Published OnlineFirst March 10, 2009; DOI: 10.1158/0008-5472.CAN-08-3357

Research Article

Central Role of c-Myc during Malignant Conversion in Human Hepatocarcinogenesis

Pal Kaposi-Novak,[1] Louis Libbrecht,[2] Hyun Goo Woo,[1] Yun-Han Lee,[1] Nathaniel C. Sears,[1] Elizabeth A. Conner,[1] Valentina M. Factor,[1] Tania Roskams,[2] and Snorri S. Thorgeirsson[1]

[1]Laboratory of Experimental Carcinogenesis, National Cancer Institute, Center for Cancer Research, NIH, Bethesda, Maryland and [2]Departments of Morphology and Molecular Pathology, University of Leuven, Leuven, Belgium

Abstract

Hepatocarcinogenesis is a multistage process in which precursor lesions progress into early hepatocellular carcinomas (eHCC) by sequential accumulation of multiple genetic and epigenetic alterations. To decode the molecular events during early stages of liver carcinogenesis, we performed gene expression profiling on cirrhotic (regenerative) and dysplastic nodules (DN), as well as eHCC. Although considerable heterogeneity was observed at the regenerative and dysplastic stages, overall, 460 differentially expressed genes were detected between DN and eHCC. Functional analysis of the significant gene set identified the *MYC* oncogene as a plausible driver gene for malignant conversion of the DNs. In addition, gene set enrichment analysis revealed global activation of the *MYC* up-regulated gene set in eHCC versus dysplasia. Presence of the *MYC* signature significantly correlated with increased expression of *CSN5*, as well as with higher overall transcription rate of genes located in the 8q chromosome region. Furthermore, a classifier constructed from *MYC* target genes could robustly discriminate eHCC from high-grade and low-grade DNs. In conclusion, our study identified unique expression patterns associated with the transition of high-grade DNs into eHCC and showed that activation of the *MYC* transcription signature is strongly associated with the malignant conversion of preneoplastic liver lesions. [Cancer Res 2009;69(7):2775–82]

Introduction

Extensive morphologic and molecular evidence indicates that carcinogenesis is a multistage process that frequently follows a dysplasia-adenoma-carcinoma sequence (1, 2). Similarly, hepatocarcinogenesis is a multistage process wherein progression is driven by the accumulation of multiple genetic and epigenetic alterations (3). In rodent models, hepatocellular carcinomas (HCC) have been clearly shown to develop from precursor lesions conventionally classified as foci and adenomas (4). Most human HCC evolves on the background of chronic liver diseases and are also preceded by premalignant lesions. The current histologic classification system divides the hepatic lesions into groups of small and/or large cell dysplastic foci and low-grade dysplastic nodules (LGDN) or high-grade DNs (HGDN; diameter < 10 mm; ref. 5). However, diagnosis of the small lesions is difficult, and the probability of malignant transformation could not be clearly determined (6). With the advancement of modern imaging techniques, an increasing number of atypical nodular lesions are encountered in cirrhotic patients, leading to a significant diagnostic and therapeutic challenge (7). Over the recent years, considerable effort has been made to decipher the molecular events of early hepatocarcinogenesis and to discover novel diagnostic markers specific for each of the consecutive phases of the disease (8). Nevertheless, the exact sequence of the molecular events leading to malignant transformation in the preneoplastic hepatic lesions is yet to be discovered, and little is known about the possible involvement of different oncogenic pathways in this process.

In the current study, we analyzed gene expression differences between the consecutive stages of hepatocarcinogenesis to identify common regulatory mechanisms orchestrating malignant transformation. High-density microarrays were used to obtain expression profiles of regenerative LGDNs and HGDNs, as well as early HCCs (eHCC). Furthermore, a comparative functional genomics approach was applied to assess the transcriptional status of different oncogene-activated expression signatures. Our data show that induction of *MYC* target genes occurred ubiquitously during malignant conversion. Moreover, a genomic predictor constructed from cross-species conserved *MYC*-induced genes was able to accurately differentiate HGDN from eHCC with high accuracy. The predictor set was further validated using an independent set of DNs and eHCCs. This large-scale genomic profiling is the first study to uncover the potentially critical role of the *MYC* transcription signature activation in the malignant conversion of preneoplastic lesions in human hepatocarcinogenesis.

