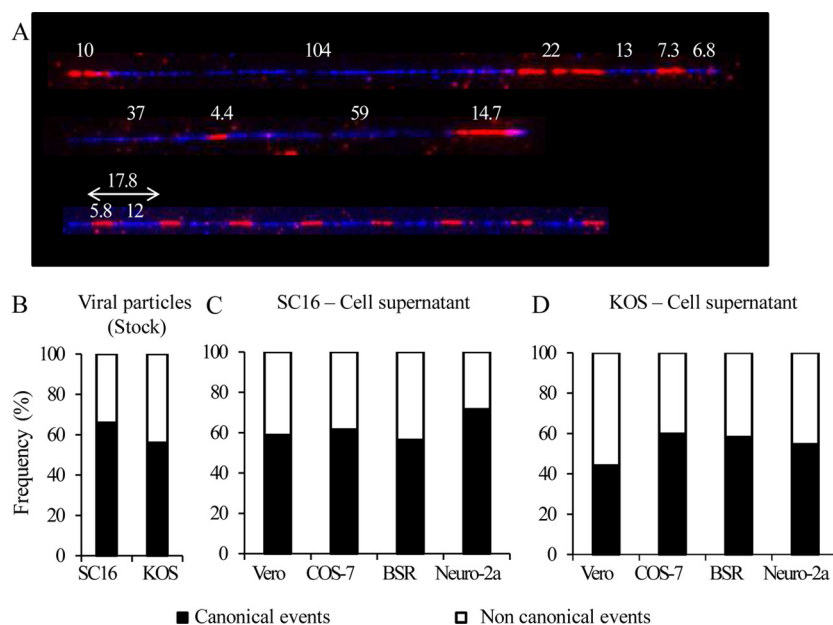


FIG 5 Detection of noncanonical HSV-1 SC16 particle DNA forms. (A) Examples of three fluorescence hybridization arrays. Scale bar, 12.5 μ m (25 kb). The sizes of the Alexa Fluor 594 and Alexa Fluor 350/AMCA fluorescence arrays are indicated in kilobases above the corresponding fluorescent signals. The arrow indicates a 17.8-kb-long repetitive array. (B to D) Histograms showing the mean frequencies of canonical and noncanonical events in SC16 and KOS viral-particle DNA produced in BSR cells (Stock) (B) and in viral particles released from production of strains SC16 (C) and KOS (D) in Vero, Cos-7, BSR, and Neuro-2a cells.

the standard genome and designated repetitive interfering particles. Among the genomes extracted from viral particles, the ratio between canonical and noncanonical events varied between viral strains and could be as high as 47% for the KOS strain of HSV-1 (Fig. 5B, C, and D).

Differential distributions of the HSV-1 isomers *in vivo*. To determine whether nonrandom isomer distributions are also observed *in vivo*, total DNA was extracted from corneas and trigeminal ganglia from mice infected with the SC16 or KOS strain. The analysis of the distributions of the canonical HSV-1 genomes revealed that the ILS isomer predominated with both strains and in both cornea and trigeminal ganglion tissues. This was never found *in vitro*. Although the numbers of isomers were low in both SC16- and KOS-infected corneas ($n = 209$ and $n = 115$, respectively), their distributions were random ($P = 0.06$ and $P = 0.6$, respectively). The distributions were similarly random in the trigeminal ganglion extracts ($P = 0.06$) (Fig. 6A) from mice infected with strain SC16, whereas the ILS isomers were overrepresented (frequency, 34%) in trigeminal ganglion extracts from mice infected with the KOS strain. This led to statistically nonequivalent distributions with a chi-square P value of 6×10^{-4} (Fig. 6B). This nonrandom distribution was distinct from those observed in *in vitro* experiments. Concatemeric DNA was rarely detected in infected cornea and trigeminal ganglion DNA extracts, suggesting a low rate of replication at 6 days postinoculation. As observed *in vitro*, 20% and 35% of the genomes detected in the corneas and trigeminal ganglia corresponded to noncanonical forms (Fig. 6C), suggesting that HSV-1 genome replication leads to relatively high levels of aberrant genomes that are probably noninfectious.

