Introduction
The E2F family of transcription factors consists of 8 members that can be organized into 3 subcategories based on function and expression patterns during the cell cycle: activators (E2F1-3), canonical repressors (E2F4–6), and atypical repressors (E2F7 and E2F8). The complexity of the E2F family is further increased by differential promoter usage and splicing, resulting in multiple E2F isoforms (1–5). Through sequential binding to target promoters, E2F activators and repressors are believed to choreograph the oscillatory nature of cell cycle–dependent gene expression (6–8).

E2Fs play a critical role in the control of cellular proliferation (9–13). In fruit flies, which have a single E2F activator (dE2F1) and repressor (dE2F2), the ablation of dE2F1 is sufficient to cause cell proliferation arrest at the larval stage, leading to embryo lethality (6, 14). In plants and mammals, which have multiple activator and repressor family members, ablation of individual E2F6 has minimal impact on cell proliferation and embryonic development (15, 16). When activator or repressor subgroups are ablated in combination, a spectrum of developmental abnormalities is observed, including wide-spread DNA damage and apoptosis, organ-specific atrophy, restricted organismal growth, and animal lethality (15, 16).

Why such diversity in the number and types of E2F components exists among eukaryotic species remains a matter of speculation.

Physical interaction of the tumor suppressor retinoblastoma 1 (RB1) with E2F transcription factors inhibits E2F-dependent gene expression and is believed to underlie RB1 tumor suppressor function (12, 16–19). In humans, germline mutations in RB1 lead to retinoblastoma at a young age and affected individuals have a strong susceptibility to a number of other malignancies (20). Inactivation of the RB pathway, either through amplification of negative regulators (CDK4, CDK6, CCND1, or CCNE2), genetic or epigenetic inactivation of positive regulators (p16INK4A, p18INK4c, and p27KIP1), or mutations in RB1 itself, is frequently detected in cancer (21, 22). This has led to the view that disruption of the RB/E2F pathway leading to increased E2F transcriptional activity is a universal requirement for cancer development (23, 24). The observation that unregulated cell proliferation as a consequence of Rb1 deficiency in mice is blocked by the ablation of E2F activators suggests that E2Fs are critical downstream effectors of RB1 tumor-suppressor function (13, 25). Consistent with these findings, deletion of the RB family or supraphysiological overexpression of E2Fs leads to hepatocellular carcinoma (HCC) in mice (26, 27). Yet despite extensive efforts, genetic alterations in E2F genes are only rarely found in human tumors, casting doubt on a direct role of E2Fs in cancer. One explanation for this paucity of genetic alterations in E2Fs may be related to the ability of RB1 to physically interact with over 150
other factors, many of which also have implicit or explicit connections to cancer (28). Thus, it is entirely possible that RB1’s tumor suppression function may reside, at least in part, outside the regulation of E2F activity and within the control of one or more of its many associated proteins.

To explore a possible role for E2Fs in human cancer, we queried recently generated public databases for genetic alterations in E2F family members. This analysis revealed very modest but significant (P < 0.001; 2-sided Student’s t test vs. control liver) increases in E2F output and a striking incidence of spontaneous HCC without affecting development at any stage, from fetal to adult. Hepatocyte-specific restoration of E2F1 copy number levels prevented spontaneous development of HCC, suggesting a cell-autonomous role for E2F1 in driving HCC. Conversely, decreased E2F output in E2f1−/− and E2f3b−/− mice protected against carcinogen-induced HCC. Further genetic and mechanistic studies identified critical molecular differences between family members that could explain the selective and potent dosage-dependent oncogenic roles of E2F1 and E2F3B. These findings expose how the precise composition of E2F family members is required to control transcriptional specificity and output in order to foster a cancer-free life span.

Results

Copy number gains in E2F1 and E2F3 in human HCC. To explore whether altered E2F signaling may be directly related to development of human cancer, we queried The Cancer Genome Atlas (TCGA; https://tcga-data.nci.nih.gov/docs/publications/tcga/) and Catalogue of Somatic Mutations in Cancer (COSMIC; http://cancer.sanger.ac.uk/cosmic) databases for genetic and epigenetic alterations in E2F family members. As previously noted by others, genetic alterations in E2Fs are infrequent in most solid tumor types. However, we found subtle but significant E2F1 and E2F3 copy number gains and E2F2 copy number losses in HCC (Table 1, Figure 1, and Supplemental Figure 1; supplemental material available online with this article; https://doi.org/10.1172/JCI87583DS1). As expected, there were also frequent mutations and gene copy losses in RB1 and TP53 tumor suppressors. These findings suggest that while relatively rare across most cancer types, modest genetic alterations in select E2F activators may be associated with development of HCC. A tumor-suppressive role for E2F2 was previously identified in B and T cell lymphomagenesis (29, 30), which may reflect copy number losses observed here in HCC. We thus focused subsequent analysis on the relevance of E2F1 and E2F3 copy number gains in cancer.

Loss of E2f1 and E2f3b protects against HCC. The E2F3 gene locus encodes 2 isoforms regulated by distinct promoters, generating proteins that differ at their N termini (E2F3A and E2F3B). We initially took a loss-of-function genetic approach to determining whether E2F1 and the 2 E2F3 isoforms contribute to HCC.