Materials and Methods

Collection of human DNs and eHCC samples. We collected biopsies of 49 nodular liver lesions, including 24 cirrhotic (regenerative) nodules (CN), 3 LGDNs, 12 HGDNs, and 10 eHCC, which were located in explanted livers of 10 patients. The freshly removed livers were serially sectioned with 1-cm intervals, and any macroscopically identified nodules were dissected out and split in half. Tissue collected for histologic analysis was fixed in 6% neutral formalin and processed following standard procedures. The other half was immediately snap-frozen in liquid nitrogen–cooled isopentane, embedded in OCT, and stored at −80°C. Histologic classification of the lesions was determined by two expert pathologists (L.L. and T.R.), who evaluated each slide according to criteria established by the International Working Party (5). Microvascular invasion was considered present when tumor cells could be identified inside of the lumen. The clinical characteristics of lesions are summarized in Supplementary Table S1. The majority of the lesions arose in male patients (8 of 10) on ethanol-induced (4 of 10) cirrhotic background; other tumors were associated with chronic HCV

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

Data deposition footnote: GEO accession number GSE12443.

Requests for reprints: Snorri S. Thorgeirsson, Laboratory of Experimental Carcinogenesis, Center for Cancer Research, National Cancer Institute, NIH, 37 Convent Drive, Room 4146, Bethesda, MD 20892. Phone: 301-496-5688; Fax: 301-496-0734; E-mail: snorri_s_thorgeirsson@mail.nih.gov.

doi:10.1158/0008-5472.CAN-08-3357

Published OnlineFirst March 10, 2009; DOI: 10.1158/0008-5472.CAN-08-3357

(3 of 10) or HBV (2 of 10) hepatitis, and one case was associated with α1-antitrypsin deficiency. Four patients also underwent preoperative chemolipidization. The average size of LGDN and HGDN was 5.3 ± 1.5 and 9.3 ± 3.3 mm, respectively (P = 0.018), and the diameter of HCC was15.6 ± 8.5 mm (P = 0.035 versus HGDN; Supplementary Fig. S1). The rate of cell proliferation progressively increased from LGDN to eHCC, although the difference in the number of MIB-1–positive cells between eHCC and HGDN did not reach statistical significance (Supplementary Fig. S1). Vascular invasion with intrahepatic metastases was observed in one eHCC, whereas four HGDNs showed bilirubin and one HGDN showed iron pigment accumulation (Supplementary Table S1).

RNA isolation, amplification, and microarray hybridization. The nodular lesions were identified and manually macrodissected. From each lesion, 5 to 10 frozen sections (8-μm thick) were cut and lysed immediately in RNA extraction buffer. Total RNA isolation, together with DNase digestion, was performed with RNA Easy Mini kit (Qiagen). Integrity of RNA was checked with gel electrophoresis. Total RNA (1 μg) was reversed transcribed and linearly amplified with Aminoallyl MessageAmp II kit (Ambion). Aminoallyl-UTP–modified aRNA product (2 μg) was indirectly labeled with either Cy3 or Cy5 fluorescent dyes. Human Operon v2 oligonucleotide library containing 22K features representing expressed sequences was printed to glass arrays in the Advanced Technology Center (National Cancer Institute). aRNA probes were fragmented and hybridized to microarray slides, following standard procedures. All samples were hybridized against a common amplified reference RNA pooled from normal liver samples. Experimental duplicates were prepared following a reverse flour design. Arrays were scanned with a GenePix 4000B laser scanner, and image analysis was performed with GenePixPro v5 suite.

