

## Регіональна специфічність вірусної інтеграції

А. В. Риндич, С. В. Зубак, Л. О. Циба, Н. Гулей<sup>1</sup>, З. В. Лазуркевич,  
Дж. Бернарді<sup>1</sup>

Інститут молекулярної біології та генетики ІАН України  
252143, Київ, вул. Заболотного, 150

<sup>1</sup>Інститут Жака Моно  
75005, Париж, пл. Жусьє, 2, 43

*Інтеграція ретровірусів у геном хазяїна до недавнього часу вважалася випадковою і на сьогодні не існує ясності, як впливають виявлені в місцях встановлених локальних ефектів в місцях інтеграції локальні ефекти на процес інтеграції в цілий геном. Використання композиційного підходу дало можливість встановити, що інтеграція ретровірусів відбувається у певні райони геному, які тотожні вірусним послідовностям за складом основ. Цю регіональну специфічність інтеграції доведено: 1) нашими експериментальними даними з локалізації ретровірусних послідовностей у різних композиційних ділянках геному; 2) результатами з інших лабораторій стосовно локалізації ретровірусних послідовностей у «відкритих» ділянках хроматину або/чи біля СrС-«острівців»; 3) нашим композиційним аналізом генів, розташованих по сусідству з інтегрованими вірусними послідовностями. Такий висновок має значення як для розуміння композиційної еволюції ретровірусних геномів, так і для генної терапії.*

**Вступ.** Інтеграція вірусних геномів як провірусів у геном клітини-хазяїна є важливим етапом в життєвому циклі ретровірусів [1, 2]. Дійсно, реплікація ретровірусів залежить від інтеграції [3]. Крім того, різні патогенні ефекти ретровірусів також пов'язані з інтеграцією. Так, інтеграція може призвести до інактивації генів клітини шляхом їхніх розривів або ж до активації генів за допомогою вірусних промоторів [4—7]. Інтегровані ретровірусні послідовності здатні впливати на появу хромосомних перебудов, що викликають інактивацію критичних генів, таких як гени-супресори пухлин [6, 8—11], модифікацію клітинних генів та формування нових трансформуючих вірусів внаслідок «захвату» деяких клітинних генів вірусним геномом [12—14]. Такі події призводять до трансформації клітин і розвитку пухлин. Все це вказує на важливість феномену ретровірусної інтеграції. Остання взагалі є досить специфічним випадком інтеграції чужорідної ДНК у геном клітини-хазяїна, і з цієї точки зору її вивчення може бути надзвичайно важливим для генної терапії [15, 16].

Для самого вірусу інтеграція є сайт-специфічним процесом, який залежить від довгих кінцевих повторів (LTR) [17, 18]. Відносно геному хазяїна довго існувала думка, що інтеграція у клітинний геном носить випадковий характер [19]. Це базувалося на даних гібридаційних аналізів ретровірусних проб з ДНК інфікованих клітин, які свідчили, що рестрикційні карти послідовностей, що оточують інтегровані послідовності, були різними для різних провірусів, та на даних з секвенування коротких послідовностей клітинних ДНК, фланкуючих інтегровані провіруси [20—23]. Ці виявлені закономірності, які вказували лише на те, що віруси не інтегрують в унікальні послідовності і що кількість сайтів інтеграції не обмежена, були використані як доказ того, що вірусна інтеграція у геном хазяїна має випадковий характер.

Ідею випадкової інтеграції недавно знову було відроджено на основі інших експериментальних підходів [24], що буде прокоментовано далі, і знов-таки взято під сумнів, виходячи з отриманих експериментальних даних, які дискутуються в цьому огляді.

Розвиток методів композиційного фракціонування (за складом основ) [25—27] призвів не лише

до відкриття того, що геном ссавців є мозаїкою ізохор — довгих композиційно гомогенних ДНК-сегментів, об'єднаних у невелике число родин, що характеризуються різним складом основ, а також до можливості локалізувати в композиційних фракціях ДНК будь-яку нуклеотидну послідовність (в тому числі і провірусну), аби для неї був виявлений відповідний зонд.

Вже на перших порах використання цього методу [28] було показано, що ретровірусна інтеграція не є випадковою. На прикладі інтеграції вірусу лейкемії бика (BLV) було продемонстровано, що вірусна інтеграція є компартменталізованою, тобто відбувається в певних районах геному хазяїна, та «ізолікнічною», бо має місце в таких послідовностях геному хазяїна, які композиційно тотожні вірусним послідовностям. Ці висновки були підтверджені на інших вірусних системах та іншими підходами, про які йтиметься в цьому огляді.

**Організація геному хазяїна.** Геном теплокровних являє собою мозаїку ізохор — довгих сегментів ДНК від 200 до 6000 тис. п. н. [29]. Ізохори досить гомогенні за складом основ (вище за 3 тис. п. н.) і належать до невеликого числа родин, які характеризуються різним рівнем GC [30].

Великі фрагменти стандартного виділення ДНК (біля 100 тис. п. н.), що утворюються в результаті неспецифічних фізичних і хімічних розривів під час екстракції, відображають склад основ ізохор, з яких вони походять (рис. 1). Такі фрагменти ДНК можна фракціонувати препаративним центрифугуванням у  $\text{Cs}_2\text{SO}_4$  у присут-

ності сиквенспецифічних ДНК-лігандів, а рівень GC кожної фракції може бути встановлений аналітичним центрифугуванням в  $\text{CsCl}$ .

Композиційний розподіл великих фрагментів ДНК геному людини, типового для ссавців [31], характеризується наявністю п'яти родин фрагментів, одержаних з відповідних ізохор (див. рис. 1). У геномі людини GC-бідні родини ізохор L1 і L2 (разом позначені як L) становлять 63 % геному, GC-багаті H1 і H2 родини — біля 24 і 7,5 % геному та дуже GC-багата H3 родина — близько 5 % геному, залишок припадає на сателітну та рибосомну ДНК [32].

Розподіли ізохор, кодуєчих послідовностей, їхньої третьої позиції кодону та інтронів за складом основ формують «геномні фенотипи» [33], які відрізняються не лише між теплокровними і холоднокровними хребетними, а й між ссавцями і птахами і навіть серед ссавців [30].

Локалізація генів у фракціях ДНК виявила існування композиційної кореляції між кодуєчими послідовностями і ізохорами, що їх несуть [34, 35]. GC-багаті і GC-бідні кодуєчі послідовності локалізовані в GC-багатих і GC-бідних ізохорах відповідно. У свою чергу це дозволило проаналізувати розподіл генів, який виявився недовільним [32, 36, 37]. Концентрація генів паралельна GC-рівню (рис. 2): будучи найнижчою в GC-бідних ізохорах, вона зростає і досягає максимуму в GC-багатих ізохорах.

Дослідження хромосомної локалізації ізохор людини показало, що GC- і ген-найбагатші ізохори розташовані в двох підгрупах R (reverse) смуг — T

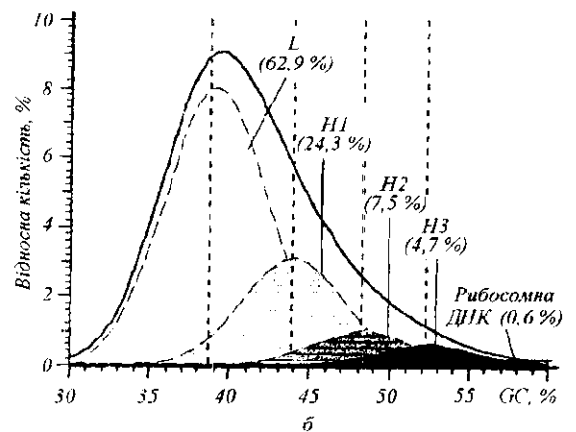
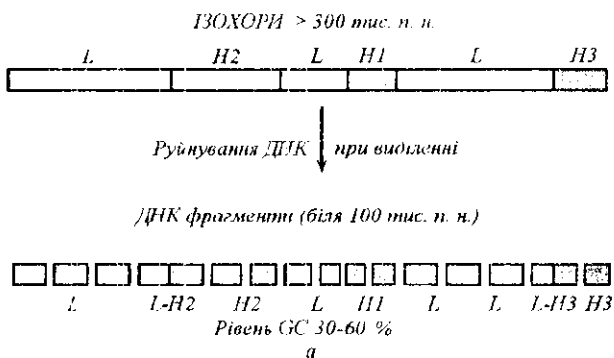


Рис. 1. Схема ізохорної організації геному людини (пояснення в тексті) (а) та декомпозиція  $\text{CsCl}$  профілю ДНК людини, яка представляє чотири основні компоненти ДНК, що відображають родини ізохор L (L1 + L2), H1, H2, H3 та рибосомну ДНК. Показано відносну кількість усіх основних компонентів ДНК у %. Сателітна ДНК не представлена (б)

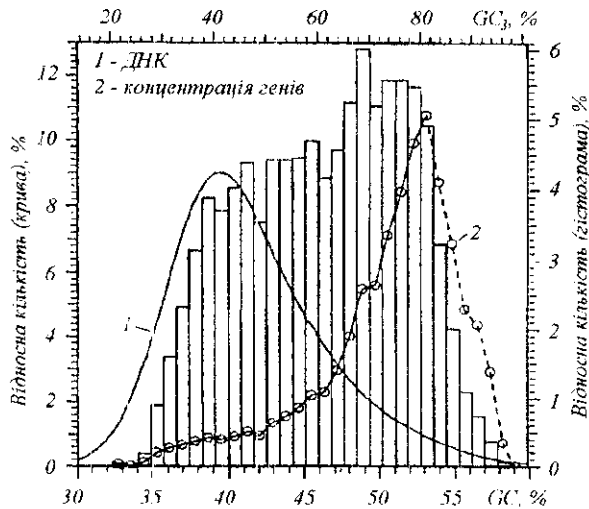


Рис. 2. Профіль концентрації генів у геномі людини. Профіль одержано шляхом ділення відносної кількості генів у кожному 2,5 % GC-інтервалі гістограми розподілу генів на відносну кількість ДНК, вираховану з CsCl профілю (суцільна лінія)

і  $T'$  (або  $H3^+$  і  $H3'$ ), перша з яких містить більше  $H3$  ізохор, інша підгрупа R-смуг ( $H3'$ ) їх не має [37–39].