To this end, HCC development was evaluated in E2f1−/−, E2f1−/−, E2f3a−/−, and E2f3b−/− mice (hereafter referred to as E2f−/−, 1−/−, 3a−/−, and 3b−/− mice, respectively; refs. 31, 32) treated with the liver-specific carcinogen diethylnitrosamine (DEN; ref. 33). The liver-to-body weight ratio of untreated mice did not differ among genetic groups (Supplemental Figure 2A). Small macroscopic liver tumors were visible in 9-month-old DEN-treated E2f−/−, 3a−/−, and 3b−/− mice, but not in 1−/− mice (Figure 2A and Supplemental Fig-
ure 2, B and C). A detailed pathological evaluation incorporating known criteria associated with disease progression and poor prognosis (Supplemental Table 1; ref. 34) showed less advanced disease in 3a–/– and 3b–/– mice than in 3a+/– or 3F3b+/– mice (Table 1 and Figure 2B). A reduction in tumor incidence and tumor burden was also observed in 14-month-old 3a1KI/– mice (Table 1, Figure 2C and D; and Supplemental Figure 2, D and E). These observations suggest that decreased dosage of specific E2Fs protect against cancer.

E2f1 and E2f3b are critical drivers of HCC. The observation that E2f1 and E2f3b have specific roles in driving cancer is interesting given our previous studies showing that E2f1, E2f3a, and E2f3b proteins are interchangeable during fetal to early adult development (32). In these studies, knockin strategies showed that expression of E2f1 or E2f3b from the E2f3a locus (referred to hereafter as 3a1KI and 3a3bKI, respectively) fostered normal fetal to adult development, suggesting redundant developmental functions for E2f activators. However, aging cohorts of knockin mice were never analyzed. Thus, we evaluated the life-long consequences of having 1 or 2 extra copies of E2f1 or E2f3b in 3a1KI and 3a3bKI knockin mice (Figure 3A). Cohorts of 18-month-old E2f1+/+ and 3a–/– control mice appeared normal, and necropsies revealed only rare age-related lesions (Table 4). In contrast, the majority of 3a1KI/– and 3a3bKI/– knockin mice developed spontaneous liver masses. Thus, despite the widespread expression of E2f1 or E2f3b from the E2f3a locus in all proliferating tissues of the mouse (32), frequent lesions were only observed in livers. We confirmed higher levels of E2f1 and E2f3b protein in livers from knockin mice than in controls (Figure 3B and Supplemental Figure 3). The majority of liver masses were characterized as highly proliferative HCC in the absence of excessive apoptosis as measured by Ki-67 and caspase 3 (Figure 3C; caspase 3 data not shown). Analysis of heterozygous 3a1KI/3a and 3a3bKI/3a knockin mice developed intermediate levels of spontaneous liver masses. By taking advantage of LoxP sites flanking the 3aKI allele (Figure 3A), we assessed spontaneous HCC in aging cohorts of mice with hepatocyte- or monocyte-specific expression of Cre (40, 41). Tumor incidence and tumor burden were increased in monocyte-ablated mice (Lys-Cre 3a1KI/– and Lys-Cre 3a3bKI/–) to an extent similar to that seen in 3a1KI/– and 3a3bKI/– mice (Figure 5A–D), indicating that E2f1 in Kupffer cells is unlikely to play a major role in HCC development. In contrast, there was a complete absence of

| Table 2. Pathologic scores ± SEM of DEN-treated E2F knockout mice at 9 months |
|---------------------------------|---------|--------|--------|----------|----------|-------------|-----|
|                                | Percentage with HCC | Average pathologic score | Average volume | Average multiplicity | Average necrosis | Average vascular invasion | n   |
| E2f3a+/+                       | 88%     | 2.24 ± 0.25 | 1.32 ± 0.18 | 0.48 ± 0.10 | 0.00 ± 0.00 | 0.44 ± 0.10 | 25  |
| 1–/–                           | 5%      | 0.05 ± 0.05 | 0.20 ± 0.06 | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 | 20  |
| 3a1KI/1KI                      | 76%     | 2.19 ± 0.41 | 1.29 ± 0.21 | 0.24 ± 0.10 | 0.19 ± 0.19 | 0.48 ± 0.11 | 21  |
| 3b1KI/3b1KI                    | 61%     | 1.30 ± 0.26 | 0.87 ± 0.18 | 0.26 ± 0.09 | 0.00 ± 0.08 | 0.17 ± 0.08 | 23  |

*P ≤ 0.05 vs. E2f3a+/+ by Fisher’s exact tests with Bonferroni’s correction for multiple tests. **P ≤ 0.05 vs. E2f3a+/+ by Wilcoxon method with Bonferroni’s correction.

| Table 3. Pathologic scores ± SEM of DEN-treated E2F knockout mice at 14 months |
|---------------------------------|---------|--------|--------|----------|----------|-------------|-----|
|                                | Percentage with HCC | Average pathologic score | Average volume | Average multiplicity | Average necrosis | Average vascular invasion | n   |
| E2f3a+/+                       | 94%     | 3.76 ± 0.45 | 2.53 ± 0.31 | 0.47 ± 0.12 | 0.35 ± 0.17 | 0.41 ± 0.12 | 17  |
| 1–/–                           | 36%     | 1.14 ± 0.47 | 0.93 ± 0.37 | 0.14 ± 0.10 | 0.00 ± 0.00 | 0.07 ± 0.07 | 14  |
| 3a1KI/1KI                      | 88%     | 3.76 ± 0.53 | 2.35 ± 0.30 | 0.41 ± 0.12 | 0.35 ± 0.17 | 0.65 ± 0.12 | 17  |
| 3b1KI/3b1KI                    | 90%     | 3.30 ± 0.54 | 1.80 ± 0.33 | 0.50 ± 0.17 | 0.20 ± 0.13 | 0.80 ± 0.13 | 10  |

*P ≤ 0.05 vs. E2f3a+/+ by Fisher’s exact tests with Bonferroni’s correction for multiple tests. **P ≤ 0.05 vs. E2f3a+/+ by Wilcoxon method with Bonferroni’s correction.

