Induction of hepatocellular carcinoma by in vivo gene targeting

Pei-Rong Wang[a,1], Mei Xu[a,1], Sara Toffanin[b,c], Yi Li[a], Josep M. Llovet[b,d,e], and David W. Russell[a,f,2]

[a]Department of Medicine and [f]Department of Biochemistry, University of Washington, Seattle, WA 98195; [b]Mount Sinai Liver Cancer Program, Division of Liver Diseases, Mount Sinai School of Medicine, New York, NY 10029; [c]Hepato-Oncology Group, Department of Surgery and Experimental Oncology, National Cancer Institute, 20133 Milan, Italy; [d]Hepatocellular Carcinoma Translational Research Laboratory, Barcelona Clinic Liver Cancer Group, Liver Unit, Hospital Clinic, Institut d'Investigacions Biomediques August Pi i Sunyer, Centro de Investigacion Biomedica en Red de Enfermedades Hepaticas y Digestivas, University of Barcelona, 08036 Catalonia, Spain; and [e]Institució 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.

adeno-associated virus | genotoxicity

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 adeno-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 ~$1/10^4$ 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*A*). For this process, 3×10^{10} 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*).

Author contributions: D.W.R. designed research; P.-R.W., M.X., S.T., and Y.L. performed research; P.-R.W., M.X., S.T., Y.L., J.M.L., and D.W.R. analyzed data; and P.-R.W., M.X., S.T., and D.W.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

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

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

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1117032109/-/DCSupplemental.

