Univerzita Karlova v Praze

Přírodovědecká fakulta

Petr Pajer

Od hledání nových onkogenů k pokusu předefinovat fenomén kancerogeneze

Disertační práce

Školitel: RNDr. Michal Dvořák, CSc.

Ústav Molekulární Genetiky

Akademie Věd České republiky

Praha, březen 2012

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, březen 2012

Poděkování

Především bych chtěl poděkovat svému školiteli Michalu Dvořákovi nejen za veškerou vědeckou i materiální pomoc a podporu, ale hlavně za trpělivost, kterou měl se mnou a mými názory. Dále musím uvést, že rozsah experimentálních výsledků je takový, že bych je nikdy sám nedokázal realizovat bez pomoci Vladimíra Pečenky (kterému též vděčím za naučení mnoha experimentálním metodikám a pragmatickému přístupu vůbec), Vítka Karafiáta (bez něhož by nebylo náročných a zdlouhavých experimentů se zvířaty a tkáňovými kulturami) a Jany Dudlové (která mi obětavě pomohla s enormním množstvím sekvenací a PCR reakcí). Zde popisované výsledky patří stejnou měrou i těmto čtyřem výjimečným lidem. Také chci poděkovat všem kolegům na pracovišti za vytvoření přátelské a plodné pracovní atmosféry (specielní dík Vladimíru Čermákovi za podíl na vzniku termínu „industáze“). Největší dík však patří mým nejbližším a rodičům za domácí prostředí, které mi vytvořili, bez něhož by tato práce vzniknout nemohla.

Úvod I. - Literární

Disertační práce je v mnoha ohledech zvláštní hybrid literárního a vědeckého textu, jehož nečetný čtenář snad může být shovívavější k nepřesnostem a chybám nežli by tomu bylo u striktně vědeckých textů. Proto se domnívám, že je to také jedinečná příležitost pro publikování názorů, myšlenek a výsledků studia, které nejsou a nemohou být z různých příčin úplné, a jejichž publikování si kandidát vědeckého titulu bude moci jinde jen s obtížemi dovolit. Zvláště pokud se jedná o myšlenky a plánované experimenty, jejichž dokončením nebo dokonce samotnou realizací v budoucnosti si nemůže být jist. Rozhodl jsem se proto využít této příležitosti a poněkud se odchýlit od běžné stavby vědeckého textu a nahradit ji dvoudílnou strukturou, kde první část je věnována nezbytným praktickým výsledkům individuálního PhD studia a druhá je pak zamyšlením nad kontextem otázek, které by měla současná věda řešit.

Jestliže první část představuje ucelený hmatatelný výsledek studia, pak část druhá je z hlediska autora podstatnější. Má odrážet výsledek skutečný – výsledek formování představ autora o studovaném tématu, tak jak se vyvinuly v průběhu studijních let. Tento vývoj je unikátní a nemůže být uspíšen. Ze své filosofické podstaty tato část spíše připomíná staré neúplné texty dochované jen ve fragmentech několikanásobných opisů. Takové fragmenty trpí několikerým zkreslením – ztrátou celých částí původního textu, chybami při opakovaném opisování (ať již úmyslnými či neúmyslnými), komolením které je nutně vyvoláváno překlady z jednoho jazyka do druhého, jakož i snahou individuálních opisovačů o aktualizaci. Předkládaný text disertační práce trpí podobnými neduhy, avšak z odlišných příčin – zkoumání chybějícího textu nebylo uzavřeno, některé chyby jsou způsobeny autorovou neznalostí a jiné specifickým výběrem či subjektivní interpretací publikovaných dat. Sepsání této práce bylo motivováno dvěma cíli. Prvním je bilancování výsledků studia (což má být pravým cílem všech disertačních prací) a druhým pak snaha zaujmout čtenáře – byť jich bude jen několik – tak, aby nepovažovali čas ztrávený čtením za bezúčelně promarněný.

Úvod II. - Formulace Otázek

V souvislosti se středoškolským studiem musím v prvé řadě zmínit klíčové - jakkoliv triviální - moudro, které nám na jedné hodině biologie předestřel učitel. Jedná se o přístup k řešení velmi komplexních problémů. Lidé si často položí Otázku, na kterou chtějí znát Odpověď (velké „O“ jsem v textu použil pro odlišení pomyslných zásadních „Otázek“ od

„otázek“ dílčích). Pokud je položená Otázka obtížná, rozdělí ji na více částí, podproblémů, které zkoumají odděleně. Pokud jsou i tyto podproblémy příliš složité, pak je opět rozdělí na menší podpodproblémy a postup opakují dokud nejsou schopni položenou podpod...podotázku vyřešit. Nevyřčeným předpokladem pak je, že soubor odpovědí na všechny podpod...podotázky bude zároveň Odpovědí na Otázku. Popsaný postup však často vede k tomu, že řešitel zapomene jak zněla původní Otázka a následně jako výsledek svého bádání zjistí detailní avšak irrelavantní odpověď. A soubor přesných odpovědí na zcela nedůležité otázky může jen stěží být hledanou Odpovědí. Jinými slovy – v důsledku zmíněného přístupu můžeme nakonec vědět všechno o ničem. Takový přístup je běžný a přes zdánlivou pejorativnost ho nelze označit za chybný. V reálném světě grantových aplikací - což píši zcela prost sarkasmu - to ani jinak dělat nelze. Spíše lze říci, že vede pouze k „mírnému pokroku v mezích zákona“ (viz. Část druhá, kapitola třetí).

Této na pohled triviální myšlence věnuji úvodní prostor proto, že jsem se opakovaně přesvědčoval o její platnosti jak na sobě, tak na svém okolí. Přijde-li například začínající student do laboratoře zkoumající nádorovou biologii, položí si pravděpodobně Otázku: „Jak vyléčit rakovinu?“ Učí se a dojde k poznání, že je mnoho druhů rakoviny a rozhodne se tedy zkoumat určitý typ nádoru. Přirozeně se vyskytující nádory ovšem lze z experimentálních i etických důvodů jen obtížně zkoumat, zvolí tedy zvířecí model, například inbrední myší kmen, který po infekci rekombinantním virem, nesoucím mutovanou formu protoonkogenu A, akutně vyvine podobný typ nádoru. Pokusí se tedy vpravit tento onkogen do buněk/zvířete jiným expresním vektorem a zjistí, že oproti očekávání tímto postupem transformaci buněk nedocílí. Proč? Po detailnějším studiu odhalí, že pro zmíněnou transformaci je zapotřebí spolupůsobení onkogenu A a minoritní spoluúčast virového proteinu B. Následující detailní studium ho přivede ke konstatování, že to je fosforylace serinu v poloze 456 virového proteinu B, která je zcela nepostradatelná pro zmíněnou transformaci onkogenem A. Tedy od původní otázky „Jak vyléčit rakovinu?“ se onen hypotetický student úspěšně dobral překvapivé, detailní a v zásadě správné odpovědi, že „pro akutní transformaci buněk specifické myší inbrední linie specifickou mutantní formou protoonkogenu A je naprosto nezbytná spoluúčast virového proteinu B fosforylovaného na serinu v poloze 456“. Fiktivní příběh připomíná variaci na dnes již klasickou knihu Stopařův průvodce po Galaxii (D. Adams), vždyť kořeny zcela odlišných jevů bývají často prosté a shodné. Smyslem takto obšírně podaného příkladu není odsoudit nebo vysmívat se zmíněnému přístupu, ale zdůraznit

význam neustálého připomínání původní Otázky, kdykoliv se člověk rozhodne položit si nějakou otázku dílčí.

A nyní se mohu vrátit k ranné formulaci Otázek. V průběhu studia na střední škole mne zaujaly tři nezodpovězené Otázky, které se mi jevily jako klíčové pro hledání smyslu života: 1) Otázka existence života, 2) Otázka smrtelnosti a 3) Otázka existence rakoviny. Zodpovězení těchto otázek jsem považoval za cíl svého studia. Pokud moje paměť sahá, byl jsem přesvědčen, že spolu souvisí a nelze je řešit odděleně, jakkoliv se mohou zdát Otázky 2 a 3 být pouhou podmnožinou té první. Asi jako každý jsem byl po krátkou dobu udiven, jak málo je známo o podstatě těchto Problémů, a zároveň přesvědčen, že musí být tak snadné vše vysvětlit. Již dávno si nemyslím, že bych se odpovědi na své Otázky dozvěděl (tím méně je sám nalezl), přesto se domnívám, že má smysl pokoušet se je zodpovědět a myslet na ně vždy, když si člověk pokládá nějakou novou podpod...podotázku. Ne pro samotné Odpovědi, ale spíše abychom se naučili klást správnější podotázky.

Domníval jsem se a stále se domnívám, že k zodpovězení těchto Otázek může člověk využít studium téměř libovolného modelu/jevu, neboť Odpovědi by měly mít obecnou platnost. Proto jsem se rozhodl pro studium nádorové transformace na zvířecích modelech. Co má paměť sahá, nikdy jsem nechtěl najít lék nebo dokonce léčit. Připadalo mi to jako problém v pořadí druhotný. Samozřejmě jsem také nikdy neměl dostatek odvahy přijmout zodpovědnost, kterou na sebe bere lékař, čelit bezmocnosti, jaké je často nucen čelit. Modely virově indukovaných nádorů kuřat se mi tedy jevily jako ideální příležitostí k experimentálním studiím, protože v sobě spojovaly nejširší spektrum problémů základního výzkumu mnohobuněčného života a nádorové transformace. Mechanismy transformace akutním onkogenem (AMV/v-myb), hledání nových protoonkogenů a studium mechanismů jejich onkogenní aktivace (MAV2), srovnávání molekulárních příčin transformace buněk různého původu (ledviny-játra-plíce) a konečně srovnávání molekulárních příčin buněčné tranformace u různých živočišných druhů (kuře-člověk-myš). Abych tedy shrnul dva zcela odlišné, ale přesto nezvratně provázané přístupy, které mne provázely studiem a které se snažím reflektovat v tomto textu: jsou to snaha nezapomenout formulaci původních Otázek (zde reprezentovaná filosofickou pasáží) a zároveň produkce praktických výsledků v reálném čase (reprezentovaná studiem buněčné transformace na konkrétních kuřecích modelech).

Část první – Ptačí modely nádorové transformace

Vzpomenu-li na historické milníky molekulární genetiky, jichž bylo dosaženo studiem ptačích modelů, vybaví se mi dva protichůdné pocity – hrdost na kontinuitu, které jsme součástí a podiv nad tím, jak jsou tyto modely dnes podceňované a zatracované. Zde mohu zmínit opravdu jen ty nejvýznačnější objevy, učiněné při studiu kuřat – objev retrovirů jako infekčních agens schopných vyvolat rakovinu [1, 2], objev reverzní transkriptázy, bez níž by molekulární genetika v současné podobě nebyla myslitelná [3, 4] a konečně objev prvního onkogenu a jeho zdroje v genomu hostitelské buňky – v-src a c-src – kterýžto objev se stal základem molekulární koncepce onkogeneze [5]. Mimo tyto jednotlivé objevy musím zmínit celá dvě vědecká odvětví, která byla od základu hluboce ovlivněná studiem ptáků - imunologie (T a B lymfocyty, imunologická tolerance) a vývojová biologie a embryologie. Z pohledu českého vědce je fascinující, inspirující a zároveň smutné, jak málo chybělo k tomu, aby za těmito průlomovými objevy stála česká jména (Hašek M., Svoboda J., Říman J.).

V posledních desetiletích je v základním výzkumu jasně patrný nástup myších modelů na úkor ptačích. Důvodů trvajícího odklonu od užívání kuřete jako modelového organismu je hned několik. Nejzjevnější jsou důvody ekonomické (cena chovu kuřat v porovnání s myší), dále pak časové (generační doba kuře vs. myš) a technické (dostupnost transgenních organismů, inbredních linií, tkáňových linií, používání citlivějších technik s menšími nároky na množství vzorku). Méně informovaný člověk by snad také argumentoval kratší evoluční vzdáleností myši od středobodu všeho zkoumání – člověka. Oproti tomu je drůbež světově největším zdrojem živočišných bílkovin pro výživu lidstva a proto se současný výzkum specializuje hlavně na šlechtění ekonomicky zajímavých kmenů drůbeže, což je prozatím asi jediná oblast kde myší model významně ztrácí.

Je zbytečné polemizovat s předchozím výčtem důvodů, jsou obecně pravdivé, ale co platí obecně mívá i své výrazné výjimky pro které má zachování dalších zvířecích modelů hluboký smysl. Například je obecně známo, že myší nádorové modely se v některých aspektech výrazně odlišují od lidských přirozeně vzniklých nádorů [6, 7]. Namátkou zmíním jednokrokovou immortalizaci myších buněk vs. dvoukrokové lidské (a kuřecí). Další je právě ona větší evoluční vzdálenost kuřete a člověka, umožňující náhled na obecné a odlišné aspekty buněčné transformace mezi ptáky a savci. A zcela na závěr ještě zmíním fenomén, ne

zrovna snadno kvantifikovatelný, ale přesto velmi významný - jev, kterým jsem tuto kapitolu započal – a tím je historická kontinuita.

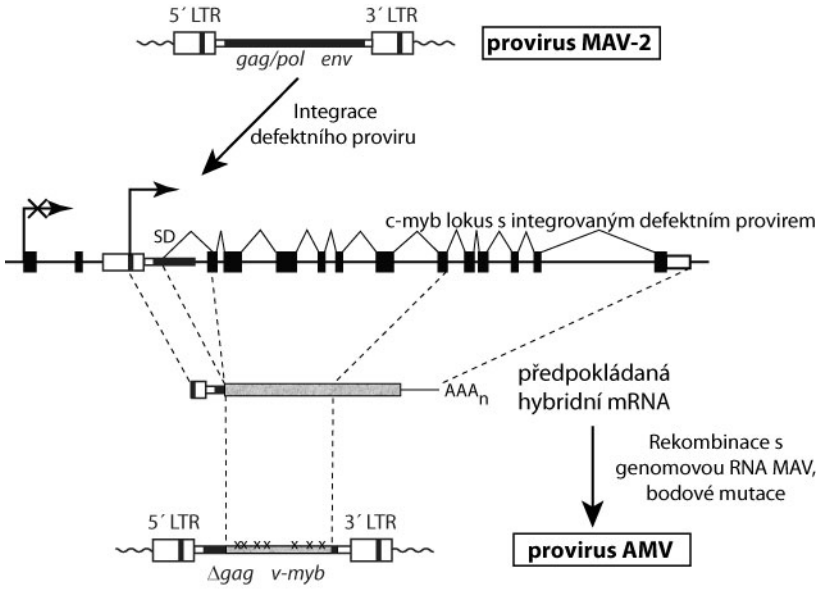
Kapitola první – Kuřecí modely inzerční mutageneze

Koncept využití transposonů a jimi vyvolané inzerční mutageneze pro identifikaci genů spojených s určitým fenotypem se objevil již počátkem 80. let (samozřejmě XX. století). Přirozeným a efektivním využitím této experimentální strategie bylo studium genetického základu tumorigeneze za použití retrovirů, jejichž unikátní vlastnosti stály u revolučních objevů na poli onkogeneze. Retroviry mohou vyvolávat vznik nádorů dvěma způsoby. Akutně, přímou transdukcí onkogenů (tj. mutantních forem normálních hostitelských genů s onkogenním potenciálem) nebo latentně, inzerční mutagenezí (tj. inzercí svého provirového genomu do genomu hostitele, což způsobí onkogenní aktivaci postiženého genu).

Rozdílné mechanismy nádorové transformace akutním a latentním retrovirem jsou dobře patrné na dvojici příbuzných kuřecích retrovirů AMV (**A**vian **M**yeloblastosis **V**irus) / MAV (**M**yeloblastosis **A**ssociated **V**irus) – obr. 1. AMV je typickým reprezentantem akutních onkogenních retrovirů, které nádorově transformují téměř každou úspěšně infikovanou cílovou buňku. Kóduje kuřecí onkogen v-myb, avšak z důvodu ztráty většiny původního virového genomu je replikačně defektní a pro své infekční šíření potřebuje pomocný (helper) virus, kterým je jeho příbuzný a zároveň předek, replikačně kompetentní MAV. Retrovirus AMV vznikl vícekrokově. V prvním kroku integrace defektního proviru MAV do intronu protoonkogenu c-myb vedla k produkci hybridní mRNA MAV-c-mybΔN (dalece N-terminálních exonů). Následně rekombinací této mRNA s další genomovou molekulou rodičovského retroviru a nahromaděním bodových mutací vznikl AMV. Jím kódovaný protein tak ztratil svůj přirozený N- i C- konec, které byly nahrazeny kousky proteinů kódovaných rodičovským virem, a získal několik bodových mutací, kteréžto změny z něj vytvořily plně aktivní onkogen v-myb [8].

Oproti AMV je MAV kompletním autonomním retrovirem, který napadené buňky hostitele nezabíjí. Ve vzácném případě, kdy se ve formě proviru náhodně zaintegruje na vhodné místo hostitelského genomu, může dojít k transformaci této hostitelské buňky, která zapříčiní její následný klonální nárůst do stádia makroskopického nádoru. Je-li tedy kuře infikováno směsí virů AMV + MAV dochází k urychlenému rozvoji polyklonální akutní

myelomonoblastické leukémie (v důsledku akutní transformace onkogenem v-myb/AMV) a hostitelské zvíře uhyne do 3 týdnů. Pokud je však kuře infikováno pouze samotným retrovirem MAV, vyvinou se po zhruba 2 – 3 měsících vícečetné avšak monoklonální (původem z jedné buňky) nádory ledvin, případně dalších orgánů.



Obr. 1 Schematické znázornění příbuzenského vztahu mezi replikačně kompetentním latentně onkogenním retrovirem MAV a od něj odvozeným replikačně defektním, avšak akutně onkogenním, retrovirem AMV. Ten vznikl integrací defektního proviru MAV do genomového lokusu kódujícího protoonkogen c-myb a následnou nehomologní rekombinací mezi transkribovanou hybridní mRNA MAV-myb a genomovou RNA MAV-2. MAV-2 i AMV jsou zobrazeny ve své provirové formě (integrované v hostitelské DNA).

Z uvedeného příkladu je patrný odlišný význam akutních a latentních retrovirů pro studium onkogenů. Akutně transformující retroviry aktivně přenášejí onkogeny a jsou tedy ideálním nástrojem pro detailní studium funkcí jednotlivých onkogenů. Oproti tomu jejich latentní protějšky svojí integrací do hostitelského genomu mimikují somatickou mutace. Pokud následkem provirové integrace dojde k onkogenní aktivaci postiženého genu, hostitelská buňka začne proliferovat a vzniká makroskopický klonální nádor (klonální selekcí). Latentně onkogenní retroviry tak zároveň svojí přítomností v klonálně

expandovaných populacích buněk označují v hostitelském genomu nové onkogenní lokusy a jsou proto účinným nástrojem jejich *de novo* identifikace. Pro podrobný rozbor mechanizmů inzerční mutageneze, historii jejich objevů a výčet experimentálních kuřecích modelů odkazuji na přiloženou kapitolu [9] a knihu Retroviruses [10].

Kapitola druhá - Kuřecí modely nádorové transformace založené na inzerční mutagenezi retrovirem MAV-2(N)

Pro disertační práci jsem zvolil studium modelu kuřecích nefroblastomů vyvolaných inzerční mutagenezí retrovirem MAV-2(N). Důvodů bylo několik; jednalo se o zavedený model [11-13], předběžné výsledky slibovaly identifikaci mnoha různých onkogenů naráz a sada nádorových vzorků obsahovala vzácně také nádory původem nejen z ledvin, ale i z jiných tkání. Nabízely se proto dvě varianty - buďto budeme moci v druhém sledu mapovat geny jejichž deregulace vede k metastazování do jiných orgánů (pokud by se jednalo o metastáze) nebo získáme možnost srovnání spektra mutovaných genů v nádorech původem z různých tkání (pokud by se jednalo o vícečetné primární nádory). V dalším textu doložím, že se jednalo o druhou variantu.

Pozorování, že infekce kuřat retroviry MAV-1 nebo MAV-2 indukuje vznik nádorů ledvin bylo učiněno již před více než 30 lety [14]. Oba retroviry se řadí mezi jednoduché pomalu onkogenní retroviry příbuzné se rodinou retrovirů ALV („**A**vian **L**eukosis **V**irus"). S ostatními ALV jsou vysoce homologní s výjimkou unikátní U3 oblasti LTR, kterýžto rozdíl je považován za hlavní příčinu odlišných onkogenních vlastností retrovirů MAV [15] (Pečenka, V., *osobní sdělení*). Mezi sebou se tyto dva virové izoláty liší v sekvenci *env* genu, což má za následek jejich odlišnou hostitelskou specifitu (liší se typem hostitelského receptoru, který využívají pro svůj vstup do buňky; MAV-1 náleží do subskupiny A, MAV-2 do subskupiny B). MAV-2 je znám ze dvou odlišných sub-izolátů MAV-2(N) a MAV-2(O), které oba vyvolávají vznik nefroblastomů ať jsou inokulovány do kuřecích embryí nebo čerstvě vylíhlých kuřat. MAV-2(O) je navíc ještě schopen indukovat akutní osteopetrózu, je-li injikován intraembryonálně [16].

Experimentální model založený na inzerční aktivaci onkogenů retrovirem MAV-2 byl na našem pracovišti zaveden a s přestávkami rozvíjen již řadu let [9, 11-13, 17-19]. Studium MAV-2 indukovaných nádorů potvrdilo klíčový podíl inzerční mutageneze na jejich vzniku (tj. náhodného vestavování provirového genomu MAV-2 do genomu hostitelské buňky). Studium velkého počtu infikovaných zvířat dále potvrdilo méně častý výskyt difúzních lézí na plicích a chrupavčitých nádorových fokusů na játrech infikovaných zvířat. V jediném vzácném případu také na ováriu. Dalším zkoumáním jsme prokázali (hybridizace Southern blotů s MAV-sondou, porovnání integračně-specifických profilů), že u zvířat, u kterých byly diagnostikovány současně nádory v různých tkáních, se nikdy nejednalo o metastáze, ale vždy

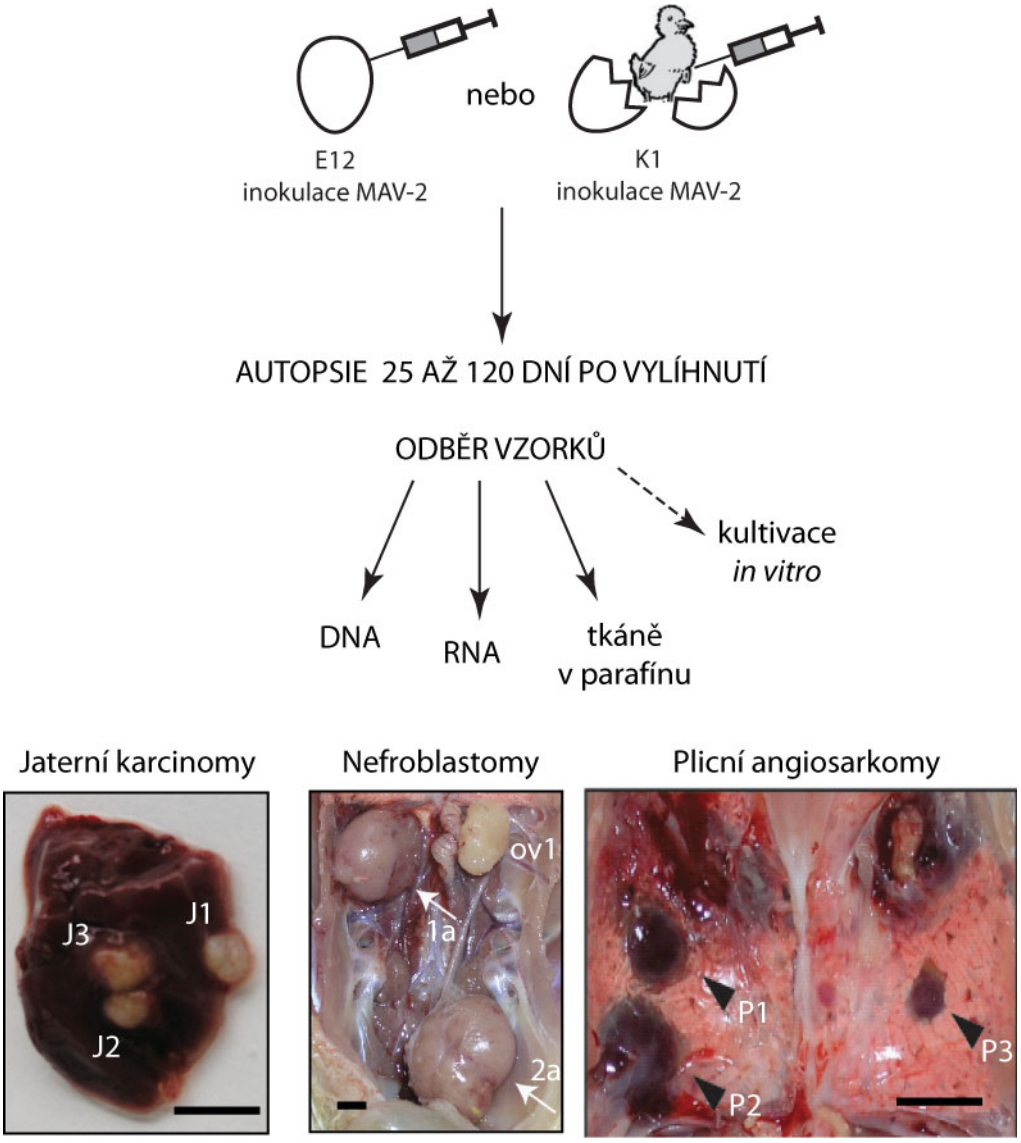
o vícečetné primární nádory. Po tomto zjištění byl původně nefroblastomový projekt rozšířen o mapování obecných míst integrace MAV-2 také na nádory plic a jater.

Kromě hledání obecných míst integrace v různých tkáních jsme v průběhu studia nádorů vyvolaných inzerční mutagenezí MAV-2 učinili ještě další významná pozorování, kterým věnuji separátní kapitoly. Jedná se o spontánní selekci MAV-2-infikovaných buněk ve tkáňových kulturách, tedy jakousi „transformaci *in vitro*“, dále jsme studovali význam struktury integrovaných provirů MAV-2 a jejich defektů pro transformaci hostitelské buňky a naším snad nejzajímavějším pozorováním bylo výrazné zvýšení četnosti plicních a jaterních nádorů pokud byla kuřata injikována nenádorovými buňkami produkujícími MAV-2 v porovnání s prostou infekcí MAV-2. Toto poslední pozorování nás nakonec vedlo k formulování teorie industáze.

V následujících částech shrnu výsledky našich nejvýznamnější pozorování a pokusím se komentovat jejich přínos a potenciál do budoucna. Cílem není detailní seznam a rozbor výsledků jednotlivých experimentů, protože ty jsou shrnuty v přiložených publikacích včetně významných literárních odkazů. Jednotlivé kapitoly také nejsou nutně experimentálně uzavřené, což je důsledkem velkého rozsahu tématu. Tato neuzavřenost popisované problematiky zároveň ukazuje nebývalý potenciál zvoleného modelu, který s každou zodpovězenou otázkou poslušně generuje mnohé další a slibuje skrývat ještě mnohé objevy pro budoucnost.

Kapitola třetí - základní pojmy, obecná charakterizace modelu, použité experimentální přístupy a metodiky

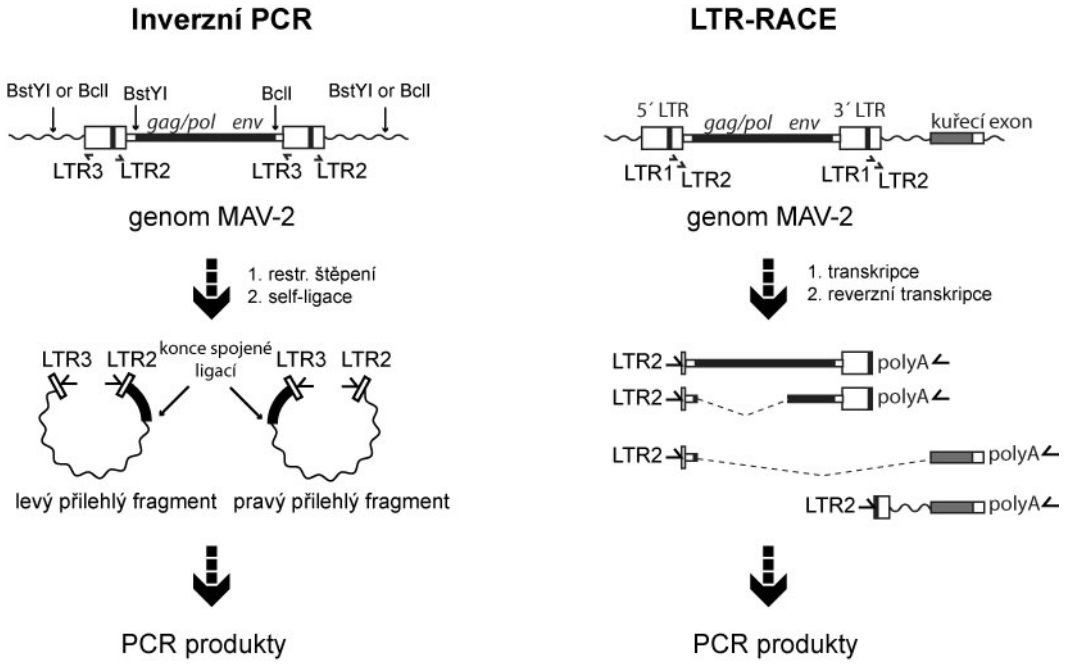
Experimentální strategie je schematicky shrnuta na obr. 2 a detailněji popsána v přiložené publikaci [19]. Virus byl inokulován do 12 dní starých kuřecích embryí (E12) nebo do čerstvě vylíhlých kuřat (K1). Tato experimentální zvířata byla následně zabita ve stáří 25 – 120 dní a pitvána s odběrem tkáňových vzorků. Každý vzorek tkáně (pokud byl dostatečně velký) byl rozdělen na tři kusy pro izolaci genomové DNA, celkové RNA a pro histologickou analýzu parafinových řezů. Příliš malé vzorky (< 3 mm) byly zpracovány pouze jedním ze zmíněných postupů. V některých případech byla část tkáně semisterilně odebrána, opláchnuta fyziologickým roztokem (PBS) a kultivována *in vitro* (viz. Kapitola sedmá).



Obr. 2 Získávání vzorků, použitých v této práci. *Kuřata jsou inokulována MAV-2 popř. buňkami produkujícími MAV-2 ve 12 dni embryogeneze (E12) nebo těsně po vylíhnutí (K1). Zvířata jsou zabita a vzorky odebírány průběžně mezi 25 a 120 dnem po vylíhnutí. Vzorky nádorové tkáně jsou odebírány a ihned zpracovávány pro izolaci celkové DNA, RNA a pro fixaci v parafínových bločcích. Část vzorků byla také semisterilně odebrána pro kultivaci in vitro. Fotografie ve spodní části obrázku ukazují typické příklady jaterních karcinomů (J1, J2, J3), nefroblastomů (1a, 2a) a plicních angiosarkomů (P1, P2, P3). Na prostřední fotografii je zachycen i unikátní ovariální karcinom (ov1). Úsečka v každém obrázku odpovídá 1 cm.*

Vzorky DNA z předpokládaných nádorových tkání byly nejprve testovány na klonalitu a počet integrovaných MAV-2 provirů (Southern blotting a následná hybridizace s MAV-2-

specifickou radioaktivně značenou sondou). Pro další analýzu byly vybírány pouze klonální vzorky s méně než 10 integrovanými proviry. U vybraných vzorků byla následně identifikována klonální místa integrací MAV-2 provirů v genomu hostitelské transformované buňky. K tomu bylo použito dvou odlišných postupů: inverzní PCR (iPCR; vychází z genomové DNA) a LTR RACE (vychází z RNA). iPCR je obecně použitelná robustní metodika vykazující minimální „bias“, která umožňuje identifikaci cca 95% všech integračních míst, avšak minimálně vypovídá o možných důsledcích jednotlivých integračních událostí. Oproti tomu metodika LTR-RACE identifikuje hybridní mRNA transkripty, které obsahují sekvence LTR z MAV-2 a pročítají se nebo jsou sestřihovány do hostitelského genomu. Tyto hybridní transkripty umožňují identifikovat jen asi čtvrtinu integračních míst provirů, ale jsou mezi nimi přednostně místa v nichž jsou vloženy defektní nebo přestavěné proviry. Je známo, že naprostá většina nádorů vzniklých inzerční mutagenezí nese (a vyžaduje) v odpovídajícím obecném místě integrace defektní provirus [9], proto je přítomnost přestavěného proviru pozitivním indikátorem kandidátního onkogenního lokusu. Struktura hybridní mRNA mapuje pravděpodobný způsob onkogenní aktivace hledaného onkogenu a v neposlední řadě může být určující pro identifikaci kandidátního onkogenu. Například tehdy, pokud se integrační místa provirů nachází v hostitelském genomu až stovky kilobazí před postiženým genem a ovlivňují jeho expresi skrze vzdálený splicing (viz. *sox6*, kapitola sedmá - selekce *in vitro*). Detaily metodik iPCR a LTR RACE jsou schematicky znázorněny v obr. 3 a podrobněji rozebrány v přiložené publikaci [19]. Shrnutí výsledků našeho mapování klonálně integrovaných provirů v kuřecích nefroblastomech, jaterních karcinomech, plicních angiosarkomech a klonálních tkáňových kulturách je obsaženo v Tab. 1.



Obr. 3 Princip metod iPCR a LTR RACE.

iPCR vychází z genomové DNA klonálních nádorů obsahující integrované MAV-2 proviry. Po restrikčním štěpení a cirkularizaci jsou pomocí PCR amplifikovány fragmenty přiléhající k pravému i levému LTR proviru. Použité restrikční endonukleázy jsou voleny s ohledem na stejnou pravděpodobnost amplifikace fragmentů přiléhajících k oběma LTR.

LTR-RACE je modifikací klasické techniky 3´- RACE. Vychází z celkové RNA klonálních nádorů a amplifikuje hybridní mRNA mající na 5´-konci MAV-2 LTR. Takto lze identifikovat hybridní mRNA vzniklé aberantním splicingem mezi MAV-2 a hostitelským genomem nebo takové, které vzniknou pročítáním slabého polyadenylačního signálu MAV-2 do hostitelského genomu (až ¼).

Vlastní vyhledávání kandidátních protoonkogenních genomových lokusů je založeno na předpokladu, že klonálně integrované proviry jsou v těchto lokusech nalezeny opakovaně v nezávislých nádorech. Tyto lokusy jsou označovány jako obecná místa integrace (**C**ommon **I**ntegration **S**ite, CIS) a je vysoká pravděpodobnost, že se v jejich okolí nachází protoonkogen, případně tumor suppressor. Se zdánlivě jednoduchou definicí CIS jsou spojeny dva problémy. Jednak jak definovat CIS a druhak kolik nezávislých nádorů je třeba analyzovat pro spolehlivé mapování genů, které se podílejí na vzniku studovaného

nádorového typu. Úskalí obou problémů jsou prakticky patrná z výsledků popsaných v následujících kapitolách, proto zde uvedu pouze stručný teoretický rozbor.

Fyzická velikost jednotlivých CIS v genomu může být velmi odlišná (od několika desítek bazí po stovky tisíc bazí). Prakticky tedy mohou být CIS definovány na základě opakované integrace proviru buď do téhož genu nebo do arbitrárně určené délkové oblasti genomu (např. 20 kbp). Je zřejmé i bez matematického aparátu, že pravděpodobnost opakované náhodné inzerce proviru do genu o velikosti 200 bp bude výrazně nižší než do genu o délce 2 x 10^6 bp, jinými slovy – pokud najdeme dvakrát integrovaný provirus v malém genu, je daleko větší pravděpodobnost, že se jedná o protoonkogen, než v případě dvou vzdálených integračních událostí v obřím genu. Oproti tomu faktor prosté fyzické vzdálenosti dvou a více integračních míst v genomu je závislý na celkovém počtu analyzovaných integračních míst [20]. Množství vzorků a integračních míst, které je třeba analyzovat pro spolehlivé mapování zúčastněných genů je dále závislé na množství různých protoonkogenů, jejichž deregulace/mutace může způsobit vznik zkoumaného nádoru a tedy četnosti, s jakou jsou v jednotlivých typech nádorů zasahovány (srovnej „kapitola čtvrtá – nefroblastomy“ a „kapitola pátá - plicní angiosarkomy“).

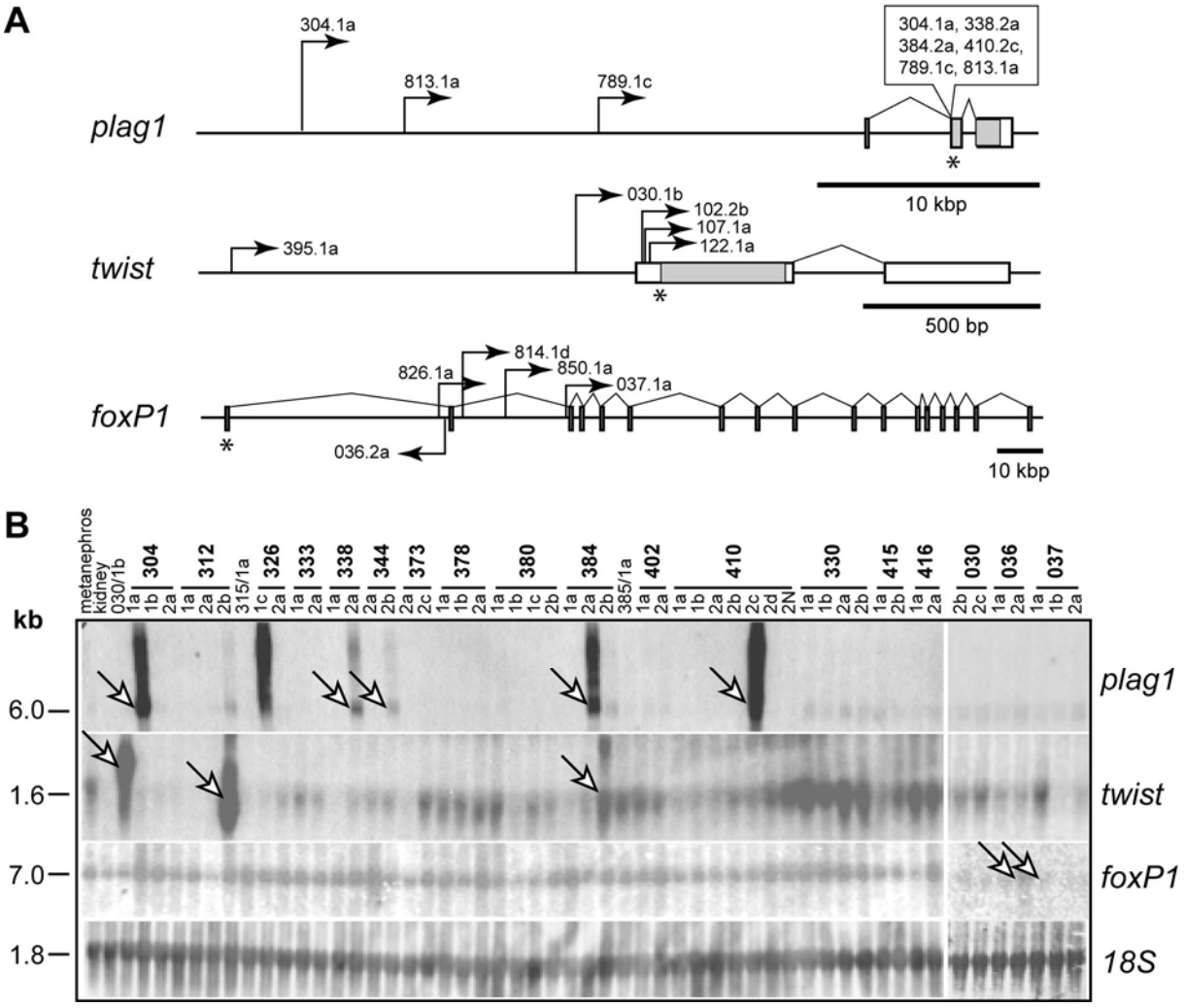
Zvolili jsme proto kombinovaný přístup; vytipovali jsme kandidátní CIS na základě minimálního požadavku nejméně dvou nezávislých integrací proviru v rozmezí 20 kbp nebo nejméně dvou nezávislých integrací proviru v anotovaném genu. Pro všechna tato kandidátní CIS jsme následně hodnotili další parametry jako celkový počet integrací, poloha v rámci genu, orientace proviru vzhledem ke kandidátnímu genu a také výsledky následných analýz exprese kandidátního genu (pokud byly provedeny). Kandidátní CIS byly seřazeny v Tab. 2 podle relevance (tj. výše dosaženého skóre, viz. legenda k Tab. 2).

Následné analýzy kandidátních genů jsou umožněny spektrem materiálu, který odebíráme z experimentálních zvířat (DNA, RNA, tkáně v parafínu). To umožňuje zpětně hodnotit hladinu exprese a strukturu mRNA kandidátních genů (RNA), strukturu a přestavby integrovaných provirů (DNA) i tkáňovou distribuci proteinů a vztah postiženého genu a tkáňové histologie (tkáňové řezy). Pro immunohistochemickou analýzu velkého množství vzorků tkání naráz jsme úspěšně zavedli techniku tkáňových mikroarrayí, umožňující zpracovávat až 1000 různých tkání na jednom mikroskopickém sklíčku [21].

Kapitola čtvrtá - nefroblastomy

Nefroblastomy jsou nádory embryonálního typu, postihující kuřata jako následek časné infekce MAV-2. Pravděpodobnost výskytu nefroblastomů u infikovaných kuřat se blíží 100%. Podle mikroskopické analýzy jsou obdobou lidských nefroblastomů – Wilmsových nádorů – které jsou nejčastější dětskou malignancí, postihující zhruba jedno dítě z deseti tisíc. Histologická analýza, klonalita, postup hledání CIS a jednotlivých kandidátních genů v kuřecích vzorcích jsou podrobně rozebrány v přiložené publikaci [19]. Oproti stavu z roku 2006 došlo ke dvěma významným změnám – navýšili jsme počet zmapovaných integračních míst v kuřecím genomu na 1260 a byl zveřejněn druhý draft kuřecího genomu (WASHUC2), což společně vedlo k preciznější identifikaci CIS a odpovídajících kandidátních genů. Aktualizované výsledky jsou shrnuty v Tab. 1 a 2.

Jedinečnost nefroblastomového modelu spočívá v neexistenci dominantního obecného místa integrace. Přes značnou podobnost morfologickou i histologickou, již první popsaná CIS naznačovala odlišné molekulární mechanismy transformace lidských a kuřecích zárodečných ledvinných buněk [15, 18]. Další plošné mapování integrovaných provirů tyto předpoklady potvrdilo. Lokusem nejčastěji postihovaným provirovou integrací je *plag1* – gen kódující transkripční faktor, jehož přestavby v lidském genomu zapřičiňují vznik pleiomorfních adenomů, lipoblastomů, hepatoblastomů a akutních myeloidních leukemiích [22]. Přesto pouze 7% nefroblastomů obsahuje integrovaný provirus právě v tomto lokusu. Frekvence integrace MAV-2 do ostatních CIS je ještě nižší; *foxP1* (5%), *twist* (4%), *Ha-ras* (4%) – obr. 4.



Obr. 4 Vybraná obecná místa provirové integrace MAV-2 v nefroblastomech.

A) *Grafické znázornění polohy a orientace integrovaných MAV-2 provirů v kuřecích genomových lokusech plag1, twist, foxP1. Exony genů jsou označeny bílými obdélníky, kódující oblasti, šedými; poloha iniciačního kodónu je označena hvězdičkou. Polohy a orientace jednotlivých provirů je označena šipkami. Čísla v obdélníku nad druhým exonem plag1 představují vzorky u nichž byl prokázán sestřih virové RNA do tohoto exonu.*

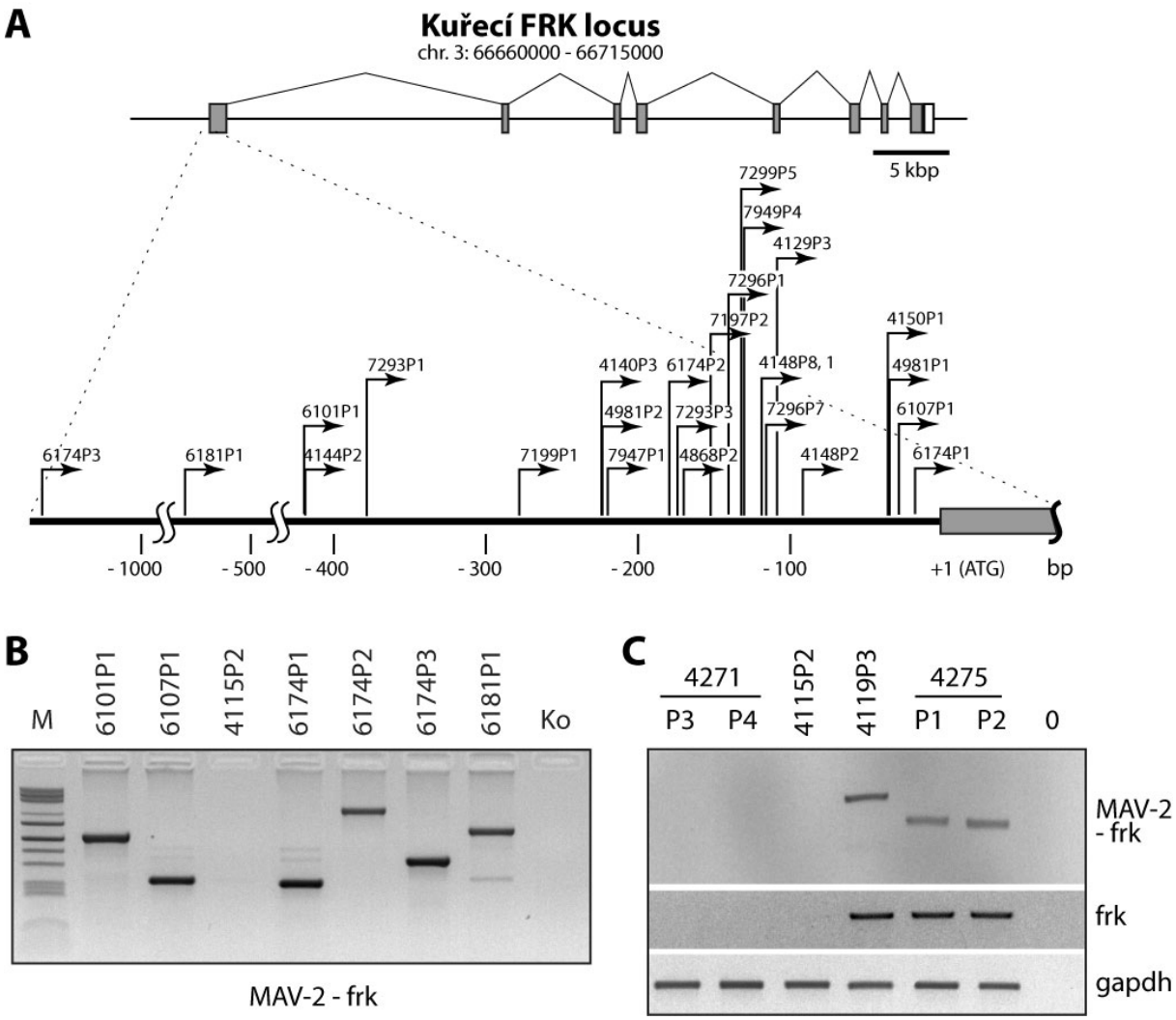
B) *Studium deregulace exprese genů plag1, twist, foxP1 hybridizací northern blotů se sondami značenými ^{32}P. Vzorky nesoucí integraci do studovaného lokusu jsou označeny šipkami. Jako kontrola množství a kvality RNA v lince je použito barvení 18S rRNA.*

Lidské orthology všech těchto genů jsou také považovány za onkogeny, avšak jejich mutace jsou u lidí spojovány s odlišnými typy nádorů či jejich vývojových stádií (např. foxP1

je kauzálně spojován s difůzním velkobuněčným lymfomem z B-buněk [23], deregulace genu *twist* je u lidí nečekaně spojována s epitelo-mesenchymovým přechodem buněčné diferenciace a tedy přímo se získáním invazivně-metastatického potenciálu transformované buňky [24] atd.). Oproti tomu, integrace do kuřecího orthologu lidského genu *wt1*, jehož mutace je spojována se zhruba 10% případů Wilmsových nádorů, nebyla vůbec zaznamenána. Jinými slovy, vysvětlení molekulárního mechanismu vzniku kuřecích nefroblastomů není jednoduché, stejně jako vysvětlení jejich příbuznosti s lidským nefroblastomem. Přesto je zřejmé, že tento experimentální model v sobě skrývá potenciál odhalit desítek buněčných protoonkogenů, jakož i zajímavé srovnání zjevně odlišných molekulárních mechanismů transformace lidských a kuřecích ledvinných buněk.

Kapitola pátá - plicní angiosarkomy

Při pitvách infikovaných zvířat byly s nízkou frekvencí nalézány také plicní hemorrhagické léze. Immunohistochemická analýza řezů těchto tkání určila jejich diagnózu jako angiosarkomy. Jejich incidence se významně zvyšuje při současné nebo následné inokulaci infikovaných zvířat volnými buňkami (viz. Kapitola osmá). Morfologicky jsou tyto plicní angiosarkomy buď lokalizované v jednotlivých ložiscích nebo agresivně kolonizují jednu či obě plíce (v těchto případech byly transformované buňky snadno molekulárně detekovány v krevním oběhu a dalších orgánech). Jak jsme prokázali, tyto nádory jsou klonálního původu, nejsou metastázemi z nefroblastomů, ani nepochází z injikovaných buněk. Analýza integračních míst MAV-2 provirů vedla k identifikaci jediného majoritního CIS, lokusu kódujícího nereceptorovou tyrosinovou kinázu *frk*, která je exprimována převážně v epiteliálních tkáních [25, 26] a předpokládá se, že se účastní v chondrogenezi a ve vývoji Langerhansových ostrůvků [27]. Tento lokus je postižen provirovou integrací ve více než 95% vzorků, které jsme zkoumali. Provirové integrace jsou lokalizovány těsně nad iniciačním kodónem a mají za následek vysokou nadexpresi *frk* mRNA (obr. 5). Detaily našich experimentů jsou popsány v přiloženém článku [17].



Obr. 5 Dominantní obecné místo integrace MAV-2 v kuřecích plicních angiosarkomech.

A) *Grafické znázornění polohy a orientace integrovaných MAV-2 provirů v kuřecím genomovém lokusu frk. Značení je stejné jako na obr. 4*

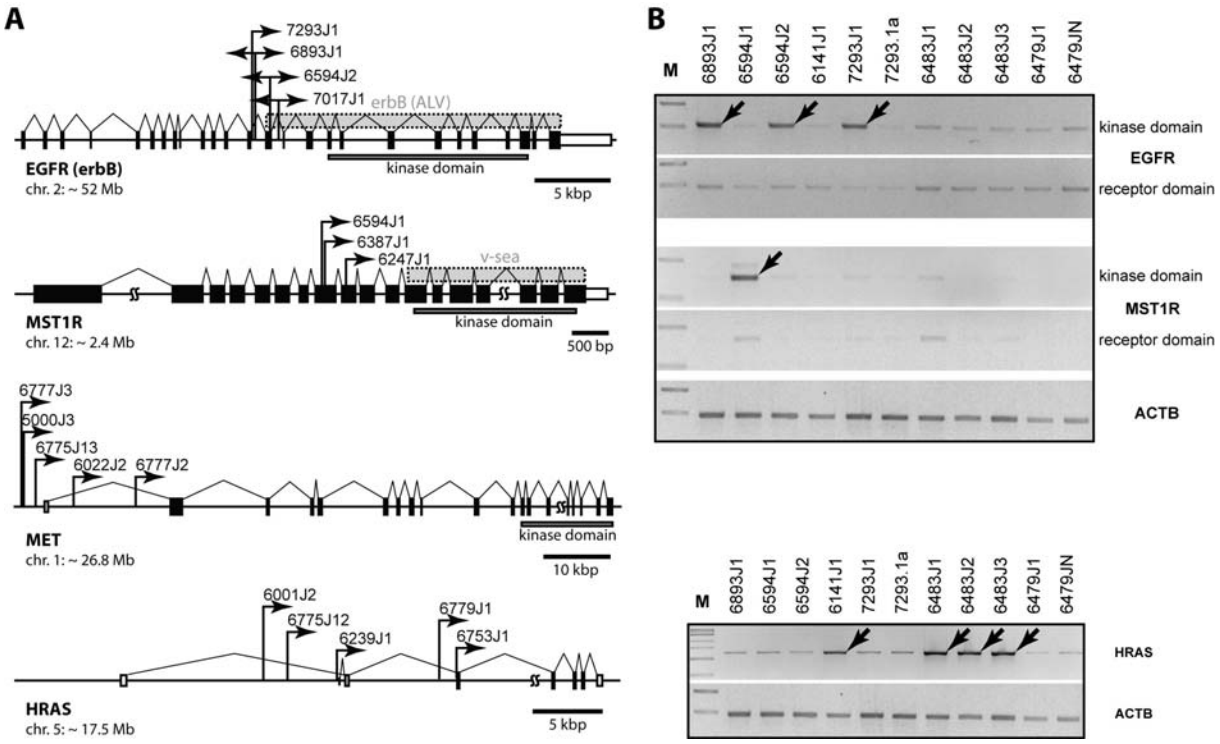
B) *Exprese hybridní mRNA MAV-2-frk ve vzorcích plicních angiosarkomů. Odlišná délka RT PCR produktu reflektuje vzdálenost partikulárního integračního místa od iniciačního kodónu frk (antisense orientovaný frk-specifický primer se nachází v prvním exonu genu, pod iniciačním kodónem).*

C) *Nadexprese frk mRNA je přímým důsledkem integrace MAV-2. RT PCR s primery MAV-2 a frk-specifickým, stejně jako s dvěma frk-specifickými primery dává produkt pouze ve vzorcích nesoucích integrovaný provirus v tomto lokusu. Jako kontrola byly použity primery specifické pro kuřecí gen gapdh.*

Nadexprese lidského orthologního genu *frk* nebyla doposud přímo spojována s onkogenní transformací, možná také proto, že lidské angiosarkomy jsou velmi vzácné a proto i nedostatečně molekulárně charakterizované. Do budoucna bude tedy velmi zajímavé stanovit hladinu exprese genu *frk* v různých lidských malignancích například za použití imunohistochemické analýzy tkáňových mikroarrayí a prokázat tak aktivní onkogenní roli deregulované exprese *frk* v lidské tumorigenezi, která je doposud pouhou spekulací [28].

Kapitola šestá - jaterní karcinomy

Dalším typem nádorů, vyskytujícím se s nízkou frekvencí (cca 5%) u kuřat infikovaných MAV-2 jsou jaterní léze, velké obvykle 3 – 5 milimetrů. Tyto byly immunohistochemickou analýzou hodnoceny jako jaterní karcinomy (biliární karcinomy, adenokarcinomy a angiosarkomy). Podobně jako nefroblastomy a plicní sarkomy jsou klonálního původu a jedná se o lokalizované primární nádory, nikoliv o metastáze z jiných orgánů. Analýza integračních míst MAV-2 provirů vedla k identifikaci čtyř majoritních CIS, z nichž ve většině zkoumaných vzorků je vždy jeden zasažen provirovou integrací (Tab. 1 a 2). Jedná se o protein kinázy *egfr* (*erbB*), *mst1r* (*ron*), *met* (*hgfr*) a GTPasu *Ha-ras* - obr. 6. Mapovaná integrační místa provirů a předpokládané důsledky v podobě deregulace cílového genu jsou ve shodě s předchozími pozorováními na kuřecích modelech (*v-erbB, ron*), ale překvapivě jsou ve velmi dobré shodě i s daty shromážděnými při studiu myších a lidských jaterních karcinomů (*Ha-ras*, *erbB*, *met*).



Obr. 6 Obecná místa provirové integrace MAV-2 v jaterních karcinomech.

A) *Grafické znázornění polohy a orientace integrovaných MAV-2 provirů v kuřecím genomovém lokusu egfr, mst1r, met a hras. Značení je stejné jako na obr. 4 U prvních tří genů jsou označeny genomové oblasti kódující jejich kinázové domény, u genů egfr a mst1r jsou též označeny genomové oblasti kódující jejich dříve popsané onkogenní varianty (erbB, v-sea). Z důvodu přehlednosti nejsou zobrazena všechna námi identifikovaná integrační místa. Obousměrné šipky v místech integrací do egfr lokusu znázorňují komplexně přestavěné defektní proviry – viz. Kapitola devátá a obr. 9c.*

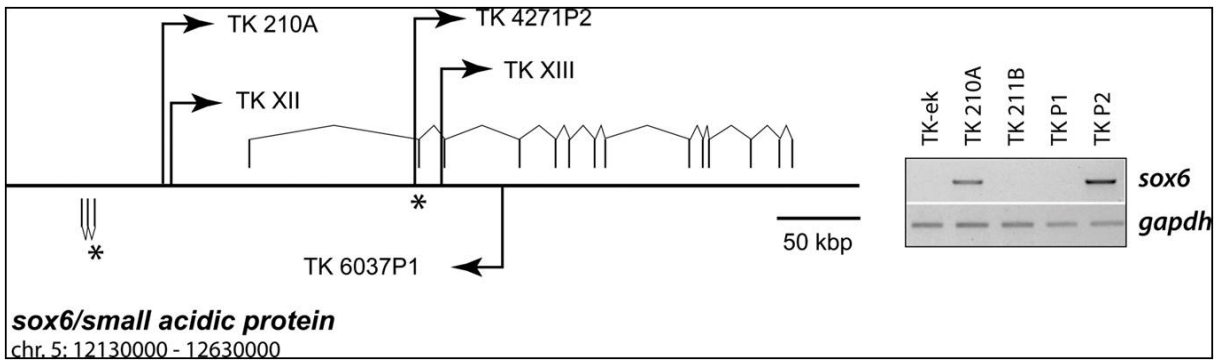
B) *RT PCR analýza hladiny exprese egfr, mst1r (kinázová a receptorová doména) a hras v jaterních nádorech. Vzorky nesoucí integraci do studovaného lokusu jsou označeny šipkami. Jako kontrola byly použity primery specifické pro kuřecí beta aktin (actb).*

Naše pozorování tedy ukazují, že jaterní karcinomy jsou velmi pravděpodobně příkladem obecných molekulárních mechanismů transformace cílových buněk, které fungují velmi podobně i u evolučně distantních zvířecích druhů. Popsaný kuřecí model rozšiřuje spektrum zvířecích modelů, na nichž je studována transformace jaterní buňky a umožňuje tak srovnání obecných a druhově-specifických charakteristik tohoto procesu. Jelikož jsou jaterní karcinomy pátým nejčastěji se vyskytujícím typem nádoru u lidí a třetí nejčastější příčinou

úmrtí pacientů s nádorovými onemocněními, je význam dalšího studia zjevný. Výsledky těchto pozorování jsou v současnosti odeslány k publikování [29].