Quantitative PCR analysis. Real-time PCR quantification of mRNA levels of selected markers, HSPA1A, *GPC3, FOXM1, HDAC2, MYC,* and *CSN5*, was performed on total RNA samples isolated from the same lesions that were used for expression profiling. All ready-to-use human primer sets were obtained from SuperArray. The real-time quantitative PCR (Q-PCR) analysis was performed with ABI Prism 7900HT (Applied Biosystems) Thermal Cycler in a 96-well plate. Melting analysis of the PCR products was conducted to validate amplification of the specific product. Expression level of human glyceraldehyde-3-phosphate dehydrogenase was used as internal reference. Relative gene expression levels were calculated with $2^{-\Delta\Delta C_t}$ method (9).

Tissue culture and small interfering RNA knockdown. Three *CSN5*-specific small interfering RNAs (siRNA; CSN5-1, CSN5-2, and CSN5-3) were designed and tested for growth inhibition in HuH-7 and HepG2 cell lines. siRNAs (15 nmol/L) were complexed with cationic lipids and added to 3×10^3 cells per well in 96-well plates. CSN5-2 siRNA was the most effective in blocking both HuH-7 and HepG2 cell proliferation (68 ± 6.0% and 77 ± 5.1%, respectively) at 4 d after transfection, as measured by 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide assay. Negative-control siRNA induced no significant inhibition of HuH-7 and HepG2 cell growth. Real-time reverse transcription–PCR analysis confirmed that treatment of HuH-7 cells with 15 nmol/L of CSN5-1, CSN5-2, and CSN5-3 siRNAs for 24 h resulted in 73 ± 5.1%, 80 ± 8.2%, and 74 ± 8.1% reduction of target mRNA, respectively, when compared with negative-control siRNA treatment. *CSN5* mRNA expression in HepG2 cells at 24 h was inhibited by 45 ± 10.1%, 46 ± 9.7%, and 51 ± 14.3%, respectively. CSN5-2 siRNA was selected for further studies.

Microarray data analysis. First image spots with a diameter of <10 or >300 μm or a signal intensity below background intensity for any of the two fluorescent channels were excluded. Only genes with at least four data points of six experiments in at least two experimental groups were selected for further data analysis. Gene expression values were normalized by median centering log ratios across all samples. For each spot, target per reference intensity ratios were $\log_2$ transformed and averaged between duplicate experiments. Differentially expressed genes between histologic predefined sample groups were identified by performing a two-sample t test with a random variance model to estimate false discovery rate. Selection criteria for individual genes included a significance level of $P < 0.001$ in the univariate t test. Hierarchical cluster analysis based on Pearson correlation distance was performed with Cluster v2.11, and results were visualized with TreeView programs.[3] For comparative genomic analysis, we selected genes represented on both the mouse and the human platform using curated mammalian orthologues from The Jackson Laboratory.[4] Common genes between human and mouse oncogene-specific signatures and early neoplastic data set were matched by Locus Link ID. Nonparametric gene set enrichment analysis (GSEA) was performed with the GSEA tool developed by Broad Institute.[5] Rank metrics were determined by performing 1,000 random permutations. We tested five class prediction algorithms, including compound covariate predictor, nearest neighbor, nearest centroid, support vector machines, and linear discriminator analysis from BRB-Array Tools[6] to differentiate between eHCC and DN using mouse-derived *MYC* signature genes. The training set contained 10 eHCCs and 12 DNs from our preneoplastic set; the independent validation set included 14 HBV-related eHCCs and 14 DNs. To build an optimized classifier list, which could estimate the probability of the identity of a particular sample, we used a leave-one-out cross-validation approach. During the cross-validation step 1, sample was removed from the analysis and the remaining samples were used to identify the most differentially expressed genes between the groups. Based on the expression of these genes, identity of the left out sample was predicted with a given algorithm. This process was repeated until each sample was left out once. The number of genes in the classifier was varied to provide the highest correct prediction rate in the training set. To estimate accuracy of the prediction model, class labels were randomly permuted and the leave-one-out cross-validation process was repeated 1,000 times. The significance level is the proportion of random permutations that gave a cross-validated error rate no greater than the cross-validated error rate obtained with the real data.