To rule out any possible contribution due to the inadvertent disruption of regulatory sequences caused by insertion of the knockin alleles into the E2f3a locus, we evaluated HCC development in 3a1KI/1KI mice lacking the endogenous E2f1 gene. This analysis revealed a dose-dependent reduction in the grade of lesions observed in knockin mice lacking 1 (1–/3a1KI/1KI) or both (1–/3a1KI/1KI) endogenous copies of E2f1 (Figure 4, A and B). Thus, we conclude that the total levels of E2f1 protein dictate tumor outcome regardless of the genomic location of the gene. These findings formally demonstrate that increased E2f1 gene dosage, culminating in higher E2f1 protein levels, leads to spontaneous and DEN-induced HCC in mice.

Oncogenic function of E2f1 is cell autonomous. Several factors contribute to HCC, including hepatitis B or C virus infection, heavy alcohol consumption, aflatoxin exposure, Western-style diet, obesity, and diabetes (37). The exact mechanism of how these varied factors promote HCC is unclear, but chronic inflammation and liver damage leading to further genomic instability likely contribute to this disease (37, 38). Kupffer cells are abundant liver-resident macrophages that have been implicated in the development of HCC by promoting inflammation, angiogenesis, and epithelial-to-mesenchymal transition (39). E2f1 and E2f3a are expressed in proliferating tissues and are particularly high in lymphoid and myeloid lineages (4). Therefore, we considered the possibility of a cell-nonautonomous oncogenic role for E2f1 in HCC. By taking advantage of LoxP sites flanking the 3a1KI allele (Figure 3A), we assessed spontaneous HCC in aging cohorts of mice with hepatocyte- or monocyte-specific expression of Cre (40, 41).
liver lesions in hepatocyte-ablated mice (Alb-Cre 3a<sup>+/+</sup> and Alb-Cre 3a<sup>3b</sup>KI/3bKI), just as in E2f<sup>+/+</sup> control mice. Together, these studies suggest a cell-autonomous oncogenic role for E2f1 in hepatocytes.

**ChIP-seq defines E2F1, E2F3A, and E2F3B chromatin-binding landscapes.** The genetic experiments described above highlight distinct roles of E2F family members in the development of cancer. To explore the underlying mechanism for these observations, we sought to identify and compare the chromatin-binding landscapes of E2F1, E2F3A, and E2F3B by ChIP followed by next-generation sequencing (ChIP-seq). E2F chromatin occupancy was mapped by sequencing ChIP-DNA libraries originating from asynchronous populations of mouse embryonic fibroblasts (MEFs) overexpressing these E2Fs (Figure 6A). DNA-sequencing tags were mapped to the mouse genome and peak summits identified using a combination of HOMER (http://homer.salk.edu/homer/) and MACS (http://liulab.dfci.harvard.edu/MACS/) algorithms. Peaks identified in MEFs were validated in a liver cancer cell line (HepG2) overexpressing each of these E2Fs (Supplemental Figure 4, A–D). The ChIP-seq analysis exposed several expected and unexpected observations. First, we found twice as many E2F1- and E2F3B-specific peaks as E2F3A-specific peaks (3,788, 4,705, and 2,000 peaks, respectively). While a distinct subset of chromatin locations was equally occupied by all 3 E2Fs, most chromatin locations were specifically enriched for E2F1 and E2F3B (Figure 6B), suggesting selectivity in the chromatin binding of E2F activators. Second, the majority of E2F1, E2F3A, and E2F3B peaks mapped near the transcriptional start site (TSS) (Figure 6C and Supplemental Table 2, 3, and 4), consistent with previous findings (13, 42). Interestingly, promoter-specific peaks (± 2 kb from the TSS) were mostly located over canonical E2F consensus-binding sequences (TTCCGCC; Figure 6D), whereas peaks mapping to other gene regions contained a variety of canonical and noncanonical E2F-binding motifs (Supplemental Table 5). There were typically more read tags per peak in promoters occupied by E2F1 and E2F3B than in promoters occupied by E2F3A (Figure 6E and Supplemental Figure 5A). Third, gene ontology using Ingenuity Pathway Analysis (IPA) software revealed that E2F1- and E2F3B-bound promoters were significantly enriched for genes encoding functions related to hepatocyte proliferation and HCC (CDKN1A, MYC, TRIB1, NFE2L2; Figure 6F and Supplemental Table 4).

### Table 4. Malignancies observed in control and knockin mice aged 12 to 18 months

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2f&lt;sup&gt;+/+&lt;/sup&gt;</td>
<td>Liver (9%), lung (13%), and stomach (4%)</td>
<td>Uterus (3%)</td>
</tr>
<tr>
<td>3a&lt;sup&gt;+/+&lt;/sup&gt;</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>3b&lt;sup&gt;+/+&lt;/sup&gt;</td>
<td>Liver (94%)</td>
<td>Liver (64%)</td>
</tr>
<tr>
<td>3a3b&lt;sup&gt;+/+&lt;/sup&gt;</td>
<td>Liver (56%) and lung (11%)</td>
<td>Liver (72%), lung (6%), and uterus (6%)</td>
</tr>
</tbody>
</table>

Malignancies include adenomas or carcinomas (liver, lung and stomach) and histiocytic sarcoma (uterus).
those that were generally related to cancer (UHFR1, TP53, MYC, BRCA1, BRCA2); Supplemental Figure 5B and Supplemental Table 7). E2F3A and E2F3B share identical DNA binding, dimerization, and transactivation domains and only differ at their N termini (2).

