N° d'ordre: 1220

50376
1984
302

50376
198H
302

THESE

présentée à

L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE I

pour obtenir le titre de

DOCTEUR DE 3e CYCLE EN BIOCHIMIE

par

Dominique LEPRINCE

BIBLIOTHEQUE INTERUNIVERSITAIRE
SECTION DE SCIENCES
LILLE

CLONAGE DU PROTO-ONCOGENE c-myb DANS L'ADN DE POULET :

_ETUDE DU GENE TRANSDUIT PAR LES RETROVIRUS AVIAIRES AMV ET E26.

_CLONAGE DE SON HOMOLOGUE HUMAIN.

Soutenue le

MM. J. MONTREUIL

J. KREMBEL

J. SAMARUT

D. STEHELIN

AVANT PROPOS

Le travail présenté dans ce mémoire a été réalisé dans le laboratoire d'Oncologie Moléculaire (Unité INSERM 186) dirigé par Dominique STEHELIN. Je le remercie sincèrement de m'avoir accueilli dans son laboratoire et confié ce travail de recherche.

Je remercie Monsieur le Président et Messieurs les Membres du Jury qui ont accepté de juger ce mémoire.

Ce travail a été rendu possible grâce à l'aide constante, aux conseils précieux et à la sympathie de tous les membres de l'équipe : qu'ils veuillent bien trouver ici, l'expresion de ma reconnaissance.

Je tiens, cependant, à remercier plus particulièrement :

- Martine ROUSSEL et Simon SAULE qui ont guidé avec bienveillance mes premiers pas dans ce travail.

- Anne GEGONNE et Anne FLOURENS-SCHNEEBERGER pour leur enthousiasme et leur collaboration aussi efficace que sympathique sur le projet E26.

- Agnès BEGUE et Christian LAGROU dont j'ai eu loisir d'apprécier tant les qualités humaines que les compétences exceptionnelles dans leur domaine respectif de la "lambdologie" et de la culture cellulaire.

- Marie Christine LAGARDERE et Nicole DEVASSINE qui ont assuré avec beaucoup de patience, de soin et de gentillesse la dactylographie de ce travail.

- Marie Berthe RAES pour la réalisation des photographies et Jean COLL pour la pointe d'humour qu'il a glissée dans ce mémoire.

- Merci, Sylvie pour ta compréhension et ton soutien moral constant.

SOMMAIRE

GENERALITES ET BUT DE CE TRAVAIL

Deux grands axes indépendants de recherche sur le cancer, l'un étudiant le mécanisme de l'induction de tumeurs chez les animaux par les rétrovirus et l'autre celui de l'induction des cancers d'origine non virales chez l'homme, ont convergé de façon surprenante durant ces dernières années vers la même cible : les proto-oncogènes cellulaires.

Le pouvoir transformant des rétrovirus oncogènes (oncovirus) est dû à la présence dans leur génome d'une séquence spécifique : l'oncogène viral (v-onc) provenant de la recombinaison entre un rétrovirus sans oncogène et un gène cellulaire, proto-oncogène cellulaire (c-onc). Certains de ces proto-oncogènes, révélés par l'étude des rétrovirus semblent être spécifiquement impliqués par des phénomènes d'amplification génique, de mutations ou de translocations chromosomiques dans certaines tumeurs animales et humaines non viro-induites.

Le travail présenté dans ce mémoire comporte trois parties : Les deux premières traitent du clonage dans l'ADN de poulet du proto-oncogène c-myb qui a été transduit de façon différente dans deux rétrovirus aviaires : le virus de la myéloblastose aviaire (AMV) et le virus érythroblastosant et myéloblastosant E26. La troisième partie traite du clonage du même proto-oncogène dans l'ADN humain. La carte physique du gène c-myb humain ainsi que l'obtention de sondes spécifiques de ce gène constituent le point de départ d'une étude systématique de son implication éventuelle dans certaines tumeurs humaines.

INTRODUCTION

PRESENTATION DES RETROVIRUS

Les rétrovirus oncogènes ou oncovirus (qui seront seuls traités dans ce mémoire) appartiennent à une famille de virus dont le génome diploïde est consitué d'ARN se répliquant par l'intermédiaire d'un ADN proviral. Cet ADN est transcrit dans les heures qui suivent l'infection, à partir de l'ARN viral par une enzyme caractéristique de ces virus. L'ADN polymérase - ARN dépendante ou rétropolymérase (Baltimore, 1970 ; Temin et Mizutani, 1970) puis intégré dans le génome de la cellule infectée. Ces gènes viraux intégrés sont alors répliqués, transcrits et traduits comme les gènes de la cellule hôte.

Les rétrovirus sont classiquement divisés en deux groupes selon leurs propriétés biologiques : les rétrovirus peu ou non transformants et les rétrovirus transformants.

Les rétrovirus peu ou non transformants

Ces virus sont incapables de transformer les cellules en culture. En revanche, ils induisent parfois chez l'animal des tumeurs, principalement des lymphomes, dont le temps d'apparition est de plusieurs mois. Comme exemple de tels virus, on peut citer le virus de la lymphomatose aviaire (ALV) ou les virus de lymphomes murins (MuLV). Leur génome comprend trois gènes de structure constituant une unité réplicative leur permettant de se propager de façon autonome (Coffin, 1979) : le gène gag codant pour les protéines de la nucléocapside virale, le gène pol pour

la rétropolymérase et le gène env pour des protéines qui après glycosylation migreront à la surface de l'enveloppe virale (Figures 1 et 2). De même, comme conséquence du mécanisme de la rétrotranscription, l'ADN proviral n'est pas une copie conforme de l'ARN. On observe aux deux extrémités du provirus une duplication de certaines séquences, le LTR ou "Longue Répétition Terminale" (Shank et al., 1978). L'analyse de la séquence nucléotidique du LTR a montré qu'il contenait les signaux de promotion, d'initiation et de terminaison de la transcription du génome viral, ainsi que des séquences répétées inversées (Figure 1). On peut donc considérer ces virus, qui s'intègrent au hasard dans l'ADN de la cellule hôte, comme des éléments génétiques mobiles (Temin, 1980 ; Varmus, 1982).

Les rétrovirus transformants

Ces virus ont la capacité de transformer des cellules en culture et d'induire chez l'animal des leucémies aigües ou des tumeurs (sarcomes ou carcinomes) dont le temps d'apparition -de l'ordre de quelques semaines- est beaucoup plus court que celui des virus non transformants (Graf et Beug, 1978 ; Graf et Stéhelin, 1982). C'est d'ailleurs à partir de tumeurs spontanées que des virus de ce type ont été isolés chez le poulet, le chat, le rat, la souris et le singe et maintenus au laboratoire. Le premier isolat d'un tel virus fut réalisé par P. Rous à partir d'un sarcome spontané de poulet, dès 1911, probablement parce que le virus du sarcome de Rous (RSV) semble être le seul rétrovirus transformant compétant pour la réplication. Dans tous les autres cas, l'acquisition de l'information génétique responsable de la transformation ou oncogène viral, v-onc, s'est effectuée au détriment d'une partie plus ou moins importante de l'unité

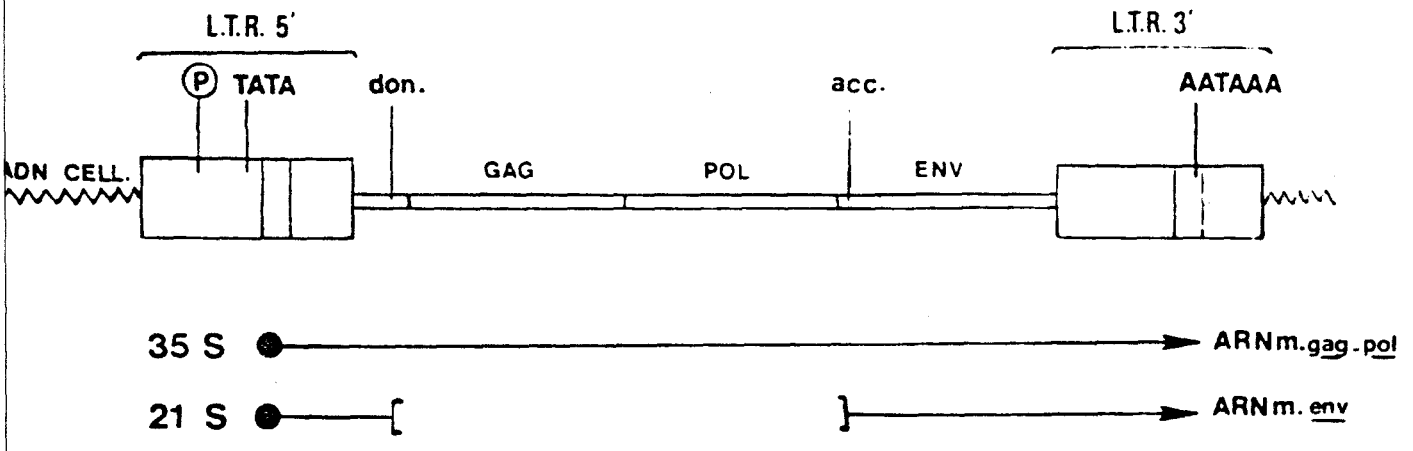


FIGURE 1 : Transcription de l'ADN proviral

P - promoteur
TATA - séquence riche en A + T qui précise l'origine de la transcription
U3 - séquence unique 3'
R - séquence répétitive
U5 - séquence unique 5'
don. - site d'épissure donneur
acc. - site d'épissure accepteur
AATAAA - signal de polyadénylation

BU LILLE

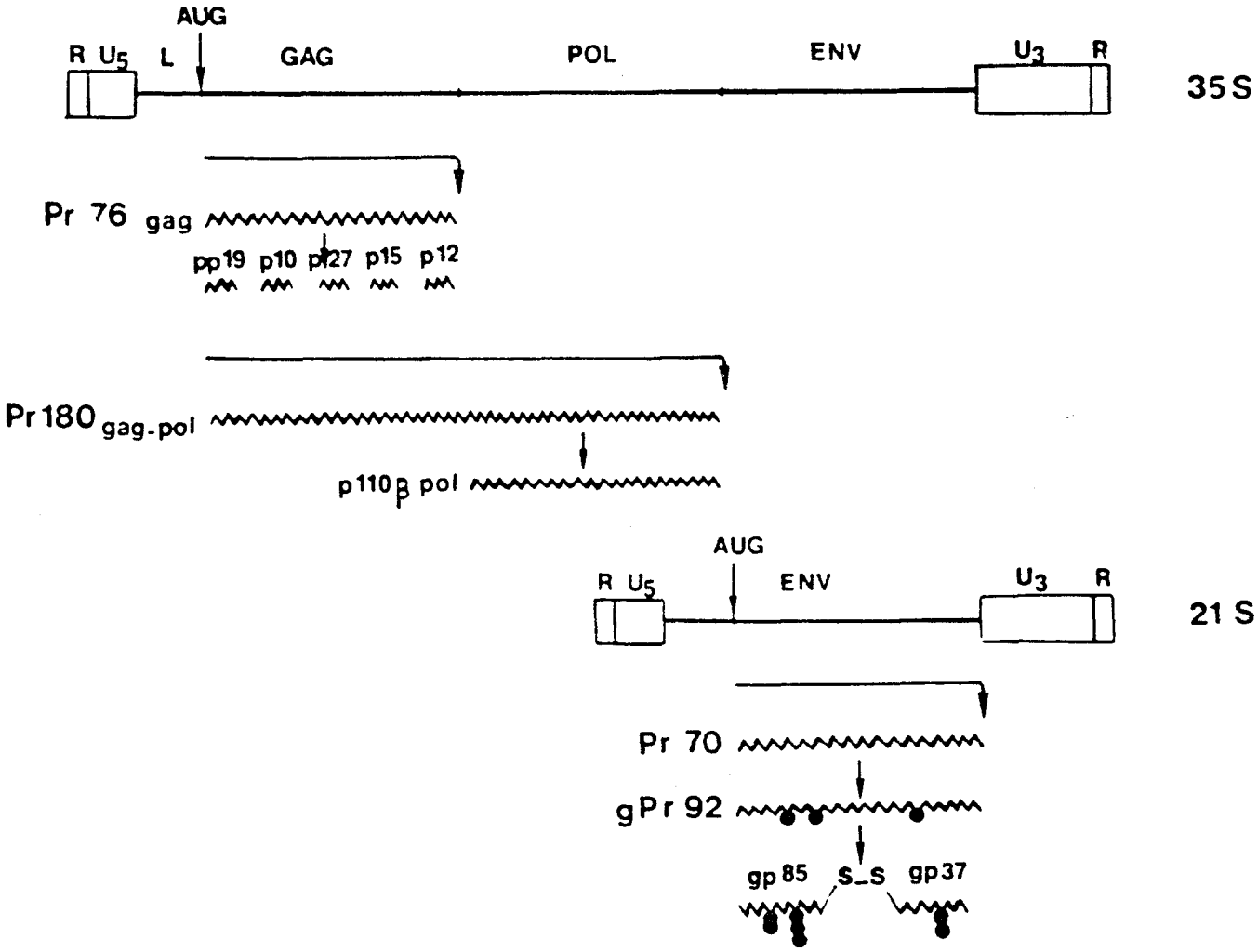


FIGURE 2 : Traduction des ARNs messagers aviaires

AUG - codon d'initiation
Pr - protéine précurseur
pp - protéine phosphorylée
gp - glycoprotéine
(●) - sucres

BU LILLE

réplicative. Cette amputation peut toucher tous les gènes de structure mais préserve néanmoins les extrémités du génome (nécessaires à la formation des LTRs) et le début du gène gag contenant les signaux nécessaires pour l'encapsidation de l'ARN . Ces virus sont donc défectifs pour la réplication et ont besoin pour se propager d'un virus auxiliaire ("helper") -généralement un virus peu ou pas transformant- qui apporte les constituants structuraux que le virus transformant ne peut synthétiser.

En plus de ces deux types de virus, existent dans le génome de leurs hôtes à l'état intégré des virus endogènes le plus souvent inoffensifs pour l'animal (Astrin et al., 1980).

Chez l'homme, les seuls rétrovirus connus jusqu'à présent ont été isolés très récemment à partir de leucémies lymphoïdes à endémies très localisées. Ces virus HTLV ("Human T Leukaemia Virus) ne possèdent apparemment pas d'oncogène mais une information génétique de structure classique gag, pol, env avec cependant à l'extrémité 3' du génome une région codante pX, dont le rôle n'est pas connu (Gallo et Wong-Staal, 1982).

ONCOGENES VIRAUX ET PROTO-ONCOGENES CELLULAIRES

Mise en évidence

Des études biologiques concernant le virus du sarcome de Rous (RSV) semblaient indiquer l'existence dans le génome de ce virus hautement transformant d'une séquence virale responsable de la transformation. Il existait en effet, des mutants thermosensibles du RSV : des fibroblastes de poulet infectés par

ces mutants pouvaient être transformés de façon réversible en fonction de la température sans affecter la production virale (Martin, 1970).

De même, lors de passages successifs en culture de stocks clonés de RSV, des auteurs ont observé la ségrégation spontanée de virus défectifs pour la transformation, td RSV (Kawai et Hanafusa, 1972). Stéhelin et al. en 1976 ont apporté la preuve directe de l'existence dans le génome du RSV d'un gène, v-src, responsable de la transformation. Après synthèse d'un cDNA radioactif complémentaire du génome de RSV et élimination des séquences communes avec un td RSV, ces auteurs ont obtenu un cDNA complémentaire d'une séquence de quelques 2000 nucléotides, trouvée exclusivement dans le génome du RSV (Stéhelin et al., 1976a ; 1976b). La protéine codée par ce gène fut ensuite mise en évidence par Brugge et Erikson (1977) grâce à des antiséra obtenus à partir de lapins porteurs de tumeurs induites par RSV. Ces antiséra immunoprécipitaient une protéine de 60 000 daltons de poids moléculaire, $pp60^{v-src}$ présente uniquement dans des cellules infectées par RSV et non immunoprécipitée par des antiséra dirigés contre les gènes de structure. L'obtention de la même protéine dans les produits de traduction in vitro du RSV leva les derniers doutes quant à son origine virale (Beemon et Hunter 1978).

L'application de la technique de préparation des sondes spécifiques à l'étude systématique des virus transformants défectifs pour la réplication a étendu la notion d'oncogène viral et a défini à l'heure actuelle 24 séquences de ce type (Table 1). (Graf et Stéhelin, 1982 ; Bishop 1983). Conventionnellement, les oncogènes viraux sont identifiés par 3 lettres rappelant, de façon simple, soit le virus d'où ils ont été isolés, le type de

maladie induit, le lieu où ces virus ont été découverts ou la personne qui les a isolés, précédées du préfixe v marquant leur nature virale (Coffin et al., 1981) (Table 1).

Les oncogènes viraux sont les produits d'une recombinaison génétique

A l'aide de la sonde spécifique, v-src, Stéhelin et al., (1976 b) ont pu montrer que dans le génome de la poule, indépendemment de tout virus endogène, existait une information cellulaire homologue à ce gène viral. Cette observation étendue ensuite aux autres séquences spécifiques des rétrovirus transformants démontre que ces virus sont le produit d'une recombinaison génétique entre un virus dépourvu d'oncogène (du type "helper") et un gène cellulaire, proto-oncogène cellulaire identifié par les trois lettres de l'oncogène viral qu'il a produit précédées cette fois du préfixe c (Coffin et al., 1981 ; Bishop, 1981). Cette origine cellulaire avait été suggérée par la méthode utilisée pour isoler certains rétrovirus transformants : le virus du sarcome de Moloney (portant l'oncogène v-mos) et le virus du sarcome de Harvey (portant l'oncogène v-Ha-ras) chez le rat sont apparus après infection des animaux par des virus murins sans oncogènes -MuLV de structure gag-pol-env- et passages successifs d'animal à animal (Harvey, 1964 ; Moloney, 1966). De même, un virus érythroblastosant ne portant que l'oncogène v-erbB -et donc clairement distinct de la souche AEV ES-4 portant les oncogènes v-erbB et v-erbA- a été obtenu à partir d'un homogénat de foie d'un poulet infecté par un ALV (Yamamoto et al., 1983a et b). Plus récemment, des lymphosarcomes thymiques naturels induits

TABLE 1 : LES ONCOGENES VIRAUX (d'après BISHOP (1983) et DUESBERG (1984))

ONCOGENE VIRAL	VIRUS PROTOTYPE	ESPECE D'ORIGINE	PROTEINE CODEE	LOCALISATION CELLULAIRE	FONCTION
v-src (3)	virus du sarcome de Rous	poulet caille	pp60src	membrane plasmique	tyrosine phosphokinase
v-yes (2)	virus du sarcome aviaire Y73	poulet	p90 gag-yes	?	tyrosine phosphokinase
v-fps (5)	virus sarcomatogène de Fujinami	poulet	p140 gag-fps	membrane plasmique	tyrosine phosphokinase
v-fes (2)	virus de sarcome félin souche Snyder-Theilen	chat	p85-115 gag-fes	membrane plasmique	tyrosine phosphokinase
v-abl (2)	virus de la leucémie murine d'Abelson	souris	p90-160 gag-abl	membrane plasmique	tyrosine phosphokinase
v-fgr	virus de sarcome félin souche Gardner-Rasheed	chat	p70 gag-fgr	?	tyrosine phosphokinase
v-ros	virus de sarcome aviaire UR2	poulet	p68 gag-ros	?	tyrosine phosphokinase
v-erbB* (2)	virus de l'erythroblastose aviaire, souche ES4	poulet	gp 74erbB	membrane plasmique	tyrosine phosphokinase
v-fms* (2)	virus sarcomatogène félin, souche Mac-Donough	chat	p120-180 gag-fms	membrane du réticulum	tyrosine phosphokinase
v-raf	virus sarcomatogène murin, souche 3611	rat	p75 gag-raf	?	?
v-mil	virus de la myelocytomatose aviaire, souche MH2	poulet	p100 gag-mil	cytoplasme	?
v-mos (2)	virus du sarcome murin de Moloney	souris	p37 mos	cytoplasme	?
v-rel	virus de la reticulo-endotheliose, souche T	dinde	p56 rel	?	?
---					
v-myc (4)	virus de la myelocytomatose aviare, souche MC29	poulet	p110 gag-myc	noyau	protéine se liant à l'ADN
v-myb (2)	virus de la myeloblastose aviaire, souche BAI-A	poulet	p45 myb	noyau	?
---					
v-Ha-ras 1 (2)	virus sarcomatogène murin de Harvey	rat	p21 H-ras	membrane plasmique	fixent le GTP et le GDP ;
v-Ki-ras 2	virus sarcomatogène murin de Kirsten	rat	p21 K-ras	membrane plasmique	Activité thréonine phosphokinase sur la protéine virale uniquement
v-bas	virus de sarcome de la souris BALB/C	souris	p21 bas	membrane plasmique	
---					
v-fos (2)	virus d'osteosarcome murin, souche FBJ	rat	p55 fos	noyau	?
v-ski	virus sarcomatogène aviaire, souche SKV770	poulet	p110gag-ski	noyau	?
v-sis (2)	virus sarcomatogène simien	singe laineux	p28 sis	cytoplasme	?
v-ets	virus erythroblastosant aviaire, souche E26	poulet	p135 gag-myb-ets	noyau	?
v-erbA	virus de l'erythroblastose aviaire, souche ES4	poulet	p75 gag-erbA	cytoplasme	?
v-kit	virus sarcomatogène félin, souche HZ4	chat	p80gag-kit	?	

* glycoproteine ; ( ) : nombre d'isolats indépendants

BU LILLE

par le virus helper félin FeLV ont révélé des provirus FeLV défectifs ayant recombiné avec le gène c-myc de chat (Neil et al., 1984).

Neanmoins, la création d'un virus hautement transformant par recombinaison -illégitime ou légitime mais en plusieurs étapes-avec un gène cellulaire approprié et transduction dans un virus infectieux demeure un phénomène d'une exceptionnelle rareté dont les bases moléculaires sont encore mal comprises. Quatre modèles, faisant intervenir soit des recombinaisons entre des intermédiaires ADN, soit des sauts de la rétropolymérase de l'ARN viral à un ARN cellulaire co-encapsidé dans un virion ont été proposés (Graf et Stéhelin, 1982 ; Bishop, 1983). Dans le cas du virus FBJ (virus d'ostéosarcome murin), la séquence de son oncogène v-fos a révélé à l'extrémité 5' et 3' de cette insertion une région de 6 nucléotides retrouvée à la fois dans le virus helper (FBJ-MuLV) et dans le gène cellulaire c-fos qui lui ont donné naissance. Ceci constitue à l'heure actuelle, l'un des rares exemples d'une possible recombinaison légitime grâce à des homologies de séquence (Van Beveren et al., 1983).

Propriétés des proto-oncogènes cellulaires

L'étude de la phylogénie de ces gènes a montré qu'ils étaient conservés tout au cours de l'évolution des vertèbrés supérieurs. En conséquence, les homologues cellulaires de gènes oncogènes viraux isolés chez l'oiseau ou le rat sont présents dans l'ADN humain (Bishop, 1983). De plus, certains proto-oncogènes cellulaires sont apparus très tôt dans l'évolution de la cellule eucaryote puisque des gènes apparentés à certains oncogènes viraux ont pu être détectés chez des

invertébrés comme Drosophila melanogaster (Shilo and Weinberg, 1981 ; Bishop, 1982) et chez des organismes eucaryotes unicellulaires comme la levure Saccharomyces cerevisia (De Feo-Jones et al., 1983 ; Gallwitz et al., 1983).

Ces proto-oncogènes cellulaires sont en général présents à l'état d'une copie par génome haploïde, avec quelques notables exceptions. Par exemple, les gènes ras constituent une famille de gènes dans l'ADN murin (Ellis et al., 1981) et dans l'ADN humain (Capon et al., 1983). De même, des équivalents tronqués et probablement non fonctionnels du gène c-myc ("pseudogènes") existent dans l'ADN humain (Dalla-Favera et al., 1982c) ainsi qu'un gène apparenté : N-myc (Kohl et al., 1983).

Les proto-oncogènes cellulaires sont exprimés à l'état constitutif dans des cellules normales à l'exception toutefois du gène c-mos. Cette expression peut se limiter à des types cellulaires très précis tel c-erb dans des précurseurs de la lignée érythroïde (Coll et al., 1983a) ou être modulée dans certains tissus en fonction de leur stade de développement, tels c-myb et c-myc dans des cellules hématopoïétiques immatures, aviaires ou humaines (Coll et al., 1983a ; Westin et al., 1982) ou les gènes c-fos, c-abl et c-fms au cours de l'embryogénèse chez la souris (Müller et al., 1982). En revanche, l'expression des gènes c-Ha-ras et c-Ki-ras dans les mêmes types cellulaires ne varierait pas au cours du développement embryonnaire (Müller et al., 1982).

Toutes ces données -forte conservation dans l'évolution et expression dans les cellules normales- suggèrent que les proto-oncogènes pourraient jouer un rôle important dans la physiologie de la cellule eucaryote et/ou dans des processus de différenciation.

Transformation par les oncogènes viraux

Si les proto-oncogènes jouent un rôle important dans le développement des cellules normales, comment expliquer leur pouvoir oncogène lorsqu'ils sont transduits par des rétrovirus ?

Deux hypothèses non mutuellement exclusives ont été proposées. L'hypothèse qualitative rend compte des différences structurelles existant entre le v-onc et le proto-oncogène ; à cette différence de structure et donc de produit pourrait correspondre une modification fonctionnelle. Alors que les c-oncs possèdent généralement une structure en mosaïque (introns et exons) au niveau de l'ADN, l'information codante du v-onc n'est jamais séparée par des introns. Ceci indique que le mécanisme de transduction s'accompagne le cas échéant d'une excision-raboutage correcte (Bishop, 1983). De plus, l'oncogène viral ne représente généralement qu'une partie du gène cellulaire équivalent, puisque sa taille est inférieure à celle de l'ARN messager mature du c-onc (Bishop, 1983). A ces deux phénomènes viennent s'ajouter des mutations ponctuelles dues à des erreurs introduites par la rétropolymérase lors de la réplication par le virus de la séquence v-onc. Cette réplication, contrairement à celle du c-onc est beaucoup plus fréquente et n'est pas soumise au contrôle strict de la cellule. Finalement, l'expression de la séquence transformante qui découle de son lieu d'insertion dans le virus peut entrainer des modifications de conformation de la protéine. En effet, si dans certains cas le v-onc est exprimé sous forme d'une protéine onc seule, traduite à partir d'un ARN messager sous génomique (Klempnauer et al., 1982 ; Saule et al., 1983), dans d'autres cas il est exprimé sous forme d'une protéine

de fusion avec des gènes de structure du virus (généralement gag) (Hayman et al., 1979). Cette liaison du v-onc au gène gag peut ne pas être neutre. Dans le cas du virus de la leucémie d'Abelson dont le produit est une p^{140} gag-abl, des mutants dans le gène gag ont perdu la capacité de transformer des cellules lymphoïdes tout en gardant celle de transformer les fibroblastes (Prywes et al., 1983).

La seconde hypothèse -hypothèse quantitative ou de dosage-propose que les produits des gènes v-onc et c-onc, quoique structurellement différents, seraient fonctionnellement identiques. L'augmentation de la quantité de la protéine onc dans la cellule infectée provoquerait la transformation de cette dernière. La transduction du proto-oncogène cellulaire par le virus élimine généralement la partie 5' du gène cellulaire et donc ses signaux normaux de promotion. Le taux de transcription du v-onc est alors placé sous le contrôle des signaux de promotion et de transcription propres au virus et contenus dans le LTR. De même, l'observation de l'activité transformante par transfection sur cellule NIH 3T3 -technique qui sera décrite plus loin- des proto-oncogènes c-mos et c-ras clonés et auxquels avaient été ligaturés des LTR viraux est en faveur de cette hypothèse (Oskarson et al., 1980 ; Blair et al., 1981 ; Chang et al., 1981). Récemment, l'étude du gène c-fos a révélé qu'une protéine cellulaire normale peut induire la transformation quand elle est exprimée dans un type cellulaire, où son expression est normalement inhibée (Miller et al., 1984).

Ces hypothèses apparaîssent particulièrement séduisantes au vu de la relation existant entre les oncogènes et les agents mitogènes (Heldin et Westermark, 1984). Un facteur de croissance, l'"Epidermal Growth Factor" (EGF) semble activer la division

cellulaire par l'intervention de phosphoprotéines kinases spécifiques de la Tyrosine, dont l'une phosphoryle une des cibles protéiques de la pp60 src (Hunter et Cooper, 1981). Il convient de rappeler, à ce propos, que les phénomènes de phosphorylation des amino acides dans les cellules normales concernent majoritairement la Sérine et la Thréonine et pour une part mineure (0.1 %) la Tyrosine alors que les produits de nombreux v-onc, dont v-src, sont des Tyrosine-phosphokinases (Table 1). De plus, les récepteurs de nombreux facteurs de croissance tels ceux de l'"EGF" et du "PDGF"("Platelet-Derived Growth Factor") possèdent une activité Tyrosine-phosphokinase. Cette activité -bien que quantitativement minoritaire- semble donc être en étroite relation avec la prolifération cellulaire induite soit par les facteurs de croissance, soit par la transformation par certains rétrovirus (Heldin et Westermark, 1984). Récemment une des cibles des Tyrosines-phosphokinases produits des oncogènes v-src et v-ros a été identifiée : ils activent le métabolisme de dérivés lipidiques de l'inositol, métabolisme dont l'activation est en rapport avec une augmentation de la prolifération cellulaire (Macara et al., 1984 ; Sugimoto et al. 1984).