Results

Changes in global gene expression profiles during early hepatocarcinogenesis. We performed detailed transcriptomic profiling of 49 liver lesions, including 24 CNs, 3 LGDNs, and 12 HGDNs, as well as 10 eHCCs. RNA, isolated from macrodissected lesions, was subjected to linear amplification and indirect fluorescent labeling and hybridized to oligonucleotide microarrays containing ~22,000 features representing ~19,000 unique tags. After data normalization, 12,994 features with at least 10 valid expression values across all samples were selected for further analysis. To assess the overall expression differences between the consecutive stages of early hepatocarcinogenesis, we performed hierarchical clustering with 1,632 genes showing at least a 2-fold regulation in >10% of the samples. This analysis revealed that most of eHCCs, with exception of two misclassified tumors, clustered together and could be readily differentiated from the remaining samples based on their distinctive expression patterns (Supplementary Fig. S2*A* and *B*). On the other hand, profiles from dysplastic and regenerative stages displayed considerable heterogeneity and were often closer to eHCC rather than to normal livers. Interestingly, the median variance of the expression levels calculated for the 12,994 genes was only slightly higher in the HCC group (0.091) than in the dysplastic (0.08) and regenerative nodules (0.062). Also, there were no substantial differences between groups in the number of genes showing 2-fold regulations relative to normal liver reference (CN 231, DN 234, eHCC 267), suggesting that transcriptomic differences across the early stages were relatively small.

[3] http://rana.lbl.gov/EisenSoftware.html
[4] http://jax.org/
[5] http://broad.mit.edu/gsea/
[6] http://linus.nci.nih.gov/BRB-ArrayTools.html

Published OnlineFirst March 10, 2009; DOI: 10.1158/0008-5472.CAN-08-3357

Stage-related expression patterns identify early markers of neoplastic progression. To evaluate the presence of unique expression patterns associated with the consecutive stages of HCC formation, we performed a supervised class comparison analysis by classifying samples according to their histology. We focused primarily on two major transition points: (*a*) from regenerative nodule to LGDN and HGDN and (*b*) from dysplastic lesions to eHCC. Significant genes ($P < 0.001$) were selected using a two-sample *t* test with a random variance model. Comparison of LGDN and HGDN with the regenerative nodules yielded 44 and 41 differentially expressed genes, respectively (Fig. 1*A*). Thirty-four of these genes had lower expression levels whereas 51 genes had higher expression levels in dysplastic livers compared with cirrhotic livers (Supplementary Table S2*A*). Functional classification of the significant gene sets based on Gene Ontology annotations showed that dysplastic lesions were characterized by overrepresentation of genes involved in cell-cell interactions (*ITGB4, CYR61, MADCAM1, DSG2*) and possessing electron transporter activity (*NDUFA1, RP11-217H1, NCF1, PDCL, PLOD1*), reaching observed/expected ratios (O/E) of 7.52 and 8.17, respectively (Supplementary Table S3*A*).