Ізохори з різною концентрацією генів відповідають різним структурам хроматину: «відкрита» — в GC- і ген-багатих ізохорах і «закрита» — в GC- і ген-бідних ізохорах. GC-багаті ізохори характеризуються ранньою реплікацією у клітинному циклі (у такому порядку:  $H3^+$ ,  $H3'$ ,  $H3'$  смуги). Висока транскрипційна активність притаманна GC-багатим регіонам, які містять більшість «house-keeping» генів, асоційованих з CpG-«острівцями». GC-бідні ізохори з низькою транскрипційною активністю мають низьку концентрацію генів, що асоційовані з TATA-бокс-промоторами і експресуються тканиноспецифічно. Нарешті, GC-багаті ізохори відзначаються високою рекомбінантною активністю на відміну від GC-бідних. Найбільш транскрипційно активні компартменти є також найбільш схильними до рекомбінації, оскільки вони локалізовані в  $T$  і  $T'$  ( $H3^+$ ,  $H3'$ ) хромосомних смугах [38, 39], де, як відомо, має місце найбільша кількість рекомбінацій. Відкрита структура хроматину характеризується доступністю для ДНКаз, більшою відстанню між нуклеосомами, відсутністю гістонів  $H1$  і ацетилюванням гістонів  $H3$  і  $H4$  [40].

Інтеграція провірусних послідовностей у геном хазяїна є компартменталізованою та «ізопікнічною». Композиційна локалізація інтегрованих послідовностей ретровірусів досліджувалася у випадку трьох GC-багатих вірусів — вірусу лейкемії бика (BLV) [28, 41], вірусу T-клітинної лейкемії людини (HTLV-I) [42], які не містять онкогенів, і вірусу саркоми Рауса (RSV) [43], який має онкоген *src*, та одного GC-бідного вірусу — вірусу пухлин молочної залози мишей (MMTV) [44]. Було встановлено деякі спільні риси провірусної інтеграції — компартменталізованість, «ізопікнічність» і кореляцію між транскрипцією провірусів та їхньою локалізацією у геномі клітини, що буде проілюстровано в огляді результатів, одержаних для різних вірусних систем в цьому та інших розділах.

**BLV.** Компартменталізовану інтеграцію провірусних послідовностей BLV (рис. 3) було продемонстровано в геномі інфікованих клітин трьох тварин (одна з персистентним лімфоцитозом і дві в пухлинному стані). В усіх шести досліджуваних зразках (з лейкоцитів, пухлин лімфатичних вузлів і селезінки) цих трьох тварин послідовності BLV виявилися інтегрованими лише у фрагменти ДНК з високим GC-вмістом (48 %), які представляють менше ніж 10 % геному бика. BLV провірус — GC-багатий (54 %), і тому така інтеграція в GC-багаті компартменти була визначена як «ізопікнічна».

**RSV.** У цьому випадку вивчали ретровірусну інтеграцію в геноми шести охарактеризованих клонів клітин пухлин хом'яка. З 13 послідовностей RSV (GC = 54 %) 12 були виявлені в ДНК-фрагментах композиційних фракцій, що відповідають GC-найбагатшим 20 % геному хом'яка з піком на 50 % GC (див. рис. 3). Більш детально, геном GC-багатого (54 %) RSV з чотирьох клітинних ліній, одержаних від різних пухлин, що містили одну—дві повні чи дефектні копії провірусу, був знайдений у фракціях з GC-рівнем 50—53 % [43]. У випадку клітинної лінії H19, що несе провірус із структурою LTR-*vsr*-LTR (GC-рівень 56,5 %) [45], провірус був локалізований в ДНК-фрагментах з нижчим GC-вмістом (49 %). Проте секвенування фланкуючих послідовностей провірусу показало вищий GC рівень (58 %), ніж у послідовностей, розташованих далі від сайту інтеграції [46]. У випадку ж, коли клітинна лінія несе багато провірусних копій як повних, так і дефектних (клітинні лінії H9 і H42), повні копії були представлені у фракціях від 47 до 54 %, тоді як дефектні копії також були знайдені в цих і в GC-бідніших фракціях.

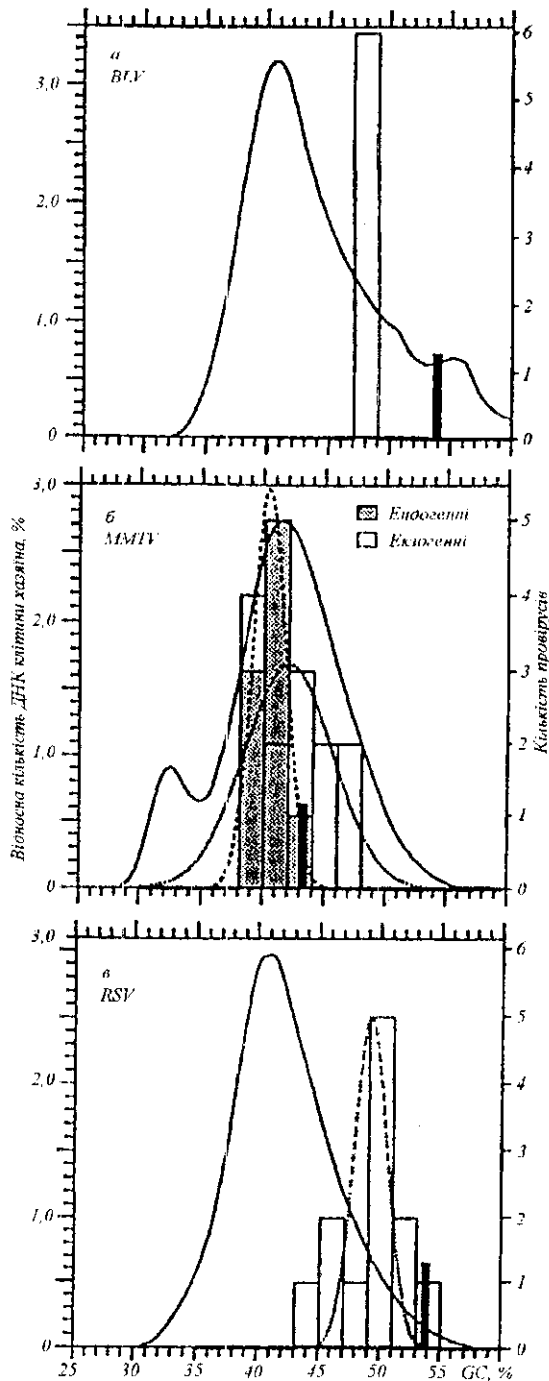


Рис. 3. Розподіл провірусів у геномах бика [28] (а) миші [44] (б) та хом'яка [43] (в).  $C_0C_1$  профіль ДНК клітини хазяїна представлений суцільною лінією, локалізація провірусних послідовностей — стовпчиками. Чорні стовпчики відображають GC-рівень вірусних генів. Пунктирні лінії показують найбільшу відповідність нормальних функцій гістограм