Hence, the identification of E2F3B-specific functions in HCC was even though targets were initially identified in MEFs. In contrast, promoters occupied by all 3 E2Fs were enriched for genes with functions in cell cycle control (CDC6, TOP2A, CCNE1, CDT1, CDC25A, MCMs), DNA damage and repair (CHEK1, RADS1, BRCA1, BRCA2), and cell death (XPO1, TP53, WEE1) as well as those that were generally related to cancer (UHFR1, TP53, MYC, BRCA1/2; Supplemental Figure 5B and Supplemental Table 7). E2F3A and E2F3B share identical DNA binding, dimerization, and transactivation domains and only differ at their N termini (2). Hence, the identification of E2F3B-specific functions in HCC was
completely unexpected. Direct comparison between E2F3A and E2F3B chromatin-binding landscapes revealed significant differences in their binding specificity (Figure 7A), and IPA illustrated the remarkable HCC-related nature of E2F3B-specific targets (Figure 7B). We considered that differences in their protein stability may underlie, at least in part, the striking qualitative and quantitative differences between E2F3A and E2F3B chromatin binding. N-terminal sequences within exon 1A of E2f3a, which are absent in E2F3B, have been shown to mediate E2F3A protein degradation (43). Thus, we compared the stability of E2F3A and E2F3B proteins following treatment of cells with cycloheximide. These time-course experiments showed that the half-life of E2F3B was significantly longer than that of E2F3A (>3-fold; Figure 7, C–E). While we have not ruled out additional roles for the N-termini of these 2 isoforms, our findings suggest that protein stability may be a contributing factor to the stronger and broader target-binding range of E2F3B. In short, genome-wide chromatin occupancy analysis exposed both shared and unique targets of E2Fs, with E2F1 and E2F3B binding preferentially to target promoters with functions that are highly related to development of HCC.

Identification of putative direct HCC-specific targets of E2F1 and E2F3B. We then hypothesized that increased chromatin binding due to copy number gains in E2F1 and E2F3B led to increased E2F transcriptional output. To test this hypothesis, we used an Affymetrix platform to profile mRNA expression in livers of 12-month-old E2F3B (normal), 3a+/− (tumor), and 3a–/– (tumor) mice. Genes differentially expressed in 3a+/− and 3a–/– tumors relative to 3a+/+ control samples are represented as heat maps (fold change ≥ 1.5 vs. P ≤ 0.05; Figure 8, A and B; and Supplemental Tables 8 and 9). Levels for most differentially expressed genes were similar between control E2F3+/− and 3a+/− samples. Upregulated mRNAs in 3a+/−/3a+/− tumors constituted a larger subset than downregulated mRNAs, consistent with a role for E2F1 in transcriptional activation. There was a relatively larger subset of downregulated mRNAs in 3a+/−/3a−/− than in 3a+/−/3a+/− tumors, possibly reflecting an additional role for E2F3B in transcriptional repression as previously described (1, 2). There was also a marked change in the expression of genes related to cancer, cell cycle control, DNA repair, and metabolic processes as well as cell survival, with a notable absence of genes that induce programmed cell death (Supplemental Figure 6A and Supplemental Table 10). We then plotted the intersection of ChIP and mRNA profiling data sets as tag-intensity maps (Figure 8, C–E; and Supplemental Tables 11 and 12). This analysis showed that 20% of dysregulated genes represent putative direct targets of E2F1 and E2F3B, but not E2F3A. The occupancy of endogenous E2F1 and E2F3 on target promoters was validated by ChIP–quantitative PCR (ChIP-qPCR) in multiple cell lines derived from human liver tumors (Figure 8F and Supplemental Figure 6B).

E2F1 and E2F3B target expression is associated with human HCC. To determine the relevance of our findings to human HCC, we queried the expression of putative E2F1 and E2F3B targets in liver-derived mRNA profiles from healthy individuals and patients with cirrhosis, dysplasia, early HCC, and advanced HCC (44). This analysis revealed a significant and progressive increase in E2F1 and E2F3B expression in patients with early and advanced HCC (Figure 9A and Supplemental Table 13). Importantly, we also found a subset of genes highly expressed in patients with early stages of liver disease (cirrhosis), the majority of which were E2F3B targets (83%). We then used IPA to query the functions of putative E2F1 and E2F3B targets (Figure 9B, Supplemental Figure 7A, and Supplemental Tables 14, 15, and 16). This analysis revealed a disproportionate enrichment of E2F1 and E2F3B upregulated targets involved in hepatocyte proliferation and cancer. There were also many E2F1- and E2F3B-specific targets with distinct functional roles. For example, many E2F3B-specific targets were uniquely enriched for functions contributing to liver fibrosis and cirrhosis, 2 major predisposing factors contributing to HCC development. In contrast, E2F1 targets were enriched for functions known to participate in canonical cell cycle, checkpoint control, and DNA repair pathways, which are often dysregulated in cancer cells. While the significance of this unexpected observation remains to be determined, we suggest that high E2F3B expression may sensitize livers to cirrhosis and HCC development, while E2F1 may act more directly by altering critical DNA integrity pathways and driving the accumulation of mutations.

Finally, we analyzed E2F1 and E2F3 gene copy number and mRNA expression in HCC patient samples from the TCGA database and found increased expression of E2F1/3 target genes in patients with amplification of E2F1 (P < 0.001) (Figure 9C, Supplemental Figure 7B, and Supplemental Tables 17 and 18). E2F3 copy number gains were correlated to increased expression of fewer target genes. Importantly, increased E2F target gene expression was associated with decreased survival of HCC patients (Figure 9D).

In summary, loss- and gain-of-function strategies in mice demonstrated a causative role for E2F1 and E2F3B in driving distinct aspects of liver disease. E2F1 and E2F3B required for cancer development (22, 23). However, there is a
The surprising paucity of genetic alterations in the main downstream transcriptional effectors of RB1, E2Fs, which has cast doubt on their direct involvement in the genesis of cancer. We identified E2F1 and E2F3 copy-number gains in human HCC patients and show in mouse models that these modest genetic alterations are sufficient to initiate HCC. Using a series of E2F alleles to finely tune E2F transcriptional output, we show that increasing output by gains in 

**Figure 5. Oncogenic functions of E2F1 and E2F3b are cell autonomous.** (A) Box plots showing the liver/body weight of 12- to 18-month-old male mice in which the knockin allele or alleles have been deleted in hepatocytes or macrophages using Alb-Cre or Lys-Cre, respectively. Wilcoxon method with Bonferroni’s correction for multiple tests. *P < 0.001 vs. control. (B) Histopathological classification of livers from A. Fisher’s exact tests with Bonferroni’s correction. ‡P = 0.001, carcinoma vs. control. (C) Box plots showing the liver/body weight of 12- to 18-month-old female mice. Wilcoxon method with Bonferroni’s correction. *P < 0.012 vs. control. (D) Histopathological classification of livers from C. Fisher’s exact tests with Bonferroni’s correction. 1P = 0.001, carcinoma vs. control.