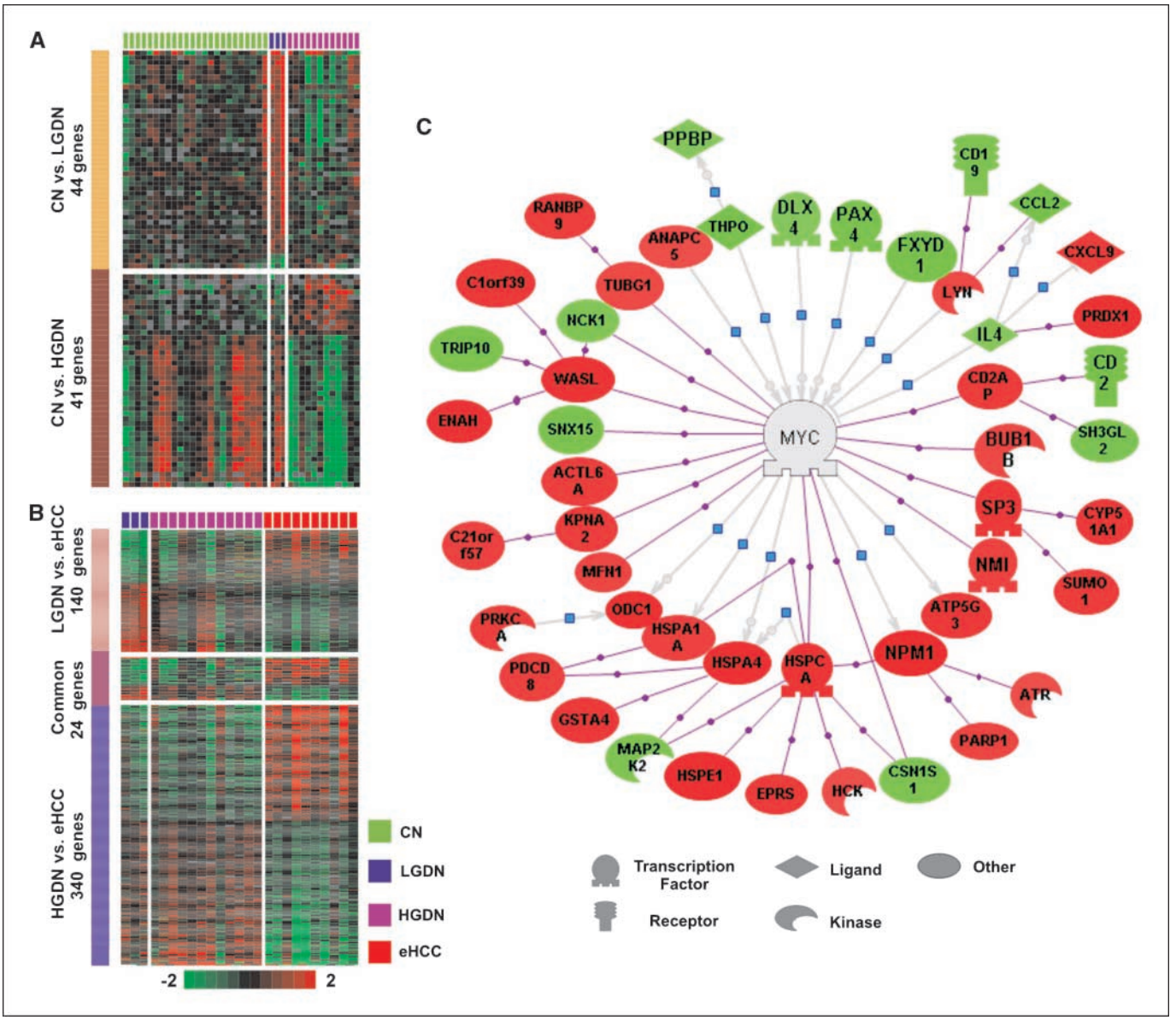

Figure 1. Expression patterns of genes differentially expressed between the consecutive stages of early human hepatocarcinogenesis. *A,* two-dimensional heat map represents profiles from 85 genes with significantly different expression levels between CNs and DNs. *B,* similar heat maps were constructed from the 460 genes differentially expressed between the dysplastic lesions and eHCCs. Progression of early hepatocarcinogenesis coincides with increase in the transcriptomic alterations between the consecutive stages. Genes were selected by setting the cutoff level in a univariate *t* test at $P < 0.001$. *Columns,* individual samples; *rows,* genes. Each cell reflects the $\log_2$-transformed expression level of an individual gene in a given sample, as indicated by the color code on the figure. The probability of having the same number of genes by chance was also determined with multivariate test after 1,000 random permutations and marked in the legends. *C,* gene connectivity network centered on the *MYC* transcription factor is identified by the PathwayAssist tool. Genes differentially expressed between the dysplastic lesions and eHCCs (Supplementary Table S2) were used to detect the dominant regulatory networks during malignant transformation. Although *MYC* itself is not differentially expressed (*gray*) between the two histologic categories, 48 of its immediate targets are either significantly more (*red*) or less expressed (*green*) in eHCC than in liver dysplasia. *Gray arrows with blue rectangle,* known transcriptional activation; *purple arrows,* direct binding between the connected nodes.