DISCUSSION

The distribution of HSV-1 genome isomers is commonly thought to be random (12, 13, 19). While herpesviruses genomes have

been studied by electron microscopy (13) and Southern blotting (2) for many years, the uniform stretching provided by molecular combing, which essentially pulls genomes into fibers like molecules that can be readily detected by fluorescent probes (36), coupled with direct visualization by differently labeled probes provides unprecedented resolution for the statistical analysis of the genome isomer distribution. The mean size of DNA molecules extracted from virions was around 30 kb after phenol-chloroform extraction (Fig. 1C), much less than one genome equivalent. To allow the analysis of the distribution of the four HSV-1 isomers, the method of extraction was improved, reaching ~70 to 80 kb as an average fiber length for DNA extracted from virions in agarose blocks. However, it was still well below one genome equivalent. This may have several origins that are not mutually exclusive. First, it has been reported that members of the *Herpesviridae* randomly produce partial genomes during their replication (22, 47). If those elements, called defective interfering particles, contain packaging signals, they are encapsidated and thus create interference during serial infection. Indeed, at a given time, more copies of the fragments than of complete genomes are reproduced, leading to a reduction in infectious particles. It is then probable that part of the viral particles extracted corresponded to accumulated defective interfering particles. Another source of HSV-1 genome fragmentation could be the activities of endonucleases, which may have cleaved concatemers from lysed infected cells. A recent work has also shown that HSV-1 genomes are susceptible to cytidine deamination by APOBEC3C (58), a reaction very rapidly followed by uracil removal by uracil *N*-glycosylase, and it is possible that packaged DNA might have some abasic sites. Finally, it has been reported that virion and replicative HSV-1 DNAs contain frequent nicks and gaps (3, 17, 26, 43, 52, 66). During DNA extraction and/or molecular combing, these sites might prove to be frag-

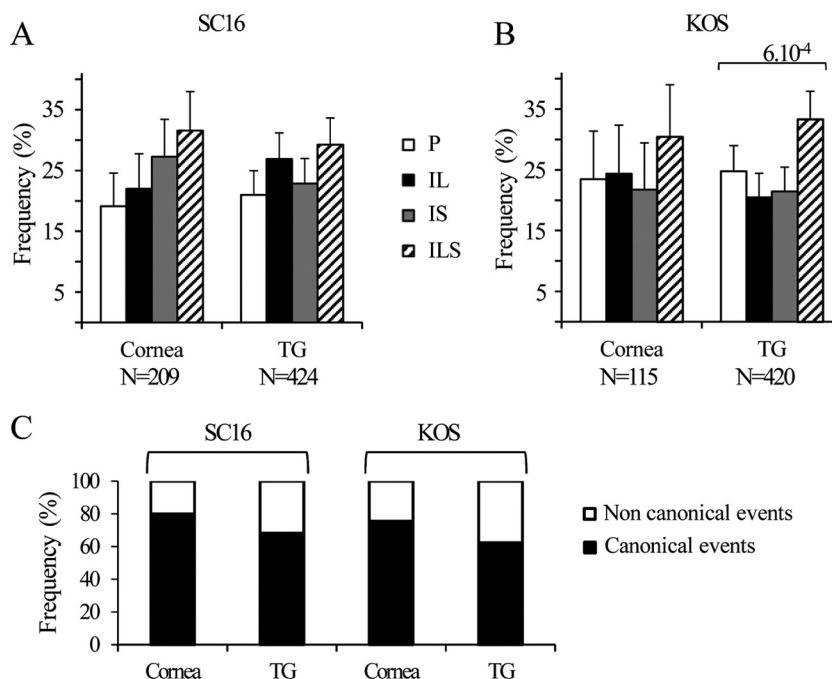


FIG 6 (A and B) Distribution of the HSV-1 isomers in right and left corneas and trigeminal ganglia (TG) of two mice infected with strain SC16 (A) and two mice infected with strain KOS (B). The histograms show the mean frequencies of the four HSV-1 genome isomers. For each condition, the number of signals analyzed (N) is indicated. *P* values of <0.05 (χ^2 test) are indicated by brackets. (C) Histograms showing the mean frequencies of canonical and noncanonical events in corneas and trigeminal ganglia of two mice infected with strain SC16 or KOS.

ile, giving rise to large numbers of subgenome fragments. What is clear is that phenol-chloroform DNA extraction results in considerable degradation of HSV-1 genomes, meaning that prior studies relying on the quality of HSV-1 DNA may have led to some erroneous findings.