Oncogenic functions of E2F1 and E2F3b. Here, we demonstrate distinct functions for E2F1, E2F3A, and E2F3B in cancer. Tissue-specific overexpression of E2F1 to supraphysiological levels using artificial transgenes has been shown to promote cancer in a variety of organs (26). In this report, we show that introduction of a single extra copy of E2F1 or E2F3B leads to a modest, yet physiologically relevant, increase in E2F1 and E2F3B protein and activity that is sufficient to induce spontaneous HCC in vivo (Figure 3). Substantial sequence differences exist between E2F1 and E2F3A that could explain their distinct biological roles in cancer. However, differences between E2F3A and E2F3B functions are less clear. Given that E2F3A was replaced with E2F3B in the 3aMKI knockin mouse model described here, we suggest that any biological differences between these 2 isoforms must be independent of regulatory sequences (promoter, intronic, 3′ UTR) and likely due to differences in their protein function. Indeed, E2F3B’s oncogenic role in vivo could be explained, at least in part, by its increased protein stability (Figure 7). While the precise mechanism remains to be elucidated, the observation that a single extra copy of E2F1 or E2F3B resulted in cancer highlights the exquisite functional specificity among E2F family activators in maintaining appropriate transcriptional output necessary for normal liver physiology. While previous knockin studies suggested that E2F1, E2F3A, and E2F3B are interchangeable during development (32), we now describe cancer as the long-term consequence associated with changing the constellation of specific E2F family members.

Precise E2F transcriptional output is regulated by both activation and repression. Interestingly, recent work from our laboratory showed that ablation of the atypical E2F repressor, E2F8, in mice also results in HCC without affecting basic liver functions (45). These observations, together with our current findings, suggest that transcriptional E2F1-3B–mediated activation and E2F8-mediated repression of target genes may be functionally linked. Indeed, we find that many of the E2F1 and E2F3B target genes upregulated in HCC (identified here) are also repressed by E2F8 (45). Thus, it appears that maintenance of normal levels of E2F activity, which require both activators and repressors, is dispensable for animal development and yet is critical for preventing HCC later in life. These findings begin to explain why higher eukaryotes evolved to have such a large and complex family of E2Fs. Evidently, substantial evolutionary currency has been spent to fine-tune E2F output and help promote a cancer-free life span.

Why is the liver particularly sensitive to increased E2F transcriptional output? Why hepatocytes may be uniquely sensitive to altered levels of E2F target gene expression is a matter of speculation. High levels of cellular proliferation and a toxic cellular environment are 2 key features of developing livers that may explain this cancer susceptibility. Postnatal livers go through a rapid phase of hepatocyte proliferation immediately following birth, which quickly subsides after 3 weeks of age. The liver is also the main site for detoxification, a process that exposes hepatocytes to continuous DNA-damaging insults. CHIP-seq and expression profiling in liver tumors identified a core set of target genes coregulated by E2F1 and E2F3B that are highly enriched for DNA repair or checkpoint activities (Aurka, Conel, Cdk1, Cdkn1a, Chekl, Myc, Rad51, and Top2a) and key components of DNA methylation machinery (Hels and Uhrf1). Interestingly, elevated expression of many of these target genes, such as Rad51, is observed in cancer and is associated with decreased survival (46, 47). It is unclear how overexpression of these proteins affects DNA repair and/or other processes involved in maintaining...
In addition to increased transcriptional output, an alternative, but not mutually exclusive, theory for the sensitivity of knockin mice to liver cancer may be related to the capacity of hepatocytes to resist a variety of genetic and/or environmental insults that induce programmed genomic integrity. But one interpretation of these observations is that increased E2F transcriptional output may alter the capacity of hepatocytes to accurately replicate and repair the genome during peak periods of proliferation in a particularly toxic environment.

Figure 6. Identification of E2F targets by ChIP sequencing. (A) Immunoblot of protein lysates from MEFs expressing the indicated E2Fs using the pBABE-Hygro (pBH) vector. E2F1 (E2f1 pBH), E2F3A (E2f3a pBH), and E2F3B (E2f3b pBH); 1–/– and 3–/– MEFs are shown as negative controls. GAPDH was used as a loading control. (B) Sequence tag-density heat map showing the distribution of tags for all E2F1, E2F3A, and E2F3B peaks. (C) Percentages of E2F1, E2F3A, and E2F3B peaks in different gene regions. Gene regions were defined by distance from TSS (TSS = 0) as follows: 5′ distal (-50 Gb to -50 kb), 5′ proximal (-50 kb to -2 kb), promoter (-2 kb to +2 kb), gene body (+2kb to end of transcript), 3′ distal (end of transcript to +30 Gb). Number of peaks is indicated above each bar. (D) Graph depicting the number of peaks for E2F1, E2F3A, and E2F3B identified by ChIP-seq and their location relative to the TSS. The promoter region (+2 kb from the TSS) is highlighted in gray, and the top binding motif identified by de novo HOMER within the promoter region is included in color inserts. (E) E2F1, E2F3A, and E2F3B occupancy on selected gene promoters shown at the same scale. (F) Gene ontology using IPA software depicts the estimated contribution of E2F1, E2F3A, or E2F3B targets to liver hyperplasia and hyperproliferation (left) or HCC-related functions (right). Bars indicate the Benjamini-Hochberg adjusted P value; the threshold of P = 0.05 is shown.
cell death. Overexpression of E2F1, either through the artificial introduction of E2F1-overexpressing transgenes or through loss of Rb and/or E2f7-8 repressors, has been shown to engage p53-dependent and -independent apoptotic programs (8, 25, 26, 48, 49). However, we show here that livers of 3a1/3b or 3a1/3b/3b mice are particularly resistant to apoptosis despite increased ectopic expression of E2F activators. Thus, we speculate that, relative to other organs in the mouse, the liver is predisposed to HCC development due to the inability of hepatocytes to engage an apoptotic response while remaining fully capable of engaging proliferative gene expression programs.