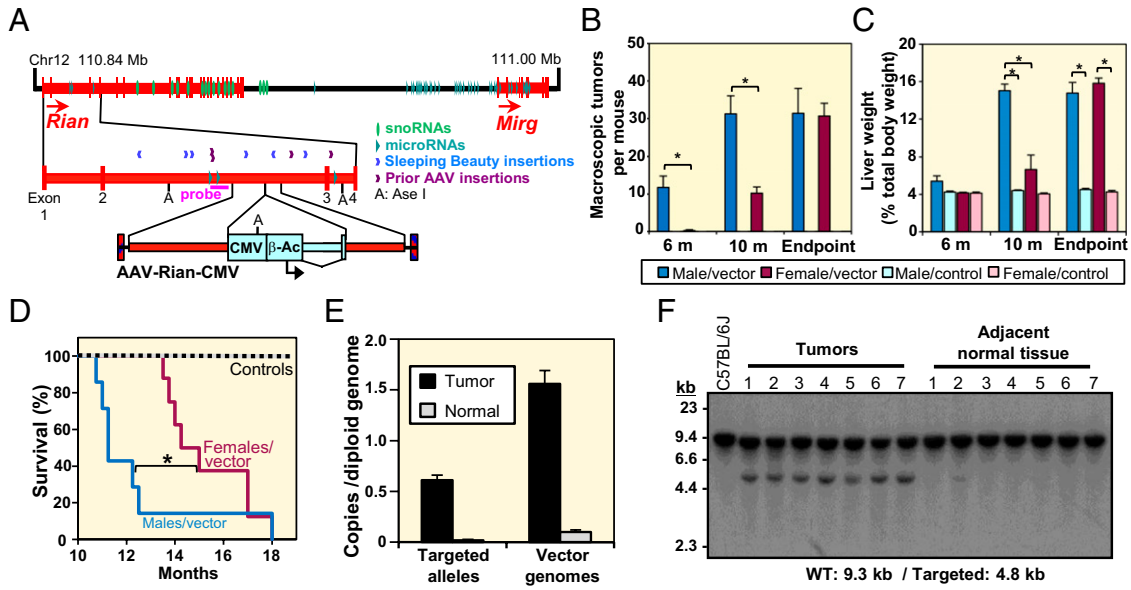

Fig. 1. *Rian* gene targeting induces liver tumors. (*A*) Map of chromosome 12 containing *Rian* and *Mirg* transcripts (exons in thick boxes) showing the target locus, AAV-Rian-CMV targeting vector, snoRNAs, microRNAs, Southern blot probe, and Ase I sites. The locations of sleeping beauty transposon and non-homologous AAV vector insertions found previously in tumors are indicated. (*B* and *C*) The number of tumors and liver weights are shown at 6 mo ($n = 4$ per group), 10 mo ($n = 4$ per group), and experimental endpoints (11–18 mo; $n \geq 7$ per group) for male and female mice (means ± SDs; $^*P < 0.05$ by two-tailed t test). (*D*) Survival of injected and control mice over time ($^*P < 0.01$ by Gehan-Breslow-Wilcoxon test). (*E*) Copy numbers of targeted alleles and vector genomes as determined by qPCR (means ± SEs). (*F*) Representative Southern blot of Ase I-digested genomic DNAs from seven tumors and adjacent normal tissues from vector-injected mice.

These findings were reflected in the shortened life span of injected mice (Fig. 1*D*). 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 littermates 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. 1*E*). This finding was confirmed by Southern blot analysis in a subset of tumors (Fig. 1*F* 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. 1*E*, and Tables S1 and S2), which could be because of minor inaccuracies in quantitation, additional random vector integrants, episomal vector genomes, or integrations 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. 2*A*). Histological analysis showed that the nodules were HCCs containing abnormal hepatocytes with irregular nuclei, mitotic figures, and giant cells (Fig. 2*B*). Many of the tumors had a trabecular pattern and vascular invasion was frequently observed (Fig. 2*B*). Staining with antibodies showed that the tumors overexpressed Pik3ca (Fig. 2*C*), which is the p110α catalytic subunit of class I PI3K and

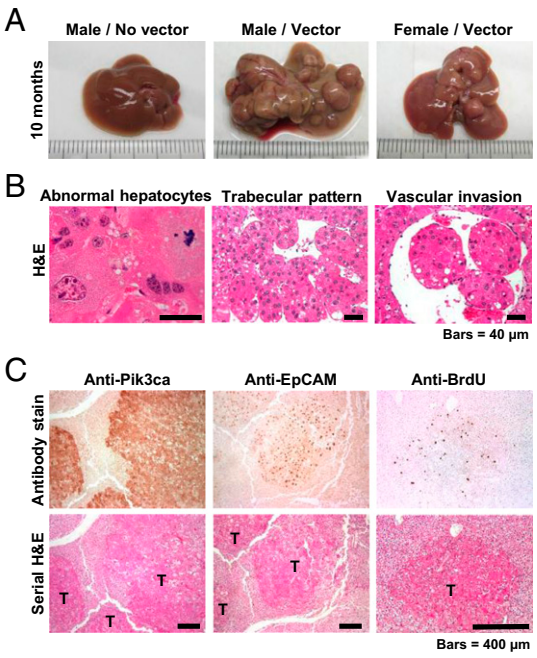

Fig. 2. Appearance and histology of liver tumors. (*A*) Photographs of gross liver specimens from 10-mo-old mice. (*B*) H&E staining of livers with HCC nodules from vector-treated mice. (*C*) Staining of liver sections with specific antibodies for Pik3ca, Epcam, or BrdU. Serial sections stained with H&E are shown below. T, tumor nodules.

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. 2*C*), 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. 2*C*). 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. 3*A*). 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. 3*B*). 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. 3*C*). 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.