D'autre part, la comparaison entre l'enchainement des amino-acides de la glycoprotéine gp 74^{erbB}, determiné à partir de la séquence nucléotidique de l'oncogène viral v-erbB et celui du récepteur de l'"EGF" indique que le produit de v-erbB est une forme tronquée du récepteur de l'"EGF" (Downward et al., 1984). L'information absente dans v-erbB correspond à l'extrémité NH2 terminale du récepteur incluant le site de fixation de l'"EGF". Une forte homologie existe également entre la séquence nucléotidique de l'oncogène v-sis du virus sarcomatogène de singe (SSV) et à celle d'un autre facteur mitogène humain, le "PDGF"

(Doolittle et al., 1983 ; Waterfield et al., 1983). De plus, le premier produit de traduction de v-sis, $p28^{sis}$, subit in vivo des étapes de maturation aboutissant à la formation de molécules structurellement et immunologiquement apparentées à la forme biologiquement active du "PDGF" (Robbins et al., 1983 ; Niman, 1984). Le gène c-sis étant présent à une copie par génome haploïde, toutes ces données ainsi que des résultats récents démontrent que le gène cellulaire c-sis est le gène codant pour le précurseur de la chaine B du "PDGF" (Johnson et al., 1984).

Combien existe-t-il d'oncogènes viraux ?

De nombreux faits expérimentaux indiquent que la liste des 24 oncogènes viraux connus à l'heure actuelle, pourrait représenter un catalogue presque exhaustif. En effet, le même oncogène peut apparaître de façon répétée dans des isolats viraux indépendants, tels v-myc trouvé dans 4 virus myélocytomatosants aviaires ou v-fps transduit par 5 virus aviaires sarcomatogènes défectifs (Table 1 ; Bishop, 1983).

Dans des espèces animales différentes, les rétrovirus peuvent transduire le même proto-oncogène cellulaire : c-fes, transduit par un virus sarcomatogène de chat est l'homologue félin du gène c-fps révélé par des virus aviaires (Hampe et al., 1982). Le gène v-bas est l'équivalent chez la souris du gène v-Ha-ras isolé dans un virus de rat (Andersen et al., 1981).

Enfin, l'étude d'isolats récents de rétrovirus révèle des oncogènes déjà connus, par ailleurs, dans d'autres virus. Les oncogènes de 2 nouveaux virus de chat étaient équivalents respectivement à l'oncogène v-sis du virus sarcomatogène de singe et à v-abl du virus de la leucémie murine d'Abelson (Bessmer et

al., 1983). L'oncogène v-mil du virus aviaire MH2 (Coll et al., 1983b) correspond à l'oncogène v-raf d'un virus sarcomatogène murin, (Jansen et al., 1984). Ceci n'exclue pas l'existence d'autres proto-oncogènes cellulaires (voir plus loin), mais le nombre de ceux compatibles avec la transduction par des rétrovirus est probablement compris entre 30 et 50 (Varmus, 1982 ; Weinberg, 1982).

Les oncogènes viraux constituent au moins 3 familles

La comparaison des séquences nucléotidiques de différents oncogènes viraux ou de leurs équivalents cellulaires chez des organismes inférieurs ont montré que des oncogènes ne présentant pas d'homologie par hybridation moléculaire et/ou ayant des propriétés tumorigéniques différentes étaient apparentés et dérivaient probablement d'un ancêtre commun. On peut distinguer:

La super famille des Tyrosines phosphokinases comprenant : v-src, v-yes, v-fps, v-fes, v-fgr, v-erbB, v-fms et v-abl . Prenant comme référence la séquence en acides aminés de la $pp60^{src}$ les homologies avec les différentes protéines sont les suivantes : yes, 82 % ; fes, 45 % ; fps 40 % et mos 25 %. Par homologie de séquence, viennent s'ajouter des oncogènes sans activité kinasique démontrée tel v-rel (Stephens et al., 1983) v-mos (Papkoff et al., 1983) et v-mil (Galibert et al., 1984). Il est à noter que toutes les protéines virales équivalentes sont à localisation membranaires sauf le produit de v-mos et de v-mil qui seraient cytosolubles (Papkoff et al., 1983 ; Bunte et al., 1983).

La famille des gènes v-Ha-ras et v-Ki-ras qui bien que faiblement apparentée au niveau des acides nucléiques, codent pour des protéines aux propriétés antigéniques analogues (Ellis et al., 1981 ; Pincus et al., 1983). Ces protéines ont une forte affinité pour les nucléotides tels le GDP ou le GTP.

Enfin, les oncogènes v-myb et v-myc codant pour des protéines nucléaires seraient apparentés entre eux et avec la protéine EIa des Adenovirus (Ralston et al., 1983). Les protéines nucléaires gag-myc des virus MC29 et CMII ont en outre la propriété de se lier à l'ADN (Donner et al., 1982 ; Bunte et al., 1983).

En admettant que des homologies de structure et de séquences impliquent des analogies fonctionnelles, il pourrait n'y avoir qu'un petit nombre de fonctions enzymatiques différentes codées par les oncogènes, chaque fonction étant caractéristique d'une famille. Il reste à déterminer si les oncogènes restants non encore classés ou ceux à découvrir, seront rangés dans l'une ou l'autre de ces familles, ou en révèleront de nouvelles (Bishop, 1983).

IMPLICATIONS DES PROTO-ONCOGENES CELLULAIRES DANS LES TUMEURS.

Les rétrovirus transformants ont constitué un système révélateur de choix permettant la détection -parmi les quelques 50 000 gènes que compte une cellule normale- d'une vingtaine de gènes qui deviennent oncogènes lorsqu'ils subissent des modifications structurelles et passent sous le contrôle de promoteurs viraux. Dans les tumeurs humaines ou animales, non

induites par des rétrovirus transformants, l'une des cibles du produit cancérigène pouvait donc être l'un des homologues cellulaires de ces oncogènes viraux et/ou d'autres proto-oncogènes cellulaires non compatibles avec la transduction dans un rétrovirus aviaire, félin ou murin.La découverte de ces gènes nécessitaient donc la mise au point et l'utilisation d'autres techniques d'identification. L'étude de tumeurs par hybridation moléculaire à l'aide de sondes spécifiques des oncogènes viraux et/ou par la technique de transfection ont respectivement permis de répondre à ces deux questions.

Les techniques d'hybridation moléculaire reposent sur la complémentarité des acides nucléiques et mesurent des phénomènes de réassociation entre un acide nucléique (ADN dissocié ou ARN) et sa séquence complémentaire marquée grâce à un isotope radioactif. Les 3 techniques existantes -en phase liquide, en phase solide ou "in situ" sur cellules ou chromosomes isolés-fournissent des données différentes et complémentaires. L'hybridation, en phase liquide, permet entre autres choses la quantification de l'acide nucléique étudié (nombre de copies par cellules ou par génome haploïde) alors que, l'hybridation en phase solide fournit essentiellement des données d'ordre structurel (taille de l'ARN messager, carte physique d'un ADN digéré par des enzymes de restriction). L'hybridation "in situ", quant à elle, permet par exemple, de détecter dans une population hétérogène les cellules exprimant un gène ou de préciser sa localisation chromosomique.

La transfection consiste à transférer en vue de son expression l'information génétique (ADN) d'une cellule dans une autre. L'ADN est introduit dans le milieu de culture sous forme d'un précipité insoluble en présence de phosphate de calcium,

précipité qui sera phagocyté par les cellules. Dans le cas des gènes transformants une lignée établie de cellules de souris, la NIH 3T3 s'est révélée efficacement transformable par transfection par certains oncogènes viraux ou certains oncogènes cellulaires activés (Cooper, 1982 ; Weinberg, 1982). En effet, lorsqu'on cultive des NIH 3T3, quelques rares foyers de cellules transformées apparaissent spontanément parmi la monocouche de cellules normales. La présence de certains gènes activés dans des tumeurs est mise en évidence par une augmentation considérable du nombre de ces foyers après transfection de leur ADN sur NIH 3T3.

Ces techniques ont révélé -outre leur transduction dans des rétrovirus- 5 autres mécanismes possibles d'activation des proto-oncogènes cellulaires soit :

* par intégration d'un rétrovirus sans oncogène au voisinage d'un proto-oncogène dans l'ADN cellulaire. Ce mécanisme est connu sous le nom de la "LTR-activation" ou mutagenèse par insertion de promoteur ;
* par des mutations spontanées au niveau des bases puriques et pyrimidiques ;
* par des translocations chromosomiques ;
* par des phénomènes d'amplification génique.
* Bien que totalement artificielle, il convient de signaler l'activation par l'attachement "in vitro" d'un LTR à un proto-oncogène cellulaire : c-mos (Blair et al., 1981) et c-Ha-ras 1 (Chang et al., 1982). La molécule chimérique ainsi obtenue provoque, contrairement au gène cellulaire normal, l'apparition de foyers de transformation après transfection dans la lignée NIH 3T3.

Chaque mécanisme est brièvement illustré ci-dessous par quelques exemples expérimentaux.

Intégration d'un rétrovirus au voisinage d'un proto-oncogène cellulaire : la "LTR activation" ou oncogénèse par insertion de promoteur.

Ce mécanisme fut découvert grâce à l'étude des tumeurs -principalement des leucémies lymphoïdes de type B- induites chez le poulet par des rétrovirus (ALV) ne possédant pas d'oncogènes mais uniquement des gènes de structure viraux nécessaires à leur réplication. Ces tumeurs apparaissent 6 à 12 mois après l'infection, temps relativement long comparé aux 2 à 4 semaines nécessaires aux virus hautement transformants pour tuer l'animal. Bien que les ALV puissent s'intégrer au hasard dans l'ADN de la cellule hôte, un provirus tronqué mais contenant au moins une copie du LTR était intégré dans 85 % de ces tumeurs, au niveau de 4 sites spécifiques situés en amont (5') du gène c-myc , l'homologue cellulaire de l'oncogène v-myc des virus myélocytomatosants aviaires (Roussel et al., 1979). Ces tumeurs exprimaient de 80 à 400 copies d'un ARN messager hybride contenant la région 5' du LTR et 2 des 3 exons du gène c-myc (Hayward et al., 1981). Par comparaison, on ne détecte que 2 à 5 copies du gène c-myc par cellule normale de la bourse de Fabricius -organe donnant naissance aux lymphocytes B- d'un poulet adulte. Le rétrovirus, comme dans le cas des éléments génétiques mobiles, était mutagène par insertion : les signaux de promotion contenus dans le "LTR" activaient la transcription du proto-oncogène adjacent. Cependant, le LTR 3' d'un provirus aviaire ne peut se comporter en promoteur de transcription

efficace s'il se trouve en aval d'un LTR (5') transcriptionnellement actif. L'absence de ce LTR (5') afin de lever cette interférence de transcription pourrait expliquer l'observation répétée de provirus tronqués, insérés en amont du gène c-myc dans les lymphomes (Cullen et al., 1984). Cette hypothèse séduisante laisse néanmoins quelques points obscurs, puisque dans certaines tumeurs le LTR est situé en 3' du gène c-myc ou en 5' mais dans la direction opposée à sa transcription suggérant dans ce cas un effet indirect sur un promoteur cellulaire cryptique (Payne et al., 1981). De plus, dans 15 % des tumeurs lymphoïdes induites par les ALV, le gène c-myc n'est pas activé.

D'autres proto-oncogènes cellulaires peuvent être touchés par ce type de mutations qui en activant leur expression aboutit à l'induction de tumeurs. Dans l'érythroblastose induite par un ALV chez une lignée de poulet, le provirus est intégré au niveau de c-erbB, l'équivalent cellulaire de l'oncogène du virus de l'érythroblastose aviaire (Fung et al., 1983).

Dans des plasmacytomes de souris, le gène c-mos est activé par un élément endogène, ressemblant à des rétrovirus -particules intracisternales de type A, "IAP"- (Cohen et al., 1983). Cet exemple est particulièrement intéressant puisqu'aucune transcription du gène c-mos n'avait pu être détectée dans des cellules normales (Bishop, 1983).

Activation par mutation ponctuelle

L'effet transformant des c-oncs pouvait être du à des mutations ponctuelles aboutissant à une altération de la structure -et peut-être de la fonction- de la protéine codée par le gène. Ce mécanisme a été le mieux documenté dans le cas des produits des gènes c-ras.

Deux lignées cellulaires -EJ et T24- établies à partir de tumeurs de la vessie contenaient des gènes transformants détectés grâce au test de transfection sur NIH 3T3 (Goldfarb et al.,1982 ; Shih et Weinberg, 1982). Après 3 cycles successifs de transfection -ceci afin de purifier par dilution le ou les gènes impliqués dans la transformation- une banque de gènes était réalisée à partir de l'ADN d'un foyer tertiaire. Le gène transformant humain a pu alors être différencié du génome de la NIH 3T3 et cloné. Le gène transformant ainsi isolé était l'homologue cellulaire humain, c-Ha-ras1, de l'oncogène viral v-Ha-ras présent dans un virus de sarcome murin (Goldfarb et al., 1982). Le fragment du gène contenant l'activité transformante a pu être défini par la construction et l'étude par transfection sur NIH 3T3 de chimères entre le c-Ha-ras 1 normal et son homologue activé.Sa séquence nucléotidique comparée à celle du gène c-Ha-ras 1 complet isolé à partir d'une banque d'ADN humain normal a révélé une seule mutation ponctuelle parmi les nucléotides séquencés. Cette mutation ne s'accompagnait d'aucune augmentation de la transcription du gène c-Ha-ras 1 dans la tumeur. Ce nucléotide différent correspondait au 3ème nucléotide du 12ème codon de la phase ouverte du gène : cette transition de G vers T entraîne le remplacement d'une Glycine par une Valine (Shih et Weinberg, 1982 ; Goldfarb et al., 1982).

A titre de comparaison, l'oncogène v-Ha-ras ne présente que 3 différences avec le gène c-Ha-ras 1 normal dont l'une touche le 12ème Acide Aminé qui est une Arginine et l'autre le 59ème où une Thréonine remplace une Alanine. Cette Thréonine serait le site d'auto-phosphorylation de la protéine virale (Capon et al., 1983). Des études de conformation indiquent que le changement du 12ème Acide Aminé provoque un bouleversement de la structure tridimensionnelle de la protéine . Ceci est à corréler avec la différence de migration électrophorétique observée entre la protéine p21c-Ha-ras normale et la p21 codée par le gène activé. Récemment, le même type du mutation ponctuelle touchant le 12ème Amino Acide du c-Ha-ras 1 a été mis en évidence chez le rat dans des carcinomes mammaires provoqués par la nitroso-urée. Ce système expérimental constituera un modèle intéressant pour l'étude de l'activation des gènes ras in vivo et pour l'étude de leurs propriétés tumorigéniques (Sukumar et al., 1983). L'étude d'autres tumeurs solides a révélé l'activation mutationnelle d'autres membres de la famille ras : le gène c-Ki-ras-2 l'homologue cellulaire de l'oncogène du virus du sarcome murin de Kirsten (Table 1) est activé dans des carcinomes du poumon et du colon, et le gène N-ras -apparenté aux gènes c-Ha-ras et c-Ki-ras mais non transduit par les rétrovirus- impliqué dans des neuroblastomes. Les données de séquence -non encore disponibles pour le gène N-ras- semblent indiquer un phénomène d'activation plus complexe dans le cas de c-Ki-ras-2. Il y aurait d'autres mutations ponctuelles, outre celles touchant le 12ème Acide Aminé, accompagnées d'un épissage différent au niveau du 4ème exon de ce gène. (Mac Grath et al., 1983).

Cependant, l'activation par mutation ponctuelle des gènes ras dans les tumeurs humaines, n'affecte ni la localisation cellulaire, ni les modification post-transcriptionnelles, ni l'affinité pour les nucléotides de la p21 (Finkel et al., 1984b). Ces résultats suggèrent plutôt une altération des intéractions de la p21 avec d'autre (s) protéine (s) cellulaire (s). Ce modèle permet en outre, d'expliquer l'activation des gènes ras à la fois par une variété de mutations structurelles ou par l'augmentation de l'expression du gène normal (Chang et al., 1982). L'une des cibles de la p21 semble être le récepteur de la transferrine (Finkel et al., 1984a).

Activation par translocations chromosomiques

L'association étroite existant entre des translocations chromosomiques et certains cancers humains est connue depuis longtemps (pour revue récente, voir Yunis, 1983). Ces translocations sont de deux types :

* les translocations constitutionnelles qui sont présentes dans chaque cellule de l'individu et le prédisposent à un type de cancer ;
* les translocations somatiques qui apparaissent dans une cellule particulière et se retrouvent dans sa descendance néoplasique.

La translocation t(9:22) (chromosome Philadelphie) dans les leucémies myéloïdes chroniques et la haute incidence des translocations t(2:8) ; t(8:14) ou t(8:22) dans les lymphomes de Burkitt étaient un argument en faveur de leur implication probable dans le processus cancéreux (Cairns, 1981).

Dans le cas des lymphomes de Burkitt, les chromosomes impliqués 2,14 et 22 portent les gènes codant pour les chaînes lourdes et légères d'immunoglobulines. Puisque ces translocations se déroulent dans des cellules productrices d'anticorps, Klein suggéra en 1981, qu'elles plaçaient un proto-oncogène non exprimé dans ces cellules sous le contrôle d'un promoteur de transcription d'Immunoglobuline. Sans entrer dans le détail de la structure très complexe des loci d'immunoglobulines et des réarrangements aboutissant à la synthèse d'une classe précise d'immunoglobuline, il convient de signaler la présence de nombreux promoteurs de transcription et d'une séquence ("enhancer") augmentant celle-ci spécifiquement dans les cellules B (Tonegewa, 1983).

La localisation du gène c-myc humain sur le chromosome 8 au niveau du point de rupture dans les tumeurs, bande 8q24, -bien qu'une bande aussi définie représente environ 10^7 paires de bases- était en faveur de l'hypothèse de Klein (Dalla-Favera et al., 1982b ; Neel et al., 1982). Néanmoins le clonage et l'étude structurelle du gène c-myc transloqué dans différentes tumeurs de Burkitt , si elle confirmait une translocation allélique de ce gène le liant à un locus d'immunoglobuline, infirma cette hypothèse dans sa forme la plus simple. En effet, dans la majorité des cas le gène c-myc et le gène d'immunoglobuline ont des orientations transcriptionnelles opposées, éliminant de ce fait la participation du promoteur de transcription. De plus la séquence "enhancer" -insensible par définition à l'orientation de la transcription- n'est généralement pas disponible, puisqu'elle est transférée sur l'autre chromosome du fait du mécanisme d'échange réciproque de la translocation (Rabbits et al.,1983 ; Hayday et al., 1983). Leder a récemment proposé un modèle de

dérégulation du gène c-myc dont la cible serait le premier exon de ce gène (Leder et al., 1983). Ce 1er exon est en fait une séquence transcrite mais supposée non traduite en protéine et contenant les 2 promoteurs de transcription normaux du gène (Battey et al., 1983). La translocation soit en éloignant cet exon des 2 exons codant de c-myc soit en y introduisant des mutations le rendrait insensible à l'action d'une protéine régulant, le taux d'utilisation des deux promoteurs de transcription du gène c-myc à la façon des systèmes opérateur-represseur bactériens. Cette protéine régulatrice pourrait d'ailleurs être codée par le gène c-myc lui même puisque récemment un cadre de lecture ouvert, pouvant diriger la synthèse d'une protéine putative de 20 kd a été mis en évidence dans le 1er exon (Gazin et al., 1984). La mise en évidence in vivo de cette protéine putative $p20^{c\text{-}myc}$ ainsi qu'une étude précise du taux de production de la protéine c-myc humaine dans les cellules normales et tumorales seront nécessaires pour confirmer ce modèle.

Dans le cas du chromosome Philadelphie, le gène c-abl est transféré du chromosome 9 sur le chromosome 22 (de Klein et al., 1982). Le point de cassure semble être situé dans le gène cellulaire c-abl en 5' de la région homologue à v-abl, l'oncogène transduit dans un virus murin (Heisterkamp et al., 1983). La situation pourrait néanmoins être plus complexe, puisque réciproquement le gène humain c-sis est transloqué du chromosome 22 sur le chromosome 9 (Heisterkamp et al., 1983).

Activation par amplification génique

De même que des cellules devenaient résistantes à une drogue comme le méthotrexate par amplification spécifique des gènes de la dihydrofolate réductase (Schimke, 1982), le phénotype cancéreux pouvait être dû à une amplification de certains proto-oncogènes cellulaires aboutissant à leur surexpression. Le proto-oncogène c-myc est ainsi amplifié de 30 à 50 fois dans la lignée cellulaire promyélocytique humaine HL60, (Collins et Groudine, 1982), de même le gène c-Ki-ras dans des cellules de tumeur adréno-corticale de souris (Schwab et al., 1983). Dans cet exemple précis, l'amplification au niveau de l'ADN est corrélée à une augmentation significative (de 30 à 60 fois) de l'ARN messager et de la protéine P21c-Ki-ras codée par le gène amplifié.

Dans certains cas ces gènes amplifiés ont pu être associés à des aberrations du chromosome en métaphase connues depuis longtemps par l'étude cytogénétique des tumeurs. Ces anomalies sont de 2 types soit des régions qui prennent les colorants chimiques de façon homogène et non sous forme de bandes alternées, HSRs ("Homogenous Staining Regions") soit des petites particules de chromatine associées par paires, DMs ("Double Minute"). Par exemple, le gène N-myc, apparenté au gène c-myc, est ainsi amplifié 50 fois dans un neuroblastome humain au niveau d'HSRs (Kohl et al., 1983).

VERS UNE DEFINITION MOLECULAIRE DES DIFFERENTES ETAPES DE LA CANCEROGENESE

Activation de plusieurs oncogènes dans la même tumeur

L'étude des oncogènes viraux et cellulaires, si elle permettait de dresser un catalogue d'agents intervenants dans la transformation cellulaire ne fournissait néanmoins aucun éclaircissement sur les différentes étapes de la cancérisation in vivo. Il apparaissait en effet difficile de concilier le processus multi-étapes de la cancérogénèse spontanée ou chimio-induite avec l'événement unique que représentait l'activation du c-Ha-ras1 par une unique mutation ponctuelle . En conséquence, cette activation pouvait ne représenter qu'une des étapes du processus cancéreux décelable sur NIH 3T3 car elle terminait le processus dans lequel était engagé cette cellule. La première confirmation de cette hypothèse vint de l'étude des lymphomes aviaires induits par les ALVs. L'ADN extrait de ces lymphomes contenait un gène transformant détecté par le test de transfection sur NIH 3T3 (Cooper et Neiman, 1980). Néanmoins, ce gène n'était pas le c-myc LTR-activé (Hayward et al., 1981) , mais un autre gène, B-lym présentant une homologie avec la famille des transferrines (Cooper et Neiman, 1980 ; Goubin et al., 1983). Cette observation pourrait être corrélée à l'apparition en deux temps de la maladie. A la suite de l'infection par l'ALV, la première évidence histologique de l'état prénéoplasique est l'apparition de follicules transformés dans la Bourse de Fabricius. La majeure partie de ces follicules régresse mais une faible proportion, environ 2 %, donne naissance à des nodules qui évolueront ensuite sous forme de leucémies

lymphoïdes disséminées (Cooper et Neiman, 1980). Ces 2 étapes dans la progression de la cellule normale vers une tumeur clonale (Hayward et al., 1981 ; Payne et al., 1981) pourrait donc concorder avec les événements moléculaires décrits ci-dessus.

Ce modèle d'activation successive de plusieurs gènes ("Hit and Run"), peut se généraliser puisque dans le cas des lymphomes humains de Burkitt -touchant également des cellules lymphoides de type B- ce n'est pas le gène c-myc transloqué qui est transformant sur NIH 3T3, mais le gène B-lym humain (Diamond et al., 1983). Il s'applique également à certaines tumeurs murines induites par le virus d'Abelson où l'on observe des remaniements du gène c-myb (Mushinski et al., 1983) ou l'activation secondaire d'un gène transformant non encore identifié (Lane et al., 1983).

Pouvoir limité d'un oncogène et coopération entre oncogènes

Cet autre aspect de la tumorigénèse multi-étapes découle de l'étude du test de transfection sur NIH 3T3 (Land et al., 1983b).

L'isolement répétitif de versions altérées de gènes ras par transfection d'ADNs extraits de lignées cellulaires établies à partir de tumeurs solides laissait planer quelques doutes quant à la fiabilité de ce test : ceci pouvait refléter soit un artéfact dû au passage de la cellule transformée en lignée établie ou être dû à l'efficacité particulière de ces gènes pour complémenter et terminer le processus tumoral dans lequel était engagé cette cellule. L'isolement du même gène ras activé par transfection sur NIH 3T3 de l'ADN extrait à partir de pièces opératoires de la tumeur primaire infirmait la première hypothèse (Land et al., 1983a et b). Cependant, bien que sélectionné par son aptitude à transformer les cellules NIH 3T3, le gène

c-Ha-ras1 activé de la lignée EJ transfecté sur des cellules normales, -des fibroblastes embryonnaires de rat (REF) en passages précoces- ne les transformait pas pleinement et ne leur conférait qu'un des critères de la transformation : la croissance en milieu semi-solide (Land et al., 1983a). En revanche, ce même proto-oncogène cellulaire activé transformait efficacement des fibroblastes fraîchement immortalisés par des carcinogènes (Newbold et Overell, 1983). Dans le cas de cellules normales, le c-Ha ras EJ ne semblait pouvoir devenir tumorigène qu'en synergie avec les phénomènes d'immortalisation des cellules puisque la cotransfection de ce gène avec les produits de certains virus à ADN capable d'induire la croissance continue des cellules en culture tels le "large T antigen" de polyome ou le E1a des adénovirus transformait pleinement les REFs alors que chaque produit pris isolément en était incapable (Land et al., 1983a ; Ruley, 1983). Ces expériences établissaient clairement les limites du pouvoir d'un oncogène donné et de plus définissaient des processus de complémentation et de potentialisation entre oncogènes. Un autre exemple de complémentation a été suggéré par l'étude de la lignée promyélocytique humaine HL60. En effet, dans cette lignée si le gène c-myc était amplifié de 30 à 50 fois, c'est un gène apparenté à la famille ras, le N-ras, qui était détecté par transfection sur NIH 3T3. La complémentation entre myc et ras a été démontrée par la cotransfection du gène v-myc du virus MC29 et du c-Ha-ras EJ qui permettait la transformation des REFS (Land et al., 1983a).

En conclusion, ces études ont réconcilié les proto-oncogènes cellulaires avec la théorie multi-étapes de la cancérogénèse et également permis une ébauche de classement des différents

oncogènes en au moins 2 classes distinctes et complémentaires. La première comprendrait les produits des gènes myc, "large T antigen" de polyome, E1a des Adénovirus et les événements permettant l'immortalisation et le passage en lignée des cellules . La deuxième comprendrait le c-Ha-ras EJ, le N-ras et le " middle T antigen". Coïncidence frappante ou réalité expérimentale, les produits des gènes du 1er groupe sont tous des protéines nucléaires tandis que ceux du 2ème sont à localisation membranaire. Il reste à savoir si ce modèle peut se généraliser et si en fait chaque tumeur peut être due à l'activation d'au moins 2 gènes quelconques appartenant à chaque groupe et permettant de toucher respectivement une cible cellulaire nucléaire et une cible membranaire.

PRESENTATION DU TRAVAIL

Les virus défectifs des leucémies aigues aviaires (DLV) ont été classés en trois groupes selon le type de néoplasme hématologique qu'ils induisent in vivo (Graf et Beug, 1978 ; Graf et Stéhelin, 1982).

Le groupe des virus myéloblastosants comprend deux souches virales isolées indépendamment : AMV (Avian Myeloblastosis Virus) (Beard, 1963) et E26 (Ivanov et al., 1962).

Le virus AMV induit des myéloblastoses, maladie qui se traduit par l'apparition dans le sang de l'animal d'un nombre élevé de précurseurs de la lignée blanche (myéloblastes) bloqués dans leur différenciation. In vitro, AMV transforme des cellules myéloïdes immatures ainsi que des macrophages matures obtenus respectivement à partir de culture de cellules embryonnaires de moelle osseuse et de sac vitellin (Beug et al., 1979 ; Boettiger et Durban, 1984). AMV semble être jusqu'à présent le seul DLV dont la capacité transformante soit restreinte à la lignée hématopoïétique (Graf et Stéhelin, 1982).