The variety of HSV-1 DNA forms detected was considerable. In cell extracts, it represented subgenomic fragments, concatemers, and noncanonical genomes. The isomer distribution was uniquely addressed in the context of one genome equivalent and concatemers. The isomer distribution was found to be equimolar *in vitro* when using the strain SC16. In contrast, statistically significant overrepresentation of P and IS isomers was observed in cell and cell supernatant DNA extracts from a culture of Vero and Neuro-2a cells infected with strain KOS. Thus, the distribution was found to be host cell and strain dependent *in vitro*, suggesting a cell- and strain-dependent mechanism. As equivalent distributions were noticed in KOS concatemeric DNA *in vitro* in all cell lines, the unbalanced distribution observed in viral particles and cell extracts was likely related to the encapsidation process. The isomerization mechanism has been previously linked to the encapsidation process, as the alternative cleavage of concatemeric molecules is part of the mechanism leading to the four genome isomers (56, 62). Even if the L and S termini both contain the signals for DNA encapsidation (57), only L termini are observed in concatemeric DNA, suggesting that encapsidation starts at the L termini and proceeds toward S (34, 69). Here, the unbalanced distribution could be linked to increased efficiency in the cleavage of the *a* sequence toward P and IS isomers or a predisposition to degradation of IL and ILS isomers after cleavage. *In vivo*, the distributions were found to be random, except with strain KOS, where the ILS isomers are overrepresented in trigeminal ganglia,

which was distinct from what was observed *in vitro*. Statistically unbalanced distributions were uniquely observed with strain KOS, a less virulent strain than SC16 (6). The KOS strain of HSV-1 is known to be modified in the ICP34.5 gene (39), which encodes a neurovirulence factor that maps in the TR_L/IR_L inverted repeats (15, 40). This, combined with the observation of the balanced distribution of putative isomers in concatemeric DNA for both SC16 and KOS, suggests that the genomic and/or epigenomic organization of the TR_L/IR_L regions could affect the isomerization process, likely at the encapsidation step. Finally, both *in vitro* and *in vivo*, the frequency of one isomer reached 40% from the null distribution, suggesting that the tendency is not to deviate too far from an equimolar distribution. However, the continued unbalanced distributions observed using a single-molecule approach suggest a controlled process. It is notable that molecular combing allowed the measurement of replication intermediates. It has been shown previously that most of the concatemeric DNA was comprised of 2 or 3 genome equivalents (14), which was confirmed here (Table 3). We also observed replication intermediates of up to 10 genome equivalents. With the power to visualize different regions of the genome, we could identify the composition of the concatemers, which suggests recombination. As they could represent ligation of monomers, dimers, or trimers or reflect nuclease isomerization within the concatemers, it is not possible to formally distinguish between the two models currently proposed to describe HSV-1 DNA replication, i.e., rolling-circle (7) and recombination-dependent replication (67) mechanisms.

We also observed large proportions of noncanonical isomers *in vitro* (30 to 60%). A large variety of patterns were detected. Overall, as suggested previously (59), recombination between HSV-1 genomes is a major event in the replication process allowing the

creation of new genetic assortments and thus may play an important role in the evolutionary process. *In vivo*, the noncanonical assortments were less abundant than *in vitro*, which suggests that some mechanisms may occur to optimize the efficacy of genome assortment and to improve the quality of encapsidated viral genomes. Moreover, the proportion of noncanonical forms was 10 to 20% higher in trigeminal ganglia than in corneas. At this stage of infection, i.e., 6 days after viral inoculation by corneal scarification, the cornea is markedly infected for several days, and most of the detected HSV-1 genomes are likely already encapsidated, if not contained in mature viral particles. In contrast the trigeminal ganglia had been reached by viral particles for only 1 or 2 days, and the replicative infection within the neuronal somata was less complete than in the cornea (data not shown). Thus, most of the HSV-1 genomes detected were likely undergoing replication, which may account for our results.

While the success of HSV-1 has never been in doubt, it seems to be achieved at a high cost in terms of DNA synthesis. While experimental factors reduce the genome size to less than one unit length, the sheer proportions of subgenomic molecules and non-canonical forms is a little surprising. They may reflect anthropocentric perceptions, as the particle- to-PFU ratio for HSV-1 is on the order of 100:1. The present data introduce another cost aspect in viral ecology, the accuracy of DNA replication of large genomes, which does not show up in particle-to-PFU ratios. By extrapolation, it is tempting to suggest that a similar situation may pertain to the replication of other herpesvirus genomes, most of which are larger than that of HSV-1. It is an open question whether genome replication of DNA viruses is similarly affected, but unique molecular combing has the potential to explore this question.