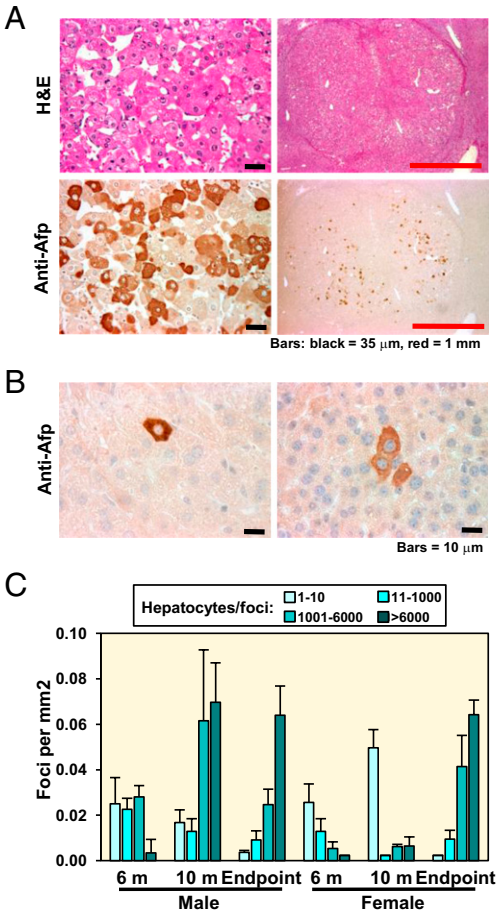

Fig. 3. Small Afp^+ foci transform into HCC. (*A*) Examples of Afp^+ tumors detected by anti-Afp antibody staining shown below serial H&E sections. (*B*) Small foci of Afp^+ hepatocytes within the normal liver tissue of mice injected with AAV-Rian-CMV. (*C*) The sizes of Afp^+ foci are shown over time for male and female mice injected with AAV-Rian-CMV (means ± SEs). The average liver surface area examined was 228 ± 60 mm^2 per mouse (n = 4 per group). No Afp^+ cells were detected in control mice.

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. 4*A*). 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*, *Begain*, 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. 4*B*). 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. 4*B*). 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. 4*C*) 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); *Ngfβ*, *Ntrk2*, *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. 4*D*) were enriched in classes associated with other types of metabolism (benzene and fatty acid metabolism) and with inflammation (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