Kapitola sedmá - selekce *in vitro*

Pro zjednodušení studia následků integrace proviru do jednotlivých kandidátních CIS lokusů jsme se rozhodli převést vybrané nádorové vzorky do tkáňových kultur a pokusit se od nich odvodit stabilní buněčnou linii. Souběžně jsme se pokoušeli izolovat buněčné klony s růstovou výhodou, vzniklé *in vitro* infekcí kuřecích fibroblastů. Primární tkáňové kultury prošly podle předpokladu růstovou krizí (dosažení Hayflickova limitu) a v některých případech po krátké době obnovily růst. Studium několika takových tkáňových kultur nás vedlo ke třem zásadním zjištěním. Tyto kultury mají výrazně prodloužený lifespan, ale nejsou immortalizované. Často jsou výrazně klonální. Pokud byly odvozeny od klonálního nádoru, spektrum jejich integrovaných provirů je naprosto odlišné od mateřského nádoru. Mapování integračních míst MAV-2 provirů vedla ke zjištění, že zhruba polovina tkáňových kultur sdílí integraci ve společném rozsáhlém CIS - v genu *sox6* (obr. 7, tab. 1 a 2). Následkem těchto provirových inzercí je silně nadexprimovaná hybridní mRNA MAV-2-*sox6*. *Sox6* patří do silně konzervované rodiny (Sry)-příbuzných transkripčních faktorů a jeho popisované funkce spadají spíše do oblasti diferenciace buněk a vývoje organismu [30]. Pozorování je o to zajímavější, že mezi více než 1500 nezávislými integračními místy původem z nádorů *in vivo* ani jedno neleží v tomto genu, přestože se sox6 CIS rozpíná přes více než 500 kbp v kuřecím genomu.



Obr. 7 Obecné místo integrace MAV-2 ve tkáňových kulturách kuřecích buněk infikovaných MAV-2.

Grafické znázornění polohy a orientace integrovaných MAV-2 provirů v kuřecím genomovém lokusu sox6. Značení je stejné jako na obr. 4. Vpravo je příklad elektroforetické analýzy RT PCR produktů s MAV-2-specifickým (sense) a sox6-specifickým (antisense) primery, ukazujícím nadexpresi hybridní mRNA ve vzorcích s integrací MAV-2 proviru v tomto lokusu (TK 210A, TK P2). Jako kontrola byly použity primery specifické pro kuřecí gen gapdh.

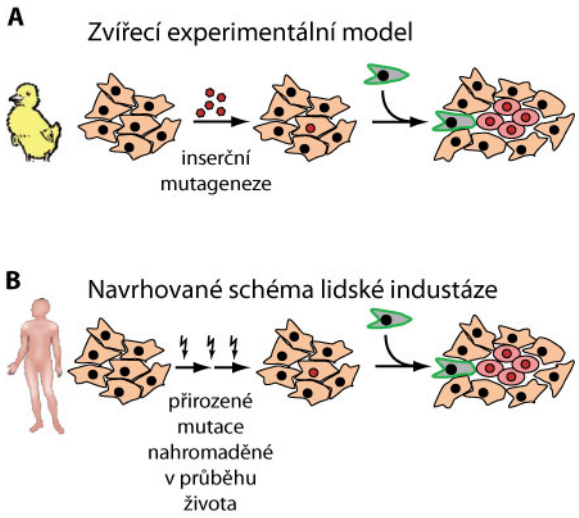
Interpretace tohoto jevu není jednoduchá a vyžaduje podrobnější výzkum, avšak poukazuje na neočekávaný „bias" při kultivaci primárních transformovaných buněk *in vitro*. Immortalizace inzerční mutagenezí *in vitro* byla popsána také u myších modelů [31]. Podobně je známo, že ektopická exprese některých onkogenů v odlišných buněčných typech vede k různým transformovaným fenotypům [32]. Naše pozorování upozorňuje na limitace studií prováděných na immortalizovaných buněčných liniích. Výsledky jsou připravovány k publikování.

Kapitola osmá - fenomén industáze

Jevem, který jsme objevili a který považujeme za nejvýznamější výsledek našeho zkoumání, je fenomén industáze (z lat. indūcere - tahat a řec. stasis – rovnováha). Jedná se o mechanismus nádorové promoce, který se uplatňuje u jednotlivých preneoplastických buněk (iniciované buňky) nebo buněčných kompartmentů (kancerizovaná pole), které již naakumulovaly potřebné kancerózní mutace, ale stále ještě zůstávají pod kontrolou tkáňové homeostáze. Bludné buňky mají schopnost narušit lokální mikroprostředí a tím porušit tuto kontrolu. To vede k plnému projevení maligního charakteru iniciované buňky. Praktické efekty tohoto jevu jsme pozorovali při studiu plicních angiosarkomů (Kapitola pátá) – kuřata infikovaná viriony MAV-2 vykazovala výskyt plicních angiosarkomů pouze asi v 5%, avšak při infekci MAV-2-produkujícími netransformovanými fibroblasty došlo k nárůstu frekvence výskytu plicních angiosarkomů až k 60%. Po prokázání podobné míry a kinetiky proinfikování inokulovaných zvířat a vlivu pořadí inokulace kuřat (viriony vs. neinfikované bludné buňky) jsme formulovali experimentální koncept industáze (obr. 8A). Detailní popis provedených experimentů viz. [17].

Předpokládáme, že u lidí může být tento mechanismus zodpovědný za vznik mnohočetných primárních nádorů tzv. nádorových duplicit. Somatické buňky akumulují

mutace v průběhu celého lidského života. V případě, že se v organismu vyvine primární nádor, je tělo zaplaveno bludnými buňkami uvolňovanými z tohoto primárního nádoru. Tyto buňky jsou považovány především za rizikový faktor vzniku metastáz, přestože většina z nich schopnost metastazovat postrádá. Tyto buňky však mohou ve svém množství představovat riziko poškozování lokální homeostatické regulace iniciovaných prekancerózních buněk (obr. 8B). Důvodů, proč tento mechanismus může snadno unikat pozornosti je hned několik, které se sčítají. 1) relativní vzácnost vícečetných primárních nádorů 2) možnost chybné diagnózy některých primárních nádorů jako metastázy 3) terapeutický – vícečetné primární nádory jsou diagnostikovány většinou u pacientů, kteří po první diagnóze projdou (chemo-, radio-) terapií, která je sama o sobě mutagenní, čemuž lze zvýšený výskyt následných primárních nádorů přičíst 4) genetický – pacienti s vícečetnými primárními nádory jsou často nositeli predisponující mutace, která výrazně zvyšuje riziko vícečetných primárních nádorů.



Obr. 8 Koncept industáze.

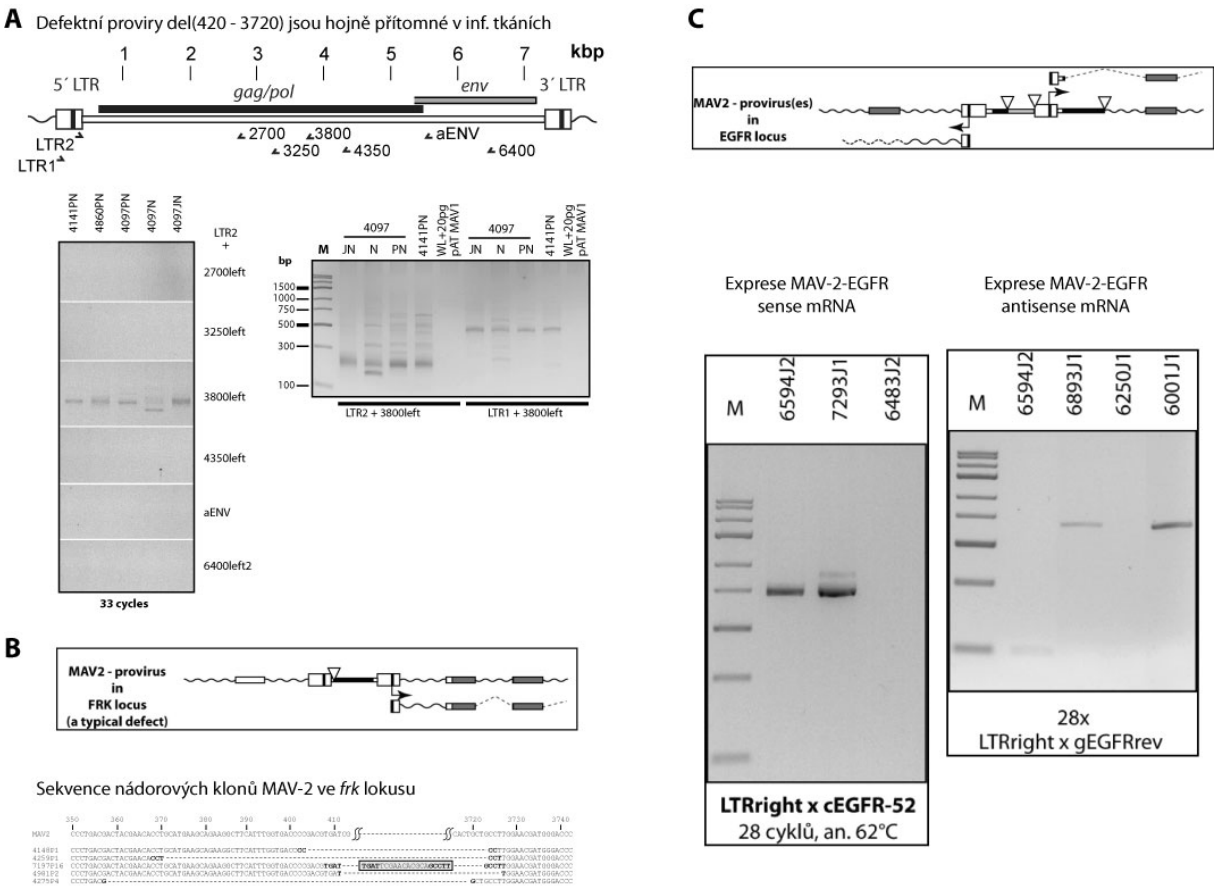
A) *U kuřecího modelu vede inzerční mutageneze MAV-2 (červené šestiúhelníčky) ke vzniku prekancerózně iniciované buňky (červené jádro), což je okamžitě následováno tumor-promočním působením injikované bludné buňky (šedá), což vede k nárůstu transformovaného buněčného klonu (červené oválné buňky)*

B) *V námi navrhovaném obecném schématu industáze se mutace v buňkách hromadí v průběhu života individua a vedou ke vzniku geneticky transformované iniciované buňky. Pokud se do její blízkosti dostane bludná netumorigenní buňka uvolněná kupříkladu z jiného primárního nádoru, tato ovlivní mikroprostředí iniciované buňky a odstartuje její maligní růst.*

Kapitola devátá - přestavěné proviry

Nejčastější příčinou inzerční onkogenní aktivace protoonkogenu je inzerce defektního proviru, který oproti intaktnímu výrazněji nebo specifičtěji ovlivňuje přilehlé hostitelské sekvence [9]. Prokázali jsme, že kuřecí nefroblastomy obsahují významný podíl klonálně integrovaných defektních provirů [12] a že proviry integrované v protoonkogenním lokusu *twist* jsou výrazně přestavěné, což má za následek silnou nadexpresi tohoto lokusu [18]. Detailní molekulární mapování integrovaných provirů v ostatních CIS lokusech odhalilo přítomnost různých defektů a přestaveb v naprosté většině zkoumaných případů. Navíc se zdá, že tyto provirové defekty jsou charakteristické pro jednotlivá CIS, pravděpodobně proto, že různé typy přestaveb mají za následek různou míru exprese přilehlých hostitelských genů.

Je známo, že většina defektních provirů vzniká v průběhu reverzní transkripce. Ukázali jsme, že většina takto „přirozeně“ vzniklých provirů obsahuje rozsáhlou interní deleci začínající těsně za 5´- LTR a končící okolo bp 3720 (obr. 9A). Přesně tento typ přestavěných provirů jsme nalezli v plicních angiosarkomech integrované v genu *frk* (obr. 9B). Oproti tomu, defektní proviry integrované v genomových lokusech *twist* nebo *egfr* vykazují rozsáhlé a komplexní přestavby. Tyto aberace mají v případě *egfr* za následek nejen zvýšenou expresi onkogenní varianty EGFR, ale dokonce i expresi potenciální antisense RNA proti druhé, normální, alele tohoto genu (obr. 9C). Pravděpodobnost vzniku takto složitých přestaveb je nízká a mechanismus jejich geneze (jednokrokový či vícekrokový, během replikace či post-replikační) je neznámý. Je možné, že některé typy molekulární transformace (např. zmíněná inzerční aktivace *egfr*) vyžadují vzácně se vyskytující přestavěný provirus, což se projevuje nízkou frekvencí výskytu jaterních karcinomů, případně jejich delší latencí.



Obr. 9 Analýza struktury defektních provirů MAV-2

A) *PCR průkaz přirozeného preferenčního výskytu defektních MAV-2 provirů s delecí cca mezi bp420-3720 v infikovaných neklonálních kuřecích tkáních. (PN – plíce, JN – játra, N – ledviny), jako negativní kontrola byl použit v bakteriích klonovaný MAV-1 ve směsi s neinfikovanou kuřecí DNA (WL + pAT MAV1). Schematické znázornění proviru (nahoře) ukazuje pozici a orientaci použitých primerů - číslo odpovídá lokalizaci bp - v MAV2.*

B) *Struktura MAV-2 provirů v klonech plicních angiosarkomů. Schema ukazuje typický defektní MAV-2 provirus integrovaný v kuřecím lokusu frk (nahoře). Příklady pěti sekvencí MAV-2 v těchto nádorech. Je patrné, že se nejedná o ostrý deleční hotspot, neboť jednotlivá místa rekombinací jsou odlišná. Sekvence v rámečku odpovídá vrekombinovanému zbytku sekvence* $tRNA^{Trp}$*, která iniciuje reverzní transkripci ve virionu.*

C) *Schematický příklad struktury komplexně přestavěného MAV-2 proviru v lokusu egfr a z něj exprimovaných hybridních transkriptů (nad a pod šipkami, které značí počátek a směr transkripce); trojúhelníčky značí delece, tmavé úseky inverze. Panely pod schematem ukazují*

RT PCR průkaz exprese ve schematu vyznačených sense a antisense mRNA (negativní vzorky 6483J2 a 6250J1 jsou kontroly bez integrace v lokusu egfr).

Tato část naší práce dosud není uzavřena, přesto získané výsledky jasně potvrzují význam defektních virů pro inzerční mutagenezi a poukazují na specifickou přítomnost různě defektních provirů v jednotlivých CIS. Nastolují také otázky doby a mechanismu vzniku komplexních přestaveb pozorovaných v některých lokusech, ale i možnost korelace četnosti výskytu jednotlivých nádorových typů s pravděpodobností výskytu odpovídajícího přestavěného proviru.

Kapitola desátá – Shrnutí poznatků získaných studiem ptačího modelu a jejich začlenění do kontextu zkoumání nádorové transformace

Sumarizovat poznatky získané během několika let studia je překvapivě jednoduché, přesto se nemohu vyhnout pohledu ze dvou různých úhlů. První pohled vyjadřuje naprostou spokojenost s výsledky studia úžasného experimentálního modelu. Molekulárně jsme popsali několik nádorových typů (ledvinné, plicní a jaterní), „pseudotransformované buňky *in vitro*“, naznačili jsme roli defektních provirů v buněčné transformaci a objevili jsme zcela nový fenomén promoce nádorového růstu bludnými buňkami. Výsledky jednotlivých kapitol nejsou samoúčelné, ale mohou být začleněny do kontextu nádorového výzkumu jak na ostatních zvířecích modelech, tak i na lidských nádorových vzorcích. Každé ze zmíněných témat otevírá samostatné pole pro další výzkum. Jestliže tedy charakteristikou dobrého a plodného výzkumného projektu je to, že s každou odpovědí nastoluje ještě více dalších otázek, pak myslím, že jsme uspěli.

Druhý úhel pohledu vykresluje odlišnou perspektivu. Pakliže jsem na počátku svého studia viděl obecné řešení problému nádorové transformace v mapování onkogenních a tumor supresorických mutací, na konci studia jsem přesvědčen, že význam genetických změn v transformované buňce je přeceňován. Můžeme se kupříkladu pozastavit nad zjištěním, že námi popsané spektrum jaterních onkogenů je identické u kuřat i lidí, ale zásadně se liší u nefroblastomů. Domnívám se, že tradiční představa vzniku nádoru jako funkce hromadění mutací a jejich selekce je v současnosti významnou brzdou dalšího výzkumu. Přestože popisování funkce jednotlivých genů, hledání nádorových markerů a molekulárních cílů pro

terapii je stále velmi přínosné, domnívám se, že je třeba výrazně měnit pohled na problematiku spojitých nádob evoluce-smrtelnost-nádorová transformace. Proto jsem se rozhodl na závěr zařadit ještě jednu spíše spekulativní nežli prakticky-experimentální knihu, která však odráží výsledky mého několikaletého studia a snad i nastiňuje možnosti nových experimentálních přístupů.

Část druhá - Fragment knihy chaosu

Kapitola první – o chaosu

Lidé rozličných kultur a náboženství ve většině případů chápali vztah „řádu“ (stav uspořádaný) a „chaosu“ (stav neuspořádaný) jako nesmiřitelné protiklady, analogicky jako třeba dualitu dobra a zla. Obě dvojice pojmů bývají v mnoha kulturách míseny a ztotožňovány. Ostatně, podobně vnímá význam těchto pojmů většina současných lidí, včetně vědců. Takovéto (dualistické) vnímání je praktickým zjednodušením, ovšem také nešťastným omylem, chceme-li řešit otázky fungování komplexnějších systémů. Chaos – deterministický chaos - je přesnější chápat ve smyslu fyziky a matematického modelování. Chceme-li kupříkladu dělit fyzikální systémy z hlediska míry vnitřní uspořádanosti, nabízí se nám jednoduché dualistické dělení na uspořádané a neuspořádané. Avšak reálnější dělení systémů by mělo tři póly: uspořádané, chaotické a neuspořádané. Hlubším zamyšlením nad tímto rozdělením dojdeme k názoru, že ve skutečném světě nejsou žádné striktně uspořádané ani neuspořádané systémy. Zbývá tedy uniformní heterogenně chaotický svět. Živé systémy mohou existovat právě jen na bázi chaosu, protože systémy uspořádané i neuspořádané představují pouze dvě odlišné, imaginární formy neživé hmoty. Ve skutečném světě tedy existují systémy s různou mírou chaotičnosti, čímž se dostáváme zpět k původnímu duálnímu pojetí – reálný chaotický svět si proto nahrazujeme (myšlenkově přijatelnějším) světem shora zmíněných (avšak nešťastně nazvaných) protikladů „řádu“ a „chaosu“. S těmito znalostmi se je tedy můžeme pokusit lépe opsat jako (chaotický systém s menší mírou vnitřní neuspořádanosti, tedy se slušnou mírou predikovatelnosti jeho chování) vs. (chaotický systém s větší mírou vnitřní neuspořádanosti, tedy obtížně predikovatelný). Rozdíl tedy není mezi „dobrým řádem“ a „zlým chaosem“, ale mezi mírou předvídatelnosti budoucího chování studovaných systémů.

Úvodní slovní cvičení nás přivádí ke konstatování, že vytváříme modely okolního světa, které nám umožňují co nejlépe předvídat jeho chaotické chování, a tyto modely jsou za různých situací různě úspěšné v závislosti na míře stability modelovaného systému v daném období. Jasně patrné je to kupříkladu na společenské úrovni. V dobách míru, socioekonomické stability a prosperity je patrný nárůstající příklon k materialistickému (vědeckému) pojetí světa, protože takové pojetí nám vcelku úspěšně pomáhá předvídat jeho budoucí chování s nízkou mírou nestability. Naopak v dobách nedostatku, krizí, pohrom a válek nabírá na významu duchovní vnímání světa reprezentované některou formou víry

(kupříkladu náboženské, ale třeba i s extremistickou politickou orientací) a hledání proroků. A opravdu se zdá, že prostá víra v turbulentních časech umožňuje lepší predikci budoucího vývoje. To nikterak nestaví víru na roveň rezignaci. V časech, kdy opakovaně selhávají i krátkodobé rigorózní předpovědi má a bude mít víra nezastupitelnou úlohu v životě jedince i společnosti a výsledky jejích predikcí budou prokazatelně lepší. Z toho dovozuji, že by naší snahou mělo být lépe chápat chování chaotických systémů a především důsledků jejich ovlivňování vnějšími vlivy (srovn. tzv. „efekt motýlího křídla“, který se může projevit jen za určitých podmínek).

Nejvýraznějšími a snad nejúžasnějšími vlastnostmi deterministického chaosu jsou škálovatelnost (z tvaru mraku nelze usuzovat na jeho velikost, z tvaru křivky nelze odvodit měřítka jednotlivých os), soběpodobnost (stejné vzorce chování se opakují v jednom systému v mnoha škálách) a univerzalita (chování naprosto nepříbuzných systémů se řídí týmiž zákonitostmi; například vývoj cen na burze, počasí, geopolitické dějiny, evoluce živých organismů (záměrně se vyhýbám užití pojmu „darwinistická“) ap.). Z předchozího shrnutí pozoruhodných vlastností chaotických systémů také vyplývá, že pro studium speciálního problému nádorové transformace může efektivně posloužit studium zákonitostí chování komplexních systémů ve zdánlivě nepříbuzných oborech; kupříkladu historie, ekonomie nebo meteorologie. Analogie z těchto oborů nám mohou posloužit pro pochopení jevů z nějakého důvodu experimentálně nepřístupných. A to je důvodem, proč se na tyto obory a příměry opakovaně odkazuji.

Kapitola druhá – alternativní vysvětlení nádorové transformace

Nyní se dostávám ke vztahu mutací a buněčné transformace a k tvrzení, že „tradiční představa vzniku nádoru jako funkce hromadění mutací a jejich selekce je v současnosti významnou brzdou dalšího výzkumu“.

Recentní nádorový výzkum se soustřeďuje na mapování genů, jejich mutací, rigidních signálních drah a jejich vzájemného ovlivňování a využívá značně simplifikované experimentální modely. To je v pořádku potud, pokud si uvědomíme, že studujeme pouze situace odpovídající chaotickým systémům ve stabilních atraktorech s minimální intervencí z vnějšku (tedy systémy se slušnou mírou predikovatelnosti jejich chování). To se paradoxně týká i nádorově transformovaných buněk, protože je studujeme v jejich konečném,

transformovaném a tedy opět stabilním atraktoru. Avšak tou vlastní a nejniternější příčinou konečného stabilizovaného stavu nebyly mutace, ale krátkodobý vysoce nestabilní stav s vysokou mírou vnitřní neuspořádanosti, který způsobil přeskok z jednoho stabilního atraktoru do druhého. To co bychom tedy měli studovat je přechodný stav mezi dvěma stabilními atraktory. Tuto epizodickou událost můžeme nejsnadněji chápat jako analogii společenské revoluce – prvotní tranzientní impulzy-osoby-ideje, které ji zdánlivě vyvolaly jsou vbrzku nahrazeny stabilním uspořádáním, povětšině značně odlišným od původních impulzů. Přeloženo: původně stabilní atraktorový stav je po průchodu nestabilním intermezzem (revoluce) nahrazen jiným stabilním atraktorovým stavem. Kupříkladu studium společenských věd se - na rozdíl od molekulární biologie – daleko intenzivněji věnuje zkoumání krátkých revolučních úseků dějin; srovnejme kupříkladu množství publikací, které se věnují období husitských válek s pozorností věnovanou období následné stabilní vlády Jagellonců. Proces onkogenní transformace je v mnohém podobný historickému přechodu mezi dvěma epochami, avšak nejpreciznějí jsou v něm zmapovány počáteční a konečný stav a o přechodných revolučních událostech mezi nimi nevíme takřka nic. Nehodlám vyvracet význam onkogenních mutací, je nesporný, domnívám se však, že interpretace jejich významu by měla být pozměněna.

Známe-li kupříkladu nějaký onkoprotein, snažíme se mapovat jeho strukturu, interagující partnery, jím ovlivněné geny a signální dráhy. Předpokládáme, že jsou příčinou výsledného transformovaného fenotypu. Avšak stejně pravděpodobné je, že přítomnost zmíněného aberantního proteinu pouze vyvolala nerovnováhu stabilního atraktoru, v němž se nacházela postižená buňka na počátku. To vedlo – nikoliv okamžitě - k nestabilnímu stavu, s původním atraktorem (a tedy i se životem) neslučitelnému (analogie revoluce). Následovala buněčná smrt nebo nalezení nového alternativního atraktoru. A pokud není v blízkém fázovém prostoru „normální“ atraktor, pak tedy k „transformovanému“. Je zjevné, že pravděpodobnost nalezení nového stabilního atraktoru bude ovlivněna více faktory, jako jeho dostupnost z původního atraktoru, hloubka a délka trvání nerovnováhy ap. Tato moje interpretace byla nedávno nepřímo podpořena popisem jevu nazvaného chromothripse [33], či dlouho známou existencí tumor-promotorů [34] i námi popsaného fenoménu industáze [17]. Pokud přijmeme tuto představu, nabízí se testovatelná hypotéza, že kupříkladu nádorovou transformaci buňky lze vyvolat nemutagením působením destabilizujícím buněčný atraktor. Dopad takového průkazu by byl ohromný nejen pro nádorovou genetiku, ale i pro evoluční biologii, úvahy o vzniku života atd.

Mapování onkogenních mutací nám tedy umožňuje slušnou míru predikovatelnosti chování systému (například při terapii), ale pro studium toho krátkého nestabilního období je téměř k nepotřebě. Z praktického hlediska by však mělo největší význam právě studium toho krátkého nestabilního období, způsobu jeho vyvolání, kvantifikace a hlavně ovlivnění jeho hloubky a doby trvání. Pokud jsou tyto předpoklady správné, mohli bychom dokázat podprahovými podněty bez přímé mutageneze transformovat, ale třeba i rediferencovat cílové buňky.

Z uvedeného vyplývá, že živé systémy mohou být pouze otevřené a vždy a jedině chaotické. Liší se však mírou predikovatelnosti jejich vývoje a to vzájemně mezi sebou i v průběhu času sami v sobě. Zjednodušeně lze říci, že rigidnější systémy jsou stabilnější (nikoliv odolnější), avšak také méně přizpůsobivé. Evoluce organismů se stabilizovaným genomem probíhá spíše postupně v duchu klasického darwinismu a například umožňuje postupné šlechtitelství. Oproti tomu u organismů s genomem méně rigidním je pravděpodobnější evoluce skoková, s dramatickými a obtížně predikovatelnými následky. A obdobné předpoklady platí i pro transformabilitu jednotlivých buněčných typů v rámci organismu.

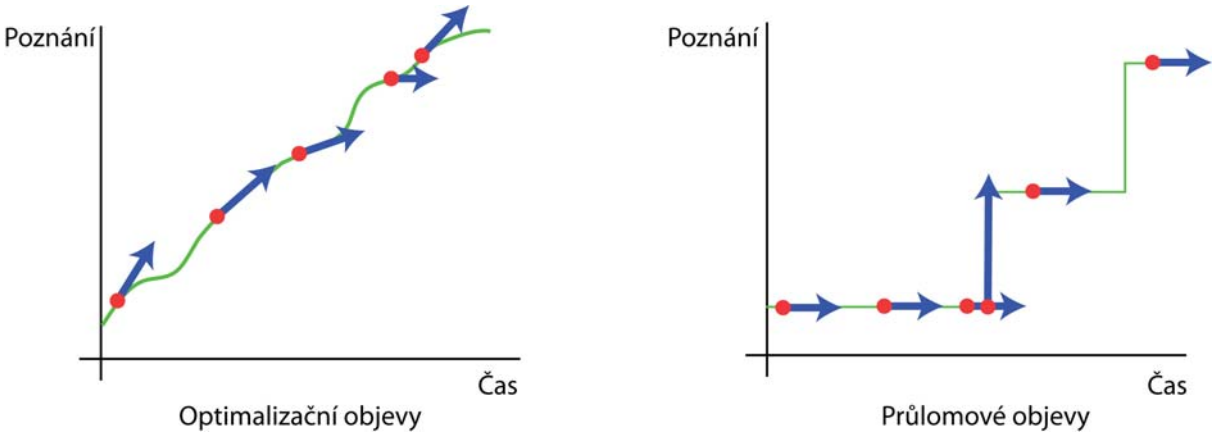
Proč tedy klást do protikladu dvě různé úrovně chaosu a jak? To, co staví v praxi do protikladu systémy méně a více chaotické jsou odlišné Otázky, které se snažíme naučit pokládat, pokud se tážeme více či méně chaotického systému. Nejsrozumitelněji by to snad mohlo být demonstrováno na příkladu dalšího chaotického systému, počasí. Pokud je aktuální míra meteorologické chaotičnosti malá, pak předpověď počasí bude poměrně přesná, naopak pokud bude vnitřní míra meteorologické chaotičnosti vysoká, předpověď bude málo spolehlivá. Většina lidí položí otázku „Jaké bude počasí?“. V případě málo chaotického systému bude tazatel s odpovědí spokojen, v druhém případě pravděpodobně nikoliv. Správná Otázka je však zcela jiná – „Jak přesná je aktuální předpověď?“ Po mém soudu je to stejný problém jako s optimalizačními a průlomovými objevy (viz. následující kapitola). U optimalizačních objevů bude tazatel pravděpodobně plně uspokojen (avšak často odpovědí na špatně položenou otázku), u průlomových objevů nikoliv. Proto je třeba věnovat velkou pozornost správné formulaci Otázek, speciálně u průlomových objevů. Ovšem zjevným oxymoronem je, že granty lze získat pouze na výzkum optimalizační, tedy na krátkodobou předpověď slunečného počasí.

Kapitola třetí – závěrečné úvahy o filosofii výzkumu

Udivující obecnou platnost Teorie chaosu, od níž jsem odvíjel své předchozí úvahy lze vidět prakticky všude. Mohli bychom se kupříkladu ptát, co je důvodem kvalitnějšího výzkumu v některých institucích či zemích než v jiných. Závisí to primárně opravdu na množství finančních prostředků? Nebo zde hrají významnou roli obtížně kvantifikovatelné faktory jako tradice nebo stabilita pracovního prostředí?

Věda, přesněji výzkum bývá rozdělován na aplikovaný a základní. Po mém soudu je toto dělení chybné a bývá užíváno spíše jako politický nástroj ve spojitosti s jeho financováním a někdy také jako simplifikující nástroj pro lenivost mysli. Jasným argumentem proti zmíněnému dělení je po mém soudu nemožnost exaktně definovat a oddělit tyto dvě kategorie, které se hluboce prolínají a v čase mění. Domnívám se, že daleko exaktnější avšak zdánlivě méně praktické je dělit výzkum na optimalizační a průlomový.

Píši o tom, neboť to zdaleka není tak triviální konstatování, jak by se mohlo na první pohled zdát. V čem je tedy zásadní odlišnost těchto dvou kategorií objevů? Domnívám se, že v předpověditelnosti výsledků. U optimalizačního výzkumu lze výsledek předvídat a pravděpodobnost úspěchu ovlivnit například množstvím vložených finančních prostředků a optimalizací jejich využívání. Jeho časový průběh lze popsat jako setrvalý nerovnoměrný avšak kontinuální vzestup (obr. 10A). U průlomových objevů výsledek předvídat nelze, jeho pravděpodobnost nelze jednoduše ovlivňovat a jejich časový průběh lze popsat jako skokový a diskontinuální (obr. 10B). S optimalizačním výzkumem se proto pojí výrazy plánovaný, financování, průběžný. S průlomovými objevy lze oproti tomu spojovat výrazy náhodný (něco nežádoucího), tradice (= něco mlhavého), génius (= něco magického), revoluce (= něco zmatečného). Triviální zjednodušení pro praxi? Na optimalizační objevy lze vypsat granty a kumulaci jejich výsledků ovlivnit. U průlomových objevů nikoliv, ale znamená to, že bychom se o ně neměli vůbec pokoušet?



Obr. 10 Schematické znázornění průběhu optimalizačního a průlomového výzkumu v čase.

Žádný z obou přístupů není lepší nebo důležitější, jsou svými protipóly. Co je zásadně odlišuje jsou jejich průběh, dopad a plánovatelnost. Z výše řečeného vyplývá, že soustředit se na průlomové objevy není možné z jejich podstaty (lze snad zvýšit jejich pravděpodobnost, nikoliv ovlivnit jejich výsledek). Naopak věnovat se pouze optimalizačnímu výzkumu je krátkozraké a čistě technologické. Jaký je tedy reálný průběh výzkumu? Nejspíše hybrid mezi oběma grafy, kde optimalizační složka zaplňuje prostor vytvořený průlomovým objevem. Paralela s chováním dalších komplexních systémů (evoluční procesy, ekonomika, společenský vývoj) je zde zcela zřejmá, avšak význam kombinace obou principů zasahuje mnohem dále, až k samotným kořenům vzniku života a k principu fungování disipačních struktur.

Významným důvodem, proč považuji za důležité rozebírat odlišnost dvou mezních výzkumných přístupů je právě tato široká analogie s dalšími významnými ději – s evolucí, chováním otevřených fyzikálních soustav, společenským vývojem nebo buněčnou transformací. Snad je z předchozího textu tato soběpodobnost dostatečně patrná. Všechny rozebírané děje mají totiž jednoho společného jmenovatele – chaos. Chaos, který je tou pravou a nejniternější podstatou evoluce, mnohobuněčného života, nádorové transformace, smrti i vědeckého zkoumání a zdaleka neplatí, že zvyšování uspořádanosti vede k lepším výsledkům. Z počátku svého studia a psaní textu této disertace jsem hledal řád přírody a původ života. Nalezl jsem chaos a dobral se představy, že tím, co má opravdu význam je studium jeho řádu.

Obr. 11 Baron Johann Konrad Richthausen (1604 – 1643)

TABULKY

Tab. 1A

Všechna klonální integrační místa (VISs)	1848
VISs v kuřecím genomu	1713
Cirkulární (neintegrované) proviry	50
Provirové sekvence	83
Jaterní VISs - všechna	143
Jaterní VISs - v kuřecím genomu	127
Plicní angiosarkomy VISs - všechna	216
Plicní angiosarkomy - VISs v kuřecím genomu	188
Nefroblastomové VISs - všechna	1337
Nefroblastomy - VISs v kuřecím genomu	1260
Tkáňové kultury - VISs všechna	130
Tkáňové kultury - VISs v kuřecím genomu	119
Ovariální tumor - VISs všechna	20
Ovariální tumor - VISs v kuřecím genomu	20
Vytipovaná OBECNÁ MÍSTA INTEGRACE	**113**

Tab. 1B

	VISs	Chr. [Mbp]	VIS/Mbp
Chr. 1	384	201	1,91
Chr. 2	271	155	1,75
Chr. 3	217	114	1,90
Chr. 4	130	94	1,38
Chr. 5	122	62	1,97
Chr. 6	48	37	1,30
Chr. 7	41	38	1,08
Chr. 8	35	31	1,00
Chr. 9	32	26	1,13
Chr. 10	26	23	1,13
Chr. 11	24	22	1,09
Chr. 12	39	21	1,86
Chr. 13	18	19	0,95
Chr. 14	17	16	1,06
Chr. 15	13	13	1,00
Chr. 16	0	0,5	0,00
Chr. 17	8	11	0,73
Chr. 18	16	11	1,45
Chr. 19	21	10	2,10
Chr. 20	13	14	0,93
Chr. 21	11	7	1,57
Chr. 22	6	4	1,50
Chr. 23	9	6	1,50
Chr. 24	8	6	1,33
Chr. 25	3	2	1,50
Chr. 26	4	5	0,80
Chr. 27	4	5	0,80
Chr. 28	10	4,5	2,22
Chr. W	1	0,3	3,33
Chr. Z	76	75	1,01

Tab. 1 Statistické shrnutí analýzy všech klonálních míst integrace MAV-2 provirů (VIS).

A) *„všechna VISs“ – všechny sekvenované unikátní fragmenty integrovaných provirů, včetně defektních provirů a cirkulárních forem MAV-2; „vytipovaná OBECNÁ MÍSTA INTEGRACE“ – počet řádků Tab. 2, viz. legenda k Tab. 2.*

B) *Rozložení počtu a hustoty klonálních VISs na jednotlivé kuřecí chromozómy (podle Gallus gallus genome assembly WASHUC2).*

Tab. 2

				X	0/1	0/1	0/1	0/1	0/1	0/1
Skóre	**Kandidátní CISs**	**Chr.**	**Pozice [MBp]**	**Počet klonál. VIS**	**Rozsah CIS [kbp]**	**Tkáň**	**Lokus**	**A**	**B**	**C**
38	Nal 3-66	3	66	33	1	P	FRK	i	s	y
18	Nal 5-17	5	17	12	18	N(3), J(9)	RASH	i	s	y
16	Nal 2-51	2	51	10	10	J(9), N(1)	EGFR	i	s	y
16	Nal 12-16	12	16	11	267	N(9), TK(2)	FOXP1	i	s	y
15	Nal 5-12	5	12	11	269	TK, PCR	SOX6	u	s	y
14	Nal 2-29	2	29	9	275	N	TWIST	i	s	y
13	Nal 2-114	2	114	8	36	N	PLAG1	u	s	y
9	Nal 12-2	12	2	3	0	J	MST1R	i	s	y
8	Nal 1-26	1	26	5	23	J, TK	MET	v	v	y
8	Nal 2-148	2	148	4	27	N(3), J(1)	mir30b, d	u	v	y
7	Nal 4-11	4	11	3	13	N	RRAG	i	s	n
7	Nal 1-13a	1	13	2	3	N	FBXL13	i	s	n
7	Nal 1-182a	1	182	2	0	N	FGF9	i	s	?
7	Nal 1-192b	1	192	2	15	N	FZD4	u	s	n
7	Nal 2-4b	2	4	2	11	P	OXSR1	i	s	n
7	Nal 2-125	2	125	4	89	J, TK, N	TPD52	i	o	n
7	Nal 2-142	2	142	2		N	CCN3	i	s	y
7	Nal 6-31	6	31	2	0	N	EIF3A	i	o	n
6	Nal 1-15	1	15	2	16	N, P	LAMB1	i	s	n
6	Nal 1-23	1	23	2	1	N	HYAL4	u	v	n
6	Nal 1-33	1	33	2	3	N, TK	LRIG3	u	o	n
6	Nal 1-116	1	116	2	4	N	BCOR	i	v	n
6	Nal 1-125	1	125	2	5	P	AP1S2	i	v	n
6	Nal 1-182b	1	182	2	16	P, TK	ZDHHC20	i	s	n
6	Nal 1-189b	1	189	2	2	N, TK	SESN3	u	s	n
6	Nal 1-190	1	190	2	28	N	FAT3	i	s	n
6	Nal 1-192a	1	192	3	80	N	TMEM135	v	o	n
6	Nal 2-28	2	28	2	5	N	TSPAN13	v	s	n
6	Nal 2-40	2	40	2	0	N, P	CMTM8	i	s	n
6	Nal 2-91	2	91	2	34	N	DNAH5	i	o	n
6	Nal 2-102	2	102	2	1	N	ARHGAP28	u	v	n
6	Nal 2-144	2	144	2	0	N	KIAA0196	i	v	n
6	Nal 2-147	2	147	2	15	N	ST3GAL1	u	v	n
6	Nal 2-151	2	151	2	3	N	GPR20	u	v	n
6	Nal 3-49a	3	49	2	8	N	SASH1	i	v	n
6	Nal 3-82	3	82	2	4	N	SH3BGRL2	v	s	n
6	Nal 3-104	3	104	2	6	N	ZP	u	v	n
6	Nal 3-110	3	110	2	14	N	EFHC1	v	s	n

6	Nal 4-46	4	46	2	32	N	BMP3	d	o	?
6	Nal 4-83	4	83	2	30	N	ABLIM2	i	s	n
6	Nal 5-2	5	2	2	57	N	GAS2	i	s	n
6	Nal 5-14	5	14	2	9	N	KCNQ1	i	v	n
6	Nal 5-38	5	38	2	7	N	EGLN3	u	v	n
6	Nal 5-50	5	50	2	27	N	EML1	u	s	n
6	Nal 5-51	5	51	2	0	N	DIO3	v	o	n
6	Nal 7-5	7	5	2	11	N	IQCA1	v	s	n
6	Nal 8-27	8	27	2	36	N	NFIA	i	o	n
6	Nal 15-2	15	5	2	0	N	???	i	s	n
6	Nal 24-5	24	5	2	24	N	PHLDB1	u	o	n
6	Nal Z-0	Z	0	2	67	N	WDR7	i	o	n
5	Nal 1-12	1	12	2	19	N	GNAI	v	v	n
5	Nal 1-38	1	38	2	145	N	TRHDE	v	s	n
5	Nal 1-108	1	108	2	22	P, N	HUNK	d	s	n
5	Nal 1-156	1	156	3	147	N	NDFIP2	v	v	n
5	Nal 1-160	1	160	2	121	N, P	DACH1	i	s	n
5	Nal 1-177a	1	177	3	225	N	DCLK1	v	v	?
5	Nal 1-177b	1	177	2	299	N	NBEA	v	s	n
5	Nal 1-184	1	184	2	16	P, N	ATM	v	o	n
5	Nal 1-186	1	186	3	107	P, N, J	DYNC2H1	v	s	n
5	Nal 1-193	1	193	3	12	P, N	C11orf82	v	v	n
5	Nal 1-197	1	197	2	142	N, P	ODZ4	i	s	n
5	Nal 2-129	2	129	2	95	N	RUNX1T1	v	o	n
5	Nal 2-130	2	130	2	18	N, P	TMEM67	v	o	n
5	Nal 2-143	2	143	2	14	N	FBXO32	v	v	n
5	Nal 3-24	3	24	2	72	N	EML4	i	v	n
5	Nal 3-30	3	30	2	78	N	DAAM2	i	v	n
5	Nal 3-53	3	53	2	58	J, N	ZDHHC14	i	s	n
5	Nal 3-56	3	56	2	6	P, N	gen	v	s	n
5	Nal 3-77	3	77	2	59	N	BACH2	v	o	n
5	Nal 3-99	3	99	2	49	N	HPCL1	v	s	n
5	Nal 3-100	3	100	2	4	P, N	TRIB2	d	v	n
5	Nal 4-90	4	90	2	75	ov, N	CTNNA2	i	s	n
5	Nal 5-15	5	15	2	22	N	DUSP8	v	s	n
5	Nal 5-34a	5	34	2	23	N	ZNF770	v	s	n
5	Nal 5-42	5	42	2	28	N	DIO2	v	o	n
5	Nal 11-21	11	21	2	10	N	ATBF1	?	?	n
5	Nal 18-5	18	5	2	5	N, J	ASPSCR1	i	v	n
5	Nal 23-1	23	1	2	17	N, P	CITED4	u	v	n
4	Nal 1-22	1	22	2	1	N				
4	Nal 1-27	1	27	2	199	N	FOXP2	v	v	n
4	Nal 1-124	1	124	2	4	N				
4	Nal 1-150	1	150	2	29	N, P	GPC6	i	v	n
4	Nal 1-169	1	169	3	41	N				

4	Nal 1-189a	1	189	2	11	N				
4	Nal 2-4a	2	4	2	0	N				
4	Nal 2-9	2	9	2	30	N	WDR60	v	v	n
4	Nal 2-19	2	19	2	40	N	PTPLA	v	v	n
4	Nal 2-110	2	110	2	52	N	DTNA	v	v	n
4	Nal 2-124	2	124	2	161	N	ZFHX4	v	v	n
4	Nal 2-136	2	136	2	50	N, TK	ZFPM2	u	v	n
4	Nal 3-49b	3	49	3	199	N(2), P(1)	UST	v	v	n
4	Nal 4-47	4	47	2	16	N	v	v	v	n
4	Nal 4-60	4	60	2	210	N, P	UNC5C	i	v	n
4	Nal 5-25	5	25	2	74	N	C11orf49	v	v	n
4	Nal 5-31	5	31	2	41	N	BMF	v	v	n
4	Nal 5-34b	5	34	2	20	P, N	STXBP6	v	v	n
4	Nal 6-25	6	25	2	135	N	SH3PXD2A	v	v	n
4	Nal 6-27	6	27	2	42	TK, N	ADD3	i	v	n
4	Nal 6-37	6	37	2	148	N	INPP5A	v	v	n
4	Nal 9-18	9	18	2	0	N				
4	Nal 11-20	11	20	2	15	N	v	v	v	n
4	Nal 12-5	12	5	2	34	N	RAF1	v	v	n
4	Nal 12-14	12	14	3	97	P(2), N(1)	LRIG	v	v	n
4	Nal 14-5	14	5	2	101	N	LMF1	v	v	n
3	Nal 1-13b	1	13	2	50	N, ov	PTPN12	v	v	n
3	Nal 1-34	1	34	2	5	N, P				
3	Nal 1-164	1	164	3	30	N, P, ov				
3	Nal 2-32	2	32	2	38	N, P	SNX10	v	v	n
3	Nal 3-92a	3	92	2	6	J, N				n
3	Nal 3-112	3	112	2	29	N, TK	RUNX2	v	v	n
3	Nal Z-9	Z	9	2	11	N, P	???	i	s	n
	Nal 1-70	1	70	2	10	N, P	gen	?	?	n
	Nal 1-137	1	137	2	6	N, TK	C2orf29	?	?	n
	Nal 4-81	4	81	2	11	TK, N	WDR1	?	?	n

Tab. 2 Obecná místa integrace (CISs) řazená podle relevance (skóre).

Jako CIS je označována oblast genomu menší než 20kbp nebo genový lokus, v němž byla identifikována 2 a více klonální(ch) VIS.

*Legenda: „Skóre“ – součet hodnot v hodnocených sloupcích [**Q**] pro dané CIS. Jednotlivá CIS jsou řazena v tabulce podle této hodnoty, podle klesající významnosti; „Kandidátní CIS“ – pracovní označení jednotlivých CIS; „chr.“ – chromozóm, „Pozice [Mbp]” – pozice na chromozómu v Mbp (G. gallus, WASHUC2);*

*“Počet klonálních VISs” – počet klonálních integrací MAV-2 v daném CIS, [**Q**; hodnota odpovídá počtu VISs]*

*“Rozsah CIS” – vzdálenost mezi nejvzdálenějšími VIS, [**Q**; Q=1 pro hodnoty <20kbp, Q=0 pro hodnoty >20kbp]*

*“Tkáň” – nefroblastom (N), plíce (P), játra (J), tkáňová kultura (TK), ov (ovárium), [**Q**; Q=1 pro jedinou tkáň, Q=0 pro vice tkání]*

*„Lokus“ – označení genového lokusu, překrývajícího se s CIS [**Q**; Q=1 pro vyplněné pole, Q=0 pro prázdné pole]*

*„A“ – lokalizace provirů v rámci zasaženého genového lokusu, u = nad transkriptem, i = uvnitř transkriptu, d = pod transkriptem, v = variabilní [**Q**; Q=1 pro hodnoty u, i, d, Q=0 pro hodnotu v]*

*„B“ – orientace provirů vzhledem k orientaci zasaženého genového lokusu, s = stejná, o = opačná, v = variabilní [**Q**; Q=1 pro hodnoty s, o, Q=0 pro hodnotu v]*

*„C“ – experimentálně ověřená deregulace postiženého lokusu, y = ano, n = ne [**Q**; Q=1 pro y, Q=0 pro n].*

Literatura

1. Ellerman, V. and O. Bang, *Experimentelle Leukämie bei Hühnern.* Zentralblatt fur Bakteriologi, Parasitenkunde, Infektionskrinkheiten und Hygiene Abteilung Originale, 1908. **46**: p. 595-609.
2. Rous, P., *A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells.* J Exp Med, 1911. **13**(4): p. 397-411.
3. Baltimore, D., *RNA-dependent DNA polymerase in virions of RNA tumour viruses.* Nature, 1970. **226**(5252): p. 1209-11.
4. Temin, H.M. and S. Mizutani, *RNA-dependent DNA polymerase in virions of Rous sarcoma virus.* Nature, 1970. **226**(5252): p. 1211-3.
5. Weiss, S.R., H.E. Varmus, and J.M. Bishop, *The size and genetic composition of virus-specific RNAs in the cytoplasm of cells producing avian sarcoma-leukosis viruses.* Cell, 1977. **12**(4): p. 983-92.
6. Frese, K.K. and D.A. Tuveson, *Maximizing mouse cancer models.* Nat Rev Cancer, 2007. **7**(9): p. 645-58.
7. Rangarajan, A. and R.A. Weinberg, *Opinion: Comparative biology of mouse versus human cells: modelling human cancer in mice.* Nat Rev Cancer, 2003. **3**(12): p. 952-9.
8. Rosson, D. and E.P. Reddy, *Nucleotide sequence of chicken c-myb complementary DNA and implications for myb oncogene activation.* Nature, 1986. **319**(6054): p. 604-6.
9. Pečenka, V., P. Pajer, V. Karafiát, and M. Dvořák, *Chicken Models of Retroviral Insertional Mutagenesis*, in *Insertional Mutagenesis Strategies in Cancer Genetics*, A.J. Dupuy and D.A. Largaespada, Editors. 2011, Springer Science+Business Media, LLC, 233 Spring Street, NY 10013, USA. p. 77-113.
10. Coffin, J.M., S.H. Hughes, and H.E. Varmus, *Retroviruses*. 2011/03/25 ed. 1997: Cold Spring Harbor (NY): CSHL Press.
11. Pecenka, V., M. Dvorak, V. Karafiat, E. Sloncova, I. Hlozanek, M. Travnicek, and J. Riman, *Avian nephroblastomas induced by a retrovirus (MAV-2) lacking oncogene. II. Search for common sites of proviral integration in tumour DNA.* Folia Biol (Praha), 1988. **34**(3): p. 147-69.
12. Pecenka, V., M. Dvorak, V. Karafiat, and M. Travnicek, *Avian nephroblastomas induced by a retrovirus (MAV-2) lacking oncogene. III. Presence of defective MAV-2 proviruses in tumour DNA.* Folia Biol (Praha), 1988. **34**(4): p. 215-32.
13. Pecenka, V., M. Dvorak, and M. Travnicek, *Avian nephroblastomas induced by a retrovirus (MAV-2) lacking oncogene. I. Construction of MAV-1 and MAV-2 proviral restriction maps and preparation of specific proviral molecular subclones.* Folia Biol (Praha), 1988. **34**(3): p. 129-46.
14. Watts, S.L. and R.E. Smith, *Pathology of chickens infected with avian nephoblastoma virus MAV-2(N).* Infect Immun, 1980. **27**(2): p. 501-12.
15. Joliot, V., C. Martinerie, G. Dambrine, G. Plassiart, M. Brisac, J. Crochet, and B. Perbal, *Proviral rearrangements and overexpression of a new cellular gene (nov) in myeloblastosis-associated virus type 1-induced nephroblastomas.* Mol Cell Biol, 1992. **12**(1): p. 10-21.
16. Smith, R.E. and C. Moscovici, *The oncogenic effects of nontransforming viruses from avian myeloblastosis virus.* Cancer Res, 1969. **29**(7): p. 1356-66.
17. Pajer, P., V. Karafiat, V. Pecenka, D. Prukova, J. Dudlova, J. Plachy, P. Kasparova, and M. Dvorak, *Industasis, a promotion of tumor formation by nontumorigenic stray cells.* Cancer Res, 2009. **69**(11): p. 4605-12.
18. Pajer, P., V. Pecenka, V. Karafiat, J. Kralova, Z. Horejsi, and M. Dvorak, *The twist gene is a common target of retroviral integration and transcriptional deregulation in experimental nephroblastoma.* Oncogene, 2003. **22**(5): p. 665-73.
19. Pajer, P., V. Pecenka, J. Kralova, V. Karafiat, D. Prukova, Z. Zemanova, R. Kodet, and M. Dvorak, *Identification of potential human oncogenes by mapping the common viral integration sites in avian nephroblastoma.* Cancer Res, 2006. **66**(1): p. 78-86.

20. Wu, X., B.T. Luke, and S.M. Burgess, *Redefining the common insertion site.* Virology, 2006. **344**(2): p. 292-5.
21. Kononen, J., L. Bubendorf, A. Kallioniemi, M. Barlund, P. Schraml, S. Leighton, J. Torhorst, M.J. Mihatsch, G. Sauter, and O.P. Kallioniemi, *Tissue microarrays for high-throughput molecular profiling of tumor specimens.* Nat Med, 1998. **4**(7): p. 844-7.
22. Van Dyck, F., J. Declercq, C.V. Braem, and W.J. Van de Ven, *PLAG1, the prototype of the PLAG gene family: versatility in tumour development (review).* Int J Oncol, 2007. **30**(4): p. 765-74.
23. Wlodarska, I., E. Veyt, P. De Paepe, P. Vandenberghe, P. Nooijen, I. Theate, L. Michaux, X. Sagaert, P. Marynen, A. Hagemeijer, and C. De Wolf-Peeters, *FOXP1, a gene highly expressed in a subset of diffuse large B-cell lymphoma, is recurrently targeted by genomic aberrations.* Leukemia, 2005. **19**(8): p. 1299-305.
24. Yang, J., S.A. Mani, and R.A. Weinberg, *Exploring a new twist on tumor metastasis.* Cancer Res, 2006. **66**(9): p. 4549-52.
25. Cance, W.G., R.J. Craven, M. Bergman, L. Xu, K. Alitalo, and E.T. Liu, *Rak, a novel nuclear tyrosine kinase expressed in epithelial cells.* Cell Growth Differ, 1994. **5**(12): p. 1347-55.
26. Lee, J., Z. Wang, S.M. Luoh, W.I. Wood, and D.T. Scadden, *Cloning of FRK, a novel human intracellular SRC-like tyrosine kinase-encoding gene.* Gene, 1994. **138**(1-2): p. 247-51.
27. Akerblom, B., C. Anneren, and M. Welsh, *A role of FRK in regulation of embryonal pancreatic beta cell formation.* Mol Cell Endocrinol, 2007. **270**(1-2): p. 73-8.
28. Hosoya, N., Y. Qiao, A. Hangaishi, L. Wang, Y. Nannya, M. Sanada, M. Kurokawa, S. Chiba, H. Hirai, and S. Ogawa, *Identification of a SRC-like tyrosine kinase gene, FRK, fused with ETV6 in a patient with acute myelogenous leukemia carrying a t(6;12)(q21;p13) translocation.* Genes Chromosomes Cancer, 2005. **42**(3): p. 269-79.
29. Pecenka, V., V. Karafiat, P. Kasparova, J. Dudlova, M. Dvorak, and P. Pajer, *Distinct cancer genes are activated in different liver tumor types in a new chicken model of liver cancerogenesis.* submitted, 2012.
30. Lefebvre, V., *The SoxD transcription factors--Sox5, Sox6, and Sox13--are key cell fate modulators.* Int J Biochem Cell Biol, 2010. **42**(3): p. 429-32.
31. Du, Y., N.A. Jenkins, and N.G. Copeland, *Insertional mutagenesis identifies genes that promote the immortalization of primary bone marrow progenitor cells.* Blood, 2005. **106**(12): p. 3932-9.
32. Ince, T.A., A.L. Richardson, G.W. Bell, M. Saitoh, S. Godar, A.E. Karnoub, J.D. Iglehart, and R.A. Weinberg, *Transformation of different human breast epithelial cell types leads to distinct tumor phenotypes.* Cancer Cell, 2007. **12**(2): p. 160-70.
33. Stephens, P.J., C.D. Greenman, B. Fu, F. Yang, G.R. Bignell, L.J. Mudie, E.D. Pleasance, K.W. Lau, D. Beare, L.A. Stebbings, S. McLaren, M.L. Lin, D.J. McBride, I. Varela, S. Nik-Zainal, C. Leroy, M. Jia, A. Menzies, A.P. Butler, J.W. Teague, M.A. Quail, J. Burton, H. Swerdlow, N.P. Carter, L.A. Morsberger, C. Iacobuzio-Donahue, G.A. Follows, A.R. Green, A.M. Flanagan, M.R. Stratton, P.A. Futreal, and P.J. Campbell, *Massive genomic rearrangement acquired in a single catastrophic event during cancer development.* Cell, 2011. **144**(1): p. 27-40.
34. Berenblum, I. and P. Shubik, *The role of croton oil applications, associated with a single painting of a carcinogen, in tumour induction of the mouse's skin.* Br J Cancer, 1947. **1**(4): p. 379-82.