ACKNOWLEDGMENTS

This work is part of the project ACTIVE, which is supported by the Strategic Industrial Innovation program of OSEO (OSEO-ISI). C.M. is supported by a grant from the Association Nationale de la Recherche et de la Technologie (ANRT).

REFERENCES

- Bataille D, Epstein A. 1994. Herpes simplex virus replicative concatemers contain L components in inverted orientation. *Virology* 203:384–388.
- Bataille D, Epstein AL. 1997. Equimolar generation of the four possible arrangements of adjacent L components in herpes simplex virus type 1 replicative intermediates. *J. Virol.* 71:7736–7743.
- Ben-Porat T, Rixon FJ, Blankenship ML. 1979. Analysis of the structure of the genome of pseudorabies virus. *Virology* 95:285–294.
- Ben-Zeev A, Weinberg E, Becker Y. 1974. Isolation of herpes simplex virus DNA from virus particles and infected cells by electrophoresis in polyacrylamide gels. *J. Gen. Virol.* 25:63–73.
- Berlemont S, Bensimon A, Olivo-Marin J-C. 2007. Detection of curvilinear objects in noisy image using feature-adapted beamlet transform. *Proc. Int. Conf. Acoustics, Speech, Signal Processing* 1:1225–1228.
- Blyth WA, Harbour DA, Hill TJ. 1984. Pathogenesis of zosteriform spread of herpes simplex virus in the mouse. *J. Gen. Virol.* 65:1477–1486.
- Boehmer PE, Lehman IR. 1997. Herpes simplex virus DNA replication. *Annu. Rev. Biochem.* 66:347–384.
- Brown SM, Subak-Sharpe JH, Harland J, MacLean AR. 1992. Analysis of intrastrain recombination in herpes simplex virus type 1 strain 17 and herpes simplex virus type 2 strain HG52 using restriction endonuclease sites as unselected markers and temperature-sensitive lesions as selected markers. *J. Gen. Virol.* 73:293–301.
- Chayavichitsilp P, Buckwalter JV, Krakowski AC, Friedlander SF. 2009. Herpes simplex. *Pediatr. Rev.* 30:119–129.
- Chou J, Roizman B. 1989. Characterization of DNA sequence-common and sequence-specific proteins binding to cis-acting sites for cleavage of the terminal a sequence of the herpes simplex virus 1 genome. *J. Virol.* 63:1059–1068.
- Chou J, Roizman B. 1985. Isomerization of herpes simplex virus 1 genome: identification of the cis-acting and recombination sites within the domain of the a sequence. *Cell* 41:803–811.
- Clements JB, Cortini R, Wilkie NM. 1976. Analysis of herpesvirus DNA substructure by means of restriction endonucleases. *J. Gen. Virol.* 30:243–256.
- Delius H, Clements JB. 1976. A partial denaturation map of herpes simplex virus type 1 DNA: evidence for inversions of the unique DNA regions. *J. Gen. Virol.* 33:125–133.
- Deshmane SL, Raengsakulrach B, Berson JF, Fraser NW. 1995. The replicating intermediates of herpes simplex virus type 1 DNA are relatively short. *J. Neurovirol.* 1:165–176.
- Dolan A, McKie E, MacLean AR, McGeoch DJ. 1992. Status of the ICP34.5 gene in herpes simplex virus type 1 strain 17. *J. Gen. Virol.* 73:971–973.
- Dutch RE, Bruckner RC, Mocarski ES, Lehman IR. 1992. Herpes simplex virus type 1 recombination: role of DNA replication and viral a sequences. *J. Virol.* 66:277–285.
- Goldstein JN, Weller SK. 1998. The exonuclease activity of HSV-1 UL12 is required for *in vivo* function. *Virology* 244:442–457.