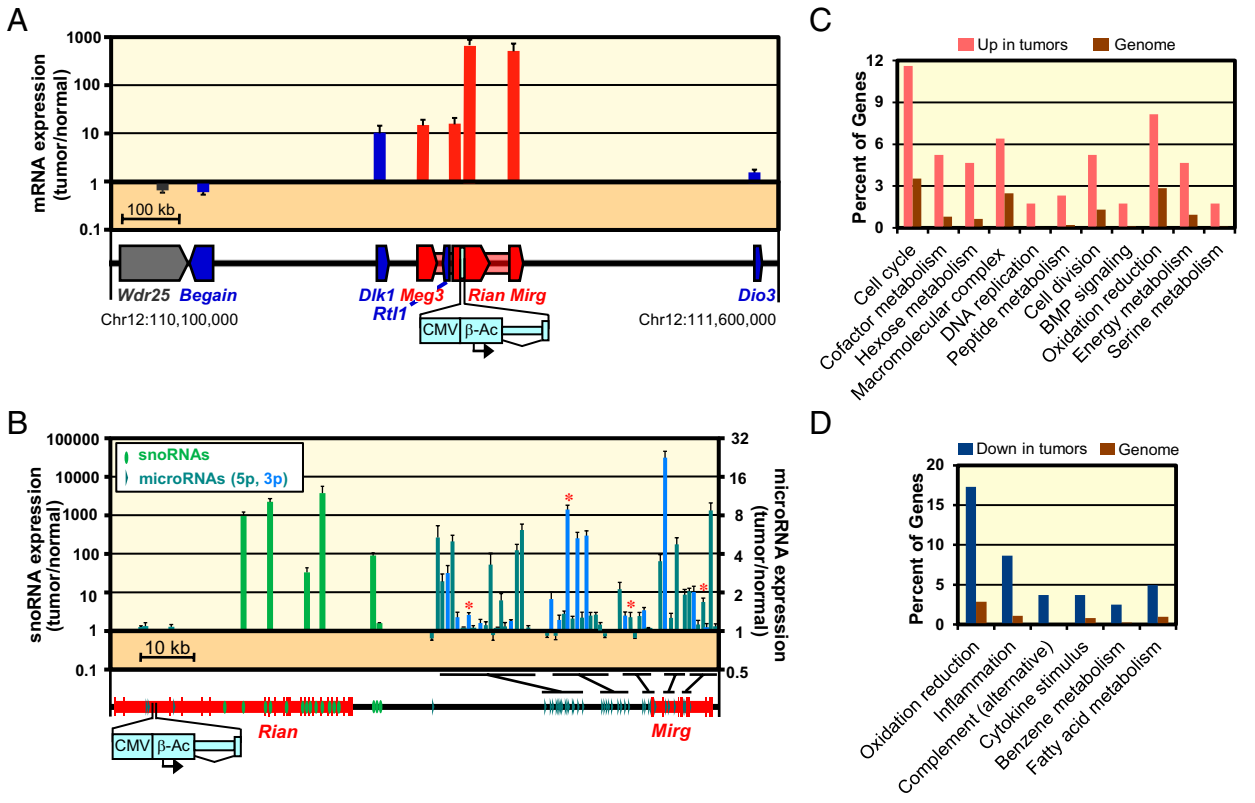

Fig. 4. Transcriptional effects of *Rian* gene targeting. (*A*) The expression levels of mRNA transcripts near the insertion site were determined quantitatively in tumors and compared with adjacent normal-appearing tissue by qRT-PCR ($n = 3$). Red and blue genes are imprinted RefSeq genes expressed from the maternal or paternal chromosomes respectively, with transparent red regions representing areas of non-RefSeq transcripts. (*B*) The expression levels of 6 snoRNAs and 42 microRNAs transcribed from the insertion site locus 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-369–5p, miR-376, miR-134, and miR-758). (*C* and *D*) Gene ontology analysis of dysregulated genes (twofold or greater change, $P < 0.05$), in comparison with all genes present in the mouse genome.

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 accurately identified at least some direct microRNA targets. Further analysis by miRanda (36) showed that 24 of the 36 down-regulated microRNA target mRNAs actually contain binding sites for multiple *Rian-Mirg* microRNAs (Fig. S5), highlighting potential redundancies in the roles of these microRNAs. This set of 36 target genes included several associated with inflammation and immunity (*Saa4*, *C8a*, *C8b*, *Susd4*, *Dclk3*, *Pigr*, and *Ccl9*), as was also noted above by gene ontology analysis (Fig. 4*D*).

The expression of snoRNAs present at the *Rian-Mirg* locus was highly elevated in gene-targeted HCCs (Fig. 4*B*). These small RNAs guide chemical modifications of other RNAs, including 2′-O-methylation of ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), and small nuclear RNAs (snRNAs). In total, 16 snoRNAs are present at the locus and 5 of these have predicted targets on the 18S rRNA subunit, 28S rRNA subunit, or U2 snRNA (37, 38) (Table S6). Overexpression of these snoRNAs could lead to increased target modification and corresponding effects on translation and splicing. The remaining 11 snoRNAs are not predicted to target rRNAs, tRNAs, or snRNAs, so these orphan snoRNAs could have unknown mRNA targets.

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 mRNAs, 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. 5*C*). A CpG island is present within the human *RTL1* gene located upstream of *MEG8* 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. 5*D*), 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.