PUBLICATIONS

Oncogene (2003) 22, 665–673

www.nature.com/onc

The *twist* gene is a common target of retroviral integration and transcriptional deregulation in experimental nephroblastoma

Petr Pajer, Vladimír Pečenka, Vít Karafiát, Jarmila Králová, Zuzana Hořejší and Michal Dvořák*

Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Flemingovo n. 2, 166 37 Prague 6, Czech Republic

The genes involved in the transformation of kidney blastema cells were searched for in avian nephroblastomas induced by the MAV2 retrovirus. The *twist* gene was identified as a common site of provirus integration in tumor cells. *Twist* was rearranged by the MAV2 provirus in three out of 76 independent nephroblastoma samples. The MAV2 integration sites were localized within 40 nucleotides of the *twist* 5′UTR region, right upstream from the ATG initiation codon. The integrated proviruses were deleted at their 5′ends. As a consequence, *twist* transcription became controlled by the retroviral 3′LTR promoter and was strongly upregulated, more than 200 times. In addition, 2–100 times elevated *twist* transcription was also detected in the majority of other nephroblastoma samples not containing MAV2 in the *twist* locus. We propose that chicken nephroblastoma originates from a single blastemic cell in which the MAV retrovirus, through its integration, has deregulated specific combinations of genes controlling proliferation and differentiation. The activation of the *twist* gene expression appears to contribute to tumorigenesis, as there is an *in vivo* positive selection of tumor cell clones containing the *twist* gene hyperactivated by MAV2 sequences inserted within the *twist* promoter.
Oncogene (2003) **22,** 665–673. doi:10.1038/sj.onc.1206105

Keywords: *Twist* gene; MAV retrovirus; common viral integration site; nephroblastoma

Introduction

The model of MAV2-induced chicken nephroblastoma serves as a valuable tool for identifying genes involved in the malignant transformation of renal cell precursors—nephrogenic blastema cells. The model can provide relevant information for understanding the etiology and pathology of the pediatric kidney nephroblastoma, Wilms tumor, as both malignancies are derived from the same cell type and share a number of similar features (Ishiguro *et al.,* 1962). The chicken model is based on the assumed ability of MAV2 retrovirus to transform cells by insertional mutagenesis, that is, by a deregulation of expression of genes hit by the proviral integration. Moreover, MAV2-encoded Env protein appears to facilitate transformation of blastema cells as it stimulates their proliferation (Joliot *et al.,* 1996).

It is assumed that macroscopic nephroblastomas arise by clonal expansion of blastema cells in which MAV2 provirus has deregulated specific genes controlling differentiation and proliferation. Each experimentally induced chicken tumor contains on average five clonally integrated MAV2 proviruses (Pečenka *et al.,* 1988a), which is in accordance with the widely accepted multihit model of cancerogenesis. The model postulates that a set of relevant genes must be affected in a specific way for the cell to achieve the fully transformed phenotype. Accordingly, alteration of several genes (including WT1) was shown to be required for Wilms tumor induction (Knudson and Strong, 1972; Dome and Coppes, 2002).

In the chicken nephroblastoma, the *nov* (nephroblastoma overexpressed) gene was suggested to be involved (Joliot *et al.,* 1992; Perbal, 2001). c-Ha-ras and c-fos proto-oncogenes have also been found to be activated by the retrovirus in chicken nephroblastoma. However, their role seems to be limited to single cases (Westaway *et al.,* 1986; Collart *et al.,* 1990).

In this work, we have cloned and characterized a number of individual MAV2 integration sites in nephroblastoma DNA and found that the *twist* 5′UTR is the common integration site of MAV2 provirus in nephroblastomas.

Twist was originally discovered as a *Drosophila* gene, whose mutation causes the characteristic 'twisted' phenotype in embryos (Thisse *et al.,* 1988). Later, mouse, human, frog and chicken homologs were cloned and characterized (Castanon and Baylies, 2002). In human, *twist* germline mutations, resulting in a reduced Twist protein level, are supposed to be responsible for some Saethre–Chotzen syndrome (SCS) cases (el Ghouzzi *et al.,* 1997; Howard *et al.,* 1997). This is also consistent with the phenotype of the mouse mutants—homozygous *twist*-null murine embryos exhibit failure of neural tube closure, while heterozygots display a moderate phenotype, including minor skull and limb abnormalities, reminding the SCS patients (Bourgeois *et al.,* 1998).

The Twist protein belongs to the basic helix–loop–helix family of transcription factors. Its expression is the highest during embryonic development in immature mesodermal cells and in organs of mesodermal origin; postnatally it declines to low levels in adult tissues. Twist has been shown to block terminal differentiation of mesodermal cells and to inhibit p53-dependent

*Correspondence: M Dvořák, E-mail: mdvorak@img.cas.cz
Received 12 July 2002; revised 30 September 2002, accepted 4 October 2002

apoptosis (Maestro *et al.,* 1999). A number of Twist-regulated genes were discovered in humans as well as in *Drosophila* and *C. elegans*, and several protein partners directly binding to the Twist protein *in vivo* were found (Castanon and Baylies, 2002).

The *twist* gene has a close homolog *dermo* showing an almost complete identity at the amino-acid level in the bHLH domain. *Dermo* is expressed in more differentiated cell types in comparison to *twist* and is hypothesized to execute similar functions as Twist (Li *et al.,* 1995).

A possible participation of *twist* in cell transformation has already been suggested. Approximately 50% of rhabdomyosarcoma samples display abnormally high levels of the Twist protein (Maestro *et al.,* 1999). In addition, the methylation status of the *twist* promoter has also been used as a sensitive marker for the detection of breast cancer cells in ductal lavage fluid (Evron *et al.,* 2001). However, no direct evidence of *twist* or *dermo* participation in cell transformation was presented so far. Our data show that overexpressed *twist* participates in malignant transformation of renal blastema cells. On the contrary, *dermo* does not seem to be involved in nephroblastoma formation.

Results

Induction of nephroblastomas, collection and characterization of samples

Nephroblastomas were induced by infecting 12-day-old embryos or 2-day-old chicks with the MAV2 retrovirus. About 100 samples of the renal tissue were collected from 45 infected animals (36 of them developed macroscopic tumors in one or both kidneys within 45–120 days postinfection).

Tumor samples represented either individual nephroblastoma foci about 1 cm in diameter or samples taken from distant parts of large tumors (ranging in diameter from 2 to 10 cm). Control tissue samples were seized from uninfected kidney and from infected but morphologically normal parts of kidneys in the vicinity of a tumor. As additional controls, mesonephros and metanephros were collected from uninfected embryos. The samples were numbered (e.g. 113/1a) to record the donor animal (the first number), the kidney from which the sample was obtained (the number after the slash refers either to the left (1) or right (2) kidney), and the individual piece of the tissue (the letter N stands for nontransformed control samples).

Genomic DNAs from tumor samples were analysed by Southern-blot hybridization to show their clonality, to determine, in each tumor sample, the number of integrated proviruses, and to single out independent clones among samples taken from one large tumor. Figure 1a shows the representative analysis of 15 tumors. DNAs were digested with restriction enzyme *Tth*111I, which does not cut the proviral DNA, and hybridized with HH25, the MAV-specific probe derived from the U3 region (Figure 2a). The result (Figure 1a) documents the clonal nature of tumor specimens. Since each DNA contains a well-defined set of MAV2 fragments with the same hybridization intensity, most of the cells in a given sample must have been derived from a single infected cell. A similar experiment was

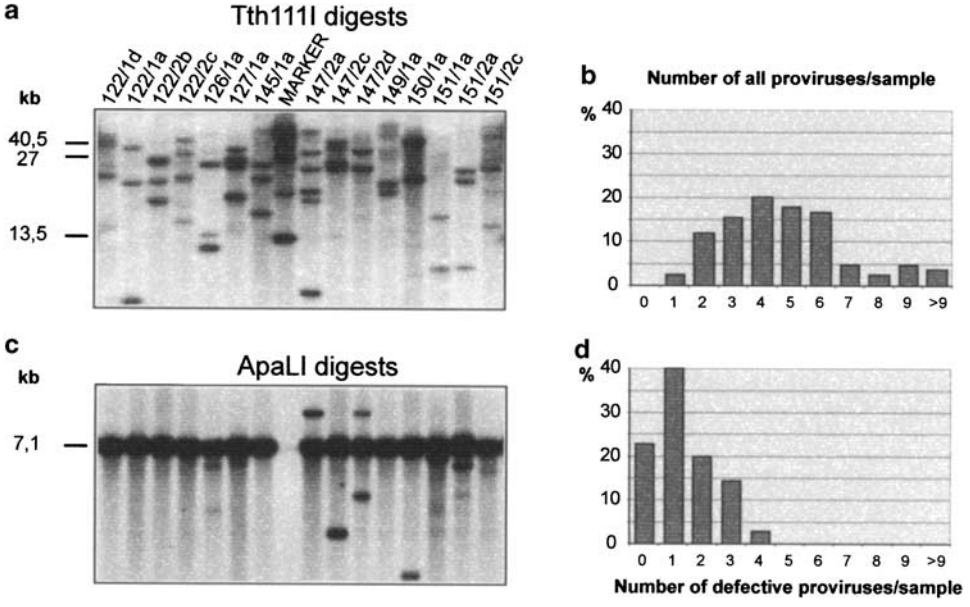


Figure 1 Nephroblastoma tumor foci are clones with a defined number of complete and deleted MAV2 proviruses: (**a**) example of tumor DNAs digested with the *Tth*111I restriction endonuclease and hybridized on Southern blots with the MAV-specific HH25 probe (see Figure 2a); (**b**) distribution of the total number of integrated proviruses in each tumor DNA detected by *Tth*111I digestion; (**c**) example of tumor DNAs digested with the *Apa*LI restriction endonuclease and hybridized on Southern blots with the retroviral 3′UTR-specific AH24 probe (see Figure 2a); (**d**) distribution of deleted proviruses detected by *Apa*LI digestion (note that to reveal all defective proviruses, the results with 5′UTR- and 3′UTR-specific probes had to be combined)

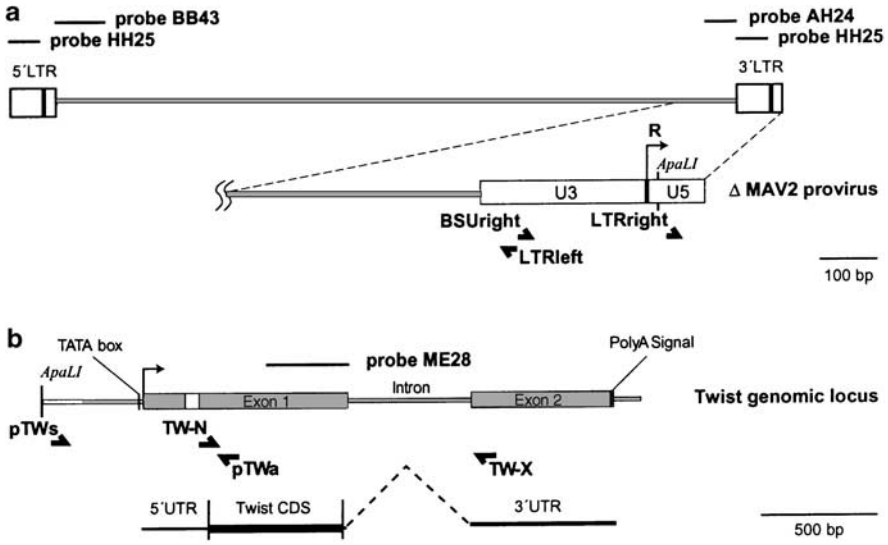


Figure 2 Schematic representation of the MAV2 provirus and the *twist* gene. (**a**) The upper scheme shows positions of hybridization probes BB43, AH24 and HH25 in the MAV2 genome. The 3′LTR has been enlarged to reveal the position of the *Apa*LI restriction site, start of transcription R and positions of PCR primers BSUright, LTRleft and LTRright. The orientations of 3′OH ends of primers are depicted by demiarrows. (**b**) Schematic representation of the exon–intron structure of the chicken *twist* locus shows the position of the *Apa*LI restriction site, the position and orientation of 3′OH ends of PCR primers, transcriptional start site downstream from the TATA box, and the position of hybridization probe ME28. The open box within the noncoding part of exon 1 shows the localization of the novel *twist* sequence whose primary structure is given in Figure 4. The structure of the *twist* mRNA is schematically depicted below with the marked 5′ (5′UTR) and 3′ (3′UTR) untranslated regions as well as the *twist* coding sequence (CDS)

performed with all tumor DNAs obtained (data not shown). In all, 76 independent clonal tumor DNAs were selected for further experiments. The numbers of proviruses in tumor clones are summarized in the histogram in Figure 1b.

To detect defective proviruses, the enzymes *Apa*LI and *Dra*I were used according to the strategy published earlier (Pečenka *et al.,* 1988b). As the criterion of defectiveness we used the appearance of MAV2 fragments different from MAV2 parental provirus. The presence of defective proviruses revealed by *Apa*LI digestion is documented by representative results in Figure 1c. The numbers of defective proviruses in all 76 tumor clones are summarized in the histogram in Figure 1d.

In summary, tumor clones contained mostly two to six proviruses. Among them, most frequently, none to three were defective.

Cloning of integration sites of defective MAV2 proviruses by inverse PCR

The clonal character of nephroblastoma foci makes it possible to isolate loci containing integrated proviruses and to identify among them those that contribute to malignant transformation. Such genes should constitute a set of 'common viral integration sites' of the MAV2 provirus.

To find common viral integration sites, we first focused on defective MAV2 proviruses as such proviruses were shown to affect surrounding genes more frequently than the complete ones (Robinson and Gagnon, 1986). Six genomic DNAs from tumor samples containing defective provirus truncated at the 5′end were digested with the restriction endonuclease *Apa*LI and size-selected on agarose gels. DNA fragments harboring defective proviruses flanked by host sequences were isolated. Inverse PCR (IPCR) was performed using LTRleft + BSUright primers (see Figure 2a and the Materials and methods section) to amplify genomic sequences adjacent to an integrated defective provirus. The resulting PCR products were cloned, sequenced and the sequences were analysed for their homology to entries in public databases (NCBI). The six clones analysed in this way contained, in addition to proviral sequences, host sequences ranging from 0.6 to 1.8 kb. The host DNA in four clones displayed no significant homology to known genes, while the fifth clone was highly homologous to a putative WD repeat-containing gene. The sixth DNA, originating from the sample 102/2b, was identical with the chicken *twist* gene except for a stretch of 50 nucleotides (see below).

The twist gene is the common integration site of MAV2 provirus in nephroblastomas

To screen our tumor samples collection for MAV2 insertions into the *twist* locus, DNAs were digested with *Hpa*I, which cuts approximately 5 kb upstream and 5 kb downstream from *twist* initiation and stop codons, yielding an 11 kb twist-containing fragment, and hybridized on Southern blots to the ME28 probe derived from the 3′part of the first *twist* exon (Figure 2b). Since MAV2 LTR contains the *Hpa*I recognition sequence, any MAV2 integration close to the *twist* coding

sequence would manifest itself by an *Hpa*I fragment shorter than 11 kb. Rearrangement in the *twist* gene was confirmed in sample 102/2b and was also found in two additional independent tumor samples 107/la and 122/la (Figure 3)

Hybridization with an MAV LTR-specific HH25 probe confirmed the presence of MAV2 sequences in the rearranged *twist* loci-the same fragments were detected by both the *twist* ME28 and MAV HH25 probes (Figure 3). The almost identical size of these rearranged *twist* DNA fragments indicates that all three MAV2 proviral integrations occurred within a very narrow region of the 5′end of the *twist*.

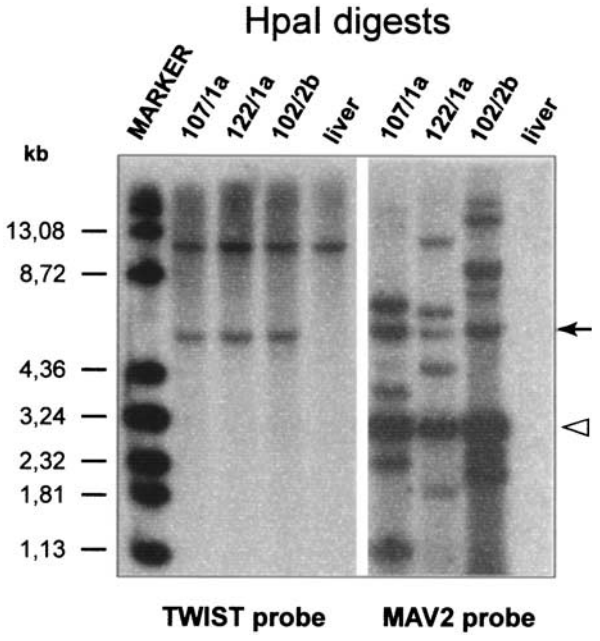


Figure 3 Detection of the MAV2 provirus integrated in the *twist* gene. Genomic DNAs from MAV2-induced tumors 107/la, 122/1a and 102/2b and from uninfected liver were digested with the *Hpa*I restriction endonuclease and Southern blots were hybridized with *twist*-specific ME28 (left panel) or MAV2-specific HH25 (right panel) probes. The arrow depicts the *twist*/*MAV2* rearranged fragment; the arrowhead marks the 2.9 kb internal fragment originating from the 5′ part of the MAV2 proviruses

In summary, about 4% of tumor samples have the *twist* gene rearranged by MAV2 integration. On the contrary, analogous experiments using a *dermo* probe have shown no rearrangement in the *dermo* gene (data not shown).

Defective MAV2 proviruses are integrated within 40 bp of the twist 5′UTR

To find out the exact position and structure of integrated defective proviruses in 122/la, 102/2b and 107/la DNAs, the virus–host DNA junctions were amplified using pTWs + LTRleft and LTRright + pTWa primers, and the PCR products were sequenced. The results confirmed the presence of the provirus within the *twist* 5′UTR in all three tumor samples. All three proviruses integrated in the *twist* loci underwent extensive recombination/deletion changes schematically represented in Figure 4. At least in the 122/1a sample, the resulting structure must have been achieved by several successive recombination steps.

The 5′end of the 102/2b provirus is flanked by 168 nucleotides that represent a duplication of the −71 to + 97 *twist* sequence in the same orientation. The provirus itself is formed by the last 40 codons of the *env* gene, followed by the 3′UTR and the complete LTR. The 107/la provirus is composed solely of the complete LTR. The 122/la provirus contains at its 5′end the incomplete and inverted LTR followed by the 3′end of *gag* and the abutting part of *pol* genes, to which the very end of *env,* 3′UTR, and the complete LTR were joined. The provirus–*twist* junctions in 107/la and 122/la DNAs are formed by the duplicated *twist* sequence (CCCTCC), while no integration site duplication was found in 102/2b DNA.

The sequence analysis of integration sites revealed the presence of the GC-rich stretch of 50 nucleotides with no homology to either MAV2 (accession number L10924) or the published *twist* genomic and cDNA sequences (accession numbers Y08261, AF093816). The integration of 102/2b, 107/la and 122/la proviruses occurred within this newly described region of the *twist* 5′UTR, at nucleotides −71, −44 and −35, respectively (Figure 5).

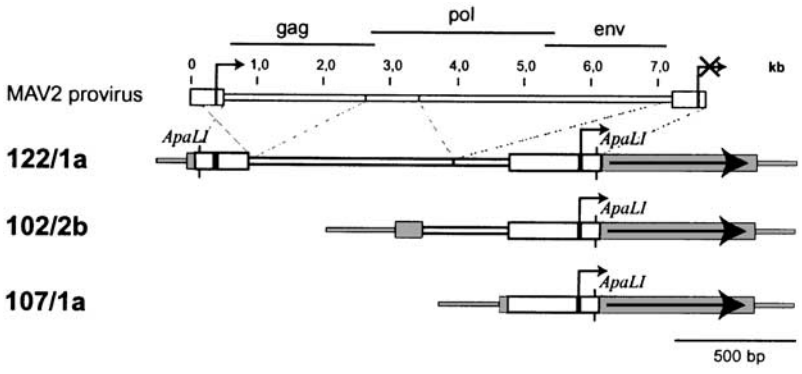


Figure 4 Schematic representations of MAV2 provirus deletions in 122/1a, 102/2b and 107/1a tumors. See the text for details

To confirm the genomic origin of novel 50 nucleotides, the sequence of the 5′end of normal chicken *twist* was determined. As a result of an extremely high GC content, this region could not be correctly amplified and sequenced by conventional PCR-based procedures. Using the modified PCR buffer containing 1.3 M betain and 1.3% DMSO (Baskaran *et al.,* 1996), we confirmed the presence of additional 50 nucleotides (nucleotides −22 through −71 in Figure 4, GenBank accession number AY126449) in the chicken genome.

In summary, the results define the sequence coding for the *twist* 5′UTR as the common integration site of MAV2 in experimental chicken nephroblastoma, and demonstrate that in all cases proviruses contain a single functional LTR that can activate transcription of the entire *twist* coding sequence.

Transcription of the twist gene is generally upregulated in nephroblastomas

To assess the levels of expression of *twist* and of other potentially relevant genes, Northern blot and RT–PCR analyses of tumor RNAs were performed (Figure 6). It was observed that in the 102/2b, 107/1a and 122/1a samples, *twist* mRNA synthesis was extremely high, more than 200 times higher in comparison with nontransformed, morphologically normal MAV2-infected samples and with uninfected kidney, where the *twist* mRNA was barely detectable. (The result with sample 102/2b is not shown, because the RNA was severely degraded and the expression estimate was based on the abundant signal of low molecular weight degradation products.) The vast majority of other tumor samples also overexpressed *twist* mRNA

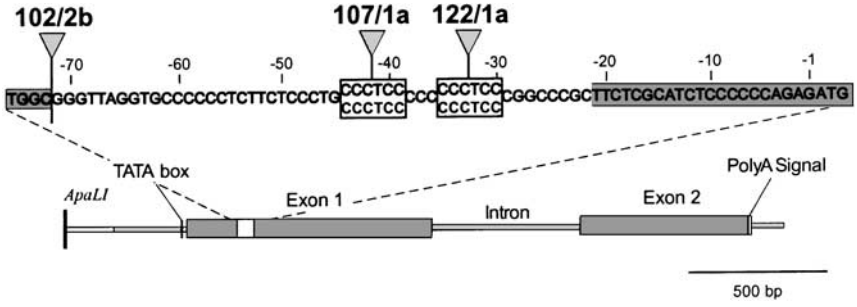


Figure 5 Nucleotide sequence of the novel *twist* DNA segment (positions −22 to −71) flanked by the so far known *twist* sequence (shadowed boxes). Bold ATG (+1 to +3) denotes the *twist* translation initiation codon. Vertical arrowheads indicate sites of MAV2 integration in 102/2b, 107/la and 122/la tumors. Duplicated CCCTCC sequences designate terminal duplications associated with integration of 107/la and 122/la proviruses. The lower scheme shows the localization of the sequence within the *twist* locus

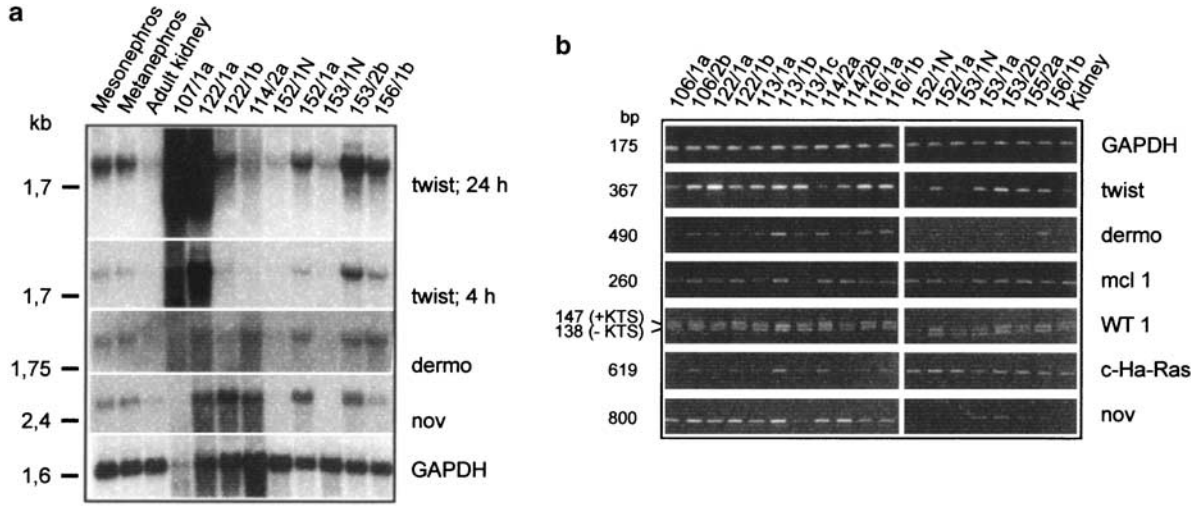


Figure 6 Comparison of mRNA levels of indicated genes in nephroblastomas and various control tissues. (**a**) Northern blot analysis. Twist 24 and 4 h panels show 24- and 4-h exposures of the blot hybridized with the ME28 *twist* probe. The blot was then stripped, rehybridized with the *dermo* probe and exposed for 4 days. After the second stripping the blot was hybridized with the *nov* probe (24 h exposure), stripped again and hybridized with the GAPDH probe. (**b**) Semiquantitative RT–PCR analysis. PCR primers were used to amplify sequences related to nephroblastomas. The number of amplification cycles used are as follows: GAPDH (24), *twist* (26), *dermo* (28), *Mcll* (25), *WT1* (26), *c-Ha-ras* (23), *nov* (30). PCR products were stained with ethidium bromide, visualized by UV illumination and photographed

(Figure 6a Twist panels), in the range of 2–100. Mesonephric and metanephric kidney contained a rather high amount of *twist* mRNA, reflecting their immature nature and a high content of primitive nephrogenic mesenchyme. In contrast, the majority of tumor samples contained low levels of the *twist* homolog, *dermo* mRNA. As reported previously, *nov* was overexpressed in most tumors and in embryonic kidney (Joliot *et al.,* 1992).

These results were confirmed by semiquantitative RT–PCR (Figure 6b). The other genes reported to be potentially involved in either chicken nephroblastoma (c-*Ha-ras*) or in human Wilms tumors (*WT1, mcl-1*) were also included; their expression showed no consistent differences between tumor and normal tissues.

In summary, out of the analysed genes only *twist* and *nov* upregulation correlates with the tumor and embryonic phenotypes.

Hyperexpression of the twist gene is driven by the MAV2 promoter

To identify the promoter that drives the unusually high expression of the rearranged *twist* gene, RNAs from nephroblastomas 107/1a and 122/1a were analysed by RT–PCR using primer pTWa in combination with either BSUright or LTRright primers. The results in Figure 7b clearly document that the *twist* mRNA synthesis starts within the sequence delimited by BSUright and LTRright primers, most probably from the 'R' site of LTR—the natural start site of proviral transcription—and is not initiated in the *twist* promoter localized upstream from the integrated proviruses. This observation confirms the prediction based on the structure of the MAV2 proviruses and the strong promoter activity of the viral LTR.

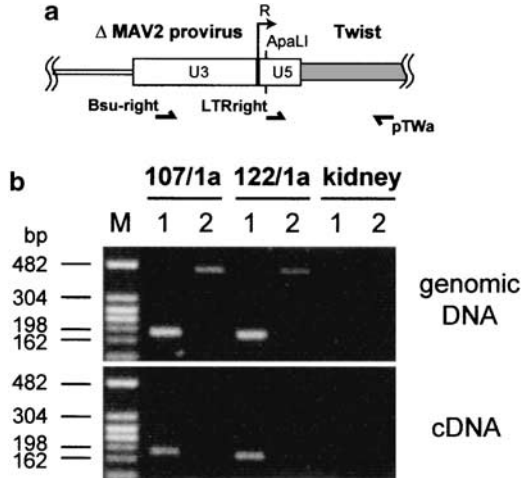


Figure 7 *Twist* transcription in tumors 107/la and 122/la is driven by MAV2 LTR. (**a**) Scheme of the integrated MAV2 3′LTR (MAV2 provirus) in the *twist* gene. The position and orientation of 3′OH ends of PCR primers BSUright, LTRright and pTWa is given by demi arrows. The R site of LTR at which the transcription probably starts is indicated; (**b**) Transcription in the tumor samples with proviral integration into the *twist* locus starts from the proviral LTR. Agarose gel electrophoresis of the genomic PCR and RT–PCR products amplified from tumor samples 107/la, 122/la and from normal uninfected kidney. Lanes: M, marker; 1, primer pair LTRright + pTWa; 2, primer pair BSUright + pTWa

Discussion

Avian nephroblastoma induced by MAV2 retrovirus has been an example of an experimental clonal tumor with a multihit etiology and diverse morphology and histopathology. This tumor is thought to be derived from nephrogenic rests-remnants of embryonic renal mesenchyme that persists in the chicken kidney for several weeks after hatching. The transformation seems mainly to overcome the homeostatic control of the kidney tissue and to support proliferation of embryonic mesenchymal cells in kidney without blocking their terminal differentiation. Nephrogenic cells in mesenchymal stroma, the major component of the nephroblastomas, give rise to abortive structures mimicking different stages of nephron development. In more developed tumors they frequently differentiate into structures resembling cartilage, bone, keratinizing epithelia or fibrosarcoma (Boni-Schnetzler *et al.,* 1985; Ishiguro *et al.,* 1962). Thus, the nephroblastomas display typical characteristics of embryonic tumors.

It has been assumed that the malignant renal cell arises as late as several specific regulatory pathways in it have been distorted by MAV2 integration. Since these pathways are constituted by cascades of functionally connected genes, the provirus does not have to hit one particular gene in order to deregulate the pathway. That is probably why the efforts to identify the crucial nephroblastoma-specific integrations had only a limited success (Westaway *et al.,* 1986; Joliot *et al.,* 1992). That is also probably why nephroblastomas display histopathological variations, since distinct deregulated genes have a different impact on the phenotype of transformed cells. Nevertheless, deregulation of some specific genes might contribute to malignant transformation more efficiently than activation/inhibition of others, and such genes should constitute a set of 'common integration sites' of the MAV2 provirus in nephroblastoma. Since retroviral integration is in principle site-unspecific (Coffin *et al.,* 1997), the existence of a common site of integration found in a limited number of independent tumor clones must be a result of the selection process: only cells in which the MAV2 provirus has hit the proper gene or combination of genes give rise to a tumor.

In our screen we have found the *twist* gene to be one of the common sites of MAV2 proviral integration in chicken nephroblastoma.

The function of *twist* in immature renal cells is not known. However, Twist has been shown to regulate the fate of mesodermal cells from *Drosophila* to vertebrates (Castanon and Baylies, 2002). For instance, it has been reported to negatively regulate differentiation in myogenesis (Spicer *et al.,* 1996) and osteogenesis (Lee *et al.,*

1999) and to inhibit p53-dependent apoptosis (Maestro *et al.,* 1999). These activities and the relatively high expression of *twist* in mesonephric and metanephric kidney may implicate this gene in the maintenance of the immature phenotype and growth of embryonic renal tissue. Owing to these properties, the abnormally upregulated *twist* could become an oncogene. This notion is strongly supported by our observations in 102/2b, 107/1a and 122/la nephroblastoma clones. First, in all three tumors, integration took place into the very narrow region of *twist* 5′UTR. In this way, the normal *twist* regulatory sequences have been replaced by the retroviral promoter. Second, the integrated proviruses underwent diverse, extensive and multiple deletions/rearrangements with the same consequence: elimination of 5′LTR, which, in the complete provirus, inhibits the strong transcriptional potential of 3′LTR (Coffin *et al.,* 1997). Thus, the provirus position together with its defects appear to ensure maximal transcription of the *twist* gene.

The aforementioned events must primarily be extremely rare. Only strong positive selection of cell clones carrying such an arrangement might explain its presence in about 4% of tumor clones. Indeed, the growth advantage of these clones seems to be strong, as 102/2b, 107/1a and 122/1 a tumor clones belonged to the largest tumors in our collection, ranging from 5 to 10 cm in diameter.

However, *twist* does not display the properties of typical avian retroviral oncogenes which transform target cells and induce tumors without a need for cooperation with other activated genes. Sequencing of overexpressed *twist* mRNAs from tumors did not reveal any mutation, and thus *twist* in nephroblastoma clones we analysed is rather the overexpressed proto-oncogene. This notion is in agreement with our preliminary experiments in which chicks were infected with retroviral expression vectors carrying *twist* cDNA. Such infection did not result in nephroblastoma induction (P Pajer, unpublished). We propose that another gene(s) must also be activated by MAV2 to induce nephroblastoma in cooperation with *twist*. Furthermore, the oncogenic activity of the *twist* proto-oncogene could become evident only at expression levels driven by 3′LTR of truncated MAV2 selected in 102/2b, 107/la and 122/la tumors. These expression levels were much higher than those achieved by a MAV-based retroviral vector (P Pajer, unpublished).

A high level of *twist* mRNA was also found in tumors with no apparent *twist* rearrangement. This might result from insertional activation of other gene(s) that directly or indirectly control(s) *twist* expression, or it might simply reflect a high proportion of twist-expressing immature blastema cells in the tumor. To distinguish between these two possibilities and to find other genes presumably cooperating with *twist,* experiments with cloning other MAV2 integration sites (shown in Figure 3) are in progress.

The role of Twist in the formation of human rhabdomyosarcoma has been suggested (Maestro *et al.,* 1999). We propose that activated *twist* might also take part in the formation of Wilms tumors. The potential role of *twist* in Wilms tumors is under investigation.

In the end, the virtually unchanged expression profile of WT1 that we have observed in chicken nephroblastomas probably reflects the differences in function of the gene in mammals and birds. This observation is also supported by the lack of the first of the two WT1 alternative exons present in its mammalian counterpart, whose participation in Wilms tumor genesis has been proposed previously (Scharnhorst *et al.,* 2001).

Materials and methods

Experimental animals

Embryos 12 days old or chicks 2 days old (outbred Brown Leghorns or closely related inbred CB and CC White Leghorns; Plachý and Hála, 1997) were infected by MAV2 viral stock injected into the chorioallantoic vein or intraperitoneally, respectively. The samples of nephroblastomas or control tissue were collected 45–120 days later. Pronephric and mesonephric kidneys were collected from 14-day-old chicken embryos.

DNA and RNA isolation, Southern- and Northern-blot hybridization

Isolation of high molecular weight genomic DNAs, restriction enzyme digestion, agarose gel electrophoresis, and Southern-blot hybridization were performed by standard methods (Ausubel *et al.,* 1993) modified as described in Pečenka *et al* (1988a), except that the Zeta Probe membrane was used instead of nitrocellulose (BioRad, Richmond CA, USA).

Total RNAs were isolated using the TRIzol reagent according to the manufacturer's instructions (Invitrogen Corporation, Carlsbad CA, USA). In all, 10 μg of total RNA per sample was fractionated by electrophoresis in 1.2% agarose gels containing formaldehyde and transferred to GeneScreen membranes (NEN, Boston, MA, USA). The membranes were prehybridized and hybridized in ULTRAhyb buffer (Ambion, Austin, Texas, USA) according to the manufacturer's instructions. Blots were then exposed to X-ray film at -70°C with an intensifying screen.

Probes used on Southern and Northern blots

HH25 (MAV2 U3 region)—nucleotides (18–271) of the MAV2 provirus

BB43 (MAV2 5′UTR)—nucleotides (396–830) of the MAV2 provirus

AH24 (MAV2 3′UTR)—nucleotides (7119–7358) of the MAV2 provirus

ME28 (*twist* first exon)—nucleotides (+298–+1217) of the *twist* coding sequence

MAV2 probes originated from pATV-MAVl (Pečenka *et al.,* 1988), ME was subcloned from pBtwist, a *twist* cDNA clone prepared by RT–PCR (P Pajer, unpublished).

The subcloning of probes, recombinant plasmid isolation and gel purification of the probes were performed by standard methods Ausubel *et al.* (1993) with modifications according to Pečenka *et al* (1988c) using the pBluescript vector and SURE

bacteria (Stratagene, La Jolla, CA, USA). The probes were labeled by nick-translation reaction using [α-^{32}P]dCTP (Amersham Pharmacia Biotech, Little Chalfont, England) according to the published procedure (Pečenka *et al.,* 1988a).

Primers

LTRleft(5′-GCATCAGGCGATTCCCTTATTTGG-3′)
LTRright(5′-GGCCGGACCGTCGATTCCCTGA-3′)
BSUright(5′-CCCATTGGTGGCGAAGGAGCGAC-3′)
pTWs(5′-AGCACCCCACAGCAGTGAGAGAAC-3′)
TW-N (5^{1}-ATGATGCAGCAGGACGAGTCAAAC-3′)
TW-X(5′-CACAACACGGTATCCAACTTCAGAG-3′)
pTWa(5′-CCGGTCCGGCTCCTCTTCGCTGTTG-3′)

Inverse PCR cloning

A total of 10 μg of genomic DNA from each selected tumor sample was digested by the *Apa*LI restriction enzyme, fractionated by agarose electrophoresis in 2 × TBE buffer + 100 mM Na-acetate; DNA fragments of the approximate length corresponding to the *Apa*LI fragment containing the defective provirus were cut out and isolated from the gel. DNAs obtained were self-ligated (typically 100 ng DNA/50 μl of ligation mixture with 0.1 Weiss units of ligase, overnight reaction at 20°C). Ligation mixtures were extracted by phenol : chloroform (1 : 1), isopropanol-precipitated, the pellet was washed with 80% ethanol and resuspended in TE buffer. PCR reactions (20 μl each) were set out according to the manufacturer's instructions (Stratagene, La Jolla, CA, USA) using 100 ng of the circularized tumor DNA, 2 U of Taq polymerase and LTRleft + BSUright primers at a final concentration of 250 nM each and 0.2 mM Mg^{2+} concentration. In all, 30 cycles (95°C 30 s, 65°C 5 min) were performed; the resulting products were resolved by agarose gel electrophoresis, gel purified and cloned into the EcoRV linearized pBluescript SK(−) vector (Stratagene, La Jolla, CA, USA).

DNA sequencing and homology searches

All sequencing reactions were performed according to the manufacturer's instructions (PE Biosystems, Warrington, England) using a BigDye Terminator Cycle Sequencing Kit (v. 2) and an ABI PRISM 310 Sequencer. Sequencing of GC-rich *twist* gene fragments was done with the addition of betain and DMSO to final concentrations of 1.3 M and 1.3%, respectively.

Sequence homology searches were conducted at the DNA and protein levels using BLAST network service at National Center for Biotechnology Information. The DNA and protein alignments were performed using MacVector (Oxford Molecular Group, Beaverton, OR, USA).

Reverse transcription and PCRs

In all, 2 μg of total RNA were reverse transcribed using the primer $(dT)_{18}$ and the reverse transcriptase system from PROMEGA (Madison, WI, USA). PCR amplifications were performed in the presence of betain and DMSO (see DNA sequencing), using Taq polymerase and buffers from ROCHE (Mannheim, Germany).

Abbreviations

MAV, myeloblastosis-associated virus; RT–PCR, reverse transcription–polymerase chain reaction; LTR, long-terminal repeat; UTR, untranslated region

Acknowledgments

The authors thank Petr Bartůněk and Daniel Elleder for the critical reading of the manuscript, helpful comments and discussion. The assistance of Šárka Takáčová in manuscript preparation is also greatly acknowledged. This work was supported by grants 301/98/K042 and 204/00/0554 from the Grant Agency of the Czech Republic.

References

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA and Struhl K. (eds). (1993). *Current Protocols in Molecular Biology*. Green Publishing Associates, Inc.

Baskaran N, Kandpal RP, Bhargava AK, Glynn MW, Bale A and Weissman SM. (1996). *Genome Res.*, **6,** 633–638.

Boni-Schnetzler M, Boni J, Ferdinand F-J and Franklin RM. (1985). *J. Virol.*, **55,** 213–222.

Bourgeois P, Bolcato-Bellemin AL, Danse JM, Bloch-Zupan A, Yoshiba K, Stoetzel C and Perrin-Schmitt F. (1998). *Hum. Mol. Genet.* **7,** 945–957.

Castanon I and Baylies MK. (2002). *Gene,* **287,** 11–22.

Coffin JM, Hughes SH and Varmus HE. (1997). *Retroviruses.* Cold Spring Harbor Laboratory Press: Plainview, NY.

Collart KL, Aurigemma R, Smith RE, Kawai S and Robinson HL. (1990). *J. Virol.,* **64,** 3541–3544.

Dome JS and Coppes MJ. (2002). *Curr. Opin. Pediatr.,* **14,** 5–11.

el Ghouzzi V, Le Merrer M, Perrin-Schmitt F, Lajeunie E, Benit P, Renier D, Bourgeois P, Bolcato-Bellemin AL, Munnich A and Bonaventure J. (1997). *Nat. Genet.,* **15,** 42–46.

Evron E, Dooley WC, Umbricht CB, Rosenthal D, Sacchi N, Gabrielson E, Soito AB, Hung DT, Ljung BM, Davidson NE and Sukumar S. (2001). *The Lancet,* **357,** 1335–1336.

Howard TD, Paznekas WA, Green ED, Chiang LC, Ma N, Ortiz de Luna RI, Garcia Delgado C, Gonzalez-Ramos M, Kline AD and Jabs EW. (1997). *Nat. Genet.,* **15,** 36–41.

Ishiguro H, Beard D, Sommer JR, Heine U, de The G and Beard JW. (1962). *J. Natl. Cancer Inst.,* **29,** 1–17.

Joliot V, Khelifi C, Wyers M, Dambrine G, Lasserre F, Lemercier P and Perbal B. (1996). *J. Virol.,* **70,** 2576–2580.

Joliot V, Martinerie C, Dambrine G, Plassiart G, Brisac M, Crochet J and Perbal B. (1992). *Mol. Cell. Biol.,* **12,** 10–21.

Knudson Jr AG, Strong LC. (1972). *J. Natl. Cancer. Inst.,* **48,** 313–324.

Lee MS, Lowe GN, Strong DD, Wergedal JE, Glackin CA. (1999). *J. Cell. Biochem.,* **75,** 566–577.

Li L, Cserjesi P and Olson EN. (1995). *Dev. Biol.,* **172,** 280–292.

Maestro R, Dei Tos AP, Hamamori Y, Krasnokutsky S, Sartorelli V, Kedes L, Doglioni C, Beach DH and Hannon GJ. (1999). *Genes. Dev.,* **13,** 2207–1722.

Pečenka V, Dvořák M, Karafiát V, Šloncová E, Hložánek I, Trávniček M and Říman J. (1988a). *Folia Biol. (Praha),* **34,** 147–169.

Pečenka V, Dvořák M, Karafiát V and Trávniček M. (1988c). *Folia Biol. (Praha),* **34,** 215–232.
Pečenka V, Dvořák M and Trávniček M. (1988c). *Folia Biol. (Praha),* **34,** 129–146
Perbal B. (2001). *Mol. Pathol.,* **54,** 57–79.
Plachý J, Hála K. (1997). *Folia Biol. (Praha),* **43,** 133–151.
Robinson HL, Gagnon GC. (1986). *J. Virol.,* **57,** 28–36.
Scharnhorst V, van der Eb AJ and Jochemsen AG. (2001). *Gene,* **273,** 141–161.
Spicer DB, Rhee J, Cheung WL and Lassar AB. (1996). *Science,* **272,** 1476–1480.
Thisse B, Stoetzel C, Gorostiza-Thisse C and Perrin-Schmitt F. (1988). *EMBO J.,* **7,** 2175–2183.
Westaway D, Papkoff J, Moscovici C and Varmus HE. (1986). *EMBO J.,* **5,** 301–309.

Identification of Potential Human Oncogenes by Mapping the Common Viral Integration Sites in Avian Nephroblastoma

Petr Pajer,[1] Vladimír Pečenka,[1] Jarmila Králová,[1] Vít Karafiát,[1] Dana Průková,[1] Zdena Zemanová,[2] Roman Kodet,[3] and Michal Dvořák[1]

[1]Institute of Molecular Genetics AS CR; [2]Institute of Physiology AS CR; and [3]Institute of Pathology and Molecular Medicine, 2nd Faculty of Medicine, Charles University, Prague, Czech Republic

Abstract

Gene deregulation is a frequent cause of malignant transformation. Alteration of the gene structure and/or expression leading to cellular transformation and tumor growth can be experimentally achieved by insertion of the retroviral genome into the host DNA. Retrovirus-containing host loci found repeatedly in clonal tumors are called common viral integration sites (cVIS). cVIS are located in genes or chromosomal regions whose alterations participate in cellular transformation. Here, we present the chicken model for the identification of oncogenes and tumor suppressor genes in solid tumors by mapping the cVIS. Using the combination of inverse PCR and long terminal repeat-rapid amplification of cDNA ends technique, we have analyzed 93 myeloblastosis-associated virus type 2–induced clonal nephroblastoma tumors in detail, and mapped >500 independent retroviral integration sites. Eighteen genomic loci were hit repeatedly and thus classified as cVIS, five of these genomic loci have previously been shown to be involved in malignant transformation of different human cell types. The expression levels of selected genes and their human orthologues have been assayed in chicken and selected human renal tumor samples, and their possible correlation with tumor development, has been suggested. We have found that genes associated with cVIS are frequently, but not in all cases, deregulated at the mRNA level as a result of proviral integration. Furthermore, the deregulation of their human orthologues has been observed in the samples of human pediatric renal tumors. Thus, the avian nephroblastoma is a valid source of cancer-associated genes. Moreover, the results bring deeper insight into the molecular background of tumorigenesis in distant species. (Cancer Res 2006; 66(1): 78-86)

Introduction

The identification of genes actively contributing to cellular transformation has been the key step in understanding the process of malignant transformation. New names are continually being added to the list of known oncogenes, tumor suppressor genes, and stability genes, reflecting the complexity of genetic changes behind the scene of malignant transformation.

Generally, tumors can be classified as hematopoietic (i.e., leukemias and lymphomas) or solid (i.e., tumors whose cells are normally immobile; ref. 1). Solid tumors prevail (roughly 90% of spontaneous human tumor cases). On the contrary, the majority of currently confirmed oncogenes (~90%) have been discovered in hematopoietic tumors (2). This disparity stems from the high complexity of genetic aberrations in solid tumors—their formation is supposed to require far more changes (chromosomal rearrangements, amplifications, submicroscopic and point mutations, epigenetic changes, etc.) compared with hematopoietic tumors. This complexity, as well as the high histologic heterogeneity of solid tumors, makes them difficult to analyze in molecular detail. Due to this complexity, the causal relationships between a genetic aberration and a phenotype of solid tumors are rather little understood.

Oncogenic retroviruses are a potent tool for the identification of cancer-causing genes as well as for further study of their oncogenic potential. These retroviruses are divided into acute (retroviruses directly transducing the mutated form of a host proto-oncogene) and nonacute (not containing a virally transduced oncogene) that induce oncogenic transformation through the insertional mutagenesis. The progressive strategy using the nonacute retroviruses for simultaneous identification of multiple candidate cancer-causing genes in a given animal tumor model is called retroviral tagging. Retroviruses integrate into the host genome almost randomly; thus, each host gene locus is hit by the provirus integration in many cells of the target tissue at different positions. Proviruses that integrate in the vicinity of a gene can influence its expression through potent viral regulatory sequences. Proviruses that integrate into a gene coding sequence can either inactivate the gene or, through gene truncation, change its function. Certain integrations (or their proper combinations) provide a cell with a growth advantage, the cell clone expands, giving rise to a tumor (3). Analysis of such tumor clones allows for the identification of integration sites of individual proviruses in the host genome. Provirus-containing loci repeatedly selected in clonal tumors (common viral integration sites—cVIS) contain genes whose alterations contribute to the cellular transformation.

Thus far, the use of retroviral tagging has been limited to only a few model tumors; comprehensive analyses have been done only on murine hematopoietic disorders (4–6). Here, we describe the model representing solid tumors: myeloblastosis-associated virus type 2 (MAV-2)–induced chicken nephroblastoma. MAV-2 is an avian replication-competent nonacute oncogenic retrovirus. In chickens infected in ovo or early after hatching it induces, with high efficiency, multiple clonal embryonic-type tumors of kidney—nephroblastomas (7).

The chicken nephroblastoma model has proved as highly efficient because multiple clonal tumors are obtained from a single infected animal (8). The simultaneous use of two independent techniques—inverse PCR and long terminal repeat-rapid amplification of cDNA

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Requests for reprints: Michal Dvořák, Institute of Molecular Genetics, Flemingovo nám. 2, Prague 6, 166 37 Czech Republic. Phone: 42-22018-3468; E-mail: mdvorak@img.cas.cz.

doi:10.1158/0008-5472.CAN-05-1728

ends (LTR-RACE)—enabled us to identify VIS in chicken nephroblastomas with >90% efficiency, and the recent completion of the chicken genome draft sequence (9) enabled precise localization of the majority of tagged VIS. This approach led to the identification of a number of cVIS in addition to the 5′-untranslated region of the *twist* gene which we have previously described (10). The expression levels of selected candidate tumor-related genes identified in this model were determined in chicken nephroblastoma and human renal tumor samples. Transcription of some genes was found abnormal in tumors from both organisms. This shows the suitability of the chicken model for the identification of human genes potentially involved in the formation of human solid tumors.

Materials and Methods

Chicken nephroblastoma induction and sample collection. Closely related inbred CB and CC White Leghorns or outbred Brown Leghorns (11) were used for *in vivo* experiments. Twelve-day-old embryos or 2-day-old chicks were infected by a MAV-2 viral stock injected into the chorioallantoic vein or i.p., respectively. Control animals were mock-infected by an identical volume of PBS. The samples of nephroblastomas or control tissues were collected 45 to 120 days later, weighed and processed immediately into DNA, RNA, and paraffin sections. Three independent samples were taken from tumors >3 cm in diameter to check the clonal uniformity of the tumor. Excessive tissue was frozen in liquid nitrogen and stored at −70°C. Control mesonephric and metanephric kidney samples were collected from 14-day-old chicken embryos.

DNA and RNA isolation, probe preparation, and Northern blot analyses. Genomic DNA was obtained by lysing the chicken tissues in DNA lysis buffer (1% SDS, 250 mmol/L EDTA, and 1 mg/mL proteinase K) and incubated at 55°C overnight. The solution was extracted once with phenol-chloroform, the water phase precipitated by an equal volume of 96% ethanol, collected using a glass capillary, rinsed in 80% ethanol and resuspended in 10 mmol/L Tris-Cl (pH 8.3) and 1 mmol/L EDTA. Restriction enzyme digestions, agarose electrophoresis, and Southern blotting were done by standard methods (12).

For the RNA preparation, the fresh chicken tissues or frozen human samples were rapidly lysed in TRIzol reagent and total RNA was isolated according to the manufacturer's instructions (Invitrogen Corporation, Carlsbad, CA). For Northern blotting, 10 μg of total RNA per sample were fractionated by electrophoresis in 1.2% agarose gels containing formaldehyde and transferred to GeneScreen membranes (NEN, Boston, MA) and fixed by UV (1,200 J/cm^2). The membranes were prehybridized and hybridized in ULTRAhyb buffer (Ambion, Austin, TX) according to the manufacturer's instructions. Blots were then exposed to X-ray film at −70°C with an intensifying screen (Kodak, Rochester, NY).

To obtain hybridization probes, gene-specific oligonucleotide primers were derived from selected human or chicken genes and used for RT-PCR amplification of gene-specific fragments 300 to 600 bp in length. Every particular PCR product was cloned into the pUC19 cloning vector (New England Biolabs, Beverly, MA) and its identity was verified by sequencing. Individual inserts were excised from the vector by the appropriate restriction endonucleases, resolved in LMP agarose gel, excised, and isolated by the standard phenol extraction procedure (12).

Paraffin-embedded samples and histologic analyses. Paraffin-embedded samples and microscopic preparations were made as described elsewhere (13), stained with H&E, and microscopically examined for the presence, quantity, and quality of the tubules, glomerules, and stromal cells (interstitium). Based on these characteristics, the samples were divided into three major classes (I-III; Fig. 1) likely representing different tumor grades.

LTR-RACE and inverse PCR. Two micrograms of total RNA from each chicken sample were reverse-transcribed using primer 3′-CDS (SMART RACE cDNA amplification kit; Clontech, Palo Alto, CA) and the reverse transcriptase from Promega (Madison, WI), resulting cDNAs were diluted to 50 μL with water. One microliter of cDNA was used as a template for the PCR (15 μL reactions). The first round of PCR was done with primers LTR1 (5′-GGTGTGCACCTGGGTTGATGGC-3′), UPM (SMART RACE cDNA amplification kit; Clontech) and AccuTaq polymerase mix (Sigma, St. Louis, MO) for 25 cycles according to the manufacturer's instructions. PCR products were resolved in LMP agarose, visible bands (typically three to five host sequence-containing fragments) were excised and ~1/100 of each was used as a template for additional 20 PCR cycles with nested primers LTR2 (5′-GGCCGGACCGTCGATTCCCTGA-3′) and NUP (SMART RACE cDNA amplification kit; Clontech), 250 nmol/L each. Resulting individual PCR products were finally resolved on LMP agarose, excised, isolated by the phenol extraction procedure (12) and directly sequenced with primer LTR2, as described below.

For inverse PCR, 500 ng of genomic DNA from each sample was double-digested with *Bst*YI and *Bcl*I restriction enzymes, self-ligated in a volume of 200 μL and linearized with the *Apa*LI digestion. One hundred nanograms of the product were used as a template for PCR reaction under the following conditions: 2 units of AccuTaq polymerase (Sigma) in 20 μL of AccuTaq buffer supplemented with 500 μmol/L of deoxynucleotide triphosphates, 500 nmol/L primers LTR2 and LTR3 (5′-GGTGCATCAGGCGAATCCCTTA-TTTGG-3′), 1.2 mol/L betain, and 1.2% DMSO. PCR cycles were as follows: 94°C for 20 seconds, 23 cycles (94°C for 20 seconds, 65°C for 8 minutes); plus additional prolonged cycles: 94°C for 20 seconds, 65°C for 12 minutes and 94°C for 20 seconds, 65°C for 20 minutes. PCR products were resolved on LMP agarose, individual bands were excised, and DNA was isolated.

DNA sequencing and homology searches. All sequencing reactions were done unidirectionally with oligonucleotide primer LTR2 according to the manufacturer's instructions (PE Biosystems, Warrington, England) using BigDye Terminator Cycle Sequencing Kit (v. 3) and ABI PRISM 310 Sequencer. High-quality noncomposite sequences were edited using the Chromas v1.42. In the case of inverse PCR, left LTR flanking fragments were resequenced with the LTR3 primer to reveal the exact sites of integration.

Sequences of chicken *plag1* and *foxP1* cDNAs were determined by standard cloning and sequencing of RT-PCR and RACE PCR products obtained from cDNA of a chicken embryonic kidney. The entire coding sequences were deposited to the National Center for Biotechnology Information (AY935990 for *plag1* and AY935991 for *foxP1*).

Sequence homology searches were conducted at the DNA and protein levels using BLASTN algorithms on the chicken genome assembly (ENSEMBL project, http://www.ensembl.org/) and on all publicly available chicken expressed sequence tags at the University of Manchester Institute of Science and Technology (14). Significant hits were considered as those having the EXPECT value $\leq 10^{-5}$. The local DNA and protein alignments were done using MacVector (Oxford Molecular Group, Beaverton, OR).

Patient samples. Surgical resection specimens were obtained from 18 patients undergoing surgery at the Motol Hospital, Prague. After resection, part of the material was immediately snap-frozen in liquid nitrogen and stored at −80°C. Twelve sections 8-μm-thick were cut from each sample in a cryostat. The first and the last sections were stained with H&E, microscopically examined and diagnosed. The remaining 10 slices were lysed in 1 mL of TRIzol reagent (Sigma) and processed to obtain total RNA as described above. Typically, 10 to 50 μg of total RNA per sample were obtained.

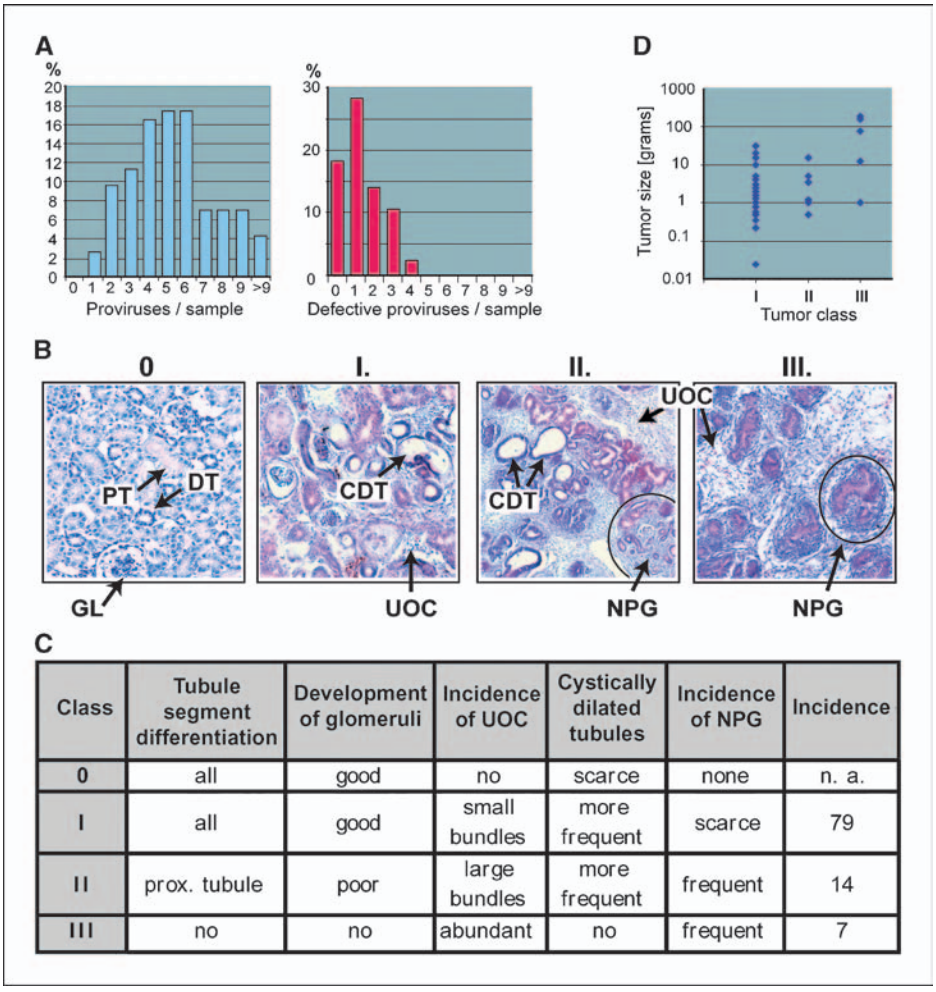
Results

The chicken nephroblastoma collection and tumor classification. About 250 nephroblastomas ranging in mass from 25 mg to >200 g were obtained from MAV-2-injected chicks. To exclude the possible effect of the strain, both the inbred White Leghorn and outbred Brown Leghorn chicks were used. Samples of each tumor were taken and used for preparation of DNA, RNA, and a paraffin-embedded tissue specimen. Their clonality and the numbers of complete and defective clonally integrated proviruses were determined by Southern blot analysis (15). Tumors most frequently contained four to six proviruses, some of which were often defective (Fig. 1*A*). Defective proviruses were pursued, because in chickens, most if not all oncogene-activating proviruses had been

shown to be defective (10, 13, 16). Ninety-three samples representing independent clonal tumors and containing eight proviruses at most were selected for further analyses, 55 coming from White Leghorn and 38 coming from Brown Leghorn chicken.