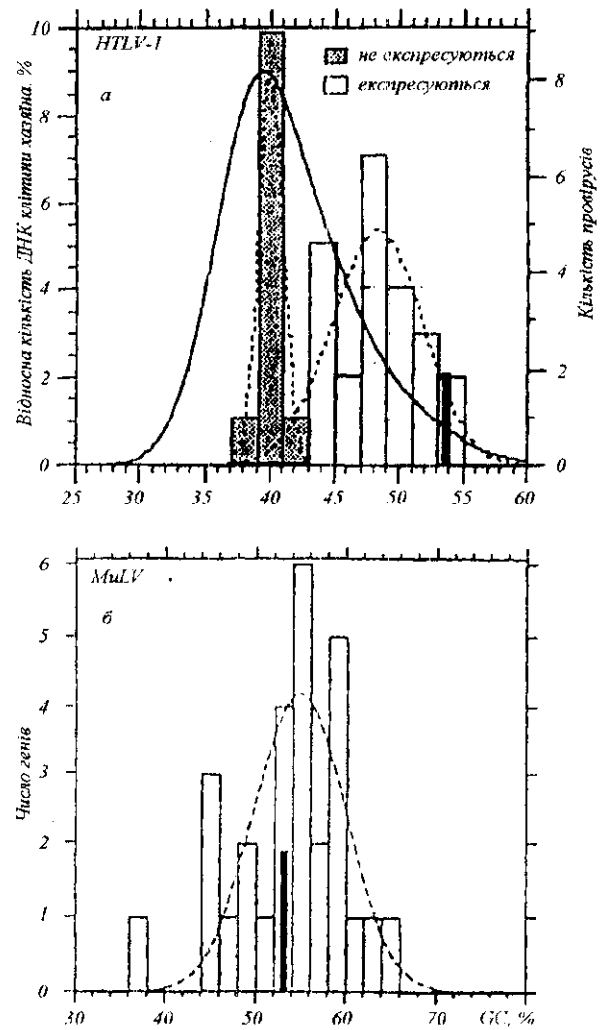


Рис. 4. Розподіл інтегрованих послідовностей вірусу HTLV-I у геномі людини (а) і розподіл генів, локалізованих біля MuLV провірусів (б)

**HTLV-I.** Локалізацію цього провірусу (рис. 4) досліджували в п'яти клітинних лініях, які містили 22 інтегрованих повних і дефектних провіруси, та в семи Т-клітинних клонах, одержаних від пацієнтів з TSP/HAM (тропічний спастичний парепрез/HTLV-I-асоційована мієлопатія), у яких хазяйські геноми містили одну—три інтегровані провірусні копії в кожному клоні, всього — 18 провірусів [42]. В усіх випадках HTLV-I послі-

довності знайдені в GC-багатших компартментах, що становлять 60 % геному людини.

*MMTV.* GC-бідні MMTV послідовності (43,3 %) були локалізовані в ДНК печінки п'яти інбредних ліній мишей і виділені з GR клітин, одержаних з первинної імплантації пухлини молочної залози [47]. Двадцять типів ендогенних MMTV послідовностей було охарактеризовано за рестрикційними особливостями і хромосомною локалізацією (всього 36 провірусів). Екзогенні послідовності з сильним (всього 13) і слабким, що відповідав вірусній інтеграції в мінорній популяції клітин (всього 6), гібридаційними сигналами, було також локалізовано. Ендогенні провіруси виявили в GC-бідних 43 % геному хазяїна з піком на 41 % GC, екзогенні показали ширше розподілення, будучи представленими в GC-бідних 60 % геному клітин з піком на 42 % GC (див. рис. 3).

Таким чином, ретровірусні послідовності (три С-типу і одна D-типу) були локалізовані в геномах чотирьох видів ссавців, два з яких (людина і бик) мають типові особливості структури геному ссавців, інші два (хом'як і миша) характеризуються відсутністю H3 родини ізохор [31, 47]. Показано, що ретровірусна інтеграція має місце лише в тих компартментах геному, які композиційно відповідають вірусним послідовностям. У кожному з досліджуваних випадків провіруси не були знайдені в сателітній чи рибосомній ДНК.

Подібні результати було одержано для інтеграції ДНК-вмісного GC-багатого вірусу гепатиту В (HBV) [48]. Вісім з дев'яти вірусних послідовностей, інтегрованих в hepatoma Alexander cell line, локалізувалися в ДНК-сегментах, що відповідали GC-найбагатшим 4 % геному людини. Аналогічні спостереження є й для інтеграції GC-багатого рекомбінанта SV40-аденовірус 5 в H3 родину ізохор [49].

Кореляція між ізохорною локалізацією інтегрованих ретровірусних послідовностей і їхньою транскрипцією. Результати композиційного розподілу HBV послідовностей [48] дали першу інформацію про можливу кореляцію між «ізопікнічною» локалізацією провірусів та їхньою транскрипцією. Дійсно, вісім HBV-послідовностей, інтегрованих в GC-найбагатші регіони, виявилися експресованими, тоді як одна, локалізована в GC-біднішому регіоні, — ні. Подібне було показано і для RSV: всь транскрипційно активні послідовності локалізувалися в GC-найбагатших ізохорах, а провіруси, експресія яких не була встановлена, — в ізохорах з нижчим GC-вмістом.

Детальніший аналіз було проведено у випадку HTLV-I інтеграції [42]. Транскрипційно активні

HTLV-I провіруси знайдені в GC-багатших регіонах, що відповідають H2 і H3 родинам ізохор геному людини, тоді як транскрипційно неактивні — в GC-багатих регіонах, що належать до H1 родини ізохор (і, можливо, частково до L ізохор) (див. рис. 4).

У випадку MMTV в GR клітинній лінії транскрибовані MTV-2 ендогенні послідовності знаходяться в GC-бідних фракціях (37—38 %). До того ж більшість ліній мишей містять як мінімум одну копію ендогенного провірусу [50], що експресується не лише в молочної залозі під час лактації [51], а й в деяких інших органах з нижчим рівнем експресії [52].

Інші докази регіональної інтеграції ретровірусних послідовностей: аналіз сайтів інтеграції у «відкритому» хроматині та/чи біля CpG-«острівців». Ряд спостережень дозволив висловити припущення, що ретровірусна інтеграція корелює з рівнем транскрипції, рекомбінації і ступенем доступності ДНК в хроматині [53—57]. В ряді незалежних дослідів було картовано інтеграцію вірусів плазминої лейкемії (ALV), та вірусу Молоні мишей (MoMuLV), біля ДНК-а I-надчутливих сайтів [53, 58, 59]. Як правило, ці сайти корелюють з експресією генів і локалізуються в районах з «відкритою» структурою хроматину [60]. Одержані результати стимулювали аналіз транскрипції клітинних послідовностей, фланкуючих інтегровані провіруси. MoMuLV був локалізований поблизу транскрипційно активних районів та/чи біля CpG-«острівців» [55, 56]. RSV провірус, що транскрибується у клітинах гризунів, був знайдений поблизу 3'-кінця CpG-«острівців», тоді як значна кількість RSV нетрансльованих послідовностей була локалізована у випадкових сайтах [57].

Ці дані, отримані для GC-багатих послідовностей ALV, MoMuLV та RSV, добре узгоджуються з викладеним вище. Справді, H2 та H3 ізохори є ізохорами, що мають найвищий рівень CpG-«острівців» [61, 62]. Оскільки CpG-«острівці», як правило, асоціюються з генами, життєво важливими для клітини, то ці композиційні компартменти, найвірогідніше, мають найвищу концентрацію генів, які конститутивно експресуються в усіх клітинах. Підсумовуючи, можна зазначити, що відомості про переважну інтеграцію GC-багатих вірусів в райони, для яких характерна наявність численних CpG-«острівців» та/чи «відкритої» структури хроматину, що асоціюється з GC-багатими районами геному ссавців [30], підтримують висновки про особливості інтеграції, виявлені для GC-багатих ретровірусів на основі експериментів із залученням композиційного фракціонування.

Інші докази регіональної інтеграції вірусних послідовностей: аналіз сайтів інтеграції біля генів. Аналіз сайтів інтеграції провірусів можна також здійснити на основі композиції генів клітини-хазяїна, локалізованих поблизу інтегрованих вірусних послідовностей.

Віруси пташиної чи мишачої лейкемії (ALV, MuLV), які індукують гематопоетичні пухлини за посередництвом комплексного непрямого механізму, що включає інсерційну активацію клітинних протоонкогенів, є у зв'язку з цим найкращими моделями для композиційного аналізу генів, локалізованих біля сайтів інтеграції. Однак використання даних з інсерційного мутагенезу, завдяки якому можна виявляти гени у ділянках, розташованих поблизу інтегрованих провірусів, повинне враховувати можливість генетичної нестабільності [63]. У випадку MuLV різні його ізоляти викликають пухлини з широкою шкалою варіюючих фенотипів. Онкогенна інсерція MuLV, що призводить до клональної експансії, спотерігалася в різні хромосоми, але часто сайти інтеграції були картовані близько один від одного на тих самих хромосомах [64—66] і були пов'язані з активацією певних онкогенів, зокрема, *c-myc* — у багатьох випадках протомоноцитної лейкемії [67, 68], *c-myc* — у Т-клітинних чи В-клітинних лімфомах [5], *Evi-1* — в мієлоїдній неоплазмі [69—71].

Композиційний аналіз генів клітини, локалізованих поблизу інтегрованого MuLV, показав, що вони розподілені в межах 44—65 % за GC-складом (див. рис. 4, таблиця), в середньому це 54,9 % GC, що наближається до GC-складу MuLV (53,1 %). Останнє вказує на те, що у випадку інсерційного мутагенезу є можливість прослідкувати «ізопікнічну» локалізацію інтегрованих провірусних послідовностей. Випадкова локалізація MuLV біля *c-Ki-ras* (37 % GC) пояснюється розташуванням у прикордонній зоні двох класів ізохор.