Published OnlineFirst March 10, 2009; DOI: 10.1158/0008-5472.CAN-08-3357

Transformation of DN into eHCC was marked by even more pronounced transcriptomic alterations (Fig. 1*B*). Class comparison identified 144 significantly different genes between LGDN and eHCC and 282 genes between HGDN and eHCC, with 34 genes being commonly significant in both comparisons. The differentially expressed set contained up-regulated (210) and down-regulated (248) genes in approximately equal abundance (Supplementary Table S2*B*). From the structural Gene Ontology categories (Supplementary Table S3*B*), actin cytoskeleton (*ARPC3, ENAH, TMSB10, CCT3, PARVB, PLS1, ACTL6A, WASL*; O/E, 2.35), microtubule (*KPNA2, TUBA1B, TUBG1, BIRC5, TUBB3*; O/E, 7.29), and ribosome (*DAP3, RPL24, NPM1, MRPL24, MRPL13, MRLP15, MRLP18*; O/E, 7.11)–related clusters were the most prominently deregulated, and the majority of these genes were more expressed in eHCC. The functional classification found overrepresentation of genes encompassing enzymes with monooxygenase activity (*CYP51A1, CYP2C8, CYP4A11, CYP1A2, CYP4V2, CYP2C9*; O/E, 5.09) or implicated in cholesterol, nucleotide, and ribonucleotide biosynthesis (Supplementary Table S3*B*).

To confirm the microarray results by an independent method and to identify sensitive early markers of malignant conversion in hepatocytic lesions, we validated expression levels of four differentially expressed genes (*HSPA1A, GPC3, FOXM1*, and *HDAC2*) by Q-PCR analysis (Supplementary Fig. S3*A–D*). The data obtained with microarray and PCR techniques were highly concordant, and all selected markers showed significantly ($P < 0.05$, Student's t test) increased expression in eHCC versus dysplastic stage.

Pathway and gene set enrichment analyses commonly implicate *MYC* transcription factor as a potential regulator of malignant conversion. To reveal the dominant intracellular signaling networks in the malignant conversion of dysplastic liver lesions, we performed pathway analysis using 460 genes differentially expressed between DNs and carcinomas. Using the PathwayAssist Analysis tool, several connectivity maps were constructed based on the previously reported interactions between members of the significant gene set. The analysis outlined an extensive regulatory network centered on the *MYC* gene (Fig. 1*C*). Overrepresentation of *MYC* targets among differentially expressed transcripts also favors that *MYC* activation might be essential for transition to eHCC.

To validate the regulatory role of *MYC* during malignant conversion, we further interrogated our human preneoplastic data set using independent MYC signatures. Previously, we obtained a prominent liver-specific *MYC* expression signature in a transgenic mouse model (10). In addition, we analyzed *MYC* signatures derived from human umbilical vascular endothelial cells (HUVEC) (11) and human breast epithelial cells (also including RAS target genes; ref. 12). Specificity of the many *MYC* target genes identified