Paraffin sections stained with H&E were examined for the presence of abnormal structures that had been already described (13). The most prominent alterations we noticed were imperfectly differentiated tubules—with or without glomeruli, smaller or larger aggregations of unorganized and apparently undifferentiated cells and unusual spherical formations not known from normal nephrogenesis but reminiscent of origins of the normal tubule formation. We call them nests of pseudonephrogenesis. In more differentiated nephrons supplemented with glomeruli, the cystic dilations of nephron tubular segments appeared frequently. Based on qualitative and quantitative representations of these structures, samples were divided into four major classes, 0, I, II, and III (Fig. 1*B* and *C*). Class 0 samples are nonclonal (as revealed by Southern blots) and represent infected tissue with a prevalence of normal renal structures and sporadic cystic dilations of tubules. The nests of abnormal nephrogenesis, which were considered the most evident symptoms of malignant transformation, were missing in class 0, but were constantly present in all other classes. Classes I and II included tumors with more or less differentiated nephrons, respectively, various numbers of cystic dilations of tubules, and a growing proportion of unorganized cells. Samples belonging to class III contained only the nests of pseudonephrogenesis and unorganized cells. There was a correlation between the tumor morphology class and the tumor size. Forty nephroblastomas were distributed into classes I to III, as described above, and ordered according to their mass (Fig. 1*D*). In general, class III members, the least differentiated nephroblastomas, clearly displayed greater size, although a rather high size variation within each class was registered. We suggest that the size variation is mainly caused by different growth rates of each individual tumor clone and not by a different time of a target cell infection because the pool of target cells for transformation (nephrogenic blastema cells) fades away rapidly within the first few days after hatching.

Identification of MAV-2 integration sites by inverse PCR and LTR-RACE. There are three principal techniques suitable for large-scale identification of retroviral integration sites—inverse PCR (17), linker-mediated PCR (18), and LTR-RACE (19). Inverse PCR and linker-mediated PCRs start with genomic DNA and amplify sequences adjoining to the site of integration. We decided to use and optimize inverse PCR reaction conditions so that LTR-flanking fragments (both left and right) of all integrated proviruses were amplified in one-step PCR reaction with equal efficiency (Materials and Methods; Fig. 2*A*). In comparison, LTR-RACE selectively amplifies only transcribed sequences downstream from the LTR promoter of the integrated provirus—the transcripts whose synthesis are driven by the LTR promoter and which contain LTR sequences. There are three main types of mRNA transcribed from integrated MAV proviruses. The first type contains only MAV



C

Class	Tubule segment differentiation	Development of glomeruli	Incidence of UOC	Cystically dilated tubules	Incidence of NPG	Incidence
0	all	good	no	scarce	none	n. a.
I	all	good	small bundles	more frequent	scarce	79
II	prox. tubule	poor	large bundles	more frequent	frequent	14
III	no	no	abundant	no	frequent	7

Figure 1. Characterization of the nephroblastoma collection. *A,* the frequency (%) of total (*blue columns*) and defective (*red columns*) proviruses integrated in individual clonal tumors. Integrated proviruses were detected in 90 tumors by Southern blot hybridization. *B,* histologic picture (×100 magnification) of MAV-2-infected nontumor renal cortex (0) and chicken nephroblastoma samples representing classes I to III. *0*, differentiated nephron segments: proximal tubule (PT), distal tubule (DT), and glomerule (Gl). *I*, tumor comprised of differentiated PT, Gl, and besides cystically dilated tubules (CDT), and clusters of unorganized cells (UOC). *II*, tumor with prevalence of UOC and "nests of pseudonephrogenesis" (NPG). *III*, tumor composed exclusively of NPG and bundles of UOC. *C,* classification of nephroblastomas according to proportions of typical histopathologic attributes. The representation of each tumor class in the studied nephroblastoma collection is expressed in percentages. *D,* the relationship between tumor class (*x*-axis) and tumor size (*y*-axis, log scale). Fifty arbitrarily selected tumors were analyzed.

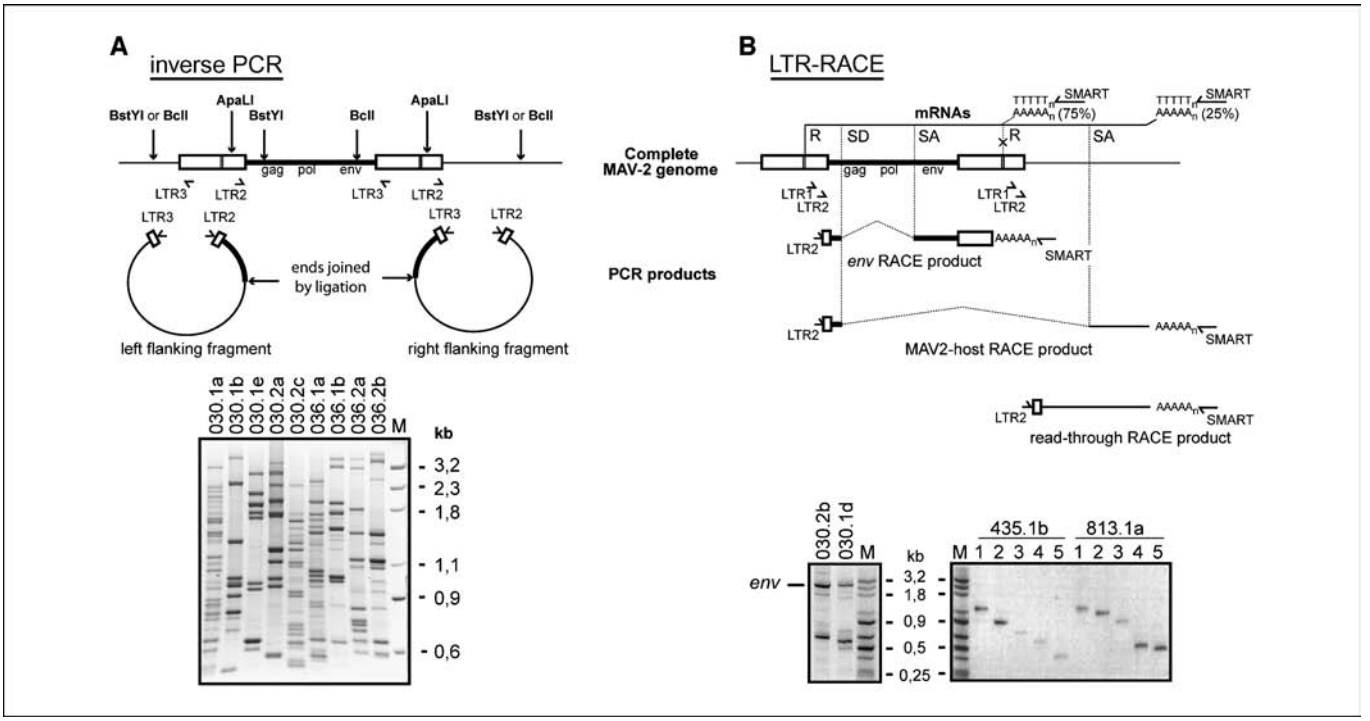


Figure 2. Schematic representations and typical results of PCR-based methods used in this work. *A,* inverse PCR: the integrated provirus (complete MAV-2 genome) containing long terminal repeats (*blank boxes*) and sequences coding for *gag, pol*, and *env* viral genes (*thick line*) is flanked by a host DNA (*thin line*). Combined *Bst*Y and *Bcl*I digestion generates fragments with compatible cohesive ends containing (*a*) the 5′-end of the provirus linked to a left-flanking fragment and (*b*) the 3′-end of the provirus linked to a right-flanking fragment. Self-ligation generates circular DNAs, which are then linearized by *Apa*LI. Using LTR3 and LTR2 oligonucleotide primers, DNAs are amplified (PCR products 1 and 2). Example of electrophoretic separation of amplified DNAs obtained from nine clonal tumors (*bottom*). In the majority of cases, each VIS is represented by two distinct PCR fragments. *B,* LTR-RACE: the integrated provirus [complete MAV-2 genome as in (*A*)] is transcribed from the R site within the left LTR. mRNAs (sketched above the MAV-2 genome) are terminated either at the right LTR termination signal (75% of transcripts) or at a termination signal in a downstream host sequence (25% of transcripts). Some transcripts are spliced by joining the splice donor site (SD) of the *gag* sequence and a splice acceptor site (SA) in *env* or within a downstream host gene. All these transcripts are converted to cDNAs using SMART RACE kit (Clontech). Nested amplification with SMART RACE kit using LTR1 and subsequently LTR2 primers yields PCR products: *env* RACE, MAV2-host RACE, and readthrough RACE. Examples of electrophoretic separation of first round PCR products (tumors 030.2b and 030.1d) and individual isolated and nested PCR reamplified products (tumors 435.1b and 813.1a; *bottom*).

sequences (complete genome or spliced *env* mRNAs). The second type is composed of MAV sequences fused to downstream host sequences. Such fusions are facilitated by the weakness of the MAV LTR polyadenylation signal that allows the read-through in 20% to 30% of transcripts.[4] A portion of fusion transcripts are further processed by a splicing between the retroviral splice donor and host splice acceptor sites. The third type are transcripts initiated by defective and rearranged proviruses containing various fragments of retroviral genomes fused to host sequences. The LTR-RACE technique we used amplifies fragments of the abovementioned mRNAs demarcated by primers LTR1 and SMART. Individual PCR products were isolated, reamplified, and directly sequenced with nested LTR2 and SMART primers (Materials and Methods; Fig. 2*B*).

Discrete PCR fragments obtained by either method have been separated on high-resolution agarose gel and directly sequenced by the LTR2 primer. Direct sequencing of the PCR products eliminates isolation of nonclonal background VIS that would emerge during the cloning procedures.

To determine the efficiency of inverse PCR and LTR-RACE, we compared the number of VIS detected by either method with the number of clonal proviruses shown by Southern blot hybridization. The LTR-RACE approach detected, on average, one-third of the integration events found by Southern blot hybridization. On the other hand, LTR-RACE amplifies proviral sequences splice-joined with sequences of the affected gene even if the VIS is separated by several kilobases from the gene. In comparison, inverse PCR enabled isolation of all VIS from the vast majority of tumor samples, but obviously provided no information about the structure of potential transcripts. Thus, the combination of inverse PCR and LTR-RACE leads to the most complete picture of retroviral integrations and their effect on gene expression in each tumor.

Genomic localization of VIS in clonal neproblastomas. To determine the precise genomic location of individual obtained sequences, the BLAST search in chicken genomic and expressed sequence tag databases was done. Left and right provirus flanking sequences were assigned to a single VIS according to two criteria: genomic position and a duplication of six nucleotides at the site of integration (20). Ninety-two percent of VIS were unequivocally positioned with an EXPECT value (the statistical significance threshold for reporting matches against database sequences) of $\leq 10^{-20}$. The remaining 8% of the retrovirally tagged VIS displayed either no significant homology or the homology was not unequivocal (when flanking sequences were short or were derived from repetitive elements). By this approach, the total of 521 independent VIS was retrieved from 93 analyzed nephroblastomas. The complete list of results is provided as Supplementary data and at http://www.img.cas.cz/nfbl.

4 P. Pajer, unpublished data.

Among 521 VIS, 18 nephroblastoma candidate loci (i.e., gene loci whose modification by MAV retrovirus is possibly required for nephroblastoma formation) were identified according to the following rules: the candidate locus is either a common VIS or it is the single clonally expanded VIS detected in the tumor by two independent techniques (Southern blot hybridization and inverse PCR).

We consider a gene locus to be common VIS when it is hit by MAV integration in at least two independent tumor clones, including tumor clones published earlier (13, 21). In addition, we consider a nongene chromosomal segment to be cVIS when hit by MAV integration in at least two independent tumor clones at positions <20 kb distant from each other. The list of candidate loci is given in Table 1. Each candidate locus has been ascribed by numbers characterizing its pertinence to a particular chromosome and its position within chromosomal sequences (e.g., Nal 1-17 is a nephroblastoma-associated locus, chromosome 1, megabase 17). If multiple candidate loci were located within a given chromosomal segment, an additional lowercase letter was added. Candidate loci overlapping known genes have been entitled by respective gene names. It is important to mention that some common VIS could arise due to chance (6) or as a result of preferential integration into some genomic regions—preferences for transcriptionally active loci (22) and GC-rich DNA regions (23) have been reported.

***twist, plag1*, and *foxP1*, the most frequent targets, are affected by the retroviral integration in different ways.** We identified three frequently targeted cVIS, each found within a distinct candidate gene locus. These three genes encoding chicken transcription factors *plag1* (accession no. AY935990), *twist* (accession no. Y08261, AY126449), and *foxP1* (accession no. AY935991) harbored MAV provirus insertions with frequencies of 6%, 4%, and 5%, respectively, suggesting that their deregulation contributes significantly to the nephroblastoma formation. The structure of each mRNA has been determined by the combination of RT-PCR and 5′- and 3′-RACEs and coding sequences have been deposited in GenBank under the accession numbers mentioned above. Figure 3*A* depicts genomic structures of *plag1, twist*, and *foxP1* drawn on the basis of comparison of genomic and cDNA sequences with marked positions and orientations of individual VIS. Figure 3*B* shows the expression of *plag1, twist*, and *foxP1* genes in selected tumors where these genes were or were not hit by integration. Closer examination of the mRNA structure and expression levels of these genes point out to different mechanisms of retroviral mutagenesis employed in each of the three loci.

In case of *plag1*, proviral integrations were found up to 30 kb upstream from the initiation codon. The resulting mRNAs are generated by the splicing which joins the *gag* gene splice donor site to the second *plag1* exon (containing initiation ATG codon) in all samples where the gene was hit. *plag1* expression is barely detectable in embryonic and adult kidney, and in all analyzed tumors except for those where the retrovirus integrated into the neighborhood of the gene; in these tumors, *plag1* is heavily overexpressed. The structure of chimeric mRNAs and their overexpression are in agreement with the reported mechanisms of oncogenic *plag1* activity in human tumors, where the *plag1* gene is often translocated under the influence of a strong promoter, which results in a high expression of the gene (24, 25).

In case of *twist*, truncated proviruses are found within the promoter region upstream of the ATG initiation codon in 4% of the samples, and drives its massive overexpression as we have previously described (10). *twist* is moderately overexpressed in the majority of other tumor clones, which resembles the

Table 1. Candidate nephroblastoma-associated loci inferred from the analysis of VIS

Candidate locus	Candidate genes	Samples	Evidence
Nal 1-17		WCC939.1a, WCBO30.1d	1
Nal 1-19	*Hyal-2*	WCB789.1b, WCB818.2a	1
Nal 1-85	*POU2/OTF1/OCT1*	BL101.1a	2, 5
Nal 1-106		WCB819.2c, WCB030.2a	1
Nal 1-145	*spry-2*	BL344.2a, BL326.1c	1, 2, 4
Nal 1-187		WCB813.1a	5
Nal 2-28	*twist*	BL102.2b, BL107.1a, BL122.1a, BL395.1a, WCB030.1b (10)	1, 2, 3, 4
Nal 2-88	*dynein*	BL395.1b, WCC036.2a	2
Nal 2-104		WCB039.1d, WCC939.1a	1
Nal 2-110	*plag1*	BL304.1a, BL338.2a, BL384.2a, BL410.2c, WCB789.1c, WCB813.1a	1, 2, 4
Nal 2-118	*atbf1*	BL326.1c, BL819.1a	2
Nal 2-132	*nov*	BL435.1b, X59284 (13)	2, 3, 4
Nal 2-145a	*kiaa0196, flj32440*	WCC037.2b, WCC042.2a	1, 4
Nal 2-145b	*enc1, fbx032*	WCB030.2c, WCC042.1b, WCC939.1c	1
Nal 3-30	*LOC116228*	WCB030.1e, BL378.2a	1
Nal 5-13	*c-Ha-ras*	BL389.2a, BL821.1a, X03578 (21)	1, 2, 3
Nal 6-28		WCB822.2b, WCC850.1c	1
Nal 12-15	*foxP1*	WCC036.2a, WCC037.1a, WCC814.1d, WCC85O.1a, WCB826.1a	1, 2, 4

NOTE: Candidate loci were selected as described in Results, based on the following criteria: 1, multiple VISs within 20 kbp; 2, multiple hits in a defined large gene locus (VIS distance >20 kbp); 3, previously described VIS (accession nos. and the corresponding references are listed in the third column along with our samples); 4, experimentally confirmed deregulated expression of the candidate gene (overexpression or presence of transcripts with an altered coding sequence); 5, single clonal retroviral integration in the sample (single integration confirmed both by Southern blot and inverse PCR analyses). The criteria that led to the selection of candidate loci are summarized in the fourth column.

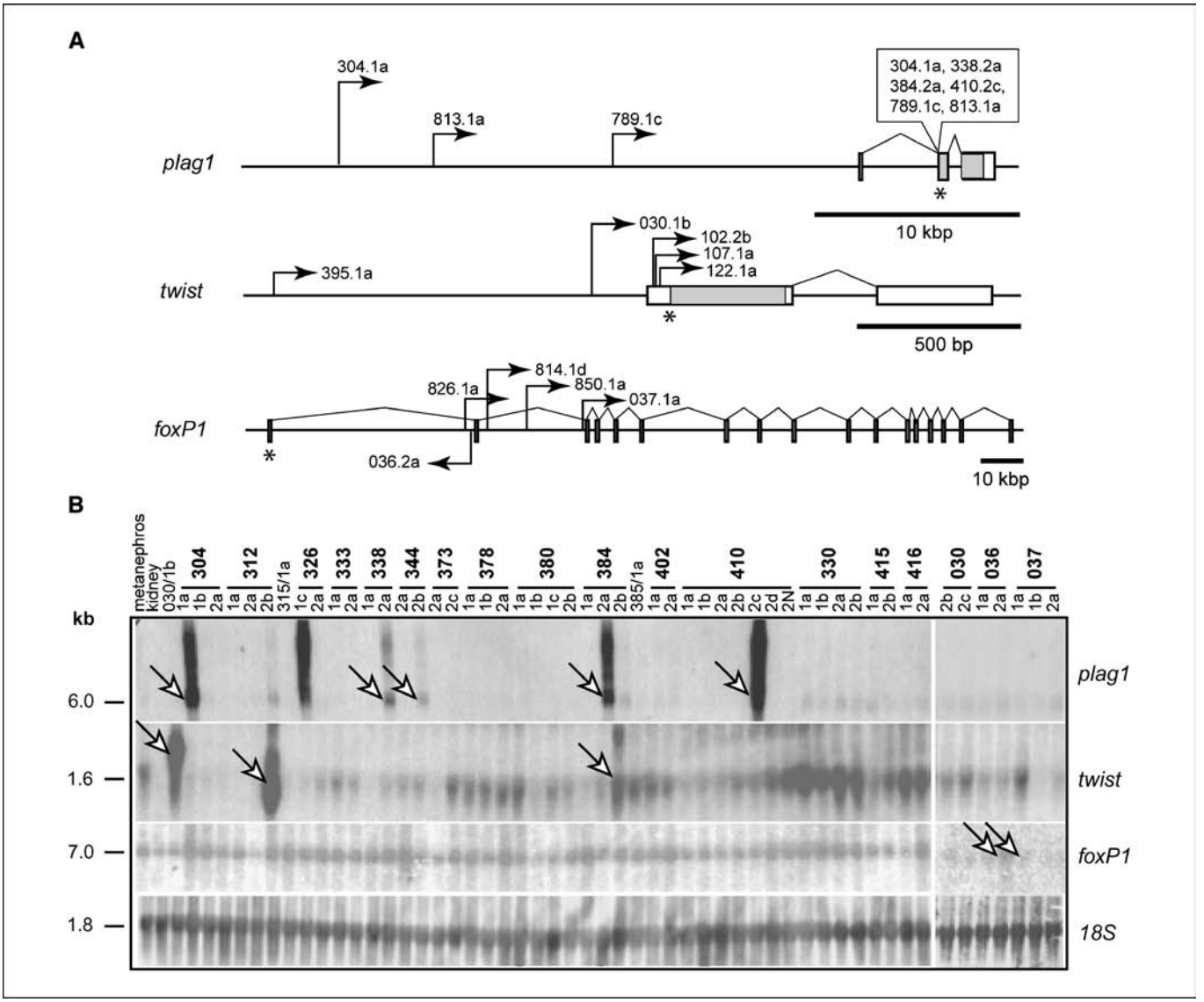


Figure 3. Provirus insertions into *plag1, twist*, and *foxP1* and expression of these genes in chicken nephroblastomas. *A,* positions of all VIS detected within *plag1, twist*, and *foxP1* gene loci are indicated by arrows above the sketched genes. *Arrows*, the direction of transcription driven by integrated proviruses (transcription of host genes proceeds from the left to the right). The numbers at arrows specify tumor clones in which the integration was found. *Thick bars* and *rectangles*, exons; *, positions of initiation ATGs. *Shaded* and *open rectangles*, coding and noncoding exons in *plag1* and *twist* genes, respectively. The precise genomic locations of MAV-2 proviruses in the *plag1* locus are known only for tumors 304.1a, 813.1a, and 789.1c. The structure of chimeric virus-*plag1* mRNAs is identical in all tumors listed in the box; the proviral *gag* donor sequence is spliced to the second *plag1* exon. *B,* a typical example of Northern blot analysis of *plag1, twist*, and *foxP1* expression. RNAs from 53 selected samples were separated by electrophoresis and blots were stained with methylene blue to reveal the amount and integrity of RNA in samples (18S) and subsequently hybridized with specific probes for chicken *plag1, twist*, and *foxP1* genes. *Open arrows*, samples harboring integration within each gene. Normal mRNA sizes are marked on the left (kb).

expression level of developing embryonic kidney. In adult kidney, this gene is not expressed. Thus, in nephroblastomas in which *twist* was not hit by the provirus, the expression probably reflects the embryonic character of the tumor tissue. Indeed, class III tumors invariantly display the most elevated *twist* expression and growth potential.

In case of *foxP1*, all the integration events are clustered around the second coding exon. Thus, in the resulting truncated proteins (predicted by conceptual translation of cloned cDNA) the putative NH_2-terminal domain is missing. It has been shown that NH_2-terminal domain deletion by alternative promoter usage modulates the activation/repression properties of the protein. A dimerization with different members of the FoxP protein family might also be affected by NH_2-terminal domain deletion (26). We propose that the retrovirally driven expression of the mentioned FoxP1 isoform can cause altered regulation of FoxP1 target genes. Such proteins would interfere with the normal function of the wild-type allele, contributing to oncogenic transformation. Surprisingly, no overexpression resulted from retroviral integration into *foxP1*. The *foxP1* mRNA level in tumors is almost uniform, with little variability whether or not the gene was hit by the provirus. It has been suggested in ref. (27) that the human wild-type *foxP1* allele has a tumor suppressor function, the virally altered *foxP1* might interfere (in a dominant-negative fashion) with a normal function of the gene and support malignant transformation.

***c-Ha-ras, nov*, and *sprouty2*: examples of rare targets for provirus integration.** *c-Ha-ras* and *nov* were earlier reported to be hit by MAV-2 integration in single cases of chicken nephroblastoma. The integration was accompanied by the elevated expression of the unaltered *c-Ha-ras* coding sequence (21) or by the overexpression of the truncated *nov* (13). In our experiments, however, the data are far less consistent, as shown below.

We observed neither overexpression nor mutation of the coding sequence of *c-Ha-ras* in either of the two tumors with retrovirally targeted *c-Ha-ras*. It might be significant, however, that all three integration events in *c-Ha-ras* (ref. 21, and this work) occurred within the region of the 5′-untranslated region in the orientation identical with the gene transcription. We suggest that the integrations disrupt a regulatory element that is not involved in transcription but rather in the control of mRNA translational availability. Alternatively, the observed integrations might cause deregulation of a distant locus with an oncogenic potential.

We have detected MAV-2 integration into *nov* in only one sample. The integration took place 125 codons upstream from the STOP codon, potentially allowing for synthesis of two mRNAs, starting either at the *nov* promoter or at the LTR of the provirus. Two faint aberrant *nov* mRNAs we found in the respective sample (Fig. 4*B*) might represent these two messages. The normal message was not detected, suggesting that the second *nov* allele was inactive in these tumors. The expression levels of unaltered *nov* mRNA in other samples in our collection fluctuated across several orders of magnitude, including samples with an undetectable level of *nov* mRNA (as in normal mature kidney). In contrast to our results, Joliot et al. have reported an ectopic high expression of *nov* in all 22 analyzed chicken nephroblastomas. We have no plausible explanation for these contradictory observations to date.

As a final example, the *sprouty2* locus was found to be hit by the MAV provirus in two clones. Both VIS are about 30 kb downstream of *sprouty2*, the closest predicted gene. They are in an opposite orientation separated by 4,550 bp from each other. The *sprouty2* mRNA level was undetectable in the sample with the provirus in the same orientation as the gene (326.1c). In contrast, in the sample with the provirus in the opposite orientation (344.2a), *sprouty2* mRNA was readily detectable (Fig. 4*C*). The expression of the gene was undetectable in the majority of other samples, but a few of them displayed a *sprouty2* transcript level comparable or even exceeding the level found in 344.2a. Because *sprouty2* can both

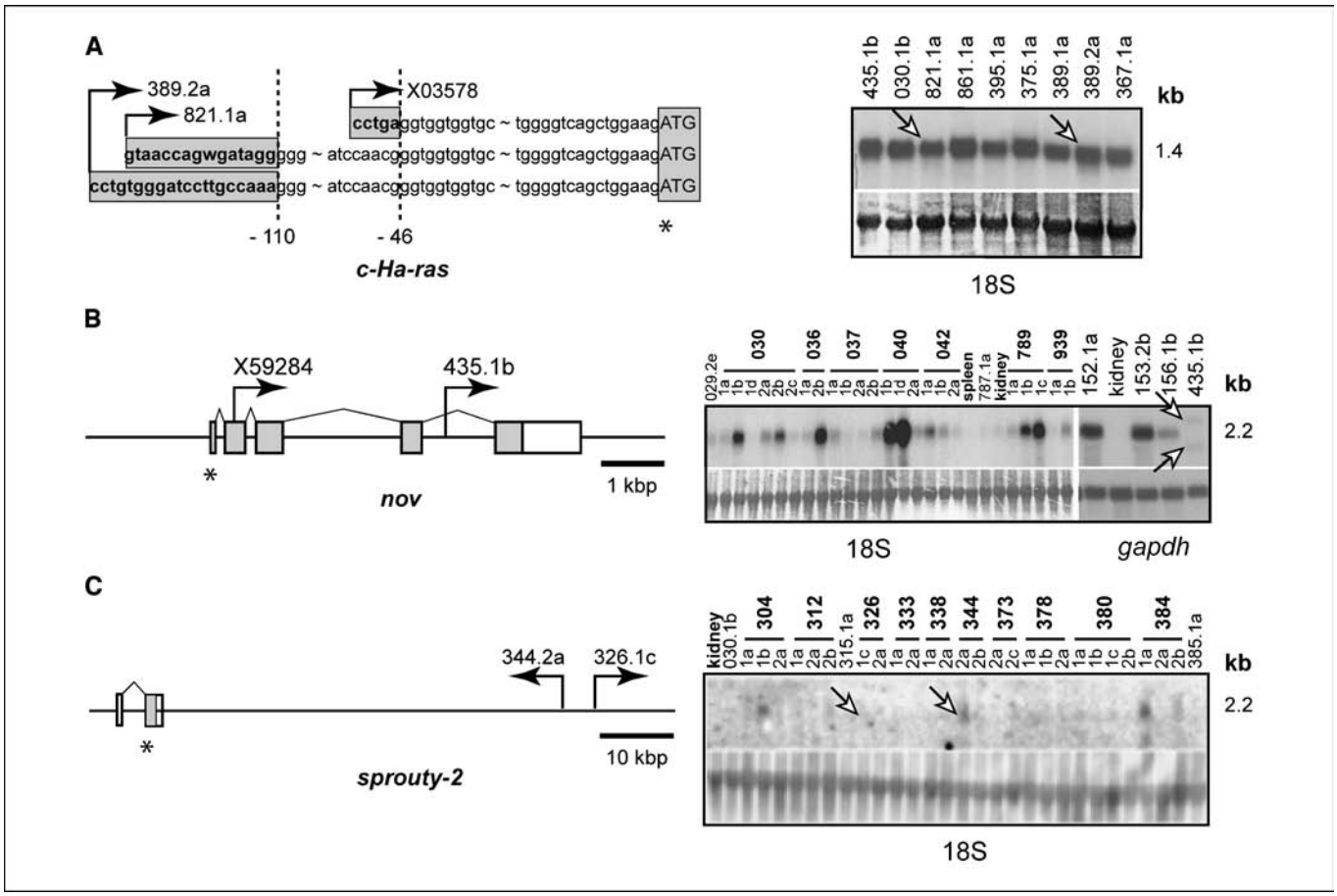


Figure 4. Analysis of cVIS in *c-Ha-ras, nov*, and *sprouty2* loci. *A,* precise structure of MAV-*c-Ha-ras* chimeric mRNAs. *Dashed lines*, the borders of *c-Ha-ras* exons; *shaded boxes*, host genomic sequences between the MAV right LTR (integration site) and the splice acceptor site; *arrows*, the direction of transcription driven by integrated proviruses; *, initiation ATGs. X03578 is the accession number of the previously described VIS. Northern blot analysis of *c-Ha-ras* expression in nine selected tumors including samples 389.2a and 821.1a (*right*). Methylene blue staining of 18S rRNA is shown. *B,* positions of VIS in the sketched *nov* locus. Northern blot analysis of *nov* expression in the tumors (including sample 435.1b) and control tissues (*right*). Methylene blue staining of 18S rRNA or glyceraldehyde-3-phosphate dehydrogenase hybridization is shown. X59284 is the accession number of the previously described VIS. *C,* positions of VIS in the sketched *sprouty2* locus. Northern blot analysis of *sprouty2* expression in the tumors (including samples 344.2a and 326.1c) and control tissues (*right*). Methylene blue staining of 18S rRNA is shown. Symbols and marks used in (*B*) and (*C*) are the same as in Fig. 3.

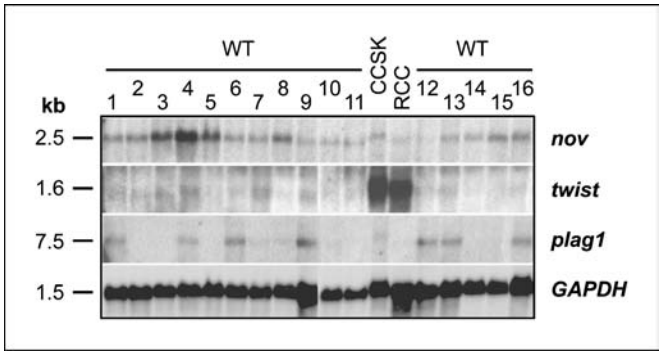


Figure 5. Northern blot analysis of *twist, plag1*, and *nov* in human renal tumors. RNAs from 16 Wilms tumors (WT), clear cell sarcoma of kidney (CCSK), and renal cell carcinomas were sequentially hybridized with the *nov, twist, plag1*, and *GAPDH* probes. mRNA sizes are indicated (kb).

inhibit and activate Ras-mediated signals (28, 29), it is a promising target for further investigation.

***Nov, twist*, and *plag1* are deregulated in some human renal tumors.** To assess the potential of the chicken nephroblastoma model for identification of genes involved in human malignancies, we analyzed human renal tumors for the expression of orthologues of chicken candidate genes *nov, twist*, and *plag1*, whose expressions are low to undetectable in normal kidney. We collected 18 samples of children's neoplasias including 16 Wilms tumors, one renal clear cell carcinoma and one clear cell sarcoma of the kidney. Wilms tumors (or human nephroblastomas) are the most frequent renal tumors of childhood in humans and are closely related to chicken nephroblastomas (30). In children, renal clear cell carcinoma and clear cell sarcoma of the kidney are highly malignant rare tumors of different etiology with a propensity for bone and brain metastases (31).

Figure 5 shows that the *twist* gene was strongly up-regulated in both non–Wilms tumors, whereas only a threshold expression was detected in Wilms tumors. *twist* was found deregulated in a number of human tumors and is assumed to be an important oncogenic contributor. It has been shown that the forced *twist* expression blocks apoptosis (32) and promotes epithelial-mesenchymal transition and tumor metastases (33). Moreover, *twist* is a negative regulator of osteogenesis and myogenesis (34). We suggest that *twist* overexpression in the two non–Wilms tumors might support their aggressive phenotype in several ways, including resistance to apoptotic stimuli, maintenance of the immature phenotype, and by providing molecular compatibility with the tissue targeted by the metastases (bone marrow). If further analyses confirm *twist* overexpression in these rare but aggressive renal tumor types, *twist* might become an important marker as well as a potential therapeutic target.

The elevated *plag1* mRNA level was detected in about half of Wilms tumors samples. *Plag1* is considered a dominant oncogene in a significant portion of pleomorphic adenomas (24) and lipoblastomas (25), and its elevated expression is associated with the development of acute myeloid leukemia (35) and hepatoblastoma (36). Our results could be the first evidence of possible involvement of *plag1* in the genesis of a subset of human renal tumors.

Finally, *nov* was expressed at diverse levels in all human tumors we analyzed, reminiscent of the pattern described above in the chicken nephroblastomas. Deregulation of *nov* in a number of human tumors has already been reported and its role in the genesis of Wilms tumors has been proposed (37).

Our experiments document the overall applicability of the chicken model for the search for genes participating in tumor formation and progression. This conclusion is supported by the fact that out of 18 candidate chromosomal loci identified in the chicken nephroblastoma screen, 5 (*twist, plag1, nov, c-Ha-ras*, and *foxP1*) clearly coincided with the human genes that have already been denoted as oncogenes/tumor suppressor genes and have been implicated in the formation of human solid tumors.

Discussion

The tumor model of chicken nephroblastoma is unique in many aspects. The original studies based on histologic and morphologic examinations led to the presumption that it is of similar origin as the Wilms tumor—the most frequent human renal tumor of childhood (30). Surprisingly, a detailed analysis presented here depicts a far more complex picture. The altered genes identified in the screen only marginally overlap with those known to participate in the genesis of Wilms tumors (i.e., *nov*), whereas no alterations of *wt1* were observed. Alterations of Wt1 tumor suppressor are the most common mutations detected in Wilms tumors. On the other hand, a number of genes previously shown to be mutated in human tumors of different origin and morphology (*plag1, twist, foxP1*, and *Ha-ras*), as well as others with yet unidentified functions, have been found. Our results are in good compliance with the observation that a similar oncogenic alteration of a gene could potentiate different tumors in different organisms (1). The *plag1* gene serves as a good example. As the most frequent cVIS, deregulated in 7% of chicken nephroblastomas, it is apparently an important inducer of chicken nephroblastoma. In humans, however, the up-regulation of *plag1*, as a result of chromosomal translocation, is known to be a prominent cause of pleomorphic adenomas and lipoblastomas (24, 25). Its participation is also presumed in the genesis of myeloid leukemias (35) and hepatoblastomas (36), but thus far, there has been no observation of its possible involvement in renal malignancies. A similar claim also fits for the other two frequent targets of MAV integration—*twist* and *foxP1*. Studying the differences in cancer induction between the species could have a fundamental effect on understanding the basics of cellular transformation and subsequent tumor formation.

The uniqueness of the chicken nephroblastoma model lies in the fact that it is currently the sole thoroughly investigated model using retroviral insertional mutagenesis for the identification of genes participating in solid tumor formation. It covers diverse aspects of cancer research—identification of responsible genes, mapping different mechanisms of their transforming abilities, and comparison of gene expression with the morphologic and histologic aspects of individual tumors. Last but not least, it gives the possibility to compare the basics of tumorigenesis in different animal classes. It is reasonable to assume that further analysis will reveal many novel candidate genes.

Acknowledgments

Received 5/25/2005; revised 8/18/2005; accepted 10/5/2005.

Grant support: AV0Z5052915-I045, A5052309 and in part by project no. AV0Z50520514 awarded by the Academy of Sciences of the Czech Republic.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Daniel Elleder for critical reading of the manuscript, helpful comments, and discussion. The assistance of Šárka Takáèová in the manuscript preparation is also greatly acknowledged.

References

1. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med 2004;10:789–99.

2. Futreal PA, Coin L, Marshall M, et al. A census of human cancer genes. Nat Rev Cancer 2004;4:177–83.

3. Coffin JM, Hughes SH, Varmus HE. Retroviruses. CSHL Press; 1997.

4. Suzuki T, Shen H, Akagi K, et al. New genes involved in cancer identified by retroviral tagging. Nat Genet 2002; 32:166–74.

5. Lund AH, Turner G, Trubetskoy A, et al. Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice. Nat Genet 2002;32:160–5.

6. Mikkers H, Allen J, Knipscheer P, et al. High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. Nat Genet 2002;32:153–9.

7. Watts SL, Smith RE. Pathology of chickens infected with avian nephroblastoma virus MAV-2(N). Infect Immun 1980;27:501–12.

8. Pecenka V, Dvorak M, Karafiat V, et al. Avian nephroblastomas induced by a retrovirus (MAV-2) lacking oncogene. II. Search for common sites of proviral integration in tumour DNA. Folia Biol (Praha) 1988;34:147–69.

9. Hillier LW, Miller W, Birney E, et al. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 2004;432:695–716.

10. Pajer P, Pecenka V, Karafiat V, Kralova J, Horejsi Z, Dvorak M. The twist gene is a common target of retroviral integration and transcriptional deregulation in experimental nephroblastoma. Oncogene 2003;22: 665–73.

11. Plachy J, Hala K. Comparative aspects of the chicken immunogenetics [review]. Folia Biol (Praha) 1997;43: 133–51.

12. Ausubel FM, Brent R, Kingston RE, et al. Current protocols in molecular biology. Green Publishing Associates, Inc.; 1993.

13. Joliot V, Martinerie C, Dambrine G, et al. Proviral rearrangements and overexpression of a new cellular gene (nov) in myeloblastosis-associated virus type 1-induced nephroblastomas. Mol Cell Biol 1992;12:10–21.

14. Boardman PE, Sanz-Ezquerro J, Overton IM, et al. A comprehensive collection of chicken cDNAs. Curr Biol 2002;12:1965–9.

15. Pecenka V, Dvorak M, Karafiat V, Travnicek M. Avian nephroblastomas induced by a retrovirus (MAV-2) lacking oncogene. III. Presence of defective MAV-2 proviruses in tumour DNA. Folia Biol (Praha) 1988;34: 215–32.

16. Robinson HL, Gagnon GC. Patterns of proviral insertion and deletion in avian leukosis virus-induced lymphomas. J Virol 1986;57:28–36.

17. Silver J, Keerikatte V. Novel use of polymerase chain reaction to amplify cellular DNA adjacent to an integrated provirus. J Virol 1989;63:1924–8.

18. Devon RS, Porteous DJ, Brookes AJ. Splinkerettes—improved vectorettes for greater efficiency in PCR walking. Nucleic Acids Res 1995;23:1644–5.

19. Valk PJ, Joosten M, Vankan Y, Lowenberg B, Delwel R. A rapid RT-PCR based method to isolate complementary DNA fragments flanking retrovirus integration sites. Nucleic Acids Res 1997;25:4419–21.

20. Fitzgerald ML, Vora AC, Zeh WG, Grandgenett DP. Concerted integration of viral DNA termini by purified avian myeloblastosis virus integrase. J Virol 1992;66: 6257–63.

21. Westaway D, Papkoff J, Moscovici C, Varmus HE. Identification of a provirally activated c-Ha-ras oncogene in an avian nephroblastoma via a novel procedure: cDNA cloning of a chimaeric viral-host transcript. EMBO J 1986;5:301–9.

22. Wu X, Li Y, Crise B, Burgess SM. Transcription start regions in the human genome are favored targets for MLV integration. Science 2003;300:1749–51.

23. Elleder D, Pavlicek A, Paces J, Hejnar J. Preferential integration of human immunodeficiency virus type 1 into genes, cytogenetic R bands and GC-rich DNA regions: insight from the human genome sequence. FEBS Lett 2002;517:285–6.

24. Kas K, Voz ML, Roijer E, et al. Promoter swapping between the genes for a novel zinc finger protein and β-catenin in pleiomorphic adenomas with t(3;8)(p21;q12) translocations. Nat Genet 1997;15:170–4.

25. Hibbard MK, Kozakewich HP, Dal Cin P, et al. PLAG1 fusion oncogenes in lipoblastoma. Cancer Res 2000;60: 4869–72.

26. Wang B, Lin D, Li C, Tucker P. Multiple domains define the expression and regulatory properties of Foxp1 forkhead transcriptional repressors. J Biol Chem 2003; 278:24259–68.

27. Banham AH, Beasley N, Campo E, et al. The FOXP1 winged helix transcription factor is a novel candidate tumor suppressor gene on chromosome 3p. Cancer Res 2001;61:8820–9.

28. Yusoff P, Lao DH, Ong SH, et al. Sprouty2 inhibits the Ras/MAP kinase pathway by inhibiting the activation of Raf. J Biol Chem 2002;277:3195–201.

29. Wong ES, Fong CW, Lim J, et al. Sprouty2 attenuates epidermal growth factor receptor ubiquitylation and endocytosis, and consequently enhances Ras/ERK signalling. EMBO J 2002;21:4796–808.

30. Heine U, De The G, Ishiguro H, Sommer JR, Beard D, Beard JW. Multiplicity of cell response to the BAI strain A (myeloblastosis) avian tumor virus. II. Nephroblastoma (Wilms' tumor): ultrastructure. J Natl Cancer Inst 1962;29:41–105.

31. Kufe D, Pollock R, Weichselbaum R, et al. Cancer medicine 6. Part VII. BC Decker; 2003.

32. Maestro R, Dei Tos AP, Hamamori Y, et al. Twist is a potential oncogene that inhibits apoptosis. Genes Dev 1999;13:2207–17.

33. Yang J, Mani SA, Donaher JL, et al. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell 2004;117:927–39.

34. Hamamori Y, Wu HY, Sartorelli V, Kedes L. The basic domain of myogenic basic helix-loop-helix (bHLH) proteins is the novel target for direct inhibition by another bHLH protein, Twist. Mol Cell Biol 1997;17: 6563–73.

35. Landrette SF, Kuo YH, Hensen K, et al. Plag1 and Plagl2 are oncogenes that induce acute myeloid leukemia in cooperation with Cbfb-MYH11. Blood 2005;105:2900–7.

36. Zatkova A, Rouillard JM, Hartmann W, et al. Amplification and overexpression of the IGF2 regulator PLAG1 in hepatoblastoma. Genes Chromosomes Cancer 2004;39:126–37.

37. Martinerie C, Huff V, Joubert I, et al. Structural analysis of the human nov proto-oncogene and expression in Wilms tumor. Oncogene 1994;9: 2729–32.

Industasis, a Promotion of Tumor Formation by Nontumorigenic Stray Cells

Petr Pajer,[1] Vít Karafiát,[1] Vladimír Pečenka,[1] Dana Průková,[1] Jana Dudlová,[1] Jiří Plachý,[1] Petra Kašparová,[2] and Michal Dvořák[1]

[1]Institute of Molecular Genetics AS CR, Prague, Czech Republic and [2]Fingerland's Department of Pathology, School of Medicine Charles University, Hradec Králové, Czech Republic

Abstract

A tumor cell is formed when a critical amount of endogenous and/or exogenous tumorigenic stimuli is exceeded. We have shown that the transient presence of nontumorigenic stray cells in tissues of experimental animals that contain cells with a subcritical set of genetic mutations can act as a tumor-promoting stimulus. To induce somatic mutations in all chicken tissues, we have used the MAV-2 retroviral insertion system that almost exclusively generates nephroblastomas. MAV-2 mutagenized animals i.v. inoculated with nonmalignant cells developed early clonal lung tumors before nephroblastomas. Importantly, the injected cells did not become a component of resultant tumors. Lung tumors displayed specific mutational signature characterized by an insertion of MAV-2 provirus into the *fyn-related kinase* (*frk*) promoter that results in the overexpression of the *frk* gene. In contrast, *plag1, foxP*, and *twist* genes were most often mutagenized in nephroblastomas. Based on such observations, we propose the mechanism termed industasis, a promotion of fully malignant phenotype of incipient tumor cell by stray cells, and hypothesize that it might be the underlying cause of human multiple primary tumors. [Cancer Res 2009;69(11):4605–12]

Introduction

Spontaneous tumors arise when multiple genetic and epigenetic changes amass in a single cell. Nongenetic factors, such as stromal microenvironment interactions (1) and the host immune response (2), have also been shown to play a role in oncogenesis. In addition, evidence has been published showing that tumor cells can remain dormant until a tumor-promoting stimulus triggers their uncontrolled proliferation (3).

Tumor cells of both metastatic and nonmetastasizing cancers are long known to circulate in the blood of patients (4, 5). Released potentially tumorigenic cells are able to persist in a second organ for an extended period of time (6, 7). In addition, normal cells can be liberated into the bloodstream as a result of an injury or surgical intervention (8). The seemingly nondeleterious ectopic presence of such stray (primary tumor or nontransformed) cells within secondary organs has been reported to induce changes within the affected local microenvironment. A fundamental link between the stromal microenvironment and behavior of transformed cells in terms of tumor development has recently been highlighted (9).

Over the last few decades, the incidence of multiple primary tumors within a single host has rapidly increased. The phenomenon is expected to become an even more serious threat in the future as a result of prolonged life span and, paradoxically, of improved healthcare. For example, individuals cured from a tumor exhibit an increased chance of developing second primary tumors in addition to the risk of metastases or a relapse of the first cured tumor (10). Two explanations have been put forward thus far: presence of an inherited genetic predisposition to tumor formation (11) and mutagenic effect of therapy for the first malignancy (12).

Retroviruses represent a potent tool for identifying cancer-related genes. Nonacute oncogenic retroviruses, such as the avian virus MAV-2, do not carry an oncogene; instead, they induce transformation through insertional mutagenesis when proviruses integrate into the host gene loci. Due to its high infectivity, the MAV-2 retroviral system ensures that essentially each host gene locus is affected through random integration in many cells of the target tissue. When a combination of mutations in a cell perturbs cellular functions critical for malignant transformation, the cell clone expands and forms a tumor. MAV-2 predominantly generates nephroblastomas after a 2- to 3-month latency period. Mutated gene loci are easily detectable as they are tagged by the proviral sequences. The genes *plag1, foxP1*, and *twist* have repeatedly been found hit by MAV-2 in nephroblastomas, thereby underlining their importance for malignant transformation of nephrogenic blastema (13, 14).

In this work, we have identified insertionally mutagenized and overexpressed *fyn-related kinase* (*frk*) gene as the salient feature of MAV-2–induced lung sarcomas. The *frk* gene encodes a nonreceptor tyrosine kinase synthesized predominantly in epithelial tissues (15, 16), which has been implicated in chondrogenesis and in the development of the islets of Langerhans (17). We have observed the substantial decrease of latency and increase in frequency of lung sarcomas when MAV-2–infected animals have been i.v. inoculated with nonmalignant cells. Activated *frk* represented the mutational signature in both late tumors induced by MAV-2 alone as well as in early tumors promoted by stray nontumorigenic cells.

Materials and Methods

Experimental Animals, Cells, and Viruses

Viruses. The MAV-2 was the MAV-2(N)–type virus isolated from the AMV-BAI-A complex stock by plaque purification as described (18).

Cells. Chick embryo fibroblasts (CEF) were prepared from CB or CC embryos, cultivated, and infected by MAV-2 virus stock as described (19).

A210 cells were prepared from kidney of 19-d-old CB White Leghorn embryo, infected at the 12th day of incubation by MAV-2 virus, and dispersed in trypsin-EDTA solution (Sigma) in PBS. The cells from each kidney were plated at concentrations of 5×10^6 to 7×10^6/100-mm Petri dish and cultivated as CEF. After 2 wk, the cells were plated in P60 tissue

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

P. Pajer and V. Karafiát contributed equally to this work.

Requests for reprints: Michal Dvořák, Institute of Molecular Genetics AS CR, v.v.i., Vídeňská 1083, 142 20 Prague, Czech Republic. Phone: 420-296443390; Fax: 420-296443586; E-mail: mdvorak@img.cas.cz.

doi:10.1158/0008-5472.CAN-08-4636

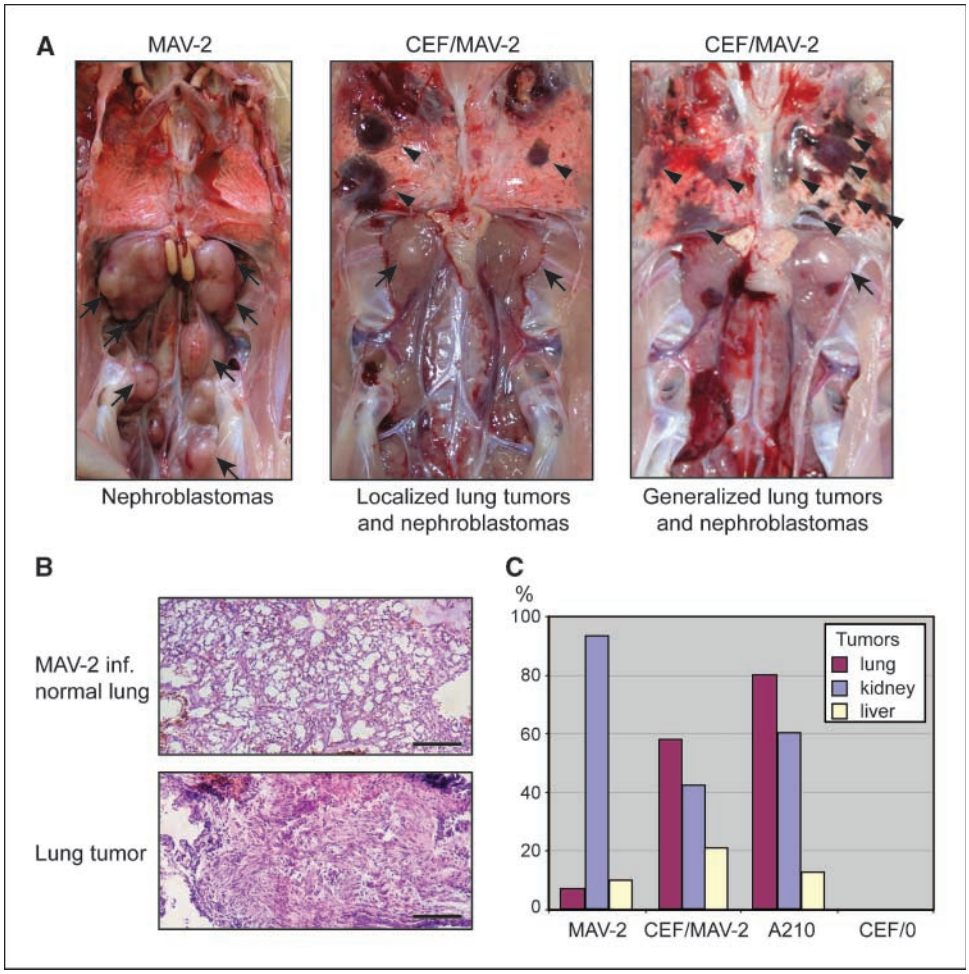


Figure 1. Injection of cells induces lung tumors. *A,* autopsy of animals injected with MAV-2 virus or CEF/MAV-2 cells. *Left, arrows,* MAV-2 alone induces nephroblastomas. Injected cells cause formation of lung tumors (*arrowheads*), either as distinct foci (*middle*) or generalized tumors (*right*). Nephroblastomas (*arrows*) also develop. *B,* histologic sections of nonmalignant (MAV-2 infected) lungs and of lung sarcoma (A210 induced). H&E staining. *Bar,* 100 μm. *C,* frequency of tumor induction in lungs, kidney, and liver by injection of MAV-2 virus, CEF/MAV-2, A210, and control CEF/0 cells. The reduction of nephroblastoma incidence in animals inoculated with the virus-producing cells is caused by earlier and frequently lethal onset of lung tumors. Groups of at least 20 animals were evaluated.

culture dishes at the density of 0.5×10^3 to 1×10^3 per one dish, and after another 2 wk, the largest foci of cells were isolated and further cultivated. Samples of cell suspension were counted in a cell counter CASY Model TTC (Schärfe Systems GmbH). The best proliferating culture, A210, turned out to be a cell clone. Cells for i.v. injections were prepared as follows: CEF, CEF/MAV-2, and A210 were grown to semiconfluency. If required, cells were treated with mitomycin C as described (19).

Animals. Chicks of inbred congenic CB and CC White Leghorns (20) were used. All procedures were performed in accordance with the Guide for the Care and Use of Laboratory Animals and approved by the Animal Care and Use Committee of the Academy of Sciences of the Czech Republic. Chicks were kept under standard laboratory conditions with free access to food and water. For cell inoculation, 1×10^5 to 2×10^5 cells detached by accutase were injected into chorioallantoic vein of 12-d-old embryo or into metatarsal vein of 1-d-old chick. MAV-2 infection was performed as described (14). Control animals were mock infected by an identical volume of PBS.

Sample Collection

The animals were sacrificed at the age between 20 and 120 d after hatching and tumor samples and control tissues were collected. The samples larger than ~4 mm in diameter were divided to thirds and processed immediately into DNA, RNA, and paraffin samples. Either genomic DNA or total RNA was isolated from the smallest tumor foci. DNA and RNA were isolated and quantitated by standard methods (14).

Southern Blot Analyses

Southern blot analyses were performed as described previously (14).

Histologic Investigations

Paraffin-embedded samples, microscopic preparations, and histologic procedures were made as described (21).

PCR, Reverse Transcription-PCR, and Quantitative PCR

cDNA was synthesized and long terminal repeat-rapid amplification of cDNA ends (LTR-RACE) was performed exactly as described previously (14). Real-time PCR was performed using DyNAmo HS SYBR Green qPCR kit (Finnzymes) on Chromo4 cycler (MJ Research/Bio-Rad) and analyzed using the included software. Semiquantitative PCR and integration site-specific PCR were performed using GoTaq polymerase system from Promega according to the manufacturer's instructions. Primers were present at 200 nmol/L, deoxynucleotide triphosphates at 0.2 mmol/L each, and Taq polymerase at 1 unit/50 μL. The standard cycling protocol was 25 cycles (95°C for 15 s, 60°C for 30 s, and 72°C for 30 s). The PCR products were resolved in agarose gels; and in cases when no product has been detected, five PCR cycles were added. Sequences of primers used in this study are in Supplementary Data.

DNA Sequencing and Homology Searches

The sequencing reactions were performed using the BigDye Terminator Cycle Sequencing kit (PE Biosystems) and resolved on ABI PRISM 310 Sequencer. Sequence homology searches were conducted using the BLAT algorithm on the chicken genome assembly database.[3]

Results

The injection of virus-producing cells instead of virus changes the spectrum of tumors. When MAV-2 retroviruses are injected into either embryonic or newborn chicks, nephroblastomas (as a

[3] Built 2, ENSEMBL project (http://www.ensembl.org/).

consequence of insertional mutagenesis) are almost exclusively induced (22) despite all tissues being similarly infected (see below). In ~5% of animals afflicted with nephroblastomas, late tumors of the lungs and liver can additionally be diagnosed. Analysis of these clonal tumors consistently revealed different proviral integration patterns when compared with nephroblastomas and as such were considered as second primary tumors. We investigated the possibility of changing the spectrum of MAV-2–induced neoplasms by injecting MAV-2–producing cells instead of virions. We used animals and cells of inbred congenic White Leghorn lines CB and CC, which differ only in the MHC(B) haplotypes (23). Differences in MHC locus could be used as molecular markers for unambiguous identification of injected cells. In preliminary experiments, a group of 20 CC 12-day-old embryos was injected with A210 cells. These are clonal virus-producing mesenchymal-type cells derived from a MAV-2–infected CB embryonic kidney. The use of A210 cells is favorable as they harbor six mapped MAV-2 provirus integration sites that can additionally aid in the cell identification following injection (Supplementary Fig. S1). Given that clonal A210 cells could potentially produce a mutant of MAV-2, a second group of embryos was injected with polyclonal unselected MAV-2–producing CB CEF cells (CEF/MAV-2). A third and fourth control group received MAV-2 virus collected from either A210 or CEF/MAV-2 cells, respectively. After injection, the birds were sacrificed and analyzed 20 to 90 days after hatching.