Подібні спостереження, по суті, мають місце і для вірусів лейкозів птахів (ALV). Для ALV-індукованих пухлин різними ALV показано інсерційну активацію протоонкогена *c-myc* [20—22, 72, 73]. Серед ALV вірусів ті ж самі вірусні послідовності можуть викликати різноманітні неоплазії шляхом активації різних онкогенів. Так, наприклад, RAV-1 викликає нефробластому, В-клітинну лімфому чи еритробластоз через активацію *c-fos*, *c-myc* чи *c-erbB* відповідно [20, 21, 74—77]. Навіть ендегенні ALV віруси (RAV-0) беруть участь в активації протоонкогена *c-bravo* в В-клітинних лімфомах. Дослідження сайтів інтеграції в цих неоплазіях, які виникли в результаті інсерції GC-багатих ALV, показали, що інтеграція відбувається біля

онкогенів з високим GC-рівнем. Навіть якщо ці інтеграційні сайти відображають властивості ALV селектованого фенотипу, наведені дані демонструють важливість інтеграції у такі геномні компартменти для створення пухлинного фенотипу, а також можливість того, що згадані мішені інтеграції є не лише специфічними для даного фенотипу, але і є властивістю ALV інтеграції. Наведені коментарі також стосуються сайтів інтеграції у MuLV інфікованих клітинах.

Ендегенні MMTV переносяться вертикально від матері до плоду і здатні викликати головним чином карциноми молочної залози [78, 83]. Сайти MMTV інтеграції було знайдено в різних хромосомах. Показано активацію цим вірусом восьми *int* генів, важливих для регуляції розвитку ембріонів мишей, для багатьох незалежних пухлин мишей: *int-1/Wnt-1* [78], *int-2/Fgf-3* [80], *int-3/Noth-4* [81], *int-4/Wnt-3* [82], *Fgf-8* [84], *int-6* [85] і *Wnt-10b* [86], *Fgf-4/hstlk-FGF* [87]. Більшість інтеграцій провірусів зустрічається перед або після *int* генів, часом на відстані понад 20 тис. п. н. [87]. Це призводить до порушення регуляції їхньої транскрипції. Оскільки *int* гени GC-багаті, то виникло припущення, що вони локалізуються в GC-багатих ізохорах, тоді як очікувалося, що сайти інтеграції MMTV повинні бути в GC-бідних ізохорах. Проте інсерційні мутації виявлялися також всередині 3'-регіону, що не транслюється [88], в інтронах, навіть у кодуючих районах гена [85, 89—91]. Всі ці пошкоджені гени — GC-багаті (54—65 %), за винятком одного — *int-6* (41 %). Така «неізопікнічна» локалізація сайтів інтеграції MMTV в GC-багаті гени підтверджує вищевикладену тезу, що інтеграція спочатку може відбуватися не тільки в регіони з однаковим складом основ. Ці випадкові інтеграції у локуси з високим GC-вмістом, що несуть GC-багаті *Wnt/int/fgf*-гени, надають ростову перевагу індивідуальним клітинам та їхній клональній експансії. Дійсно, це є випадок, у якому специфічність сайтів інтеграції може обумовлюватися селекцією, викликаною перевагою в рості інфікованих клітин.

Оскільки дослідження інтеграції GC-бідного MMTV для різних варіантів, наведених вище, виявило «ізопікнічну» картину інтеграції, інсерція MMTV біля GC-багатих генів *int/Wnt/fgf* у пухлинах молочної залози, ймовірно, є унікальною властивістю, пов'язаною з патогенезом. Це підтверджується й тим, що гени (за виключенням *hst*), що асоційовані з MMTV, звичайно не активуються в інших типах пухлин.

Інші фактори, що беруть участь в інтеграції ретровірусних послідовностей. Дослідження, при-

## Композиційний склад генів, розташованих поблизу інтегрованих провірусів MuLV, ALV, MMTV

Вірус Ген/Регіон	GC, %	Посилання	Вірус Ген/Регіон	GC, %	Посилання
MuLV, 53 % GC			MuLV, 53 % GC		
<i>Gfi-1</i>	64,90	[132]	<i>Nfr-2</i>	47,90	[123]
<i>N-myc</i>	63,24	[104]	<i>Mtvi-1/pvt-1</i>	45,98	[124]
<i>collagen type 1</i>	60,18	[105]	<i>Evi-1</i>	45,50	[69, 125]
<i>pim-1</i>	58,81	[106]	<i>c-Ki-ras</i>	36,86	[126]
<i>spi-1</i>	58,80	[107]	ALV, 53 % GC		
<i>Gfi-2</i>	58,10	[109]	<i>c-fos</i>	70,56	[76]
<i>c-myc</i>	58,03	[110]	<i>c-myc</i>	62,99	[20, 127]
<i>bmi-1 (mel-18)</i>	56,85	[111]	<i>c-myb</i>	57,25	[128]
<i>vin-1</i>	56,75	[112]	<i>c-Ha-ras</i>	53,16	[129]
<i>p-53</i>	55,75	[113]	<i>c-rel</i>	46,7	[77]
<i>lck (p-56 tck)</i>	55,42	[114]	MMTV, 43,32 % GC		
<i>Dsi-1</i>	54,50	[115]	<i>Wnt-1/int-1</i>	65,32	[78]
<i>c-fms</i>	54,18	[116]	<i>Fgf-8</i>	65,3	[84]
<i>c-Ha-ras</i>	53,68	[117]	<i>int-3/Noth--4</i>	65 (61)	[81]
<i>fli-1</i>	53,64	[118]	<i>hst-1 (Kfgf, fgf-4), FGF</i>	64,53	[79]
<i>Tpl-1/ets-1</i>	52,90	[119]	<i>int-2/jgf-3</i>	63	[130]
<i>Tpl-2</i>	50,50	[120]	<i>Wnt-10b</i>	61	[86]
<i>GM-CSF</i>	49,65	[121]	<i>int-4/Wnt-3</i>	60,86	[82]
<i>fli-2</i>	48,29	[122]	<i>int-6</i>	41	[133]

св'язані іншим аспектам ретровірусної інтеграції і прямо не пов'язані з основним питанням, що розглядається у даній публікації, а саме: з регіональною інтеграцією ретровірусних геномів, будуть лише коротко згадані.

Цікавим є спостереження, що інтеграція HIV відбувається з більшою частотою в нуклеосомну ДНК, ніж у вільну [92—94] і що сайти інтеграції кластеризовані в найбільш деформовані регіони нуклеосомного кора.

Детальний аналіз *in vitro* інтеграційних сайтів у великих виборках інтеграційних подій дозволяє показати, що специфічність інтеграції пов'язана з особливими структурами ДНК [93, 95]. Структурні елементи ДНК, які пов'язані з її напруженістю, є сайтами ретровірусної інтеграції. Ці елементи задіяні в багатьох біохімічних процесах [96] і можуть бути асоційованими з відповідними білками. Показано, що реакції інтеграції *in vitro* можуть блокуватися білками-репресорами транскрипції, які зв'язуються із специфічними ДНК-сайтами, можливо,

заважаючи доступу інтеграційного апарату [92]. До того ж інші клітинні білки, що деформують ДНК під час зв'язування, можуть сприяти інтеграції [97, 98]. Таким чином, важливою є загальна конформація ДНК-білкового комплексу, а не лише напруженість ДНК, що забезпечує доступність регіонів ДНК для взаємодії з інтегразою.

Вивчення селекції сайтів інтеграції *in vitro* також виявило, що інтеграція в «голу» ДНК не є випадковою, при цьому передбачалася деяка перевага особливих нуклеотидних послідовностей [93]. Перевага для А/Т-нуклеотидів спостерігалася у коротких регіонах, оточуючих сайти інтеграції ряду GC-багатих ретровірусів, таких як MuLV [93], BLV і HTLV-1 [99], SNV [100].

У відповідності з цими даними на вибір сайту інтеграції можуть впливати послідовності та конформація ДНК клітини і білки хазяїна [96]. Більш того, специфічна взаємодія між ДНК-зв'язуючими білками хазяїна і інтегразою може спрямовувати інтеграцію в певні регіони [131].

**Висновки.** У представленому огляді продемонстровано, що геном теплокровних хребетних є не лише структурно мозаїчним, являючи собою набір компартментів чи ізохор, але й функціонально мозаїчним. Дійсно, 1) концентрація генів різночисна — будучи дуже низькою в GC-бідних ізохорах, вона значно зростає у GC-багатих ізохорах; 2) райони геному, багаті генами, мають вищу транскрипційну і рекомбінантну активність у порівнянні з GC-бідними районами. В свою чергу, це пов'язано з різницею в структурі хроматину: «відкритою» в першому випадку і «закритою» — в другому. У зв'язку з цим виникає питання відносно розподілу інтегрованих послідовностей у мозаїчному геномі хазяїна.

