Induction of hepatocellular carcinoma by in vivo gene targeting

Pei-Rong Wang1, Mei Xu1, Sara Toffanin3, Yi Li1, Josep M. Llovet4,5, and David W. Russell2,6,7,

1Department of Medicine and 2Department of Biochemistry, University of Washington, Seattle, WA 98195; 3Mount Sinai Liver Cancer Program, Division of Liver Diseases, Mount Sinai School of Medicine, New York, NY 10029; 4Hepato-Oncology Group, Department of Surgery and Experimental Oncology, National Cancer Institute, 20133 Milan, Italy; 5Hepatocellular Carcinoma Translational Research Laboratory, Barcelona Clinic Liver Cancer Group, Liver Unit, Hospital Clinic, Institut d’Investigacions Biomèdiques August Pi i Sunyer, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas, University of Barcelona, 08036 Catalonia, Spain; and 6Institució Catalana de Recerca i Estudis Avançats, Barcelona, 08028 Catalonia, Spain

Edited by Neal G. Copeland, Methodist Hospital Research Institute, Houston, TX, and approved April 9, 2012 (received for review October 17, 2011)

The distinct phenotypic and prognostic subclasses of human hepatocellular carcinoma (HCC) are difficult to reproduce in animal experiments. Here we have used in vivo gene targeting to insert an enhancer-promoter element at an imprinted chromosome 12 locus in mice, thereby converting ~1 in 20,000 normal hepatocytes into a focus of HCC with a single genetic modification. A 300-kb chromosomal domain containing multiple mRNAs, snoRNAs, and microRNAs was activated surrounding the integration site. An identical domain was activated at the syntenic locus in a specific molecular subclass of spontaneous human HCCs with a similar histological phenotype, which was associated with partial loss of DNA methylation. These findings demonstrate the accuracy of in vivo gene targeting in modeling human cancer and suggest future applications in studying various tumors in diverse animal species. In addition, similar insertion events produced by randomly integrating vectors could be a concern for liver-directed human gene therapy.

Existing animal cancer models are limited in their ability to analyze specific oncogenic mutations as they naturally occur within rare cells present in normal individuals. Many types of vectors have been used for transgene delivery or insertional mutagenesis, but these methods do not reproducibly introduce defined chromosomal changes. Transgenic animals have been engineered to undergo precise chromosomal modifications, but these typically occur in an entire organism or class of cells (for example, by tissue-specific Cre-mediated recombination). Here we describe an alternative method for generating tumors through in vivo gene targeting, which introduces a specific chromosomal mutation into a subset of cells, allowing genetically identical tumors to develop within an otherwise normal organ.

Hepatocellular carcinoma (HCC) is a major worldwide health problem, with over 750,000 new cases diagnosed each year (1). Specific subclasses of HCC have been identified based on pathological features and gene-expression patterns, with prognostic implications (2, 3). Several mouse HCC models exist, including chemical tumor induction and transgenic strains that express oncogenes, growth factors, or viral genes (4). In general, the relationship of these mouse models to specific human HCC types is not clearly established, and a global gene-expression analysis of multiple mouse HCC models only showed that some were similar to either better or poorer survival groups of human HCCs (5). In vivo gene targeting may more accurately mimic a particular human HCC subclass by reproducibly introducing specific oncogenic mutations into a small number of hepatocytes in normal animals.

We previously found that mice that received a randomly integrating adenovirus-associated virus (AAV) vector developed HCCs that contained integrated vector genomes within the Rian gene (6), suggesting that these particular integration events somehow led to HCC. A later study of sleeping beauty transposition also found HCCs with integrations at this locus (7). Both studies highlight the potential genotoxicity of vector integration in hepatocytes, but their significance remains controversial because other reports have shown that animals do not develop HCC after AAV vector injections (8–10). The integration site locus contains a complex set of imprinted genes that are uniquely dysregulated after reprogramming to pluripotency (11), and two noncoding RNAs (Rian and Mirg) that contain multiple snoRNAs and microRNAs (12). These small RNAs could regulate a large number of target genes, and the human homologs of some of these microRNAs have been proposed to both stimulate and inhibit tumorigenesis in other types of malignancies (13–16). Two recent studies found that a subset of human HCCs have elevated expression of this microRNA cluster (17, 18).