As expected, the majority of chicks injected directly with MAV-2 virions (third and fourth control groups) developed numerous focal nephroblastomas that were evident within 2 to 3 months. Analysis of older animals revealed additional rare tumor foci that developed on the lungs and liver. Surprisingly, chicks injected with A210 or CEF/MAV-2 cells were found to develop, in addition to nephroblastomas, early lung sarcomas in most animals. As early as 27 days after hatching, macroscopic tumor foci were apparent on the lungs of chicks and in ~30% of the aforementioned animal tumors accounted for the majority of lung tissue (Fig. 1*A*; Supplementary Fig. S2). The same frequency of lung tumor formation was recorded when CB CEF/MAV-2 cells were injected into CB animals, suggesting that the MHC haplotype plays no role. The histology of lung sarcomas (Fig. 1*B*) was independent of whether A210 or CEF/MAV-2 cells were injected.

No tumors were induced in animals injected with non–virus-producing cells alone. The combined results of several independent experiments are summarized in Fig. 1*C*. In addition to the lung tumors, we also observed, with less frequency, tumors of the liver and ovary.

The resultant lung tumors are clonal, host derived, and frequently invasive. To determine whether the resultant lung tumors originated from the injected A210 or CEF/MAV-2 cells, Southern blot analysis of tumor cell DNA was performed to detect the proviral integration sites. The Southern blots revealed that lung tumors as well as nephroblastomas and less common liver and ovary tumors were formed by single-cell clones. Various tumor types of a single animal represented unrelated clones, suggesting that these were independently originated tumors, not metastases

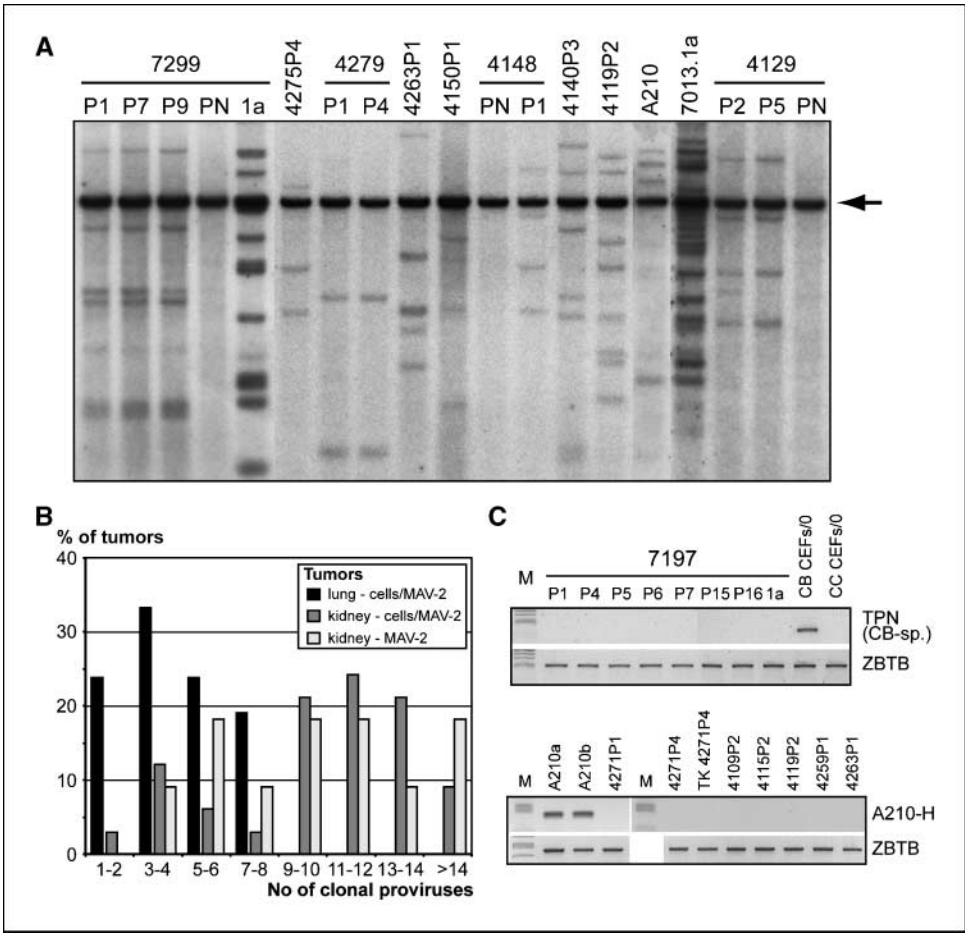


Figure 2. Lung tumors are clonal, host derived, and frequently invasive. *A,* the representative Southern blot analysis of 13 lung tumor foci (*P1–P5, P7*, and *P9*), 3 nonmalignant lung tissue (*PN*), 2 nephroblastomas (*1a*), and A210 cells. The identical integration pattern in various lung foci (*P1, P7*, and *P9* from the animal 7299 and *P2* and *P5* from the animal 4129) documents a frequent invasivity of the primary tumor clone. Normal lung tissue (*PN*) displays nonclonal pattern; nephroblastomas (7299.1a and 7013.1a) are different clones. *Arrow,* internal fragment of MAV-2 common to all integrated proviruses. *B,* distribution of clonally integrated MAV-2 proviruses (determined by Southern blot hybridization with the MAV-2–specific probe) in 20 lung tumors and in 33 and 11 nephroblastomas collected from animals injected with either CEF/MAV-2 or MAV-2, respectively. *C, top,* the representative result of PCR detection of CB cell–specific sequences (from *tpn* locus) in lung tumors (*P*) and nephroblastoma (*1a*) isolated from CC chick 7197 and in control fibroblasts (*CEF*) from both strains. The primer pair designed to check the amount and quality of DNA amplifies the fragment of the *zbtb* locus identical in both CB and CC strains. *Bottom,* representative result of detection of the A210-H integration site in A210 cells (*A210a*, 5th passage of stabilized cell culture; *A210b*, 20th passage) and in lung tumors from experimental animals (4271, 4109, 4115, 4119, 4259, and 4263). Cells from the tumor focus 4271P4 were grown in tissue culture for several passages (*TK 4271P4*) before analysis. *zbtb* control as in the top. The principle of the used PCR method is schematically depicted in Supplementary Fig. S3*A*.

from a single primary tumor. Intriguingly, the lung tumor cells displayed a high invasive potential that frequently resulted in the dissemination of one clone into several foci in both lungs. In contrast, nephroblastomas or tumors of other tissues were non-invasive and adjacent foci were always formed from independent clones (Fig. 2*A*; ref. 14).

Close examination of various tumors isolated from animals injected with either A210 or CEF/MAV-2 cells revealed that none of the tumors was derived from the injected cells. This was first confirmed by PCR analysis of lung tumors from CC chicks injected with CB cells (A210 or CEF/MAV-2). Attempts to amplify genomic DNA isolated from lung tumors with a CB-specific primer pair consistently yielded negative results (Fig. 2*C, top*). Additional PCR analysis of tumor DNA with primer combinations that should identify the specific proviral integration sites unique to A210 cells was also negative as illustrated in Fig. 2*C* (*bottom*).

Further experiments were designed to rule out the possibility that tumors arose from a fusion of an A210 and host cell (when only a limited amount of tumorigenesis-related genetic material of A210 cells might remain in the resulting tumor cell). Southern blot analysis of DNAs isolated from A210-induced lung tumors of 20 different animals revealed that patterns of proviral integration were different from those in A210 cells and different in each animal. Notably, not even a single band on the blots (representing particular provirus insertion site) was shared between A210 and the tumors as well as between tumors (Fig. 2*A*).

To ensure that A210 cells do not contain a minute population of cells from which lung tumors could derive, we identified the genomic positions of viral integration sites in two lung tumors and designed PCR primers for their detection. We were unable to detect any shared proviral integration sites neither between the two tumors nor between tumors and A210 cells (Supplementary Fig. S3; Supplementary Table S1; data not shown). Taken together, the data conclusively show that the first malignant cell that eventually gives rise to a tumor originates from the host and exclude the possibility that the tumor cell arises from the fusion of an injected and host cell.

The tumor-promoting effect of injected cells is transient and independent of their abilities to divide and to infect surrounding cells. Figure 3*B* summarizes incidence of tumors in

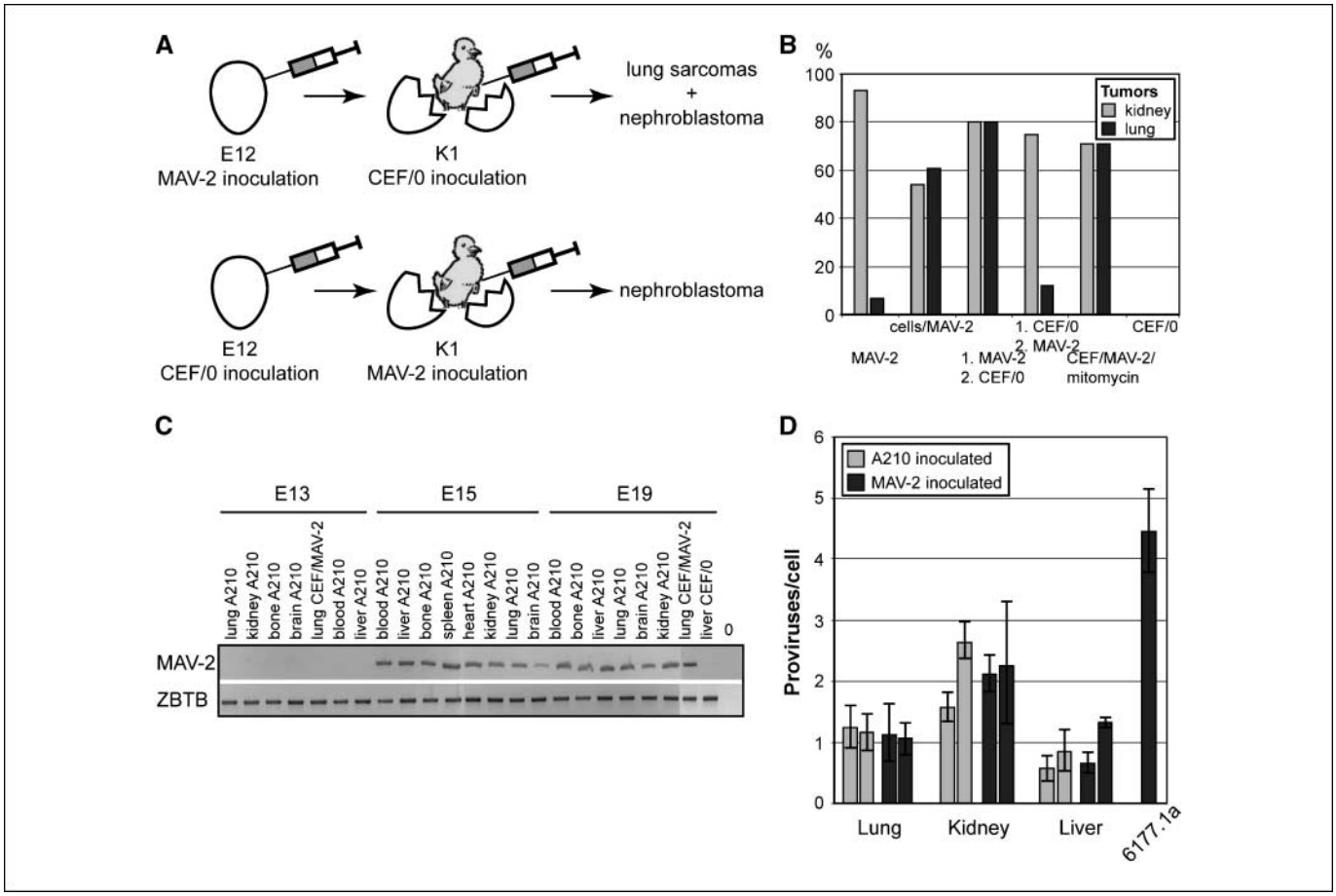


Figure 3. Contribution of the MAV-2 virus and cells to lung tumor induction and the extent of proviral integrations in different tissues. *A,* a schematic representation of successive injections of experimental chicks with virus (*MAV-2*) and cells (*CEF/0*) at E12 and at the first day after hatching (*K1*), and resulting tumors. *B,* incidence of tumors in chicks (groups of at least 10 animals) injected by MAV-2 virus at E12 (*MAV-2 columns*), A210 or CEF/MAV-2 cells at E12 (*cells/MAV-2 columns*), MAV-2 at E12 and CEF/0 cells at K1 (*1. MAV-2 2. CEF/0 columns*), CEF/0 at E12 and MAV-2 at K1 (*1. CEF/0 2. MAV-2 columns*), CEF/MAV-2 cells treated with mitomycin C at E12 (*CEF/MAV-2/mitomycin columns*), and noninfected CEF at E12 (*CEF/0 columns*). *C,* MAV-2 integration into various organs of chick embryos injected with MAV-2–producing cells. The virus-producing cells were injected at E12 into several embryos and DNAs from indicated tissues were analyzed by semiquantitative PCR using MAV-2–specific primers. Primers specific for the chicken *zbtb* locus (two copies per cell) were used as a control. *D,* quantification of MAV-2 proviruses integrated in nonmalignant tissues of 30-d-old experimental animals and in nephroblastoma 6177.1a. The marked tissues were isolated from two MAV-2–infected and two A210-infected chicks and subjected to quantitative real-time PCR with MAV-2–specific primers.

chicks injected by MAV-2 virus or A210, CEF/MAV-2, or CEF/0 cells. To distinguish whether the dividing capability of injected cells is necessary for tumor formation, we injected MAV-2–infected CEFs that had been treated with mitomycin, a compound that permanently blocks cell division. After treatment, the cells were still able to promote tumor formation (Fig. 3*B*), indicating that the proliferative potential of an injected cell is dispensable. The result also further supports the notion of host origin of lung tumors.

To understand more precisely the role of stray cells and virus in lung tumor formation, a series of experiments were carried out in which virus-free CEFs and MAV-2 virus were injected stepwise in varying orders. The first agent was applied at embryonic day 12 (E12), whereas the second was injected after hatching (8–9 days after first injection, K1). The animals were sacrificed 40 to 60 days later and their lungs were examined. Interestingly, lung tumors were only evident in animals that received virus as the first treatment (Fig. 3*A* and *B*). This observation suggests that stray cells can promote tumor formation in tissues that have formerly been mutagenized and that the tumor-promoting potential of the stray cells persists not longer than several days in our model. The tumor-promoting capacity of virus-free cells injected into animals that have already been systemically infected also confirms that the effect of injected cells is independent of the dissemination of virus infection by injected cells.

Injected cells have no effect on the spectrum of infected tissues or the level of infection. To examine whether the injected virus-producing cells might affect the spectrum of infected tissues, we measured steady-state levels of proviral sequences in the genomic DNA of several organs. This analysis revealed that all tissues, including the lungs, are fully infected 3 days after the cell injection (Fig. 3*C*) similarly as if animals were injected with MAV-2 virus (data not shown). The average number of proviruses in 2-month-old chicken tissues was the same independent of whether the animals were injected by virions or by virus-producing cells (Fig. 3*D*). To further investigate the possibility that the increased incidence of lung tumors resulted from a high local level of infection (and thus a high mutation load) caused by the residing virus-producing stray cells, we compared the number of integrated proviruses in lung tumors and nephroblastomas. The Southern blot analysis of 20 randomly selected lung tumor clones and 33 nephroblastomas from the same animals showed a significantly lower average number of integrations in lung tumors compared with nephroblastomas, contradicting the idea of high local level of infection in the lung inflicted by the A210 or CEF/MAV-2 cells. Importantly, the average number of integrations in nephroblastomas was independent of whether the animals were injected by virions or by virus-producing cells (Fig. 2*B*).

The injected cells preferentially reside in liver, kidney, and lung embryonic tissues. To elucidate the precise localization of injected cells in tissues, MAV-2–producing CB CEFs were metabolically labeled with [^{35}S]methionine before i.v. injection into embryonic chicks. After 1 or 3 days after injection, several embryonic tissues were collected and both protein lysates and paraffin sections were prepared. The radioactivity in lysates was measured and specific radioactivity was calculated. Radioactive label was found in all tissues analyzed, with markedly the highest levels in the liver, kidney, and lungs (Fig. 4*A*). Very low radioactivity was found in the blood already 20 hours after injection, indicating prompt passage of the cells into the target tissues. Examination of the paraffin-embedded sections by autoradiography revealed similar observations (Fig. 4*B*; Supplementary Fig. S4). These results correlate well with the localization of the eventual tumors and show that the injected cells survive for at least 3 days in target tissues, sufficient time to influence the local microenvironment.

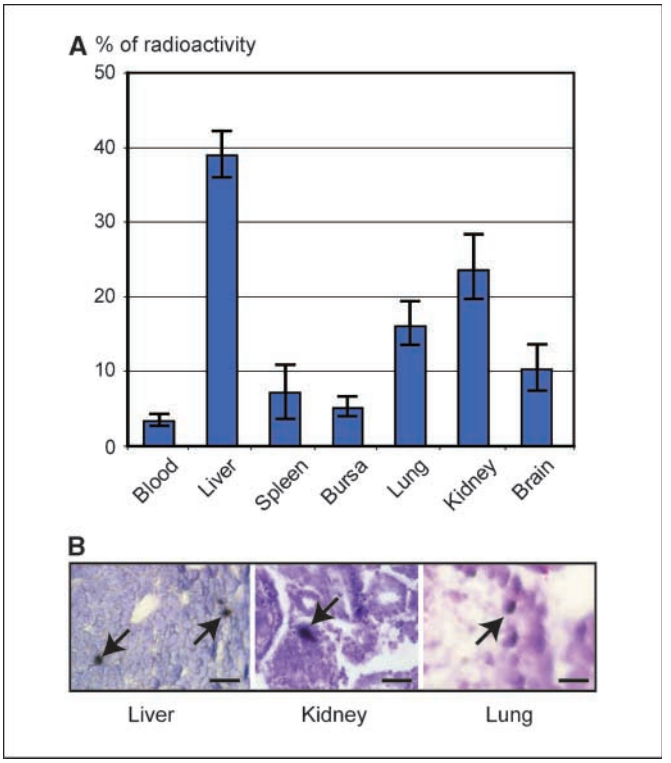


Figure 4. Homing and distribution of *in ovo*–injected MAV-2–producing cells in embryonic tissues. *A,* the distribution of injected [^{35}S]methionine-labeled cells was determined by calculating specific radioactivity in various embryonic tissues 20 h after injection. *B,* examples of radiolabeled cells (*arrows*) homed in liver, kidney, and lungs. For details, see Supplementary Fig. S4. *Bars,* 20 μm.

The recurrent target of MAV-2 insertional mutagenesis in lung tumors is the *frk* gene. The above observations are consistent with the hypothesis that a stray circulating cell can emit a tumor-promoting signal in a tissue with "silent tumor cells" that contain cancerous mutations, such as oncogene activation or tumor suppressor gene inactivation. If there are cancerous mutations typical for MAV-2–induced lung tumors, then common sites of MAV-2 integration should be found in tumor DNAs.

To identify loci hit by the provirus, we analyzed 17 chicks with cell promoted and 2 chicks with late MAV-2 lung tumors using the LTR-RACE method (14, 24). The study revealed recurrent integrations within the promoter/5′-untranslated region of the *frk* gene in 15 animals injected with cells and both animals injected only with MAV-2 (Supplementary Table S1). Proviruses were found to integrate in the same transcriptional orientation as the *frk* mRNA sequence within the narrow region of 1077 to 14 nucleotides upstream of the FRK initiation codon (Fig. 5*A*; Supplementary Table S2). Additional analysis by PCR, reverse transcription-PCR (RT-PCR), and Southern blotting confirmed the aforementioned results. As a consequence of proviral integration, the hybrid mRNA (starting in the proviral 3′-LTR and proceeding into the *frk* coding sequence) was highly expressed in all positive samples (Fig. 5*B* and *C*). The hybrid transcripts contained protein coding sequence identical with the proposed endogenous *frk* transcript. No *frk* expression was detectable neither in nontumor lung tissue nor in the two tumors with unaffected *frk* locus (Fig. 5*B* and *C*).

Two conclusions can be drawn from the presence of an activated *frk* gene in ~90% of the MAV-2–induced lung tumors, both cell promoted and MAV-2-only induced. First, it indicates that this dominant mutation is critically important for the tumor formation. Second, it suggests that injected cells promote tumor formation from the same pool of dormant mutagenized cells that also give rise to the rare long latency lung tumors in animals injected with MAV-2 only.

Discussion

Common presence of cancer-related mutations in cells of normal tissue has been well documented. The mutant cells are subject to microenvironmental and systemic regulation that enables them to suppress any potential malignant characteristics of the cells. This homeostasis can only be maintained as long as a critical amount of tumorigenic stimuli is not exceeded. As proven by cancer patients that relapse years after a successful therapy, a potential cancer cell can be maintained and suppressed for a long period of time. The experimental systems typically used to study the mechanisms that push potential cancer cells beyond the control of tissue homeostasis have been based on the use of tumor promoters such as phorbol esters that, although not mutagenic in themselves, induce tumors in animals previously mutagenized by chemicals (25).

We have used an experimental model that takes advantage of the ability of MAV-2 to randomly integrate its provirus into the host genome to create somatic mutations. In chicks, the process of infection and proviral integration is essentially completed within a few days after infection. Additional insertions are strongly limited due to a viral interference mechanism (26). However, before interference develops, permissive cells can acquire multiple proviruses. This model mimics the slow accumulation of naturally occurring somatic mutations. An advantage of the retroviral system is that mutated loci contain retroviral sequences so that they can readily be identified in tumor clones (27). Furthermore, the specific integration patterns allow independent tumor clones to be distinguished one from another.

In chickens, it is well established that the predominant tumors caused by MAV-2 infection are clonal nephroblastomas (22). Occasionally, we have also observed the formation of other tumor types predominantly of lung and liver origin that arose independently of

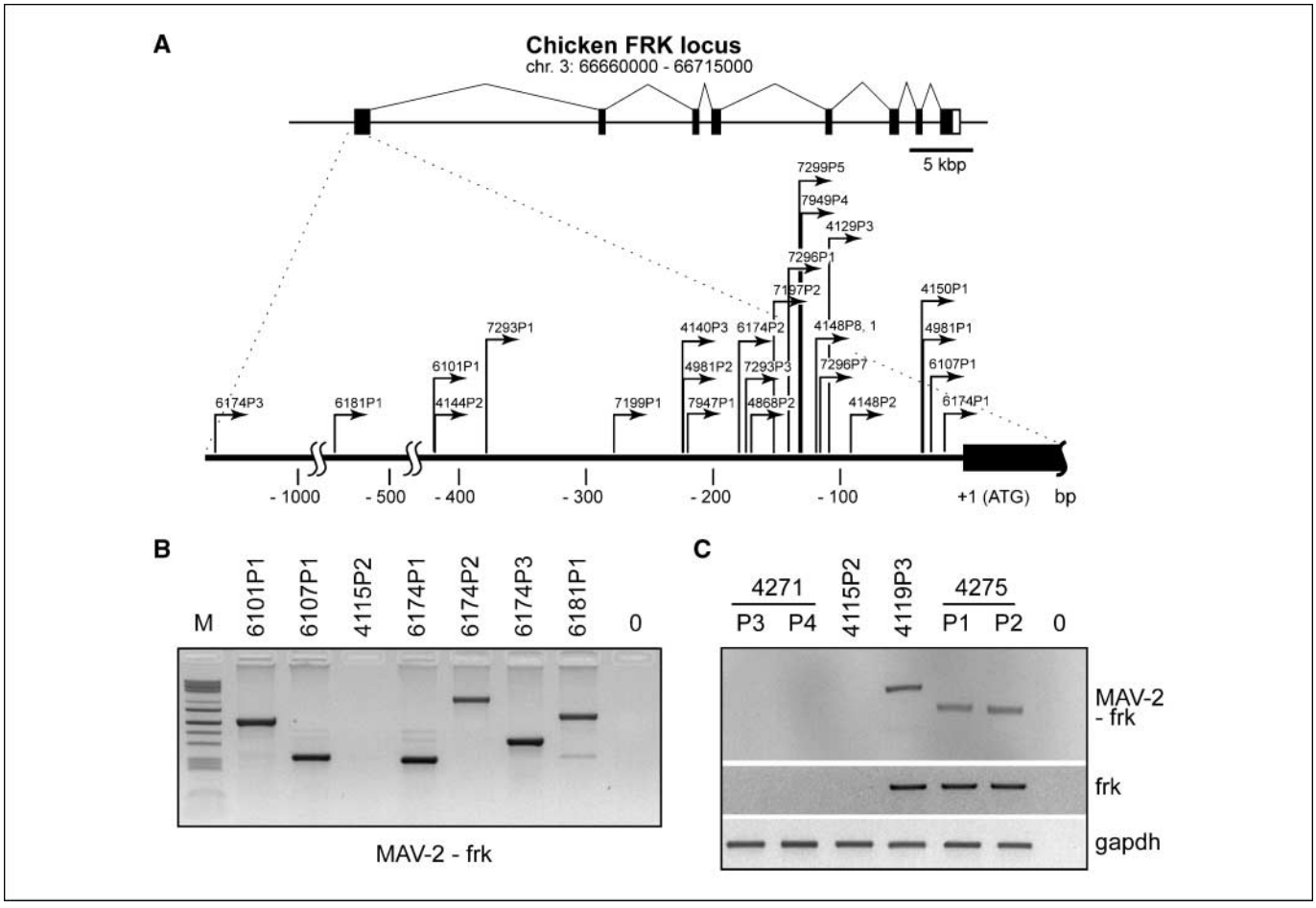


Figure 5. Common site of proviral integration in the *frk* locus. *A,* the exon-intron structure of chicken *frk* genomic locus and sites of proviral integrations upstream from ATG initiation codon. Arrows indicating integration sites and orientation of transcription are marked by the code number of lung tumors in which they were found. *B,* an example of detection of hybrid MAV-2-frk mRNAs synthesized as a result of MAV-2 integration, by integration site-specific RT-PCR (described in Supplementary Fig. S3*A*). The variation in lengths of PCR products reflects the distance of individual integration sites from the *frk*-derived primer. The tumor 4115P2 has no integration in *frk* gene. *0,* reaction without cDNA. *C,* RT-PCR analyses of *frk* expression using MAV-2 and frk (*top*) or only frk primers (*middle*). The majority of *frk* mRNA is produced from the proviral promoter. The tumors without integration in the *frk* locus do not express the *frk* gene. *Bottom,* control *gapdh* expression.

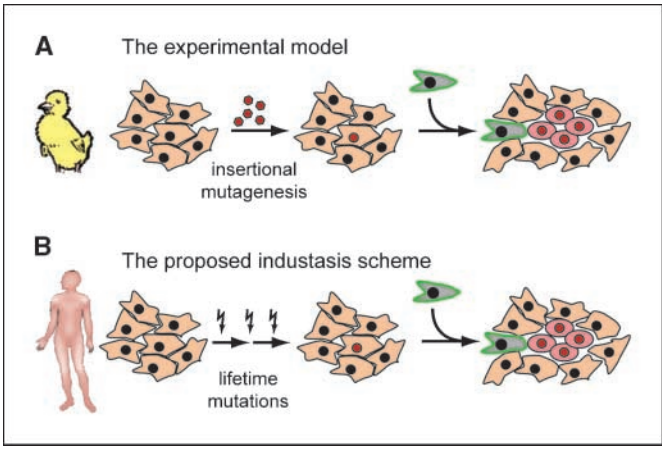


Figure 6. Industasis concept. *A,* in the chicken model (*top scheme*), the insertional mutagenesis by MAV-2 (*red hexagons*) results in the formation of a precancerous initiated cell (*red nucleus*) and is immediately followed by a tumor-promoting effect of an injected stray nontumorigenic cell (*gray*), resulting in the formation of a tumor clone (*red oval cells*). *B,* in the proposed general model (*bottom scheme*), mutations accumulate throughout the life span of an individual forming the genetically transformed initiated cell. A stray nontumorigenic cell liberated for instance from a primary tumor affects the microenvironment of an initiated cell promoting its malignant progression.

the nephroblastomas. As such, these tumors induced by MAV-2 insertional mutagenesis represent multiple primary tumors. Compared with nephroblastomas, these second primary tumors displayed a different mutation signature. In nephroblastomas, the *twist*, *plag1*, and *foxP1* loci represent dominant common sites of proviral integration (14). In contrast, lung tumors had very frequent insertions in the promoter region of the *frk* gene, never found in nephroblastoma. This observation for the first time documents oncogenic capability of *frk* overexpression, previously being only suspected (28). This result supports the concept that distinct mutational signatures can exist in cancers derived from different tissues (29).

The injection of cells into the blood circulation of MAV-2 mutagenized animals resulted in the appearance of numerous early lung tumors with the same mutation signature as the rare late lung tumors in animals injected with the MAV-2 alone (i.e., proviral integration in the *frk* gene). Thus, the presence of stray cells changed the formerly rare and late second tumors to frequent and early primary neoplasias. We suggest that the preferential promotion of lung tumors by stray cells is due to the hemodynamics, specific characteristics of the walls of fine pulmonary capillaries, and tissue-specific activity of the retroviral promoter/enhancer.

Based on the above observations, we propose the concept of industasis (Fig. 6). This mechanism of cancer promotion may take place in single preneoplastic cells or cell compartments (cancerized fields) that have accumulated cancerous mutations yet remain under the control of tissue homeostasis. Stray cells may interfere with the regulation of the local microenvironment that maintains tissue homeostasis through cell-cell adhesion/communication and cell-extracellular matrix interactions. By locally weakening homeostasis, stray cells may reduce the amount of tumorigenic alterations required for predisposed cells to express their malignant character. Eventually, presence (or death) of stray cells may elicit signals stimulating proliferation normally associated with wound healing. Once the tumor-predisposed cells start to proliferate, they lose contact with suppressive microenvironment and may become self-sustaining by creating supportive microenvironment for themselves, including recruitment of supportive stromal cells (30). The proposed mechanism could be the final step in genesis of several human multiple primary tumors (Fig. 6*B*). Human cells accumulate numerous somatic mutations during their lengthy life span, and it is well documented that the suppression of genetically transformed cells by the surrounding microenvironment represents an important defense against tumor outburst (3). The organism hosting one primary tumor is often flooded by stray cells liberated from the tumor (31). Their role is currently supposed to constitute a substantial risk of forming metastases (32). However, these stray cells could disturb the suppressive defense mechanisms, allowing pretransformed cells to grow.

Our hypothesis suggests that a substantial portion of human second primary tumors might be provoked by industasis. Such cases might be frequent as there is evidence that several presumed metastases could be in fact misdiagnosed second primary tumors (33).

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Received 12/9/08; revised 2/25/09; accepted 3/30/09.

Grant support: Grant Agency of the Czech Republic grant 204-06-1728, Grant Agency of the Academy of Sciences of the Czech Republic grants AV0Z50520514 and A500520608, and Ministry of Education, Youth and Sports grant LC06061 (M. Dvořák).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Prof. Jan Svoboda for critical comments and Dr. Alicia Corlett for the help with manuscript preparation.

References

1. Karnoub AE, Dash AB, Vo AP, et al. Mesenchymal stem cells within tumour stroma promote breast cancer metastasis. Nature 2007;449:557–63.
2. Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. Nature 2008;454:436–44.
3. Rubin H. Contact interactions between cells that suppress neoplastic development: can they also explain metastatic dormancy? Adv Cancer Res 2008;100:159–202.
4. Ashworth TR. A case of cancer in which cells similar to those in the tumors were seen in the blood after death. Aust Med J 1869;14:146.
5. Butler TP, Gullino PM. Quantitation of cell shedding into efferent blood of mammary adenocarcinoma. Cancer Res 1975;35:512–6.
6. Meng S, Tripathy D, Frenkel EP, et al. Circulating tumor cells in patients with breast cancer dormancy. Clin Cancer Res 2004;10:8152–62.
7. Podsypanina K, Du YC, Jechlinger M, Beverly LJ, Hambardzumyan D, Varmus H. Seeding and propagation of untransformed mouse mammary cells in the lung. Science 2008;321:1841–4.
8. Raa ST, Oosterling SJ, van der Kaaij NP, et al. Surgery promotes implantation of disseminated tumor cells, but does not increase growth of tumor cell clusters. J Surg Oncol 2005;92:124–9.
9. Tlsty TD, Hein PW. Know thy neighbor: stromal cells can contribute oncogenic signals. Curr Opin Genet Dev 2001;11:54–9.
10. Hayat MJ, Howlader N, Reichman ME, Edwards BK. Cancer statistics, trends, and multiple primary cancer analyses from the Surveillance, Epidemiology, and End Results (SEER) Program. Oncologist 2007;12:20–37.
11. Dong C, Hemminki K. Multiple primary cancers of the colon, breast and skin (melanoma) as models for polygenic cancers. Int J Cancer 2001;92:883–7.
12. Haddy N, Le Deley MC, Samand A, et al. Role of radiotherapy and chemotherapy in the risk of secondary leukaemia after a solid tumour in childhood. Eur J Cancer 2006;42:2757–64.
13. Pajer P, Pecenka V, Karafiat V, Kralova J, Horejsi Z, Dvorak M. The twist gene is a common target of retroviral integration and transcriptional deregulation in experimental nephroblastoma. Oncogene 2003;22:665–73.
14. Pajer P, Pecenka V, Kralova J, et al. Identification of potential human oncogenes by mapping the common viral integration sites in avian nephroblastoma. Cancer Res 2006;66:78–86.
15. Cance WG, Craven RJ, Bergman M, Xu L, Alitalo K,

Liu ET. Rak, a novel nuclear tyrosine kinase expressed in epithelial cells. Cell Growth Differ 1994;5:1347–55.
16. Lee J, Wang Z, Luoh SM, Wood WI, Scadden DT. Cloning of FRK, a novel human intracellular SRC-like tyrosine kinase-encoding gene. Gene 1994;138: 247–51.
17. Akerblom B, Anneren C, Welsh M. A role of FRK in regulation of embryonal pancreatic β cell formation. Mol Cell Endocrinol 2007;270:73–8.
18. Pecenka V, Dvorak M, Karafiat V, et al. Avian nephroblastomas induced by a retrovirus (MAV-2) lacking oncogene. II. Search for common sites of proviral integration in tumour DNA. Folia Biol Praha 1988;34:147–69.
19. Karafiat V, Dvorakova M, Pajer P, et al. The leucine zipper region of Myb oncoprotein regulates the commitment of hematopoietic progenitors. Blood 2001;98:3668–76.
20. Plachy J. The chicken—a laboratory animal of the class Aves. Folia Biol Praha 2000;46:17–23.
21. Bonifacino JS, Dasso M, Harford JB, Lippincott-Schwartz J, Yamada KM. Current protocols in cell biology. Hoboken (NJ): John Wiley and Sons, Inc.; 2007.
22. Watts SL, Smith RE. Pathology of chickens infected with avian nephoblastoma virus MAV-2(N). Infect Immun 1980;27:501–12.
23. Plachy J, Pink JR, Hala K. Biology of the chicken MHC (B complex). Crit Rev Immunol 1992;12:47–79.
24. Valk PJ, Joosten M, Vankan Y, Lowenberg B, Delwel R. A rapid RT-PCR based method to isolate complementary DNA fragments flanking retrovirus integration sites. Nucleic Acids Res 1997;25:4419–21.
25. Berenblum I, Shubik P. The role of croton oil applications associated with a single painting of a carcinogen in tumor induction of the mouse's skin. Brit J Cancer 1947;1:379–82.
26. Coffin JM, Hughes SH, Varmus HE. Retroviruses. Cold Spring Harbor (NY): CSHL Press; 1997.
27. Silver J, Keerikatte V. Novel use of polymerase chain reaction to amplify cellular DNA adjacent to an integrated provirus. J Virol 1989;63:1924–8.
28. Hosoya N, Qiao Y, Hangaishi A, et al. Identification of a SRC-like tyrosine kinase gene, FRK, fused with ETV6 in a patient with acute myelogenous leukemia carrying a t(6;12)(q21;p13) translocation. Genes Chromosomes Cancer 2005;42:269–79.
29. Yeang CH, McCormick F, Levine A. Combinatorial patterns of somatic gene mutations in cancer. FASEB J 2008;22:2605–22.
30. Kalluri R, Zeisberg M. Fibroblasts in cancer. Nat Rev Cancer 2006;6:392–401.
31. Allard WJ, Matera J, Miller MC, et al. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. Clin Cancer Res 2004;10:6897–904.
32. Mocellin S, Keilholz U, Rossi CR, Nitti D. Circulating tumor cells: the 'leukemic phase' of solid cancers. Trends Mol Med 2006;12:130–9.
33. Albanese I, Scibetta AG, Migliavacca M, et al. Heterogeneity within and between primary colorectal carcinomas and matched metastases as revealed by analysis of Ki-ras and p53 mutations. Biochem Biophys Res Commun 2004;325:784–91.

Chapter 4
Chicken Models of Retroviral Insertional Mutagenesis

Vladimír Pečenka, Petr Pajer, Vít Karafiát, and Michal Dvořák

The concept of transposon tagging or insertional mutagenesis as a strategy for fishing out genes connected with the phenotype of interest was emerging since the early 1980s. Study of genetic basis of tumorigenesis is one of the fields where insertional mutagenesis proved to be exceptionally powerful. Crucial elements of this experimental approach have been retroviruses whose unique properties have revolutionized the work in the field of oncogenesis. Retroviruses contributed to our knowledge of tumor formation in two ways. First, some of them transduce oncogenes—mutants of normal cellular genes with an oncogenic potential. And it was the comparison of viral and cellular alleles of these genes that allowed comprehending the principles of oncogenic activation of genes. Second, retroviruses not carrying oncogenes can induce tumors by affecting host genes. Through integration of their proviral DNA into chromosomes they can activate tumorigenic potential of oncogenes or inactivate tumor suppressor genes. The mechanism is referred to as oncogenesis by insertional mutagenesis. The insertional mutagenesis by retroviruses is very efficient. Perhaps each locus of a host genome can be hit by the provirus insertion in many cells of an infected tissue. If any of these insertions or their combinations incites malignant transformation, the touched cell outgrows and can give rise to a tumor. Affected host gene loci can be easily identified since they are tagged by integrated proviral sequences.

Chicken retroviruses and chickens as experimental animals laid the grounds to the entire field. Chicken B–lymphoma induced by the avian leukosis virus (ALV) was the first tumor type where this mechanism of tumorigenesis was reported in 1981. In these tumors the cellular oncogene *c–myc* was found activated by a provirus integration. In coming years, these pioneering works were followed by further papers that included other retroviruses and animal models and found other insertionally activated oncogenes in other tumor types, thus establishing insertional mutagenesis as a new paradigm in retroviral oncogenesis. The speed of further data accumulation was increasing with continuous improvement of the techniques

V. Pečenka (✉)
Department of Molecular Virology, Institute of Molecular Genetics, Prague, CZ–14220, Czech Republic

A.J. Dupuy, D.A. Largaespada (eds.), *Insertional Mutagenesis Strategies in Cancer Genetics*, DOI 10.1007/978-1-4419-7656-7_4,

Table 4.1 Insertionally activated oncogenes in chicken cancer models

Tumor	Virus	Oncogene	Gene type	Frequency of activation (%)	References
B–lymphoma	RAV	***c–myc***	transcription factor	> 95	[126]
B–lymphoma	RAV	***Bic/miR–155***	miRNA	14 (50% in metastasis)	[23]
B–lymphoma (short latency)	RAV, EU–8	***c–myb***	transcription factor	∼100	[23] [119]
B–lymphoma (short latency)	RAV	***tert***	telomerase catalytic subunit	∼14	[165]
erythoblastosis	RAV	***c–erbB/EGFR***	receptor tyr–kinase	> 95	[92] [123]
nephroblastoma	MAV	***foxP1***	transcription factor	∼6	[108]
nephroblastoma	MAV	***plag1***	transcription factor	∼5	[108]
nephroblastoma	MAV	***twist***	transcription factor	∼3	[108]
nephroblastoma	MAV	***c–Ha–ras***	G–protein	∼2	[108]
lung sarcoma	MAV	***frk***	tyrosine kinase	> 95	[106]
hepatocarcinoma	MAV	***c–Ha–ras***	G–protein	∼39	Pajer et al., in preparation
hepatocarcinoma	MAV	***c–erbB/EGFR***	receptor tyr–kinase	∼32	Pajer et al., in preparation
hepatocarcinoma	MAV	***c–ron/MST1R***	receptor tyr–kinase	∼11	Pajer et al., in preparation
hepatocarcinoma	MAV	***c–met/HGFR***	receptor tyr–kinase	∼7	Pajer et al., in preparation

for isolating provirus-flanking sequences (inverse PCR and related techniques) and for their analysis (automatic sequencing machines). The final explosion of data started after the complete genome sequences of individual model organisms had become available. From that point it has been a relatively easy task to assort individual provirus insertion sites along the chromosomes, to pick up common sites of integration and associate them with suspect genes.

In this chapter we will describe several chicken models of tumorigenesis through retroviral insertional mutagenesis that have been thoroughly investigated (for the overview, see Table 4.1). More space will be given to the recently advanced models: chicken nephroblastomas, lung sarcomas, liver carcinomas and the industasis phenomenon. We will focus on unique features of the chicken models including mechanistic details of gene activation by defective proviruses. Insertional inactivation of tumor suppressor genes will not be discussed since it has not been proven in chicken system so far. For general issues we refer to other parts of this book and to the reviews [75, 150]. In the end, we will discuss the future of chicken models in the era of high throughput oncogene screening in mouse models and large scale sequencing of human cancer genomes.

4.1 The Beginning of the Story: the Case of Chicken Bursal Lymphomas

Bursal lymphoma (malignant lymphoma of the bursa of Fabricius named also lymphoid leukosis or lymphomatosis) is the most common neoplasia of domestic chickens. The viral etiology of bursal lymphomas has been known since 1908 when cell-free extract was shown to transmit the neoplasia [41]. In farms, the causative agent, now called avian leukosis virus (ALV), is spread mainly by congenital infection from hens to embryos [16]. Experimentally, the virus can be transmitted by intraperitoneal injection of plasma of infected animal into newly hatched susceptible chickens. Several weeks after infection many microscopic transformed follicles arise in bursa of Fabricius (see also Section 4.7). Within months, one or more macroscopically observable nodules—bursal lymphomas—develop. The animals die after progressive metastatic tumors develop in many organs [4, 28].

ALVs are typical representatives of simple slowly oncogenic retroviruses carrying only viral replicative genes gag, pol and env to which no substantial transforming potential could be ascribed [25]. The first suggestion about the mechanism by which these transforming gene-lacking viruses could induce neoplasia emerged from the study of recombinants between oncogenic ALVs (RAV–1 and RAV–2, Rous associated viruses subgroups A and B) and their nononcogenic homolog (RAV–0, Rous associated virus subgroup E). The oncogenicity of the recombinant always correlated with the presence of U3 region derived from RAV–1 or RAV–2 [31, 149]. The U3 region just precedes the start of viral mRNA transcription at the 5′ end of the integrated DNA provirus. Not surprisingly, sequence analyses have localized typical promoter sequences for RNA polymerase II in the U3 region [164].

Since the U3 region, as part of the long terminal repeat (LTR), is also present at the 3′ end of the provirus, it could drive transcription not only of retroviral genes, but also of downstream located host genes (the mechanism named "promoter insertion", see Fig. 4.1a). Hybrid RNAs containing both viral and cellular sequences have indeed been found in infected cells [122]. The non-oncogenicity of RAV–0

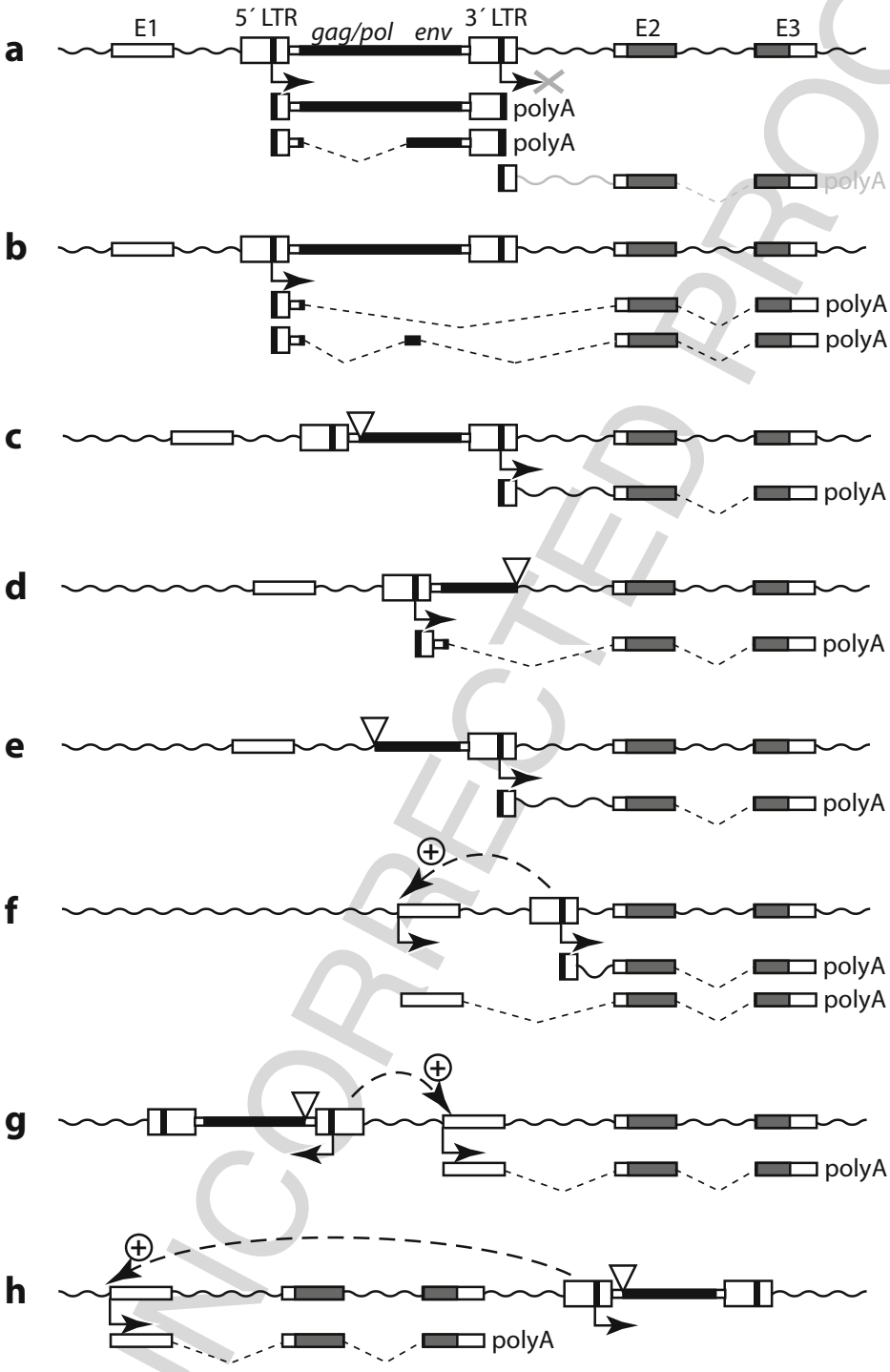


Fig. 4.1 Schemes of gene activation by promoter insertion (a–f) and enhancer insertion (f–h). a Nondeleted provirus. Predominant transcripts driven by proviral LTRs are full-length and spliced copies of the provirus initiated at the 5′LTR promoter. Transcription starting in the 5′LTR and proceeding through the 3′LTR arrests the promoter activity of the 3′LTR.

and of recombinants carrying RAV–0-derived U3 could be the consequence of the weakness of RAV–0 promoter which correlated with the slower replication of these viruses [149]. Thus, all the facts were consistent with the hypothesis of insertional activation of an adjacent cellular gene.

The confirmation came when several groups [27, 45, 54, 55, 96, 112] reported frequent provirus integration next to the same discrete cellular sequence in B-lymphomas (denominated as common integration site—CIS). In vast majority of B-lymphomas one of the integrated proviruses was found just upstream of the *c–myc* gene, at the time the already known cellular oncogene. The provirus position and orientation were exactly as required for *c–myc* insertional activation; high levels of hybrid viral/*c–myc* mRNAs were detected in tumors. The same situation was found in chicken B-lymphomas induced by CSV strain of REV [100]. (CSV is another leukemogenic chicken retrovirus lacking an oncogene, unrelated to ALV but classified with mammalian C-type retroviruses).

4.2 Efficient Activation of *C–MYC* Requires Defective Provirus: the Transcriptional Interference and Related Phenomena

More detailed analysis of mechanisms of insertional activation indicated that the original scheme (Fig. 4.1a) was oversimplified. The scheme, where both the 5′LTR

Fig. 4.1 (continued) **b** Nondeleted provirus–readthrough mechanism. Some transcripts initiated at the 5′LTR promoter continue through the weak polyA site in the 3′LTR and are processed at cellular polyA sites. Splicing takes place between the gag splice donor and the cellular splice acceptor. In some instances, a miniexon corresponding to the beginning of *env* is included. **c** Deleted provirus–internal deletion. Deletion of the sequence element downstream of 5′LTR disables 5′LTR promoter activity (see the proposed mechanism in Section 4.2 and Fig. 4.2a). Transcription of downstream sequences is driven by the released 3′LTR promoter. **d** Deleted provirus – 3′LTR deleted. Transcription starts in the 5′LTR. No readthrough is required. The env splice acceptor competing with the cellular splice acceptor is frequently also eliminated. **e** Deleted provirus–5′LTR deleted. Transcription starts in the 3′LTR promoter that was unblocked by 5′LTR deletion. **f** Deleted provirus–single LTR left. Transcription of a downstream cellular gene is driven by proviral promoter (promoter insertion). Alternatively, the authentic cellular gene promoter is stimulated by the proviral enhancer (enhancer insertion). In case the LTR has inserted inside the gene structure both transcription readthrough and splicing take place. **g,h** Enhancer insertion mechanism. Provirus is located downstream of the gene in the same transcriptional orientation or upstream of the gene in the reverse orientation. Transcription starting at the authentic gene promoter is stimulated by an enhancer in the proximate LTR. Also proviruses activating by enhancer insertion carry deletions (e.g. the internal one, as shown here); for details see Section 4.2 and Fig. 4.2b. Only schemes validated experimentally are presented here. Obviously, further schemes combining elements of those shown above can be conceived. Open and black boxes represent noncoding and coding parts of exons, respectively (except for the R region inside the LTRs which, though noncoding, is also black). Wavy lines are introns and non-transcribed regions of genomic DNA, cranked arrows indicate starts of transcription, broken dashed lines indicate splicing; open triangles show the points where deletion has occurred; dashed arrows symbolize stimulation of a promoter by an enhancer

and the 3′LTR promoters were supposed to be simultaneously active was in contradiction with the phenomenon of transcriptional interference, originally named promoter occlusion [1]. Particularly in retroviruses it was shown that an efficient initiation in the upstream promoter (i.e. 5′LTR) and progression through the downstream promoter (i.e. 3′LTR) prevents assembling of an initiation complex at the downstream promoter and abolishes its activity [33].

The distinction between 5′ and 3′LTR is further increased by the presence of positive regulatory elements outside the LTRs. In ALV, one such accessory element, called *gag* enhancer, is located in 5′ part of the *gag* gene and acts preferentially on 5′LTR [3, 129]. In CSV, other accessory element lies immediately downstream of 5′LTR but does not behave like an enhancer—it is operative only when located downstream of the promoter and only if present in the proper orientation [10]. That suggests the element performs on the transcript rather than on DNA level, alike TAR element in HIV [25]. Since both elements selectively activate promoter in 5′LTR, they, indirectly, through transcriptional interference, reinforce 3′LTR promoter inhibition.

Accordingly, when the structure of LTR-initiated transcripts in unselected population of ALV-infected cells was carefully examined, less than 2% of them have been found to initiate in 3′LTRs [58]. There was ca 15% of retroviral transcripts that contained both viral and cellular sequences but these hybrid RNAs have been generated by initiation at 5′LTR and by transcription passing through the leaky 3′LTR polyA signal into adjacent host sequences; polyadenylation took place at a distant cellular poly A signal [58]. Hence, the expression of *c–myc* oncogene in B-lymphomas could be, in theory, mediated by hybrid readthrough RNA instead of 3′LTR-initiated RNA. The readthrough RNA would have to be further processed by splicing; if unprocessed, it would code for the viral gag/pol proteins only (see Fig. 4.1b). The splicing would join first six codons of gag with the *c–myc* coding exon; incidentally, such splicing would properly, without a frameshift, join gag and *c–myc* reading frames. However, the readthrough mechanism of *c–myc* activation was not observed in B-lymphomas (see below). Why this mechanism is not effective in the case of *c–myc* in B-lymphomas (while it works quite well with other oncogenes, see next sections) can be explained by the fact that it involves two relatively inefficient steps: 3′LTR readthrough and splicing to *c–myc* second exon splice acceptor while skipping the env splice acceptor. Apparently, the *c–myc* expression levels obtained would not reach the threshold level required for chicken B-lymphomagenesis [109].

Final resolution of activation mechanisms followed from detailed analyses of ALV/*c–myc* arrangement in B-lymphomas. Extensive mapping showed that none out of nearly one hundred of *c–myc* activating proviruses under study was complete [49, 111, 126]. The most frequent defects were internal deletions from 0.5 to 6 kb long that included sequences in the 5′region of the provirus immediately downstream of the LTR. Such deletions result from reverse transcriptase switching the template during the provirus synthesis [168]. Other defects were 5′end or 3′end deletions, some of them incurred during or after the integration since the deletion extended into adjacent cellular sequences. Sometimes nearly all proviral sequences

were found deleted leaving only solitary LTR, probably the result of homologous recombination between the two LTRs.

Frequent generation of defects during retrovirus propagation has been well documented [72, 79, 135, 155]. That, however, does not explain why every provirus in *c–myc* locus carries a deletion while majority of other proviruses in the same cells are nondefective. It had to be postulated that only defective proviruses could activate *c–myc* and only cells with such insertion were selected during oncogenesis. Indeed, it can be demonstrated that defects in activating proviruses eliminate obstructions to efficient activation. When 3′ LTR is deleted (Fig. 4.1d), inefficient readthrough step is obviated—all transcripts proceed from 5′LTR into the host sequences; moreover, competing env splice acceptor is often also deleted which makes the splicing from gag donor to the *c–myc* second exon acceptor much more efficient. When 5′ LTR is deleted (Fig. 4.1e) or there is only solitary LTR left (Fig. 4.1f), transcription starts in the remaining LTR whose activity was released from transcriptional interference. In all the above cases the authentic *c–myc* promoter is still present upstream of the inserted LTR but its activity is low and does not interfere with the LTR activity.

Internal deletions in proviruses (Fig. 4.1c) had somewhat unexpected consequences. These deletions completely turned off 5′LTR while 3′LTR became fully active [49]. The very same situation (i.e. disabling 5′LTR by deletion of adjacent downstream sequences) was found in *c–myc* activating proviruses in chicken B-lymphomas induced by CSV [147]. This provokes a question of why 5′LTR becomes completely inactive after losing its downstream accessory elements while 3′LTR that also does not have these elements turns fully active. Presently we can only hypothesize that regulatory interrelations in retroviral provirus are more complex then we have supposed and that any disruption of the provirus integrity can modify activity of both LTRs in a way we are still unable to predict. The existence of additional proviral regulatory elements and interactions including mutual interactions of both LTRs has already been suggested by [101].