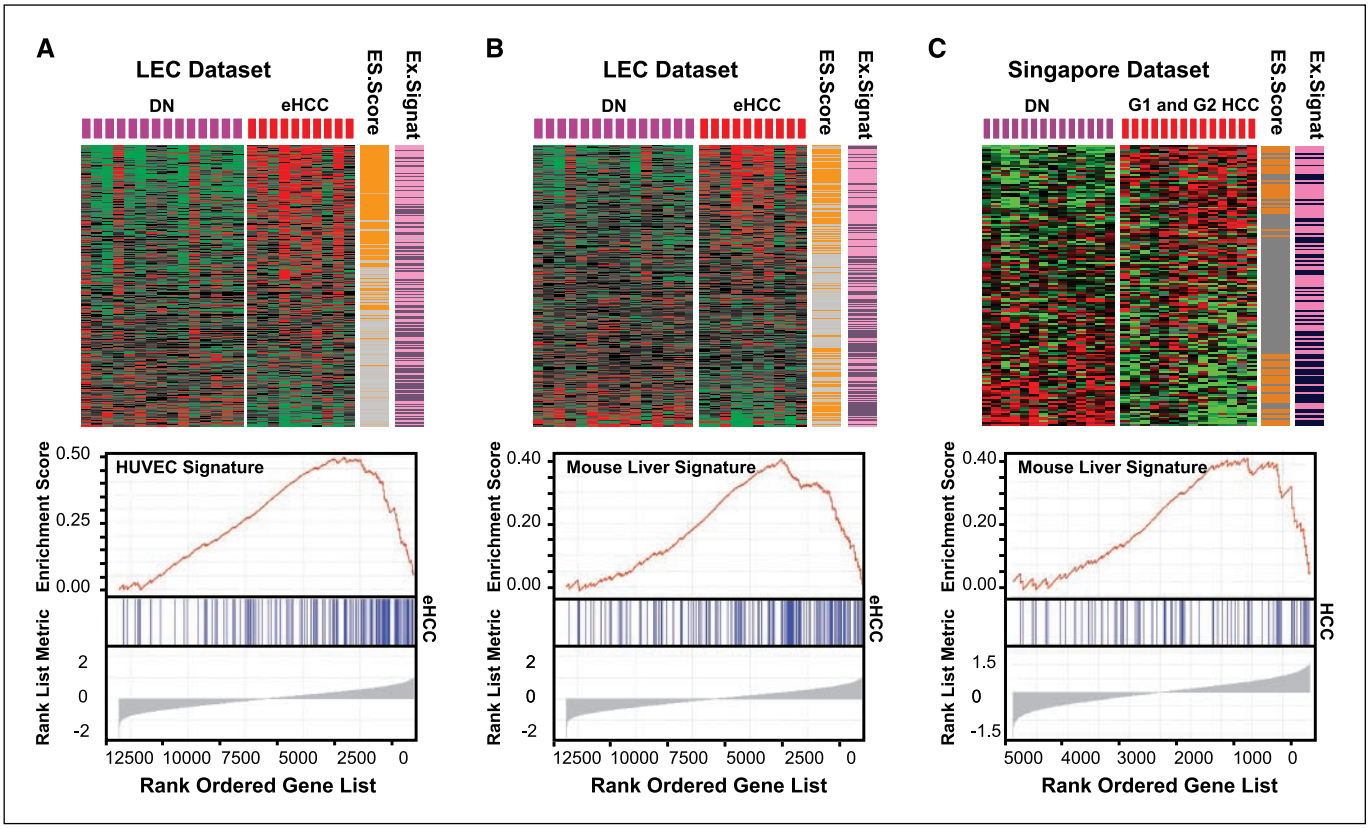

Figure 2. Enrichment analysis of the *MYC* expression signatures in the eHCCs. Four *MYC*-regulated gene sets (two up-regulated and two down-regulated) were identified in *MYC* transgenic mice and human endothelial cells transfected with an adenovirus-MYC vector. Heat maps representing gene expression in the early hepatocarcinogenesis data set (*LEC*) revealed that both the HUVEC-derived (*A*) and mouse liver–derived (*B*) *MYC* signatures show increased activation in the eHCC samples versus DNs. The liver-specific *MYC* signature displayed similar activation (*C*) in an independent data set containing dysplastic lesions and early carcinomas from HBV-infected patients. *Columns,* samples; *rows,* $\log_2$-transformed expression ratios of individual genes, as defined by the color code on the figure. *Colored bars,* genes with significant core enrichment (*orange*). Up-regulation (*pink*) and down-regulation (*dark blue*) of the gene in the original signatures. Distribution and ES of the rank-ordered *MYC* up-regulated genes in HUVEC, LEC, and HBV-related data sets. Both up-regulated gene sets are significantly enriched in the eHCC compared with the DNs.