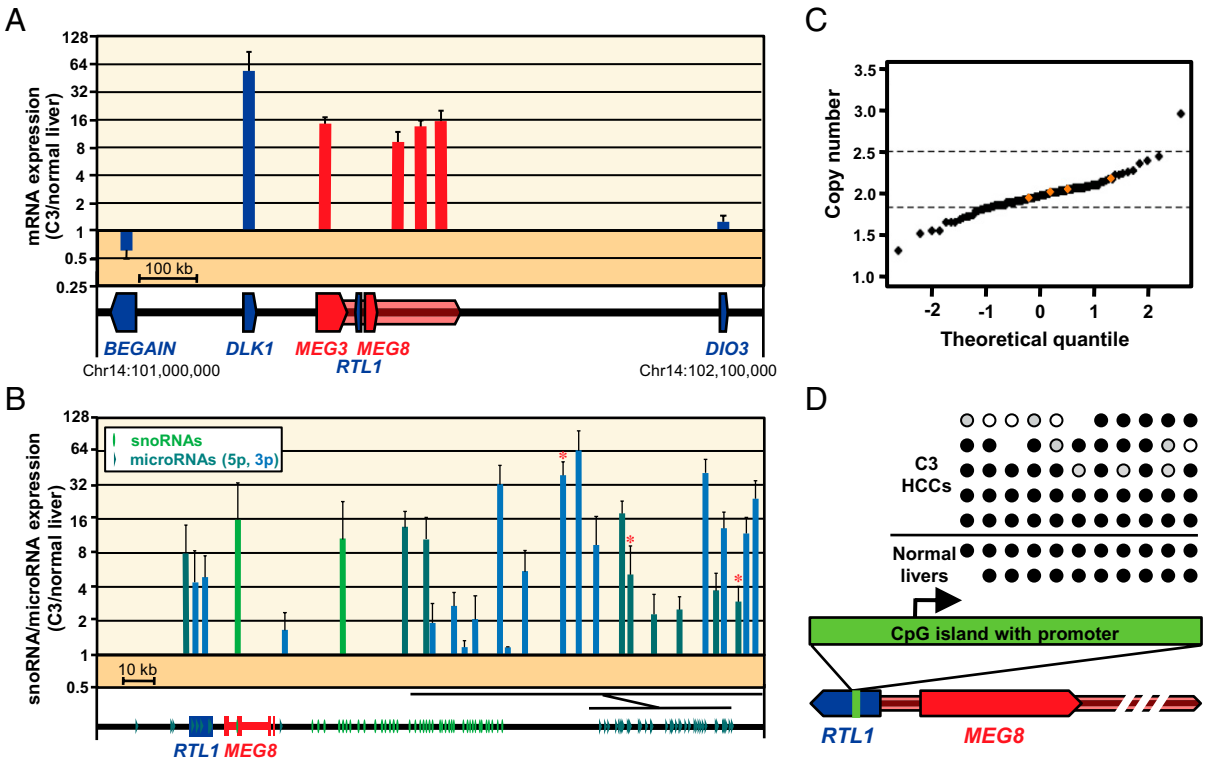

Fig. 5. Transcription profile of subclass C3 human HCCs. (*A*) The expression levels of mRNA transcripts near the insertion site were determined quantitatively in C3 HCC samples ($n = 5$) and compared with normal liver samples ($n = 10$) by microarray analysis. (*B*) The expression levels of two snoRNAs and 26 microRNAs transcribed from the insertion site locus were determined from the same samples as in *A* by qRT-PCR and microarray analysis, respectively. Human homologs of the microRNAs highlighted in Fig. 4*B* are marked with asterisks. (*C*) Copy number analysis as measured by SNP array. Each dot represents the mean copy number in a genomic region including *MEG3*, *RTL1*, *MEG8* and the snoRNA-microRNA cluster on human chromosome 14 in 103 human HCC samples plotted against standardized theoretical quantiles. Dashed lines show the range of control values from nonmalignant cirrhotic samples. Orange symbols indicate 4 C3 HCC samples. (*D*) Methylation status of 11 CpG sites present in a CpG-rich region is shown for five C3 HCC samples and two normal liver specimens. The location of the CpG Island in the *RTL1* gene and its predicted transcription start site (arrow) are indicated. CpG dinucleotides are represented as circles (black, methylated; white, unmethylated; gray, partial methylation). Missing circles were excluded because of aberrant base calls.

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. 1*E*). 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 it contains 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. 1*A*). 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 whether the paternal or maternal imprinted alleles were dysregulated by gene targeting, this could not be determined in the inbred mouse strain used in our experiments.

The HCCs produced by *Rian* gene targeting were remarkably similar to the C3 subclass of human HCCs identified by microRNA profiling (17, 18), demonstrating the accuracy of in vivo gene targeting for modeling human cancer. These similarities included an aggressive phenotype with vascular invasion, AFP and EPCAM expression, activated PI3K/AKT signaling, and an identical domain of increased gene expression within the imprinted *BEGAIN-DIO3* locus. Although insertion of an entire promoter/enhancer would be an unusual type of spontaneous mutation, epigenetic dysregulation could have produced the human tumors, because we observed reduced cytosine methylation within a CpG island adjacent to the syntenic insertion site in some C3 HCCs. Furthermore, our findings raise the possibility that unintended vector integration at the human locus might also lead to C3 HCC in gene therapy trials, despite numerous preclinical studies demonstrating safety in other animal models (8–10, 43).