- Harmenberg J, Oberg B, Spruance S. 2010. Prevention of ulcerative lesions by episodic treatment of recurrent herpes labialis: a literature review. *Acta Derm. Venereol.* 90:122–130.
- Hayward GS, Jacob RJ, Wadsworth SC, Roizman B. 1975. Anatomy of herpes simplex virus DNA: evidence for four populations of molecules that differ in the relative orientations of their long and short components. *Proc. Natl. Acad. Sci. U. S. A.* 72:4243–4247.
- Hill TJ, Field HJ, Blyth WA. 1975. Acute and recurrent infection with herpes simplex virus in the mouse: a model for studying latency and recurrent disease. *J. Gen. Virol.* 28:341–353.
- Honess RW, Buchan A, Halliburton IW, Watson DH. 1980. Recombination and linkage between structural and regulatory genes of herpes simplex virus type 1: study of the functional organization of the genome. *J. Virol.* 34:716–742.
- Huang AS. 1973. Defective interfering viruses. *Annu. Rev. Microbiol.* 27:101–117.
- Huang KJ, Ku CC, Lehman IR. 2006. Endonuclease G: a role for the enzyme in recombination and cellular proliferation. *Proc. Natl. Acad. Sci. U. S. A.* 103:8995–9000.
- Huang KJ, Zemelman BV, Lehman IR. 2002. Endonuclease G, a candidate human enzyme for the initiation of genomic inversion in herpes simplex type 1 virus. *J. Biol. Chem.* 277:21071–21079.
- Jackson SA, DeLuca NA. 2003. Relationship of herpes simplex virus genome configuration to productive and persistent infections. *Proc. Natl. Acad. Sci. U. S. A.* 100:7871–7876.
- Jacob RJ, Roizman B. 1977. Anatomy of herpes simplex virus DNA. VIII. Properties of the replicating DNA. *J. Virol.* 23:394–411.
- Jenkins FJ, Martin JR. 1990. Role of the herpes simplex virus 1 internal repeat sequences in pathogenicity. *Intervirology* 31:129–138.
- Jenkins FJ, Roizman B. 1986. Herpes simplex virus 1 recombinants with noninverting genomes frozen in different isomeric arrangements are capable of independent replication. *J. Virol.* 59:494–499.
- Kowalczykowski SC. 2000. Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem. Sci.* 25:156–165.
- Labetoulle M, et al. 2005. Incidence of herpes simplex virus keratitis in France. *Ophthalmology* 112:888–895.
- Labetoulle M, et al. 2000. Neuronal propagation of HSV1 from the oral mucosa to the eye. *Invest. Ophthalmol. Vis. Sci.* 41:2600–2606.
- Liesegang TJ. 2001. Herpes simplex virus epidemiology and ocular importance. *Cornea* 20:1–13.
- Martin DW, Weber PC. 1996. The a sequence is dispensable for isomerization of the herpes simplex virus type 1 genome. *J. Virol.* 70:8801–8812.
- Martinez R, Sarisky RT, Weber PC, Weller SK. 1996. Herpes simplex virus type 1 alkaline nuclease is required for efficient processing of viral DNA replication intermediates. *J. Virol.* 70:2075–2085.
- McVoy MA, Nixon DE, Adler SP, Mocarski ES. 1998. Sequences within the herpesvirus-conserved pac1 and pac2 motifs are required for cleavage and packaging of the murine cytomegalovirus genome. *J. Virol.* 72:48–56.
- Michalet X, et al. 1997. Dynamic molecular combing: stretching the whole human genome for high-resolution studies. *Science* 277:1518–1523.