Le virus E26 fut isolé en 1962 à partir d'une érythroblastose ("E") spontanée (Ivanov et al., 1962). Néanmoins, l'étude in vitro des cellules de moelle osseuse transformées par E26 révélait la présence de cellules myéloïdes (Beug et al., 1979) alors que les études in vivo confirmaient la nature érythroïde de la maladie : les cellules transformées par E26 présentes dans le sang circulant des poulets leucémiques réagissaient positivement avec un anticorps dirigé contre un marqueur antigénique érythroïde, l'Histone H5 (Sotirov et al., 1981). Ce paradoxe a été résolu récemment par la mise au point de milieux de culture permettant la croissance spécifique de cellules myéloïdes ou de cellules érythroïdes (Beug et al., 1982

; Radke et al., 1982 ; Moscovici et al., 1983). E26 induit *in vivo* une leucémie mixte érythroïde et myéloïde, à prédominance érythroïde, mais des cellules transformées appartenant à chaque lignée peuvent être mises en évidence selon les conditions de culture utilisées. Les cellules-cibles de E26 seraient recrutées parmi des cellules-souches pluripotentes (myéloïdes et érythroïdes) et parmi des précurseurs déjà engagés dans l'une ou l'autre des lignées de différenciation (Moscovici et al., 1983). De plus, E26 peut transformer, *in vitro*, des fibroblastes de caille -mais non de poulet- (Graf et al., 1979).

Aux différences biologiques présentées par ces deux virus s'ajoutent des différences de structure et de contenu génétique. Le clonage de l'ADN proviral d'AMV (Klempnauer et al., 1982) a montré que son ARN génomique d'une taille de 7.2 kb est constitué d'un gène *gag* complet et fonctionnel, de la majeure partie du gène *pol* et d'une séquence transformante spécifique, *v-myb* (Roussel et al., 1979 ; Gonda et al., 1981) à la place du gène *env*. Dans toutes les cellules infectées par AMV, on détecte un ARN viral sous génomique de 2.1 kb produit par épissure et ne contenant que *v-myb* et des signaux de régulation provenant des extrémités 5' et 3' du génome. La séquence nucléotidique de *v-myb* (Rushlow et al., 1982 ; Klempnauer et al., 1982) révèle que cet ARN code pour une protéine de 45-48 Kilodaltons (kd), $p45^{v\text{-}myb}$ (Boyle et al., 1983 ; Klempnauer et al., 1983a). Cette protéine $p45^{v\text{-}myb}$ présente une localisation nucléaire dans les myéloblastes transformés mais acquiert une localisation cytoplasmique lorsque ces derniers sont induits à se différencier en macrophages (Klempnauer et al., 1984b).

L'ARN viral de E26, quant à lui, hybride à 60 % avec une sonde AMV v-myb, ce qui explique son affiliation au groupe des virus myéloblastosants (Roussel et al., 1979 ; Stehelin et al., 1980). Le seul ARN détecté dans des cellules infectées par E26 est un ARN génomique de 5.7 kb comprenant de 5' vers 3', un gène gag tronqué, environ 60 % de l'oncogène viral AMVv-myb, une autre séquence spécifique d'origine cellulaire, v-ets, (Nunn et al., 1983 ; Leprince et al., 1983) et finalement une partie du gène env. La protéine transformante de ce virus serait une protéine de 135 kd, de structure gag-onc traduite à partir de l'ARN génomique : $p135^{gag\text{-}myb\text{-}ets}$ (Bister et al., 1982 ; Beug et al., 1982 ; Boyle et al., 1983 ; Klempnauer et al., 1984a). Cette protéine présente également une localisation nucléaire (Klempnauer et al., 1984b). En outre, sa synthèse est spécifiquement augmentée dans les myéloblastes transformés par l'action directe ou indirecte d'un facteur de croissance hématopoïétique exogène (Beug et al., 1982 ; Radke et al., 1982).

Le travail présenté dans ce mémoire traitera du clonage dans l'ADN aviaire et humain des séquences homologues à l'oncogène viral AMV v-myb : c-myb (poulet) et c-myb (homme).

D'autre part, l'étude de la transduction à partir de l'ADN poulet du même proto-oncogène cellulaire par un autre rétrovirus, E26, sera envisagée.

RESULTATS ET DISCUSSION

1ère partie : DETECTION ET CLONAGE DU PROTO-ONCOGENE CELLULAIRE c-myb DANS L'ADN DE POULET.

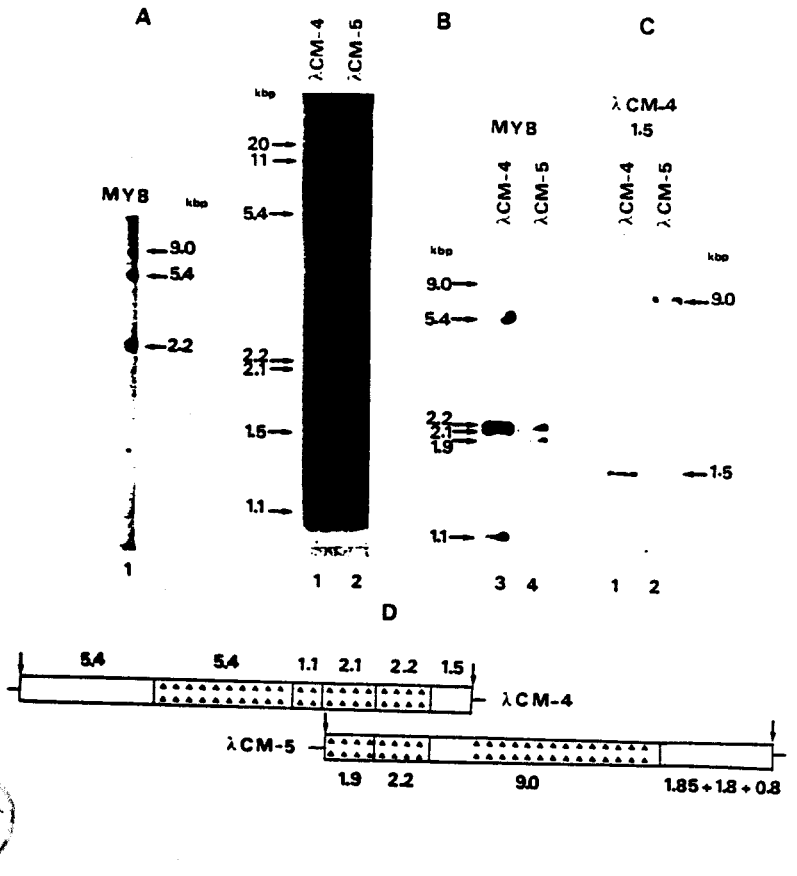
Avant d'aborder le clonage proprement dit la structure du locus c-myb dans l'ADN total de poulet a été étudiée par hybridation en phase solide.

Après digestion d'ADN haut poids moléculaire de fibroblastes de poulet, par l'enzyme de restriction EcoRI, les fragments obtenus sont séparés en fonction de leur taille par électrophorèse en gel d'agarose, dénaturés "in situ" dans le gel par traitement à la soude, transferrés sur une feuille de nitrocellulose puis fixés irréversiblement par "cuisson" à 80° C (Southern, 1975). Sur cette réplique du gel ou "Southern blot", les fragments d'ADN sont alors accessibles à l'hybridation avec des sondes spécifiques radioactives. Une sonde v-myb préparée à partir de l'ARN viral AMV BAI-A (Roussel et al., 1979 ; Gonda et al., 1981) révèle dans l'ADN total de poulet 3 fragments EcoRI positifs par hybridation et de taille respective : 9.0, 5.4 et 2.2 milliers de paires de bases (kbp : "kilo-base pair") (Figure 3A). Ces fragments appartiennent bien au gène c-myb et ne représentent pas une information virale endogène puisque :

* la sonde utilisée n'hybride qu'à 3 % avec de l'ARN d'un ALV "helper" par hybridation liquide.

* une sonde représentive du génome d'un ALV, détecte dans le même échantillon d'ADN des fragments de taille différente.

Ces séquences c-myb ont été isolées à partir d'une banque de fragments d'ADN (banque de gènes ou "génothèque") d'érythrocytes de poulet. Après digestion partielle par les enzymes de restriction AluI et HaeIII, les fragments d'ADN d'une taille de


A
λCM-4
λCM-5
kbp
20
11
5.4
2.2
2.1
1.5
1.1
MYB
kbp
9.0
5.4
2.2
1
1 2
B
MYB
λCM-4
λCM-5
kbp
9.0
5.4
2.2
2.1
1.9
1.1
3 4
C
λCM-4
1.5
λCM-4
λCM-5
kbp
9.0
1.5
1 2
D
5.4
5.4
1.1
2.1
2.2
1.5
λCM-4
λCM-5
1.9
2.2
9.0
1.85+1.8+0.8

BU LILLE

FIGURE 3 : Les bactériophages recombinants lambda CM-4 et lambda CM-5 contiennent le locus c-myb (poulet).

(A) Hybridation entre le cDNA myb et de l'ADN de poulet digéré par l'enzyme de restriction EcoRI, séparé par électrophorèse en gel d'agarose et transféré sur nitrocellulose selon la technique de Southern (1975).

(B) Visualisation par fluorescence du Bromure d'Ethidium sous lumière UV (canaux 1 et 2) et hybridation avec le cDNA myb (canaux 3 et 4) de l'ADN des bactériophages recombinants lambda CM-4 et lambda CM-5 digéré par EcoRI.

(C) Le chevauchement des insertions cellulaires contenues dans les clones lambda CM-4 et lambda CM-5 est clairement démontré par l'hybridation du fragment EcoRI 1.5 kbp du clone lambda CM-4 utilisé comme sonde radioactive avec le fragment EcoRI 9.0 kbp du clone lambda CM-5 (canaux 1 et 2).

(D) Organisation des séquences cellulaires de poulet clonées dans les bactériophages lambda CM-4 et lambda CM-5

▭: Fragments d'ADN obtenus après digestion par EcoRI, les chiffres indiquant leur taille exprimée en kilo-base paires (kbp)

↓: sites de coupure artificiels pour l'enzyme de restriction EcoRI dus à l'adjonction de séquences "EcoRI linkers" lors de la construction de la banque de gènes.

▲ ▲: fragment hybridant avec le cDNA myb.

10 à 20 kbp sont sélectionnés, ligaturés à des extrémités cohésives EcoRI ("linkers") et insérés entre deux fragments d'ADN ("bras") contenant l'information génétique nécessaire et suffisante à l'encapsidation, à la réplication et à la propagation du bactériophage vecteur Charon 4A (Dodgson et al., 1979). En tenant compte de la taille des fragments d'ADN utilisés tout le génome du poulet d'une complexité de $5x10^8$ paires de bases est statistiquement contenu dans 200 000 bactériophages. Lors d'une expérience de criblage de la génothèque environ 250 000 bactériophages sont mis en contact avec leurs bactéries-hôtes et amplifiés sur milieu gélosé (60 000 bactériophages/boîte). Après croissance, une partie des bactériophages est transférée sur une feuille de nitrocellulose, orientée grâce à des repères sur la boîte. Après hybridation à l'aide d'une sonde v-myb, lavage, autoradiographie et repérage des points d'hybridation on isole à partir de la boîte de culture des bactériophages recombinants contenant des fragments du gène c-myb parmi un "mélange" de clones non hybridants. Des étapes successives de purification par dilution de ce mélange et hybridation permettent de cloner un bactériophage contenant l'insertion désirée. L'étude physique par enzymes de restriction de cette insertion cellulaire clonée est alors possible après une étape d'amplification et de production du bactériophage en milieu liquide dont le rendement est d'environ 150-200 microgrammes d'ADN de bactériophage par litre de culture.

Le criblage de la génothèque à l'aide de la sonde v-myb a permis l'isolement de deux clones différents portant des insertions cellulaires chevauchantes et contenant les 3 fragments EcoRI cellulaires détectés dans l'ADN de poulet. Le clone lambda CM-4 digéré par EcoRI fournit outre les "bras" du vecteur (19,8

et 10,9 kbp) des fragments de taille 5.4, 2.2, 2.1, 1.5 et 1.1 kbp (Figure 3B, canal 1) visualisés par fluorescence sous lumière UV du Bromure d'Ethidium intercalé dans l'ADN. Le clone lambda CM-5 quant à lui révéle dans les mêmes conditions des bandes de 9.0, 2.2, 1.9, 1.85, 1.8 et 0.8 kbp (Figure 3B, canal 2).

Les fragments Eco RI 5.4, 2.2, 2.1 et 1.1 kbp du lambda CM-4 ainsi que les 9.0, 2.2 et 1.9 kbp du lambda CM-5 hybrident avec la sonde v-myb (Figure 3B, canal 3 et 4). La carte physique ainsi que l'organisation des fragments EcoRI dans chaque insertion cellulaire a été déterminée par digestion à l'aide d'enzymes de restriction dont les sites de coupure dans l'ADN du bactériophage Charon 4A sont connus, séparation électrophorétique en gel d'agarose, transfert sur feuille de nitrocellulose et hybridation avec la sonde v-myb ou avec chaque fragment marqué au dCTP ^{32}P par la technique d'excision-resynthèse ("Nick Translation"). Dans chaque bactériophage, deux fragments -représentant la jonction entre l'insertion cellulaire et l'ADN phagique- ont un site EcoRI artificiel dû à la digestion de l'ADN par AluI et HaeIII et à l'adjonction de "linkers" EcoRI lors de la construction de la génothèque : des sondes radioactives de ces fragments détectent donc dans l'ADN total digéré par EcoRI des équivalents ayant une taille supérieure.

Dans le clone lambda CM-4, ces deux fragments de jonction sont le 5.4 et le 1.5 kbp. Ce fragment de 5.4 kbp était négatif par hybridation avec la sonde v-myb indiquant donc l'existence d'un doublet de deux fragments -l'un positif et l'autre négatif-ayant apparemment la même taille. Ces deux fragments pouvaient aussi être différenciés par une double digestion, soit EcoRI et HindIII soit EcoRI et BamHI (voir carte du locus c-myb). Finalement, l'ordre des fragments dans le clone est le suivant :

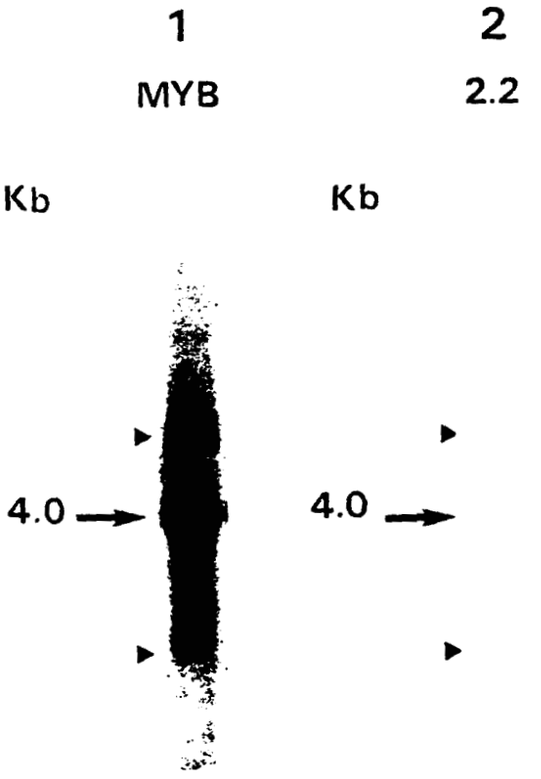
5.4.- 5.4- 1.1- 2.1- 2.2- 1.5 (Figure 3D). L'ordre est donné de 5' vers 3' par référence à la séquence AMV v-myb (voir Discussion) et les fragments positifs avec la sonde v-myb sont soulignés. Le fragment 1.1 kbp de par sa position interne n'est donc pas un fragment tronqué dû à la construction de la banque de gène mais un vrai fragment EcoRI du locus c-myb qui n'avait pas été détecté dans l'ADN cellulaire total (Figure 3A). De plus, ceci démontre que la bande à 2.2 kbp détectée dans l'ADN total à l'aide de la sonde v-myb est en fait constituée d'un doublet de 2 fragments EcoRI de 2.1 et 2.2 kbp qui n'avaient pas été séparés à cause d'un manque de résolution du gel utilisé, dans ces zones de taille (Figure 3A).

L'ordre des fragments dans le clone lambda CM-5 déterminé par la même stratégie est le suivant : 1.9-2.2-9.0-(1.85,1.8,0.8), l'ordre de ces 3 derniers fragments, négatifs avec la sonde v-myb, n'ayant pas été étudié (Figure 3D). La bande EcoRI 1.9 kbp positive avec la sonde v-myb, est un fragment de jonction et représente une version tronquée du 2.1 kbp EcoRI du clone lambda CM-4 (Figure 3B, canal 4 et Figure 3D). De même, le fragment 1.5 kbp du clone lambda CM-4 hybride avec la bande interne EcoRI 9.0 kbp (Figure 3C, canaux 1 et 2). Ces résultats démontrent clairement que les deux clones contiennent des insertions cellulaires chevauchantes couvrant donc un seul locus génétique dans l'ADN de poulet.

De plus, le gène c-myb semble être présent à 1 copie par génome de poulet, puisqu'aucun phage positif avec la sonde v-myb mais ayant une structure différente n'a pu être isolé lors de nombreux criblages de la banque de gène.

FIGURE 4 : Détection de l'ARN cellulaire c-myb.

L'ARN polyadénylé provenant de la lignée cellulaire MSB1 a été dénaturé, soumis à électrophorèse sur gel d'agarose, transféré et hybridé avec le cDNAmyb (canal 1) et le fragment EcoRI 2.2 kbp du clone lambda CM-4 marqué au dCTP ^{32}P (canal 2) . La taille du transcrit c-myb est indiquée en kilobases(kb). Les têtes de flèches indiquent la localisation d'ARNs ribosomaux 28S (5.5 kb) et 18S (2.2 kb) utilisés comme marqueurs de taille.

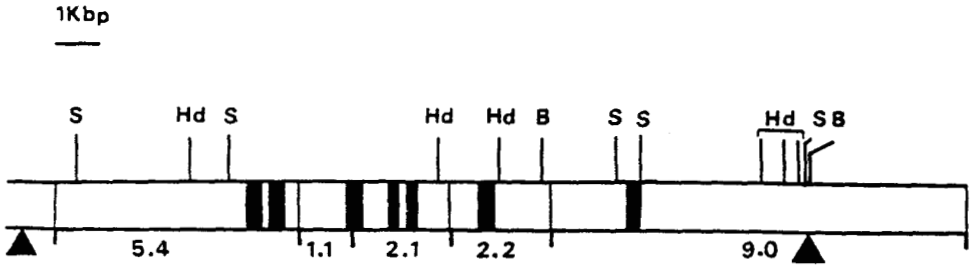


BU LILLE

Néanmoins, cet ADN cloné grâce à son homologie avec l'oncogène viral AMV v-myb ne représentait donc pas obligatoirement la totalité du locus c-myb (poulet) mais uniquement la fraction de ce gène cellulaire transduite dans le virus. La détermination de la taille du transcrit mature (ARN messager polyadenylé) d'un gène cellulaire constituant une bonne approche de sa complexité, nous avons donc étudié une lignée de lymphoblastes T transformés par le virus de Marek où on détecte par hybridation moléculaire en phase liquide de 30 à 50 copies du gène c-myb par cellule (Coll et al., 1983a). La sonde v-myb détecte dans l'ARN polyadénylé extrait de ces cellules séparé par électrophérèse en gel d'agarose et transféré sur nitrocellulose un ARN messager de 4.0 kb qui est également détecté par le fragment EcoRI 2.2 kbp présent dans les clones lambda CM-4 et CM-5 (Figure 4, canaux 1 et 2).

DISCUSSION

L'ADN cellulaire de poulet homologue à l'oncogène viral v-myb du virus de la myéloblastose aviaire (AMV) a été cloné dans deux phages recombinants isolés à partir d'une banque de fragments d'ADN d'érythrocytes de poulet normal. Après digestion de l'ADN de poulet par EcoRI, les séquences hybridant avec la sonde virale v-myb sont localisées dans 5 fragments de taille respective : 5.4-1.1-2.1-2.2-9.0 orientés de 5' vers 3' par référence à AMV v-myb. En effet, une sonde spécifique de l'extrémité 5' de v-myb préparée à partir d'un clone d'ADN



BU LILLE

FIGURE 5 : Carte physique du locus c-myb (poulet)

Les chiffres représentent la taille des fragments obtenus par l'enzyme de restriction EcoRI, exprimée en kilo-base paires (kbp). L'orientation 5' vers 3' est définie par rapport à l'oncogène viral AMV v-myb.

▌: régions d'homologie entre AMV v-myb et son équivalent cellulaire définies par séquencage nucléotidique (Klempnauer et al., 1982).

▲ : frontières putatives de l'unité transcriptionnelle complète du gène c-myb d'après Gonda et al. (1983).

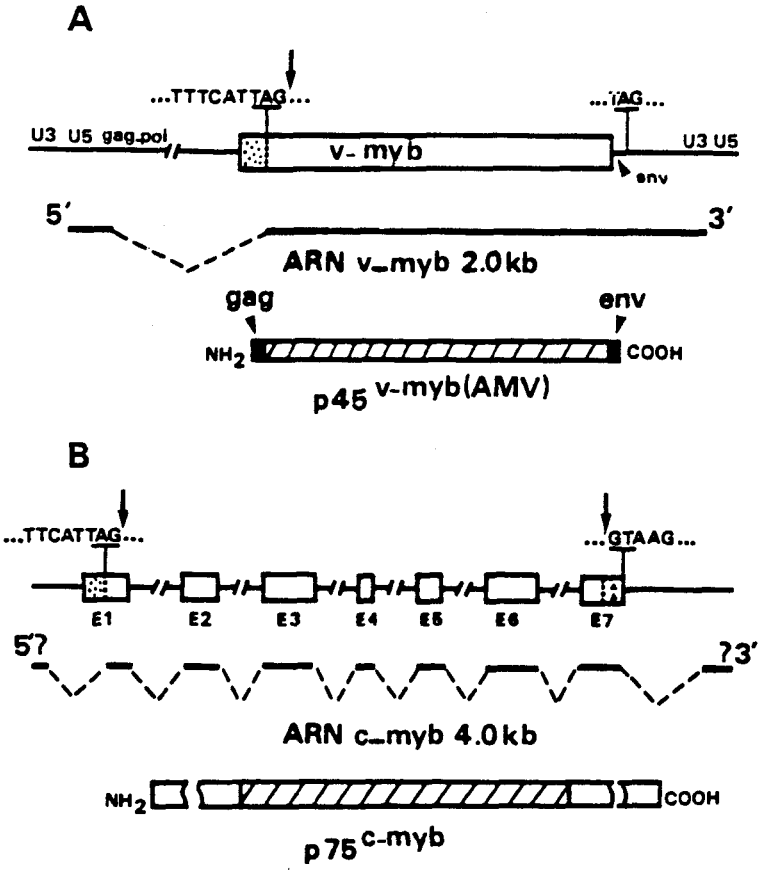
Abréviation des endonucléases de restriction utilisées :

B : BamHI ; Hd ; HindIII ; S : SstI.

proviral hybride avec le fragment EcoRI 5.4 kbp alors qu'une sonde de l'extrémité 3' de v-myb révèle le 9.0 kbp (Perbal et al., 1983 ; Klempnauer et al.,1982) (Figure 5).

Les régions d'ADN cellulaire hybridant avec la sonde v-myb sont séparées par des régions non homologues à v-myb, indiquant que le gène c-myb présente une structure en mosaïque d'"introns" et d'"exons" comme la majorité des gènes eucaryotes. Le virus n'a transduit qu'une faible partie du gène cellulaire, puisque le produit de transcription de c-myb, un ARN messager de 4.0 kb, a une taille bien supérieure à celle de la séquence virale.

Ce travail est venu confirmer les résultats de deux groupes américains qui ont cloné des bactériophages recombinants ayant une structure analogue (Perbal et al., 1983 ; Klempnauer et al., 1982). L'isolement indépendant par trois groupes à partir de banque de gènes différentes , de clones rendant compte du même locus est un argument de poids en faveur de l'existence d'un locus c-myb unique dans l'ADN de poulet. Certains de ces auteurs ont exploré beaucoup plus en détail la structure du gène cellulaire c-myb puisqu'ils ont obtenu la séquence nucléotidique d'AMV v-myb et de son équivalent cellulaire (Klempnauer et al., 1982). AMV v-myb peut être divisé en 7 régions, notées de "E1" à "E7" de l'extrémité 5' vers l'extrémité 3' et dont les équivalents cellulaires sont localisés dans les fragments EcoRI du gène c-myb de la façon suivante (Figures 5 et 6) : "E1" et "E2" dans le fragment EcoRI 5.4 kbp ; "E3" se trouve dans les fragments 1.1 kbp et 2.1 kbp. Le site EcoRI séparant ces 2 fragments est équivalent à celui situé au tiers 5' de v-myb dans l'ADN proviral (Klempnauer et al., 1982 ; Rushlow et al., 1982) (Figure 5 et 6). Le fragment de 1.1 kbp ne contient que 20 nucléotides homologues à AMV v-myb : ceci explique sa


A
...TTTCATTAG...
...TAG...
U3 U5 gag-pol
v-myb
U3 U5
env
5'
3'
ARN v-myb 2.0kb
gag
env
NH2
COOH
p45 v-myb(AMV)
B
...TTCATTAG...
...GTAAG...
E1
E2
E3
E4
E5
E6
E7
5'
3'
ARN c-myb 4.0kb
NH2
COOH
p75 c-myb

BU LILLE

Figure 6 : Modèles de transcription et de traduction de v-myb dans l'ADN proviral d'AMV (A) et du gène cellulaire c-myb (B) (d'après Klempnauer et al., 1983a).

A) La structure de l'ADN proviral d'AMV est représentée schématiquement U3, U5, gag, pol et env représentent les informations virales provenant du virus "helper" (voir Figure 1). La position de deux codons stop de translation (TAG) flanquant le cadre de lecture ouvert de v-myb est indiquée. La séquence contenant le TAG au début de v-myb correspond à un site accepteur d'épissage (↓) utilisé dans la génèse de l'ARN sous-génomique v-myb. Cette région (▪ ▪ ▪) correspond à un fragment d'" intron" du gène c-myb (voir (B)).

Les régions jointes par épissage sont représentées par un trait épais, relié par des pointillés. L'ARN sous génomique est utilisé pour la synthèse de la $p45^{v\text{-}myb}$. Cette protéine est codée par l'oncogène viral (zone hachurée) et par quelques acides-aminés des gènes viraux gag et env (zone noire).

B) Représentation schématique d'une partie du gène c-myb. Les régions "E1" à "E7" représentent les régions d'homologie entre AMV v-myb et c-myb. "E2", "E3", "E4", "E5" et "E6" correspondent à de vrais "exons" du gène cellulaire c-myb. "E1" correspond à un "exon" cellulaire flanqué en 5' par une région intronique (▪ ▪). "E7" représente un "exon" du gène c-myb, dont une partie (▲) n'a pas été transduite dans AMV. Seuls les signaux d'épissage (AG et GT) ainsi que les sites d'épissage (↓) en 5' de "E1" et en 3' de "E7" ont été représentés.

Ces 7 régions après épissage correct constituent une fraction de l'ARN 4.0 kb du gène c-myb qui doit donc comporter des séquences non encore localisées en 5' et/ou en 3'. La protéine $p75^{c\text{-}myb}$ est

donc constituée d'un domaine commun avec la p45$^{v\text{-}myb}$ (zone hachurée) et de domaines (30 kilodaltons) propres à l'ARN cellulaire (rectangle ouvert).

non-détection dans l'ADN cellulaire ainsi que la faible hybridation de ce fragment purifié dans le clone lambda CM-4. Finalement, les régions "E6" et "E7" sont localisées respectivement dans les fragments de 2.2 et 9.0 kbp.

Le cadre de lecture ouvert utilisé pour la synthèse de la protéine $p45^{v\text{-}myb}$ est le même que celui du gène cellulaire. De plus, les régions "E2", "E3", "E4", "E5" et "E6" représentent des vrais "exons" du gène cellulaire c-myb raboutés par épissure correcte des séquences introniques. Immédiatement en 5' et 3' de ces séquences codantes colinéaires avec l'oncogène viral se trouvent les signaux canoniques d'épissure dans la cellule eucaryote et notamment les dinucléotides AG en 5' et GT en 3' (Mount, 1982) (Figure 6).