Hepatocyte-specific ablation of the extra copies of E2F1 (3a1/3b) during early development using the Alb-Cre system suggested a cell-autonomous oncogenic role for E2F1 in the initiation of HCC (Figure 5). Whether E2F1 and E2F3 have a role beyond the initiation of HCC remains to be determined. Some insights can be gained from a pilot study using an inducible Cre system (Mxi-Cre) where the oncogenic 3a1/3b alleles can be deleted in adult mice after tumor formation (Supplemental Figure 8). In this study, magnetic resonance imaging of established liver tumors in live mice showed that tumor volume continued to increase following the Mxi-Cre-mediated deletion of the 3a1/3b knock-in alleles, suggesting that extra copies of E2F1 may be dispensable for tumor growth (Supplemental Figure 8). The mechanism of tumor initiation may involve alterations in checkpoint and DNA repair activities specific to the physiology of the liver as discussed above. While not necessary for continued growth, we cannot rule out a role for oncogenic E2F1 in progression to a more malignant phenotype and a poorer outcome.

The genomic landscape of HCC differs from that of most other cancer types since it is void of dominant driving genetic alterations (50). Whole genome analyses of HCC revealed a large number of “heterogeneous” mutations, each contributing a small portion of the entire patient cohort studied. Consistent with this view, we show that not all HCC patients suffer E2F1 and/or E2F3 copy number gains (Figure 1 and Figure 9C), but surprisingly, virtually all tumor samples analyzed exhibit increased expression of E2F1/E2F3B target genes. We propose that dysregulation of the E2F program is sufficient to drive HCC development in mice and is likely a downstream and necessary consequence of most if not all genetic alterations associated with HCC development.

Methods

Supplemental Methods are available online with this article; https://doi.org/10.1172/JCI87583DS1.

Mouse care and models. Mice were housed under normal conditions. E2f1–/–, E2f3a–/–, E2f3b–/–, E2f3a1/3b, E2f3a1/3b/3b, Lys-Cre, and Alb-Cre mouse lines were described previously (31, 32, 40, 41). The backgrounds of the E2f1–/–, E2f3a–/–, and E2f3b–/– mice were FVB/NT fifth, fifth, and seventh generations, respectively. The E2f3a1/3b and E2f3a1/3b/3b mice were on a mixed background (FVB/NT, 129/Sv, and C57BL/6J). Genotyping was performed on tail or liver DNA using standard PCR techniques. Primer sets used are listed in Supplemental Table 19. Dissected livers were flash-frozen in liquid nitrogen and stored at –80°C for further analysis or fixed in 10% formalin (Fisher Scientific) for histologic analyses. For DEN (Sigma-Aldrich) treatment, mice received an intraperitoneal injection of DEN (20 mg/1 kg body mass) at 20 days of age (33, 51).

Histology and diagnosis of HCC. Formalin-fixed mouse tissues were processed and stained for H&E using standard protocols. Histopathological analysis was used to determine HCC by pathologists based on cellular morphology (34). Additional analysis was conducted on a subset of samples in which tumor volume, burden, necrosis, and vascular invasion were scored as described in Supplemental Table 1. For immunohistochemistry against Ki-67 (Abcam, 16667) and caspase 3 (Cell Signaling, 9662), slides were stained using a Bond Rx autostainer (Leica) per the manufacturer’s instructions.
Ghoshal (The Ohio State University). MEFs, SNU-449, PLC/PRF5, HepG2, Huh7, and SK-Hep-1 cells were cultured in DMEM with 10% FBS. E2Fs were overexpressed in MEFs and HepG2 cells after transduction with the retroviral vector pBABE-Hygro (pBH, Addgene) containing the sequence for murine $E2f1$, $E2f3a$, or $E2f3b$.

**ChIP-seq and ChIP-qPCR.** For all ChIPs, proliferating cells were fixed with 1% formaldehyde. Nuclear lysates were sonicated and incubated with 10 μg α-E2F1 (Santa Cruz Biotechnology Inc., c20), α-E2F3 (Santa Cruz Biotechnology Inc., c18), or normal rabbit IgG (Santa Cruz Biotechnology Inc.) overnight. Antibody-protein-DNA complexes were recovered using protein G agarose beads (Millipore), and then decrosslinked.

**Immunoblotting.** Immunoblots were utilized to verify expression of the proteins as described (52). Whole-cell protein lysates were collected at the indicated time points after treatment, separated by SDS-PAGE, and transferred to PVDF membranes. Blots were probed with α-E2F1 (Santa Cruz Biotechnology Inc., C-20), α-E2F3 (Santa Cruz Biotechnology Inc., C-18), α-MYC (Santa Cruz Biotechnology Inc., 9E10), α-tubulin (Santa Cruz Biotechnology Inc., B-7), α-actin (Santa Cruz Biotechnology Inc., I-19), or α-GAPDH (Santa Cruz Biotechnology Inc., FL-335).