37. Mocarski ES, Roizman B. 1982. Herpesvirus-dependent amplification and inversion of cell-associated viral thymidine kinase gene flanked by viral α sequences and linked to an origin of viral DNA replication. *Proc. Natl. Acad. Sci. U. S. A.* 79:5626–5630.
38. Nasserli M, Mocarski ES. 1988. The cleavage recognition signal is contained within sequences surrounding an α -junction in herpes simplex virus DNA. *Virology* 167:25–30.
39. Perng GC, et al. 2002. Herpes simplex virus type 1 mutants containing the KOS strain ICP34.5 gene in place of the McKrae ICP34.5 gene have McKrae-like spontaneous reactivation but non-McKrae-like virulence. *J. Gen. Virol.* 83:2933–2942.
40. Perry LJ, McGeoch DJ. 1988. The DNA sequences of the long repeat region and adjoining parts of the long unique region in the genome of herpes simplex virus type 1. *J. Gen. Virol.* 69:2831–2846.
41. Poffenberger KL, Tabares E, Roizman B. 1983. Characterization of a viable, noninverting herpes simplex virus 1 genome derived by insertion and deletion of sequences at the junction of components L and S. *Proc. Natl. Acad. Sci. U. S. A.* 80:2690–2694.
42. Pogue-Geile KL, Spear PG. 1986. Enhanced rate of conversion or recombination of markers within a region of unique sequence in the herpes simplex virus genome. *J. Virol.* 58:704–708.
43. Puvion-Dutilleul F, Pichard E, Laithier M, Puvion E. 1989. Cytochemical study of the localization and organization of parental herpes simplex virus type I DNA during initial infection of the cell. *Eur. J. Cell Biol.* 50:187–200.
44. Roizman B. 1979. The structure and isomerization of herpes simplex virus genomes. *Cell* 16:481–494.
45. Sarisky RT, Weber PC. 1994. Requirement for double-strand breaks but not for specific DNA sequences in herpes simplex virus type 1 genome isomerization events. *J. Virol.* 68:34–47.
46. Sauer A, Wang JB, Hahn G, McVoy MA. 2010. A human cytomegalovirus deleted of internal repeats replicates with near wild type efficiency but fails to undergo genome isomerization. *Virology* 401:90–95.
47. Schroder CH, Furst B, Weise K, Gray CP. 1984. A study of interfering herpes simplex virus DNA preparations containing defective genomes of either class I or II and the identification of minimal requirements for interference. *J. Gen. Virol.* 65:493–506.
48. Schurra C, Bensimon A. 2009. Combing genomic DNA for structural and functional studies. *Methods Mol. Biol.* 464:71–90.
49. Severini A, Morgan AR, Tovell DR, Tyrrell DL. 1994. Study of the structure of replicative intermediates of HSV-1 DNA by pulsed-field gel electrophoresis. *Virology* 200:428–435.
50. Severini A, Scraba DG, Tyrrell DL. 1996. Branched structures in the intracellular DNA of herpes simplex virus type 1. *J. Virol.* 70:3169–3175.
51. Sheldrick P, Berthelot N. 1975. Inverted repetitions in the chromosome of herpes simplex virus. *Cold Spring Harb. Symp. Quant. Biol.* 39:667–678.
52. Shlomai J, Friedmann A, Becker Y. 1976. Replication intermediates of herpes simplex virus DNA. *Virology* 69:647–659.
53. Slobodman B, Zhang X, Simmons A. 1999. Herpes simplex virus genome isomerization: origins of adjacent long segments in concatemeric viral DNA. *J. Virol.* 73:810–813.
54. Smiley JR, Duncan J, Howes M. 1990. Sequence requirements for DNA rearrangements induced by the terminal repeat of herpes simplex virus type 1 KOS DNA. *J. Virol.* 64:5036–5050.
55. Smiley JR, Fong BS, Leung WC. 1981. Construction of a double-jointed herpes simplex viral DNA molecule: inverted repeats are required for segment inversion, and direct repeats promote deletions. *Virology* 113:345–362.
56. Smiley JR, Lavery C, Howes M. 1992. The herpes simplex virus type 1 (HSV-1) α sequence serves as a cleavage/packaging signal but does not drive recombinational genome isomerization when it is inserted into the HSV-2 genome. *J. Virol.* 66:7505–7510.
57. Stow ND, McMonagle EC, Davison AJ. 1983. Fragments from both termini of the herpes simplex virus type 1 genome contain signals required for the encapsidation of viral DNA. *Nucleic Acids Res.* 11:8205–8220.
58. Suspène R, et al. 2011. Genetic editing of herpes simplex virus 1 and Epstein-Barr herpesvirus genomes by human APOBEC3 cytidine deaminases in culture and in vivo. *J. Virol.* 85:7594–7602.
59. Thiry E, et al. 2005. Recombination in alphaherpesviruses. *Rev. Med. Virol.* 15:89–103.
60. Trus BL, et al. 2007. Allosteric signaling and a nuclear exit strategy: binding of UL25/UL17 heterodimers to DNA-filled HSV-1 capsids. *Mol. Cell* 26:479–489.
61. Umene K, Fukumaki Y. 2011. DNA genome of spontaneously occurring deletion mutants of herpes simplex virus type 1 lacking one copy of the inverted repeat sequences of the L component. *Arch. Virol.* 156:1305–1315.
62. Varmuza SL, Smiley JR. 1985. Signals for site-specific cleavage of HSV DNA: maturation involves two separate cleavage events at sites distal to the recognition sequences. *Cell* 41:793–802.
63. Wadsworth S, Jacob RJ, Roizman B. 1975. Anatomy of herpes simplex virus DNA. II. Size, composition, and arrangement of inverted terminal repetitions. *J. Virol.* 15:1487–1497.
64. Weber PC, Challberg MD, Nelson NJ, Levine M, Glorioso JC. 1988. Inversion events in the HSV-1 genome are directly mediated by the viral DNA replication machinery and lack sequence specificity. *Cell* 54:369–381.
65. Whitley RJ. 2006. Herpes simplex encephalitis: adolescents and adults. *Antiviral Res.* 71:141–148.
66. Wilkie NM. 1973. The synthesis and substructure of herpesvirus DNA: the distribution of alkali-labile single strand interruptions in HSV-1 DNA. *J. Gen. Virol.* 21:453–467.
67. Wilkinson DE, Weller SK. 2003. The role of DNA recombination in herpes simplex virus DNA replication. *IUBMB Life* 55:451–458.
68. Wohlrab F, Chatterjee S, Wells RD. 1991. The herpes simplex virus 1 segment inversion site is specifically cleaved by a virus-induced nuclear endonuclease. *Proc. Natl. Acad. Sci. U. S. A.* 88:6432–6436.
69. Zhang X, Efsthathiou S, Simmons A. 1994. Identification of novel herpes simplex virus replicative intermediates by field inversion gel electrophoresis: implications for viral DNA amplification strategies. *Virology* 202:530–539.