Матеріал, наведений в огляді, можна розділити на три великі групи: 1) власні експериментальні дані з локалізації ряду ретровірусних послідовностей, інтегрованих у різні родини ізохор; 2) дані інших лабораторій відносно локалізації ретровірусних послідовностей у районах «відкритого» хроматину та/чи біля CpG-«острівців» та 3) власні дані з композиційного аналізу генів, розташованих біля інтегрованих провірусів.

Ці три групи даних дають дуже схожу картину, навіть якщо найбільша їхня кількість стосується транскрибованих інтегрованих послідовностей і тільки деякі — нетранскрибованих послідовностей. Відносно інтегрованих послідовностей, що транскрибуються, найперші роботи з локалізації BLV провірусів у геномі бика показали, що ці послідовності компартменталізовані в GC-багатих ізохорах, які відповідають високому GC-рівню провірусів («ізопікнічна» інтеграція).

Компартменталізована «ізопікнічна» інтеграція ретровірусних послідовностей була потім показана для всіх інших досліджених GC-багатих транскрибованих провірусів, так само як і для інтегрованих послідовностей GC-багатого ДНК-вмісного вірусу. Ці висновки підтверджуються іншими незалежними дослідженнями.

Відносно нетранскрибованих і часто дефектних вірусних послідовностей було показано, що вони, як правило, інтегруються як «ізопікнічно», так і більш широко розповсюджені в геномі хазяїна. Слід відзначити, що ці основні спостереження викликають декілька запитань.

Перше запитання стосується найпоширенішого випадку, коли ми маємо справу з GC-багатими транскрибованими провірусами, а саме: чому ці послідовності виявляються в невеликому композиційному компартменті геному хазяїна, багатому на гени, який також, як і провірус, є GC-багатим?

Аналіз онкогенів, розташованих поблизу про-

вірусів, дає можливість відповісти на перше запитання. Дійсно, більшість онкогенів — GC-багаті і локалізовані в GC-багатих і багатих на гени ізохорах. Очевидно, транскрипційна *cis*-активація під впливом провірусів може відбуватися лише тоді, коли провірус інтегрований в GC-багату ізохору, яка має онкоген. Це, в свою чергу, дозволяє припустити, що набута в результаті селекції перевага в темпах росту може пояснити компартменталізовану інтеграцію. Згадане пояснення було підкріплено даними, отриманими при дослідженні таких незвичайних випадків, коли GC-багаті вірусні послідовності виявлялися по сусідству з GC-бідними онкогенами або, навпаки, коли GC-бідні вірусні послідовності (такі як MMTV) було знайдено поблизу GC-багатих онкогенів (*int*-родина). Однак селекція необов'язково асоціюється з впливом вірусної послідовності через активацію сусіднього онкогена. Переважність інтеграції в транскрипційно активні райони геному може відобразити селективність провірусної експресії та передіснуючу різницю в транскрипційних властивостях районів, у які відбулась інтеграція [102].

Друге питання стосується мішені вірусних послідовностей. Якби інтеграція вірусної послідовності була явищем випадковим, то інтеграція послідовностей біля (GC-багатого) онкогена була б мало ймовірною, тому що GC-багаті ізохори представлені у ссавців (і у птахів) малою фракцією. Це призводить до думки, що навіть початок інтеграції є не випадковим процесом, але відбувається в ті компартменти геному, які є GC-багатими, багатими на гени, транскрипційно та рекомбінантно активні. Ідея «ізопікнічності» мішені, здається, підкріплюється деякими даними: 1) GC-багаті вірусні послідовності, що не транскрибуються, також відають перевагу цим районам; 2) родини мобільних елементів геному також переважно локалізовані в композиційно відповідних ізохорах [48, 101, 108]; 3) провірусні послідовності, вірогідно, найстабільніші в композиційно відповідному оточенні. Дійсно, гени хазяїв композиційно відповідають хромосомному середовищу, що ясно вказує на наявність композиційної кореляції між кодуєчими послідовностями та ізохорами, в яких вони локалізуються; необхідно підкреслити, що докази такої кореляції встановлені еволюційно [30].

Третє питання пов'язане з бімодальним композиційним розподілом ретровірусних послідовностей. Нами показано, що ретровірусні геноми розподіляються на два класи [103]: GC-бідний та GC-багатий класи. До першого належать всі лентівіруси і спумаіруси, що не мають онкогенів, але мають



гени регуляторних білків, та онковіруси В-типу і деякі онковіруси типу D (подібно MMTV), які також не мають онкогенів. До другого класу належать всі онковіруси, що містять онкогени (подібно RSV) та деякі онковіруси, що не мають онкогенів (подібно BLV та HTLV-I), за винятком В-типу та деяких представників D-типу. Виявлений бімодальний композиційний розподіл ретровірусних геномів супроводжується вражаючою композиційною гомогенністю генів (*gag*, *pol* та *env*, а також онкогенів та регуляторних генів) всередині кожного ретровірусного геному. Навіть довгі кінцеві повтори (LTRs) є або GC-багаті, або GC-бідні в залежності від рівня GC-основ у вірусних послідовностях. Можливо, ці вірусні послідовності коєволюціонували композиційно з послідовностями геному хазяїна, від деяких частин якого вони походять. Тому можна зробити висновок, що композиційні зміни, які призводять до утворення GC-багатих ізохор, призводять також і до композиційних змін інтегрованих вірусних послідовностей [103].

А. В. Ряндич, С. В. Зубак, Л. А. Циба, Н. Гулей,  
З. В. Лазуркевич, Дж. Бернарди

Региональная специфичность вирусной интеграции

Резюме

Интеграция ретровирусов в геном хозяев до недавнего времени считалась случайной и на сегодня еще нет ясности в отношении влияния установленных локальных эффектов в местах интеграции на процесс интеграции в целый геном. Использование композиционного подхода позволило установить, что интеграция ретровирусов происходит в определенные районы генома, подобные вирусным последовательностям по составу оснований. Такая региональная специфичность интеграции доказана: 1) нашими экспериментальными данными по локализации ретровирусных последовательностей в различных композиционных участках генома; 2) результатами других лабораторий по локализации ретровирусных последовательностей в «открытых» участках хроматина и/или в CpG-«островках»; 3) нашим композиционным анализом генов, расположенных по соседству с интегрированными вирусными последовательностями. Этот вывод важен как для понимания композиционной эволюции ретровирусных геномов, так и для генной терапии.

А. В. Rynditch, С. В. Zoubak, Л. О. Tsyba, Н. Guley,  
З. В. Lozurkevitch, G. Bernardi

Regional specificity of retroviral integration

Summary

The long-held general opinion was that retroviral integration into the cellular genome occurs at random and it is not clear how the local features of integration can be account for the pattern of integration over the whole genome. Using the compositional approach it was shown that viral integration takes place in some regions of the host genome which show a compositional match with viral sequences. This regional specificity of retroviral integration has been demonstrated by: (i) our experimental localization of a number of viral sequences integrated in different compositional genomic compartments; (ii) results from other laboratories concerning the localiza-

tion of retroviral sequences in open chromatin regions and/or next to CpG islands; (iii) our compositional analysis of genes in the neighborhood of integrated viral sequences. Such conclusion have implications concerning the compositional evolution of retroviral genomes and gene therapy.