The possible applications of in vivo gene targeting for modeling cancer are diverse. The complex locus targeted in this study can be further dissected by altering the location, orientation, and types of mutations introduced, or other candidate HCC loci could be targeted. The experiments can be extended to a variety of cell types and species, because multiple AAV serotypes are available for efficient in vivo delivery to different organs and animals (44). This availability may allow specific tumors to be studied in large animals that cannot otherwise be genetically manipulated, and provide outstanding models for developing human anticancer treatments. 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 snoRNAs, 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 FP7-2010-Health European Commission, SAF2010-16055 from the National Institute of Health of Spain, and the Samuel Waxman Cancer Research Foundation (to J.M.L.).

1. Parkin DM, Bray F, Ferlay J, Pisani P (2005) Global cancer statistics, 2002. *CA Cancer J Clin* 55:74–108.
2. Hoshida Y, et al. (2009) Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res* 69:7385–7392.
3. Boyault S, et al. (2007) Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. *Hepatology* 45:42–52.
4. Heindryckx F, Colle I, Van Vlierberghe H (2009) Experimental mouse models for hepatocellular carcinoma research. *Int J Exp Pathol* 90:367–386.
5. Lee JS, et al. (2004) Application of comparative functional genomics to identify best-fit mouse models to study human cancer. *Nat Genet* 36:1306–1311.
6. Donsante A, et al. (2007) AAV vector integration sites in mouse hepatocellular carcinoma. *Science* 317:477.
7. Dupuy AJ, et al. (2009) A modified sleeping beauty transposon system that can be used to model a wide variety of human cancers in mice. *Cancer Res* 69:8150–8156.
8. Bell P, et al. (2005) No evidence for tumorigenesis of AAV vectors in a large-scale study in mice. *Mol Ther* 12:299–306.
9. Nathwani AC, et al. (2011) Long-term safety and efficacy following systemic administration of a self-complementary AAV vector encoding human FIX pseudotyped with serotype 5 and 8 capsid proteins. *Mol Ther* 19:876–885.
10. Li H, et al. (2011) Assessing the potential for AAV vector genotoxicity in a murine model. *Blood* 117:3311–3319.
11. Stadtfeld M, et al. (2010) Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* 465:175–181.
12. Seitz H, et al. (2004) A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome Res* 14:1741–1748.
13. Meng F, Wehbe-Janek H, Henson R, Smith H, Patel T (2008) Epigenetic regulation of microRNA-370 by interleukin-6 in malignant human cholangiocytes. *Oncogene* 27:378–386.
14. Hwang-Verslues WW, et al. (2011) miR-495 is upregulated by E12/E47 in breast cancer stem cells, and promotes oncogenesis and hypoxia resistance via downregulation of E-cadherin and REDD1. *Oncogene* 30:2463–2474.
15. Saito Y, et al. (2006) Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell* 9:435–443.
16. Swarbrick A, et al. (2010) miR-380-5p represses p53 to control cellular survival and is associated with poor outcome in MYCN-amplified neuroblastoma. *Nat Med* 16:1134–1140.
17. Toffanin S, et al. (2011) MicroRNA-based classification of hepatocellular carcinoma and oncogenic role of miR-517a. *Gastroenterology* 140:1618.e16–1628.e16.
18. Luk JM, et al. (2011) DLK1-DIO3 genomic imprinted microRNA cluster at 14q32.2 defines a stemlike subtype of hepatocellular carcinoma associated with poor survival. *J Biol Chem* 286:30706–30713.
19. Russell DW, Hirata RK (1998) Human gene targeting by viral vectors. *Nat Genet* 18:325–330.
20. Miller DG, et al. (2006) Gene targeting in vivo by adeno-associated virus vectors. *Nat Biotechnol* 24:1022–1026.