En revanche, le cadre de lecture ouvert du gène c-myb se prolonge au-delà de l'extrémité 3' de la région virale " E7". Puisque cette région est flanquée en 5' de signaux d'épissure, elle constitue donc un "exon" du gène cellulaire (tronqué dans sa partie 3' terminale) (Klempnauer et al., 1983b) . Dans le virus, cette séquence "E7" est jointe aux 11 derniers codons du gène env, qui fournit donc le signal (TAG) d'arrêt de la traduction de la $p45^{v\text{-}myb}$. De même la séquence "E1" est un "exon" complet flanqué en 5' d'une région intronique cellulaire. Cet "intron" a deux rôles à jouer dans le virus AMV. Il fournit le site accepteur d'épissure qui avec le site donneur situé quelques nucléotides en aval du gène gag permet la formation de l'ARN sous génomique v-myb de 2.1 kb. Grâce à cette épissure, le codon initiateur (AUG) et six codons du gène gag sont joints dans la phase ouverte de lecture de AMV v-myb et contribuent à la synthèse de la $p45^{v\text{-}myb}$ (Klempnauer et al., 1983a ; Schwartz et al., 1983).

Cet "intron" cellulaire est également joint en phase avec les gènes gag et pol d'AMV et fournit quelques amino-acides et le terminateur de traduction du $Pr180^{gag-pol}$. La partie carboxy-terminale de cette poly-protéine virale est donc dérivée d'un "intron" du gène c-myb : ceci pourrait expliquer que le précurseur $Pr180^{gag-pol}$ ne puisse être processé en protéines matures (Duesberg et al., 1980). Un mécanisme de recombinaison en deux étapes entre le gène c-myb et un virus sans oncogène pour donner naissance à AMV a également été proposé (Klempnauer et al., 1982 ; 1983b). Dans une première étape, le virus se serait intégré avec perte de sa partie 3' (env-LTR3') au niveau de l'"intron" "E1" , un ARN hybride contenant de 5' en 3 ' : U5, gag, pol, la partie intronique de "E1", et le reste du gène c-myb épissé correctement aurait alors été produit. Cet ARN contenant les séquences 5' et le gène gag pouvait donc être co-empaqueté avec un ARN de structure classique U5, gag, pol, env, U3 dans une particule virale. La deuxième étape soit par recombinaison entre ces deux ARNs ou par saut de la rétropolymérase d'une molécule à l'autre lors du processus de rétrotranscription aurait remplacé la partie du gène c-myb située en 3' de "E7" par les séquences env et U3 donnant ainsi naissance au génome d'AMV.

D'autre part, les différences existant entre la protéine transformante de l'oncogène viral AMV v-myb et la protéine du gène cellulaire c-myb ayant un rôle à jouer dans le développement de la cellule normale ont pu être définies.

Le produit du gène c-myb cellulaire est une protéine de 75 kilodaltons (kd) , $p75^{c-myb}$ (Klempnauer et al., 1983a ; Boyle et al., 1983) plus grande que la protéine virale, $p45^{v-myb}$ (Figure 6) mais également à localisation nucléaire (Klempnauer et al., 1984b). Cette protéine cellulaire nécessiterait donc une capacité

codante en ARN d'environ 2,25 kb ce qui laisse environ 1.75 kb de séquences non traduites en 5' et/ou en 3' de l'ARN messager 4,0 kb du gène c-myb.

Toute l'information de cet ARN c-myb a d'ailleurs pu être grossièrement localisée au voisinage immédiat de la région d'ADN cellulaire transduite dans le virus. Une sonde cADN radioactive synthétisée par rétrotranscription d'une fraction d'ARN messagers cellulaires sélectionnée en taille à 4.0 kb -et donc enrichie en ARN c-myb-hybride uniquement avec des fragments contenus dans les clones lambda CM-4 et lambda CM-5 (Gonda et al., 1983).Le locus complet c-myb pourrait donc avoir une taille n'excédant pas 30 kbp d'ADN (Gonda et al., 1983 ; Figure 5).

En conclusion, trois hypothèses non mutuellement exclusives peuvent rendre compte du pouvoir oncogène de v-myb :

* v-myb est une version tronquée du gène cellulaire c-myb ;

* la protéine $p45^{v\text{-}myb}$ est associée à son extrémité NH2 terminale à 6 Amino-Acides du gène viral gag et à son extrémité Carboxy-terminale à 11 Acides Aminés du gène env. De plus, 21 différences d'acides aminés, existent entre sa séquence et celle de l'homologue cellulaire. Ces différences de structure pourraient pertuber la redistribution de la protéine myb au cours de la différenciation myéloïde normale (Klempnauer et al., 1984b) ;

* la protéine v-myb est produite en grande quantité dans les cellules cibles infectées grâce aux signaux de transcription fournis par le LTR viral.

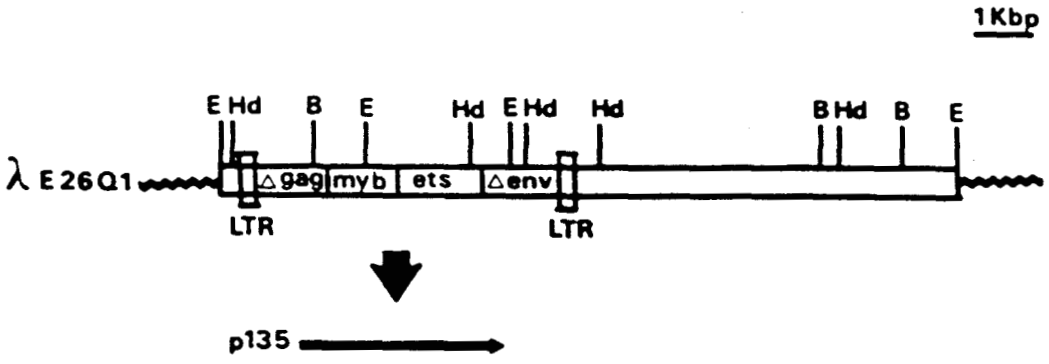
2ème partie : TRANSDUCTION DU PROTO-ONCOGENE c-myb (POULET) DANS UN AUTRE RETROVIRUS AVIAIRE : E26.

L'ADN proviral du virus E26 a été cloné à partir d'une banque de fragments d'ADN de fibroblastes de caille, E26QC1, transformés par ce virus . Un bactériophage recombinant hybridant à la fois avec la sonde v-myb et des sondes représentatives des gènes viraux de structure a été isolé. La carte physique et génétique de ce clone, lamba E26-Q1 obtenue après digestion de son ADN par différentes enzymes de restriction et hybridation avec les sondes U5, gag, pol, v-myb, env et U3 démontre qu'il contient un provirus entier. De plus, sa structure LTR-gag-myb-ets-env-LTR est en accord avec celle suggérée précédemment par l'étude de l'ARN viral en hybridation liquide (Roussel et al., 1979 ; Stéhélin et al., 1980 ; Bister et al., 1982) (Figure 7). Les séquences apparentées à AMV v-myb, présentes dans E26, ont pu être localisées par hybridation dans 2 fragments obtenus après digestion par les enzymes de restriction BamHI et EcoRI :

* un fragment BamHI-EcoRI de 1.0 kbp contient la fin du gène gag tronqué et la partie 5' de E26 v-myb.

* un fragment EcoRI de 2.5 kbp contient respectivement de l'extrémité 5' vers 3' (définies par rapport au génome viral) la partie centrale et 3' terminale de E26 v-myb, la séquence spécifique v-ets et le début du gène env tronqué. Ces deux fragments ont été reclonés dans des plasmides et étudiés plus en détail.

La détermination de la séquence nucléotidique de E26 v-myb et sa comparaison avec celle d'AMV v-myb et du gène cellulaire c-myb constitueraient l'approche la plus fine de l'étude des



BU LILLE

Figure 7 : Carte physique et organisation génétique du provirus E26 cloné dans le bactériophage recombinant lambda E26Q1.

Le rectangle ouvert représente l'insertion cellulaire et les lignes brisées les "bras" du vecteur phagique. Les LTRs sont symbolisées par des rectangles ouverts alors que les jonctions entre chaque domaine (gag et myb ; myb et ets ; ets et env) dont les séquences nucléotidiques sont connues sont représentées par des barres verticales.

La flèche épaisse représente l'information virale codant pour la protéine transformante putative de E26, $p135^{gag\text{-}myb\text{-}ets}$.

Enzymes de restriction : B : BamHI ; E : EcoRI ; Hd : HindIII.

propriétés de cette séquence et des modalités de sa transduction. Néanmoins, deux stratégies peuvent permettre l'obtention beaucoup plus rapide des principales caractéristiques de E26 v-myb :

* l'établissement de la carte physique de E26 v-myb à l'aide d'enzymes de restriction dont les sites de coupure dans AMV v-myb sont connus (Klempnauer et al., 1982 ; Rushlow et al., 1982).

* la comparaison des fragments du gène c-myb (poulet) cloné, qui hybrident d'une part avec la sonde AMV v-myb et d'autre part avec chaque fragment de E26 v-myb utilisé comme sonde.

La digestion par les enzymes de restriction AvaII et EcoRI du fragment BamHI-EcoRI de E26 v-myb fournit un fragment AvaII-AvaII de 55 paires de bases et un fragment EcoRI-AvaII de 225 paires de bases trouvés également dans AMV v-myb. De même à l'extrémité 5' du fragment EcoRI de 2.5 kbp, une séquence de 350 paires de bases définie par les enzymes EcoRI et SalI semble colinéaire à la partie centrale d'AMV v-myb puisqu'elle contient la même succession de 5 sites de coupures d'enzymes de restriction (EcoRI, PstI, SmaI, PstI et SalI) (Figure 8A). Les parties 3' terminales de AMV et E26 v-myb n'ont pu être comparées de cette manière puisqu'aucun site de restriction simple n'était disponible.

Ces données ont été confirmées et étendues par l'étude des clones recombinants lambda CM-4 et CM-5 à l'aide de sondes spécifiques de E26 v-myb . La sonde E26 v-myb BamHI-EcoRI de 1,0 kbp détecte dans le clone lambda CM-4 digéré par EcoRI les fragments c-myb 5.4 kbp et très faiblement le 1.1 kbp -ce fragment ne contenant que 20 nucléotides d'homologie avec AMV v-myb n'a pu être visualisé qu'après une autoragiographie très longue- et aucun fragment dans le lambda CM-5 (Figure 8B, canaux 1 et 2). Ceci démontre que le site EcoRI présent dans E26 v-myb

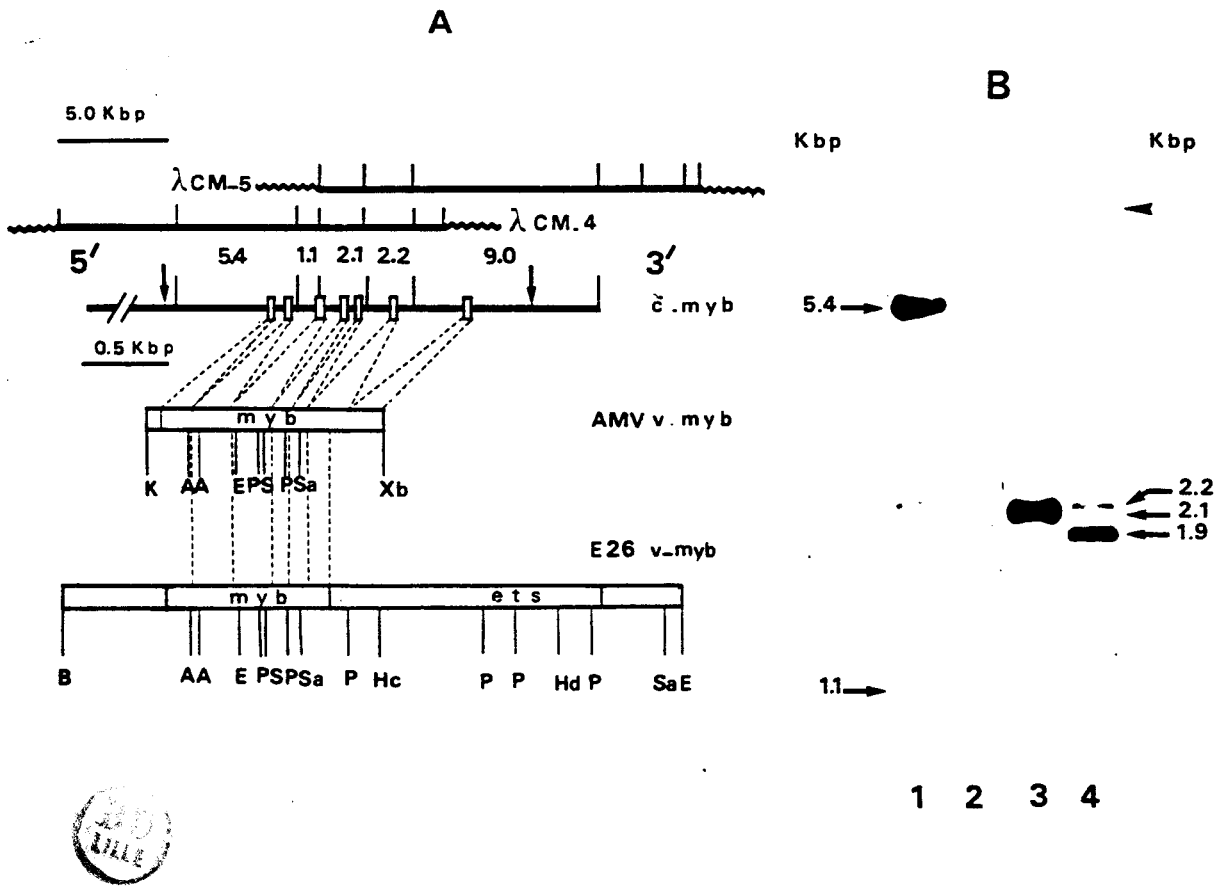

A
B
5.0 Kbp
λCM-5
λ CM.4
5′
5.4
1.1
2.1
2.2
9.0
3′
č.myb
0.5 Kbp
myb
AMV v.myb
K
AA
EPS
PSa
Xb
E26 v-myb
myb
ets
B
AA
E
PSPSa
P
Hc
P
P
Hd P
SaE
Kbp
Kbp
5.4
2.2
2.1
1.9
1.1
1
2
3
4

Figure 8 : E26 v-myb représente une forme tronquée de l'oncogène viral AMV v-myb.

(A) Organisation du locus c-myb (poulet) et comparaison des cartes physiques de AMV v-myb et E26 v-myb. De bas en haut : pour les phages recombinants lambda CM-4 et lambda CM-5 couvrant le locus c-myb complet, les barres verticales représentent dans l'insertion cellulaire de poulet (ligne épaisse) les fragments -dont la taille en kbp est indiquée en dessous- générés par l'enzyme de restriction EcoRI.

La troisième ligne représente l'unité transcriptionnelle c-myb complète dont les limites définies par Gonda et al. (1983) sont symbolisées par les 2 flèches. L'orientation 5' et 3' est donnée par référence à l'oncogène viral v-myb dans le virus AMV. De même, les rectangles ouverts représentent les 7 régions d'homologie définies entre AMV v-myb et son homologue cellulaire (Klempnauer et al., 1982).

Ces 7 régions sont fusionnées dans le virus AMV pour former l'oncogène v-myb inclus dans un fragment KpnI - XbaI de 1.1 kbp flanqué en 5' de 104 nucléotides du gènes pol (1) et en 3' de 33 nucléotides du gènes env (non représentés).

Enzymes de restriction utilisés : A : AvaII ; E : EcoRI ; P : PstI ; K : KpnI ; Sa : SalI ; S : SmaI et Xb : XbaI.

La carte physique des deux fragments du provirus E26 positifs avec la sonde v-myb : BamHI - EcoRI 1.0 kbp et EcoRI-EcoRI 2.5 kbp (voir figure 7) a été établie avec les endonucléases de restriction suivantes : A : AvaII pour le premier fragment et Hc : HincII ; Hd : HindIII ; P : PstI ; Sa : SalI et S : SmaI pour le second fragment. Autres abréviations utilisées : B: BamHI et E : EcoRI.

Les barres verticales représentent les jonctions définies précisemment grâce à la séquence nucléotidique entre les domaines gag et E26 v-myb d'une part et v-ets et env d'autre part.

(B) Hybridation de E26 v-myb avec les phages recombinants lambda CM-4 et lambda CM-5.

Les deux phages recombinants portant le locus c-myb poulet (voir figure 8A) lambda CM-4 (canaux 1 et 3) et lambda CM-5 (canaux 2 et 4) ont été digérés par EcoRI séparé par électrophorèse en gel d'agarose, transféré sur nitrocellulose et hybridé avec les 2 fragments E26 v-myb marqué au dCTP ^{32}P par excision-resynthèse : BamHI-EcoRI 1.0 kbp (canaux 1 et 2) et EcoRI-EcoRI 2.5 kbp (canaux 3 et 4).

La tête de flèche (canal 4) indique la position du fragment EcoRI 9.0 kbp du locus c-myb poulet contenant la région d'homologie "E7" équivalente à l'extrémité 3' terminale d'AMV v-myb.

est équivalent à celui situé au tiers 5' de AMV v-myb ou dans l'ADN cellulaire à la jonction entre les fragments EcoRI 1.1 et 2.1 kbp du gène c-myb. D'autre part, une digestion par PstI du fragment 5.4 kbp EcoRI cloné dans un plasmide permet d'individualiser les "exons" "E1" et "E2" dans deux fragments hybridant tout deux avec la sonde E26 v-myb ; tout "E2" et tout ou partie de "E1" ont donc été transduits dans E26. Néanmoins, il est raisonnable de penser que la séquence intronique de "E1" -fournissant le site accepteur d'épissure utilisé dans la génèse du messager 2.1 kb v-myb- a été tronquée puisqu'aucun ARN viral sous-génomique n'est détecté dans les cellules infectées par E26 (Klempnauer et al., 1984). Il reste à savoir si cette amputation touche également la partie codante et exonique de "E1". La séquence nucléotidique de E26 v-myb permettra de répondre à cette question (voir Discussion).

Le fragment EcoRI 2.5 kbp de E26 v-myb hybride quant à lui, avec le doublet 2.1-2.2 kbp dans le clone lambda CM-4 (Figure 8B, canal 3). Cette hybridation est imputable à chaque fragment du doublet, puisque dans le clone lambda CM-5 la même sonde révèle les fragments EcoRI 1.9 kbp et 2.2 kbp équivalents respectivement aux fragments 2.1-2.2 du clone lambda CM-4 (Figure 8A et 8B, canal 4). Le fragment EcoRI 9.0 kbp contenant la région d'homologie "E7" définie entre AMV v-myb et c-myb (poulet) n'est pas détecté dans le clone lambda CM-5 par la sonde E26 v-myb 2.5 kbp EcoRI (Figure 8B, canal 4). Cette région n'ayant pas été transduite dans E26, E26 v-myb est donc une forme tronquée au moins dans sa partie 3' terminale d'AMV v-myb. De plus, la très faible hybridation entre la sonde E26 v-myb et le fragment 2.2 kbp, bien individualisé dans le clone lambda CM-5, suggère que cette information absente dans E26 v-myb englobe non seulement

"E7" mais également une partie de la séquence "E6" (Figure 8B, canal 4). En effet dans les mêmes conditions, l'hybridation obtenue entre le fragment EcoRI 2.2 kbp et la sonde AMV v-myb, contenant tout "E6", est beaucoup plus intense (Figure 3B, canal 4).

DISCUSSION : Grâce au clonage moléculaire d'un provirus E26, la partie du gène cellulaire c-myb (poulet) transduite par ce virus a pu être en partie définie et comparée à celle transduite par AMV. Les cartes physiques de E26 et AMV v-myb sont fortement similaires dans la partie 5' et centrale de la séquence transformante, ce qui est en bon accord avec les 60 % d'hybridation obtenus entre l'ARN de E26 et un cDNA myb par hybridation liquide (Roussel et al., 1979). La différence observée entre AMV v-myb et E26 v-myb a pu être définie par l'étude des clones cellulaires c-myb(poulet) à l'aide de sondes radioactives des fragments E26 v-myb : la portion du gène cellulaire c-myb homologue à l'extrémité 3' distale de AMV v-myb, tout "E7" et une partie de "E6", n'a pas été transduite par E26. L'étendue de cette amputation a été précisément définie par l'obtention de la séquence nucléotidique de E26 v-myb (Nunn et al., 1983 ; B. Debuire, communication personnelle). Les 128 nucléotides de l'extrémité 3' terminale de l'"exon" "E6" ainsi que tout "E7" (120 nucléotides) sont absents de E26 v-myb. Ceci représente l'information génétique permettant la synthèse du quart carboxy-terminal de la $p45^{v\text{-}myb}$. De même, la séquence nucléotidique de la partie 5' terminale de E26 v-myb a confirmé l'absence de séquences appartenant à "E1" et notamment du signal d'épissure. Ce signal devait être obligatoirement éliminé

puisqu'aucun ARN viral sous-génomique n'a été détecté dans les cellules transformées par E26 (Klempnauer et al., 1984). Cette amputation a également touché les 10 premiers codons de la partie codante de "E1". Le domaine v-myb dans la protéine transformante $p135^{gag\text{-}myb\text{-}ets}$ de E26 est donc une forme tronquée dans sa partie Amino et Carboxy-terminale de la $p45^{v\text{-}myb}$.

D'autre part, quelques mutations ponctuelles existent entre les régions communes d'AMV v-myb et E26 v-myb bien que le cadre de lecture utilisé dans ces 2 virus soit identique à celui du gène cellulaire c-myb (Boyle et al., 1983 ; Klempnauer et al., 1983a). Ces différences reflètent des erreurs introduites par la rétropolymérase après la transduction et la réplication indépendantes de la même séquence cellulaire dans deux isolats de rétrovirus.

Le mécanisme de recombinaison qui a donné naissance au virus E26 en joignant en phase avec le gène viral gag, deux séquences d'origine cellulaires provenant de deux loci distincts c-myb et c-ets, reste inconnu pour l'instant. On peut imaginer un phénomène en deux étapes par recombinaison successive avec les deux gènes et disparition d'un virus intermédiaire "fossile" qui ne possédait que l'une ou l'autre des séquences cellulaires. L'intervention du virus AMV comme intermédiaire dans la formation de E26 peut cependant être exclue. Il est difficilement convevable de créer un virus de structure gag-myb-ets-env à partir d'un virus de structure gag-pol-myb. De plus AMV et son helper naturel MAV (Myeloblastosis-Associated Virus) possèdent une région U3 d'un type particulier (Klempnauer et al., 1982) qui n'est pas retrouvée dans E26.Alternativement, une unique étape de

recombinaison dans une cellule où les deux gènes cellulaires auraient été fusionnés par un remaniement chromosomique aurait pu donner naissance à E26.

Le rôle respectif de chaque séquence cellulaire dans le spectre oncogène induit par E26 -clairement distinct ce celui d'AMV -est également obscur. Deux hypothèses peuvent être proposées :

* chaque domaine pourrait être responsable de la transformation de cellules appartenant à une lignée hématopoïétique : ets dans le cas des cellules érythroïdes, myb dans le cas des cellules myéloïdes (E26 v-myb pourrait alors représenter la partie minimale du gène c-myb responsable de la transformation myéloïde).

* les propriétés transformantes de myb pourrait en être augmentées grâce à l'action de ets. E26 tout comme AEV (Avian Erythroblastosis Virus) ou erbA est un potentiateur de l'oncogène viral erbB constituerait un exemple "in vivo" de complémentation entre oncogènes, définie "in vitro" par le test de transformation des NIH 3T3 (Frykberg et al., 1983 ;Land et al., 1983b) pour d'autres oncogènes.

3ème Partie : CLONAGE DU PROTO-ONCOGENE c-myb DANS L'ADN HUMAIN.

Des fragments d'ADN total de placenta humain obtenus par l'enzyme de restriction EcoRI ont été séparés par électrophérèse sur gel d'agarose, transférés sur nitrocellulose selon la technique de Southern et hybridés avec la sonde v-myb en conditions relachées. Ces conditions -augmentation de la concentration en NaCl dans le tampon d'hybridation et de lavage

et diminution de la température de lavage- sont nécessaires pour stabiliser les hybrides obtenus entre la séquence humaine et virale qui présentent une divergence phylogénétique de quelques 300 millions d'années. Dans ces conditions, la sonde AMV v-myb détecte dans l'ADN humain deux fragments EcoRI de taille respective, 2.8 et 2.0 kbp (Figure 9A, canal 1).

Le gène c-myb humain a été cloné à partir d'une banque de fragments d'ADN de foie foetal obtenus par digestion partielle avec les enzymes de restriction AluI et HaeIII et insertion dans les "bras" du vecteur phagique Charon 4A. Lors de l'expérience de criblage de la banque de gènes (300 000 bactériophages pour 4 boîtes), en conditions non-strictes, trois candidats ont été obtenus. Néanmoins, les étapes de purification ultérieures ont démontré que deux de ces candidats représentaient des artéfacts dus à une hybridation non spécifique dans les conditions de criblage utilisées. Seul un clone, lambda HM-1, hybridant de façon significative et reproductible avec la sonde virale a été étudié plus en détail. L'ADN de ce clone, après digestion par EcoRI révèle, outre les "bras", du vecteur des fragments de taille respective 5,0 ; 2,8 ; 2,0 ; 1,2 et 0,9 kbp (Figure 9B, canal 1). La carte physique de cette insertion cellulaire a été construite grâce à une stratégie analogue à celle décrite pour les clones c-myb (poulet). L'ordre des fragments ainsi déterminé est le suivant : 0,9 ; 2,8 ; 1,2 ; 2,6 ; 2,0 ; 5,0 (Figure9D) ; les fragments 5,0 et 0,9 kbp représentant les fragments de jonction entre l'ADN humain et le vecteur. La sonde AMV v-myb hybride avec 5 fragments EcoRI du clone lambda HM-1 : 5,0 ; 2,8 ; 2,6 ; 2,0 et 1,2 kbp (Figure 9B, canal 2). La bande de 2.8 kbp visualisée dans l'ADN total est donc constituée d'un doublet de deux fragments -2,8 et 2,6 kbp-non séparés sur le gel (Figure 9A,

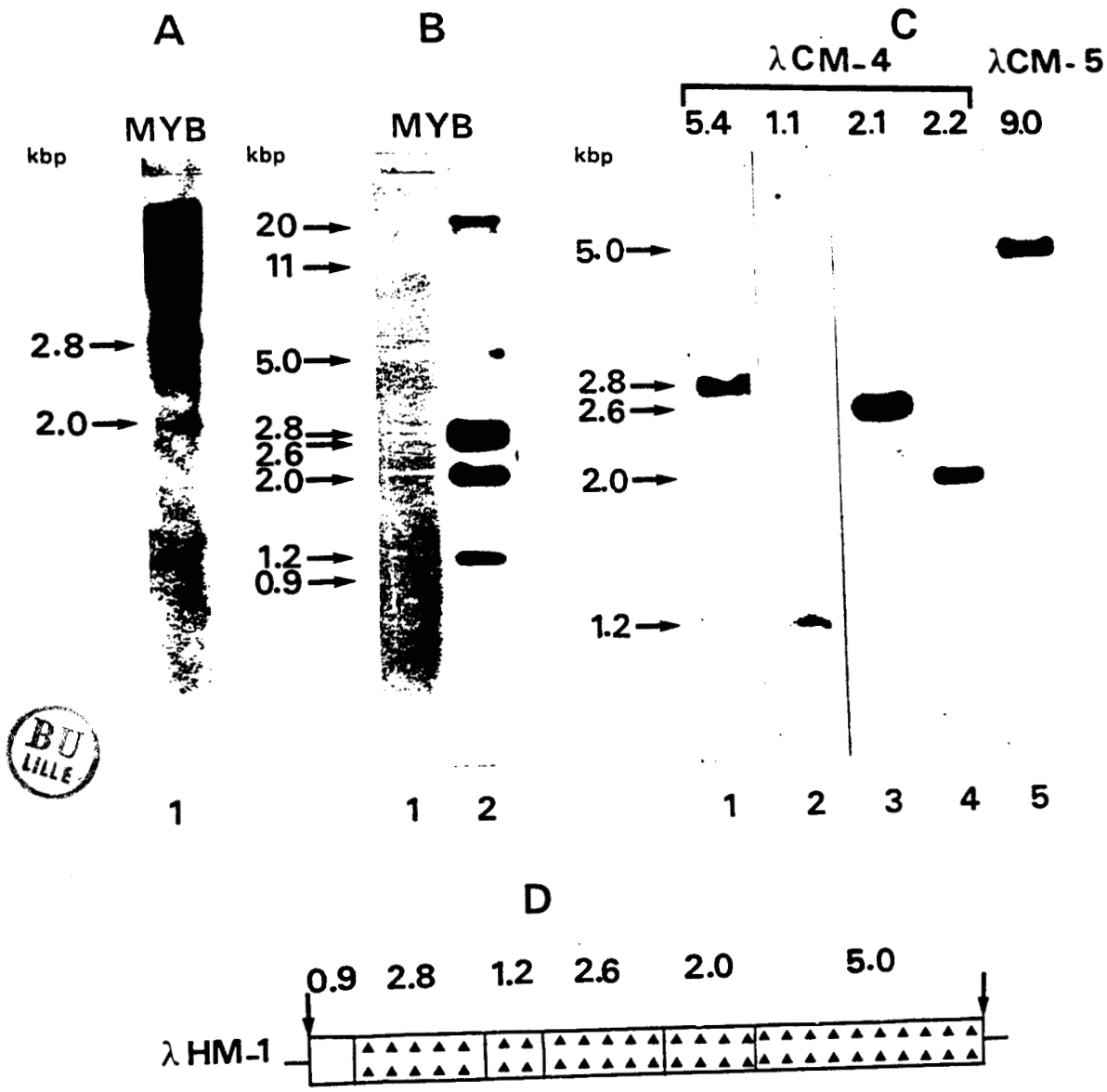

A
MYB
kbp
2.8
2.0
1
B
MYB
kbp
20
11
5.0
2.8
2.6
2.0
1.2
0.9
1
2
C
λCM-4
λCM-5
5.4
1.1
2.1
2.2
9.0
kbp
5.0
2.8
2.6
2.0
1.2
1
2
3
4
5
D
0.9
2.8
1.2
2.6
2.0
5.0
λHM-1

Figure 9 : Le locus c-myb humain

(A) Des fragments d'ADN de placenta humain digéré par l'enzyme de restriction EcoRI ont été séparés par électrophorèse en gel d'agarose, transférés sur nitrocellulose et hybridé avec le cDNA myb.