СПИСОК ЛІТЕРАТУРИ

1. Temin H. M. The DNA provirus hypothesis // *Science*.—1976.—192.—P. 1075—1080.
2. Varmus H. E. The molecular genetics of cellular oncogenes // *Ann. Rev. Genet.*—1984.—18.—P. 553—612.
3. Goff S. P. Genetics of retroviral integration // *Ibid.*—1992.—26.—P. 527—544.
4. Kung H. J., Baerkoel C., Carters T. H. Retroviral mutagenesis of cellular oncogenes: a review with insights into the mechanisms of insertional activation // *Curr. Top. Microbiol. Immunol.*—1991.—171.—P. 1—25.
5. Tsielhis P. N., Lazo P. A. Virus-host interactions and the pathogenesis of murine and oncogenic retroviruses // *Ibid.*—P. 95—171.
6. Athas G. B., Starkey C. R., Levy L. S. Retroviral determinants of leukemogenesis // *Crit. Rev. Oncol.*—1994.—5.—P. 169—199.
7. Jankers J., Berns A. Retroviral insertional mutagenesis as a strategy to identify cancer genes // *Biochim. et biophys. acta.*—1996.—1287.—P. 29—57.
8. Largaespada D. A., Shanghnessy J. D. Jr., Jenkins J. A., Copeland N. G. Retroviral integration at the *Evi-1* locus in BXH-2 myeloid leukemia cell lines disrupts NGF-I expression without changes in steady-state Ras-GTP levels // *J. Virol.*—1995.—69.—P. 5095—5102.
9. Shen-Ong G. L. C., Morse H. C., Petter M., Mushinski F. Two models of *c-myc* activation in virus-induced mouse myeloid tumors // *Mol. Cell. Biol.*—1986.—6.—P. 380—392.
10. Weinstein Y., Cleveland J. L., Askew D. S. et al. Insertion and truncation of *c-myc* by murine leukemia virus in a myeloid cell line derived from cultures of normal hematopoietic cells // *J. Virol.*—1987.—61.—P. 2339—2343.
11. Setoguchi M., Higuchi Y., Yoshida S. et al. Insertional activation of *N-Myc* by endogenous Moloney-like murine retroviral sequences in macrophage cell lines derived from myeloma cell like-macrophage hybrids // *Mol. Cell. Biol.*—1989.—9.—P. 4515—4522.
12. Bishop J. M. Enemies within: The genesis of retrovirus oncogenes // *Cell.*—1980.—23.—P. 5—6.
13. Neil J. C. Defective avian sarcoma viruses // *Curr. Top. Microbiol. Immunol.*—1983.—103.—P. 51—75.
14. Hughes D. H. Synthesis, integration and transcription of the retroviral provirus // *Ibid.*—P. 23—51.
15. Bushman F. Targeting retroviral integration // *Science*.—1995.—276.—P. 1443—1444.
16. Anderson W. F. Human gene therapy // *Nature*.—1998.—392.—P. 25—30.
17. Varmus H. E., Brown P. Retroviruses // *Mobile DNA American Society for Microbiology / Eds D. E. Berg, M. M. Howe.*—Washington: DC., 1989.—P. 53—108.
18. Grangenet D. P., Munn S. R. Unraveling retrovirus integration // *Cell.*—1990.—60.—P. 3—4.
19. Weinberg R. A. Integrated genomes of animal viruses // *Annu. Rev. Biochem.*—1980.—49.—P. 197—226.
20. Hayward W. S., Neel B. G., Austrin S. M. AV1-induced lymphoid leukosis: activation of a cellular onc gene by promoter insertion // *Nature*.—1981.—290.—P. 475—480.
21. Payne G. S., Bishop J. M., Varmus H. E. Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas // *Ibid.*—1982.—295.—P. 209—214.

22. Brown P. O., Bowerman B., Varmus H. E., Bishop J. M. Correct integration of retroviral DNA *in vitro* // *Cell*.—1987.—49.—P. 347—356.
23. Fujiwara T., Mizuuchi K. Retroviral DNA integration: structure of an integration intermediate // *Ibid.*—1988.—54.—P. 497—504.
24. Withers-Ward E. S., Kitamura Y., Barnes J. P., Coffin J. M. Distribution of targets for avian retrovirus DNA integration *in vivo* // *Genes and Develop.*—1994.—8.—P. 1473—1487.
25. Filipski J., Thiery J. P., Bernardi G. An analysis of the bovine genome by Cs<sub>2</sub>SO<sub>4</sub>-Ag density gradient centrifugation // *J. Mol. Biol.*—1973.—80.—P. 177—197.
26. Thiery J. P., Macaya G., Bernardi G. An analysis of eukaryotic genomes by density gradient centrifugation // *Ibid.*—1976.—108.—P. 219—235.
27. Macaya G., Thiery J. P., Bernardi G. An approach to the organization of eukaryotic genomes at a macromolecular // *Ibid.*—P. 237—254.
28. Kettmann R., Meunier-Fotival M., Cortadas J. et al. Integration of bovine leukemia virus DNA in the bovine genome // *Proc. Nat. Acad. Sci. USA*.—1979.—76.—P. 4822—4826.
29. De Sario A., Roizes G., Allegre N., Bernardi G. A compositional map of the cen-q21 region of human chromosome 21 // *Gene*.—1997.—194.—P. 107—113.
30. Bernardi G. The human genome: Organization and evolutionary history // *Annu. Rev. Genet.*—1995.—29.—P. 445—476.
31. Sabeur G., Macaya G., Kadi F., Bernardi G. The isochore patterns of mammalian genomes and their phylogenetic implications // *J. Mol. Evol.*—1993.—37.—P. 93—108.
32. Zoubak S., Clay O., Bernardi G. The gene distribution of the human genome // *Gene*.—1996.—174.—P. 95—102.
33. Bernardi G., Bernardi G. Compositional constraints and genome evolution // *J. Mol. Evol.*—1986.—24.—P. 1—11.
34. Bernardi G., Olofsson B., Filipski J. et al. The mosaic genome of warm-blooded vertebrates // *Sciences*.—1985.—228.—P. 938—958.
35. Clay O., Caccio S., Zoubak S. et al. Human coding and non-coding DNA: compositional correlations // *Mol. Phylogenet. Evol.*—1996.—5.—P. 2—12.
36. Mouchiroud D., D'Onofrio G., Aissani B. et al. The distribution of genes in the human genome // *Gene*.—1991.—100.—P. 181—187.
37. Saccone C., Caccio S., Ferani P. et al. Compositional mapping of mouse chromosomes and identification of the gene-rich regions // *Chromosome Res.*—1997.—5.—P. 293—300.
38. Saccone S., De Sario A., Della Valle G., Bernardi G. The highest gene concentrations in the human genome are in T-bands of metaphase chromosomes // *Proc. Nat. Acad. Sci. USA*.—1992.—89.—P. 4913—4917.
40. Tazi J., Bird A. P. Alternative chromatin structure at CpG islands // *Cell*.—1990.—60.—P. 909—920.
41. Kettmann R., Cleuter Y., Mammerickx M. et al. Genomic integration of bovine leukemia provirus: comparison of persistent lymphocytosis with lymph node tumor form of enzootic bovine leukemia // *Proc. Nat. Acad. Sci. USA*.—1980.—77.—P. 2577—2581.
42. Zoubak S., Richardson J., Rynditch A. et al. Regional specificity of HTLV-I proviral integration in the human genome // *Gene*.—1994.—143.—P. 155—163.
43. Rynditch A., Kadi F., Geryk J. et al. The isopycnic, compartmentalized integration of Rous sarcoma virus sequences // *Ibid.*—1991.—106.—P. 165—172.
44. Salinas J., Zerial M., Filipski J. et al. Non-random distribution of MMTV proviral sequences in the mouse genome // *Nucl. Acids Res.*—1987.—15.—P. 3009—3022.
45. Bodor J., Svoboda J. The LTR, v-src, LTR provirus generated in the mammalian genome by src mRNA reverse transcription and integration // *J. Virol.*—1989.—63.—P. 1015—1018.
46. Machon O., Hejnar J., Hajkova P. et al. The LTR, v-src, LTR provirus in H-19 hamster tumor cell line is integrated adjacent to the negative regulatory region // *Gene*.—1996.—174.—P. 9—17.
47. Salinas J., Zerial M., Filipski J., Bernardi G. Gene distribution and nucleotide sequence organization in the mouse genome // *Eur. J. Biochem.*—1986.—160.—P. 469—478.
48. Zerial M., Salinas J., Filipski J., Bernardi G. Genomic localization of hepatitis B virus in a human hepatoma cell line // *Nucl. Acids Res.*—1986.—14.—P. 8373—8386.
49. Romani M., Casciano J., Querzola F. et al. Analysis of a viral integration event in a GC-rich region at the 1p36 human chromosomal site // *Gene*.—1993.—135.—P. 153—160.
50. Marcus S. L., Smith S. W., Sarkar N. H. Quantitative of murine mammary tumor virus-related RNA in mammary tissues of low- and high-mammary-tumor-incidence mouse strains // *J. Virol.*—1981.—40.—P. 87—95.
51. Choi Y. W., Henrard D., Lee I., Ross S. R. The mouse mammary tumor virus long terminal repeat directs expression in epithelial and lymphoid cells of different tissues in transgenic mice // *Ibid.*—1987.—61.—P. 3013—3019.
52. Henrard D., Ross S. R. Endogenous mouse mammary tumor virus is expressed in several organs in addition to the lactating mammary gland // *Ibid.*—1988.—62.—P. 3046—3049.
53. Vijaya S., Steffen D. L., Kozak C., Robinson H. L. Acceptor sites for retroviral integrations map near DNA I-hypersensitive sites in chromatin // *Ibid.*—1986.—60.—P. 683—692.
54. Rohdewohld H., Weinher H., Reik W. et al. Retrovirus integration and chromatin structure: Moloney murine leukemia proviral integration sites map near DNA I-hypersensitive sites // *Ibid.*—1987.—61.—P. 336—343.
55. Scherdin V., Rhodes K., Brendl M. Transcriptionally active genome regions and preferred targets for retrovirus integration // *Ibid.*—1990.—64.—P. 907—912.
56. Mooslehner K., Karl U., Harbers K. Retroviral sites in transgenic M<sub>0</sub>v mice frequently map in the vicinity of transcribed DNA region // *Ibid.*—1990.—64.—P. 3056—3058.
57. Fincham V. J., Wyke J. A. Differences between cellular integration sites of transcribed and non transcribed Rous sarcoma proviruses // *Ibid.*—1991.—65.—P. 461—463.
58. Robinson H. L., Gagnon G. Patterns of proviral insertion and deletion in avian leukosis virus-induced lymphomas // *Ibid.*—1986.—57.—P. 28—36.
59. Schubach W., Groudine M. Alteration of c-myc chromatin structure by avian leukosis virus integration // *Nature*.—1984.—307.—P. 702—708.
60. Edmonson D. G., Roth S. Y. Chromatin and transcription // *FASEB J.*—1996.—10.—P. 1173—1182.
61. Aissani B., Bernardi G. CpG islands, genes, isochores in the genome of vertebrates // *Gene*.—1991.—106.—P. 185—195.
62. Aissani B., Bernardi G. CpG islands: features and distribution in the genome of vertebrates // *Ibid.*—P. 173—183.
63. Ozturk M. Chromosomal rearrangements and tumor suppressor genes in primary liver cancer // *Primary liver cancer: etiological and progression factors* / Ed. C. Brechat.—Boca Raton: CRC press, 1994.—P. 269—281.
64. Lazo P. A., Lee J. S., Tschlis P. N. Long-distance activation of the Myc protooncogene by provirus insertion in Mlv-1 or Mlv-4 in rat T-cell lymphomas // *Proc. Nat. Acad. Sci. USA*.—1990.—87.—P. 170—173.
65. Bartholomew C., Ihle J. N. Retroviral insertions 90 kilobases proximal to the Evi-1 myeloid transforming gene activate transcription from the normal promoter // *Mol. Cell. Biol.*—1991.—11.—P. 1820—1828.