Here we have studied the role of this locus in HCC formation by intentionally introducing a promoter-enhancer element into the mouse Rian gene through in vivo gene targeting with AAV vectors. Previous studies have shown that in addition to their potential for random, nonhomologous integration, AAV vectors can efficiently and accurately introduce mutations into homologous chromosomal target sequences (19). This process occurs in ~1/106 hepatocytes after in vivo vector delivery to the liver (20, 21). We reasoned that this gene-targeting frequency would be adequate to initiate multiple foci of HCC because of dysregulated gene expression after targeted promoter-enhancer insertion. Here we show that this occurs, and we describe the development of these tumors, their gene expression patterns, and their similarity to a specific subclass of human HCC.

Results

Gene-Targeted Liver Cells Form HCCs. We constructed an AAV gene targeting vector to introduce a “CAG” enhancer/promoter consisting of the CMV enhancer and chicken β-actin promoter-intron fragment into intron 2 of the mouse Rian gene where prior nonhomologous integration events were associated with liver tumors (6, 7) (Fig. 1). For this process, 3 × 1010 genome-containing vector particles were injected into newborn C57BL/6J mice via the temporal vein and cohorts were analyzed at 6 mo, 10 mo, and the time of killing because of tumor growth (11–18 mo of age). All of the vector-injected male mice from each cohort had liver tumors, but there was a relative delay in tumor formation in female mice, with 50%, 75%, and 100% of females developing tumors at 6 mo, 10 mo, and the time of being killed, respectively. The delay in females may have been due to decreased transduction (22) or hormonal effects on HCC development (23). The number of tumors increased with time, as did liver weight, also with a delay in females (Fig. 1 B and C).


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1P.-R.W. and M.X. contributed equally to this work.

2To whom correspondence should be addressed. E-mail: drussel@u.washington.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1117032109/-/DCSupplemental.
These findings were reflected in the shortened life span of injected mice (Fig. 1D). Peripheral blood analysis was consistent with the progressive development of liver cancer, showing elevated bilirubin and liver enzyme levels, and decreased albumin and total protein levels (Fig. S1). None of the 32 control littersmates that did not receive vector developed tumors. We also examined the spleen, pancreas, lungs, brain, kidney, heart, prostate, testes, skeleton, large muscles, and gastrointestinal tract of each vector-injected mouse and did not observe primary tumors in these organs. Although this could be due to a liver-specific effect of targeting this locus, it may also reflect the liver-tropic nature of the vector after intravenous delivery.

Several tumors were dissected from these injected animals along with adjacent normal tissue for molecular analyses (Tables S1 and S2). Quantitative PCR (qPCR) with one primer in the CAG promoter and another in flanking mouse genomic DNA outside the region of vector homology showed that the tumors contained an average of 0.6 targeted Rian alleles per diploid genome, but adjacent normal tissue had much lower levels (Fig. 1E). This finding was confirmed by Southern blot analysis in a subset of tumors (Fig. 1F and Table S1). These values are consistent with the presence of at least one targeted allele in every tumor cell, because mouse hepatocytes are usually polyploid (24), and the samples also contain DNA from other cell types, such as endothelial cells and Kupffer cells (25). The low levels of targeted alleles in adjacent normal tissue are presumably due to infiltration by tumor cells not appreciated on gross dissection. We also measured the total amount of vector genomes in the same samples by using two qPCR primers in the CAG promoter. These values were two- to threefold higher than the copy numbers of targeted alleles (Fig. 1E, and Tables S1 and S2), which could be because of minor inaccuracies in quantitation, additional random vector integrants, episomal vector genomes, or integration of vector multimers at the Rian locus.