(B) L'ADN du clone lambda HM-1 a été traité comme décrit ci-dessus : visualisation des fragments après coloration au bromure d'éthidium (canal 1) et hybridation avec le cDNA myb (canal 2)

(C) Des aliquots de l'ADN du clone lambda HM-1 traité comme en (B) ont été hybridés avec les fragments EcoRI du clone lambda CM-4 marqués au dCTP ^{32}P : 5.4 kbp (canal 1) ; 1.1 kbp (canal 2) ; 2.1 kbp (canal 3) et 2.2 kbp (canal 4). De même, le canal 5 a été hybridé avec le 9.0 kbp du clone lambda CM-5

(D) Organisation du locus c-myb humain dans le phage recombinant lambda HM-1.

▭: taille en kbp des fragments d'ADN obtenus après digestion par EcoRI

↓: sites EcoRI artificiels (voir légende de la Figure 3)

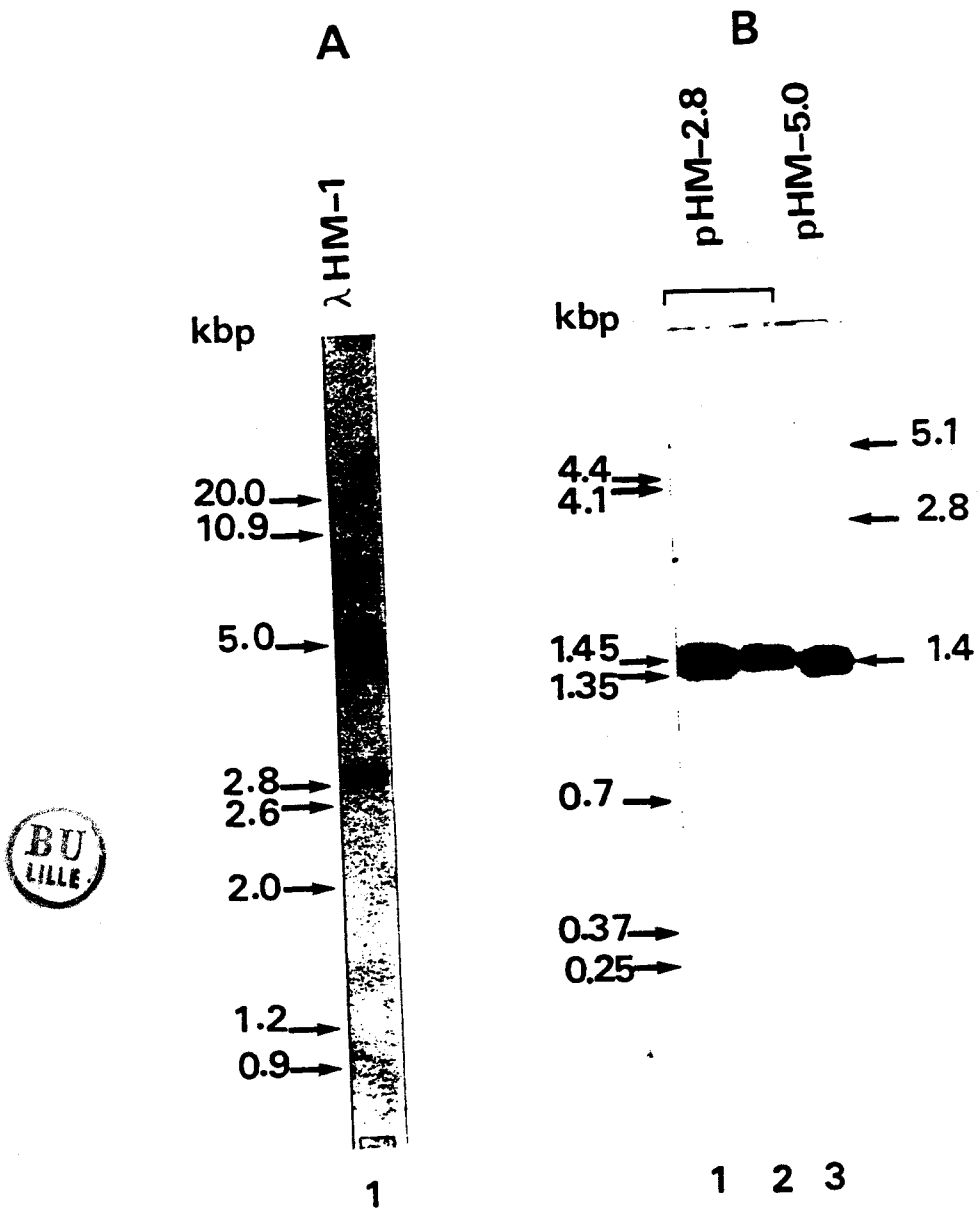
▲ ▲: fragments hybridant avec le cDNA myb

canal 1). De même, la sensibilité accrue de ce test, utilisant un ADN hautement purifié permet de détecter deux autres fragments du gène c-myb humain hybridant avec la sonde AMV v-myb : un fragment EcoRI de 1,2 kbp et un fragment EcoRI "artificiel" de 5.0 kbp. Ces fragments pourraient donc contenir des régions d'homologie avec la séquence virale de petite taille et /ou ayant fortement divergées, ce qui expliquerait leur non-détection dans l'ADN total.

Chaque fragment EcoRI du gène c-myb humain a été cloné dans des plasmides et hybridé avec la sonde AMV v-myb après digestion par diverses enzymes de restriction. Ces résultats ont montré que des séquences homologues à AMV v-myb sont séparées par des régions non hybridantes, suggèrant une organisation en mosaïque pour le gène c-myb humain (Figure 11).

Finalement, l'homologie entre les loci cellulaires aviaire et humain a été étudiée par hybridation entre les fragments du gène c-myb (poulet) utilisés comme sonde et une digestion EcoRI du clone lambda HM-1. Les deux loci apparaissent colinéaires -au moins dans la zone équivalente à l'oncogène viral v-myb- puisque chaque fragment du gène c-myb poulet hybride avec un seul fragment EcoRI situé en position analogue dans le gène humain (Figures 9C, 9D et 11).

Il restait néanmoins à démontrer que le clone lambda HM-1 contient toutes les séquences homologues à la partie 3' d'AMV v-myb dans le fragment EcoRI "artificiel" -et donc tronqué- de 5.0 kbp. Ce fragment hybride à la fois avec la sonde virale v-myb et avec le fragment EcoRI 9.0 kbp du gène c-myb (poulet) contenant la partie 3' distale de AMV v-myb. De plus, le fragment SstI 2.7 kbp -obtenu à partir du 9.0 kbp EcoRI du lambda CM-5 et contenant les séquences cellulaires du gène c-myb (poulet)


A
λHM-1
kbp
20.0
10.9
5.0
2.8
2.6
2.0
1.2
0.9
1
B
pHM-2.8
pHM-5.0
kbp
4.4
4.1
1.45
1.35
0.7
0.37
0.25
5.1
2.8
1.4
1 2 3

BU LILLE

Figure 10 : Localisation de séquences humaines répétitives dans le clone lambda HM-1

(A) Hybridation entre l'ADN du phage HM-1 digéré par EcoRI et la sonde répétitive humaine

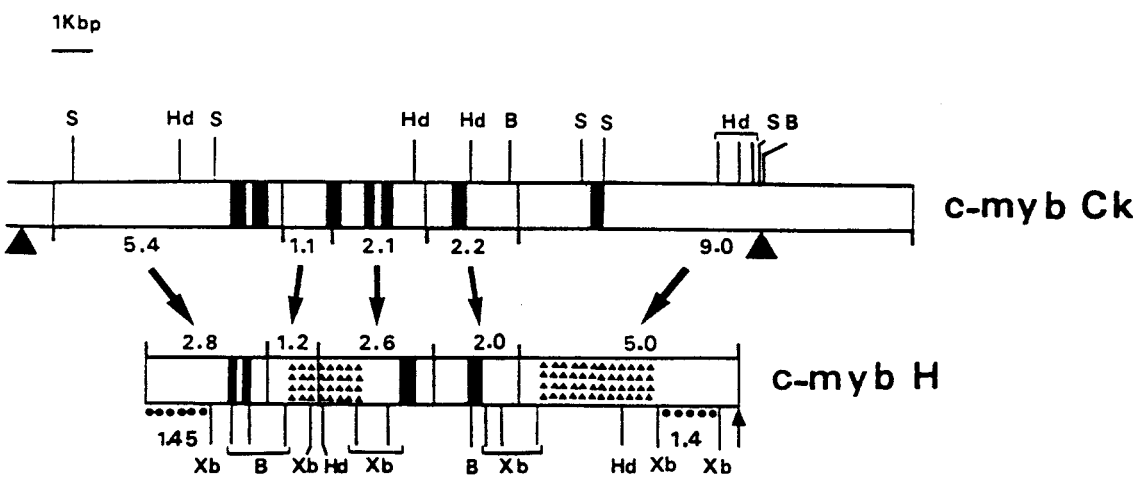
(B) Les fragments EcoRI 5.0 et 2.8 kbp, contenant les séquences répétitives ont été clonées dans les plasmides pHM-5.0 et pHM-2.8 respectivement. L'ADN de pHM-2.8 a été digéré soit par EcoRI et XbaI (canal 1) ou par EcoRI, XbaI et BamHI, soumis à une électrophorèse en gel d'agarose, transféré sur nitrocellulose et hybridé avec la sonde répétitive humaine. La même expérience a été réalisée avec une digestion par XbaI du clone pHM-5.0 (canal 3).

Les fragments BamHI-XbaI 0.7 kbp du clone pHM-2.8 et XbaI-XbaI 2.8 kbp du clone pHM-5.0 positifs avec le cDNA myb (voir Figure 11) ne contiennent pas de séquences répétées.

situées en aval (3') de la région homologue virale "E7"- hybride également avec le fragment EcoRI 5.0 kbp du clone lambda HM-1 au niveau du fragment XbaI 1.45 kbp (Voir Figure 11). Toutes ces données démontrent donc clairement que le clone lambda HM-1 contient toutes les séquences humaines homologues à l'oncogène aviaire AMV v-myb.

Afin de définir des sondes spécifiques du gène c-myb humain utilisables dans des expériences d'hybridation en phase solide avec des ADNs ou ARNs humains extraits de différentes tumeurs, il fallait au préalable localiser les séquences répétitives humaines dans le locus c-myb. La présence de telles séquences -communes et dispersées dans tout l'ADN humain- dans un fragment utilisé comme sonde permettrait la détection de nombreux fragments d'ADN non apparentés au gène c-myb et rendrait les résultats obtenus ininterprétables. Une sonde préparée par excision resynthèse d'ADN humain se réassociant rapidement (Cot inférieur à 400 m x l x s^{-1}) et donc présent dans le génome humain à plusieurs copies -rappelons à ce propos que le Cot 1/2 d'un gène humain copie unique est de 3000 m x l x s^{-1}- détecte de telles séquences répétées dans les fragments EcoRI 5.0 et 2.8 kbp du clone lambda HM-1 (Figure 10A, canal 1). La même expérience réalisée avec ces fragments reclonés dans des plasmides et digérés par les enzymes de restriction EcoRI, BamHI et XbaI dans le cas du fragment 2.8 kbp et XbaI dans le cas du fragment 5.0 kbp indique que ces séquences répétées ne sont pas localisées dans des régions hybridants avec la sonde AMV v-myb (Figure 10B, canaux 1 à 3 et Figure 11).

Le fragment BamHI-XbaI 0.7 kbp préparé à partir du fragment EcoRI 2.8 kbp et dépourvu de séquences répétitives a été choisi comme sonde spécifique du gène c-myb humain. En effet, il


1Kbp
S
Hd S
Hd
Hd B
S S
Hd SB
c-myb Ck
5.4
1.1
2.1
2.2
9.0
2.8
1.2
2.6
2.0
5.0
c-myb H
1.45
1.4
Xb B Xb Hd Xb
B Xb
Hd Xb Xb

BU LILLE

Figure 11 : Comparaison des loci c-myb (poulet) et c-myb (homme). Les cartes physiques détaillées des loci c-myb (poulet) et c-myb (homme) ont été construites à l'aide des enzymes de restriction suivantes :

B : BamHI ; Hd : HindIII ; S : SstI et Xb : XbaI (dans le cas du locus c-myb (homme) uniquement). L'orientation 5' vers 3' est donnée par rapport à l'oncogène viral AMV v-myb.

Autres symboles utilisés :

▭ : taille en kbp des fragments d'ADN obtenu après digestion par EcoRI.

▲ : limites putatives de l'unité transcriptionnelle c-myb (poulet) (Gonda et al., 1983).

▌ : régions d'homologie entre AMV v-myb et c-myb (poulet) Klempnauer et al., 1982).

Dans le cas du locus c-myb (homme) :

▌ : régions d'homologie entre le locus cellulaire et AMV v-myb définies à la fois par hybridation (ce mémoire) et par des études de microscopie électronique (Franchini et al., 1983).

▲ ▲ : régions d'homologie avecc AMV v-myb détectées uniquement par hybridation (ce mémoire).

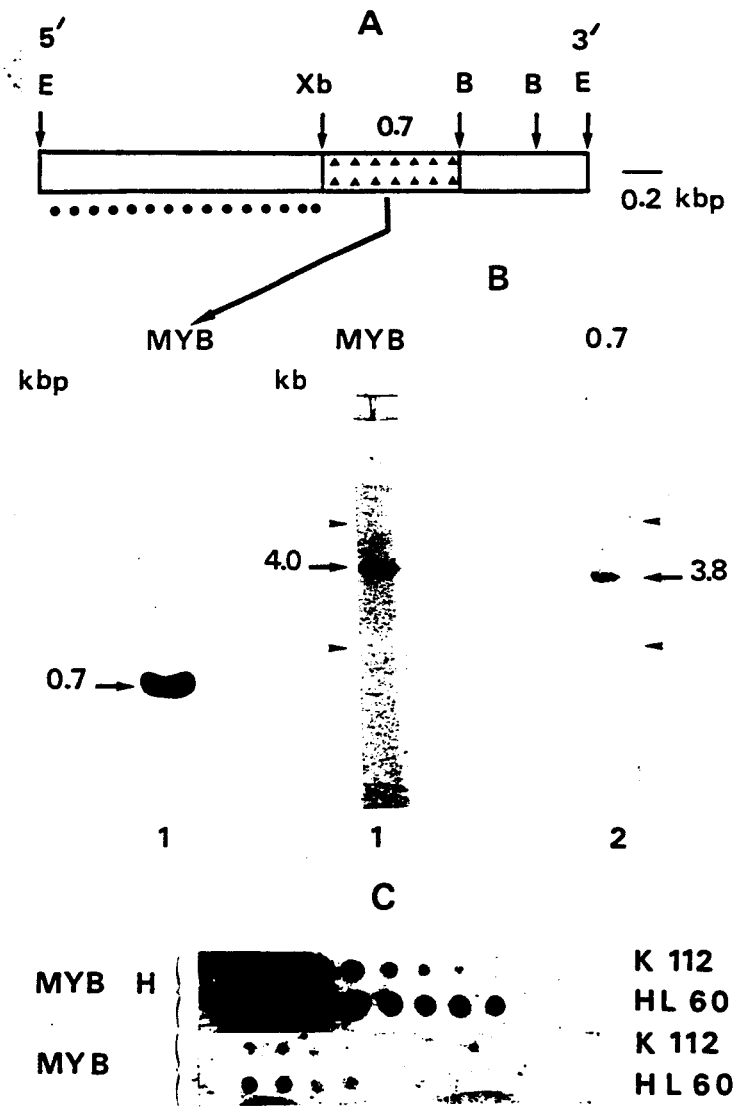
• • • : fragments contenant les séquences répétitives humaines

↙ : hybridation détectée entre le fragment du locus aviaire et le fragment homologue du locus humain.

contient toutes les séquences humaines homologues à la partie 5' de AMV v-myb ("E1" et "E2") soit environ 415 nucléotides de séquence virale codante pour une taille de 700 paires de bases (Figure 12A canal 1). Ceci donne un rapport séquences codantes ("exons") sur séquences non codantes ("introns") très élevé de l'ordre de 60 %. Cette sonde détecte dans différentes cellules hématopoïétiques humaines -et notamment dans la lignée de cellules érythroïdes, K562- un ARN messager de 3.8 kb, taille très voisine de celle du gène c-myb chez le poulet (Figure 12B, canaux 1 et 2).

De même, cette sonde utilisée en hybridation avec des dilutions successives d'ARN extraits de tumeurs et fixés sur nitrocellulose ("Dot-Blot" ; Thomas, 1981) se révèle de 5 à 10 fois plus sensible que la sonde virale v-myb, même lorsque cette dernière est utilisée en conditions non-strictes (Figure 12 C).

DISCUSSION. Les séquences cellulaires homologues à l'oncogène viral v-myb du virus de la myéloblastose aviaire (AMV) ont été clonées dans un phage recombinant lambda HM-1. Ces séquences présentent une structure en mosaïque répartie sur environ 13,6 kbp d'ADN cellulaire humain. Tout comme son homologue aviaire, ce locus semble être l'unique locus c-myb humain puisque récemment Franchini et al., (1983) ont cloné à partir d'autres banques de gènes humains des bactériophages recombinants contenant le même gène. Le clonage de c-myb dans deux espèces différentes nous a permis d'envisager la conservation phylogénétique de ce gène au cours de 300 millions d'années d'évolution. Le gène c-myb n'a apparemment subi aucun remaniement de structure et présente une structure colinéaire dans les deux espèces avec les "exons" distribués de façon similaire. Seule la faible hybridation entre


A
5′
3′
E
Xb
B
B
E
0.7
0.2 kbp
B
MYB
MYB
0.7
kbp
kb
4.0
3.8
0.7
1
1
2
C
MYB H
MYB
K 112
HL 60
K 112
HL 60

BU LILLE

Figure 12 : Transcription du gène c-myb humain

(A) l'ADN du plasmide pHM-2.8 a été digéré par les enzymes de restriction EcoRI (E), BamHI (B) et XbaI (Xb). Un unique fragment (▲ ▲) BamHI-XbaI 0.7 kbp hybride avec la sonde cDNA myb (canal 1). Ce fragment a pu être utilisé comme sonde humaine c-myb spécifique car il ne contient pas de séquences humaines répétitives (● ● ●).

(B) L'ARN c-myb humain est détecté dans une lignée érythroïde humaine k562, à l'aide de cette sonde 0.7 BamHI-XbaI (canal 2). La taille (3.8 kb) est très similaire à celle du transcrit aviaire détecté par le cDNA myb dans la lignée lymphoïde aviaire MSB1 (canal 1). Les têtes de flèches indiquent la position des ARN ribosomaux 18S et 28S utilisés comme marqueurs de taille.

(C) Des quantités décroissantes (facteur de dilution 2.5) d'ARN polyadenylé des lignées cellulaires hématopoïétiques humaines K112 et HL60 ont été déposées sur nitrocellulose selon la technique de Thomas (1980). L'ARN fixé a été hybridé en conditions strictes avec la sonde 0.7 kbp BamHI-XbaI. Après une autoradiographie de 24 h, la nitrocellulose a été déshybridée puis rehybridée avec le cDNA myb, en conditions relachées, et autoradiographiée pendant 24 h.

la sonde v-myb et le fragment EcoRI de 5.0 kbp -comprenant les séquences homologues à la région 3' terminale de l'oncogène viral- laisse supposer une moins bonne conservation de cette partie du gène (Figure 9B, Canal 2 ; Franchini et al., 1983). Néanmoins, cette homologie de structure ne semble pas limitée uniquement à la région du gène transduite dans le virus aviaire -région qui pourrait constituer la partie essentielle, donc la mieux conservée du gène- puisque les ARNs de 3,8 et 4,0 kb sont de taille très voisines, indiquant donc une conservation du locus entier. Les séquences présentes dans l'ARN messager cellulaire sont comme dans le cas du gène c-myb poulet situées immédiatement en 3' et 5' des séquences homologues à l'oncogène viral v-myb. (Franchini et al., 1983). Grâce à une sonde du fragment EcoRI 2.0 kbp, le gène c-myb humain a pu être localisé sur le chromosome 6 bandes q22-q24, région qui semble être transloquée dans différents cas de leucémies lymphocytaires ou de lymphomes cellulaires T (Dalla-Favera et al., 1982 ; Harper et al., 1983). De plus, le taux de transcription du gène c-myb humain, élevé dans les précurseurs hématopoïétiques immatures diminue avec la différenciation de ces cellules (Westin et al., 1982).

La connaissance de la carte physique et de la localisation chromosomique de c-myb ainsi que la disponibilité de fragments clonés permettant une détection sensible et reproductible de la transcription de ce gène devrait permettre une étude systématique de différentes tumeurs. Ceci devrait à long terme permettre de définir le rôle possible de c-myb dans l'induction de néoplasmes humains.

BIBLIOGRAPHIE

- ANDERSEN,P.R., DEVARE,S.G., TRONICK,S.R., ELLIS,R.W., ARONSON,S.A. et SCOLNICK,E.M.
Generation of Balb-MuSV and Ha-MuSV by type C retrovirus transduction of homologous transforming genes from different species.
Cell (1981) **26** : 129-134

- ASTRIN,S.M., ROBINSON,H.L., CRITTENDEN,L.B., BUX,E.G., WYBAN,J. et HAYWARD,W.S.
Ten genetic loci in the chicken that contain strucural genes for endogenous Avian Leukosis Virus.
Cold Spring Harb. Symp. Quant. Biol. (1980) 44 : 1105-1109

- BALTIMORE,D.
RNA-dependent DNA polymerase in virions of RNA tumor viruses.
Nature (1970) **226** : 1209-1211

- BATTEY,J., MOULDING,C., TAUB,R., MURPHY,W., STEWART,T., POTTER,H., LENOIR,G. et LEDER,P.
The human c-myc oncogene : structural consequences of translocation into the IgH locus in Burkitt lymphoma.
Cell (1983) 34 : 779-787

- BEARD,J.W.
Avian virus growths and their etiologic agents.
Adv. Cancer Res. (1963) 7 : 1-127

- BEEMON,K. et HUNTER,T.

Characterization of Rous Sarcoma Virus src gene products synthesized in vitro.

J. Virol (1978) **30** : 190-200

- BESMER,P., HARDY Jr, W.D., ZUCHERMAN,E.E., BERGOLD,P., LEDERMAN,L. et SNYDER Jr, H.W.

The Hardy-Zuckerman 2-FeSV : a new feline retrovirus with oncogene homology to Abelson-MuLV.

Nature (1983) **303** : 825-827

- BEUG,H., VON KIRCHBACH,A., DOEDERLEIN,G., COSCIENCE,J.F. et GRAF,T.

Chicken hematopoietic cells transformed by seven strain of defective avian leukemia viruses display three distinct phenotypes of differenciation.

Cell (1979) **18** : 375-390

- BEUG,H., HAYMAN,M.J. et GRAF,T.

Myeloblasts transformed by the avian acute leukemia virus E26 are hormone-dependent for growth and for expression of a putative myb containing protein p135 E26.

EMBO J. (1982) **1** : 1069-1073.

- BISHOP,J.M.

Enemies within : the genesis of retrovirus oncogenes.

Cell (1981) **23** : 5-6

- BISHOP,J.M.

Cellular oncogenes and retroviruses.

Ann. Rev. Biochem. (1983) **52** : 301-354

- BISTER,K., NUNN,M., MOSCOVICI,C., PERBAL,B., BALUDA,M.A. et DUESBERG,P.H.
Acute leukemia viruses E26 and avian myeloblastosis virus have related transformation specific RNA sequences but different genetic structures, gene products and oncogenic properties.
Proc. Natl. Acad. Sci. USA. (1982) **79** : 3677-3681

- BLAIR,D.G., OSKARSSON,M., WOOD,T.G., McCLEMENTS,W.L., FISHINGER,P.J. et VAN DE WOUDE,G.F.
Activation of the transforming potential of a normal cell sequence : A molecular model for oncogenesis.
Science (1981) **212** : 941-943

- BOETTIGER,D. et DURBAN,E.
Target cells for avian myeloblastosis virus in embryonic yolk sac and relationship of cell differenciation to cell transformation.
J. Virol. (1984) **49** : 841-847

- BOYLE,W.J., LIPSICK,J.S., REDDY,E.P. et BALUDA,M.A.
Identification of the leukemogenic protein of avian myeloblastosis virus and of its normal cellular homologue.
Proc. Natl. Acad. Sci USA (1983) **80** : 2834-2838.

- BRUGGE,J.S. et ERIKSON,R.L.
Identification of a transformation-specific antigen induced by an avian sarcoma virus.
Nature (1977) **269** : 346-348.

- BUNTE,T., GREISER-WILKE,I. et MOELLING,K.
The transforming protein of the MC29- related virus CMII is a nuclear DNA-binding protein whereas MH2 codes for a cytoplasmic RNA-DNA binding protein.
The EMBO J. (1983) 2 : 1087-1092

- CAIRNS,J.

The origin of human cancers.

Nature (1981) **289** : 353-357.

- CAPON,D.J., CHEN,E.Y., LEVINSON,A.D., SEEBURG,P.H. et GOEDDEL,D.V.

Complete nucleotide sequences of the T24 human bladder carcinoma oncogene and its normal homologue.

Nature (1983) **302** : 33-37.

- CHANG,E.H., FURTH,M.E., SCOLNICK,E. et LOWY,D.R.

Tumorigenic transformation of mammalian cells induced by a normal human gene homologous to the oncogene of Harvey murine sarcoma virus.

Nature (1982) **297** : 479-482.

- COFFIN,J.M.

Structure, replication and recombination of retroviruses genomes : some unifying hypotheses.

J. Gen. Virol. (1979) **42** : 1-26

- COFFIN,J.M., VARMUS,H.E., BISHOP,J.M., ESSEX,M., HARDY JR,W.D., MARTIN,G.S., ROSENBERG,N.E., SCOLNICK,E.M., WEINBERG,R.A. et VOGT,P.K.

A proposal for naming host cell-derived inserts in retrovirus genomes.

J. Virol. (1981) **40** : 953-957

- COHEN,J.B., UNGER,T., RECHAM,G., CANAANI,E. et GWOL,D.

Rearrangement of the oncogene c-mos in mouse myeloma NSI and hybridomas.

Nature (1983) **306** : 797-799

- COLL,J., SAULE,S., MARTIN,P., RAES,M.B., LAGROU,C., GRAF,T., BEUG,H., SIMON,I.E. et STEHELIN,D.
The cellular oncogenes c-myc, c-myb and c-erb are transcribed in defined types of avian hematopoietic cells.
Exp. Cell. Res. (1983a) **149** : 151-162

- COLL,J., RIGHI,M., de TAISNE,C., DISSOUS,C., GEGONNE,A. et STEHELIN,D.
Molecular cloning of the avian acute transforming retrovirus MH2 reveals a novel cell-derived sequence v-mil in addition to the myc oncogene.
EMBO J. (1983b) **2** : 2189-2194

- COLLINS,S. et GOUDINE,M.
Amplification of endogenous myc related DNA sequences in a human myeloid leukaemia cell line.
Nature (1982) **298** : 679-681.