**Cell culture.** MEFs were isolated from day-13.5 embryos and immortalized using the 3T3 method (53). SNU-449, PLC/PRF5, HepG2, Huh7, and SK-Hep-1 cells were obtained from Kalpana Ghoshal (The Ohio State University). MEFs, SNU-449, PLC/PRF5, HepG2, Huh7, and SK-Hep-1 cells were cultured in DMEM with 10% FBS. E2Fs were overexpressed in MEFs and HepG2 cells after transduction with the retroviral vector pBABE-Hygro (pBH, Addgene) containing the sequence for murine $E2f1$, E2F3a, or E2F3b.

**ChIP-seq and ChIP-qPCR.** For all ChIPs, proliferating cells were fixed with 1% formaldehyde. Nuclear lysates were sonicated and incubated with 10 μg α-E2F1 (Santa Cruz Biotechnology Inc., c20), α-E2F3 (Santa Cruz Biotechnology Inc., c18), or normal rabbit IgG (Santa Cruz Biotechnology Inc.) overnight. Antibody-protein-DNA complexes were recovered using protein G agarose beads (Millipore), and then decrosslinked.

**Figure 8. Intersection of gene-expression profiling and chromatin binding identifies E2F1 and E2F3B targets.** (A) Heat map of Affymetrix microarray data showing differentially expressed genes in $3a^{+/+}$ liver tumors when compared with normal liver samples from $E2f^{+/+}$ and $3a^{-/-}$ age-matched controls. Differentially expressed genes are defined as having a fold change of 1.5 or more (P ≤ 0.05) relative to $3a^{-/-}$ samples. (B) Heat map of Affymetrix microarray data showing differentially expressed genes in $3a^{+/+}$ liver tumors when compared with normal liver samples from $E2f^{+/+}$ and $3a^{-/-}$ age-matched controls. Differentially expressed genes are defined as having a fold change of 1.5 or more (P ≤ 0.05) relative to $3a^{-/-}$ samples. (C) Venn diagram illustrating the overlap of E2F1-specific promoter peaks with upregulated or downregulated genes in $3a^{+/+}$ liver tumors identified in A. (D) Venn diagram illustrating the overlap of E2F3B-specific promoter peaks with upregulated or downregulated genes in $3a^{+/+}$ liver tumors identified in B. (E) Sequence tag-density heat map showing the distribution of E2F1, E2F3A, and E2F3B binding to targets identified in C and D (overlapping groups). (F) ChIP-qPCR validation using E2F1, E2F3, or IgG antibodies in SNU-449 and PLC/PRF5 HCC-derived cells. Occupancy of E2Fs on selected target promoters is shown. A nonpromoter region of TUBA4A (TUBA4A neg) was used as a negative control. Primers were designed to amplify ChIP-seq–identified peak regions.
For ChIP-seq experiments, E2F1, E2F3A, and E2F3B ChIPs were performed in MEFs stably overexpressing the proteins and in 1−/− and 3−/− MEFs. E2F-ChIP-DNA was processed to generate libraries using TruSeq ChIP-DNA library prep kit (Illumina). Briefly, 10 ng of pull-down DNA was used to ligate with barcoded adaptors and amplified with 14 cycles of PCR; then a 200–300–bp library was collected using E-gel (Invitrogen). Library quality was confirmed using High Sensitive DNA Assay (Agilent) and then sequenced by the OSUCCC.

Figure 9. Association of the expression of E2F1 and E2F3B targets with human HCC. (A) Heat map showing the expression of E2F1 and E2F3B upregulated targets in normal and diseased (cirrhosis, dysplasia, early or advanced HCC) human livers. Genes are grouped based on median expression values per patient group. Genes with the highest median expression in advanced HCC and cirrhosis are denoted on the left. (B) Gene ontology using IPA software showing the estimated contribution of E2F1 and E2F3B targets identified in Figure 8 to functions related to cancer, cell cycle, proliferation, fibrosis, and cirrhosis. Bars indicate the Benjamini-Hochberg adjusted P value; the threshold of P = 0.05 is shown. (C) Expression of E2F1 and E2F3B targets in human HCC samples with normal or increased copy numbers of E2F1 and E2F3. E2F1 diploid, 122 samples; gain, 62 samples. E2F3 diploid, 102 samples; gain, 80 samples. Wilcoxon rank sum test Benjamini-Hochberg corrected P values are indicated. (D) Kaplan-Meier plots evaluating the survival time of patients with low to high E2F1/3B target expression. Patients were divided into 3 categories based on target expression. Low expression (0%–5% targets upregulated; black line; n = 136); moderate expression (5%–10% targets upregulated; gray line; n = 113); or high expression (10%–45% targets upregulated; red line; n = 113). Log-rank test P values are shown.
GenChip-qPCR for endogenous (E2F1 and E2F3) or exogenous
(E2F1, E2F3A, and E2F3B) genes was performed in human liver
cells as previously described (45). Antibodies used were α-E2F1 (San-
ta Cruz Biotechnology Inc., clone C-20), α-E2F3 (Santa Cruz Biotech-
nology Inc., clone C-18), MYC (Abcam, clone 9132), or normal rabbit
IgG (Santa Cruz Biotechnology Inc.). Primers were designed to ampli-
fy regions containing peaks and are listed in Supplemental Table 20.

ChIP-seq data analysis. E2F1, E2F3A, and E2F3B ChIP-seq reads were
mapped against the reference genome (mm9) using Bowtie2
(54). Alignments from each ChIP experiment and input control were
processed using MACS1.4 (55) with -t and -c parameters to detect
peaks. Peaks from the 1/M1 MEFs were subtracted from the E2F1 ChIP,
while 3/M1 peaks were subtracted from the E2F3A and E2F3B ChIPs.
Promoter regions (±2 kb from the TSS) were utilized for downstream
analysis. Motif finding was done using the HOMER motif-finding tool
(Salk Institute). IPA (QIAGEN) was used for functional analysis of
reads. For individual genes were visualized using the Integrative
Genomics Viewer (http://software.broadinstitute.org/software/
igv/). Gene lists were compared using Venny 2.0.2 (http://bioinfogp.
cnb.csic.es/tools/venny/index2.0.2.html).