Résumé :

Le virus de l'herpès simplex de type 1 (HSV1) établit une infection latente dans le système nerveux de l'homme, au cours de laquelle un type de transcrits, appelés LATs (pour *latency associated transcripts*), s'accumule dans les neurones infectés. Le rôle clef des LATs dans le contrôle de la latence virale est reconnu. Cependant, depuis leur découverte dans les années 80, leur mécanisme d'action reste non élucidé.

Le gène des LATs est transcrit en un LAT primaire de 8,3kb, qui est épissé, conduisant à la formation de deux LATs stables : le LAT_{2kb} et le LAT_{1.5kb}. De façon remarquable, le LAT_{2kb} et le LAT_{1.5kb} sont des introns. Leur stabilité est la conséquence d'un branchement non canonique qui se traduit par le maintien de la structure en lariat.

Au cours de ce travail, nous avons pu que les LATs sont un nouveau type d'éléments mobiles dotés de propriétés uniques qui sont à la fois capables de modifier les génomes viraux et cellulaires. Ces propriétés des LATs ouvrent de nouveaux horizons sur la biologie du virus, en particulier sur la régulation de latence.

Enfin, l'identification de cet élément mobile viral, qui partage des similitudes fonctionnelles et structurales avec des éléments mobiles procaryotes et eucaryotes, soulève des questions importantes dans le domaine de l'évolution.

Summary :

The herpes simplex virus type 1 (HSV1) establishes a lifelong latent infection in the human nervous system, during which a subtype of transcripts, called *Latency Associated Transcripts* (LATs), accumulate in infected neurons. The key role of LATs in the control of HSV1 latency is well acknowledged. However, their mechanism of action remains unknown since their discovery in the 80's.

The LAT gene encodes an 8.3 kb primary transcript, which is spliced, leading to the formation of two stable LATs, LAT_{2kb} and LAT_{1.5kb} (2 kb and 1.5 kb in size, respectively). Remarkably, LAT_{2kb} and LAT_{1.5kb} are stable introns, as a consequence of a non-consensus branch-point that results in the maintenance of a lariat structure.

During this work, we showed that LATs are a new type of mobile elements with unique properties which are able to modify both the viral and the host genomes. This unique property of LATs opens new horizons on HSV1 biology, particularly on latency regulation.

Finally, the identification of this viral mobile element, that shares functional and structural similarities with both prokaryotic and eukaryotic mobile elements, raises major questions about evolution.

Mots-clés : Herpes simplex de type 1, Elément mobile, Latence, Epissage inverse, Rétrotransposition Neurones, Réactivation.

Laboratoire de rattachement :

Virologie Moléculaire et Structurale

UPR CNRS 3296, bat 14B

1 avenue de la Terrasse

91198 Gif sur Yvette France

tel: +33 (0)1 69 82 38 35

fax: +33 (0)1 69 82 43 08

PÔLE : MICROBIOLOGIE / THERAPEUTIQUES ANTIINFECTIEUSES

UNIVERSITÉ PARIS-SUD 11

UFR «FACULTÉ DE PHARMACIE DE CHATENAY-MALABRY »

5, rue Jean Baptiste Clément

92296 CHÂTENAY-MALABRY Cedex