- COOPER,G.M. et NEIMAN,P.E.
Transforming genes of neoplasms induced by avian lymphoid leukosis viruses.
Nature (1980) **287** : 656-659

- COOPER,G.M.
Cellular transforming genes.
Science (1982) **218** : 801-806

- CULLEN,B.R., LOMEDICO,P.T. et JU,G.
Transcriptionnal interference in avian retroviruses - implications for the promotor insertion model of leukaemogenesis.
Nature (1984) **307** : 241-245

- DALLA-FAVERA,R., FRANCHINI,G., MARTINOTTI,S., WONG-STAAL,F., GALLO,R.C. et CROCE,C.M.
Chromosomal assignement of the human homologous of feline sarcoma virus and avian myeloblastosis virus onc genes.
Proc. Natl. Acad. Sci. USA. (1982a) **79** : 4714-4717

- DALLA-FAVERA,R., BREGNI,M., ERIKSON,J., PATTERSON,D., GALLO,R.C. et CROCE,C.M.
Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells.
Proc. Natl. Acad. Sci. USA (1982b) **79** : 7824-7827

- DALLA-FAVERA,R., GELMANN,E.P., MARTINOTTI,S., FRANCHINI,G., PAPAS,T.S., GALLO,R.C. et WONG-STAAL,F.
Cloning and characterization of different human sequences related to the onc gene (v-myc) of avian myelocytomatosis virus (MC29).
Proc. Natl. Acad. Sci. USA. (1982c) **79** : 6497-6501.

- De FEO-JONES,D., SCOLNICK,E., KOLLER,R. et DHAR,R.
ras-related gene sequences identified and isolated from Saccharomyces cerevisae.
Nature (1983) **306** : 707-709

- DIAMOND,A., COOPER,G.M., RITZ,J. et LANE,M.A.
Identification and molecular cloning of the human B-lym transforming gene activated in Burkitt's lymphomas.
Nature (1983) **305** : 112-116.

- DODGSON,J.B., STROMMER,J. et ENGEL,D.J.
Isolation of the chicken B globin gene from a chicken DNA recombinant library.
Cell (1979) **17** : 879-887

- DONNER,P., GREISER-WILKE,I. et MOELLING,K.
Nuclear localization and DNA binding of the transforming gene product of avian myelocytomatosis virus.
Nature (1982) **296** : 262-266.

- DOWNWARD,J., YARDEN,Y., MAYES,E., SCRACE,G., TOTTY,N., STOCKWELL,P., ULLRICH,A., SCHLESSINGER,J. et WATERFIELD,M.D
Close similarity of epidermal growth factor receptor and v-erbB oncogene protein sequences.
Nature (1984) **307** : 521-525

- DUESBERG,P.H., BISTER,K. et MOSCOVICI,C.
Genetic structure of avian myeloblastosis virus released from transformed myeloblasts as a defective virus particle.
Proc. Natl. Acad. Sci. USA. (1980) **77** : 5120-5124.

- DUESBERG,P.H.
Retroviral transforming genes in normal cells.
Nature (1983) **304** : 219-226.

- ELLIS,R.W., De FEO,D., SHIH,T.Y., GONDA,M.A., YOUNG,H.A., TSUCHIDA,N., LOWY,D.R. et SCOLNICK,E.M.
The p21 src genes of Harvey and Kirsten sarcoma viruses originate from divergent members of a family of normal vertebrate genes.
Nature (1981) **292** : 506-511.

- FINKEL,T. et COOPER,G.M.
Detection of a molecular complex between ras proteins and transferrin receptor.
Cell (1984a) **36**, 1115-1121

- FINKEL,T., DER,C.J. et COOPER,G.M.
Activation of ras genes in human tumor does not affect localization, modification or nucleotide binding properties of p21.
Cell (1984b) 37 : 151-158

- FRANCHINI,G., WONG-STAAL,F., BALUDA,M.A., LENGEL,C. et TRONICK,S.R.
Structural organization and expression of human DNA sequences related to the transforming gene of avian myeloblastosis virus.
Proc. Natl. Acad. Sci. USA. (1983) 80 : 7385-7389

- FUNG,Y.K., LEWIS,W.G., CRITTENDEN,L.D. et KUNG,H.J.
Activation of the cellular oncogene c-erbB by LTR insertion : Molecular basis for induction of erytroblastosis by avian leukosis virus.
Cell (1983) 33 : 357-368

- FRYKBERG,L., PALMIERI,S., BEUG,H., GRAF,T., HAYMAN,M.J. et VENNSTRÖM,B.
Transforming capacities of avian erythroblastosis virus mutants deleted in the erbA or erbB oncogenes.
Cell (1983) 32 : 227-238

- GALIBERT,F., DUPONT de DINECHIN,S., RIGHI,M. et STEHELIN,D.
The second oncogene mil of avian retrovirus MH2 is related to the src gene family.
The EMBO J. (1984) 3 : 1333-1338

- GALLO,R.C. et WONG-STAAL,F.

Retroviruses as etiologic agents of some animal and human leukaemias and lymphomas and as tools for elucidating the molecular mechanisms of leukemogenesis.

Blood (1982) **60** : 545-557.

- GALLWITZ,D., DORATH,C. et SANDER,C.

A yeast gene encoding a protein homologous to the human c-has/bas proto-oncogene product.

Nature (1983) **306** : 704-707

- GAZIN,C., DUPONT de DINECHIN, HAMPE,A., MASSON,J.M., MARTIN,P., STEHELIN,D. et GALIBERT,F.

Nucleotide sequence of the human c-myc locus : provocative open reading frame within the first exon.

The EMBO J. (1984) 3 : 383-387

- GOLDFARB,M., SHIMIZU,K., PERUCHO,M. et WIGLER,M.

Isolation and preliminary characterization of a human transforming gene from T24 bladder carcinoma cells.

Nature (1982) **296** : 404-409

- GONDA,T.J., SHEINESS,D.K., FANSHIER,L., BISHOP,J.M., MOSCOVICI,C. et MOSCOVICI,M.G.

The genome and intracellular RNAs of avian myeloblastosis virus.

Cell (1981) Cell **23** : 279-290.

- GONDA,T.J. et BISHOP,J.M.

Structure and transcription of the cellular homolog (c-myb) of the avian myeloblastosis virus transforming gene (v-myb).

J. Virol. (1983) **46** : 212-220.

- GOUBIN,G., GOLDMAN,D.S., LUCE,J., NEIMAN,P.E. et COOPER,G.M.
Molecular cloning and nucleotide sequence of a transforming gene detected by transfection of chicken B-cell lymphoma DNA.
Nature (1983) **302** : 114-119.

- GRAF,T. et BEUG,H.
Avian leukaemia viruses. Interaction with their target cells in vivo and in vitro.
Bioch. Bioph. Acta (1978) **516** : 269-299.

- GRAF,T., OKER-BLOM,N., TODOROV,T.G. et BEUG,H.
Transforming capacities and defectiveness of avian leukaemia viruses OK10 and E26.
Virology (1979) **99** : 431-436.

- GRAF,T. et STEHELIN,D.
Avian leukaemia viruses oncogenes and genome structure.
Bioch. Bioph. Acta (1982) **651** : 245-271.

- HAMPE,A., LAPREVOTTE,I., GALIBERT,F., FEDELE,L.A. et SHERR,C.J.
Nucleotide sequences of feline retroviral oncogenes (v-fes) provide evidence for a family of tyrosene-specific protein kinase genes.
Cell (1982) **30** : 775-785.

- HAMPE,A., GOBET,M., SHERR,C.J. et GALIBERT,F.
Nucleotide sequence of the feline retroviral oncogene v-fms shows unexpected homology with oncogenes encoding tyrosine-specific protein kinases.
Proc. Natl. Acad. Sci. USA. (1984) **81** : 85-89.

- HARPER,M.E., FRANCHINI,G., LOVE,J., SIMON,M.I., GALLO,R.C. et WONG-STAAL,F.
Chromosomal sublocalization of human c-myb and c-fes cellular onc genes.
Nature (1983) **304** : 169-171.

- HARVEY,J.J.
An unidentified virus which causes the rapid production of tumours in mice.
Nature (1964) **204** : 1104-1105.

- HAYDAY,A.C., GILLIES,S.D., SAITO,H., WOOD,C., WIMAN,K., HAYWARD,W.S. et TONEGAWA,S.
Activation of a translocated human c-myc gene by an enhancer in the immunoglobulin heavy chain locus.
Nature (1983) **307** : 334-340.

- HAYMAN,M.J., ROYER-POKORA,B. et GRAF,T.
Defectiveness of avian erythroblastosis virus : synthesis of a 75 k gag-related protein.
Virology (1979) **92** : 31-45.

- HAYWARD,W.S., NEEL,B.G. et ASTRIN,S.M.
Activation of a cellular oncogene by promotor insertion in ALV induced lymphoid leukosis.
Nature (1981) **290** : 475-480.

- HELDIN,C.H. et WESTERMAERK,B.
Growth Factors : Mechanism of Action and relation to oncogenes.
Cell (1984) **37** : 9-20

- HEISTERKAMP,N., STEPHENSON,J.R., GROFFEN,J., HANSEN,P.F., de KLEIN,A., BARTRAM,C.R. et GROSVELD,G.
Localization of the c-abl oncogene adjacent to a translocation breakpoint in chronic myelocytic leukaemia.
Nature (1983) **306** : 239-242.

- HUNTER,T. et COOPER,J.A.
Epidermal growth factor induces rapid tydrosine phosphorylation of proteins in A431 human tumor cells.
Cell (1981) **24** : 741-752.

- IVANOV,X., MLADENOV,Z., NEDYALKOV,S. et TODOROV,T.G.
Experimental investigations into avian leukosis. Transmision experiments of certain diseases of the avian leukosis complex found in Bulgaria.
Bulgarian Acad. Sci. Bull. Inst. Pathol. Comp. Animaux Domestiques (1962) 9 : 5-36.

- JANSEN,H.W., LURZ,R., BISTER,K., BONNER,T.I., MARK,G.E. et RAPP,U.R.
Homologous cell-derived oncogenes in avian carcinoma virus MH2 and murine sarcoma virus 3611.
Nature (1984) **307** : 281-284.

- JOHNSON,A., HELDIN,C.H., WASTESON,A., WESTERMARK,B., DEUEL,D.F., HUANG,J.S., SEEBURG,P.H., GRAY,A., ULLRICH,A., SCRACE,G., STROOBANT,P. et WATERFIELD,M.D.
The c-sis gene encodes a precursor of the B chain of platelet-derived growth factor.
The EMBO J. (1984) 3 : 921-928

- KAWAI,S. et HANAFUSA,H.
Genetic recombination with avian tumor virus.
Virology (1972) **49** : 37-44.

- KLEIN, G.

The role of gene dosage and genetic transpositions in carcinogenesis.

Nature (1981) **294** : 313-318.

- de KLEIN,A., van KESSEL,A.G., GROSVELD,G., BARTRAM,C.R., HAGEMEYER,A., BOOTSMA,D., SPURR,N.K., HEISTERKAMP,N., GROFFEN,J. et STEPHENSON,J.R.

A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukemia.

Nature (1982) **300** : 765-767.

- KLEMPNAUER,K.H., GONDA,T.J. et BISHOP,J.M.

Nucleotide sequence of the retroviral leukaemia gene v-myb and its cellular progenitor c-myb : the architecture of a transduced oncogene.

Cell (1982) **31** : 453-463.

- KLEMPNAUER,K.H., RAMSAY,G., BISHOP,J.M., MOSCOVICI,M.G., MOSCOVICI,C., MAC GRATH,J.P. et LEVINSON,A.D.

The product of the retroviral transforming gene v-myb is a truncated version of the protein encoded by the cellular oncogene c-myb.

Cell (1983a) **33** : 345-355.

- KLEMPNAUER,K.H. et BISHOP,J.M.

Transduction of c-myb into avian myeloblastosis virus : locating point of recombination within the cellular gene.

J. Virol. (1983b) **48** : 565-572.

- KLEMPNAUER,K.H. et BISHOP,J.M.

Neoplastic transformation by E26 leukemia virus is mediated by a single protein containing domains of gag and myb genes.

J.Virol. (1984a) **50** : 280-283

- KLEMPNAUER,K.H., SYMONDS,G., EVAN,G.I. et BISHOP,J.M.
Subcellular localization of proteins encoded by oncogenes of avian myeloblastosis virus and avian leukemia virus E26 and by the chicken c-myb gene.
Cell (1984b) **37** : 537-547.

- KOHL,N.E., KANDA,N., SCHRECK,R., BRUNS,G., LATT,S.A., GILBERT,F. et ALT,F.W.
Transposition and amplification of oncogene related sequences in human neuroblastomas.
Cell (1983) **35** : 359-367.

- LAND,H., PARADA,L.F. et WEINBERG,R.A.
Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes.
Nature (1983a) **304** : 596-602.

- LAND,H., PARADA,L.F. et WEINBERG,R.A.
Cellular oncogenes and multistep carcinogenesis.
Science (1983b) Science **222** : 771-778.

- LANE,M.A., NEARY,D. et COOPER,G.M.
Activation of a cellular transforming gene in tumours induced by Abelson murine leukaemia virus.
Nature (1982) **300** : 659-661.

- LEDER,P., BATTEY,J., LENOIR,G., MOULDING,C., MURPHY,W., POTTER,H., STEWART,T. et TAUB,R.
Translocations among antibody genes in human cancer.
Science (1983) **222** : 765-771.

- LEPRINCE,D., GEGONNE,A., COLL,J., de TAISNE,C., SCHNEEBERGER,A., LAGROU,C. et STEHELIN,D.
A putative second cell-derived oncogene of the avian leukaemia retrovirus E26.
Nature (1983) **306** : 395-397.

- MACARA,I.G., MARINETTI,G.V. et BALDUZZI,P.C.
Transforming protein of avian sarcoma virus UR2 is associated with phosphatidylinositol kinase activity : possible role in tumorigenesis.
Proc. Natl. Acad. Sci USA (1984) **81** : 2728-2732.

- MacGRATH,J.P., CAPON,D.J., SMITH,D.H., CHEN,E.Y., SEEBURG,P.H., GOEDDEL,D.V. et LEVINSON,A.D.
Structure and organization of the human Ki-ras proto-oncogene and a related processed pseudogene.
Nature (1983) **304** : 501-506.

- MARTIN,G.S.
Rous Sarcoma Virus : a fonction required for the maintenance of the transformed state.
Nature (1970) **227** : 1021-1023.

- MILLER,D.A., CURRAN,T. et VERMA,I.M.
c-fos protein can induce cellular transformation : a novel mechanism of activation of a cellular oncogene.
Cell (1984) **36** : 51-60.

- MOLONEY,J.B.
A virus induced rhabdomyosarcoma of mice.
Natl. Cancer Inst. Monogr. (1966) **22** : 139-142.

- MOSCOVICI,M.G., JURDIC,P., SAMARUT,J., GAZZOLO,L., MURA,C.V. et MOSCOVICI,C.
Characterization of the hemopoietic target cells for the avian leukemia virus E26.
Virology (1983) **129** : 65-78.

- MOUNT,S.M.
A catalogue of splice junction sequences.
Nucl. Acids. Res. (1982) **10** : 459-472.

- MÜLLER,R., SLAMON,D.J., TREMBLAY,J.M., CLINE,M.J. et VERMA,I.M.
Differential expression of cellular oncogenes during pre and post natal development of the mouse.
Nature (1982) **299** : 640-644.

- MUSHINSKI,J.F., POTTER,M., BAUER,S.R. et REDDY,E.P.
DNA rearrangement and altered RNA expression of the c-myb oncogene in mouse plasmacytoid lymphosarcomas.
Science (1983) **220** : 795-798.

- NEEL,B.G., SURESH,C.J., CHAGANTI,R.S.K. et HAYWARD,W.S.
Two human c-onc genes are located on the long arm of chromosome 8.
Proc. Natl. Acad. Sci. USA (1982) **79** : 7842-7846.

- NEIL,J.C., HUGHES,D., Mc FARLANE,R., WILKIE,N.M., ONIONS,D.E., LEES,G. et JAREET,O.
Transduction and rearrangement of the myc gene by feline leukaemia virus in naturally occuring T-cell leukaemias
Nature (1984) **308** : 814-820.

- NEWBOLD,R.F. et OVERELL,R.W.
Fibroblast immortality is a prerequesite for transformation by EJ c-Ha-ras oncogene.
Nature (1983) **304** : 648-651.

- NUNN,M.F., SEEBURG,P.H., MOSCOVICI,C. et DUESBERG,P.H.
Tripartite structure of the avian erythroblastosis virus E26 transforming gene.
Nature (1983) **306** : 391-395.

- OSKARSSON,M., McCLEMENTS,W.L., BLAIR,D.G., MAIZEL,J.V. et VAN DE WOUDE,G.F.
Properties of a normal mouse cell DNA sequences (sarc) homologous to the src sequence of Moloney sarcoma virus.
Science (1980) **207** : 1222-1224.

- PAPKOFF,J., NIGG,E.A. et HUNTER,T.
The transforming protein of Moloney Murine Sarcoma Virus is a soluble cytoplasmic protein.
Cell (1983) **33** : 161-172.

- PAYNE,G.S., BISHOP,J.M. et VARMUS,H.E.
Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas.
Nature (1982) **295** : 209-214.

- PERBAL,B., CLINE,J.M., HILLYARD,R. et BALUDA,M.A.
Organization of chicken DNA sequences homologous to the transforming gene of avian myeloblastosis virus.
J. Virol. (1983) **45** : 925-940.

- PRYWES,R., FOULKES,J.G., ROSENBERG,N. et BALTIMORE,D.
Sequences of the A-MuLV protein needed for fibroblast and lymphoid cell transformation.
Cell (1983) **34** : 569-579.

- RABBITS,T.H., FORSTER,A., BAER,R. et HAMLYN,P.H.
Transcription enhancer identified near the human Cu immunoglobulin heavy chain gene is unavailable to the translocated c-myc gene in a Burkitt lymphoma.
Nature (1983) **306** : 806-809.

- RADKE,K., BEUG,H., KORNFELD,S. et GRAF,T.
Transformation of both erythroid and myeloid cells by E26, an avian leukemia virus that contains the myb gene.
Cell (1982) **31** 643-653.

- RALSTON,R. et BISHOP,J.M.
The protein products of the myc and myb oncogenes and adenovirus E1a are structurally related.
Nature (1983) **306** : 803-806.

- ROBBINS,K.C., ANTONIADES,H.N., DEVARE,S.G., HUNKAPILLER,M.W. et AARONSON,S.A.
Structural and immunological similarities between simian sarcoma virus gene product(s) and human platelet-derived growth factor.
Nature (1983) **305** : 605-608.

- ROUSSEL,M., SAULE,S., LAGROU,C., ROMMENS,C., BEUG,H., GRAF,T. et STEHELIN,D.
Three new types of viral oncogene of cellular origin specific for hematopoietic cell transformation.
Nature (1979) **281** : 452-455.

- ROWLEY,J.
Human oncogene locations and chromosome aberrations.
Nature (1983) **301** : 290-291.

- RUSHLOW,K.E., LAUTENBURGER,J.A., PAPAS,T.S., BALUDA,M.A., PERBAL,B., CHIRIKJIAN,J.G. et REDDY,E.P.
Nucleotide sequence of the transforming gene of Avian Myeloblastosis Virus.
Science (1982) **216** : 1421-1423.

- RULEY,H.E.

Adenovirus early region 1A enables viral and cellular transforming genes to transform primary cells in culture.

Nature (1983) **304** : 602-606.

- SAULE,S., COLL,J., RIGHI,M., LAGROU,C., RAES,M.B. et STEHELIN,D.

Two different types of transcription for the myelocytomatosis viruses MH2 and CMII.

EMBO J. (1983) **2** : 805-809.

- SCHIMKE,R.T.

In gene amplification R.T. Schimke ed. (New York : Cold Spring Harbor Laboratory press) (1982) : 327-333.

- SCHWAB,M., ALITALO,K., VARMUS,H.E., BISHOP,J.M. et GEORGE,D.

A cellular oncogene (c-Ki-ras) is amplified, overexpressed and located within karyotypic abnormalities in mouse adrenocortical tumor cells.

Nature (1983) **303** : 497-501.

- SCHWARTZ,D.E., TIZARD,R. et GILBERT,W.

Nucleotide sequence of Rous Sarcoma Virus.

Cell (1983) **32** : 853-869.

- SHANK,P.R., HUGHES,S.H., KUNG,H.J., MAJORS,J.E., QUINTRELL,N., GUNTAKA,R.V., BISHOP,J.M. et VARMUS,H.E.

Mapping unintegrated avian sarcoma virus DNA : Termini of linear DNA bear 300 nucléotides present once or twice in two species of circular DNA.

Cell (1978) **15** : 1383-1395.

- SHILO,B.Z. et WEINBERG,R.A.

Genes homologous to vertebrate oncogenes are conserved in Drosophila melanogaster.

Proc. Natl. Acad. Sci (1981) **78** : 6789-6792.

- SOTIROV,N.

Histone H5 in the immature blood cells of chickens wit leukosis induced by avian leukosis virus strain E26.

J. Nat. Cancer Inst. (1981) **66** : 1143-1149.

- SOUTHERN,E.M.

Detection of specific sequences among DNA fragments separated by gel electrophoresis.

J. Mol. Biol. (1975) **38** : 503-517.

- STEHELIN,D., GUNTAKA,R.V., VARMUS,H.E. et BISHOP,J.M.

Purification of DNA complementary to nucleotide sequences required for neoplastic transformation of fibroblasts by avian sarcoma viruses.

J. Mol. Biol. (1976a) **101** : 349-365.

- STEHELIN,D., VARMUS,H.E., BISHOP,J.M. et VOGT,P.K.

DNA related to the transforming gene(s) of avian sarcoma virus is present in normal avian DNA.

Nature (1976b) **260** : 170-173.

- STEHELIN,D., SAULE,S., ROUSSEL,M., SERGEANT,A., LAGROU,C., ROMMENS,C. et RAES,M.B.

Three new types of viral oncogenes in defective avian leukemia viruses : I Specific nucleotide sequences of cellular origin correlate with specific transformation.

Cold Spring Harbor Symp. Quant. Biol. (1980) **44** : 1215-1223

- STEPHENS,R.M., RUE,N.R., HIEBSCH,R.R., BOSE Jr, H.R. et GILDEN,R.V.
Nucleotide sequence of v-rel : the oncogene of reticuloendotheliosis virus.
Proc. Natl. Acad. Sci. USA. (1983) **80** : 6229-6233.

- SUGIMOTO,Y., WHITMAN,M., CANTLEY,L.C. et ERIKSON,R.L.
Evidence that the Rous sarcoma virus transforming gene product phosphorylates phosphotidylinositol and diacylglycerol.
Proc. Natl. Acad. Sci. USA. (1984) **81** : 2117-2121.

- SUKUMAR,S., NOTARIO,V., MARTIN-ZANCA,D. et BARBACID,M.
Induction of mammary carcinomas in rats by nitros-methylurea involves malignant activation of H-ras-1 locus by single point mutation.
Nature (1983) **306** : 658-661.

- TEMIN,H.M. et MIZUTANI,S.
RNA directed DNA polymerase in virions of Rous Sarcoma Virus.
Nature (1970) **226** : 1211-1213.

- TEMIN,H.M.
Origin of retroviruses from cellular movable genetic elements.
Cell (1980) **21** : 599-600.

- TONEGAWA,S.
Somatic generation of antibody diversity.
Nature (1983) **302** : 575-581.

- VAN BEVEREN,C.,VAN STRAATEN,F., CUVRAN,T., MÜLLER,R. et VERMA,I.M.
Analysis of FBJ-MuSV provirus and c-fos (mouse) gene reveals that viral and cellular fos gene products have different carboxy termini.
Cell (1983) **32** : 1241-1255.

- VARMUS,H.E.
Form and function of retroviral proviruses.
Science (1982) **216** : 812-820.

- WATT,R., STANTON,L.W., MARCU,K.B., GALLO,R.C., CROCE,C.M. et ROVERA,G.
Nucleotide sequence of cloned cDNA of human c-myc oncogene.
Nature (1983) **303** : 725-728.

- WATERFIELD,M.D., SCRACE,G.T., WHITTLE,N., STROOBANT,P., JOHNSON,A., WASTESON,A., WESTERMARK,B., HELDIN,C.H., HUANG,J.S. et DEUEL,T.F.
Platelet derived growth factor is structurally related to the putative transforming protein $p28^{sis}$ of simian sarcoma virus.
Nature (1983) **304** : 35-39.

- WEINBERG,R.A.
Fewer and fewer oncogenes.
Cell (1982) **30** : 3-4.

- WESTIN,E.H., GALLO,R.C., ARYA,S.K., EVA,A., SOUZA,L.M., BALUDA,M.A., AARONSON,S.A. et WONG-STEAL,F.
Differential expression of the amv gene in human hematopoietic cells.
Proc. Natl. Acad. Sci (1982) **79** : 2194-2198.

- YAMAMOTO,T., HIHARA,H., NISHIDA,T., KAWAI,S. et TOYOSHIMA,K.
A new avian erythroblastosis virus, AEV-H, carries erbB gene responsible for the induction of both erythroblastosis and sarcomas.
Cell (1983a) **34** : 225-232.

- YAMAMOTO,T., NISHIDA,T., MIYAJIMA,N., KAWAI,S., OOI,T. et TOYOSHIMA,K.
The erbB gene of avian erythroblastosis virus is a member of the src gene family.
Cell (1983b) **35** : 71-78.

- YUNIS,J.J.
The chromosomal basis of human Neoplasia.
Science (1983) 221 : 227-236

The EMBO Journal Vol.2 No.7 pp.1073–1078, 1983

The human DNA locus related to the oncogene *myb* of avian myeloblastosis virus (AMV): molecular cloning and structural characterization

D.Leprince, S.Saule, C.de Taisne, A.Gegonne, A.Begue, M.Righi and D.Stehelin*

INSERM U 186, Oncologie Moleculaire, Institut Pasteur, 15 rue C.Guérin, 59019 Lille Cedex, France

Communicated by D.Stehelin
Received on 17 March 1983; revised on 18 April 1983

Chicken and human cell DNA contains sequences homologous to the avian myeloblastosis virus oncogene, *v-myb*. These cellular sequences, *c-myb* (human) and *c-myb* (chicken), were isolated from libraries of human or chicken cell DNA fragments, generated by partial digestion with the restriction enzymes *Alu*I and *Hae*III, and compared. The chicken *c-myb* locus isolated from two distinct overlapping recombinant phages, contained five contiguous *Eco*RI fragments of 5.4, 1.1, 2.1, 2.2 and 9 kbp, accounting for all the bands seen with a *v-myb* probe in a complete *Eco*RI digest of chicken cellular DNA. Likewise, the screening of the human library yielded a recombinant phage hybridizing with the *v-myb* specific probe, that contained five *Eco*RI fragments of 2.8, 2.6, 2.0, 1.2 and 5.0 kbp (the last ending with an artificial *Eco*RI site, due to the construction of the library) belonging to the *c-myb* (human) locus. Probes using the *Eco*RI chicken DNA cloned fragments revealed corresponding contiguous *Eco*RI fragments in the human clone. Subsequent analyses of cellular polyadenylated mRNA extracted from human and chicken cells allowed the identification of single RNA species of 3.8 and 4.0 kb, respectively, as the representative transcripts of the *c-myb* locus in the two species. Thus, *c-myb* appears as a single locus in man and chicken, conserved with a similar structure in the two distantly related species. Our preparation of a specific human *c-myb* probe with an increased sensitivity on DNA/RNA blots should facilitate analyses concerning this gene in human normal or tumour cells or tissues.

Key words: avian myeloblastosis virus/viral oncogene/cellular oncogene/DNA cloning/RNA blotting

Introduction

Acute transforming retroviruses are responsible for the onset of different types of cancers in many vertebrate species. The avian defective leukemia viruses (DLV), whose cellular targets are primarily hematopoietic cells, is a particularly well studied group. When propagated *in vivo*, these viruses induce different types of cancers within a few weeks of inoculation and they transform hematopoietic cells *in vitro*; some strains also transform fibroblasts and epithelial cells (Graf and Beug, 1978; Graf and Stehelin, 1982). They are all defective for replication and need, for their propagation, helper virus containing the normal viral structural genes (*gag, pol* and *env*) (Graf and Stehelin, 1982). Their genomic RNA contains both helper virus-related sequences and specific sequences called *v-onc*, implicated in the transformation process. These *v-onc* sequences have normal cellular counterparts, *(c-onc)*, in the DNA of vertebrates and possibly invertebrates (Shilo and Weinberg, 1981). These highly transforming retroviruses are thought to have arisen from genetic recombination between poorly oncogenic helper viruses and phylogenetically stable cellular genes. How such recombination occurred is not yet well understood (Graf and Stehelin, 1982; Bishop, 1982).

The eight available DLV isolates have been divided into several groups according to their transforming potentials and *v-onc* genes. The avian myeloblastosis group includes avian myeloblastosis virus (AMV) and E26, two myeloblastosis-inducing virus strains which have been obtained as independent isolates (Graf and Beug, 1978). AMV (Beard, 1963; Moscovici, 1975) incudes myeloblastosis in chickens and transforms immature myeloid cells (Gazzolo *et al.*, 1979) as well as mature macrophages (Durban and Boettiger, 1981) from chick embryo yolk sac cultures; AMV is so far the only DLV strain which has a transforming capacity restricted to the hematopoietic lineage (Graf and Stehelin, 1982).