The explanation can ensue when the original promoter occlusion/transcriptional interference hypothesis is modified to comply with the present-day knowledge of how distant regulatory regions on chromosome execute their effects. Transcriptionaly active regions including retroviral proviruses form dynamic chromatin loop structures. Enhancer and promoter regions, together with all associated regulatory proteins, are attached to each other at the loop base creating so called active chromatin hub. After a pioneer round of transcription, a terminator region is also juxtaposed to the promoter. In this way a transcriptional unit is demarcated and recycling of transcriptional apparatus in this unit is facilitated [35, 40, 116, 159]. In the integrated provirus, however, the situation is more intricate: there are two identical sets of enhancer, promoter and terminator (i.e. 5′LTR and 3′LTR), preceded by original *c–myc* promoter plus enhancer and followed by original *c–myc* terminator. Thus, several alternative looping structures may form; the final structure is a result of mutual competition for complex formation between individual elements. We suggest that accessory element in 5′part of the complete provirus assist preferential joining of 5′ and 3′LTRs into the loop, presumably by itself forming part of the complex

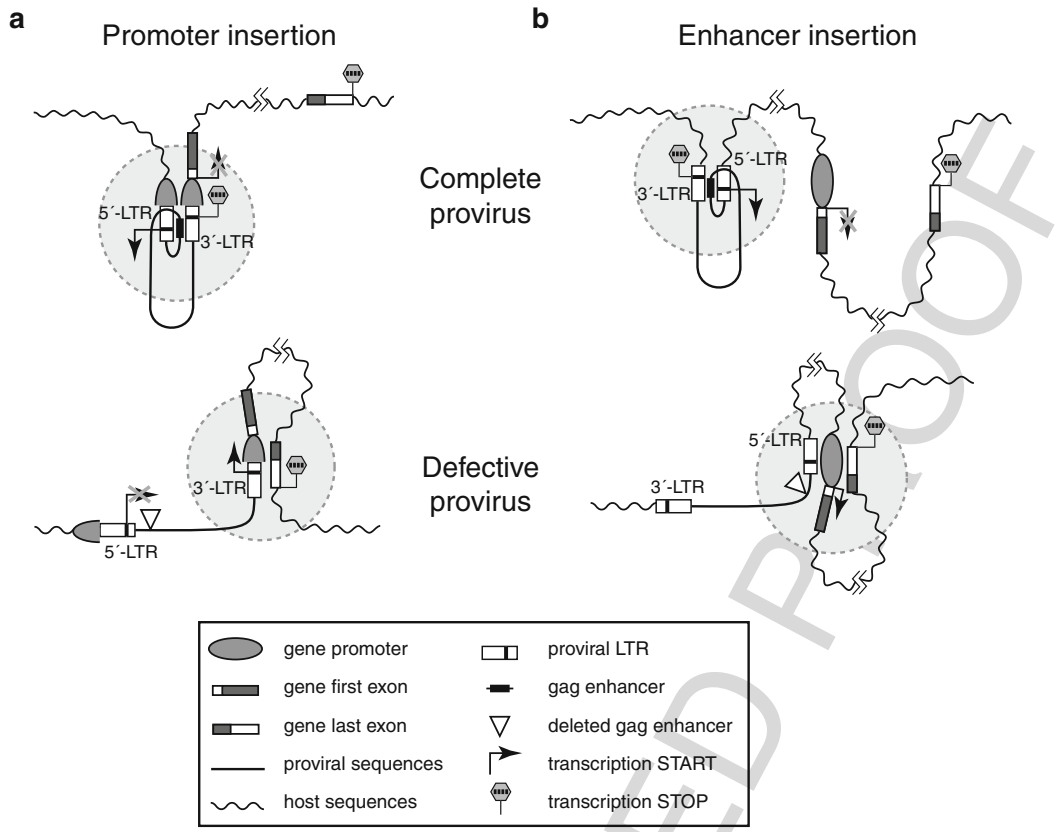


Fig. 4.2 Organization of transcription units controlled by promoters and enhancers of nondeleted and internally deleted integrated proviruses: A hypothesis. **a** Promoter insertion. The complete integrated provirus forms an isolated highly active transcription unit through the formation of a compact chromatin loop structure (active chromatine hub–encircled *gray area*) that is stabilized by interaction of the accessory gag enhancer with both LTRs. In this loop the transcriptional machinery is permanently recycled from the 3′LTR to the 5′LTR. When the gag enhancer is deleted (defective provirus), the proviral loop is destabilized, 5′LTR turns inactive and 3′LTR can form the active transcriptional loop with downstream sequences. **b** Enhancer insertion. The complete provirus forms the active transcriptional unit in which the proviral enhancer is engaged, as in **a**. Therefore the proviral enhancer exerts no influence on regulation of neighboring genes. However, when the provirus chromatine loop is destabilized by deletion (defective provirus), the enhancer is free to interact with neighboring host sequences and stimulates the authentic host gene promoter. All the contacts are mediated by protein complexes associated with promoter, enhancer and terminator regions (not shown in this schematic drawing)

(Fig. 4.2a). The accessory element would thus share some functional properties with the promoter targeting sequence (PTS) described in Drosophila which promotes association of a specific enhancer with a specific distant promoter even when an insulator/enhancer blocker element is present between them [80, 170].

The 5′LTR—3′LTR chromatine hub may be arranged so that an initiation complex is formed only in 5′LTR promoter; however, the existence of the loop itself is sufficient to determine which of the two proviral promoters remains active. When transcription starts in 3′LTR, the transcription complex leaves the loop and new round of promoter selection begins; when transcription starts in 5′LTR, the promoter

in 3′LTR is silenced by transcriptional interference and transcription machinery is permanently recycled from 3′LTR back to 5′LTR. In such arrangement, enhancers in both LTRs may stimulate transcription from a single, 5′LTR promoter—the effect observed by [101]. When internal proviral accessory element is deleted as it is the case in most of the *c–myc* activating proviruses the proviral transcriptional unit is destabilized and other looping structures preferentially form including the 3′LTR—*c–myc* polyA loop that demarcates new, highly active c–myc transcriptional unit (Fig. 4.2a).

In several tumors [49, 111, 126], the provirus was located upstream of *c–myc* in the reverse transcriptional orientation or downstream of *c–myc* in the same transcriptional orientation (Fig. 4.1g and 1h). Such arrangement did not conform with the promoter insertion model as transcription starting from any provirus LTR could not proceed to *c–myc* sequences. Nevertheless, substantially increased levels of the *c–myc* mRNA have been found. Its synthesis started at the authentic *c–myc* promoter the activity of which was, apparently, stimulated by juxtaposed powerful enhancer in proviral LTR. This configuration has been denominated "enhancer insertion". It has been reported that also isolated LTR can activate by the enhancer insertion mechanism [81]: in two B-lymphoma lines carrying a single LTR insertion in the first *c–myc* intron in the same transcriptional orientation the LTR-contained enhancer stimulated transcription from the original *c–myc* promoter while the LTR-contained promoter was silenced by transcriptional interference (Fig. 4.1f).

Also proviruses activating *c–myc* by enhancer insertion have been shown to carry internal deletions [49, 126]. That is not surprising when we accept the model presented in previous paragraphs: to be able to attach and stimulate *c–myc* promoter, the relevant proviral enhancer must first be released from the complex with other proviral elements. Deletion of accessory elements in 5′ part of the provirus may be the easiest way of such enhancer liberation (Fig. 4.2b). Obviously, previous reasoning presumes that proviral enhancer cannot make a complex with and stimulate two promoters simultaneously. Such assumption is quite acceptable considering very condensed arrangement of the enhancer into less than 100 bp; it is also in compliance with the provirus tendency to form isolated highly active transcriptional unit.

The support for the above model can be found in a way how MLV and FeLV activate oncogenes. They do it predominantly by enhancer insertion mechanism; however, in contrast to ALV, activating MLV and FeLV proviruses are generally not defective [97, 143]. At first sight it may seem to contradict the proposed model. However, the opposite is true. The above model can, in fact, explain one peculiarity of oncogenesis by MLVs and FeLVs: absolute majority of activating proviruses carry enhancer duplications or even higher level multiplications. For example, lymphomagenesis in some mice strains like AKR relies on spontaneous genesis of oncogenic recombinants (named MCF, mink cell focus-forming virus) from endogenous MLV proviruses present in the strain. Indispensable component of the oncogenic MCF recombinant is a presence in LTR of two copies (in tandem) of strong enhancer acquired from an endogenous provirus in which only single copy of the enhancer is present [143]. Highly lymphomagenic exogenous

MLV strains isolated from naturally occurred tumors, like Mo–MLV, already contain duplicated enhancer [25]. In tumorigenic exogenous FeLVs there is only a single copy of an enhancer; however, in induced T-lymphomas, where *c–myc* is activated by the enhancer insertion mechanism, all activating proviruses carry tandem enhancer duplication acquired always de novo during infection of the animal [97].

The enhancer duplication was shown to have only mild effect on virus replication and provirus transcription. When single-enhancer MLV or FeLV viruses are propagated in culture the anticipated variants with duplication do not arise (or, more accurately, are not positively selected for). Hence, rather than being important for the virus spread in the animal, the duplication is required for the mechanism of oncogenesis itself [97, 143]. We suggest that presence of tandem copies of an enhancer enables provirus 5′LTR to form two chromatin loops simultaneously. While one enhancer is engaged in the formation of proviral transcriptional unit the other one may contact oncogene promoter; no provirus defect is needed for the oncogene activation.

To summarize, distinct ways of how to make provirus capable of strongly affecting host genes are being employed during lymphomagenesis by avian versus mammalian retroviruses. In the ALV model, compactness of the proviral transcriptional unit is broken by different types of provirus deletions, mostly by deletion of the accessory element outside the LTR. Transcription of proximate host genes is then driven by the promoter of defective provirus (mechanism of promoter insertion). In MLV and closely related FeLV, however, no internal accessory elements have been reported. It is possible that compactness of the MLV and FeLV transcriptional units is not dependent on internal accessory elements; thus it cannot be so easy disrupted by provirus deletion. Instead, another way to generate strongly activating proviruses is realized: enhancer multiplication that capacitates provirus to boost transcription of neighboring genes from the genes' own promoters at a distance (mechanism of enhancer insertion). Accordingly, while the mechanism of oncogene activation by enhancer insertion is found only exceptionally in chicken models, it is the dominating mechanism in MLV and FeLV models.

In this context it would be interesting to find out how gene activation works in MMTV model, where enhancer insertion is also the prevalent mechanism. Though the data concerning structural changes in oncogene-activating MMTV proviruses are scarce, it seems that in mammary carcinomas neither deletion nor enhancer duplication takes place [102] while in variant MMTV-induced T-lymphomas enhancer multiplications may play an important role [6, 14]

For the sake of completeness we have to mention additional possible effects of provirus insertion that have not been discussed above. First, an important part of insertional activation mechanism may be separating negative regulatory elements like silencers from the oncogene promoter. Eventually, cellular transcription pausing sites may be shifted away from the transcription unit by the provirus insertion. Presence of negative elements is easy to imagine in cases where the activated gene has been completely inactive before provirus integration. These mechanistic

details, however, are commonly not addressed by the investigators. Second, the range of effects incited by the provirus integrating into the gene structure is not limited to the level of transcription. For example, any change in mRNA structure (truncation, fusion with viral sequences, activation of cryptic promoters and splicing or polyadenylation sites etc.) may have profound impact on its export from the nucleus, stability, and rate of translation. In mouse models, the most common mechanism of this type may be the frequent truncation of 3′untranslated regions containing destabilizing AT rich sequences as well as target sequences for miRNAs which downregulate both stability and translatability of affected mRNA. Such effects during insertional activation have occasionally been suspected or even documented, see e.g. [20, 61, 132–134, 152], but only recently have also been more thoroughly analyzed [34].

4.3 Readthrough Activation of *C–MYB* in B-lymphomas: the Case of "Superactivating ALV"

When 10–14 day old embryos, instead of newly hatched chickens, were intravenously injected with ALV, some of them developed, primarily in livers, highly aggressive B-lymphomas after latency of only several weeks (from now on these tumors will be called short latency lymphomas as opposed to long latency lymphomas arising after infection of newborn chicks—see Section 4.2). In most of short latency lymphomas ALV provirus was detected around (upstream, inside or downstream) the first exon of the *c–myb* gene, a cellular progenitor of retrovirally transduced oncogene *v–myb* [23, 117, 118]. The proviruses had no defects and were integrated in the same transcriptional orientation as the *c–myb* gene. Chimeric ALV/*c–myb* mRNAs were synthesized by a typical readthrough mechanism (see Section 4.2 and Fig. 4.1b): they were initiated in 5′LTR, polyadenylated at *c–myb* polyA site and processed by splicing the ALV gag donor to the *c–myb* second exon acceptor. Gag and *c–myb* were not coded by the same reading frame in this message. Moreover, gag start codon was followed by three in frame stop codons at the beginning of *c–myb*; translation of *c–myb* relied on reinitiating at the codon 21. The produced slightly truncated protein was strongly oncogenic, in contrast to the wt protein [63].

Profound effect of infection timing suggests that target cells for short latency lymphomas are different from target cells for long latency lymphomas (presumably prebursal stem cells versus bursal stem cells) and that the first ones are absent in chickens after hatching [118]. In consequence, different target cells are sensitive to activation of different oncogenes (*c–myb* versus *c–myc*). Consistently, no lymphomas develop when infection of the bursa is postponed to the time when bursal stem cells are no more present (ca 3 weeks after hatching), even in animals with subsequent strong lifelong viremia [15].

The frequency of short latency lymphomas and *c–myb* activations were considerably increased when specific recombinant ALV strains were used for infection.

In these "superactivating" strains an element in the 5′part of gag gene, called negative regulator of splicing (NRS), was knocked out by either deletion or mutation [8, 71, 119, 136, 138]. NRS has two roles in ALV both of which help to fulfill peculiar requirements of retroviral replication: to leave substantial proportion of retroviral RNA unspliced and, at the same time, to ensure polyadenylation at the 3′end of such RNA (which is generally coupled to splicing); for details see [24, 84]. To this end NRS first acts as a pseudo-splice donor, forming nonproductive splicing complexes with downstream acceptors thus blocking them from splicing to authentic donor sites. Second, by binding factors that interact with polyadenylation machinery, NRS helps to recruit and stabilize processing complexes at the 3′LTR polyA site. The second effect is dependent on the first one, since only the formation of an abortive splicing complex between NRS and downstream splicing acceptor brings NRS close enough to 3′LTR polyA site to exert its effect on polyadenylation [84].

As shown by [93] on RSV model, deletion of NRS has serious consequences to viral RNA processing. First, 3′LTR readthrough is strongly elevated (up to over 50%). Second, while the frequency of splicing from the gag donor to the env acceptor is, surprisingly, little affected, the splicing to a more distant acceptor is boosted ca four times. Similar (though less conspicuous) effects were observed for "superactivating" ALVs with the mutated NRS [105, 138]. Thus, when such ALV integrates upstream of a cellular gene, the spliced readthrough mRNA coding for the cellular gene is produced with much higher efficiency compared to wt ALV.

4.4 Insertional Activation Can Be Accompanied by Extensive Alterations of the Oncogene Structure and by Formation of an Oncogene-Transducing Virus: the Case of Chicken Erythroblastosis

Besides short- or long-latency B-lymphomas, chickens infected with ALV can infrequently develop other types of neoplasia (see Section 4.8). Among them, the most thoroughly analyzed was erythroblastosis, a disease characterized by uncurbed proliferation of erythroblasts with arrested differentiation. In sensitive chicken strains high amounts of erythroblasts appear in blood ca 2 months after infection with ALV; at the same time they start to massively infiltrate spleen and liver causing rapid death of the animal. Paradoxically, erythroblastosis is often accompanied by anemia since feedback mechanisms triggered by the flood of transformed erythroblasts block further development of nontransformed red blood cell precursors.

In 1983, recurrent insertional activation of the *c–erbB* protooncogene (or *EGFR*—epidermal growth factor receptor gene) in chicken erythroblastosis was first reported [46]. Further papers [50, 92, 99, 123] described peculiar molecular details of this activation. In all analyzed cases, the full-length ALV provirus

was integrated in the same intron in the middle of *c–erbB* gene in the same transcriptional orientation. High levels of *c–erbB* specific mRNAs were produced. The mechanism of activation was much the same as activation of *c–myb* in short latency B-lymphomas (Section 4.3): readthrough transcript starting in 5′LTR and polyadenylated at one of the two authentic *c–erbB* polyA sites was processed by splicing that brought *c–erbB* coding sequences close to the start of mRNA (Fig. 4.1b). There were, however, two differences. First, mRNA processing was more complex. An alternative short exon demarcated by env splice acceptor and cryptic splice donor 159 bp downstream was included in high proportion of messages. Second, the extent of oncogene truncation was much more extensive then in the case of *c–myb*. The protein encoded by the spliced mRNAs consisted of 6 AA of gag (plus 53 AA of env in some messages) fused to the C-terminal part of c–erbB. In both protein versions the N-terminal part—the EGF binding extracellular domain—was missing. Such truncation is known to result in a constitutive kinase activity of the receptor [77]. It was suggested that the 53 AA+ version of activated c–erbB may be the major mediator of oncogenesis. The alternative short exon codes for a signal peptide that normally targets env protein for membrane translocation. This function was retained also in a fusion with truncated c–erbB, where it substituted for missing authentic c–erbB signal peptide. During the processing the 53 AA^+ protein had the signal sequence cleaved off, was fully glycosylated and exposed on the cell surface, in contrast to the 53 AA^- version, which was only partially glycosylated and was not exposed on the cell surface [85].

The frequency of erythroblastosis is just several percent in most chicken strains including outbred flocks; there are however inbred strains that develop ALV-induced erythroblastosis with penetrance ca 80%. This sensitivity is a dominant trait not connected with the ability of ALV to activate *c–erbB*, but rather determining sensitivity of the animal to such activation [127].

In high proportion of tumors the situation was obscured by formation of *c–erbB*-transducing recombinant virus. While oncogene transduction is generally a rare event [25], highly leukemogenic viruses carrying truncated version of *c–erbB* can be isolated from up to 50% of chickens with ALV-induced erythroblastosis [92, 124]. That has made erythroblastosis an attractive model and rich source of material for studying mechanisms of transduction. For nearly all other isolates of acutely transforming retroviruses the original tumor where the virus has originated is not available and the virus has further evolved during repeated passages since then. Here, by comparing sequences of the transducing viruses with sequences of the insertionally activated *c–erbB* genes and fusion mRNAs synthesized in the tumors it was possible to confirm basic scheme of transduction as suggested earlier. Details of the mechanism are beyond the scope of this book; for them, see [25, 59, 92, 124, 128, 146]. The important point is that the first step of transduction is the synthesis of hybrid mRNA containing both the retroviral and oncogene sequences, i.e. the transduction is preceded by an insertional mutagenesis. Hence it is not surprising that the sites of truncation in insertionally activated *c–erbB* and in the transduced *v–erbB* precisely coincided.

4.5 Chicken Nephroblastomas Induced by MAV: Complex Model Suitable for High Throughput Oncogene Screening

In previous chapters we have described mechanisms of insertional mutagenesis by retroviruses in hematopoietic cells, which led to the identification of several oncogenes. Each of the models has been dominated by a single CIS. Hematopoietic malignancies are generally characterized by a less diverse pattern of mutated genes than solid tumors. Solid tumors, especially carcinomas, display high complexity and variability of karyotypic alterations, which seriously complicates the search for oncogenes and tumor suppressor genes in these tumors [94]. The complexity probably results from the complex homeostatic control of cells in solid tissues which must be overcome. Thus, finding appropriate model tumors which would enable a large-scale search for solid tumor oncogenes and tumor suppressor genes has been a challenge. The models based on MAV retroviruses are among the successful ones.

MAV–1 and MAV–2 (myeloblastosis-associated viruses 1 and 2) are simple slowly oncogenic retroviruses classified with ALVs. They differ by their env gene (and, consequently, by host specificity), MAV–1 belonging to subgroup A, MAV–2 to subgroup B. They are highly homologous to other ALVs except one region: the U3 part of LTR, the region containing proviral enhancer/promoter and playing crucial role in insertional activation (see Sections 4.1 and 4.2). It is this MAV-specific region that confers very distinct oncogenic potential on MAV retroviruses [67]. When MAV–1 or MAV–2 are injected into 12 day old embryos or newly hatched chicks they induce nephroblastomas with efficiency 70–100% depending on chicken strain. Multiple tumors become visible on kidney of an infected animal as early as 2 months after infection and they rapidly grow into massive size [68, 140, 158, 160].

The chicken nephroblastomas are embryonic tumors derived from nephrogenic blastema cells which persist in the newborn kidney for several days after hatching. They consist mainly of undifferentiated mesenchyme and aberrant differentiating epithelial renal elements reminiscent of developing kidney. They are comparable to human nephroblastomas—Wilms' tumors both by their histology and presumed target cells of tumorigenesis [7, 11, 56, 62].

First attempts to identify CISs in chicken nephroblastomas had only limited success. Two groups reported two different loci hit by proviral integration in chicken nephroblastomas—*Ha–ras* [160] and *nov/ccn3* [68]. Although deregulation of these genes in the chicken or mammalian cells have been proven to result in a transformed phenotype, each of them have been found hit only in a single case and therefore they could not be classified as CIS.

Restriction mapping of integration sites in clonal nephroblastomas using Southern blot analyses revealed unexpectedly complex pattern from which existence of CISs could not be definitely inferred [114]. Three interpretations have been put forward. Firstly, there may be a major CIS but it represents region too large to be disclosed by restriction mapping on the background of several further integrated proviruses in each nephroblastoma clone. Secondly, there might be no major CIS but rather numerous less frequent ones. Thirdly, activating proviruses could carry

deletions similarly to the situation in B-lymphomas which would invalidate the used mapping strategy.

Later on, an improvement of the techniques for isolation and analysis of provirus-flanking sequences enabled more comprehensive survey of VISs ("Viral Integration Sites") in nephroblastomas. As described in Section 4.2, defective proviruses are very potent mutagens and deletion in the provirus may be a prerequisite for oncogene activation. Frequent occurrence of defective proviruses in nephroblastomas has been documented before [115, 140, 160]. Therefore, to start with, molecular cloning of VISs harboring defective proviruses in six randomly chosen nephroblastomas have been carried out, leading to identification of the insertionally activated gene *twist*. Subsequent screening of a panel of chicken nephroblastomas have shown that the *twist* gene is the first true CIS in these tumors—it was hit in approximately 4% of analyzed tumors [107].

These results strongly favored the hypothesis of numerous rather infrequent CISs in the chicken nephroblastomas. To prove it, high throughput screening of VIS in hundreds of nephroblastoma samples had to be performed. To achieve a maximum coverage of VISs, two independent and overlapping methods for the VIS identification were employed [108]: inverse PCR (iPCR) and LTR-based rapid amplification of 3′cDNA ends (LTR RACE). The iPCR technique provided genomic sequences flanking the integrated provirus. The technique was optimized to the level when flanking sequences of all clonally integrated proviruses could be gel-isolated after single PCR reaction; routine VISs recovery was 90–100% in each tumor. LTR–RACE technique provided sequences of mRNA species containing provirus LTR at its 5′end, including hybrid mRNAs generated by readthrough transcription or by transcription starting in the promoter of a defective provirus (see Section 4.2 and Fig. 4.1). LTR–RACE analysis covered up to 30% of VISs. The lower yield achieved by LTR–RACE is compensated by the fact that found VISs were the most relevant ones—those where high expression of cellular sequences was driven by an integrated provirus. At the same time, the structure of LTR–RACE products revealed details of the mechanism of host gene activation. Importantly, both techniques detected only insertions present within a substantial proportion of tumor cells as opposed to currently used sequencing strategies that record also oncogenically unselected integrations present even in a single cell.

To date more than 1100 unique VISs have been identified in 187 MAV-induced nephroblastoma clones [108] and (Pajer, unpublished). Out of them 69 candidate tumor-related loci (denominated NALs, "nephroblastoma associated loci") have been selected according to any of the following three criteria:

(a) two and more VISs in a close genomic distance; the distance limit was set arbitrary to 20 kb;
(b) two and more VISs within one annotated gene;
(c) VISs from the tumors harboring only a single integrated provirus as demonstrated both by Southern blot hybridization and iPCR; to date only three such samples have been found.

It should be emphasized that not all candidate NALs necessarily have to represent cancer-related genes, especially those hit by a provirus in two samples only. It has been shown that using the sole criterion of multiple insertions into the same locus in retroviral oncogene screens may lead to high proportion of false positives in consequence of retroviral integration preferences [163]. The most relevant loci (those hit in at least four tumors) were the *foxP1,* the *plag1*, the *twist*, and the *c–Ha–ras*. *FoxP1* (forkhead box P1) encodes a widely expressed transcription factor important for heart, lung and lymphocyte development; it was shown to function either as an oncogene or a tumor suppressor gene depending on cell type [76]. *Plag1* codes for a zinc finger transcription factor in humans expressed only in embryonal tissues; its inapropriate activation is responsible for induction of several types of human tumors [151]. *Twist* encodes a basic helix-loop-helix transcription factor regulating epithelial–mesenchymal transition during embryogenesis and metastasis and supporting survival of cancer cells [2, 22].

Analysis of produced mRNAs and provirus alignment in the selected CISs showed that the molecular mechanism of activation was different in each case (see also Fig. 4.3). All *twist*-activating proviruses were integrated in the same transcriptional orientation in a narrow region just upstream of the *twist* initiation codon and most of them carried various deletions or rearrangements, alike the ALV proviruses activating the *c–myc* gene in chicken B-lymphomas (Section 4.2, Fig. 4.1c–f). Very high expression of *twist* mRNA was driven mostly by 3′LTR promoter. Neither normal adult kidney nor nephroblastomas without MAV insertion in the *twist* locus did contain detectable amount of *twist* mRNA while moderate level of *twist* mRNA was detected in the embryonic kidney.

Similarly, in tumors with activated *plag1* gene very high levels of the *plag1* mRNA were found. This RNA was generated by a readthrough mechanism followed by splicing between the MAV gag donor site and the *plag1* second exon acceptor site, much like during *c–myb* activation in short latency lymphomas (Section 4.3, Fig. 4.1b). The hybrid mRNA coded for full-length *plag1* protein.

A strikingly different situation was found in the case of *foxP1*. Proviruses were inserted downstream of the first or second *foxP1* coding exons mostly in the same orientation as the gene's transcription. No gross change in overall expression of *foxP1* mRNA has been observed in any tumor harboring affected allele of the gene. Hybrid MAV/*foxP1* mRNAs, however, could be detected and analyzed using RT–PCR; they coded for the N-terminally truncated FoxP1 proteins. Further analysis of the effect of *foxP1* truncation was complicated by the existence of alternative starting points, alternative exons and alternative polyA signals in *foxP1* gene. Moreover, it has been observed that both the amount and subcellular localization of *FoxP1* protein was highly variable in individual nephroblastoma samples, including those not harboring viral integration in *foxP1* gene.

As to the *c–Ha–ras* activation, provirus insertions have been found in 5′untranslated region of the gene in the same transcriptional orientation. Both hybrid MAV/*c–Ha–ras* mRNAs and normal *c–Ha–ras* mRNAs were observed; the hybrid mRNAs coded for normal c–Ha–ras protein but they were missing 5′untranslated region present in normal *c–Ha–ras* mRNA. The overall level of *c–Ha–ras* mRNA

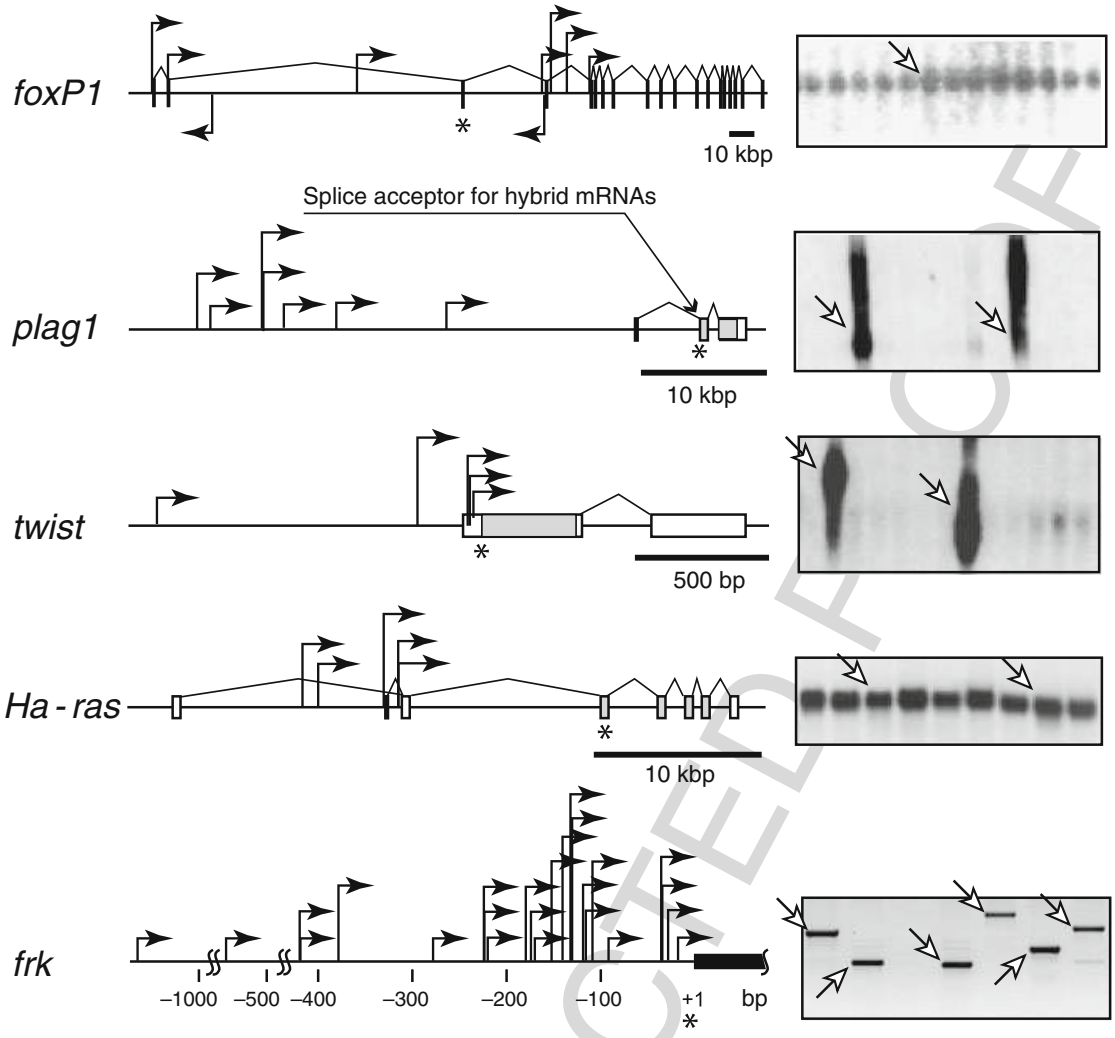


Fig. 4.3 Genomic organization, distribution of insertion sites and mRNA expression in selected CISs. The CISs were identified in chicken nephroblastomas (*foxP1, plag1, twist and Ha–ras*) or lung sarcomas (*frk*). Sites and orientations of individual proviral insertions are indicated by *arrows*. Positions of translation initiation are marked by asterisks. Panels to the right show Northern blot hybridization profile of a few representative tumor samples, except for the *frk* locus where RT–PCR profile is shown. The samples harboring a provirus in the given locus are marked by *arrows*

was the same in nephroblastomas with or without insertion in *c–Ha–ras* locus and in the normal kidney. No point mutations, the archetypal mode of *Ha–ras* activation in human and mouse tumors, have been found (Pajer, unpublished). How such "activation", which affects neither mRNA level nor the structure of the produced protein can participate in nephroblastoma induction is not clear. We suggest that deleting 5′untranslated region might result in the increased translation of the hybrid mRNA. Alternatively, *Ha–ras* activation could be important only during the induction stage in nephroblastema cells, the target cells of oncogenesis—for instance, if basal level of *Ha–ras* expression is low in these cells (hit and run hypothesis).

The results draw a highly complex picture of virus-induced mutations in nephroblastomas. Out of 69 candidate NALs only four were hit in a reasonable

percentage of samples: the *foxP1* (in about 6% of tumors), the *plag1* (ca 5% of tumors), the *twist* (ca 3% of tumors) and the *c–Ha–ras* (ca 2% of tumors). All the identified NALs together would explain for induction of ca 130 out of 197 nephroblastomas (ca 65%). Evidently, many alternative cancer-related genes (or combinations of them) may be deregulated by the provirus integration in nephroblastomas.

The special feature of the chicken nephroblastoma model is the high efficiency of infection of nephrogenic blastema cells [114, 160]. Each cell acquires numerous integrations thus generating enormous amount of combinations of insertional mutations. These facts establish the chicken nephroblastoma as a perfect model for the study of oncogene/tumor suppressor gene cooperation during multistep tumor induction.

4.6 Extension of Spectrum of MAV-Induced Tumors By Local Homeostasis Perturbation: Lung Sarcomas, Liver Carcinomas and the Industasis Phenomenon

In about 7% of MAV–2-infected animals late lung sarcomas or liver carcinomas developed in addition to nephroblastomas. These tumors were independent primary clonal outgrowths, not metastases originated from the nephroblastoma as shown by distinct patterns of proviral integrations. While nephroblastomas and liver carcinomas appeared as nodules clearly separated from normal tissue, the lung sarcomas were highly invasive: one tumor clone was frequently disseminated into several foci in both lungs and in some animals these foci constituted majority of lung tissue [106]. Presence of considerable amount of tumor cells could even be detected by PCR in macroscopically normal lung tissue samples (Pajer, unpublished). Both types of late tumors would represent new interesting models for oncogene screening if it were not for the crucial drawback: they were too rare to provide enough samples for thorough analysis.

The obstacle of low penetrance was cleared by discovery that injection of MAV–2 producing chicken cells instead of virions intravenously into the embryos or newly hatched chickens changes formerly rare tumors into frequent ones. In the case of lung tumors the process of tumorigenesis was also markedly accelerated—the lung tumors appeared in most animals within 1–2 months after cell injection, prior to nephroblastomas and liver tumors, the latency of which did not change. The cell-assisted tumors did not differ from the rare ones neither by gross morphology nor by histology. Preventing the division of injected MAV–2 producing cells by treatment with mitomycin did not abolish their ability to promote tumor formation. Consistently, the cell-assisted lung and liver tumors were shown to be of the host origin and not to contain detectable amounts of the injected cells [106] and (Pajer et al., in preparation).

Surprisingly, the same tumor-promoting effect was observed when animals were injected first with MAV–2 virus at embryonal day 12 and then with uninfected

chicken embryonal fibroblasts 9 days later, shortly after hatching. At that time all tissues of the animal including lungs and liver were already fully infected, suggesting that injected cells do not function only as reservoirs spreading virus in the animal. This conclusion was further supported by the fact that injected cells affected neither the spectrum of infected tissues nor the level of their infection. However, when injections were carried out in the opposite order, i.e. the virus was injected 9 days after the cells, no tumor promotion was observed [106]. This fact suggests that the tumor-promoting effect of the injected cells persists not longer than several days.

To follow the fate of the injected cells, [35S]methionine-labelled embryonal fibroblasts were used and monitored in individual embryonic tissues both by total tissue radioactivity and by paraffin-embedded section autoradiography. The experiments showed a prompt passage of the injected cells from blood into the tissues where they settled as individual stray cells. 20 h after injection the cells were detected in all analyzed embryonic tissues including brain; the highest levels were found in the liver, kidney, and lungs while the residual levels in the blood were very low.

Based on these observations the concept of industasis was proposed—a promotion of fully malignant phenotype of an incipient tumor cell by a stray cell through a disruption of local homeostasis; for details see [106]. It was suggested that the phenomenon of industasis might be the underlying cause of many human multiple primary tumors when a stray cell released by an advanced tumor promotes tumor formation from another, mutated and potentially malignant cell that was kept under control by tissue homeostasis.

Besides its potential impact on our understanding of mechanisms of oncogenesis in general, the industasis fenomenon advanced the model of MAV–2 insertional mutagenesis as it enabled a large scale analysis of lung and liver tumors. Both rare virus-only-induced and frequent cell-assisted tumors have been searched for CISs using the same methods as described in Section 4.5, i.e. iPCR and LTR RACE. Quite dissimilar pattern of CISs was revealed in the lung versus liver tumors; importantly, however, there was no difference between the rare and the cell-promoted tumors [106].

In contrast to nephroblastomas, where numerous low frequency CISs have been found, lung sarcomas were dominated by a single one. In more than 95% lung tumors a provirus insertion was observed in the gene *frk/rak* [106] and (Pajer et al., unpublished); no other CIS was identified. The *frk* (fyn-related-kinase gene) encodes a tyrosine kinase expressed in humans predominantly in epithelial tissues [19]. It was shown to act as a tumor suppressor gene that negatively regulates PI3K/Akt pathway through stabilisation of PTEN, another tumor suppressor and negative regulator of PI3K [13, 167]. Surprisingly, *frk*-knockout mice displayed very mild phenotype and no increase in tumor incidence [21]. The model of chicken lung sarcomas has demonstrated for the first time the oncogenic capability of *frk* overexpression, previously being only suspected [60].

The provirus/*frk* arrangement was much alike provirus/*c–myc* arrangement in B-lymphomas: defective proviruses had integrated in the same transcriptional orientation into a very narrow area within the promoter/5′UTR region in front

of the gene's coding sequences. Most defects were internal deletions comprising the enhancer element in 5′part of *gag* (Pecenka et al., in preparation). High levels of MAV–2/*frk* hybrid mRNA initiating in the proviral 3′LTR have been detected in all samples carrying provirus insertion in the *frk* locus. The protein coding sequence was not afflicted in any sample. No *frk* expression was detectable neither in tumors with unaffected *frk* locus nor in the non-tumor lung tissue [106].

In liver tumors, the pattern of CISs was more variable and different from the pattern in kidney or lung tumors (Pajer et al., in preparation) and (Pecenka et al., in preparation). The vast majority of liver tumors carried MAV–2 provirus inserted within one of four genes: *c–Ha–ras* (39% of tumors), *c–erbB/EGFR* (32%), *c–ron/c–stk/c–sea/MST1R* (11%) or *c–met/HGFR* (7%); no other CIS was recorded. The last three genes encode receptor tyrosine kinases expressed in many types of epithelial cells and implicated in several carcinoma classes in humans [47, 104, 156]. Interestingly, they all control, among others, signaling pathways converging on the *c–Ha–ras,* the most frequently activated gene in chicken liver tumors. That suggests that activation of the *c–Ha–ras* signaling is the pivotal disturbance in all MAV–2 induced liver tumors and that insertions into four alternative genes might be just four different ways how to affect *c–Ha–ras* and pathways downstream of it. Consistently, simultaneous activation of two or more of these genes in the same tumor was never observed, though, on mere statistical basis, *c–Ha–ras* should be hit in about one third of tumors with *c–erbB* activation and vice versa. It must be pointed out, however, that the screen did not comprise genes connected with progression steps since all samples analyzed were early stages of tumor development, according to terminology of [73] classified as preneoplastic nodules or adenomas (Pajer et al., in preparation).

Similarly to nephroblastomas, *c–Ha–ras* gene was activated by provirus insertion into the gene's 5′untranslated region; the *c–Ha–ras* coding sequences were preserved. Alike in nephroblastomas, no point mutations have been found. Unlike nephroblastomas, however, insertions resulted in substantial overexpression of hybrid MAV–2/*c–Ha–ras* mRNA initiating in the proviral 3′–most LTR. All activating proviruses have been heavily rearranged including large deletions, duplications and inversions (Pecenka et al., in preparation).

The mode of *c–erbB* activation resembled the situation in ALV-induced erythroblastosis (Section 4.4): all proviruses had integrated into the same intron in the middle of the gene. That resulted in high expression of hybrid mRNA encoding the N-terminally truncated c–erbB/EGFR fused to the first 6 AA of *gag* gene. Contrary to erythroblastosis the *c–erbB*-activating proviruses had been heavily rearranged much alike the *c–Ha–ras*-activating ones. Due to the provirus rearrangement, variant mRNA containing *env* mini-exon described in erythroblastosis model could not be generated.

The *c–ron* activation paralleled the *c–erbB* activation in all aspects described above including the truncation of gene's 5′part that codes for the receptor's extracellular ligand-binding domain. Alike in case of c–erbB, N-terminal truncation confers constitutive kinase activity on c-ron protein [82].

The details of *c–met* activation could not be conclusively established since RNA from these tumors was not available. Based on DNA analysis it seems that the integrated defective proviruses could drive overexpression of the full-length protein—they were located upstream of the gene's initiation codon (in the promoter or 5′untranslated region), in the same transcriptional orientation.

4.7 Evidence of Multistage Cancerogenesis in Chicken Models

Tumors are end products of successive evolution through selection of progressively more malignant cell subclones which is reflected by complex pattern of accumulated mutations of oncogenes and tumor suppressor genes ("cancer genes"). Such complex pattern is typical for human tumors and has also been found in mouse tumor models. In majority of chicken models the mutational pattern seems to be quite simple (see previous Sections); that may, however, be just due to the absence of thorough high throughput analyses. Nevertheless, the multistep nature of oncogenesis in chicken models can be evidenced by studying progressive evolution of tumors.

The examination of early stages of B-lymphoma development (Section 4.1) showed that cells with a proviral integration in the *c–myc* locus appear early after ALV infection. The chicken Bursa of Fabricius is compartmentalized into ca 10 000 follicles which are colonized by B-cell progenitors during late embryonic development [103]. One month after ALV infection of young chicks, up to 100 hyperplastic transformed follicles are observable in bursa. Each of them represents a clonal outgrowth of a different cell with activated *c–myc* [9, 48]. The hyperplastic state is not a consequence of accelerated proliferation but of blockage of lymphoblast differentiation and emigration from bursa [12]. Nearly all transformed follicles disappear during bursa involution, only one or a few (if any) of them progress to the stage of B-lymphoma. Distinct and (to a lesser extent) overgrown clones of cells with activated *c–myc* gene have also been observed 1 month after infection in tissues where primary tumors do not develop (spleen, bone marrow). Thus, *c–myc* activation is the early step in oncogenesis, but it is not sufficient for B-lymphoma formation; further steps must follow. Indeed, when chicken B-lymphomas were searched for additional CISs, the gene *bic* coding for regulatory RNA miR–155 was found insertionally activated in a significant proportion of chicken B-lymphomas, most frequently in metastases together with activated *c–myc* [23, 38]. Undoubtedly, an extensive analysis of chicken B-lymphomas would discover further activated genes cooperating with *c–myc* during tumor development much like in mouse MLV-induced T-lymphomas where many such genes have already been found [91, 142].

A very similar situation was observed during short latency lymphoma development (Section 4.3). Here the *c–myb* gene was found activated in hyperplastic follicles in bursa early after infection. Contrary to long latency B-lymphomas, emigration of the transformed lymphoblasts was not fully blocked. Cells harboring activated *c–myb* also infiltrated neighbouring follicles (without compromising the structure of bursa) so that substantial proportion of bursa was formed by clusters

of hyperplastic follicles [98, 117]. While bursa mostly stayed in a hyperplastic state without tumors, the lymphoblasts released from it pervaded liver (plus bone marrow and other organs) and formed B-lymphomas there. Even in chicks which eventually developed no lymphoma, the bursa of Fabricius was mostly overgrown by oligoclonal population of cells with activated *c–myb*. Obviously, lymphoblast clones with activated *c–myb* had a strong advantage when populating bursa, but absolute majority of them did not progress to form tumors. Thus, again, *c–myb* activation is only one event in a multistep process of oncogenesis. The additional steps, however, seem mostly not to involve insertional activation since a high proportion of short latency B-lymphomas contain only a single ALV provirus—the one in the *c–myb* locus [71].

Recently, a frequent insertional activation of telomerase reverse transcriptase gene (*TERT*) was discovered in short latency B-lymphomas [165]. Activation of *TERT* and *c–myb,* however, seem to represent alternative, not cooperating events since tumor clones carrying activated *TERT* did not contain activated *c–myb* and vice versa.

Also the development of erythroblastosis (Section 4.4) shows the same signs of multistep tumorigenesis: multiple distinct clones of cells with insertionally activated *c–erbB* originate and overgrow in hematopoetic tissues but only one or few clones in spleen or bone marrow go on to form a tumor [48]. So far, no search for *c–erbB*-cooperating genes in chicken erythroblastosis was performed. Similarly, no systematic attempts to identify cooperating oncogenes in MAV-induced tumors have been carried out yet.

It must be emphasized that some steps in a multistage retroviral oncogenesis may have no relation to provirus insertions. Spontaneous mutations and chromosomal rearrangements, the major mechanisms operating during human tumorigenesis, may be involved, especially when we realize the huge reservoir of premalignant cells created by the first step of tumorigenesis, i.e. by the insertional activation of an oncogene. However, to discover genes affected by stochastic processes is incomparably more difficult than to pick up genes tagged by integrated proviruses and relevant technologies are very labor-intensive and costly. Consequently, no data are available about possible involvement of point mutations or chromosomal rearrangements in the models described above. Limited searches for Ha–ras mutations in chicken nephroblastomas and liver tumors (Pajer, unpublished) provided negative results.

4.8 The Future of Models Based on Chicken Retroviruses

During the last years the effort to identify genes implicated in cancer induction and development (“cancer genes”) has intensified. High throughput forward genetic screens based on insertional mutagenesis have been employed and transposons as a new versatile insertional mutagens have been introduced [37, 150]. Supremely informative data started to be acquired by genome-wide analyses of human tumors [29, 144]. All this effort is driven by hopes that with detailed knowledge of “genetic

landscape of cancer" it will be possible to design new therapeutics targeted directly against the crucial mutated cancer genes. Such therapeutics are expected to be more efficient and have less detrimental side effects compared to the conventional ones that are directed nondiscriminatingly against all rapidly dividing cells.

Accumulating results show that the number of genes that can be implicated even in one particular type of cancer is formidable albeit in a single tumor only several of them are mutated. Thus, except for a few commonest oncogenes and tumor suppressor genes, each potential cancer gene is mutated only in a very small proportion of the tumors. To be able to design personalized therapy in every cancer case enormous numbers of targeted therapeutics would have to be developed. Fortunately, majority of the discovered cancer genes fall into a handful of signaling pathways which are, to a large extent, shared between different tumor classes. With little exceptions only one member of the pathway is mutated in one tumor reaffirming the notion that deregulation of a particular pathways, not of a particular genes, is what matters in tumorigenesis. Consequently, it should suffice to develop therapeutics targeted against the selected key components of the pathways, not against all their members.

The genome-wide analyses of human tumors must comprise huge sample collections due to large heterogeneity of tumors, large number of alternative cancer genes and presence, in each tumor, of immense numbers of fortuitous passenger mutations not related to cancerogenesis. The results have to be processed using statistical analyses and the screens provide rather suspect than undoubted cancer genes [162]. Moreover, these analyses reveal only static endpoint-set of genetic alterations not showing the history and dynamics of tumor development; also the fundamental importance of stromal cells and tumor microenvironment is missed out. All what was said underscores the imperative that extensive cancer gene screens were followed by experiments—to validate the gene's causative role in cancerogenesis, show in what step they act, establish their grouping into individual signaling pathways, elucidate their mutual cooperation and study the role of microenvironment. Since all this is impossible to pursue clinically, animal models are indispensable.

While mouse models of human cancerogenesis are now generally accepted [44] a question remains if and to what extent the models based on evolutionary more distant species like chickens are relevant to humans. The answer is certainly positive as majority of organ systems are shared between chickens, humans and mice, their tissues have the same organisational principles and contain the same types of intercommunicating cells. The signal transduction pathways including those involved in tumorigenesis are conserved over much larger evolutionary distances, beyond the phylum chordata. Nevertheless, different elements of the pathway or even different pathways are often preferentially mutated in different species during oncogenic transformation of a particular cell type. Such predisposition suggests that differences in regulatory networks between different vertebrate species do exist or, at least, that in some pathways different elements in different species are the rate-limiting ones. Well known and frequently encountered difference is a varying pattern of tissue specific expression of functionally redundant genes in different species. An illustrative example is induction of retinoblastoma in humans versus mice. In humans, retinoblastoma originates after *Rb1* gene has been homozygously inactivated in a

single cell of developing retina. To induce retinoblastoma in mice, both *Rb1* and its relative *p107* must be inactivated. Different sensitivity to *Rb1* inactivation results from the simple fact that both *Rb1* and *p107* are expressed in developing mouse retina and can compensate for each other's inactivation while only *Rb1* is expressed during retinal development in humans [36].

As a matter of fact, certain cancerogenesis-influencing differences in genetic background exist even between strains of the same species and between individuals from non-inbred population, which includes human population as well. All the more significant differences must be expected between members of different animal classes like between humans and chickens. Faithful modeling of a specific human cancer is challenging even in mouse system [44, 125], therefore chickens do not represent a suitable system for this type of modeling. When we use the term "animal model" in the next paragraphs, we mean nothing more than productive and convenient instruments to search for cancer genes and to study their function and cooperation.

Cancer gene screens in animals, especially those based on mouse retroviruses, have been very successful in gathering data applicable for outlining genetic landscape of cancer. For a long time, however, the investigation has been limited to hematopoetic and mammary tumors due to a limited tissue tropism of the used mouse retroviruses (MLV and MMTV). To progress further it was essential to extend the range of analyzed tumor types and, ideally, to create models corresponding to the most relevant human tumors. This objective was fulfilled with great success by the development of universal transposon-based technology [26, 73, 141].

The tropism of avian retroviruses is not as strictly limited as is the case for MLV and MMTV. In addition to neoplasias described above (B-lymphoma, erythroblastosis, nephroblastoma, lung sarcoma and hepatocarcinoma) simple avian retroviruses are capable of inducing many other tumor types, e.g. myeloid leukosis, histiocytic sarcoma, angiosarcoma, fibrosarcoma, renal carcinoma, mesothelioma, and pancreatic adenocarcinoma [15, 110, 113, 121, 137, 153]. Though the penetrance of these malignancies is mostly very low, the fundamental ability to transform a wide range of target cells is apparent. To carry out cancer gene screens, however, the frequency of rare tumor classes have to be increased. Below we suggest two ways how to achieve that; no doubt further possibilities can be conceived or will come up over time.

Nongenetic non-cell-autonomous factors that influence tissue homeostasis, such as wound healing and local inflammation, have long been known to play an important role in both experimental and human oncogenesis [69, 74, 86, 154]. The industasis phenomenon (Section 4.6) illustrates that these factors may as well operate during oncogenesis through insertional mutagenesis. Classical tumor promoters like phorbol esters fall into the same sort of factors. Many different tumor promoters have been described, some of them tissue specific [87]. Hence, the penetrance of the rare retrovirus-induced cancer classes can be increased by tumor promoters either through their tissue-specific delivery or by using the tissue-specific ones. The important point is that assistance of tumor promoters does not bring any additional genetic changes into the cells and that tumors originate from the same pool

of dormant mutagenized cells that give rise to the rare tumors when no promotion is applied (see Section 4.6). In addition to the increased penetrance, tumor promoters could also reduce latency thus facilitating study of tumor progression.

The oncogenic spectrum of simple retroviruses is primarily determined by the promoter/enhancer sequences in provirus LTR—see, for example, the properties of MAV/RAV recombinants [67]. Consistently, enhancer mutations have been shown to change oncogenic spectrum of the retrovirus [39, 139]. This is not surprising since tissue-specific activity of promoter/enhancer not only determines the tissue where the virus will replicate best, but also the cell type where the integrated provirus will affect neighboring genes most vigorously. Manipulating the enhancer is thus another way how to shift retroviral oncogenic spectrum. For example, segments of known tissue specific promoters/enhancers may be inserted into viral U3 region. Definitely, problems with the proper enhancer design can be anticipated, resulting from our lack of understanding how combined enhancer segments interact with each other. These problems, however, may be surmounted by generating many alternative variants (e.g. by shotgun approach when assembling the LTR or by employing random mutagenesis) and selecting the successful variants in cell culture or animals. Similar approaches have already been utilized, though with varying success, see e.g. [5, 42, 43, 52, 88].

To understand fully the relevance of results obtained in animal cancer gene screens, we must contemplate also another factor, independent of evolutionary distance, responsible for the incomplete overlap of cancer genes mutated in experimental model versus human cancer. It is the dissimilar molecular mechanisms of cancer gene mutation/deregulation. In humans the cancer genes are affected largely by point mutations and chromosome rearrangements while in animal screens they are altered by provirus/transposon insertions. This introduces bias to the spectrum of mutated cancer genes because (i) oncogenic changes of a cancer gene introduced by point mutations or chromosome rearrangement cannot always be accurately simulated by proviral insertion and (ii) the starting distributions (prior to oncogenic selection) of chromosome rearrangements, point mutations and proviral insertions, respectively, throughout the genome are uneven and differ from each other.

(i) Evidently, the effect of human cancer gene amplification (with resulting overexpression) or, conversely, the effect of gene knockout/deactivation can be reproduced by provirus/transposon insertion. Also the effect of chromosomal translocation can mostly be simulated by insertional mutagenesis, see c–myc activation by translocation in human Burkitt lymphomas [32] as well as by insertion in chicken B-lymphomas or mouse T-lymphomas [30, 126]. In these cases the full-length cancer gene is overexpressed as a result of being placed under the control of strong non-authentic promoter/enhancer. Frequently, however, the translocated human cancer gene is not only overexpressed, but also truncated and fused to another gene which results in production of a chimeric protein [94]. Effects of such translocations can be mimicked by insertional mutagenesis only in those instances when the overexpression and/or truncation alone, without a fusion, is sufficient to activate

the gene's oncogenic potential or when the cancer gene's fusion partner can be substituted by retrovirus/transposon-derived sequences like in the case of *bcr–abl* and *gag–abl* fusions [161]. Point mutations, the typical mode of oncogenic activation of many cancer genes in humans, cannot, obviously, be inflicted by provirus/transposon insertion. In spite of it, many of these genes repeatedly emerge in insertional mutagenesis screens. It is so because they can also be activated by other mechanisms than point mutation, e.g. by amplification/overexpression as exemplified by *Ha–ras* gene [83, 108, 120, 145, 148] or by truncation as exemplified by *B–raf* gene [26, 90]. Certainly, cancer genes mutated by different mechanisms may have not fully equivalent oncogenic properties; that, however, is irrelevant since we are discussing forward genetic screening, not modeling of a particular human tumor.

(ii) Even if particular gene can be, in principle, oncogenically activated by different mechanisms (for example by chromosomal translocation in humans and by provirus/transposon insertion in experimental animals) the relative frequency of such events may be very different. The primary choice of chromosomal recombination sites as well as selection of integration sites are far from random. Particularly the frequency of breakage and recombination in different parts of human genome is highly non-uniform as manifested by the existence of hot spots of recombination and chromosomal fragile sites. This bias is caused by specific local DNA and chromatin features that are unrelated to function of the residing genes. Different features are relevant for chromosomal recombination versus provirus/transposon insertion [17, 18, 166, 169]. In consequence, when particular pathway is to be deregulated, different elements of the pathway will be preferentially mutated in humans versus experimental animals.

The propensity to mutate preferentially different genes exists even between different insertional mutagens. For instance, *H–ras* gene is frequently activated by retroviral proviruses [83, 108, 145, 148] but never by Sleeping Beauty transposon [141]. Even closely related retroviruses that induce identical subclass of tumors but differ in their enhancer sequences may activate very different sets of genes [65]. Most plausible explanation of this phenomenon is that there exists significant (yet so far overlooked) interaction of inserted enhancer sequences with the host regulatory sequences at the site of insertion; this interaction may be productive in case of one enhancer and counterproductive in case of another one.

There are some practical advantages of chicken system worth to mention. For example manipulating and monitoring chicken embryo is much more feasible and convenient. A great advantage of chicken system is that gene activation is preferentially carried out by the promoter insertion mechanism, which is characterized by a narrow region of insertions in front of the activated gene. This eliminates dilemma frequently encountered in mouse models: which of the genes in the CIS locus is the cancer-related one, see [53, 131].

One disadvantage of chicken models must also be mentioned here. It is the absence of the huge variety of strains predisposed to specific tumor types due to the presence of transgenic or knockout cancer gene alleles, which are available

in mouse system. Insertional mutagenesis in genetically engineered animals is a powerful method for studying cancer genes cooperation. The first generation of genetically engineered mice had, however, certain week sides: the oncogenic mutation was expressed already during embryogenesis and in all cells of the target tissue, which is very different condition compare to spontaneous tumor origination and may significantly influence course and outcome of tumorigenesis [44, 130, 171]. To preclude this drawback, new generations of genetically engineered mice are being constructed in which the oncogene/tumor suppressor gene is activated/deleted later during the lifetime and only in sporadic cells [66, 89, 95, 157].

The technology for generating transgenic chickens is now at hand [51, 70]; however, the range of genetically modified chicken strains can never come close to the variety available in mouse system. Fortunately, production of predisposed genetically modified chickens can be conveniently substituted by using in vivo gene transfer mediated by retroviral or transposon-based vectors that carry an activated oncogene, inhibitory RNA directed against tumor suppressor gene or dominant-negative form of tumor suppressor gene. This technique, moreover, does not comprise the drawbacks of the transgenic approach mentioned above. Combination of insertional mutagenesis and in vivo gene transfer would enable not only to investigate cancer genes cooperation in chicken models but also to redirect the retrovirus tumorigenic activity into selected tissue either by tissue-specific vector delivery or through the tissue-specific capacity of the delivered cancer gene. Efficacy of in vivo gene transfer strategies has already been proven in mouse system [57, 64, 73, 78].

The conclusion of the above discussion is that no cancer gene screening model fully reproduces corresponding human cancer. Multiple models including multiple model animals must be combined to cover all players involved in the disease. Comparing and integrating results obtained in different models and in genome-wide screening of human tumors may prove to be especially beneficial for delineating the genetic landscape of cancer. Thus, we are convinced that models based on chicken retroviruses can still considerably contribute to our knowledge of genes, pathways and networks implicated in cancerogenesis.

AQ2

References

1. Adhya, S., & Gottesman, M. (1982). Promoter occlusion: Transcription through a promoter may inhibit its activity. *Cell*, *29*, 939–944.
2. Ansieau, S., Bastid, J., Doreau, A., Morel, A. P., Bouchet, B. P., Thomas, C., et al. (2008). Induction of EMT by twist proteins as a collateral effect of tumor-promoting inactivation of premature senescence. *Cancer Cell*, *14*, 79–89.
3. Arrigo, S., Yun, M., & Beemon, K. (1987). cis-acting regulatory elements within gag genes of avian retroviruses. *Molecular and Cellular Biology*, *7*, 388–397.
4. Baba, T. W., & Humphries, E. H. (1985). Formation of a transformed follicle is necessary but not sufficient for development of an avian leukosis virus-induced lymphoma. *Proceedings of the National Academy of Sciences of the USA*, *82*, 213–216.
5. Bahrami, S., Duch, M., & Pedersen, F. S. (2004). Change of tropism of SL3-2 murine leukemia virus, using random mutational libraries. *Journal of Virology*, *78*, 9343–9351.