66. Jiang X., Villeneuve L., Tarmel C. et al. The Myb and Abi-1 genes are physically very closely linked on mouse chromosome 10 // *Mam. Genome.*—1994.—5.—P. 142—148.
67. Shen-Ong G. L., Wolff L. Moloney murine leukemia virus-induced myeloid tumors in adult BALB/c mice: requirement of *c-myc* activation but lack of *v-abl* involvement // *J. Virol.*—1987.—61.—P. 3721—3725.
68. Wolff L., Koller R., Davidson W. Acute myeloid leukemia induction by amphotropic murine retrovirus (4070A): clonal integrations involve *c-myc* in some but not all leukemias // *Ibid.*—1991.—65.—P. 3607—3616.
69. Bordereaux D., Fichelson S., Sola B. et al. Frequent involvement of the *fm-3* region in Friend murine leukemia virus-induced mouse myeloblastic leukemias // *Ibid.*—1987.—61.—P. 4043—4045.
70. Mushinski J. F., Mountz J. D., Pierce J. H. et al. Expression of the murine proto-oncogene *bcl-2* is stage specific and cell-type specific // *Curr. Top. Microbiol. Immunol.*—1988.—141.—P. 332—336.
71. Morishita K., Parker D. S., Mucenski M. L. et al. Retroviral activation of a novel gene encoding a zinc finger protein in IL-3-dependent myeloid leukemia cell lines // *Cell.*—1988.—54.—P. 831—840.
72. Pacht C., Schabach W., Eisenman R., Linial M. Expression of *c-myc* RNA in bursal lymphoma cell lines: identification of *c-myc*-encoded proteins by hybrid-selected translation // *Ibid.*—1983.—33.—P. 335—344.
73. Boerkoel C. F., Kung H. J. Transcriptional interaction between retroviral long terminal repeats (LTRs): mechanism of 5' LTR suppression and 3' LTR promoter activation of *c-myc* in avian B-cell lymphomas // *J. Virol.*—1992.—66.—P. 4814—4823.
74. Fung Y.-K. T., Lewis W. G., Crittenden L. B., Kung H. J. Activation of the cellular oncogene *c-erbB* by LTR insertion: molecular basis for induction of erythroblastosis by avian leukosis virus // *Cell.*—1983.—33.—P. 357—368.
75. Nilsen T. W., Maroney P. A., Goodwin R. G. et al. *c-erbB* activation in ALV-induced erythroblastosis: novel RNA processing and promoter insert on result in expression of an amino-truncated EGF receptor // *Ibid.*—1985.—41.—P. 719—726.
76. Collart K. L., Aurigemma R., Smith R. E. et al. Infrequent involvement of *c-fos* in avian leukosis virus-induced nephroblastoma // *J. Virol.*—1990.—64.—P. 3541—3544.
77. Kabrun N., Bumstead N., Hayman M. J., Enrietto P. J. Characterization of a novel promoter insertion in the *c-rel* locus // *Mol. Cell. Biol.*—1990.—10.—P. 4788—4794.
78. Nusse R., Varmus H. E. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome // *Cell.*—1982.—31.—P. 99—109.
79. Peters G., Brookes S., Smith R. et al. The mouse homolog of the *hst/k-FGF* gene is adjacent to *int-2* and is activated by proviral insertion in some virally induced mammary tumors // *Proc. Nat. Acad. Sci. USA.*—1989.—86.—P. 5678—5682.
80. Morris D. W., Dutra J. C. Identification of a MMTV insertion mutation within the coding region of the *Fgf-3* protooncogene // *Virology.*—1997.—238.—P. 161—165.
81. Gullahan D., Callahan R. Mammary tumorigenesis in feral mice—identification of a new *int* locus in mouse mammary tumor-virus (czech II)-induced mammary tumors // *J. Virol.*—1987.—61.—P. 66—74.
82. Roelink H., Wagenaar E., Lopes da Silva S., Nusse R. *Wnt-3*, a gene activated by proviral insertion in mouse mammary tumors, is homologous to *int-1/Wnt-1* and is normally expressed in mouse embryos and adult brain // *Proc. Nat. Acad. Sci. USA.*—1990.—87.—P. 4519—4523.
83. Peters G., Brookes S., Placzek M. et al. A putative *int* domain for mouse mammary tumor virus on mouse chromosome 7 is a 5' extension of *int-2* // *J. Virol.*—1989.—63.—P. 1448—1450.
84. MacArthur C. A., Shankar D. B., Shackelford G. M. *Fgf-8*, activated by proviral insertion, cooperates with the *Wnt-1* transgene in murine mammary tumorigenesis // *Ibid.*—1995.—69.—P. 2501—2507.
85. Marchetti A., Butitta F., Miyazaki S. et al. *Int-6*, a highly conserved, widely expressed gene, is mutated by mouse mammary tumor virus in mammary preneoplasia // *Ibid.*—1993.—P. 1932—1938.
86. Lee F. S., Lane T. F., Kuo A. et al. Insertional mutagenesis identifies a member of the *Wnt* gene family as a candidate oncogene in the mammary epithelium of *int-2/Fgf-3* transgenic mice // *Proc. Nat. Acad. Sci. USA.*—1995.—92.—P. 2268—2272.
87. Morris D. W., Barry P. A., Bradshaw H. D. Jr., Cardiff R. D. Insertion mutation of the *int-1* and *int-2* loci by mouse mammary tumor virus in pre-malignant and malignant neoplasms from the GR mouse strain // *J. Virol.*—1990.—64.—P. 1794—1802.
88. Dickson C., Smith R., Brookes S., Peters G. Proviral insertions within the *int-2* gene can generate multiple anomalous transcripts but leave the protein-coding domain intact // *Ibid.*—N 2.—P. 784—793.
89. Miyazaki S., Kozak C. A., Marchetti A. et al. The chromosomal location of the mouse mammary tumor gene *Int-6* and related pseudogenes in the mouse genome // *Genomics.*—1995.—27.—P. 420—424.
90. Robbins J., Blondel B. J., Callahan D., Callahan R. Mouse mammary tumor gene. *Int-3*: a member of the Notch gene family transforms mammary epithelial cells // *J. Virol.*—1992.—66.—P. 2594—2599.
91. Diella F., Levi G., Callahan R. Characterization of the INT6 mammary tumor gene product // *DNA Cell Biol.*—1997.—16.—P. 839—847.
92. Pryciak P. M., Varmus H. E. Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection // *Cell.*—1992.—69.—P. 769—780.
93. Pryciak P. M., Sil A., Varmus H. E. Retroviral integration into minichromosomes *in vitro* // *EMBO J.*—1992.—11.—P. 291—303.
94. Pryciak P. M., Muller H.-P., Varmus H. E. Simian virus 40 minichromosomes as targets for retroviral integration *in vivo* // *Proc. Nat. Acad. Sci. USA.*—1992.—89.—P. 9237—9241.
95. Kitamura Y., Lee Y. M. H., Coffin J. M. Non-random integration of retroviral DNA *in vitro*: effect of CpG methylation // *Ibid.*—P. 5532—5536.
96. Muller H. P., Varmus H. E. DNA bending creates favored sites for retroviral integration: an explanation for preferred insertion sites in nucleosomes // *EMBO J.*—1994.—13.—P. 4704—4714.
97. Pruss D., Bushman F. D., Wolffe A. P. Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core // *Proc. Nat. Acad. Sci. USA.*—1994.—91.—P. 5913—5917.
98. Bor Y.-C., Bushman F. D., Orgel L. E. *In vitro* integration of human immunodeficiency virus type 1 cDNA into targets containing protein-induced bends // *Ibid.*—1995.—92.—P. 10334—10338.
99. Chou K. S., Okayama A., Su I.-J. et al. Preferred nucleotide sequence at the integration target site of human T-cell leukemia virus type 1 from patients with adult T-cell leukemia // *Int. J. Cancer.*—1996.—65.—P. 20—24.
100. Shimotohno K., Temin H. M. No apparent nucleotide sequence specificity in cellular DNA juxtaposed to retrovirus proviruses // *Proc. Nat. Acad. Sci. USA.*—1980.—77.—P. 7357—7361.