The 7.2-kb AMV genomic RNA contains a complete functional *gag* gene, most of the *pol* gene and a transforming sequence, *v-myb*, in place of *env* sequences (Souza *et al.*, 1980). The expression of *v-myb* is achieved through the presence of a subgenomic spliced RNA of 2.1 kb found in all AMV-transformed cells (Gonda *et al.*, 1981). This spliced RNA contains *v-myb* as well as regulatory sequences derived from the 5′ (U5) and 3′ (U3) part of the genomic RNA (Graf and Stehelin, 1982). Sequencing of *v-myb* revealed a single open reading frame, which might code for a protein of 30 000 daltons (265 amino acids) (Rushlow *et al.*, 1982) or, alternatively, of 45 000 daltons, if an upstream initiation codon is used (Klempnauer *et al.*, 1982).

E26 is another myeloblastosis virus independently isolated from a field case of 'erythroblastosis' (Ivanov *et al.*, 1962), that also harbours part of the *v-myb* sequences (Roussel *et al.*, 1979). E26 induces the proliferation of both erythroid and myeloid cells *in vivo* (Sotirov, 1981). A similar situation is observed with bone marrow cells transformed *in vitro* by E26 and seeded in erythroid- or myeloid-stimulating growth medium (Beug *et al.*, 1982; Radke *et al.*, 1982). Furthermore, E26 also transforms quail but not chicken fibroblasts in culture (Graf *et al.*, 1979), therefore showing an oncogenic spectrum different from that of AMV. In addition to *v-myb*, E26 contains a so far unidentified sequence (X) which could play a role in erythroid (and/or) fibroblastic transformation (Roussel *et al.*, 1979; Stehelin *et al.*, 1980). In contrast to AMV, the putative transforming protein of E26, designated p135, is translated from the 5.7-kb genomic length RNA (Bister *et al.*, 1982), as a fusion protein containing *gag* and *myb* determinants, as well as probably X-determinants (D.Leprince in preparation). E26 and AMV illustrate the two distinct types of expression used by retroviruses, they specify either a fusion protein (generally *gag-onc*) translated from a genomic length RNA, or an *onc* protein translated from a subgenomic RNA. The *v-myb* sequences have a cellular counterpart in chicken and mammalian DNA, called *c-myb* (Roussel *et al.*, 1979; Bergmann *et al.*, 1981). Using DNA from mouse-human somatic cell hybrids, this sequence was recently assigned to human chromosome 6 (Dalla Favera *et al.*, 1982).

*To whom reprint requests should be sent.

© IRL Press Limited, Oxford, England.

We describe our attempts to investigate the structure of the human *c-myb* DNA as compared with its chicken counterpart. In chicken DNA, digested by *Eco*RI, the *c-myb* locus was represented by five *Eco*RI fragments spanning ~20 kbp of cellular DNA. From a chicken random DNA library, we cloned two overlapping phages carrying all these *v-myb*-positive, *Eco*RI fragments and representing a single chicken DNA locus. From a human random DNA library, we found one recombinant DNA phage with a structure highly homologous to the chicken locus. Using cellular probes, we also showed that *c-myb* was transcribed in human and chicken hematopoietic cells as a single RNA transcript with a size of 3.8 and 4.0 kb, respectively. Finally, we prepared a human *c-myb* probe yielding a high signal on blots of human DNA/RNA that should facilitate studies of this oncogene in human cells and tissues.

Results

Isolation of recombinant phages containing c-myb sequences of human or chicken DNA

*Eco*RI digests of human and chicken DNA were separated by gel electrophoresis and transferred to nitrocellulose. The blots were hybridized with a *v-myb*-labelled cDNA probe prepared as described previously (Roussel *et al.*, 1979; Gonda *et al.*, 1981). Several positive fragments were detected in both human (2.8 and 2 kbp, see Figure 2A) and chicken (9.0, 5.4 and 2.2 kbp, see Figure 1A) DNA, indicating that both species contained cellular DNA *(c-myb)* related to the *v-myb* probe. To compare the structure of *c-myb* sequences in man and chicken, we screened, with our *v-myb* probe, libraries of random *Alu*I-*Hae*III human or chicken DNA fragments inserted in the λ phage Charon 4A after the attachment of synthetic *Eco*RI linkers (Maniatis *et al.*, 1978; Dodgson *et al.*, 1979). One human DNA phage (λ HM-1) containing *c-myb* sequences was isolated, whereas two distinct chicken DNA phages (λ CM-4 and CM-5) accounted for the *c-myb* locus. These three phages were characterized and compared as described below.

Characterization of the chicken c-myb locus

Analyses performed on purified *Eco*RI-digested phage DNAs by blot-hybridization using as probes *v-myb* or *Eco*RI fragments of the isolated phages showed that the two clones λ CM-4 and λ CM-5 overlapped and accounted for the *c-myb* *Eco*RI fragments seen in chicken cellular DNA (Figure 1A). The *Eco*RI pattern of these clones is shown in Figure 1B together with their hybridization with the *v-myb* probe. These fragments were ordered by restriction mapping as shown in Figure 1D. The overlapping was confirmed by the hybridization of the ^{32}P-labelled 1.5-kbp fragment (clone λ CM-4, Figure 1D) with the 9.0-kbp fragment of clone λ CM-5 (Figure 1C). The chicken *c-myb* locus was recently described by Klempnauer *et al.* (1982), and our two phages λ CM-4 and CM-5, although slightly different from theirs, cover the same locus and confirm their results.

Structure of the c-myb locus in human DNA

The human *c-myb* DNA phage (λ HM-1) restricted with *Eco*RI (Figure 2B lane 1) contained five fragments (5.0, 2.8, 2.6, 2.0 and 1.2 kbp) positive with the *v-myb* probe (Figure 2B, lane 2). This indicated that the strong band of 2.8 kbp seen in a total human DNA digest (Figure 2A) was in fact a doublet (2.8 and 2.6 kbp) not resolved in the gel and that two

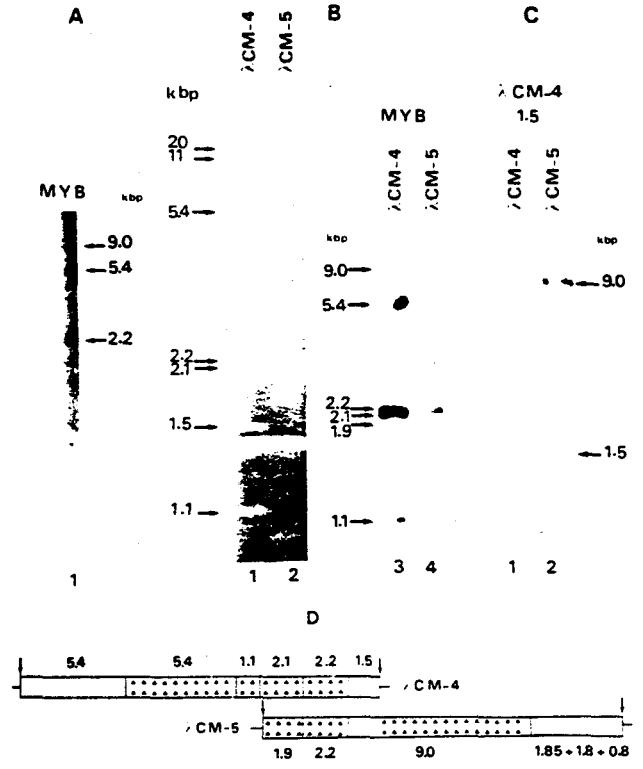
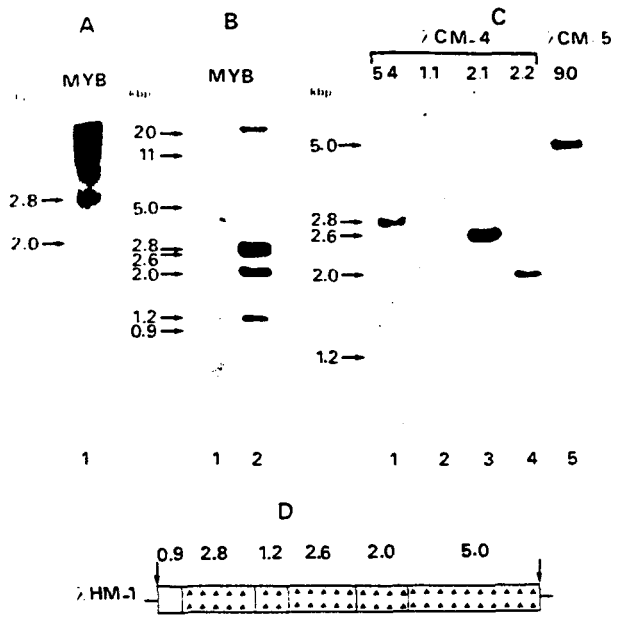


Fig. 1. Mapping of the chicken *c-myb* locus. **(A)** *Eco*RI-digested chicken DNA was size-separated by gel electrophoresis, transferred onto nitrocellulose and hybridized with cDNA myb. Adenovirus 2 DNA labelled *in vivo* with [^{14}C]thymidine was digested by *Eco*RI and run on the same gel as size marker (data not shown). **(B)** DNA from recombinant phages λ CM-4 **(lane 1)** and CM-5 **(lane 2)** was isolated, digested with *Eco*RI, subjected to 1% agarose gel electrophoresis and stained with ethidium bromide. The *Eco*RI DNA fragments separated as described above were blotted to nitrocellulose filters and hybridized to cDNA myb: λ CM-4 **(lane 3)** and λ CM-5 **(lane 4)**. **(C)** The overlapping of the phages was demonstrated by the hybridization of the 1.5-kbp fragment of clone λ CM-4 **(lane 1)** with the 9.0-kbp fragment of clone λ CM-5 **(lane 2)**. **(D)** Organization deduced from the above analyses of the chicken DNA sequences cloned in the recombinant phages λ CM-4 and CM-5. Boxes represent the *Eco*RI DNA fragments with their size indicated in kbp. ↓ Artificial *Eco*RI site due to the *Eco*RI DNA linker used to construct the library. (▲▲) Fragments hybridizing with cDNA myb.

additional *c-myb* *Eco*RI DNA fragments (5.0 and 1.2 kbp) were not initially detected. After the determination of the structure of clone λ HM-1 (Figure 2D), we analyzed the homology between *c-myb* (human) and *c-myb* (chicken) loci and the transcriptional orientation of *c-myb* (human) by using size-separated *Eco*RI digests of clone λ HM-1 and different ^{32}P-labelled *Eco*RI fragments of *c-myb* (chicken) (Figure 1D), as probes. Figure 2C (lanes 1–5), shows that each chicken *c-myb* fragment hybridized with a single counterpart fragment of clone λ HM-1 (see also Figures 2D and 4). Thus, the *c-myb* (human) and *c-myb* (chicken) loci appeared co-linear and of similar organization. Therefore, the orientation of *c-myb* (human) was as follows: 5′-2.8-1.2-2.6-2.0-5.0-3′ by reference to the viral AMV RNA defined as the 'plus' strand. The 5.0-kbp band was in fact an artificial *Eco*RI fragment. It hybridized both with the *v-myb* probe (Figure 2B, lane 2) and the 9.0-kbp chicken DNA fragment containing the 3′-terminal sequences homologous to *v-myb* (Figure 2C, lane 5). Therefore, this *c-myb* (human) clone



ig. 2. The human *c-myb* locus shares structural homology with its chicken ounterpart. **(A)** Human placenta DNA digested by *Eco*RI and blotted om agarose gel was hybridized with ^{32}P-labelled cDNA myb. **(B)** λ HM-1 hage DNA was digested by *Eco*RI. Fragments were separated on a 1% garose gel visualized by ethidium bromide staining **(lane 1)** then trans-rred to nitrocellulose and hybridized with ^{32}P-labelled cDNA myb **(lane**). **(C)** Aliquots of clone λ HM-1 DNA digested with *Eco*RI were parated as in **(B)** and hybridized with the following labelled *Eco*RI agments of the chicken clone CM-4: 5.4 kbp **(lane 1)**; 1.1 kbp **(lane 2)**; .1 kbp **(lane 3)**; 2.2 kbp **(lane 4)**. **Lane 5** was hybridized similarly with the -kbp fragment of the chicken clone λ CM-5. **(D)** Organization deduced om the above analyses for the *c-myb* (human) locus in recombinant hage λ HM-1. Boxes represent the *Eco*RI DNA fragments with their size dicated in kbp as in Figure 1. ↓ Artificial *Eco*RI site due to the *Eco*RI NA linker used to construct the library. (▲▲) Fragments hybridizing ith cDNA myb.

robably represents all of the human DNA sequences omologous to *v-myb*.

Further analyses with the fragments of the *c-myb* (human) ocus subcloned into plasmids (pKH 47, Hayashi, 1980), igested by restriction nucleases, and hybridized with viral DNA myb allowed us to localize more precisely the regions f homology with the virus ('exon-like' fragments) in the uman *c-myb* locus (Figure 4). We next examined whether he human *c-myb* locus contained repetitive DNA sequences. Human DNA was sheared, alkali-denatured, reassociated at a C_0t of 400 mol.s/l and fractionated on hydroxylapatite: the apidly reassociating DNA was nick-translated and used as a ^{32}P-labelled probe to localize repetitive sequences in phage λ HM-1. The results (Figure 3A lane 1) indicated the presence f such sequences only in the 5.0- and 2.8-kbp *Eco*RI ragments. These were localized more precisely by similar ex-eriments performed on subclones of these fragments in lasmids. They were found in the 2.8-kbp fragment within a ingle *Eco*RI-*Xba*I 1.45-kbp band (Figure 3B lanes 1 and 2) nd in the 5.0-kbp fragment within a *Xba*I-*Xba*I 1.4-kbp and (Figure 3B lane 3). Such bands were negative with a *v-nyb* probe (Figure 4).

In conclusion, *c-myb* (human) and *c-myb* (chicken) showed highly conserved structure indicating that no extensive gene earrangement had occurred in this locus during speciation.

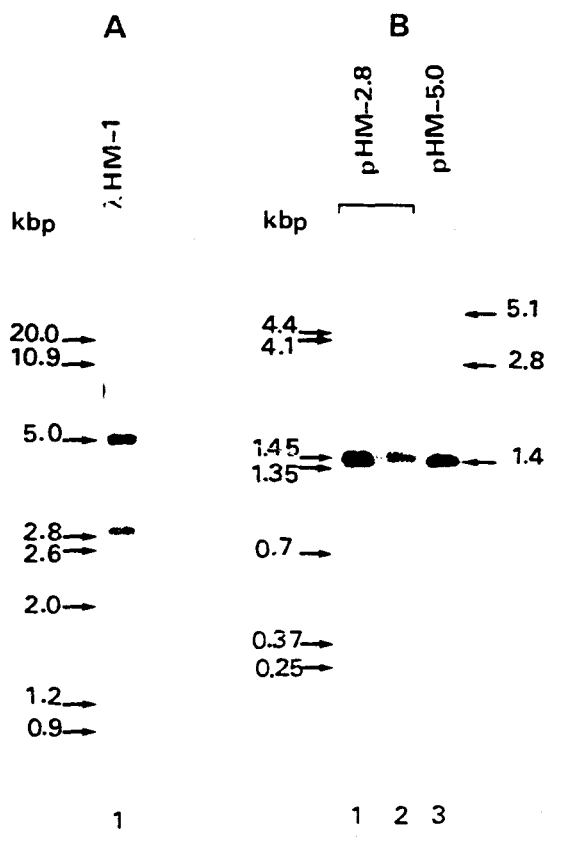


Fig. 3. Localization of human repetitive DNA sequences in clone λ HM-1. **(A)** λ HM-1 phage DNA was digested by *Eco*RI. Fragments were separated on a 1% agarose gel then transferred onto nitrocellulose and hybridized **(lane 1)** with ^{32}P-labelled rapidly reassociating total human DNA (repetitive DNA). **(B)** The 5.0- and 2.8-kbp *Eco*RI fragments positive with the human repetitive probe were subcloned into plasmid pKH 47 to yield pHM-5.0 and pHM-2.8, respectively. pHM-2.8 was digested either by *Eco*RI and *Xba*I **(lane 1)** or by *Eco*RI, *Xba*I and *Bam*HI **(lane 2)**, submitted to gel electrophoresis, blotted onto nitrocellulose and hybridized with the [^{32}P]human repetitive DNA. The same experiment was performed with a *Xba*I digest of pHM-5.0 **(lane 3)**. Note that the 0.7-kbp *Bam*HI-*Xba*I fragment of pHM-2.8 and the 2.8-kbp *Xba*I-*Xba*I fragment of pHM-5.0 positive with cDNA myb (see Figure 4) did not contain repetitive sequences.

Detection of human and chicken RNA transcripts with cloned human or chicken c-myb probes

The cellular analog of the AMV transforming gene, *c-myb*, is transcribed in many hematopoietic immature chicken cells as a single mRNA species of 4.0 kb (Gonda *et al.*, 1982) transcribed from the same strand as the viral RNA in AMV (Coll *et al.*, 1983). This result was confirmed in the course of the present study by analysis of the cellular RNAs from a chicken lymphoid cell line (MSB1) transformed by Marek disease-virus (MDV) (Figure 5B, lane 1). Moreover, the same hybridization pattern was obtained with the cloned 2.2-kbp *Eco*RI fragment from λ CM-4 (Figure 5B, lane 2).

We decided to prepare a human *c-myb* probe from the plasmid pHM-2.8 containing the 2.8-kbp *Eco*RI fragment of λ HM-1. When digested by *Eco*RI, *Xba*I and *Bam*HI (Figure 5A), the subclone pHM-2.8 yielded a single 0.7-kbp fragment positive with *v-myb* (Figure 5A, lane 1), and devoid of the human repetitive sequences (Figure 3B, lane 2) that, when used as a probe, allowed the detection in the human erythroid

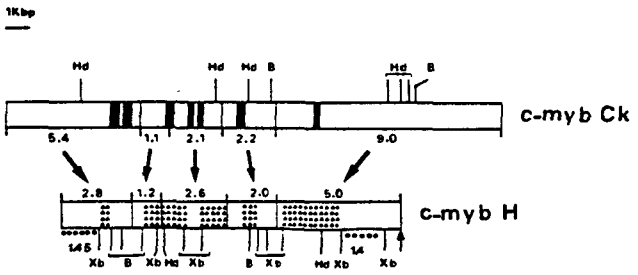


Fig. 4. *c-myb* (chicken) and *c-myb* (human) loci. Detailed restriction maps of the *c-myb* (chicken) and *c-myb* (human) loci were established as described for *Eco*RI with the following endonucleases: *Bam*HI (B), *Hind*III (Hd) and *Xba*I (Xb) (the last used only for the human clone). *c-myb* (chicken) exons (■) have been mapped according to Klempnauer *et al.*, 1982. Regions of homology with cDNA myb (▲▲) were localized in the *c-myb* (human) locus by double digestion of λ phage HM-1 or of plasmid clones containing human *c-myb* *Eco*RI fragments followed by blot hybridization. Repeated sequences (● ●) were determined similarly using a ^{32}P-labelled rapidly reassociating human DNA (see Materials and methods). Arrows indicate the human *Eco*RI fragments detected with the homologous chicken *Eco*RI fragment used as a probe.

cell line K562 of a single RNA species of ~3.8 kb (Figure 5B, lane 3). Comparative studies done on Northern blots (data not shown) or dot blots (Figure 5C) of RNA from human hematopoietic cell lines indicated that our selected partial human *myb* probe gave a much stronger signal (5- to 10-fold) than the *v-myb* probe.

Discussion

We have molecularly cloned and characterized a human *c-myb* locus in the recombinant λ phage HM-1. It spreads over at least the 13.6 kbp of DNA sequences which share homology with a *v-myb* probe. The structure of this human *c-myb* locus appears similar to its chicken equivalent which we have cloned for comparative studies; at the resolution of the techniques we used (see Figure 4), the exons are distributed similarly in the human and chicken genes. Thus, the organization of the *c-myb* locus seems to be conserved and sequence analyses will be required to evaluate more precisely this phylogenetic relatedness.

All our evidence favours a single *c-myb* locus in human as well as chicken cell DNA: our phages account for all the bands seen by *Eco*RI digestion of cell DNA and no phage with an alternate structure was found during the screening of the libraries used. Other authors have also concluded that there is a single *c-myb* locus in the chicken DNA (Klempnauer *et al.*, 1982; Perbal *et al.*, 1982b), and our results extend this conclusion to human DNA; *c-myb* differs from other *c-onc* genes which appear to belong to gene families (*c-ras*, Ellis *et al.*, 1981; *c-src*, Kitamura *et al.*, 1982) or to give rise to pseudogenes (*c-myc*, Dalla-Favera *et al.*, 1982).

The *c-myb* locus is transcribed in several immature chicken hematopoietic cells (Gonda *et al.*, 1982) as a single RNA species of 4.0 kb. Our studies indicate a similar single transcript of 3.8 kb in human cells (K562 cell line, Westin *et al.*, 1982). Thus, in both species, the size of the *c-myb* RNA is considerably larger than the 1196 nucleotides representing *v-myb* in AMV. The *v-myb*-transduced sequences lack both initiation and termination codons necessary to make the *v-myb* protein, these signals being provided by the virus (Klempnauer *et al.*, 1982). It follows that the *c-myb* locus in the two species must extend into nucleotide sequences both upstream and downstream of the homologous region to yield the 3.8-or

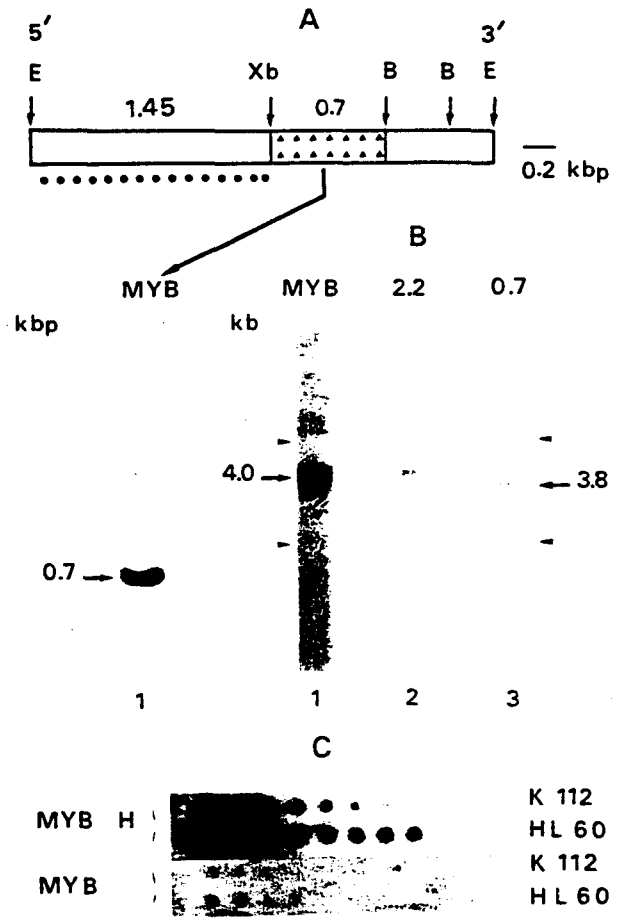


Fig. 5. Chicken and human *c-myb* mRNA transcripts. **(A)** The 2.8-kbp *Eco*RI fragment of clone λ HM-1 was subcloned into plasmid pKH 47: pHM-2.8. This plasmid DNA was digested simultaneously by *Eco*RI (E), *Bam*HI (B) and *Xba*I (Xb), separated by agarose gel electrophoresis and transferred to nitrocellulose. A single 0.7-kbp fragment (▲▲) hybridized with cDNA myb **(lane 1)** and was used as a human *c-myb*-specific probe hereafter. It contained no repetitive sequences (● ●), unlike the *Eco*RI-*Xba*I 1.45-kbp adjacent fragment. **(B)** Polyadenylated RNA from the chicken lymphoid cell line MSB1 **(lanes 1 and 2)** was denatured with glyoxal, separated by agarose gel electrophoresis, bound to activated DBM paper and hybridized with cDNA myb **(lane 1)** or with ^{32}P-labelled *c-myb* (chicken) 2.2-kbp *Eco*RI fragment of λ CM-5 **(lane 2)**. Similarly, RNA from the human erythroid cell line K562 was treated as above and hybridized with ^{32}P-labelled 0.7-kbp *c-myb* (human) fragment described in **(A) (lane 3)**. Arrow-heads indicate the 28S and 18S rRNA markers. **(C)** Decreasing amounts (2.5 time dilutions) of polyadenylated RNA from human hematopoietic cell lines K112 and HL60 were applied to nitrocellulose paper as described in Materials and methods. The bound RNA was hybridized with nick-translated 0.7-kbp DNA fragment (10^6 c.p.m.) from plasmid pKH HM-2.8, in stringent conditions (24 h at 41°C, hybridization buffer containing 3 x SSC, washed in 0.1 x SSC at 50°C, 4 times, 10 min) and the blot dried and exposed to X-ray film for 1 day at −70°C (MYB H). After dehybridization, the blot was rehybridized with viral cDNA myb (10^6 c.p.m.) but in non-stringent conditions (24 h 41°C, hybridization buffer containing 7.5 x SSC, washed in 2 x SSC at 41°C, 4 times, 10 min), then dried and exposed similarly (MYB).

4.0-kb RNA. We do not know if those portions of the *c-myb* locus have been phylogenetically conserved. We also do not understand how a cellular gene like *c-myb* became an oncogene in the virus. Research will have to concentrate on three possible mechanisms, not mutually exclusive: (1) an increased level of transcription due to the viral promoter; (2) a virally truncated *c-myb* gene; (3) a *c-myb* gene mutated within the virus.

The human *c-myb* RNA was revealed using a cloned man *c-myb* probe (0.7 kbp *Bam*HI-*Xba*I fragment of the age λ HM-1, see Figure 4) devoid of repetitive sequences, at gave at least a 5 – 10 times better signal as compared with e *v-myb* probe (see Figure 5). This human *c-myb*-specific obe should prove to be a valuable and sensitive reagent to alyse the size and accumulation of *c-myb* transcripts in ny types of human cells or tumours by Northern blotting *in situ* hybridization. It should also facilitate studies at the NA level: for example, the *c-myb* locus has recently been apped to human chromosome 6 q22-24 (Harper *et al.*, 83). Translocations/deletions possibly involving this gene ve been described in acute lymphoid leukemia (ALL) pa- nts (deletions in bands 6q21-q25 or in ovarian carcinomas anslocation 6:14 with break point at 6q21) (Rowley, 1983). ie availability of human specific *c-myb* probes will facilitate ture studies.

aterials and methods

lls, viruses and cloned DNA fragments

Plasma from AMV-infected birds was obtained from J.Beard (Life Science :.).

MDCC-MSB1 is a chicken lymphoid T-type cell line obtained by *in vitro* nsformation with MDV (Nazerian and Sharma, 1979), K562 a human throid cell line (Lozio and Lozio, 1975), HL60, a promyelocytic cell line ollins *et al.*, 1977) and K112, an Epstein-Barr virus-positive cell line ginating from an acute granulocytic leukemia patient (A.Karpas, personal mmunication). The chicken and human DNA libraries were obtained ough the courtesy of Dodgson *et al.* (1979) and Maniatis *et al.* (1978), pectively.

eparation of ^{32}P virus-specific cDNAs

Representative radioactive cDNA (cDNArep) was synthesized on a 70S td B viral RNA template, using exogenous AMV reverse transcriptase Beard, Life Science Inc., through the auspices of the Office of Program sources and Logistics, NCI) and sonicated calf thymus DNA as primer. NA myb specific for AMV sequences, was selected from AMV [^{32}P]BAI-A presentative cDNA by subtractive hybridization as previously described oussel *et al.*, 1979; Gonda *et al.*, 1981).

olation of chicken and human, v-myb homologous recombinant phages

Random libraries of chicken or human DNA fragments obtained by partial ul-*Hae*III digestion, and inserted with artificial *Eco*RI linkers in the *Eco*RI of the EK2 certified vector Charon 4A, were screened with ^{32}P-labelled NA myb and cDNA rep probes according to Benton and Davis (1977). th-positive, rep-negative plaques were purified twice (Maniatis *et al.*, 1978) d grown to high titer on *Escherichia coli* strain DP50.

eparation of DNAs

Normal chicken high mol. wt. (HMW) DNA was obtained from primary ltures of chicken embryo fibroblasts prepared from 11-day-old fertile eggs own Leghorn, Institut Gustave Roussy, Villejuif, France) following a ssical procedure (Saule *et al.*, 1982). Human HMW DNAs was obtained by same procedure from total human placenta. High yields of plasmid DNA re recovered from chloramphenicol (150 μg/ml) treated bacteria (12 h at C) after induction of cell lysis by lysozyme (1 mg/ml) and Triton X-100 1 w/v) treatment (Perbal *et al.*, 1982a). Recombinant phage DNAs were pared as described by Saule *et al.* (1982).

eparation of ^{32}P-labelled DNA fragments

Chicken or human *c-myb* fragments were purified twice by preparative gel ctrophoresis after digestion with the appropriate enzyme(s), recovered from gel by one cycle of freeze-thawing at −70°C and labelled by nick-nslation in the presence of [^{32}P]dCTP using a New England Nuclear nick-nslation kit. Specific activities of the probes were 50 – 100 x 10^6 c.p.m./μg DNA.

uman repetitive DNA probe

DNA from human placenta was sonicated to an average size of 4 – 6S, naturated and re-annealed to a C_0t of 400 mol.s/l, then fractionated on droxylapatite. The double-stranded DNA recovered was dialyzed, ethanol-cipitated and used for nick-translation labelling.

loning of c-myb EcoRI DNA fragments

After digestion of cloned recombinant phage DNAs by *Eco*RI, the fragments to be subcloned were purified by preparative agarose gel electrophoresis. *Eco*RI-digested plasmid pKH47 was treated with calf alkaline phosphatase. Fragment and vector pKH47 DNAs were ligated by T4 DNA ligase (Bethesda Research Laboratories Inc, USA) and used to transform calcium-treated cells of *E. coli* strain HB 101. Transformants containing the appropriate recombinant plasmids were detected by colony hybridization using different ^{32}P-labelled fragments as probes and purified twice on agar plates before large scale cultures.