Gross inspection revealed multiple nodules in the livers of vector-injected mice, with more present in males (Fig. 2A). Histological analysis showed that the nodules were HCCs containing abnormal hepatocytes with irregular nuclei, mitotic figures, and giant cells (Fig. 2B). Many of the tumors had a trabecular pattern and vascular invasion was frequently observed (Fig. 2B). Staining with antibodies showed that the tumors overexpressed Pik3ca (Fig. 2C), which is the p110α catalytic subunit of class I PI3K and
a key component of the PI3K/Akt signaling pathway often associated with HCC progression (26). Epithelial cell adhesion molecule (Epcam) was also expressed in the tumor cells, which is a characteristic of some types of HCC (27). Epcam expression was predominantly nuclear in tumors (Fig. 2C), but membranous in the biliary epithelial cells present in normal liver, consistent with the proposed mitogenic role of the Epcam intracellular domain after its nuclear translocation (28). The tumor nodules had increased DNA synthesis compared with adjacent normal tissue based on their uptake of BrdU (Fig. 2C). In one animal, lung metastases of HCC were found that contained targeted Rian genes (Fig. S2). These findings are all consistent with invasive, multifocal HCC produced by homologous recombination and promoter insertion at the Rian locus.

Human HCC often results in elevated serum α-fetoprotein (AFP) levels. Although reagents were not available to measure serum AFP levels in mice, we were able to show that the HCC nodules expressed AFP by staining liver sections with an anti-AFP antibody (Fig. 3A). Adjacent normal tissue did not express AFP. Interestingly, we also found small foci of apparently normal AFP+ hepatocytes that were only present in vector-injected mice (Fig. 3B). The growth of AFP+ foci over time suggested that these small, normal foci were the precursors of HCCs, because they were replaced by large malignant foci at later timepoints (Fig. 3C). The total frequency of AFP+ foci remained relatively constant, ranging from 0.046 to 0.12 per square millimeter (corresponding to 2.4–6.4 × 10^−5 hepatocytes). This value is similar to the gene targeting frequency of ~0.03 foci/mm^2 measured previously for an integrated lacZ reporter gene or a mutant β-Glucuronidase gene in mice that received an equivalent AAV targeting vector injection (20), suggesting that each Rian targeting event leads to an AFP+ focus. The data also suggest that each targeted AFP+ hepatocyte eventually forms a focus of HCC, because all AFP+ foci of >1,000 cells contained malignant, abnormal hepatocytes.

**Gene Expression Analysis.** We examined the expression of genes flanking the vector insertion site by quantitative RT-PCR (qRT-PCR) using exonic primers, and found 10- to 20-fold increases in mRNA transcripts upstream of the vector-encoded CAG promoter (Dlk1, Meg3, and Rian-upstream), and 500- to 800-fold increases in downstream transcripts (Rian-downstream and Mirg) (Fig. 4A). This finding can be explained by modest CAG enhancer-dependent activation of both downstream and upstream chromosomal promoters, combined with significantly higher levels of transcription initiating at the CAG promoter present within the Rian gene and extending into downstream regions including the Mirg gene. We found evidence for such fusion transcripts by sequencing of RT-PCR products (Fig. S3). Transcription of the Wdr25, Bega7, and Dio3 genes was not significantly changed, so the direct cis-acting effects of gene targeting did not extend to these more distant genes located 650–750 kb away. The Rian and Mirg genes contain multiple microRNA genes within their introns, the levels of which were assayed by microRNA array analysis. This finding showed that 18 of the 696 microRNAs interrogated by the array were expressed at >twofold higher levels in tumors, and all of these were transcribed from the Rian-Mirg locus (Fig. 4B). Several microRNAs located upstream of the Rian gene were not overexpressed (Fig. S4). There were also significant increases in the expression of snoRNAs present at the targeted locus as measured by qRT-PCR (five of six snoRNAs examined) (Fig. 4B). These results are consistent with increased transcription through the Rian-Mirg gene cluster in gene-targeted loci, including the intergenic region between Rian and Mirg, coupled with RNA processing into mature microRNAs and snoRNAs.