Cycloheximide treatment. Populations of MEFs stably overexpress-
ing MYC-tagged E2F3A or E2F3B were synchronized through serum
withdrawal and treated with 10 μg/ml of cycloheximide (Sigma-
Aldrich). Protein levels were quantified using ImageJ (NIH), and
E2F3A and E2F3B intensity values were normalized using tubu-
lin. Half-life was calculated using an online calculator (http://www.
calculator.net/half-life-calculator.html).

cDNA microarray. Normal and tumor total RNA were extracted
from the livers of mice at 12 months of age using TRIZol (Invitrogen).
RNA (5 μg) was submitted to the OSUCCC Genomics Shared Resource
for generation of cDNA and hybridization to Mouse Genome 430 2.0
(MOE4302.0) Affymetrix chips. GeneChip Scanner 3000 (Affymetrix)
was used for data acquisition. Data were processed using Expression
and Transcriptome Analysis Consoles (Affymetrix). Significantly upreg-
ulated probe sets were defined as having a 1.5-fold or greater increase
between knockin (3a<sup>−/−</sup> or 3a<sup>−/−</sup>/3a<sup>−/−</sup>) tumor samples and normal
control liver (3a<sup>−/−</sup>) with P ≤ 0.05 (determined by the software using ANO-
VA). Human HCC gene expression analysis was performed on a public
data set (GSE6764; ref. 44) downloaded from the Gene Expression
Omnibus (GEO) and processed using expression and transcriptome
analysis consoes. Expression heat maps were generated using R.

Database analysis and data mining. Analysis of copy number vari-
ation mutation rates in human cancer was evaluated using the TCGA
cancer dataset (LIHC with the ONCOMINE browser (https://
www.oncomine.org/resource/login.html) and the COSMIC database
on 07/28/2015. For correlation of copy number and gene expression,
patient data from the Liver Hepatocellular Carcinoma (TCGA, provi-
sional) project with putative copy-number alterations from GISTIC and
RNA-seq V2 RSEM values were downloaded from cbioPortal (http://
www.cbioportal.org/). Kaplan-Meier plots showing HCC patient sur-
vival was produced by R package survplot (ref. 56; upregulated expres-
sion was defined as a z-score greater than 2). The log-rank test was per-
formed to show significance between different patient groups.

Data deposition. All original microarray data were deposited in the
NCBI’s GEO (GEO GSE71383).

Statistics. Statistical analyses were performed after consultation
with The Ohio State University Center for Biostatistics using JMP (SAS
Institute) software. Values were found to be significant at P ≤ 0.05.
Individual P values and tests used are described in the figure legends.
Briefly, 2-tailed Student’s t tests were used for comparisons between
2 groups. When more than 2 groups were compared, Wilcoxon tests
with Bonferroni’s correction for multiple tests were used. Fisher’s
exact tests with Bonferroni’s correction for multiple tests were used
to analyze categorical data. For gene ontology analysis, Benjamin-
ni-Hochberg adjusted P value was calculated by the IPA software.
Significance for ChIP-seq and microarray data was calculated by the
analysis software as explained above. Significant changes in gene
expression between control and diseased human liver samples were
determined by Wilcoxon method with Benjamini–Hochberg corre-
pction. Wilcoxon rank sum tests were used to test changes in expres-
sion of individual genes due to E2F1 or E2F3 copy number changes.
To determine whether E2F1/3B targets as a group were increased
due to copy number variations, the rank sums of expression values
for the targets were calculated for E2F1 or E2F3 copy gained patients
and compared with 1,000 sets of randomly sampled genes with the
same size. Significance was determined using the tail probability val-
ue after comparing these 2 groups. For Kaplan-Meier plots, log-rank
tests were used to determine significant changes in patient survival.
For box plots, the boxes represent the first and third quartiles and the
whiskers represent the highest and lowest values. The center lines are
the medians, and outliers are represented with gray dots.

Study approval. Mouse protocols were approved by the institutional
review boards at The Ohio State University. All human data was derived
from publicly available data sets therefore no additional consent is needed.

Author contributions
LNK, SB, SYT, and GL designed the experiments. All authors
performed the experiments and collected and analyzed data. GL
supervised the studies. LNK and GL wrote the manuscript with
input from all authors.

Acknowledgments
We thank J. Bice and D. Bryant for assistance with histology. We
also thank J. McElroy for suggestions on the statistical analysis
of data and Kalpana Ghoshal for the liver cell lines. We are thankful to
Mike Ostrowski, Denis Guttridge, Lawrence Kirschner, and Chris-
tin Burd for critical suggestions. This work was supported by the
OSUCCC Genomics Shared Resource and funded by NIH grants
(R01CA121275 and CA097189 to GL; R01CA098956 to JMP) as well
as a grant from the Dutch Cancer Society (UU 2013-5777 to BW and
ABD). LNK and SB were recipients of an NCI T32 in Cancer Genetics.
CK was a recipient of an NCI T32 in Mouse Models of Human Dis-
ease. CKM was a recipient of a Pelotonia Postdoctoral Fellowship; CK
was a recipient of a Pelotonia Graduate Fellowship; GCM, EP, and YH
were recipients of Pelotonia Undergraduate Fellowships.

Address correspondence to: Gustavo Leone, Medical University
of South Carolina, Hollings Cancer Center, 86 Jonathan Lucas
Street, MSC 955, Charleston, South Carolina 29425, USA. Phone:
843-792-1164. E-mail: Gustavo.leoneg@musc.

GL’s present address is: The Medical University of South Carolina,
Charleston, South Carolina, USA.
RESEARCH ARTICLE