Restriction mapping

Restriction endonucleases were obtained from Boehringer (Mannheim, FRG) and BRL Inc. Digested DNA fragments were size separated by electrophoresis in horizontal agarose gels and transferred to nitrocellulose according to Southern (1975).

Hybridization of blots to ^{32}P-labelled DNA, washing, and autoradiography at −70°C using Kodak X-ray films, Dupont Lightning Plus X-ray intensifying screens, were performed as previously reported (Saule *et al.*, 1982).

Agarose gel electrophoresis and transfer-hybridization of polyadenylated cellular RNAs

After extraction of total cellular RNA, polyadenylated RNAs were isolated by chromatography on oligo(dT)-cellulose (T3, Collaborative Research). RNA samples were then denatured by a glyoxal/dimethyl sulphoxide treatment and submitted to electrophoresis in 1% agarose horizontal gels in phosphate 0.02 M pH 6.8 buffer. Size-separated RNAs were transferred to diazobenzyloxymethyl (DBM) paper, hybridized with ^{32}P-labelled DNA fragments, washed and submitted to autoradiography (Saule *et al.*, 1982).

Dot blot hybridizations

These were performed essentially as described by Thomas (1980). 2.5 serial RNA dilutions were spotted onto nitrocellulose, the first spot representing ~0.8 μg of polyadenylated RNA. Hybridization, washing and autoradiography were carried out as described above for RNA transferred from gels.

Acknowledgements

We thank J.Dodgson, J.Strommer and D.Engel for providing the chicken DNA library; T.Maniatis for the human DNA library; A.Karpas for cell line K 112; C.Lagrou and A.Schneeberger for expert technical help; M.B.Raes for help in the preparation of this manuscript and N.Devassine for patient typing. This work was supported by grants from Institut National de la Santé et de la Recherche Médicale (CRL 812043, CRL 822033), Centre National de la Recherche Scientifique (ATP 98270), Délégation Générale à la Recherche Scientifique et Technique (81 L 0725) and the Pasteur Institute of Lille.

References

Beug,H., Hayman,M.J. and Graf,T. (1982) *EMBO J.*, **1**, 1069-1073.
Beard,J.W. (1963) *Adv. Cancer Res.*, **7**, 1-127.
Benton,W.D. and Davis,R.W. (1977) *Science (Wash.)*, **196**, 180-182.
Bergmann,D.G., Souza,L.M. and Baluda,M.A. (1981) *J. Virol.*, **40**, 450-455.
Bishop,J.M. (1982) in *The Molecular Biology of Tumour Viruses*, Volume **II**, published by Cold Spring Harbor Laboratory Press, NY.
Bister,K., Nunn,M., Moscovici,C., Perbal,B., Baluda,M.A. and Duesberg, P.H. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 3677-3681.
Coll,J., Saule,S., Martin,P., Raes,M., Graf,T., Beug,H., Simon,J.E., Lagrou,C. and Stehelin,D. (1983) *Exp. Cell Res.*, in press.
Collins,S.J., Gallo,R.G. and Gallagher,R.E. (1977) *Nature*, **270**, 347-349.
Dalla Favera,R., Franchini,G., Martinotti,S., Wong-Staal,F., Gallo,R.C. and Croce,C.M. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4714-4717.
Dodgson,J.B., Strommer,J. and Engel,J.D. (1979) *Cell*, **17**, 879-887.
Durban,E.M. and Boettiger,D. (1981) *J. Virol.*, **37**, 488-492.
Ellis,R.W., Defeo,D., Shih,T.Y., Gonda,M.A., Young,H.A., Tsuchida,N., Lowy,D.R. and Scolnick,E.M. (1981) *Nature*, **292**, 506-510.
Gazzolo,L., Moscovici,C., Moscovici,M.G. and Samarut,J. (1979) *Cell*, **16**, 627-638.
Gonda,T.J., Sheiness,D.K., Fanshier,L., Bishop,J.M., Moscovici,C. and Moscovici,M.G. (1981) *Cell*, **23**, 279-290.
Gonda,T.J., Sheiness,D.K. and Bishop,J.M. (1982) *Mol. Cell. Biol.* **2**, 617-624.
Graf,T., Okerblom,N., Todorov,T.G. and Beug,H. (1979) *Virology*, **99**, 431-436.
Graf,T. and Beug,H. (1978) *Biochim. Biophys. Acta*, **516**, 269-278.
Graf,T. and Stehelin,D. (1982) *Biochim. Biophys. Acta*, **651**, 245-271.
Harper,M.E., Simon,M.I., Franchini,G., Gallo,R.C. and Wong-Staal,F. (1983) *Proc. Natl. Acad. Sci. USA*, in press.
Hayashi,K. (1980) *Genet.*, **11**, 109-115.

Ivanov,X., Mladenov,Z., Nedyalkov,S. and Todorov,T.G. (1962) *Bull. Inst. Pathol. Comp. Anim. Domest.*, **9**, 5-36.
Kitamura,N., Kitamura,A., Toyoshima,K., Hirayama,Y. and Yoshida,M. (1982) *Nature*, **297**, 205-208.
Klempnauer,K.H., Gonda,T.J. and Bishop,J.M. (1982) *Cell*, **31**, 453-463.
Lozio,C.B. and Lozio,B.B. (1975) *Blood*, **45**, 321-334.
Maniatis,T., Hardison,R.C., Lacy,E., Lauer,J., O'Connel,C., Quon,D., Sim,G.K. and Efstradiatis,A. (1978) *Cell*, **15**, 687-701.
Moscovici,C. (1975) *Curr. Top. Microbiol. Immunol.*, **71**, 79-101.
Nazerian,K. and Sharma,J.M. (1979) *J. Natl. Cancer Inst.*, **54**, 277-279.
Perbal,B. and Baluda,M.A. (1982a) *J. Virol.*, **41**, 250-257.
Perbal,B. and Baluda,M.A. (1982b) *J. Virol.*, **44**, 586-594.
Radke,K., Beug,H., Kornfeld,S. and Graft,T. (1982) *Cell*, **31**, 643-653.
Roussel,M., Saule,S., Lagrou,C., Rommens,C., Beug,H., Graf,T. and Stehelin,D. (1979) *Nature*, **281**, 452-455.
Rowley,J.D. (1983) *Nature*, **301**, 290-291.
Rushlow,K.E., Lautenberger,J., Papas,T.S., Baluda,M., Perbal,B., Chirikjian,J.G. and Reddy,E.P. (1982) *Science (Wash.)*, **216**, 1421-1423.
Saule,S., Sergeant,A., Torpier,G., Raes,M.B., Pfeifer,S. and Stehelin,D. (1982) *J. Virol.*, **42**, 71-82.
Shilo,B.Z. and Weinberg,R.A. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 6789-6792.
Sotirov,N. (1981) *J. Natl. Cancer Inst.*, **66**, 1143-1147.
Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503-517.
Souza,L.H., Strommer,J.N., Hillyard,R.L., Komaromy,M.C. and Baluda, M.A. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5177-5181.
Stehelin,D., Saule,S., Roussel,M., Sergeant,A., Lagrou,C., Rommens,C. and Raes,M.B. (1980) *Cold Spring Harbor Symp. Quant. Biol.*, **44**, 1215-1223.
Thomas,P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201-5205.
Westin,E.H., Gallo,R.C., Arya,K.S., Eva,A., Souza,L.M., Baluda,M.A., Aaronson,S.A. and Wong-Staal,F. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 2194-2198.

Reprinted from Nature, Vol. 306, No. 5941, pp. 395-397, 24 November 1983
© Macmillan Journals Ltd., 1983

A putative second cell-derived oncogene of the avian leukaemia retrovirus E26

D. Leprince, A. Gegonne, J. Coll, C. de Taisne, A. Schneeberger, C. Lagrou & D. Stehelin*

Laboratoire d'Oncologie Moleculaire, INSERM Unite 186, Institut Pasteur, 15 Rue C. Guerin, 59019 Lille Cedex, France

The acute avian leukaemia retroviruses AMV and E26 both induce myeloblastosis *in vivo* and transform myeloblasts *in vitro*[1-5]. Both viruses contain the oncogene v-*myb* first described for AMV[6,7]. Unlike AMV, E26 has the additional capacity to induce erythroblastosis *in vivo* and to transform erythroblasts[4,5]. Previous analyses indicated that the genome of E26 also contained nucleotide sequences distinct from v-*myb* and unrelated to viral replicative genes[6,8,9]. Using a molecularly cloned E26 provirus, we have now identified a novel nucleotide sequence designated v-*ets* (for *E*-twenty-six specific)[10] of ~1.5 kilobase pairs (kbp) located next to v-*myb*. v-*ets* possesses all the structural characteristics of a putative new oncogene: it has a conserved cellular counterpart c-*ets* which is transcribed in some normal chicken cells as a major 7.5-kb polyadenylated RNA. Although our results now await elucidation of their biological significance, we propose that v-*ets* could be a new oncogene accounting for the additional transforming properties of E26, or potentiating the transforming properties of the v-*myb* oncogene.

A quail recombinant gene library was prepared in λ phage Charon 4A by partial *Eco*RI cleavage of DNA from an E26-transformed quail fibroblast cell clone (E26 QC1). The library was screened with a viral *myb* probe[6,11] or probes of structural viral genes (*gag*, *pol*, *env*)[1,12] or the long terminal repeat (LTR) of ASV[13]. A recombinant phage (λ E26 Q1) was shown to contain the E26 provirus (Fig. 1). A *Hin*cII–*Sal*I fragment of 1.6 kbp (Fig. 1, fragment *a*) hybridized with none of the above probes, except faintly in its 3′ end (*Pst*I–*Sal*I region) with the *env* probe (data not shown). This fragment thus contained E26-specific nucleotide sequences, denoted v-*ets* according to published nomenclature[10], that were located downstream of the E26 v-*myb* sequences.

We then examined if v-*ets* had a normal cellular counterpart. The 1.6-kbp *Hin*cII–*Sal*I fragment was used as a v-*ets* probe (probe *a*, Fig. 1). Normal chicken DNA digested with *Eco*RI hybridized with probe *a* as shown in Fig. 2, lane 2. Similar studies with quail and human DNA also exhibited several positive bands (Fig. 2, lanes 3 and 4, respectively). Therefore, v-*ets* had a cellular counterpart c-*ets* in all the vertebrate DNAs tested. This pattern was different from those described for the cellular oncogenes (c-*myb*, c-*erbA*, c-*erbB* and c-*myc*) corresponding to the other avian defective leukaemia viruses[1,14-17]. v-*ets* also failed to hybridize in non-stringent conditions to murine v-*onc* genes *abl*[18] and *ras*[19] reported possibly to induce erythroid transformation (data not shown). We thus concluded that c-*ets* represented a distinct set of conserved nucleotide sequences.

It then became important to document whether the *ets* and *myb* sequences found contiguous in E26, represented two distinct cellular domains, or belonged to a single large locus in chicken DNA. Indeed, it was reported recently that the c-*myb* locus (contained within the arrows in Fig. 3*A*, third line) in the chicken genome was considerably larger than the truncated fragment of *myb* found in AMV[14,20,21]. Thus, v-*ets* could be part of the c-*myb* locus, a larger portion of which would have been transduced by E26 than by AMV. This was not the case as the *Eco*RI restriction fragments of c-*ets* (Fig. 2, lane 2) differed in size from those contained within the c-*myb* locus (Fig. 3*A*, third line). Incidentally, the E26 v-*myb* domain appeared even smaller than in AMV (more truncated at least at the 3′ end: see lines 4 and 5 in Fig. 3*A*, and Fig. 3*B*). We conclude that E26 contains two sets of cell-derived nucleotide sequences, one (E26 v-*myb*) derived from the c-*myb* locus, colinear with the AMV v-*myb* gene transduced by AMV but more truncated at least at its 3′ end, and the other (v-*ets*) from a distinct cellular domain c-*ets* unrelated so far to any c-*onc* gene tested.

* To whom reprint requests should be addressed.

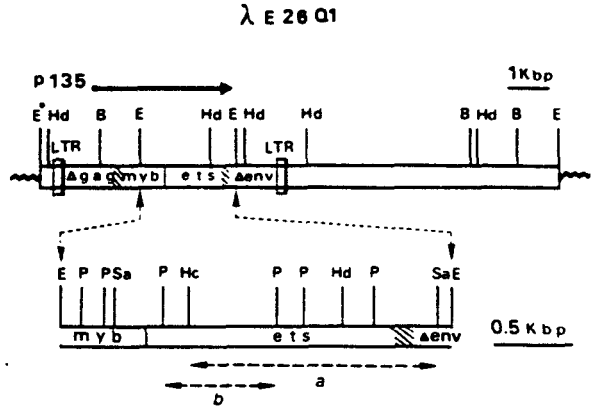


Fig. 1 Map of E26 molecularly cloned provirus. The clone λ E26 Q1, isolated as described in the text, was analysed by restriction endonuclease fragmentation followed by Southern blotting and hybridization with cDNA *myb*[6,11], and cloned probes representing viral genes (*gag*, *pol*, *env*)[6,12] or the LTR of a cloned ASV DNA (SRA-2) as described in ref. 13. Restriction endonucleases were: B, *Bam*HI; E, *Eco*RI; Hc, *Hin*cII; Hd, *Hin*dIII; P, *Pst*I; Sa, *Sal*I. The *Eco*RI 2.5-kbp fragment contains an additional 0.07-kbp *Pst*I fragment not yet mapped. E: *Eco*RI site expected from the construction of the gene library but not found in this clone. Wavy lines, Charon 4A DNA. ≈, the putative boundaries not precisely determined between *gag* and *myb* or between *env* and *ets* domains. ◂- - -▸, Fragments used as v-*ets* probes in Figs 2 and 4. ∘→, Indicates the fusion protein p135 known for this virus to initiate into the *gag* gene[8,24], which calculations indicate must extend through the *myb* sequences and terminate around the 3′ end of the *ets* sequences.

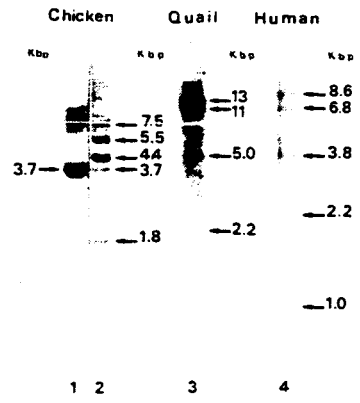


Fig. 2 v-*ets* has a cellular counterpart c-*ets*. The v-*ets* probe (probe *a*, see Fig. 1) was hybridized in non-stringent conditions to Southern blots of normal cellular DNAs restricted by *Eco*RI (lanes 2–4). DNAs used were from chick embryo fibroblasts (lanes 1, 2); Japanese quail fibroblasts (lane 3) and human placenta (lane 4). For normal chicken DNA, known to contain endogenous *env*-related sequences (as shown in lane 1 with a ^{32}P-labelled *env* probe), we checked that the v-*ets* probe, containing also some *env* sequences, yielded additional *ets*-specific bands (lane 2) not seen in lane 1.

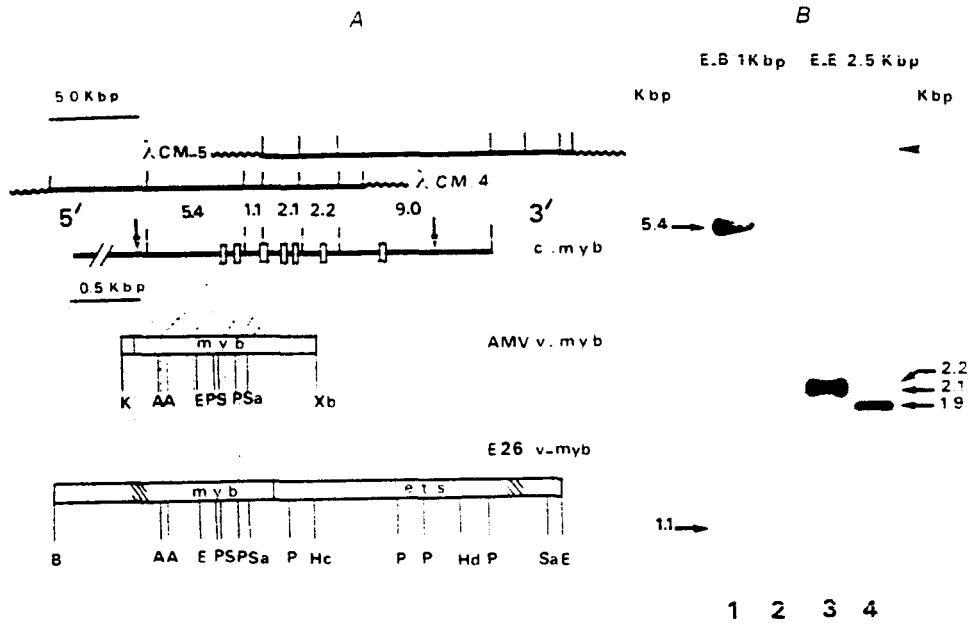


Fig. 3 E26 v-*myb* is a truncated (at least in its 3′ part) form of AMV v-*myb*. A. Organization of the complete chicken c-*myb* locus and comparison of AMV v-*myb* and E26 v-*myb* restriction maps. From top to bottom: λ CM-4 and λ CM-5 are two phages[15] isolated from a library of chicken DNA fragments inserted into the lambda phage Charon 4A. Vertical bars represent the *Eco*RI sites in the chicken DNA insert (solid line). Wavy lines, Charon 4A DNA. These two clones cover the complete c-*myb* locus, with the arrows defining the boundaries of the c-*myb* transcriptional unit[14] and the open boxes the seven regions of homology (named E1 to E7 from 5′ to 3′)[21] found between AMV v-*myb* and its cellular homologue. The restriction map of the two v-*myb* positive fragments in clone λ E26 Q1 (*Bam*HI–*Eco*RI 1 kb and *Eco*RI–*Eco*RI 2.5 kb; see Fig. 1) was determined with the following restriction endonucleases: *Ava*II (A) (first fragment) and *Hin*cII (Hc), *Hin*dIII (Hd), *Pst*I (P), *Sal*I (Sa) and *Sma*I (S) (second fragment). The 3′ boundary between E26 v-*myb* and v-*ets* was precisely defined (▽) by sequencing analyses of the *Sal*I–*Hin*cII 460 bp fragment (B. Debuire, unpublished results) and comparison with the unpublished AMV v-*myb* sequence[7,20,21]. The 3′ end (128 nucleotides) of E6 as well as all of E7 (120 nucleotides)[21] were absent in E26 v-*myb*, and replaced by *ets* sequences in the same open reading frame (B. Debuire, manuscript in preparation). In contrast, the 5′ boundary of E26 v-myb (with the *gag* gene) and the boundary between v-*ets* and *env* have been only roughly determined (▨). *B*. Hybridization of E26 v-*myb* with c-*myb* recombinant phages. The two chicken c-*myb* recombinant phages (see panel *A*) λ CM-4 (lanes 1, 3) and λ CM-5 (lanes 2, 4) were digested with *Eco*RI, separated by agarose gel electrophoresis, transferred on to nitrocellulose and hybridized with the two ^{32}P-labelled E26 v-*myb* fragments. In case of the *Bam*HI–*Eco*RI 1-kbp probe, the blot was overexposed to visualize the 1.1-kbp fragment (this fragment is difficult to detect even with AMV v-*myb* probe since it contains only a 26-bp region of homology with v-*myb*[21]). The faint 2.2–2.1-kb band visualized in λ CM-4 (lane 1) is probably nonspecific because the equivalent fragments in λ CM-5 (2.2 and 1.9 kbp) did not hybridize (lane 2). The 2.5-kbp *Eco*RI probe detected the 2.2- and 2.1-kbp c-*myb* *Eco*RI fragment in λ CM-4 (lane 3). These fragments are better separated as 2.2- and 1.9-kb fragments in λ CM-5 (lane 4). The weak hybridization of the 2.2-kbp fragment is expected, as only half of E6 (140 nucleotides) is present in this probe. The arrowhead for lane 4 indicates the position of the 9-kbp c-*myb* *Eco*RI fragment containing the 3′ distal 'exon' E7 not detected by the E26 v-*myb* probe.

Finally, we showed that c-*ets* sequences were transcribed in chicken fibroblasts, in chicken bone marrow cells or MSB1 cells (a chicken line of T lymphoblasts transformed by Marek disease virus[22]). In all cases we found a RNA species of 7.5 kb (as well as in some instances two faint 2.2- and 2.0-kb RNAs which are under investigation, Fig. 4, lanes 1–3). Indeed, analysis of the same RNA blots with a v-*myb* probe did not detect such bands but only the expected 4.0-kb c-*myb* mRNA species[14,23], as shown, for example, with MSB1 cells (Fig. 4, lane 4).

In conclusion, c-*ets* and c-*myb* represent two distinct loci in chicken cellular DNA, both transduced by a yet unknown mechanism into the E26 virus where they are translated from a genomic size viral RNA (no subgenomic RNA was demonstrated for E26, our unpublished results) as a triple hybrid (*gag–myb–ets*) protein of 135,000 molecular weight[8,24,25], a situation never seen before in acute transforming avian retroviruses. Although we cannot exclude the possibility that the complex oncogenic spectrum of E26 versus AMV could involve the *myb* gene alone (found truncated and perhaps mutated in E26), our experiments favour another alternative. We propose that the ability of E26 to cause erythroleukaemia in addition to myeloblastosis could result from the presence of a second oncogene, v-*ets*, within the viral genome. v-*ets* shares the properties found so far for other oncogenes[10] (viral sequences translated into a protein, that are of cellular origin, conserved with evolution and are expressed as discrete mRNA species at some stages in normal cells). So far, all specific sequences with these properties found in retroviruses turned out to be oncogenes. That two distinct oncogenes are found co-existing in a same virus may not be coincidental as we recently described another retrovirus with two cell-derived specific sequences[26]. A first such example, avian erythroblastosis virus (AEV) was shown to contain two *onc* genes (*erbA*, *erbB*) transcribed as two distinct mRNAs[16] and translated into different proteins[27] (75,000 and 68,000 molecular weight, respectively), where *erbA* potentiates the transforming activity of *erbB* within the same erythroid lineage[28]. In this respect in E26, v-*ets* could potentiate the transforming activity of E26 v-*myb*, as E26-transformed myeloblasts seem to be more tightly blocked in their differentiation than AMV-transformed myeloblasts[4]. In addition, the two specific genes described here for E26 may confer on this virus the property to transform cells already committed into both

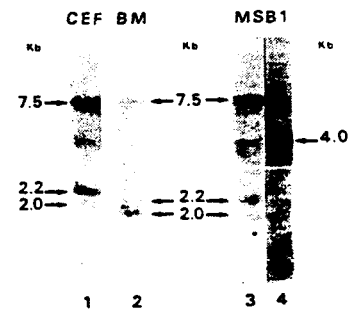


Fig. 4 Size of c-*ets* mRNA in chicken cells. Northern blots were prepared with polyadenylated mRNA and hybridized in stringent conditions (0.45 M NaCl, 41 °C, 50% formamide) with a v-*ets*-specific probe (the 720-bp *Pst*I fragment, probe *b* in Fig. 1) (lanes 1–3). RNAs tested were: CEF, chicken embryo fibroblasts (lane 1); BM, bone-marrow (lane 2); MSB1, a cell line of T lymphoblasts transformed by Marek disease virus (MDV) (lane 3) that also expresses c-*myb* RNA at an elevated level[22,23] as shown in lane 4 with a *myb* probe.

myeloid or erythroid lineages[4] or uncommitted precursor cells still able to differentiate into these two lineages[5]. For further analyses along these lines, we are currently attempting site-specific mutagenesis in the E26 *myb* or *ets* sequences. Clearly, several experimental systems involving complementation of or synergy between more than one oncogene[29–31] are emerging, which may lead to a better understanding of multi-step carcinogenesis.

We thank Drs H. Beug and T. Graf for providing the E26 QC1 cells; Dr B. Debuire for providing data before publication; Drs C. Dissous, M. Righi and S. Saule for helpful advice and discussion; Dr D. Sheiness for comments on this manuscript and N. Devassine for patient typing. This work was supported by the Institut National de la Santé et Recherche Médicale (U186), Centre National de la Recherche Scientifique (AI 033695), Association pour la Recherche sur le Cancer (6124) and the Pasteur Institute of Lille.

Received 8 July; accepted 21 September 1983.

1. Graf, T. & Stehelin, D. *Biochim. biophys. Acta* **651**, 245–271 (1982).
2. Beard, J. W. *Adv. Cancer Res.* **7**, 1–127 (1963).
3. Ivanov, X., Mladenov, Z., Nedyalkov, S. & Todorov, T. G. *Bull. Inst. Path. comp. Anim. Domest.* **9**, 5–36 (1962).
4. Radke, K., Beug, H., Kornfeld, S. & Graf, T. *Cell* **31**, 643–653 (1982).
5. Moscovici, M. G. *et al. Virology* **129**, 65–78 (1983).
6. Roussel, M. *et al. Nature* **281**, 452–455 (1979).
7. Souza, L. M., Strommer, J. N., Hilyard, R. L., Komaromy, M. C. & Baluda, M. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5177–5181 (1980).
8. Bister, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 3677–3681 (1982).
9. Stehelin, D. *et al. Cold Spring Harb. Symp. quant. Biol.* **44**, 1215–1223 (1980).
10. Coffin, J. M. *et al. J. Virol.* **40**, 953–957 (1981).
11. Gonda, T. J. *et al. Cell* **23**, 279–290 (1981).
12. Saule, S. *et al. J. Virol.* **42**, 71–82 (1982).
13. Delorbe, W. J., Luciw, P. A., Goodman, H. M., Varmus, H. E. & Bishop, J. M. *J. Virol.* **36**, 50–61 (1980).
14. Gonda, T. J. & Bishop, J. M. *J. Virol.* **46**, 212–220 (1983).
15. Leprince, D. *et al. EMBO J.* **2**, 1073–1078 (1983).
16. Vennstrom, B. & Bishop, J. M. *Cell* **28**, 135–143 (1982).
17. Sheiness, D., Hughes, S. H., Varmus, H. E., Stubblefield, E. & Bishop, J. M. *Virology* **105**, 415–424 (1980).
18. Wang, J. Y. J. & Baltimore, D. *Molec. cell. Biol.* **3**, 773–779 (1983).
19. Ellis, R. W. *et al. Nature* **292**, 506–511 (1981).
20. Rushlow, K. E. *et al. Science* **216**, 1421–1423 (1982).
21. Klempnauer, K. H., Gonda, T. J. & Bishop, J. M. *Cell* **31**, 453–463 (1982).
22. Nazerian, K. & Sharma, J. M. *J. natn. Cancer Inst.* **54**, 277–279 (1979).
23. Gonda, T. J., Sheiness, D. K. & Bishop, J. M. *Molec. cell. Biol.* **2**, 617–624 (1982).
24. Beug, H., Hayman, M. J. & Graf, T. *EMBO J.* **1**, 1069–1073 (1982).
25. Boyle, W. J., Lipsick, J. S., Reddy, P. E. & Baluda, M. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2834–2838 (1983).
26. Coll, J. *et al. EMBO J.* **2** (in the press).
27. Hayman, M. J. *et al. Cell* **32**, 579–588 (1983).
28. Frykberg, L. *et al. Cell* **32**, 227–238 (1983).
29. Land, H., Parada, L. F. & Weinberg, R. A. *Nature* **304**, 596–602 (1983).
30. Ruley, H. E. *Nature* **304**, 602–606 (1983).
31. Newbold, R. F. & Overell, R. W. *Nature* **304**, 648–651 (1983).

BIBLIOTHEQUE INTERUNIVERSITAIRE
SECTION DE SCIENCES
LILLE