101. Meunier-Rotival M., Scirano P., Cuny G. et al. Sequence organization and genomic distribution of the major family of interspersed repeats of mouse DNA // *Ibid.*—1982.—79.—P. 355—359.
102. Peckham J., Sobel S., Comer J. et al. Retrovirus activation in embryonal carcinoma cells by cellular promoters // *Genes Dev.*—1989.—3.—P. 2052—2071.
103. Zoubak S., Rynditch A., Bernardi G. Compositional bimodality and evolution of retroviral genomes // *Gene.*—1992.—119.—P. 207—213.
104. Boiocchi M., Dolcetti R., Maestro R. et al. A coordinated proto-oncogene expression characterizes MCF 247 murine leukemia virus-induced T-cell lymphomas irrespectively of proviral insertion affecting *myc* loci // *Leukemia Res.*—1990.—14.—P. 549—558.
105. Kratochwil K., von der Mark K., Kollar E. J. Retrovirus-induced insertional mutation in *Mov13* mice affects collagen I expression in a tissue-specific manner // *Cell.*—1989.—57.—P. 807—816.
106. Selten G., Cuypers H. T., Berns A. Proviral activation of the putative oncogene *Pim-1* in MuLV induced T-cell lymphomas // *EMBO J.*—1985.—4.—P. 1793—1798.
107. Moreau-Gachelin F., Ray D., de-Both N. J. et al. *Spi-1* oncogene activation in Rauscher and Friend murine virus-induced acute erythroleukemias // *Leukemia.*—1990.—4.—P. 20—23.
108. Soriano P., Meunier-Rotival M., Bernardi G. The distribution of interspersed repeats is non-uniform and conserved in the mouse and human genomes // *Proc. Nat. Acad. Sci. USA.*—1983.—80.—P. 1816—1820.
109. Flubacher M. M., Bear S. E., Tschlis P. N. Replacement of interleukin-2 (IL-2)-generated mitogenic signals by a mink cell focus-forming (MCF) or xenotropic virus-induced IL-9 dependent autocrine loop: implications for MCF virus-induced leukemogenesis // *Virology.*—1994.—68.—P. 7709—7716.
110. O'Donnell P. V., Fleissner E., Lonial H. et al. Early clonality and high-frequency proviral integration into the *c-myc* locus in AKR leukemias // *J. Virol.*—1985.—55.—P. 500—503.
111. Van-Lohuizen M., Verbeek S., Scheijen B. et al. Identification of cooperating oncogenes in *E mu-myc* transgenic mice by provirus tagging // *Cell.*—1991.—65.—P. 737—752.
112. Tremblay P. J., Kozak C. A., Jolicoeur P. Identification of a novel gene, *Vin-1*, in murine leukemia virus-induced T-cell leukemias by provirus insertional mutagenesis // *J. Virol.*—1992.—66.—P. 1344—1353.
113. Wolf D., Admon S., Oren M., Rotter V. Abelson murine leukemia virus-transformed cells that lack p53 protein synthesis express aberrant p53 mRNA species // *Mol. Cell. Biol.*—1984.—4.—P. 552—558.
114. Voronova A. F., Sefton B. M. Expression of a new tyrosine protein kinase is stimulated by retrovirus promoter insertion // *Nature.*—1986.—319.—P. 682—685.
115. Vijaya S., Steffen D. L., Kozak C., Robinson H. L. *Dsi-1*, a region with frequent proviral insertions in Moloney murine leukemia virus-induced rat thymomas // *J. Virol.*—1987.—61.—P. 1164—1170.
116. Gisselbrecht S., Fichelson S., Sola B. et al. Frequent *c-fms* activation by proviral insertion in mouse myeloblastic leukemias // *Nature.*—1987.—329.—P. 259—261.
117. Ihle J. N., Smith-White B., Sisson B. et al. Activation of the *c-H-ras* proto-oncogene by retrovirus insertion and chromosomal rearrangement in a Moloney leukemia virus-induced T-cell leukemia // *J. Virol.*—1989.—63.—P. 2959—2966.
118. Ben-David Y., Giddens E. B., Letwin K., Bernstein A. Erythroleukemia induction by Friend murine leukemia virus insertional activation of a new member of the *ets* gene family. *FLI-1*, closely *L1* *ked* to *c-ets-1* // *Genes Dev.*—1991.—5.—P. 908—918.
119. Bear S. E., Bellacosa A., Lazo P. A. et al. Provirus insertion in *Tpl-1*, an *Ets-1*-related oncogene, is associated with tumor progression in Moloney murine leukemia virus-induced rat thymic lymphomas // *Proc. Nat. Acad. Sci. USA.*—1989.—86.—P. 7495—7499.
120. Makris A., Patriotis C., Bear S. E., Tschlis P. N. Genomic organization and expression of *Tpl-2* in normal cells and Moloney murine leukemia virus-induced rat T-cell lymphomas: activation by provirus insertion // *J. Virol.*—1993.—67.—P. 4283—4289.
121. Shimada Y., Migliaccio G., Ruscetti S. et al. Expression of the interleukin-3 and granulocyte-macrophage colony-stimulating factor genes in Friend spleen focus-forming virus-induced erythroleukemia // *Blood.*—1992.—79.—P. 2423—2431.
122. Ben-David Y., Bani M. R., Chabot B. et al. Retroviral insertions downstream of the heterogeneous nuclear ribonucleoprotein A1 gene in erythroleukemia cells: evidence that A1 is not essential for cell growth // *Mol. Cell. Biol.*—1992.—12.—P. 4449—4455.
123. Lu S. J., Rowan S., Bani M. R., Ben D. Y. Retroviral integration within the *Fli-2* locus results in inactivation of the erythroid transcription factor NF-E2 in Friend erythroleukemias: evidence that NF-E2 is essential for globin expression // *Proc. Nat. Acad. Sci. USA.*—1994.—91.—P. 8398—8402.
124. Matthews E. A., Vasmel W. L., Schoenmakers H. J., Melief C. J. Retrovirally induced murine B-cell tumors rarely show proviral integration in sites common in T-cell tumors // *Int. J. Cancer.*—1989.—43.—P. 1120—1125.
125. Morishita K., Parker D. S., Mucenski M. L. et al. Retroviral activation of a novel gene encoding a zinc finger protein in IL-3-dependent myeloid leukemia cell lines // *Cell.*—1988.—54.—P. 831—840.
126. George D. L., Glick B., Trusko S., Freeman N. Enhanced *c-Ki-ras* expression associated with Friend virus integration in a bone marrow-derived mouse cell line // *Proc. Nat. Acad. Sci. USA.*—1986.—83.—P. 1651—1655.
127. Noori-Daloui M. R., Swift R. A., Kung H. J. et al. Specific integration of REV proviruses in avian bursal lymphomas // *Nature.*—1981.—294.—P. 574—576.
128. Kanter M. R., Smith R. E., Hayward W. S. Rapid induction of B-cell lymphomas: insertional activation of *c-myc* by avian leukosis virus // *J. Virol.*—1988.—62.—P. 1423—1432.
129. Westaway D., Papkoff J., Moscovici C., Varmus H. E. Identification of a provirally activated *c-ha-ras* oncogene in an avian nephroblastoma via a novel procedure: cDNA cloning of a chimeric viral-host transcript // *EMBO J.*—1986.—5.—P. 301—309.
130. Dickson C., Smith R., Brookes S., Peters G. Tumorigenesis by mouse mammary tumor virus: proviral activation of a cellular gene in the common integration region *int-2* // *Cell.*—1984.—37.—P. 529—536.
131. Chalker D. L., Sandmeyer S. B. *Ty3* integrates within the region of DNA polymerase III transcription initiation // *Genes Dev.*—1992.—6.—P. 117—128.
132. Gilks C. B., Bear S. E., Grimes H. L., Tschlis P. N. Progression of interleukin-2 (IL-2) dependent rat T cell lymphoma lines to IL-2 independent growth following activation of a gene (*Gfi-1*) encoding a novel zinc finger protein // *Mol. Cell. Biol.*—1993.—13.—P. 1759—1768.
133. Marchetti A., Buttitta F., Miyazaki S. et al. *Int-6*, a highly conserved, widely expressed gene, is mutated by mouse mammary tumor virus in mammary preneoplasia // *J. Virol.*—1995.—69.—P. 1932—1938.

Надійшла до редакції 01.04.98