6. Ball, J. K., Diggelmann, H., Dekaban, G. A., Grossi, G. F., Semmler, R., Waight, P. A., et al. (1988). Alterations in the U3 region of the long terminal repeat of an infectious thymotropic type B retrovirus. *Journal of Virology*, *62*, 2985–2993.
7. Beard, J. W., Chabot, J. F., Beard, D., Heine, U., & Houts, G. E. (1976). Renal neoplastic response to leukosis virus strains BAI A (avian myeloblastosis virus) and MC29. *Cancer Research*, *36*, 339–353.
8. Beemon, K. L., O'Reilly, M. M., Smith, M. R., Smith, R. E., Dunkel, I., & Hayward, W. S. (1997). Rapid induction of B-cell lymphomas by avian leukosis virus. *Leukemia*, *11*(Suppl 3), 179–182.
9. Bird, K. J., Semus, H. L., & Ruddell, A. (1999). Resistance to avian leukosis virus lymphomagenesis occurs subsequent to proviral c-myc integration. *Oncogene*, *18*, 201–209.
10. Boerkoel, C. F., & Kung, H. J. (1992). 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. *Journal of Virology*, *66*, 4814–4823.
11. Boni-Schnetzler, M., Boni, J., Ferdinand, F. J., & Franklin, R. M. (1985). Developmental and molecular aspects of nephroblastomas induced by avian myeloblastosis-associated virus 2-O. *Journal of Virology*, *55*, 213–222.
12. Brandvold, K. A., Ewert, D. L., Kent, S. C., Neiman, P., & Ruddell, A. (2001). Blocked B cell differentiation and emigration support the early growth of Myc-induced lymphomas. *Oncogene*, *20*, 3226–3234.
13. Brauer, P. M., & Tyner, A. L. (2009). RAKing in AKT: A tumor suppressor function for the intracellular tyrosine kinase FRK. *Cell Cycle*, *8*, 2728–2732.
14. Broussard, D. R., Mertz, J. A., Lozano, M., & Dudley, J. P. (2002). Selection for c-myc integration sites in polyclonal T-cell lymphomas. *Journal of Virology*, *76*, 2087–2099.
15. Brown, D. W., Blais, B. P., & Robinson, H. L. (1988). Long terminal repeat (LTR) sequences, env, and a region near the 5′ LTR influence the pathogenic potential of recombinants between Rous-associated virus types 0 and 1. *Journal of Virology*, *62*, 3431–3437.
16. Burmester, B. R., & Waters, N. F. (1955). The role of the infected egg in the transmission of visceral lymphomatosis. *Poultry Science*, *34*, 1415–1429.
17. Bushman, F., Lewinski, M., Ciuffi, A., Barr, S., Leipzig, J., Hannenhalli, S., et al. (2005). Genome-wide analysis of retroviral DNA integration. *Nature Reviews Microbiology*, *3*, 848–858.
18. Buttel, I., Fechter, A., & Schwab, M. (2004). Common fragile sites and cancer: Targeted cloning by insertional mutagenesis. *Annals of the New York Academy Science*, *1028*, 14–27.
19. Cance, W. G., Craven, R. J., Bergman, M., Xu, L., Alitalo, K., & Liu, E. T. (1994). Rak, a novel nuclear tyrosine kinase expressed in epithelial cells. *Cell Growth and Differentiation*, *5*, 1347–1355.
20. Ceci, J. D., Patriotis, C. P., Tsatsanis, C., Makris, A. M., Kovatch, R., Swing, D. A., et al. (1997). Tpl-2 is an oncogenic kinase that is activated by carboxy-terminal truncation. *Genes and Development*, *11*, 688–700.
21. Chandrasekharan, S., Qiu, T. H., Alkharouf, N., Brantley, K., Mitchell, J. B., & Liu, E. T. (2002). Characterization of mice deficient in the Src family nonreceptor tyrosine kinase Frk/rak. *Molecular and Cellular Biology*, *22*, 5235–5247.
22. Cheng, G. Z., Zhang, W., & Wang, L. H. (2008). Regulation of cancer cell survival, migration, and invasion by Twist: AKT2 comes to interplay. *Cancer Research*, *68*, 957–960.
23. Clurman, B. E., & Hayward, W. S. (1989). Multiple proto-oncogene activations in avian leukosis virus-induced lymphomas: Evidence for stage-specific events. *Molecular and Cellular Biology*, *9*, 2657–2664.
24. Cochrane, A. W., McNally, M. T., & Mouland, A. J. (2006). The retrovirus RNA trafficking granule: From birth to maturity. *Retrovirology*, *3*, 18.
25. Coffin, J. M., Hughes, S. H., & Varmus, H. E. (1997). *Retroviruses*. New York: Cold Spring Harbor Laboratory Press.

26. Collier, L. S., Carlson, C. M., Ravimohan, S., Dupuy, A. J., & Largaespada, D. A. (2005). Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. *Nature*, *436*, 272–276.
27. Cooper, G. M., & Neiman, P. E. (1981). Two distinct candidate transforming genes of lymphoid leukosis virus-induced neoplasms. *Nature*, *292*, 857–8.
28. Cooper, M. D., Payne, L. N., Dent, P. B., Burmester, B. R., & Good, R. A. (1968). Pathogenesis of avian lymphoid leukosis. I. Histogenesis. *Journal of the National Cancer Institute*, *41*, 373–378.
29. Copeland, N. G., & Jenkins, N. A. (2009). Deciphering the genetic landscape of cancer - from genes to pathways. *Trends in Genetics*, *25*, 455–462.
30. Corcoran, L. M., Adams, J. M., Dunn, A. R., & Cory, S. (1984). Murine T lymphomas in which the cellular myc oncogene has been activated by retroviral insertion. *Cell*, *37*, 113–122.
31. Crittenden, L. B., Hayward, W. S., Hanafusa, H., & Fadly, A. M. (1980). Induction of neoplasms by subgroup E recombinants of exogenous and endogenous avian retroviruses (Rous-associated virus type 60). *Journal of Virology*, *33*, 915–919.
32. Croce, C. M., & Nowell, P. C. (1985). Molecular basis of human B cell neoplasia. *Blood*, *65*, 1–7.
33. Cullen, B. R., Lomedico, P. T., & Ju, G. (1984). Transcriptional interference in avian retroviruses–implications for the promoter insertion model of leukaemogenesis. *Nature*, *307*, 241–245.
34. Dabrowska, M. J., Dybkaer, K., Johnsen, H. E., Wang, B., Wabl, M., & Pedersen, F. S. (2009). Loss of MicroRNA targets in the 3′ untranslated region as a mechanism of retroviral insertional activation of growth factor independence 1. *Journal of Virology*, *83*, 8051–8061.
35. de Laat, W., & Grosveld, F. (2003). Spatial organization of gene expression: The active chromatin hub. *Chromosome Research*, *11*, 447–459.
36. Donovan, S. L., Schweers, B., Martins, R., Johnson, D., & Dyer, M. A. (2006). Compensation by tumor suppressor genes during retinal development in mice and humans. *BMC Biology*, *4*, 14.
37. Dupuy, A. J., Jenkins, N. A., & Copeland, N. G. (2006). Sleeping beauty: A novel cancer gene discovery tool. *Human Molecular Genetics*, *15*(Spec No 1), R75–R79.
38. Eis, P. S., Tam, W., Sun, L., Chadburn, A., Li, Z., Gomez, M. F., et al. (2005). Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proceedings of the National Academy of Sciences of the USA*, *102*, 3627–3632.
39. Ejegod, D., Sorensen, K. D., Mossbrugger, I., Quintanilla-Martinez, L., Schmidt, J., & Pedersen, F. S. (2009). Control of pathogenicity and disease specificity of a T-lymphomagenic gammaretrovirus by E-box motifs but not by an overlapping glucocorticoid response element. *Journal of Virology*, *83*, 336–346.
40. El Kaderi, B., Medler, S., Raghunayakula, S., & Ansari, A. (2009). Gene looping is conferred by activator-dependent interaction of transcription initiation and termination machineries. *Journal of Biological Chemistry*, *284*, 25015–25025.
41. Ellerman, V., & Bang, O. (1908). Experimentelle Leukämie bei Hühnern. *Zentralblatt fur Bakteriologi, Parasitenkunde, Infektionskrinkheiten und Hygiene Abteilung Originale*, *46*, 595–609.
42. Ethelberg, S., Sorensen, A. B., Schmidt, J., Luz, A., & Pedersen, F. S. (1997). An SL3-3 murine leukemia virus enhancer variant more pathogenic than the wild type obtained by assisted molecular evolution in vivo. *Journal of Virology*, *71*, 9796–9799.
43. Feuer, G., & Fan, H. (1990). Substitution of murine transthyretin (prealbumin) regulatory sequences into the Moloney murine leukemia virus long terminal repeat yields infectious virus with altered biological properties. *Journal of Virology*, *64*, 6130–6140.

44. Frese, K. K., & Tuveson, D. A. (2007). Maximizing mouse cancer models. *Nature Reviews Cancer*, *7*, 645–658.
45. Fung, Y. K., Fadly, A. M., Crittenden, L. B., & Kung, H. J. (1981). On the mechanism of retrovirus-induced avian lymphoid leukosis: Deletion and integration of the proviruses. *Proceedings of the National Academy of Sciences of the USA*, *78*, 3418–3422.
46. Fung, Y. K., Lewis, W. G., Crittenden, L. B., & Kung, H. J. (1983). Activation of the cellular oncogene c-erbB by LTR insertion: Molecular basis for induction of erythroblastosis by avian leukosis virus. *Cell*, *33*, 357–368.
47. Gentile, A., Trusolino, L., & Comoglio, P. M. (2008). The Met tyrosine kinase receptor in development and cancer. *Cancer Metastasis Reviews*, *27*, 85–94.
48. Gong, M., Semus, H. L., Bird, K. J., Stramer, B. J., & Ruddell, A. (1998). Differential selection of cells with proviral c-myc and c-erbB integrations after avian leukosis virus infection. *Journal of Virology*, *72*, 5517–5525.
49. Goodenow, M. M., & Hayward, W. S. (1987). 5′ long terminal repeats of myc-associated proviruses appear structurally intact but are functionally impaired in tumors induced by avian leukosis viruses. *Journal of Virology*, *61*, 2489–2498.
50. Goodwin, R. G., Rottman, F. M., Callaghan, T., Kung, H. J., Maroney, P. A., & Nilsen, T. W. (1986). c-erbB activation in avian leukosis virus-induced erythroblastosis: Multiple epidermal growth factor receptor mRNAs are generated by alternative RNA processing. *Molecular and Cellular Biology*, *6*, 3128–3133.
51. Han, J. Y. (2009). Germ cells and transgenesis in chickens. *Comparative Immunology, Microbiology and Infectious Diseases*, *32*, 61–80.
52. Hanecak, R., Pattengale, P. K., & Fan, H. (1991). Deletion of a GC-rich region flanking the enhancer element within the long terminal repeat sequences alters the disease specificity of Moloney murine leukemia virus. *Journal of Virology*, *65*, 5357–5363.
53. Hanlon, L., Barr, N. I., Blyth, K., Stewart, M., Haviernik, P., Wolff, L., et al. (2003). Long-range effects of retroviral insertion on c-myb: Overexpression may be obscured by silencing during tumor growth in vitro. *Journal of Virology*, *77*, 1059–1068.
54. Hayward, W. S., Neel, B. G., & Astrin, S. M. (1981a). Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis. *Nature*, *290*, 475–480.
55. Hayward, W. S., Neel, B. G., Fang, J., Robinson, H. L., & Astrin, S. M. (1981b). Avian lymphoid leukosis is correlated with the appearance of discrete new RNAs containing viral and cellular genetic information. *Hematology and Blood Transfusion*, *26*, 439–444.
56. Heine, U., De The, G., Ishiguro, H., Sommer, J. R., Beard, D., & Beard, J. W. (1962). Multiplicity of cell response to the BAI strain A (myeloblastosis) avian tumor virus. II. Nephroblastoma (Wilms' tumor): Ultrastructure. *Journal of the National Cancer Institute*, *29*, 41–105.
57. Hemann, M. T., Fridman, J. S., Zilfou, J. T., Hernando, E., Paddison, P. J., Cordon-Cardo, C., et al. (2003). An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. *Nature Genetics*, *33*, 396–400.
58. Herman, S. A., & Coffin, J. M. (1986). Differential transcription from the long terminal repeats of integrated avian leukosis virus DNA. *Journal of Virology*, *60*, 497–505.
59. Herman, S. A., & Coffin, J. M. (1987). Efficient packaging of readthrough RNA in ALV: Implications for oncogene transduction. *Science*, *236*, 845–848.
60. Hosoya, N., Qiao, Y., Hangaishi, A., Wang, L., Nannya, Y., Sanada, M., et al. (2005). Identification of a SRC-like tyrosine kinase gene, FRK, fused with ETV6 in a patient with acute myelogenous leukemia carrying a t(6;12)(q21;p13) translocation. *Genes Chromosomes and Cancer*, *42*, 269–279.
61. Huppi, K., Siwarski, D., Shaughnessy, J. D., Jr., & Mushinski, J. F. (1993). Co-amplification of c-myc/pvt-1 in immortalized mouse B-lymphocytic cell lines results in a novel pvt-1/AJ-1 transcript. *International Journal of Cancer*, *53*, 493–498.
62. Ishiguro, H., Beard, D., Sommer, J. R., Heine, U., De The, G., & Beard, J. W. (1962). Multiplicity of cell response to the BAI strain A (myeloblastosis) avian tumor virus. II.

Nephroblastoma (Wilms' tumor): Gross and microscopic pathology. *Journal of the National Cancer Institute*, *29*, 1–17.

63. Jiang, W., Kanter, M. R., Dunkel, I., Ramsay, R. G., Beemon, K. L., & Hayward, W. S. (1997). Minimal truncation of the c-myb gene product in rapid-onset B-cell lymphoma. *Journal of Virology*, *71*, 6526–6533.
64. Johansson, F. K., Brodd, J., Eklof, C., Ferletta, M., Hesselager, G., Tiger, C. F., et al. (2004). Identification of candidate cancer-causing genes in mouse brain tumors by retroviral tagging. *Proceedings of the National Academy of Sciences of the USA*, *101*, 11334–11337.
65. Johnson, C., Lobelle-Rich, P. A., Puetter, A., & Levy, L. S. (2005). Substitution of feline leukemia virus long terminal repeat sequences into murine leukemia virus alters the pattern of insertional activation and identifies new common insertion sites. *Journal of Virology*, *79*, 57–66.
66. Johnson, L., Mercer, K., Greenbaum, D., Bronson, R. T., Crowley, D., Tuveson, D. A., et al. (2001). Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature*, *410*, 1111–1116.
67. Joliot, V., Khelifi, C., Wyers, M., Dambrine, G., Lasserre, F., Lemercier, P., et al. (1996). The noncoding and surface envelope coding sequences of myeloblastosis-associated virus are respectively responsible for nephroblastoma development and renal hyperplasia. *Journal of Virology*, *70*, 2576–2580.
68. Joliot, V., Martinerie, C., Dambrine, G., Plassiart, G., Brisac, M., Crochet, J., et al. (1992). Proviral rearrangements and overexpression of a new cellular gene (nov) in myeloblastosis-associated virus type 1-induced nephroblastomas. *Molecular and Cellular Biology*, *12*, 10–21.
69. Jones, F. S., & Rous, P. (1914). On the cause of the localization of secondary tumors at points of injury. *J Exp Med*, *20*, 404–412.
70. Kalina, J., Senigl, F., Micakova, A., Mucksova, J., Blazkova, J., Yan, H., et al. (2007). Retrovirus-mediated in vitro gene transfer into chicken male germ line cells. *Reproduction*, *134*, 445–453.
71. Kanter, M. R., Smith, R. E., & Hayward, W. S. (1988). Rapid induction of B-cell lymphomas: Insertional activation of c-myb by avian leukosis virus. *Journal of Virology*, *62*, 1423–1432.
72. Kawai, S., & Yamamoto, T. (1970). Isolation of different kinds of non-virus producing chick cells transformed by Schmidt-Ruppin strain (subgroup A) of Rous sarcoma virus. *Japanese Journal of Experimental Medicine*, *40*, 243–256.
73. Keng, V. W., Villanueva, A., Chiang, D. Y., Dupuy, A. J., Ryan, B. J., Matise, I., et al. (2009). A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma. *Nature Biotechnology*, *27*, 264–274.
74. Kluwe, J., Mencin, A., & Schwabe, R. F. (2009). Toll-like receptors, wound healing, and carcinogenesis. *Journal of Molecular Medicine*, *87*, 125–138.
75. Kool, J., & Berns, A. (2009). High-throughput insertional mutagenesis screens in mice to identify oncogenic networks. *Nature Reviews Cancer*, *9*, 389–399.
76. Koon, H. B., Ippolito, G. C., Banham, A. H., & Tucker, P. W. (2007). FOXP1: A potential therapeutic target in cancer. *Expert Opinion on Therapeutic Targets*, *11*, 955–965.
77. Lax, I., Kris, R., Sasson, I., Ullrich, A., Hayman, M. J., Beug, H., et al. (1985). Activation of c-erbB in avian leukosis virus-induced erythroblastosis leads to the expression of a truncated EGF receptor kinase. *EMBO Journal*, *4*, 3179–3182.
78. Lewis, B. C., Klimstra, D. S., Socci, N. D., Xu, S., Koutcher, J. A., & Varmus, H. E. (2005). The absence of p53 promotes metastasis in a novel somatic mouse model for hepatocellular carcinoma. *Molecular and Cellular Biology*, *25*, 1228–1237.
79. Li, T., & Zhang, J. (2002). Intramolecular recombinations of Moloney murine leukemia virus occur during minus-strand DNA synthesis. *Journal of Virology*, *76*, 9614–9623.
80. Lin, Q., Chen, Q., Lin, L., Smith, S., & Zhou, J. (2007). Promoter targeting sequence mediates enhancer interference in the Drosophila embryo. *Proceedings of the National Academy of Sciences of the USA*, *104*, 3237–3242.

81. Linial, M., & Groudine, M. (1985). Transcription of three c-myc exons is enhanced in chicken bursal lymphoma cell lines. *Proceedings of the National Academy of Sciences of the USA*, *82*, 53–57.
82. Lu, Y., Yao, H. P., & Wang, M. H. (2007). Multiple variants of the RON receptor tyrosine kinase: Biochemical properties, tumorigenic activities, and potential drug targets. *Cancer Letters*, *257*, 157–164.
83. Lund, A. H., Turner, G., Trubetskoy, A., Verhoeven, E., Wientjens, E., Hulsman, D., et al. (2002). Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice. *Nature Genetics*, *32*, 160–165.
84. Maciolek, N. L., & McNally, M. T. (2007). Serine/arginine-rich proteins contribute to negative regulator of splicing element-stimulated polyadenylation in rous sarcoma virus. *Journal of Virology*, *81*, 11208–11217.
85. Maihle, N. J., Raines, M. A., Flickinger, T. W., & Kung, H. J. (1988). Proviral insertional activation of c-erbB: Differential processing of the protein products arising from two alternate transcripts. *Molecular and Cellular Biology*, *8*, 4868–4876.
86. Mantovani, A., Allavena, P., Sica, A., & Balkwill, F. (2008). Cancer-related inflammation. *Nature*, *454*, 436–444.
87. Marks, F., Schwartz, M., & Furstenberg, G. (1995). Promotion and cocarcinogenesis. In J. C. Arcos (ed.), *Chemical induction of cancer: Modulation and combination effects. An inventory of the many factors which influence carcinogenesis* (pp. 123–184.). Boston: Burkhauser.
88. McGee-Estrada, K., Palmarini, M., Hallwirth, C., & Fan, H. (2005). A Moloney murine leukemia virus driven by the Jaagsiekte sheep retrovirus enhancers shows enhanced specificity for infectivity in lung epithelial cells. *Virus Genes*, *31*, 257–263.
89. Meuwissen, R., Linn, S. C., van der Valk, M., Mooi, W. J., & Berns, A. (2001). Mouse model for lung tumorigenesis through Cre/lox controlled sporadic activation of the K-Ras oncogene. *Oncogene*, *20*, 6551–6558.
90. Michaloglou, C., Vredeveld, L. C., Mooi, W. J., & Peeper, D. S. (2008). BRAF(E600) in benign and malignant human tumours. *Oncogene*, *27*, 877–895.
91. Mikkers, H., Allen, J., Knipscheer, P., Romeijn, L., Hart, A., Vink, E., et al. (2002). High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. *Nature Genetics*, *32*, 153–159.
92. Miles, B. D., & Robinson, H. L. (1985). High-frequency transduction of c-erbB in avian leukosis virus-induced erythroblastosis. *Journal of Virology*, *54*, 295–303.
93. Miller, J. T., & Stoltzfus, C. M. (1992). Two distant upstream regions containing cis-acting signals regulating splicing facilitate 3′-end processing of avian sarcoma virus RNA. *Journal of Virology*, *66*, 4242–4251.
94. Mitelman, F., Johansson, B., & Mertens, F. (2007). The impact of translocations and gene fusions on cancer causation. *Nature Reviews Cancer*, *7*, 233–245.
95. Muzumdar, M. D., Luo, L., & Zong, H. (2007). Modeling sporadic loss of heterozygosity in mice by using mosaic analysis with double markers (MADM. *Proceedings of the National Academy of Sciences of the USA*, *104*, 4495–4500.
96. Neel, B. G., Hayward, W. S., Robinson, H. L., Fang, J., & Astrin, S. M. (1981). Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: Oncogenesis by promoter insertion. *Cell*, *23*, 323–334.
97. Neil, J. C., Fulton, R., Rigby, M., & Stewart, M. (1991). Feline leukaemia virus: Generation of pathogenic and oncogenic variants. *Current Topics in Microbiology and Immunology*, *171*, 67–93.
98. Neiman, P. E., Grbic, J. J., Polony, T. S., Kimmel, R., Bowers, S. J., Delrow, J., et al. (2003). Functional genomic analysis reveals distinct neoplastic phenotypes associated with c-myb mutation in the bursa of fabricius. *Oncogene*, *22*, 1073–1086.
99. Nilsen, T. W., Maroney, P. A., Goodwin, R. G., Rottman, F. M., Crittenden, L. B., Raines, M. A., et al. (1985). c-erbB activation in ALV-induced erythroblastosis: Novel RNA processing

and promoter insertion result in expression of an amino-truncated EGF receptor. *Cell*, *41*, 719–726.

100. Noori-Daloii, M. R., Swift, R. A., Kung, H. J., Crittenden, L. B., & Witter, R. L. (1981). Specific integration of REV proviruses in avian bursal lymphomas. *Nature*, *294*, 574–576.
101. Norton, P. A., & Coffin, J. M. (1987). Characterization of Rous sarcoma virus sequences essential for viral gene expression. *Journal of Virology*, *61*, 1171–1179.
102. Nusse, R., Theunissen, H., Wagenaar, E., Rijsewijk, F., Gennissen, A., Otte, A., et al. (1990). The Wnt-1 (int-1) oncogene promoter and its mechanism of activation by insertion of proviral DNA of the mouse mammary tumor virus. *Molecular and Cellular Biology*, *10*, 4170–4179.
103. Olah, I., & Glick, B. (1978). The number and size of the follicular epithelium (FE) and follicles in the bursa of Fabricius. *Poultry Science*, *57*, 1445–1450.
104. Olayioye, M. A., Neve, R. M., Lane, H. A., & Hynes, N. E. (2000). The ErbB signaling network: Receptor heterodimerization in development and cancer. *EMBO Journal*, *19*, 3159–3167.
105. O'Sullivan, C. T., Polony, T. S., Paca, R. E., & Beemon, K. L. (2002). Rous sarcoma virus negative regulator of splicing selectively suppresses SRC mRNA splicing and promotes polyadenylation. *Virology*, *302*, 405–412.
106. Pajer, P., Karafiat, V., Pecenka, V., Prukova, D., Dudlova, J., Plachy, J., et al. (2009). Industasis, a promotion of tumor formation by nontumorigenic stray cells. *Cancer Research*, *69*, 4605–4612.
107. Pajer, P., Pecenka, V., Karafiat, V., Kralova, J., Horejsi, Z., & Dvorak, M. (2003). The twist gene is a common target of retroviral integration and transcriptional deregulation in experimental nephroblastoma. *Oncogene*, *22*, 665–673.
108. Pajer, P., Pecenka, V., Kralova, J., Karafiat, V., Prukova, D., Zemanova, Z., et al. (2006). Identification of potential human oncogenes by mapping the common viral integration sites in avian nephroblastoma. *Cancer Research*, *66*, 78–86.
109. Parghi, S. S., Brandvold, K. A., Bowers, S. J., Neiman, P. E., & Ruddell, A. (2004). Reduced Myc overexpression and normal B-cell differentiation mediate resistance to avian leukosis virus lymphomagenesis. *Oncogene*, *23*, 4413–4421.
110. Payne, L. N. (1992). Developments in avian leukosis research. *Leukemia*, *6*(Suppl 3), 150S–152S.
111. Payne, G. S., Bishop, J. M., & Varmus, H. E. (1982). Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas. *Nature*, *295*, 209–214.
112. Payne, G. S., Courtneidge, S. A., Crittenden, L. B., Fadly, A. M., Bishop, J. M., & Varmus, H. E. (1981). Analysis of avian leukosis virus DNA and RNA in bursal tumours: Viral gene expression is not required for maintenance of the tumor state. *Cell*, *23*, 311–322.
113. Payne, L. N., & Fadly, A. M. (1997). Leukosis/sarcoma group. In H. J. Calnek C. W. Barnes, C. W. Beard, L. R. McDougald, & Y. M. Saif (ed.), *Diseases of poultry* (10th ed.). Ames, IA: Iowa State University Press.
114. Pecenka, V., Dvorak, M., Karafiat, V., Sloncova, E., Hlozanek, I., Travnicek, M., et al. (1988a). Avian nephroblastomas induced by a retrovirus (MAV-2) lacking oncogene. II. Search for common sites of proviral integration in tumour DNA. *Folia Biologica (Praha)*, *34*, 147–169.
115. Pecenka, V., Dvorak, M., Karafiat, V., & Travnicek, M. (1988b). Avian nephroblastomas induced by a retrovirus (MAV-2) lacking oncogene. III. Presence of defective MAV-2 proviruses in tumour DNA. *Folia Biologica (Praha)*, *34*, 215–232.
116. Perkins, K. J., Lusic, M., Mitar, I., Giacca, M., & Proudfoot, N. J. (2008). Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. *Molecular Cell*, *29*, 56–68.
117. Pizer, E. S., Baba, T. W., & Humphries, E. H. (1992). Activation of the c-myb locus is insufficient for the rapid induction of disseminated avian B-cell lymphoma. *Journal of Virology*, *66*, 512–523.

AQ3

118. Pizer, E., & Humphries, E. H. (1989). RAV-1 insertional mutagenesis: Disruption of the c-myb locus and development of avian B-cell lymphomas. *Journal of Virology*, *63*, 1630–1640.
119. Polony, T. S., Bowers, S. J., Neiman, P. E., & Beemon, K. L. (2003). Silent point mutation in an avian retrovirus RNA processing element promotes c-myb-associated short-latency lymphomas. *Journal of Virology*, *77*, 9378–9387.
120. Pulciani, S., Santos, E., Long, L. K., Sorrentino, V., & Barbacid, M. (1985). ras gene Amplification and malignant transformation. *Molecular and Cellular Biology*, *5*, 2836–2841.
121. Purchase, H. G., Okazaki, W., Vogt, P. K., Hanafusa, H., Burmester, B. R., & Crittenden, L. B. (1977). Oncogenicity of avian leukosis viruses of different subgroups and of mutants of sarcoma viruses. *Infection and Immunity*, *15*, 423–428.
122. Quintrell, N., Hughes, S. H., Varmus, H. E., & Bishop, J. M. (1980). Structure of viral DNA and RNA in mammalian cells infected with avian sarcoma virus. *Journal of Molecular Biology*, *143*, 363–393.
123. Raines, M. A., Lewis, W. G., Crittenden, L. B., & Kung, H. J. (1985). c-erbB activation in avian leukosis virus-induced erythroblastosis: Clustered integration sites and the arrangement of provirus in the c-erbB alleles. *Proceedings of the National Academy of Sciences of the USA*, *82*, 2287–2291.
124. Raines, M. A., Maihle, N. J., Moscovici, C., Crittenden, L., & Kung, H. J. (1988). Mechanism of c-erbB transduction: Newly released transducing viruses retain poly(A) tracts of erbB transcripts and encode C-terminally intact erbB proteins. *Journal of Virology*, *62*, 2437–2443.
125. Rangarajan, A., & Weinberg, R. A. (2003). Opinion: Comparative biology of mouse versus human cells: Modelling human cancer in mice. *Nature Reviews Cancer*, *3*, 952–959.
126. Robinson, H. L., & Gagnon, G. C. (1986). Patterns of proviral insertion and deletion in avian leukosis virus-induced lymphomas. *Journal of Virology*, *57*, 28–36.
127. Robinson, H. L., Miles, B. D., Catalano, D. E., Briles, W. E., & Crittenden, L. B. (1985). Susceptibility to erbB-induced erythroblastosis is a dominant trait of 151 chickens. *Journal of Virology*, *55*, 617–622.
128. Roebroek, A. J., Schalken, J. A., Onnekink, C., Bloemers, H. P., & Van de Ven, W. J. (1987). Structure of the feline c-fes/fps proto-oncogene: Genesis of a retroviral oncogene. *Journal of Virology*, *61*, 2009–2016.
129. Ryden, T. A., de Mars, M., & Beemon, K. (1993). Mutation of the C/EBP binding sites in the Rous sarcoma virus long terminal repeat and gag enhancers. *Journal of Virology*, *67*, 2862–2870.
130. Sage, J., Miller, A. L., Perez-Mancera, P. A., Wysocki, J. M., & Jacks, T. (2003). Acute mutation of retinoblastoma gene function is sufficient for cell cycle re-entry. *Nature*, *424*, 223–228.
131. Sauvageau, M., Miller, M., Lemieux, S., Lessard, J., Hebert, J., & Sauvageau, G. (2008). Quantitative expression profiling guided by common retroviral insertion sites reveals novel and cell type specific cancer genes in leukemia. *Blood*, *111*, 790–799.
132. Schubach, W. H., & Horvath, G. (1988). Alternate structures and stabilities of c-myc RNA in a bursal lymphoma cell line. *Nucleic Acids Research*, *16*, 11171–11186.
133. Selten, G., Cuypers, H. T., & Berns, A. (1985). Proviral activation of the putative oncogene Pim-1 in MuLV induced T-cell lymphomas. *EMBO Journal*, *4*, 1793–1798.
134. Shaughnessy, J. D., Jr., Owens, J. D., Jr., Wiener, F., Hilbert, D. M., Huppi, K., Potter, M., et al. (1993). Retroviral enhancer insertion 5′ of c-myc in two translocation-negative mouse plasmacytomas upregulates c-myc expression to different extents. *Oncogene*, *8*, 3111–3121.
135. Shtivelman, E., Zakut, R., & Canaani, E. (1984). Frequent generation of nonrescuable reorganized Moloney murine sarcoma viral genomes. *Proceedings of the National Academy of Sciences of the USA*, *81*, 294–298.

136. Simon, M. C., Neckameyer, W. S., Hayward, W. S., & Smith, R. E. (1987). Genetic determinants of neoplastic diseases induced by a subgroup F avian leukosis virus. *Journal of Virology*, *61*, 1203–1212.
137. Simon, M. C., Smith, R. E., & Hayward, W. S. (1984). Mechanisms of oncogenesis by subgroup F avian leukosis viruses. *Journal of Virology*, *52*, 1–8.
138. Smith, M. R., Smith, R. E., Dunkel, I., Hou, V., Beemon, K. L., & Hayward, W. S. (1997). Genetic determinant of rapid-onset B-cell lymphoma by avian leukosis virus. *Journal of Virology*, *71*, 6534–6540.
139. Sorensen, K. D., Quintanilla-Martinez, L., Kunder, S., Schmidt, J., & Pedersen, F. S. (2004). Mutation of all Runx (AML1/core) sites in the enhancer of T-lymphomagenic SL3-3 murine leukemia virus unmasks a significant potential for myeloid leukemia induction and favors enhancer evolution toward induction of other disease patterns. *Journal of Virology*, *78*, 13216–13231.
140. Soret, J., Dambrine, G., & Perbal, B. (1989). Induction of nephroblastoma by myeloblastosis-associated virus type 1: State of proviral DNAs in tumor cells. *Journal of Virology*, *63*, 1803–1807.
141. Starr, T. K., Allaei, R., Silverstein, K. A., Staggs, R. A., Sarver, A. L., Bergemann, T. L., et al. (2009). A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. *Science*, *323*, 1747–1750.
142. Stewart, M., Mackay, N., Hanlon, L., Blyth, K., Scobie, L., Cameron, E., et al. (2007). Insertional mutagenesis reveals progression genes and checkpoints in MYC/Runx2 lymphomas. *Cancer Research*, *67*, 5126–5133.
143. Stoye, J. P., Moroni, C., & Coffin, J. M. (1991). Virological events leading to spontaneous AKR thymomas. *Journal of Virology*, *65*, 1273–1285.
144. Stratton, M. R., Campbell, P. J., & Futreal, P. A. (2009). The cancer genome. *Nature*, *458*, 719–724.
145. Suzuki, T., Shen, H., Akagi, K., Morse, H. C., Malley, J. D., Naiman, D. Q., et al. (2002). New genes involved in cancer identified by retroviral tagging. *Nature Genetics*, *32*, 166–174.
146. Swanstrom, R., Parker, R. C., Varmus, H. E., & Bishop, J. M. (1983). Transduction of a cellular oncogene: The genesis of Rous sarcoma virus. *Proceedings of the National Academy of Sciences of the USA*, *80*, 2519–2523.
147. Swift, R. A., Boerkoel, C., Ridgway, A., Fujita, D. J., Dodgson, J. B., & Kung, H. J. (1987). B-lymphoma induction by reticuloendotheliosis virus: Characterization of a mutated chicken syncytial virus provirus involved in c-myc activation. *Journal of Virology*, *61*, 2084–2090.
148. Theodorou, V., Kimm, M. A., Boer, M., Wessels, L., Theelen, W., Jonkers, J., et al. (2007). MMTV insertional mutagenesis identifies genes, gene families and pathways involved in mammary cancer. *Nature Genetics*, *39*, 759–769.
149. Tsichlis, P. N., & Coffin, J. M. (1980). Recombinants between endogenous and exogenous avian tumor viruses: Role of the C region and other portions of the genome in the control of replication and transformation. *Journal of Virology*, *33*, 238–249.
150. Uren, A. G., Kool, J., Berns, A., & van Lohuizen, M. (2005). Retroviral insertional mutagenesis: Past, present and future. *Oncogene*, *24*, 7656–7672.
151. Van Dyck, F., Declercq, J., Braem, C. V., & Van de Ven, W. J. (2007). PLAG1, the prototype of the PLAG gene family: Versatility in tumour development (review). *International Journal of Oncology*, *30*, 765–774.
152. van Lohuizen, M., Breuer, M., & Berns, A. (1989). N-myc is frequently activated by proviral insertion in MuLV-induced T cell lymphomas. *EMBO Journal*, *8*, 133–136.
153. Venugopal, K. (1999). Avian leukosis virus subgroup J: A rapidly evolving group of oncogenic retroviruses. *Research in Veterinary Science*, *67*, 113–119.
AQ4 154. Virchow, R. (1863). Aetiologie der neoplastischen Geschwülste/Pathogenie der neoplastischen Geschwülste. Die krankhaften Geschwülste Verlag von August von (57–101 pp) Hirschwald: Berlin.

155. Vogt, P. K. (1971). Spontaneous segregation of nontransforming viruses from cloned sarcoma viruses. *Virology*, *46*, 939–946.
156. Wagh, P. K., Peace, B. E., & Waltz, S. E. (2008). Met-related receptor tyrosine kinase Ron in tumor growth and metastasis. *Advance in Cancer Research*, *100*, 1–33.
157. Wang, W., Warren, M., & Bradley, A. (2007). Induced mitotic recombination of p53 in vivo. *Proceedings of the National Academy of Sciences of the USA*, *104*, 4501–4505.
158. Watts, S. L., & Smith, R. E. (1980). Pathology of chickens infected with avian nephoblastoma virus MAV-2(N. *Infection and Immunity*, *27*, 501–512.
159. West, A. G., & Fraser, P. (2005). Remote control of gene transcription. *Human Molecular Genetics*, *14*(Spec No 1), R101–R111.
160. Westaway, D., Papkoff, J., Moscovici, C., & Varmus, H. E. (1986). Identification of a provirally activated c-Ha-ras oncogene in an avian nephroblastoma via a novel procedure: cDNA cloning of a chimaeric viral-host transcript. *EMBO Journal*, *5*, 301–309.
161. Wong, S., & Witte, O. N. (2004). The BCR-ABL story: Bench to bedside and back. *Annual Review of Immunology*, *22*, 247–306.
162. Wood, L. D., Parsons, D. W., Jones, S., Lin, J., Sjoblom, T., Leary, R. J., et al. (2007). The genomic landscapes of human breast and colorectal cancers. *Science*, *318*, 1108–1113.
163. Wu, X., Luke, B. T., & Burgess, S. M. (2006). Redefining the common insertion site. *Virology*, *344*, 292–295.
164. Yamamoto, T., de Crombrugghe, B., & Pastan, I. (1980). Identification of a functional promoter in the long terminal repeat of Rous sarcoma virus. *Cell*, *22*, 787–797.
165. Yang, F., Xian, R. R., Li, Y., Polony, T. S., & Beemon, K. L. (2007). Telomerase reverse transcriptase expression elevated by avian leukosis virus integration in B cell lymphomas. *Proceedings of the National Academy of Sciences of the USA*, *104*, 18952–18957.
166. Yant, S. R., Wu, X., Huang, Y., Garrison, B., Burgess, S. M., & Kay, M. A. (2005). High-resolution genome-wide mapping of transposon integration in mammals. *Molecular and Cellular Biology*, *25*, 2085–2094.
167. Yim, E. K., Peng, G., Dai, H., Hu, R., Li, K., Lu, Y., et al. (2009). Rak functions as a tumor suppressor by regulating PTEN protein stability and function. *Cancer Cell*, *15*, 304–314.
168. Zhang, J., & Ma, Y. (2001). Evidence for retroviral intramolecular recombinations. *Journal of Virology*, *75*, 6348–6358.
169. Zhang, Y., & Rowley, J. D. (2006). Chromatin structural elements and chromosomal translocations in leukemia. *DNA Repair (Amst)*, *5*, 1282–1297.
170. Zhou, J., & Levine, M. (1999). A novel cis-regulatory element, the PTS, mediates an anti-insulator activity in the Drosophila embryo. *Cell*, *99*, 567–575.
171. Zhu, Y., Ghosh, P., Charnay, P., Burns, D. K., & Parada, L. F. (2002). Neurofibromas in NF1: Schwann cell origin and role of tumor environment. *Science*, *296*, 920–922.

Distinct cancer genes are activated in different liver tumor types in a new chicken model of liver cancerogenesis

Pečenka V.[1], Karafiát V.[1], Kašparová P.[2], Dudlová J.[3], Dvořák M.[1] and Pajer P.[1*]

[1] Institute of Molecular Genetics AS CR, v.v.i., Vídeňská 1083, 142 20 Prague, Czech Republic

[2] Fingerland´s Department of Pathology, School of Medicine Charles University, Sokolská 581, 500 05 Hradec Králové, Czech Republic

[3] Chambon s.r.o., Evropská 176/16, 160 00, Prague 6, Czech Republic

[*]**Contact information:** correspondence should be addressed to Peter Pajer, e-mail: petr.pajer@img.cas.cz, phone: +420-296443391, fax: +420-296443586

Keywords: hepatocellular carcinoma, cholangiocarcinoma, hemangiosarcoma, insertional mutagenesis, industasis, MAV-2

Abbreviations: MAV-2, myeloblastosis-associated virus-2; VIS, virus integration site; CIS, common integration site; RT-PCR, reverse transcription-polymerase chain reaction; iPCR, inverse polymerase chain reaction; LTR, long-terminal repeat; EGFR, epidermal growth factor receptor; MST1R, macrophage stimulating-1 receptor; HGFR, hepatocyte growth factor receptor; HCC, hepatocellular carcinoma; ICC, intrahepatic cholangiocarcinoma; HHS, hepatic hemangiosarcoma; HBV, hepatitis B virus; HCV, hepatitis C virus;

Financial support: This work was supported by grants AV0Z50520514 from Academy of Sciences of the Czech Republic and 301/09/1727 from the Grant Agency of the Czech Republic to M.D.

Abstract

We have developed a new model of liver cancer based on insertional mutagenesis by the chicken retrovirus MAV-2 and tumor promotion through local homeostasis disruption by stray non-tumor cells. Liver lesions induced in this model developed within 2–3 months and closely resembled three types of human liver cancer: hepatic hemangiosarcomas, hepatocellular carcinomas, and intrahepatic cholangiocarcinomas. In over 85% of these tumors one of four genes (*HRAS*, *ERBB1/EGFR*, *RON/MST1R* and *MET/HGFR*) were insertionally activated. There was a clear correlation between the mutated gene and tumor type – *HRAS* was activated in hemangiocarcinomas, *EGFR* in hepatocarcinomas, and *MST1R* in cholangiocarcinomas or combined hepato-cholangiocarcinomas. The histopathology of lesions with activated *HGFR* was not performed due to their very small size. The genes were activated through either overexpression of the wt allele (*HRAS* and *HGFR*) or truncation of exons encoding the extracellular ligand-binding receptor domain thus creating a constitutively active form of the receptor kinase (*EGFR* and *MST1R*). All the genes are known human oncogenes. Most of them have already been implicated in human liver cancerogenesis; here we show, for the first time, deregulation of *MST1R* as the causative event in cholangiocarcinomas. Similarly as in humans, growth of the great majority of liver lesions ceased at the stage of small nodules, suggesting that the mutations discovered were early initiating events. *Conclusion:* This model, due to very short latency, provides an opportunity to model the progression of three liver tumor types, study the interplay of genetic and nongenetic factors in their development, and identify new targets for therapeutic development.

Introduction

Liver tumors are among the most frequent and most fatal of human malignancies, with no effective treatment options. The major types of liver cancer are hepatocellular carcinomas (HCC) and the less frequent bile duct carcinomas – intrahepatic cholangiocarcinomas (ICC). Other histotypes, including hepatic hemangiosarcomas (HHS), are rare (1, 2).

An absolute majority of HCC cases are linked to persistent liver damage and inflammation caused by HBV or HCV infections, chronic intoxication with alcohol or aflatoxins, and some metabolic diseases. Recurring regenerative hyperplasia, fibrosis, the sustained activation of inflammatory and wound healing pathways, and inflammation-related increases in mutation rate create an effective combination of tumor-promoting conditions (3). Early stages (dysplastic nodules) are being found in a high proportion of individuals with chronic liver disease. Their progression through the sequential

overgrowth of more aggressive subclones may eventually lead to untreatable highly malignant tumors. The inherent capability of progressed HCC to metastasize intrahepatically is the primary cause of regular recurrences and poor prognoses. The development of HCC is very slow however, mostly lasting tens of years, and the majority of dysplastic nodules never advance to the stage of malignancy, similarly to precursors of other types of carcinomas (4, 5). Concerning genetics, HCCs are highly heterogeneous. Extensive studies have uncovered a plethora of cancer genes and pathways that could participate in tumor development, and the list is steadily growing (6, 7). Despite enormous effort a great deal still remains to be done to comprehensively describe the genomic landscape of HCC.

All major attributes of HCC also apply to ICC, including the frequent occurrence of preneoplastic lesions, strong promotion by chronic liver damage/inflammation, and high mortality. The HCC-promoting liver diseases specified above increase the incidence of ICC as well, but there are additional distinctive ICC-promoting conditions like hepatolithiasis, sclerosing cholangitis, and infection with certain parasitic flatworms endemic in south-east Asia. However, a large proportion of ICCs has no obvious etiology (8). The genetics of ICC has been less investigated than that of HCC but seems to be equally complex (9).

HHS has been studied very little so far. Exposure to certain tumor promoting or mutagenic agents (e.g. vinyl chloride) is the known risk factor for the development of HHS. There are no recognized preneoplastic precursors to hemangiosarcoma and the knowledge of underlying genetic defects is very limited (10).

To screen effectively for further cancer genes, validate their causative role in cancerogenesis, and elucidate the interplay and chronology of individual events as well as study interactions between the tumor and its microenvironment, animal models are indispensable. A list of existing liver cancer models is extensive but contains mostly rodent models of HCC (11, 12).

Recently we established a chicken model based on the unique tumorigenic properties of myeloblastosis-associated virus-2 (MAV-2) (13, 14). Infection of chickens with MAV-2 results in tumor formation in three organs: kidney (nephroblastomas), lung (hemangiosarcomas) and liver. In contrast to HBV- and HCV-related tumorigenesis, tumor induction by MAV-2 does not rely on viral genes´ function. Instead, provirus integration (an obligatory part of the retroviral life cycle) and the capacity of regulatory sequences residing in provirus LTRs to significantly influence the expression of adjacent host genes are responsible for the activation of resident cellular oncogenes or inactivation of tumor suppressor genes. Since retroviruses infect the target tissue with high efficiency and integrate into the genome almost randomly, essentially each locus of a host genome can be hit by a provirus insertion in some cells. If any of these insertions or their combinations incites malignant transformation, the affected cell outgrows and gives rise to a tumor. Perturbed host gene loci can be easily identified since they are tagged by integrated proviral sequences. Recurrent provirus insertion into the same locus in independent tumors (a so-called CIS, Common Integration Site) is considered proof of the presence of a cancer-related gene in the locus. This mechanism is referred to as

oncogenesis through insertional mutagenesis and has proven to be the most effective technique of searching for cancer-related cellular genes (11, 15).

We have already identified cancer genes insertionally mutagenized in MAV-2-induced nephroblastomas (*FOXP1*, *PLAG1*, *TWIST*, *HRAS* etc.) and lung hemangiosarcomas (*FYN*-related kinase – *FRK*) (13, 14). In our model we have also made the original observation that non-tumor cells released into the blood circulation and passing further into tissues strongly promote the formation of a tumor from latent malignant cells primed by MAV-2 insertional mutagenesis (13). Based on this observation, we have proposed the concept of industasis – the promotion of a fully malignant phenotype of an incipient tumor cell by a stray non-tumorigenic cell through a disruption of local homeostasis.

Here we present histological and molecular analyses of chicken liver lesions induced by MAV-2 with the help of industasis. These analyses revealed *HRAS*, *EGFR*, *MST1R* and *HGFR* as driver genes specific for individual tumor types and characterized the mechanisms of their tumorigenic activation.

Materials and methods

Animals: Outbread Brown Leghorns and inbred congenic CB and CC White Leghorns (16) were used. All procedures were performed in accordance with the "Guide for the Care and Use of Laboratory Animals" (NIH 1985) and approved by the Animal Care and Use Committee of the Academy of Sciences of the Czech Republic.

Viruses: The MAV-2 was the MAV-2(N)-type virus isolated from the AMV-BAI-A complex stock by plaque purification (14).

Cells: Chicken embryo fibroblasts (CEF) were prepared from CB embryos and infected with MAV-2 in culture. Semistable MAV-2-producing clonal fibroblastoid cell lines were prepared from kidney of MAV-2-infected CB embryos (A210, B211) or from lung of adult MAV-2-infected CB chicken (P4) (13).

Tumor induction and sample collection

12-day chicken embryos were injected with $1x10^5$ MAV-2-producing cells or with cell-free MAV-2 stock ($4x10^6$ PFU) into the chorioallantoic vein. The animals were sacrificed and examined 5 to 15 weeks post hatching. Macroscopically discernible liver nodules were resected for histological analysis and isolation of DNA and RNA.

Analysis of samples

Histological analysis and isolation of DNA and RNA were performed as described (14). For details of immunohistochemistry, iPCR, RT-PCR, sequencing and homology searches see Supplementary Methods.

Supplementary methods

Histochemistry:

Immunohistochemical analysis was performed on a Ventana BenchMark ULTRA Advanced Staining System (Ventana Medical Systems, Inc., USA) with the ultraView Universal DAB Detection Kit by heat induced epitope retrieval at 95°C in cell conditioning medium CC1 for 36 or 32 minutes, respectively, and incubation at 37°C for 24 or 32 minutes, respectively, with primary antibody diluted 1:100 (mouse monoclonal anti-CK antibody, clone AE1/AE3, Dako or mouse monoclonal anti-SMA antibody, clone 1A4, Ventana, RTU, respectively).

Inverse PCR:

1 μg of genomic DNA was digested with BstYI plus BclI restriction enzymes, circularized with T4 DNA ligase in a volume of 500 μl and linearized with enzymes cutting in the middle of the LTR (ApaLI, Eco81I or SphI; alternative enzymes was used to avoid systematically missing VISs containing target sites for individual enzymes). Each reaction was carried out overnight under conditions according to manufactures instructions and purified by phenol/chloroform extraction and isopropanol precipitation between individual steps. 100 ng of final product was used as template for PCR reactions under the following conditions: 1 U Taq polymerase + 0,005 U DeepVent polymerase (New England Biolabs, Beverly, MA) per 20 μl in AccuTaq buffer (Promega, Madison, WI) supplemented with 500 uM dNTPs and 500 nM primers. Mg^{2+} concentration was adjusted to 2.5 mM with $MgSO_4$. PCR cycles were as follows: 95°C for 20 seconds, 22 cycles (95°C for 20 seconds, 65°C for 12 minutes) plus additional prolonged cycles: 95°C for 20 seconds, 65°C for 18 minutes and 95°C for 20 seconds, 65°C for 30 minutes. Three separate PCR reactions (using different pairs of primers) were run for each tumor DNA providing 5´LTR junction fragments (primers L + 5´), 3´LTR junction fragments (primers R + 3´) and both (primers L + R). The PCR products were resolved overnight on 3% NuSieve GTG agarose (Lonza, Basel, Switzerland) in TBE supplemented with 50 mM sodium acetate; individual bands were excised and placed into 100 μl of 100 mM KCl, 0,1 mM EDTA. The samples were melted 5 minutes at 70°C and 1 μl of the melted samples were reamplified in 10 μl reactions under the same conditions as above except using 15 cycles instead of 22. The reaction mixture was purified with ExoSAP IT (USB, Cleveland, OH) before sequencing. A more detailed protocol is available upon request.

RT PCR:

cDNA was synthesized as described previously (14). CIS-specific RT-PCR was performed using GoTaq polymerase (Promega, Madison, WI) according to manufacturer instructions. For fusion *HGFR* mRNA detection, AccuTaq polymerase (SIGMA, St. Louis, MO) was used. The standard cycling

protocol was 25 cycles (95°C for 15 seconds, 60°C for 30 seconds and 72°C for 30 seconds). When no product could be seen on the agarose gel, 5 PCR cycles were added. The PCR products were resolved on 2% NuSieve GTG agarose in TBE; for sequencing, individual bands were excised, reamplified and purified as above.

Primers:

For position and orientation of retrovirus-specific primers see Fig. S2; the position of gene-specific primers is indicated by the exon number (as shown in Fig. 3), their orientation by ± sign.

iPCR primers: L (GGTGCATCAGGCGATTCCCTTATTTGG)

R (CCTGGGTTGATGGCCGGACCG)

5´ (AAGGCTTCATTTGGTGACCCCGACG)

3´ (TAAGACTACATTTCCCCCTCCCTATGCAAAAGC)

RT-PCRprimers

total *HRAS* mRNA:	**HrasE3+** (GCGTCTCCATGACCGAGTACAAGCTGGTG)
	/ **HrasE6–** (CGTCGACACCTCTCCAAGTCAGG)
fusion *HRAS* mRNA:	**R** / **HrasE6–**
total *5´EGFR* mRNA:	**EgfrE8+** (TGTGCAATGGCATTGGAATAGGTGAAC)
	/ **EgfrE10–** (GCTGCTTGGTTCGGCCTCGGATA)
total *3´EGFR* mRNA:	**EgfrE16+** (TATCGTGCGGAAGCGCACCCTG)
	/ **EgfrE19–** (CCAGCAAGCGGCACACATGAGGA)
fusion *EGFR* mRNA:	**R** / **EgfrE17– (**GCCTGGTTTGGTGCCTCCCCACTG)
total *5´MST1R* mRNA:	**MstrE4+** (CCACCCTCAGTGGCTTCGTCTTCG)
	/ **MstrE6–** (CAGGCTGGTGGCAGCAGGAGCTG)
total *3´MST1R* mRNA:	**MstrE16+** (TACTACAGCATCCGGCAGCACCG)
	/ **MstrE17–** (GCAGGTAGCGGGCCATGTCGTAG)
fusion *MST1R* mRNA:	**R** / **MstrE17–**
total *HGFR* mRNA:	**HgfrE14+** (CCTGGACAGGCTGGTGAGTGCAAG)
	/ **HgfrE15–** (CCACTGGACAGAATTGGGGAGAGGTC)
fusion *HGFR* mRNA:	**R** / **HgfrE15–**
total *AKT1* mRNA:	**AktE11+** (ACTATGGTCGTGCAGTGGACTGGTG)
	/ **AktE12–** (CTTGGCATCATCAGGACCGCCTC)
fusion *AKT1*mRNA:	**R** / **AktE12–**
control *ACTB* mRNA:	**ACTBsense** (CCGTAAGGATCTGTATGCCAAC)
	ACTBantisense (GTGTGGGTGTTGGTAACAGTCC)

DNA sequencing and homology searches

Sequencing reactions were performed using the BigDye Terminator Cycle Sequencing Kit (PE Biosystems, Warrington, England) and resolved on an ABI PRISM 310 Sequencer. Primers L and R were used for sequencing 5´ and 3´ junction fragments, respectively. High-quality noncomposite

sequences were edited using Chromas v1.42. Sequence homology searches were conducted using the BLAT algorithm on the chicken genome assembly database (built 3, ENSEMBL project, http://www.ensembl.org/).

Efficiency of the retrieval of insertion sites

In ca. 12% of the analyzed tumors, no insertion into any of the four major CISs was detected by iPCR. We suggest that at least in a majority of these cases, the CIS-insertions were present but were missed due to several possible reasons. We tested that highly GC-rich sequences and long stem-loop structures were no obstacle for PCR under our conditions (very long initiation/elongation steps at slightly decreased temperature). Nevertheless, some junction fragments were lost due to the presence of recognition sites for linearization enzymes ApaLI, Eco81I or SphI in the region to be amplified. Since the average frequency of these target sites in the chicken genome is ca. one per 6 kb while the frequency of BstYI+BclI target sites is ca. one per 800 bp, the expected probability of successful amplification is ca. 87% for each junction fragment. In case of nondefective proviruses where either a 5′ or 3′ junction fragment could be detected, the average total yield would thus be ca. 98%. However, as a result of extensive rearrangement, the majority of cancer gene-activating proviruses produced only one LTR-containing junction fragment, the other LTR (if present) was buried inside the provirus structure and did not produce a genomic sequence-containing iPCR fragment (Pečenka, in preparation). Furthermore, ca. 3% of junction fragments are expected to be longer than 5 kb, which is above the size limit of PCR under our conditions. Finally, 21 out of 234 unique VISs (ca. 9%) could not be unambiguously aligned since they were located in repetitive sequences or gaps of the current genomic assembly or the host part of the iPCR product was too short.

Similarly, in ca. 11% of samples CIS-specific RT-PCR failed to detect activation of any of the four major cancer genes. We suggest this was mostly due to the presence of non-tumor samples among those analyzed. We found that some of the smallest liver lesions (ca. 10% of isolates) repeatedly produced no distinct iPCR fragments, indicating the absence of any clonal cell population. These lesions were considered non-tumor outgrowths, e.g. fibrotic and regenerative nodules, and excluded from further analysis. Samples analyzed only by CIS-specific RT-PCR, however, could not be preselected this way.

Results

Industatic promotion of liver tumors

Liver tumors were induced by intravenous injection of MAV-2-producing mesenchymal cells into 12-day chicken embryos. This treatment coupled virus infection and the introduction of tumor-promoting cells, and had previously been shown to have the same effect as stepwise injection of

virions and cells (13). Control embryos were injected with cell-free MAV-2 preparations. The animals were sacrificed and examined 5 to 15 weeks post hatching. Macroscopically discernible liver nodules were resected for histological analysis and isolation of DNA and RNA. As shown in Fig. 1, industatic promotion increased the incidence of lung and liver tumors from 12 and 15 % to 69 and 38 % respectively; renal tumors were not promoted.

Three types of tumors develop in MAV-2-mutagenized liver: hemangiosarcomas, hepatocarcinomas and cholangiocarcinomas

All liver tumors appeared as well-circumscribed nodules 0.5–30 mm in diameter with no specific lobular distribution. Two types of nodules could be recognized: off-white to yellow-white with the texture of fibrous tissue, and rather soft pinkish-red or red-white-speckled (Fig. S1.). Histopathological analysis (Fig. 2) revealed a heterogeneous spectrum of histotypes falling into three groups: HHS, HCC (including preneoplastic stages), and ICC. HHSs mostly corresponded to the pinkish/red-white nodules, while HCCs and ICCs were off-white/yellow-white and were indistinguishable from each other macroscopically.

HHSs contained large cystic vascular spaces often filled with red blood cells, surrounded by bulky and dense masses of unorganized poorly differentiated spindle-shaped endothelial cells bulging into the vascular lumen in some places. The tumors were accompanied by occasional hemorrhage and histologically resembled the high grade stage of human hemangiosarcomas (17).

HCCs were composed of densely packed well to poorly differentiated hepatocytes with pleomorphic nuclei and prominent nucleoli. The appearance of tumors was trabecular with minimal stroma, sometimes with a pseudoacinar pattern within the trabeculae, occasionally with pseudoglandular structures accompanied by pronounced stroma. Architectural atypia of HCCs included high cell density, a multilayer nature of epithelial structures, and an absence of triads plus hypovascularity in many places. Unpaired arteries and stromal invasion were present in progressed tumors. According to the criteria and terminology introduced by the ICGHN (18), these tumors ranged from high grade dysplastic nodules to early HCC.

ICCs comprised disorganized ducts, cell nests and cords as well as individualized cells, surrounded by rich amounts of fibrous stroma. The cells of ICC were cuboidal, columnar, or pleomorphic with small nuclei and less prominent nucleoli compared to that of HCC. Occasionally there was a small amount of inflammatory cell infiltrate. Overall the chicken ICCs resembled the mass forming ductular type of human ICC (19).

Two tumors displayed features of both HCC and ICC, reminiscent of combined hepato-cholangiocarcinomas described in humans (20).

Clonal status of the tumors, intraorgan metastasis, and tumor invasion

To carry out high throughput identification of VISs, we optimized the iPCR technique so that for each tumor we obtained even amplification of all clonal junction fragments in a single reaction (Supplementary Methods and Fig. S2). The clonality and clonal homogeneity of lesions were

confirmed by distinct iPCR fragments with uniform intensity and by identical iPCR patterns in DNAs isolated from distant parts of a nodule (see nodule 6781Li1, Fig. S2B). By contrast, iPCR patterns in liver, lung, and kidney tumors from the same animal were completely diverse, implying independently originated tumors.