We compared the global mRNA expression patterns of gene-targeted tumors and surrounding normal tissue by microarray analysis: 199 genes were ≥ twofold up-regulated in tumors and 100 genes were ≥ twofold down-regulated (Tables S3 and S4). Gene ontology analysis (Fig. 4C) showed that the up-regulated genes were enriched in classes associated with cell proliferation (e.g., cell cycle, DNA replication, and cell division) as expected for tumor cells, and in classes associated with diverse metabolic pathways (hexose, energy, and peptide metabolism). These metabolic pathways do not have clear oncological significance, and their altered expression may simply reflect the abnormal function of transformed hepatocytes. Specific up-regulated genes with relevance for HCC included: Dlk1 from the insertion site, a marker of fetal hepatic progenitor cells and HCC (29–31); Nfkb1, Nfk2, Fos, and Hspb1, which regulate the MAPK signaling pathway that is frequently activated in HCC (32); H19 and Igf2, which are imprinted genes from a chromosome 7 locus associated with HCC development (33); and Birc5 (Survivin), an inhibitor of apoptosis (34). Genes down-regulated in tumors (Fig. 4D) were enriched in classes associated with other types of metabolism (benzene and fatty acid metabolism) and with inflammation and inflammatory response, complement activation, and cytokine responses). The latter class could allow tumor cells to escape destruction by immune and inflammatory cells.

A central question is how overexpression of the small RNAs at the targeted locus leads to global changes in gene expression, and which RNAs are the direct targets of these regulatory RNAs. We approached this question by an in silico analysis of potential
microRNA binding sites in the dysregulated genes found in tumors, with the expectation that in some cases the mRNA targets of overexpressed microRNAs would have reduced expression levels. DIANA-miRExTra analysis (35) showed that a total of eight microRNAs have statistically robust predicted binding sites on at least one of the 62 down-regulated mRNAs processed by the program. Four of these microRNAs are expressed from the Rian-Mirg locus and have binding sites on 36 of these 62 mRNAs (58%) (Table S5). In contrast, zero or one of the Rian-Mirg microRNAs had predicted effects on the 62 most up-regulated and unchanged mRNAs, respectively, suggesting that this in silico analysis has had predicted effects on the 62 most up-regulated and unchanged mRNAs (58%) (Table S5). In contrast, zero or one of the Rian-Mirg microRNAs had predicted effects on the 62 most up-regulated and unchanged mRNAs (58%) (Table S5).

Gene Ontology Analysis of Dysregulated Genes. Analysis of gene expression was performed using the DAVID Gene Ontology Tool (34). DIANA-miRExTra analysis (35) showed that a total of 6 snoRNAs and 42 microRNAs transcribed from the insert site were determined from the same samples as in A by qRT-PCR and microarray analysis, respectively. The four microRNAs with binding sites in down-regulated genes (Table S5) are indicated by asterisks (miR-369, miR-376, miR-134, and miR-758). C (C and D) Gene ontology analysis of dysregulated genes (two-fold or greater change, \( P < 0.05 \)), in comparison with all genes present in the mouse genome.

Gene-Targeted HCCs Model a Specific Type of Human HCC. Humans have an imprinted locus on chromosome 14 syntenic to mouse chromosome 12, and molecular profiling has identified a poor prognosis subclass of human HCCs (type C3) characterized by increased microRNA expression from this locus (17, 18). These human C3 HCCs produced high serum AFP levels, expressed DLK1 and an Epcam-related gene signature, activated PI3K/AKT signaling, and exhibited frequent vascular invasion, demonstrating their similarity to mouse tumors produced by Rian gene targeting. We compared human C3 HCCs to normal human liver samples and found that the microRNAs, snoRNAs, and microRNAs present in this locus were overexpressed as a single domain flanked by the BEGAIN and DIO3 genes, just as in the mouse tumors (Fig. 5 A and B). An analogous comparison with other types of human HCCs demonstrated the specificity of these changes for the C3 subclass of HCC (Fig. S6).

To understand how the human locus could be activated in HCC, we analyzed copy number variation and epigenetic changes in 103 human HCC samples. A single non-C3 HCC sample that did not overexpress the snoRNA-microRNA cluster had increased copy number in this region, but no abnormalities were found in the four C3 HCC samples analyzed (Fig. 5C). A CpG island is present within the human RTL1 gene located upstream of MEG3 and the syntenic insertion site, and it is methylated in normal hepatocytes. Bisulfite sequencing showed that three of the five C3 HCC samples had decreased cytosine methylation in this region (Fig. 5D), suggesting that in some cases epigenetic changes could have altered gene expression at this locus. This CpG island contains a predicted promoter element and transcription start site. Methylation levels at two nearby CpG islands in the MEG3 and DLK1 promoters were unchanged in C3 HCC samples.
**Discussion**

Here we have shown that normal mice that received an intravenous injection of an AAV gene targeting vector developed HCCs containing a precise promoter/enhancer insertion at the *Rian* locus. This in vivo gene targeting approach reproducibly activated a 0.3-Mb imprinted domain of chromosome 12 that is not normally expressed in adult hepatocytes, and produced more than 30 genetically identical, macroscopic HCC foci per liver. Ultimately, all mice that received the vector developed terminal liver failure.