21. Paulk NK, et al. (2010) Adeno-associated virus gene repair corrects a mouse model of hereditary tyrosinemia in vivo. *Hepatology* 51:1200–1208.
22. Davidoff AM, Ng CY, Zhou J, Spence Y, Nathwani AC (2003) Sex significantly influences transduction of murine liver by recombinant adeno-associated viral vectors through an androgen-dependent pathway. *Blood* 102:480–488.
23. Naugler WE, et al. (2007) Gender disparity in liver cancer due to sex differences in MyD88-dependent IL-6 production. *Science* 317:121–124.
24. Epstein CJ, Gatens EA (1967) Nuclear ploidy in mammalian parenchymal liver cells. *Nature* 214:1050–1051.
25. Arias IM, ed (1994) *The Liver* (Raven Press, New York, NY), 3rd Ed, pp 6–7.
26. Nakanishi K, Sakamoto M, Yamasaki S, Todo S, Hirohashi S (2005) Akt phosphorylation is a risk factor for early disease recurrence and poor prognosis in hepatocellular carcinoma. *Cancer* 103:307–312.
27. Yamashita T, et al. (2008) EpCAM and alpha-fetoprotein expression defines novel prognostic subtypes of hepatocellular carcinoma. *Cancer Res* 68:1451–1461.
28. Maetzel D, et al. (2009) Nuclear signalling by tumour-associated antigen EpCAM. *Nat Cell Biol* 11:162–171.
29. Tanimizu N, Nishikawa M, Saito H, Tsujimura T, Miyajima A (2003) Isolation of hepatoblasts based on the expression of Dlk/Pref-1. *J Cell Sci* 116:1775–1786.
30. Huang J, et al. (2007) Up-regulation of DLK1 as an imprinted gene could contribute to human hepatocellular carcinoma. *Carcinogenesis* 28:1094–1103.
31. Yanai H, et al. (2010) Dlk-1, a cell surface antigen on foetal hepatic stem/progenitor cells, is expressed in hepatocellular, colon, pancreas and breast carcinomas at a high frequency. *J Biochem* 148:85–92.
32. Schmidt CM, McKillop IH, Cahill PA, Sitzmann JV (1997) Increased MAPK expression and activity in primary human hepatocellular carcinoma. *Biochem Biophys Res Commun* 236:54–58.
33. Haddad R, Held WA (1997) Genomic imprinting and Igf2 influence liver tumorigenesis and loss of heterozygosity in SV40 T antigen transgenic mice. *Cancer Res* 57:4615–4623.
34. Tamm I, et al. (1998) IAP-family protein survivin inhibits caspase activity and apoptosis induced by Fas (CD95), Bax, caspases, and anticancer drugs. *Cancer Res* 58:5315–5320.
35. Alexiou P, et al. (2010) The DIANA-mirExTra web server: From gene expression data to microRNA function. *PLoS ONE* 5:e9171.
36. Betel D, Koppal A, Agius P, Sander C, Leslie C (2010) Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol* 11:R90.
37. Yang JH, Shao P, Zhou H, Chen YQ, Qu LH (2010) deepBase: A database for deeply annotating and mining deep sequencing data. *Nucleic Acids Res* 38(Database issue): D123–D130.
38. Lowe TM, Eddy SR (1999) A computational screen for methylation guide snoRNAs in yeast. *Science* 283:1168–1171.
39. Tward AD, et al. (2007) Distinct pathways of genomic progression to benign and malignant tumors of the liver. *Proc Natl Acad Sci USA* 104:14771–14776.
40. Cavaillé J, Seitz H, Paulsen M, Ferguson-Smith AC, Bachellerie JP (2002) Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. *Hum Mol Genet* 11:1527–1538.
41. Seitz H, et al. (2003) Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene. *Nat Genet* 34:261–262.
42. Kishore S, Stamm S (2006) The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science* 311:230–232.
43. Mingozzi F, High KA (2011) Therapeutic in vivo gene transfer for genetic disease using AAV: Progress and challenges. *Nat Rev Genet* 12:341–355.
44. Wu Z, Asokan A, Samulski RJ (2006) Adeno-associated virus serotypes: Vector toolkit for human gene therapy. *Mol Ther* 14:316–327.
45. Manno CS, et al. (2006) Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat Med* 12:342–347.