Multiple liver nodules found in some animals mostly also exhibited diverse patterns of MAV-2 insertions. In a few cases, however, two or three of these nodules displayed identical patterns, indicating intrahepatic metastasizing; all the respective nodules fell into the HHS group. In such cases only one nodule was further analyzed except two pairs of nodules that were clonally related but not identical (6775Li2/Li3 and 6239Li2/Li3). The latter pair of nodules in fact did not represent metastasis but rather tumor invasion, since the apparently separate nodules 6239Li2/Li3 (Fig. S1) were connected by a thin thread of white tenacious tissue inside the liver body. These nodules also demonstrate how clonal evolution can be traced using iPCR. There were four common iPCR fragments in the nodules plus several Li3-specific ones (Fig. S2B). The Li3-specific fragments (and one common) were weaker, suggesting that Li3 was a mixture of at least two subclones sharing three common fragments. Apparently, Li3 was the parental nodule and Li2 was founded by one of the Li3 subclones.

Identification of insertionally activated cancer genes

To identify VISs, nearly 400 iPCR-produced MAV-2 junction fragments from 52 liver tumors were gel-isolated, PCR-reamplified and directly sequenced. Sequences were aligned to the chicken genome (see Supplementary Methods). Twenty fragments contained only MAV-2 sequences. The remaining fragments represented 234 unique VISs (on average 4.5 clonal proviruses per tumor, ranging from 1 to 11 proviruses per tumor). Positions in the chicken genome were reliably assigned to 213 VISs (Tab. S1). Only 204 aligned VISs were truly independent since there were 9 pairs of identical VISs contained in the clonally related samples. Among these we identified four major CISs that fulfilled the strictest criteria of the CIS definition – the clusters of insertions were 1 to 20 kb wide and contained insertions from 5 or more tumors (Tab. S1 and Fig. 3). Each major CIS could be unequivocally associated with a single gene encoding a member of classical signaling cascades known to be causally linked to cancerogenesis: small signaling GTPase Ha-ras, and the receptor tyrosine kinases Erb-B/EGFR, Ron/MST1R and Met/HGFR. These gene loci were hit in 15, 18, 6 and 5 independent tumors, respectively (Fig. 3 and Tab. 1, S1 and S2). Simultaneous integration in two or more major CISs was not found in any tumor. In 6 tumors (12 %) no major CIS was found to be hit.

The *AKT1* locus represented a minor CIS, hit in three independent tumors (Fig. S4, Tab. S1 and S2). These three tumors also carried the provirus in one of the major CISs (*HRAS*, *HGFR* and *MST1R*, respectively). Provirus integrations in LIMS1, EPAS1 and HS1BP3 loci, resp., each in two independent tumors (Tab. S1), were not statistically significant enough to be considered a CIS.

Modes of cancer genes deregulation: overexpression versus gene truncation

Next we examined the expression of the identified cancer genes using CIS-specific RT-PCR that detected either all gene-specific mRNAs or only hybrid mRNAs containing sequences of both the

MAV-2 LTR and a candidate cancer gene. This extended our analysis to samples from which only mRNA was available and revealed the impact of inserted provirus on CIS-associated genes. The RT-PCR confirmed recurrent provirus integration into the major CISs and demonstrated the significantly increased expression of candidate cancer genes in the affected loci (Fig. 4, Tab. 1). The VISs positions and the structure of provirus-induced transcripts, revealed by sequencing of iPCR and RT-PCR products as well as mapping of the provirus structure (Pecenka *et al.*, in preparation), allowed us to propose detailed mechanisms of the activation of individual cancer genes (Fig. 3 and S3).

In the *HRAS* locus, all provirus insertions occurred inside the introns upstream of the second or third exon. All closely analyzed activating proviruses were heavily rearranged, including large deletions, duplications, and inversions. Insertions resulted in the overexpression of hybrid MAV-2/*HRAS* mRNA containing the end of the proviral 3´LTR and the downstream host sequences. Commonly, a cryptic splice donor downstream of the LTR was used to adjoin the first downstream *HRAS* exon. Complete *HRAS*-coding sequences including the start codon were preserved. No mutations were found in the entire coding sequence of several LTR-driven *HRAS* transcripts (data not shown). This suggests that normal Ha-ras protein was overproduced in these tumors.

A different mode of activation was found in the *EGFR* locus. All proviruses were defective and had integrated within the same region in the middle of the gene. Insertions resulted in the truncation of *EGFR* exons coding for the N-terminal ligand-binding extracellular domain, and triggered a high expression of mRNA encoding only the C-terminal regions. Respective mRNAs were initiated at the proviral 5´LTR and processed by splicing between the MAV-2 donor site and the first downstream *EGFR* exon. The coding region of the spliced mRNAs consisted of the first six gag codons (including the start codon) fused in-frame to sequences encoding transmembrane plus cytoplasmic EGFR domains.

The *MST1R* activation in some aspects paralleled the *EGFR* activation – the insertions of defective proviruses occurred in the middle of the gene and truncated the exons encoding the receptor´s extracellular ligand–binding domain. The proviruses, however, were positioned inside exons (the chicken *MST1R* gene has uncommonly short introns), the mRNA expression was driven by the 3´LTR, and its translation relied on the internal in-frame *MST1R* AUG codon.

The arrangement of activated *HGFR* resembled the situation in the *HRAS* locus – defective proviruses integrated within the promoter region or in the first intron, upstream of the second (i.e. the first coding) exon, and drove overexpression of fusion mRNA encoding full–length HGFR protein.

All analyzed cases corresponded to the promoter insertion mechanism where the transcription of a resident cellular gene is driven by the inserted proviral promoter (21). No case of enhancer insertion (where provirus regulatory sequences boost, at a distance, a gene´s transcription from its own promoter) or formation of recombinant oncogene-transducing retrovirus was observed.

We conclude that the constitutive overexpression of either normal (*HRAS* and *HGFR*) or truncated forms (*EGFR* and *MST1R*) of cancer genes propelled the oncogenic transformation of affected cells.

The involvement of *AKT1* in oncogenesis was unlikely. Contrary to the major CISs, insertions into *AKT1* were scattered throughout the locus and there was no increase of *AKT1* mRNA expression (Fig. S4) or any detectable LTR-*AKT1* fusion mRNA (data not shown) in two analyzed samples.

Correlations between insertionally mutagenized genes, tumor types, and tumor growth

Next, we asked if there is a connection between a particular CIS and the induced tumor type. The correlation proved to be unexpectedly straightforward: *HRAS* insertional mutations were present only in HHS, *EGFR* mutations in HCC, and *MST1R* mutations in ICC or combined HCC-ICC (Tab. 1 and S2). Histological characterization of samples containing mutated *HGFR* could not be reliably performed due to their small size. These lesions might be a discrete group since, contrary to other liver lesions, their size never exceeded 2 mm (Fig. 5B and 5C).

Unexpectedly, the frequency of tumors was not significantly higher in older animals (data not shown). Moreover, a great majority of nodules were small and the proportion of small nodules did not decrease with the animals' age (Fig. 5B, C and D), which is not what one would expect if small nodules were just a transient stage on the way to large tumors. Clearly, the growth of most lesions ceased at the stage of a small nodule. Since the *HGFR* group represented only 14% of small nodules, the phenomenon necessarily also involved the other groups.

Discussion

Tumor development depends crucially on the interaction of cancer cells with the tumor microenvironment and even with distant parts of the body (22). Liver tumors fully conform to this concept, as exemplified by their strong dependency on chronic liver wounding and inflammation in humans, by the remarkable promotion of liver tumors by non-genotoxic hepatotoxins in mouse models, and by the common existence of preneoplastic stages that progress to a malignant tumor only infrequently despite already carrying liver cancer-related genetic defects (3, 5, 23, 24). Our experimental model takes advantage of industasis – strong tumor promotion mediated by non-tumorigenic stray cells (13).

Three types of liver tumors were identified in our model: HHSs, HCCs and ICCs (plus rare cases of mixed HCC-ICC histotypes). There were no extrahepatic metastases and only a few cases of intrahepatic ones, all involving HHSs which are inherently aggressive due to the lack of natural barriers to their spread. The majority of tumors were small, in the case of HCC representing mostly preneoplastic stages (dysplastic nodules). Large tumors were a minority but were present in each histological group.

In over 85% of chicken liver tumors one of four cancer genes was found to be insertionally activated. All these genes (*HRAS*, *ERBB1/EGFR*, *RON/MST1R* and *MET/HGFR*) are constituents of

known signaling pathways participating in cancerogenesis. No tumor carried simultaneous activation of two or more of these genes, and there was no other recurrent insertionally mutated gene.

Two modes of insertional activation were found. Insertions in *HRAS* and *HGFR* loci did not affect the genes´ coding regions and caused overexpression of mRNAs encoding normal proteins. By contrast, proviruses in *EGFR* and *MST1R* loci were positioned in the middle of the genes, thus eliminating exons encoding ligand-binding extracellular domains. This resulted in the expression of "receptors" containing only transmembrane and cytoplasmic domains and displaying high constitutive kinase activity, analogous to protein products of the transduced viral oncogenes *v-erb-B* and *v-sea* (*SEA* is the original name of the chicken ortholog of *RON/MST1R*) (25-28).

There was a perfect correlation between the mutated gene and tumor type: *HRAS* activation was found in HHSs, *EGFR* in HCCs, and *MST1R* in ICCs or combined HCC-ICCs. *HGFR* activation invariantly produced only very small lesions; their histotype and capacity to progress further into malignant stages remain to be determined. *HGFR* is a close relative of *MST1R* both in terms of sequence homology and shared downstream signaling pathways, and has a proven role in human HCC and ICC (6, 9, 29). We suggest that the different impact of insertions into *HGFR* versus *MST1R* ensued from different modes of the genes´ activation. While the insertional activation of *MST1R* resulted in ligand-independent firing, activation of *HGFR* only rendered the cells hypersensitive to HGF but signaling still relied on the ligand supply from mesenchymal cells, which might become insufficient after the nodule reaches a certain size.

The activation of the same (and no other) cancer gene in each particular tumor type in our model is in striking contrast to human HCCs or ICCs, which carry multiple and diverse genetic defects that, when combined, represent a majority of known cancer-related pathways (6, 7, 9). We suggest that this difference in genetic landscape complexity does not arise from evolutionary distance but results from features typical for experimental models. First, human liver tumors have many possible etiologies, e.g. infection by HBV or HCV, alcohol abuse, or exposure to mutagens like aflatoxins in the case of HCC. Each of the primary agents induces specific oncogenic changes in the cell: introduction of transforming genes encoded by HBV or HCV, preferred mutations induced by aflatoxins or inflammation-related reactive molecules, etc. This subsequently makes the cell sensitive to a limited set of further cooperating genetic defects, most likely different for different primary insults (30). By contrast, the chicken model has a single tumor-inducing agent that acts as an insertional mutagen with its own mutation target preferences (21). Second, human liver tumors develop in very diverse conditions regarding the tumor microenvironment; different conditions may support tumors with different genetic defects. On the other hand, the conditions in experimental animals are almost identical. There is no chronic liver damage. The chicken tumors thus parallel human tumors that originated without preceding liver disease or inflammation. Third, human liver tumors seem to have diverse cells of origin (mature differentiated cells and resident or circulating stem/progenitor cells), each of them sensitive to the deregulation of different cancer genes. The target cells for tumorigenesis

in chicken embryonal liver are not known, but may not be as diverse as in the case of human liver cancerogenesis. Fourth, the majority of analyzed chicken tumors represent early stages of tumor development, "waiting" for further oncogenic changes. Thus, only the initiating genetic events were revealed in our screen. Further tumor progression would likely erode the apparent simplicity of the genetic landscape of chicken liver tumors.

Tumors in our model originate in embryos during a short period of massive MAV-2 integrations and stray cell promotion. Most of them remain under the control of tissue homeostasis and stop growing at the stage of a small nodule. Such a situation is also typical for human liver tumors (5) and mirrors the multistage nature of tumor development. Further progression is thought to depend on the acquisition of additional hallmarks of cancer. This is unlikely to happen through additional provirus insertions during tumor development because of superinfection resistance (31). Eventual additional insertions accounting for infrequent large tumors would have to be present in the founder cell from the beginning. However, no additional CISs have been found in the progressed tumors compared to the small ones. The only revealed additional CIS, the *AKT1* locus, was hit in three tumors (together with insertions in *HRAS*, *HGFR* and *MST1R*, respectively), but all three lesions belonged the smallest size category; moreover the integrations had no effect on *AKT1* expression. We suggest that tumor progression in our system depends on spontaneous mutations/chromosome rearrangements and/or spontaneous epigenetic shifts in tumor cells. The infrequent formation of large tumors corresponds to the short time available for such events. Since spontaneous *β-catenin* mutations have been shown to be the secondary hits in a mouse HCC model (32), we looked for *β-catenin* mutations or rearrangements by exon 3 sequencing and long-range PCR in a series of tumors; the results were negative (data not shown).

In many aspects, the chicken liver tumors appear to be faithful models of human ones. Their histopathology is so similar that they can be assigned to specific subtypes of human liver tumors, they are similarly dependent on microenvironment perturbation, similarly develop arrested early stages of cancerogenesis, and (with the exception of *MST1R)* the involved cancer genes/pathways are also implicated in corresponding human liver tumors (6, 9, 33). This all provides proof-of-principle for the validity of the model. Simultaneously, the model demonstrated, for the first time, the *MST1R* deregulation as causative initiating event in cholangiocarcinomas.

Dissimilarities between chicken versus human liver tumors are rather insignificant, e.g. the issue of gender disparity: while males are more liver tumor-prone in mammalian species (34), the penetrance of each liver tumor type in hens and roosters is identical (data not shown).

Several areas of cancer research can be investigated using our model in addition to the traditional screening for cancer genes. First of all, genetic and non-genetic factors capable of driving further progression of small liver nodules could be sought and the course of tumorigenesis reconstructed, capitalizing on short tumor latency, the possibility of concurrently analyzing three different tumor types, their simple and uniform genetic landscapes and relatively easy search for the involved

insertionally activated cancer genes. To achieve this, however, the development of aggressive lung hemangiosarcomas that limit the animals´ survival must be prevented. This could be accomplished by using liver-specific tumor promotion with nongenotoxic hepatotoxins (carbon tetrachloride, phenobarbital, buthoxyethanol etc.) (23, 24) instead of injected mesenchymal cells.

Second, the origins of hemangiosarcoma (for which there are currently very limited data and no good experimental model) can be investigated. Our data already obtained show organ-specific sensitivity to the deregulation of a particular cancer gene: *HRAS* in liver hemangiosarcomas (this paper), *FRK* in lung hemangiosarcomas (13) and *MST1R* in heart hemangiosarcomas (data not shown).

In conclusion, we believe that the chicken model described here has the potential to resolve some important and controversial issues in cancerogenesis, leading to data that could translate to human medicine.

References

1. Bioulac-Sage P, Laumonier H, Laurent C, Blanc JF, Balabaud C. Benign and malignant vascular tumors of the liver in adults. Semin Liver Dis 2008;28:302-314.
2. Walther Z, Jain D. Molecular pathology of hepatic neoplasms: classification and clinical significance. Patholog Res Int 2011;2011:403929.
3. Sanyal AJ, Yoon SK, Lencioni R. The etiology of hepatocellular carcinoma and consequences for treatment. Oncologist 2010;15 Suppl 4:14-22.
4. Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. Nat Med 2011;17:320-329.
5. Teoh NC. Proliferative drive and liver carcinogenesis: too much of a good thing? J Gastroenterol Hepatol 2009;24:1817-1825.
6. Whittaker S, Marais R, Zhu AX. The role of signaling pathways in the development and treatment of hepatocellular carcinoma. Oncogene 2010;29:4989-5005.
7. Zender L, Villanueva A, Tovar V, Sia D, Chiang DY, Llovet JM. Cancer gene discovery in hepatocellular carcinoma. J Hepatol 2010;52:921-929.
8. Gatto M, Bragazzi MC, Semeraro R, Napoli C, Gentile R, Torrice A, Gaudio E, et al. Cholangiocarcinoma: update and future perspectives. Dig Liver Dis 2010;42:253-260.
9. Fava G. Molecular mechanisms of cholangiocarcinoma. World J Gastrointest Pathophysiol 2010;1:12-22.
10. Andersen NJ, Froman RE, Kitchell BE, Duesbery NS: Clinical and Molecular Biology of Angiosarcoma. In: Derbel F, ed. Soft Tissue tumors. Rijeka, Croatia: InTech, 2011; 149-174.

11. Keng VW, Villanueva A, Chiang DY, Dupuy AJ, Ryan BJ, Matise I, Silverstein KA, et al. A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma. Nat Biotechnol 2009;27:264-274.
12. Li Y, Tang ZY, Hou JX. Hepatocellular carcinoma: insight from animal models. Nat Rev Gastroenterol Hepatol 2011;9:32-43.
13. Pajer P, Karafiat V, Pecenka V, Prukova D, Dudlova J, Plachy J, Kasparova P, et al. Industasis, a promotion of tumor formation by nontumorigenic stray cells. Cancer Res 2009;69:4605-4612.
14. Pajer P, Pecenka V, Kralova J, Karafiat V, Prukova D, Zemanova Z, Kodet R, et al. Identification of potential human oncogenes by mapping the common viral integration sites in avian nephroblastoma. Cancer Res 2006;66:78-86.
15. Dupuy AJ, Largaespada DA. Insertional Mutagenesis Strategies in Cancer Genetics. New York: Springer, 2011.
16. Plachy J, Hala K. Comparative aspects of the chicken immunogenetics (review). Folia Biol (Praha) 1997;43:133-151.
17. Weiss SW, Goldblum JR, Folpe AL: Malignant Vascular Tumors. In: Enzinger and Weiss's Soft Tissue Tumors, 5th edition. St. Louis, MO: Mosby, 2007; 703-732.
18. ICGHN. Pathologic diagnosis of early hepatocellular carcinoma: a report of the international consensus group for hepatocellular neoplasia. Hepatology 2009;49:658-664.
19. Nakanuma Y, Sato Y, Harada K, Sasaki M, Xu J, Ikeda H. Pathological classification of intrahepatic cholangiocarcinoma based on a new concept. World J Hepatol 2010;2:419-427.
20. Kim H, Park C, Han KH, Choi J, Kim YB, Kim JK, Park YN. Primary liver carcinoma of intermediate (hepatocyte-cholangiocyte) phenotype. J Hepatol 2004;40:298-304.
21. Pecenka V, Pajer P, Karafiat V, Dvorak M: Chicken Models of Retroviral Insertional Mutagenesis. In: Dupuy AJ, D.A. L, eds. Insertional Mutagenesis Strategies in Cancer Genetics. New York: Springer, 2011; 77-112.
22. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144:646-674.
23. Beer S, Komatsubara K, Bellovin DI, Kurobe M, Sylvester K, Felsher DW. Hepatotoxin-induced changes in the adult murine liver promote MYC-induced tumorigenesis. PLoS One 2008;3:e2493.
24. Cohen SM, Storer RD, Criswell KA, Doerrer NG, Dellarco VL, Pegg DG, Wojcinski ZW, et al. Hemangiosarcoma in rodents: mode-of-action evaluation and human relevance. Toxicol Sci 2009;111:4-18.
25. Lax I, Kris R, Sasson I, Ullrich A, Hayman MJ, Beug H, Schlessinger J. Activation of c-erbB in avian leukosis virus-induced erythroblastosis leads to the expression of a truncated EGF receptor kinase. Embo J 1985;4:3179-3182.
26. Huff JL, Jelinek MA, Borgman CA, Lansing TJ, Parsons JT. The protooncogene c-sea encodes a transmembrane protein-tyrosine kinase related to the Met/hepatocyte growth factor/scatter factor receptor. Proc Natl Acad Sci U S A 1993;90:6140-6144.

27. Lu Y, Yao HP, Wang MH. Multiple variants of the RON receptor tyrosine kinase: biochemical properties, tumorigenic activities, and potential drug targets. Cancer Lett 2007;257:157-164.
28. Raines MA, Lewis WG, Crittenden LB, Kung HJ. c-erbB activation in avian leukosis virus-induced erythroblastosis: clustered integration sites and the arrangement of provirus in the c-erbB alleles. Proc Natl Acad Sci U S A 1985;82:2287-2291.
29. Wagh PK, Peace BE, Waltz SE. Met-Related Receptor Tyrosine Kinase Ron in Tumor Growth and Metastasis. Adv. Cancer Res. 2008;100:1-33
30. Schlaeger C, Longerich T, Schiller C, Bewerunge P, Mehrabi A, Toedt G, Kleeff J, et al. Etiology-dependent molecular mechanisms in human hepatocarcinogenesis. Hepatology 2008;47:511-520.
31. Coffin JM, Hughes SH, Varmus HE. Retroviruses. New York: Cold Spring Harbor Laboratory Press, 1997.
32. Tward AD, Jones KD, Yant S, Cheung ST, Fan ST, Chen X, Kay MA, et al. Distinct pathways of genomic progression to benign and malignant tumors of the liver. Proc Natl Acad Sci U S A 2007;104:14771-14776.
33. Weihrauch M, Bader M, Lehnert G, Koch B, Wittekind C, Wrbitzky R, Tannapfel A. Mutation analysis of K-ras-2 in liver angiosarcoma and adjacent nonneoplastic liver tissue from patients occupationally exposed to vinyl chloride. Environ Mol Mutagen 2002;40:36-40.
34. Naugler WE, Sakurai T, Kim S, Maeda S, Kim K, Elsharkawy AM, Karin M. Gender disparity in liver cancer due to sex differences in MyD88-dependent IL-6 production. Science 2007;317:121-124.

Acknowledgements: The authors are indebted to Dr. Alisdair M. Wood, AHVLA Lasswade, Edinburgh, UK, for help with histopathological examination of tumours and Dr. David W. Hardekopf, Chambon s.r.o., Prague, Czech Republic, for his critical comments on the manuscript.

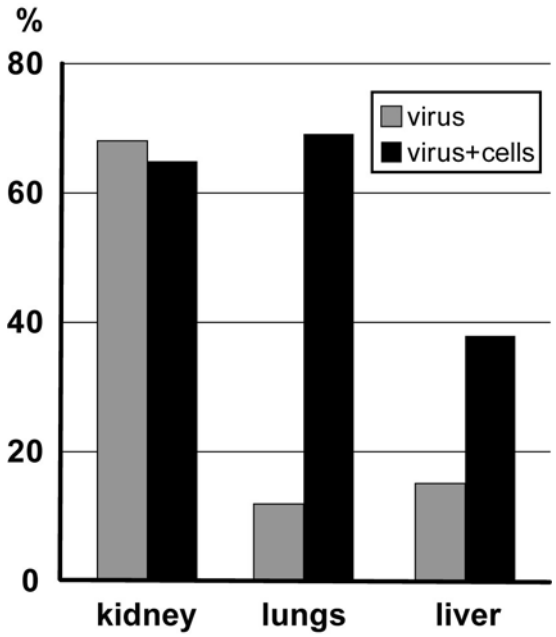


Fig. 1. Industasis – non-tumorigenic stray cell-assisted cancerogenesis. The proportion of chickens carrying tumors in different organs after *in ovo* injection with either MAV-2-producing chicken mesenchymal cells (151 animals) or MAV-2 virions (47 animals) is shown. The numbers represent the sum of data on animals sacrificed at different times post infection (5 to 15 weeks). The figure underestimates the efficacy of industasis since animals injected with MAV-2-producing cells had to be sacrificed earlier (on average 7 weeks P.I.) than animals infected with MAV-2 virions (on average 10 weeks P.I.) due to the early wasting of many animals from the first group caused by aggressive lung hemangiosarcomas. This also explains the slightly lower frequency of nephroblastomas in the first group.

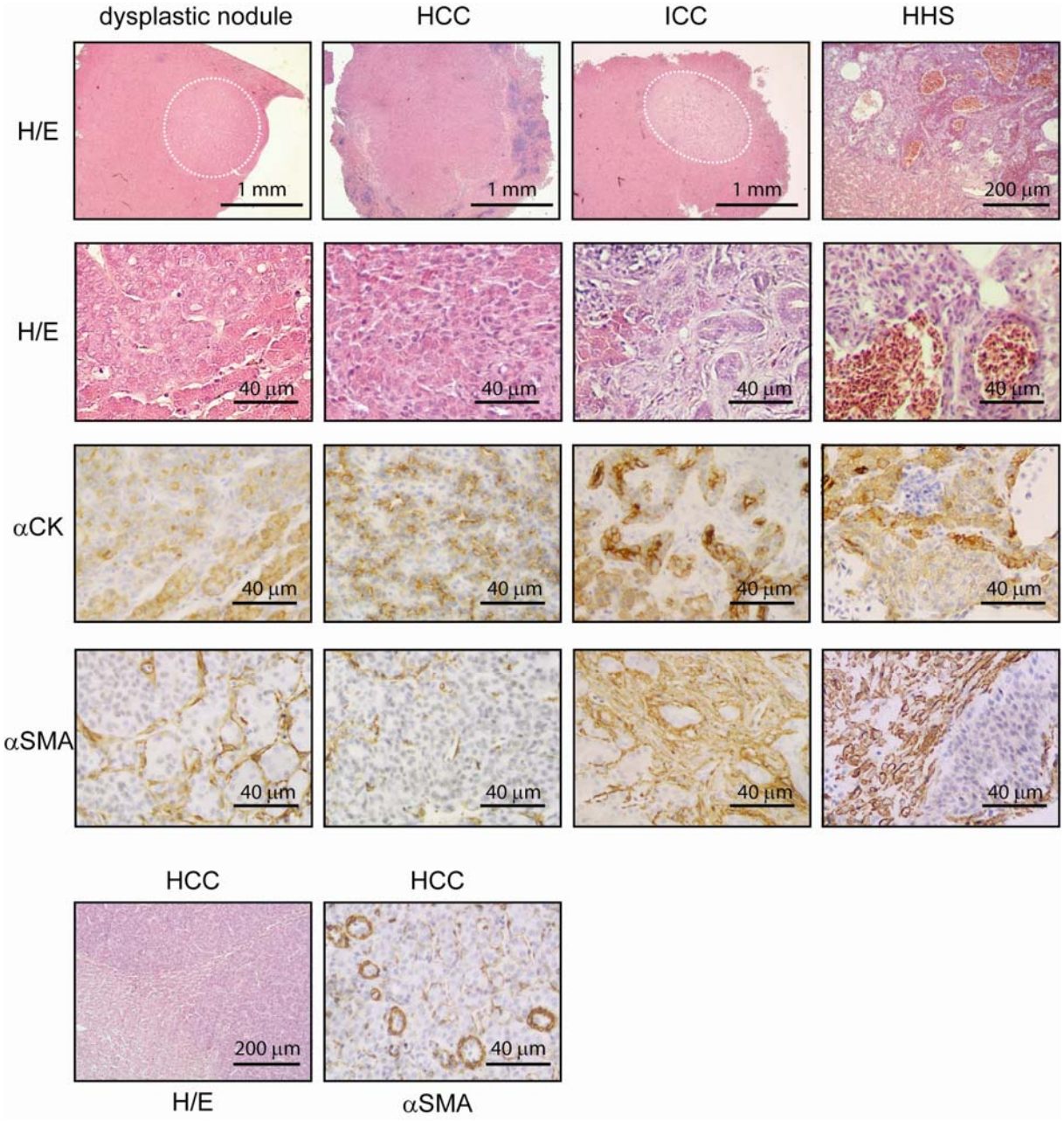


Fig. 2. Histopathology of chicken liver tumors. Representative histological sections of tumors were stained with haematoxylin-eosin (H/E), anti-cytokeratin antibody (αCK) to visualize epithelial cells, or anti-α-smooth muscle actin antibody (αSMA) to visualize mesenchymal cells (smooth muscle cells, myofibroblasts and pericytes). The two pictures on the bottom exemplify tumor invasion (H/E staining) and unpaired arteries (αSMA staining) characteristic for early HCC. Dysplastic nodule is a preneoplastic stage of HCC.

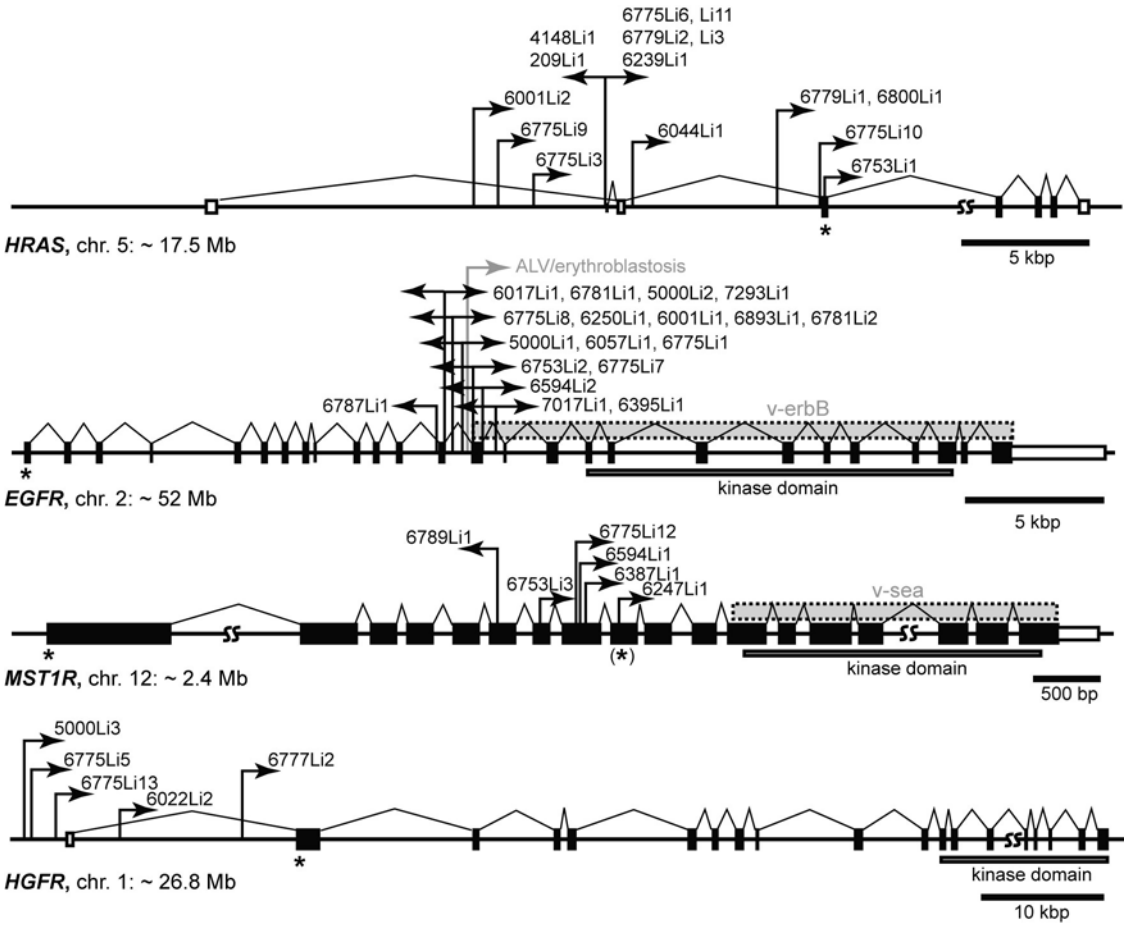


Fig. 3. Clusters of insertions into major CISs. The exon-intron structure of candidate cancer genes and position of provirus insertions in tumors are shown. Gaps in the chicken genome assembly (WASHUC2) are indicated by the double squiggly lines. Open and full rectangles represent noncoding and coding exons, respectively; transcriptional orientation is set from left to right. Bent arrows with tumor numbers indicate individual integration sites and orientation of the provirus LTR inferred from iPCR and RT-PCR results. Bidirectional arrows reflect the fact that some proviruses (nearly all in the EGFR locus) are rearranged in such a way that they carry LTRs oriented in opposite directions. Sample ALV/erythroblastosis (*EGFR* locus) represents a typical provirus integration site in RAV-1-induced chicken erythroblastosis (28). Asterisks mark the position of initiation codons; the asterisk in parentheses is the *MST1R* internal AUG codon expected to be used as an alternative start of translation in the provirus-driven fusion MAV-2/*MST1R* mRNA. Sequences retained in viral transduced oncogenes *v-erbB* and *v-sea* are denoted by a gray rectangle above these genes.

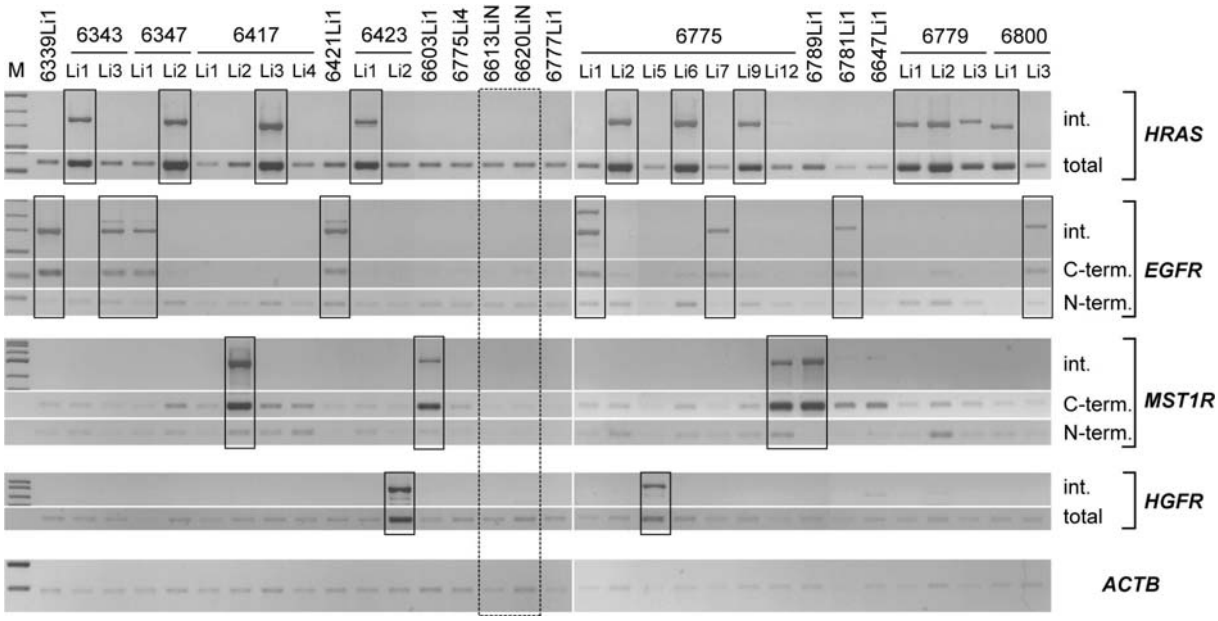


Fig. 4. Transcriptional activation of candidate cancer genes. Representative results of semiquantitative CIS-specific RT-PCR are shown. For each candidate cancer gene two types of RT-PCR were carried out: RT-PCR displaying integration-specific hybrid mRNAs, marked int. (using a primer specific for the provirus against a primer specific for the gene´s exon downstream of the VISs cluster) and RT-PCR displaying all transcripts of the gene, labeled "total" (using a pair of primers specific for the gene´s exons). For *EGFR* and *MST1R*, expression of 3´ parts of the gene (using a pair of primers specific for exons downstream of the VISs cluster) and 5´ parts of the gene (using a pair of primers specific for exons upstream of the VISs cluster) are separately displayed, labeled "C-term." and "N-term.", respectively. Strong overexpression of one of the four genes is evident in most of the samples (highlighted by boxes). In case of *EGFR* and *MST1R* activation, only the 3´part of the genes was overexpressed as a result of intragene integrations. For details on the structure of integration-specific mRNAs see Fig. S3. Neither integration-specific products nor the genes´ overexpression can be seen in samples LiN – non-tumor livers from infected animals (highlighted by the large dotted box). The length variation of integration-specific products reflects the variable position of individual integration sites; occasional multiple bands reflect alternative splicing of provirus-adjacent chicken intronic sequences. The primer pair specific for chicken *β-actin* was used to check the amount and quality of RNA (labeled "*ACTB"*).

A

Gene	Tumor type		
	HCC	ICC	HHS
HRAS	0	0	**5**
EGFR	**6**	0	0
MST1R	(1)	**2** + (1)	0
none	1	0	0

B

Gene	Nodule size [mm]		
	< 2	2 - 5	> 5
HRAS	10	6	6
EGFR	13	6	3
MST1R	3	1	4
HGFR	5	0	0
none	5	2	2
not anal.	34	3	1

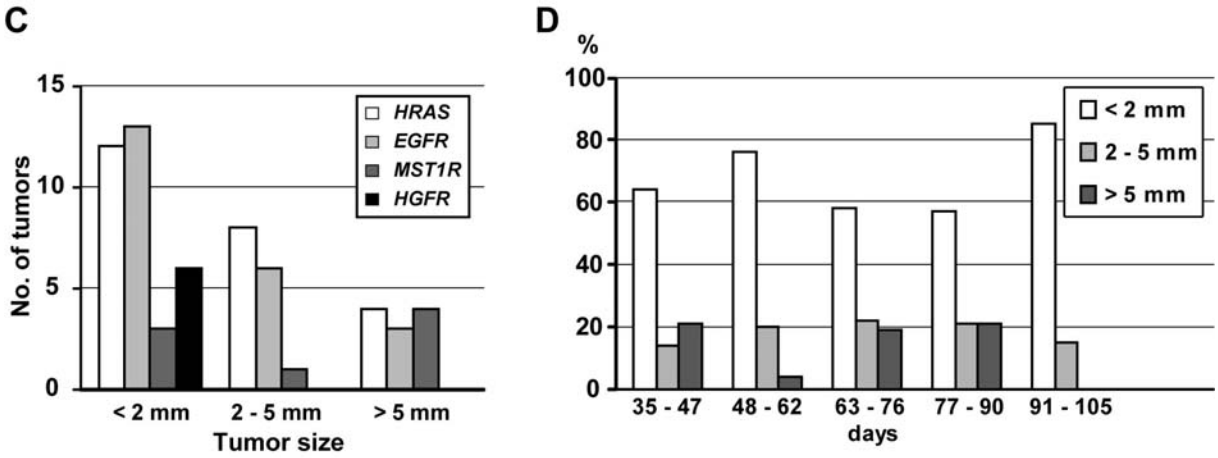


Fig. 5: Correlations between mutated genes, tumor types and nodule size.

A. Correlation between the mutated gene and tumor type. The number of tumors of different histotypes carrying insertions in individual cancer genes are shown. Only samples analyzed both molecularly and histologically are included. The number in parentheses represents a tumor of mixed HCC-ICC histotype. Samples containing mutated *HGFR* could not be histologically characterized because of their small size.

B. Size distribution of nodules in individual CIS-based categories. The numbers of tumors carrying insertions in individual CISs divided into three size categories are shown. While molecular analysis was performed in most large tumors only a minority of small ones were analyzed.

C. Frequency of mutated genes in individual size categories. The data from B displayed as a column chart. Note the biased distribution of nodules bearing a proviral insertion in the *HGFR* gene.

D. Time development of nodule sizes. Size distribution of liver tumor nodules in animals sacrificed at different times post infection is shown. Note that the proportion of each tumor size category did not change significantly during the time course of the experiment.

	iPCR	RT-PCR
HRAS	15 (30%)	15 (40%)
EGFR	18 (36%)	11 (30%)
MST1R	6 (12%)	5 (14%)
HGFR	5 (10%)	2 (5%)
none	6 (12‘%)	4 (11%)

Tab. 1. Frequency of insertions into the major CISs. The number and percentage of tumors carrying provirus insertion in individual CISs are shown as detected by iPCR and RT-PCR. No tumor carried insertions into two or more CISs simultaneously.

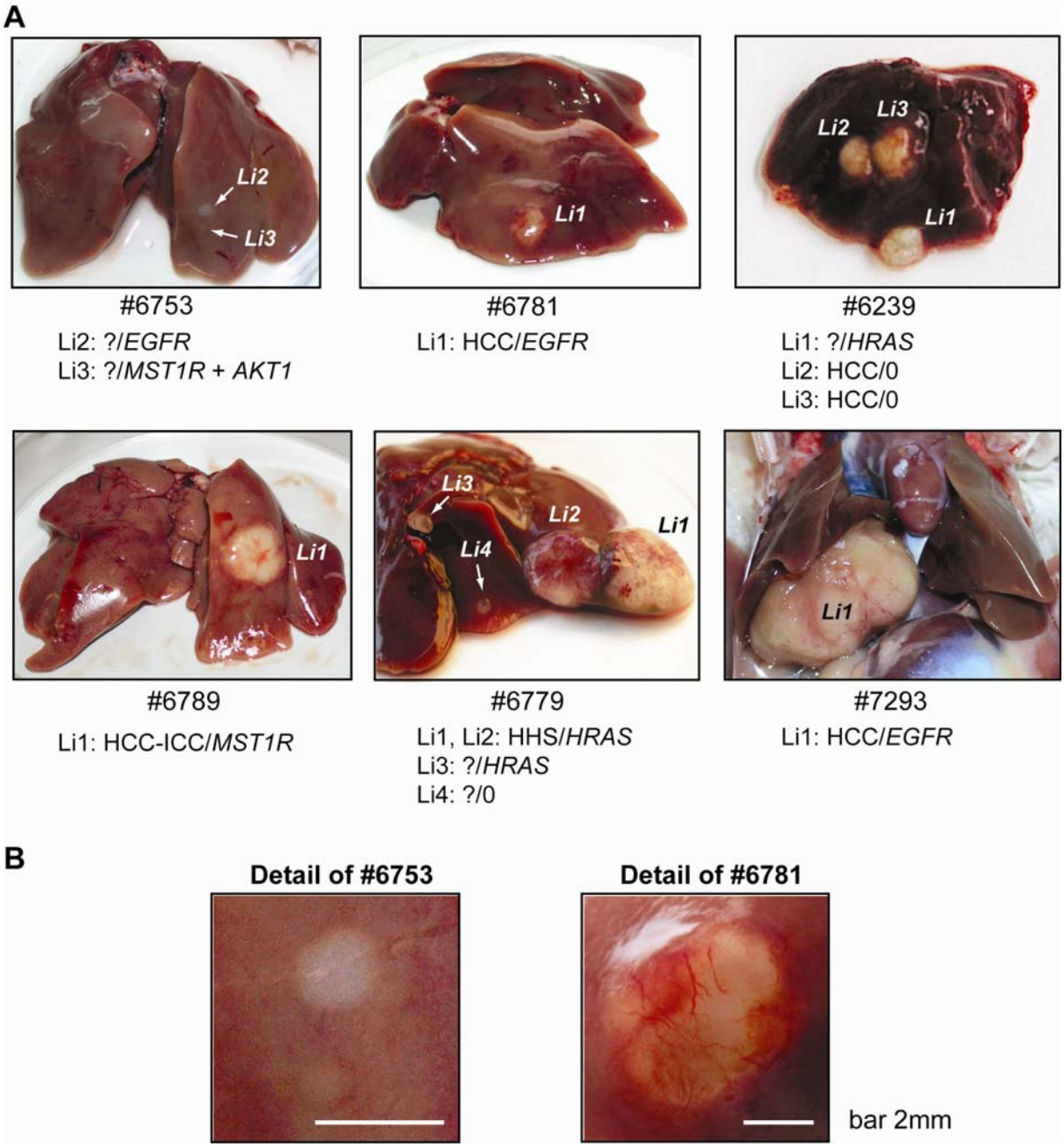
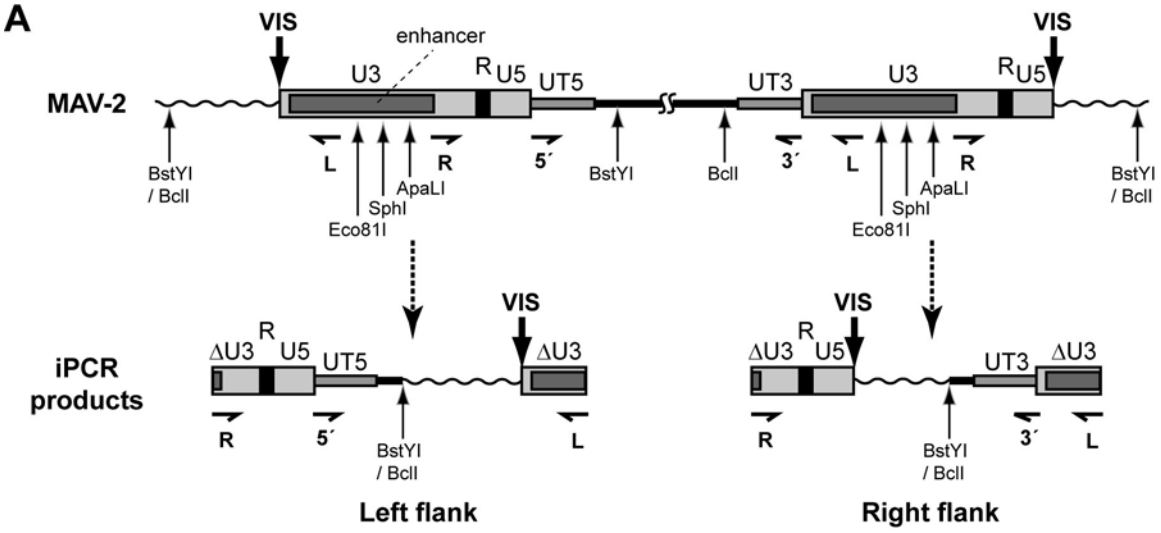


Fig. S1. Gross morphology of chicken liver tumors. **A.** Tumors representing different size categories are shown; histotype/activated cancer gene are indicated. ? – not done. **B.** A detailed view of tumors without (left) and with (right) vascularization.


A
MAV-2
VIS
enhancer
U3
R
U5
UT5
UT3
BstYI / BclI
L
R
5´
BstYI
BclI
3´
SphI
ApaLI
Eco81I
iPCR products
ΔU3
ΔU3
Left flank
Right flank

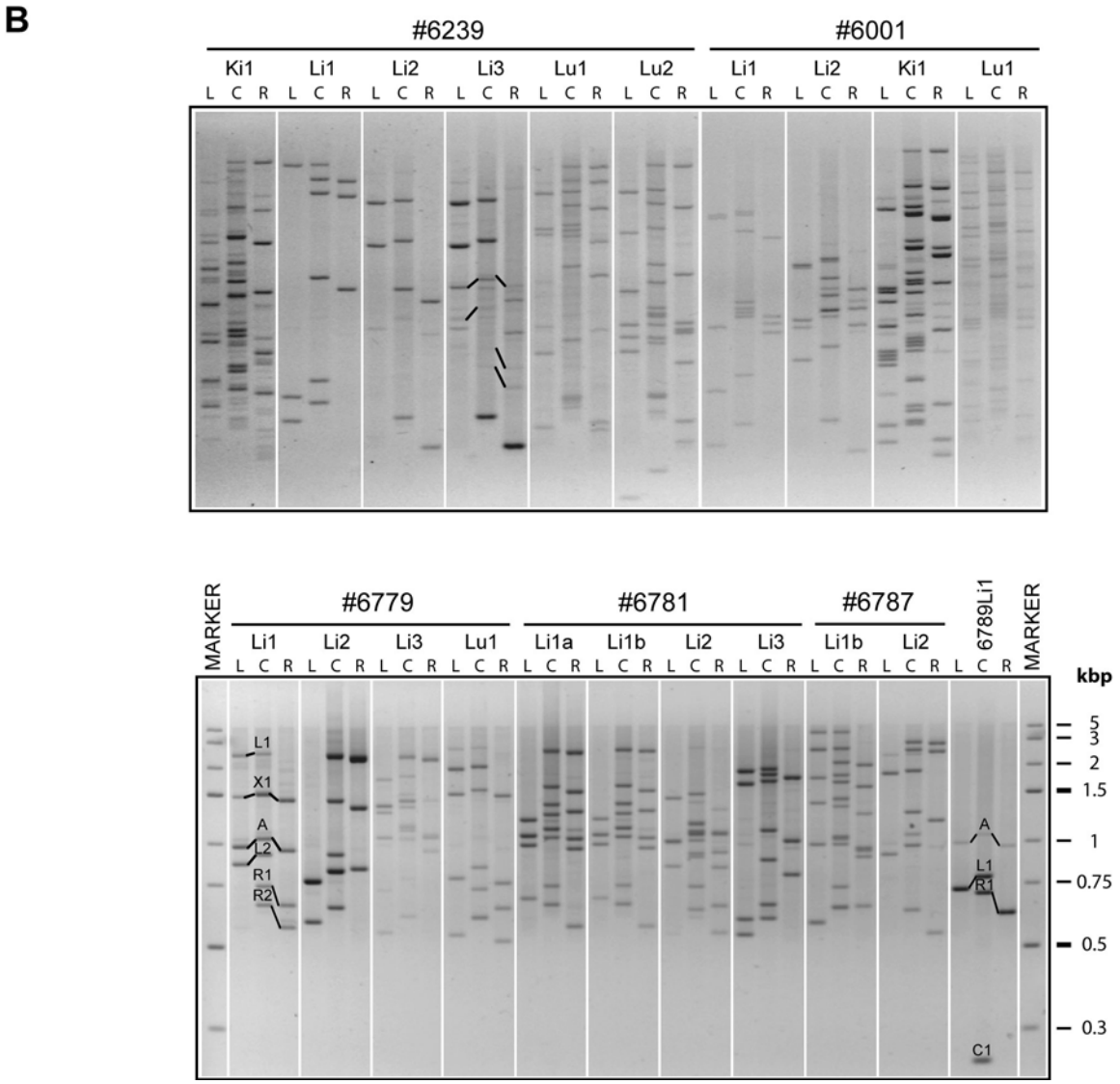

B
#6239
#6001
Ki1
Li1
Li2
Li3
Lu1
Lu2
L C R
MARKER
#6779
#6781
#6787
Li1a
Li1b
6789Li1
kbp
5
3
2
1.5
1
0.75
0.5
0.3
L1
X1
A
L2
R1
R2
C1

Fig. S2. Mapping of provirus insertion sites in liver tumors.

A. Scheme of iPCR. Long terminal repeats of the integrated provirus (LTRs) are shown as rectangles divided into U3, R and U5 parts. UT5 and UT3 are untranslated parts of the provirus adjacent to 5′ and 3′LTR, resp. Wavy lines represent chicken genomic sequences. iPCR fragments corresponding to left (5′) and right (3′) provirus/host junctions were produced by combined BstY and BclI digestion, circularization by self-ligation with T4 DNA ligase, linearization with ApaLI, Eco81I or SphI and finally PCR. Of note, the proviral enhancer (box inside the LTR box) is a necessary component of insertional activation and should be retained even in the most defective cancer gene-activating proviruses.

B. Examples of electrophoretic separation of iPCR products. Three iPCR reactions were carried out for each sample using different combinations of primers and providing for left junctions (lanes L; primers L + 5′), right junctions (lanes R; primers R + 3′) or both (lanes C; primers L + R). A typical VIS is represented by a pair of iPCR fragments: in lane C plus lane L or R; fragments in lanes L and R are shorter by 59 bp and 87 bp, resp. Exceptions reflecting the existence of defective proviruses occurred occasionally. Proviruses with deleted/mutated sequence complementary to primers 5′ or 3′ provided a fragment only in lane C (sample 6789Li1). Rearranged proviruses in which an LTR was surrounded on both sides by viral sequences (UT3 and UT5) produced iPCR fragments in all three lanes (labeled with X, sample 6779Li1). Similar arrangement is also present in non-nintegrated circular proviruses containing either one LTR or two LTRs in tandem (42) that manifested itself by ca. 1kb product of variable intensity in all three lanes in most samples (labeled with A - artificial). Of note, the proviral enhancer (the box inside the LTR box) is the necessary component of insertional activation and should be retained even in the most defective cancer gene-activating proviruses. **Upper picture:** Comparison of iPCR patterns in different tumor types in the same animal. Li – liver tumors, Lu – lung tumor, Ki – kidney tumor. Sample 6239Li2 is a product of tumor invasion from sample 6239Li3; junction fragments unique for 6239Li3 are indicated. **Lower picture:** Comparison of iPCR patterns in multiple liver tumors in the same animal. Samples 6781Li1a and Li1b are distant parts of the same tumor. The system of junction fragments labeling is suggested in samples 6779Li1 and 6789Li1.

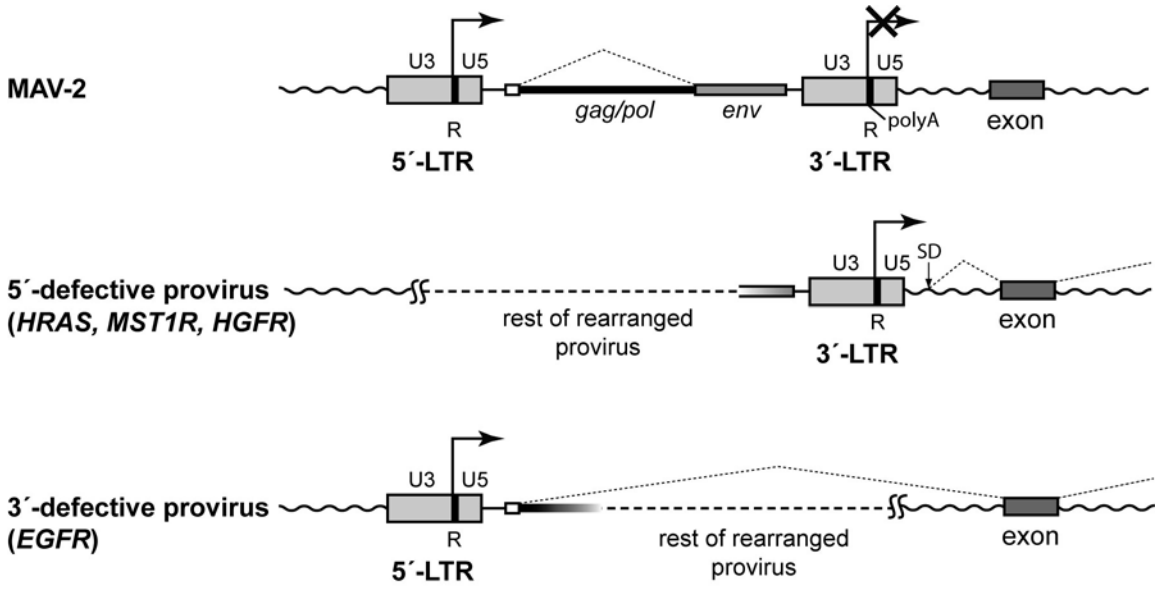


Fig. S3. Modes of insertional activation of candidate cancer genes. Structure of MAV-2 provirus plus sites of start (bent arrow), splicing (dotted lines) and termination (polyA) of its transcript are shown (upper drawing). The mechanism of formation of hybrid MAV-2/cancer gene mRNAs is suggested (lower drawings). The activity of the 3´LTR promoter in a non-defective provirus is blocked by transcriptional interference (20). Efficient activation of the adjacent gene is achieved either through deletion/rearrangement of 5´part of the provirus (thus releasing 3´LTR promoter from transcriptional interference) or by deletion of 3´LTR (thus eliminating proviral polyA signal). Structure of many cancer gene-activating proviruses is far more complex than on this simplified drawing; e. g. nearly all *EGFR*-activating proviruses carry additional viral sequences upstream of 5´LTR that contain 3´LTR oriented in the opposite direction. SD denotes an intronic cryptic splice donor site (or the authentic one in the case of MST1R).

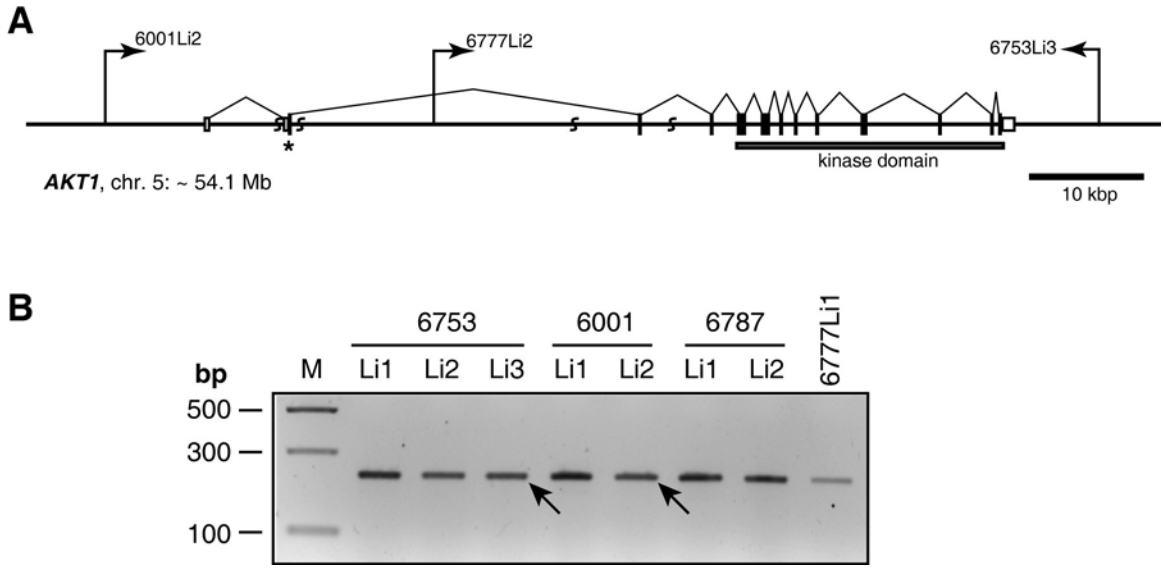


Fig. S4. Provirus insertions in AKT1 locus

A. Exon-intron structure of *AKT1* and position of provirus insertions. Design and symbols are as in Fig. 3.

B. Semiquantitative *AKT1*-specific RT-PCR. A pair of the *AKT1* exons-derived specific primers was used to detect all *AKT1* mRNAs. No increase in *AKT1* expression can be seen in samples carrying provirus insertions in the locus (marked by arrows).