Vector-treated mice initially contained small foci of normal-appearing Afp+ hepatocytes that were subsequently replaced by larger foci of HCC (Fig. 3). The frequency of Afp+ foci was similar to that observed when targeting other genes in prior AAV experiments (20, 21), suggesting that each targeting event produced a single Afp+ hepatocyte that proliferated to form an HCC focus. If additional spontaneous oncogenic mutations were required they must have consistently occurred before each ~1,000 cell malignancy developed. Nonhomologous vector integration events at other oncogenic loci were also unlikely, because on average each targeted cell contained at most one or two additional vector copies that could have been present at random locations (Fig. 1E). Therefore, promoter insertion at the *Rian* locus appears to be sufficient to transform a normal hepatocyte. This locus is ideally suited for single-step transforming mutations, because the many snoRNA and microRNA genes that contain could each have multiple possible targets, and there are also protein-coding genes that could be oncogenic, such as *Dlk1* encoding a marker of fetal hepatoblasts (29). A requirement for coordinated, simultaneous activation of multiple genes may explain the tight, central clustering of integrants observed in HCCs formed by insertional mutagenesis at this locus (Fig. 1A). A recent report showed that c-Met overexpression in transgenic mice can also induce Afp+ HCCs that express Dlk1 and a subset (23 of 42) of the *Rian-Mirg* microRNAs (18). However, in this model additional secondary mutations are required for HCC formation (39), and c-Met was not overexpressed in the mouse tumors we produced by gene targeting (Table S3).

Our experiments provide a glimpse into the complex genetic changes induced by this single chromosomal alteration. The phenotype of gene-targeted hepatocytes suggests that they in part reverted to a more fetal developmental stage, with Dlk1 and Afp expression, as well as activation of the target locus snoRNA and microRNA genes that are normally expressed during embryogenesis (40, 41). A combination of in silico analysis and global gene expression profiling identified a set of down-regulated genes that are likely targets of the microRNAs overexpressed after gene targeting. This in silico analysis is certainly incomplete, because we applied rigorous criteria to identify target binding sites, and many of the microRNAs presumably regulate their targets at the translational level. The snoRNAs present at the target locus could also have profound effects on gene expression, including global changes in translation and splicing because of modification of their known rRNA and snRNA targets, as well as other effects related to as yet unidentified target RNAs. Although snoRNAs have not been shown to play a role in HCC, orphan snoRNAs can regulate specific mRNA targets (42), suggesting that analogous mRNA targets of *Rian-Mirg* snoRNAs could be contributing...
to HCC formation. Although it would be interesting to establish the potential risks of human gene therapy can also be addressed. Several ongoing or planned clinical trials involve the delivery of AAV vectors to the liver (43, 45), and future gene-targeting experiments could help identify the safest promoter/enhancer elements to be used in these types of trials.

Materials and Methods

See SI Materials and Methods for details regarding vector production, animal care, human samples, histology, DNA isolation and analysis, RNA isolation, qRT-PCR of mRNAs and snRNAs, copy number analysis, methylation analysis, microRNA analysis, microarray analysis of mRNAs, and online microarray datasets. See Table S7 for a list of primers used. All animal procedures were approved by the University of Washington Animal Care and Use Committee.

ACKNOWLEDGMENTS. We thank E. Olson for adeno-associated virus stock preparation, Y. Jiang for help with mouse injections, and Y. Hoshida for help with microarray analysis. This research was funded by National Institutes of Health Grants AR48328 and DK55759 (to D.W.R.) and DK76986 (to J.M.L.), as well as Grants 259744-2 from the FFP-2010-Health European Commission, SAF2010-16055 from the National Institute of Health of Spain, and the Samuel Waxman Cancer Research Foundation (to J.M.L.).


Wang et al.