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

Petr Pajer,1 Vladimir Pečenka,1 Jarmila Králová,1 Vít Karafiát,1 Dana Průková,1 Zdena Zemanová,2 Roman Kodet,2 and Michal Dvorák1

1Institute of Molecular Genetics AS CR, 1Institute of Physiology AS CR, and 2Institute 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 Dvorák, Institute of Molecular Genetics, Flemingovo nám. 2, Prague 6, 166 37 Czech Republic. Phone: 42-22018-3468; E-mail: mdvorak@img.cas.cz.

©2006 American Association for Cancer Research. 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 eVIS 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 chooroallantoic vein or ip., 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, the DNA pellet washed once with 70% 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 in a volume equal to 96% of 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).

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 plUC19 cloning vector (New England Biolabs, Beverly, MA) and its identity was verified by sequencing. A number of cVIS in addition to the 5′-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 eVIS 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 chooroallantoic vein or ip., 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/cm²). 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, electrophoresed, and Southern blotting were done by standard methods (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, glomeruli, and stromal cells (interstitium). Based on these characteristics, the samples were divided into three major classes (1-3, 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) 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 each was used as a template for additional 40 PCR cycles with nested primers LTR2 (5′-GGGCGAAGCTTGATCTCCCGTGA-3′) and NUP (SMART RACE cDNA amplification kit; Clontech), 250 mmol/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 BstYI and BclI restriction enzymes, self-ligated in a volume of 200 μL and linearized with the ApaI 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 mmol/L primers LTR2 and LTR3 (5′-GGTGCAATCGGGAATCCCTCCCTT- TTTGG-3′), 1.2 mmol/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 noncompositional 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 plaq1 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 plaq1 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 ≤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. 1A). 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. 1B 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. 1D). 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 largescale 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. 2A). 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

![Image](cancerres.aacrjournals.org)
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. 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 a splicing between the retroviral splice donor and 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).

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 BstY and BclI digestion generates fragments with compatible cohesive ends containing (a) the 5′-end of the provirus linked to a left-flanking fragment and (b) the 5′-end of the provirus linked to a right-flanking fragment. Self-ligation generates circular DNAs, which are then linearized by ApaLI. 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).

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 nephroblastomas. 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 ≤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. 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 3A 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 3B 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**

<table>
<thead>
<tr>
<th>Candidate locus</th>
<th>Candidate genes</th>
<th>Samples</th>
<th>Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nal 1-17</td>
<td>Hyal-2</td>
<td>WCC939.1a, WCB030.1d</td>
<td>1</td>
</tr>
<tr>
<td>Nal 1-19</td>
<td>POI2/OTF1/OCT1</td>
<td>WCB789.1b, WCB818.2a</td>
<td>1</td>
</tr>
<tr>
<td>Nal 1-85</td>
<td>spry-2</td>
<td>BL101.1a</td>
<td>2, 5</td>
</tr>
<tr>
<td>Nal 1-106</td>
<td>twist</td>
<td>WCB819.2c, WCB030.2a</td>
<td>1</td>
</tr>
<tr>
<td>Nal 1-145</td>
<td>dynein</td>
<td>BL344.2a, BL326.1c</td>
<td>1, 2, 4</td>
</tr>
<tr>
<td>Nal 1-187</td>
<td>twist</td>
<td>WCB813.1a</td>
<td>5</td>
</tr>
<tr>
<td>Nal 2-28</td>
<td>twist</td>
<td>BL102.2b, BL107.1a, BL122.1a, BL395.1a, WCB030.1b (10)</td>
<td>1, 2, 3, 4</td>
</tr>
<tr>
<td>Nal 2-88</td>
<td>FoxP1</td>
<td>BL395.1b, WCC036.2a</td>
<td>2</td>
</tr>
<tr>
<td>Nal 2-104</td>
<td>FoxP1</td>
<td>WCB039.1d, WCC939.1a</td>
<td>1</td>
</tr>
<tr>
<td>Nal 2-110</td>
<td>plag1</td>
<td>BL304.1a, BL338.2a, BL384.2a, BL410.2c, WCB789.1c, WCB813.1a</td>
<td>1, 2, 4</td>
</tr>
<tr>
<td>Nal 2-118</td>
<td>abf1</td>
<td>BL326.1c, BL819.1a</td>
<td>2</td>
</tr>
<tr>
<td>Nal 2-132</td>
<td>nov</td>
<td>BL435.1b, X59284 (13)</td>
<td>2, 3, 4</td>
</tr>
<tr>
<td>Nal 2-145a</td>
<td>kiaa0196, fij32440</td>
<td>WCC037.2b, WCC492.2a</td>
<td>1, 4</td>
</tr>
<tr>
<td>Nal 2-145b</td>
<td>enc1, fbx032</td>
<td>WCB030.2c, WCB042.1b, WCC939.1c</td>
<td>1</td>
</tr>
<tr>
<td>Nal 3-30</td>
<td>LOC116228</td>
<td>WCB030.1e, BL378.2a</td>
<td>1</td>
</tr>
<tr>
<td>Nal 5-13</td>
<td>c-Ha-ras</td>
<td>BL389.2a, BL821.1a, X03578 (21)</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Nal 6-28</td>
<td>WCC822.2b, WCC850.1c</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nal 12-15</td>
<td>foxP1</td>
<td>WCC036.2a, WCC037.1a, WCC814.1d, WCC850.1a, WCB826.1a</td>
<td>1, 2, 4</td>
</tr>
</tbody>
</table>

**NOTE:** Candidate loci were selected as described in Results, based on the following criteria: 1, multiple VISs within 20 kb; 2, multiple hits in a defined large gene locus (VIS distance >20 kb); 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.
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 NH2-terminal domain is missing. It has been shown that NH2-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 NH2-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. 4B) 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 nephroblasomas. 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. 4C). 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.](image-url)
Common Altered Genes in Human and Chicken Renal Tumors

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: AV0Z50520514 awarded by the Academy of Sciences of the Czech Republic.

References

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 Sárka Takařová in the manuscript preparation is also greatly acknowledged.
Identification of Potential Human Oncogenes by Mapping the Common Viral Integration Sites in Avian Nephroblastoma

Petr Pajer, Vladimír Pecenka, Jarmila Králová, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/66/1/78

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2006/01/16/66.1.78.DC1

Cited articles
This article cites 34 articles, 16 of which you can access for free at:
http://cancerres.aacrjournals.org/content/66/1/78.full.html#ref-list-1

Citing articles
This article has been cited by 8 HighWire-hosted articles. Access the articles at:
/content/66/1/78.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.