Published OnlineFirst March 10, 2009; DOI: 10.1158/0008-5472.CAN-08-3357

in the HUVECs were validated with microarray, Q-PCR, and chromatin immunoprecipitation assays (11), whereas breast epithelial signature was successfully applied for the classification of human lung, breast, and ovarian carcinomas (12). Based on the list of curated homologous mouse and human UniGene clusters, we matched corresponding features from the different platforms. Next, the expression of six orthologous oncogene-specific signatures, including the *MYC* up-regulated and down-regulated gene sets identified in transgenic mice and in HUVECs, as well as *RAS* up-regulated and down-regulated gene sets detected in the breast epithelial cells, was assessed with GSEA (13). This nonparametric method calculates enrichment of a signature in a group of samples based on the location of its members in the rank-ordered gene list. We found that both HUVEC-derived [enrichment score (ES) = 0.5, $P < 0.03$] and liver-specific (ES = 0.4, $P < 0.05$) *MYC* up-regulated gene sets were significantly enriched in the eHCC versus the dysplastic samples (Fig. 2*A* and *B*). In contrast, neither MYC down-regulated sets (Supplementary Fig. S4*A*–*C*) nor RAS up-regulated genes (ES = 0.12, $P < 1$) showed enrichment in any group (Supplementary Table S4*A*). Family-wise error rates *P* values and normalized ES (NES) calculated in the multivariate analysis also suggested enrichment of the *MYC* up-regulated sets in the carcinomas (mouse NES = 1.46, $P < 0.15$; HUVEC NES = 1.53, $P < 0.08$), although they did not reach significance probably due to the small size of the gene sets. In contrast, we could not observe apparent enrichment of any of the down-regulated sets. Furthermore, none of the oncogene-dependent signatures displayed enrichment when GSEA was performed to compare dysplastic versus regenerative nodules (data not shown). Together, these results argue convincingly that activation of *MYC* target genes has a decisive role in the malignant transformation of dysplastic lesions.

***MYC* up-regulated gene expression signature can distinguish eHCC from dysplastic lesions.** To test the universal significance of our findings, we performed GSEA on an independent group of 14 DNs and 14 eHCCs (grades 1 and 2) samples developed on a background of chronic HBV hepatitis (Fig. 2*C*). After overlapping the gene sets, we verified that *MYC* up-regulated gene sets were indeed enriched in eHCC versus hepatic dysplasia. However, the transgenic mouse liver–derived specific signature (ES = 0.47, $P < 0.05$) displayed the most significant enrichment (Supplementary Table S4*B*).

We also tested whether the *MYC* signature genes could differentiate between hepatic dysplasia and eHCC in a prediction model. Because mouse liver–specific signature showed the most consistent enrichment pattern across all early neoplastic data sets, we used these genes to construct a *MYC*-dependent classifier. Five different supervised prediction algorithms, including compound covariate predictor, nearest neighbor, nearest centroid, support vector machines, and linear discriminator analysis methods were used to build a prediction model, following a leave-one-out cross-validation strategy. The training set included 10 eHCCs and 12 DNs from our collection. The optimal classifier, producing the highest correct classification rate in the training set, contained 36 liver-specific *MYC* target genes (Supplementary Table S5*A*). Among the classifier genes, ribosomal structural proteins were most abundant (O/E = 3.2). The gene list could achieve 82% to 86% correct classification rate in the training set, depending on the algorithm (Supplementary Table S5*B*). We also applied the same classifier to the validation set of 14 DNs and 14 HCCs (grades 1 and 2) associated with chronic HBV hepatitis. The prediction rates were somewhat lower in the validation set, but the compound covariate predictor (85%) and support vector machine (79%) could successfully classify eHCCs from DNs (Supplementary Table S5*C*). These results not only confirmed the almost universal activation of the *MYC* signature at the eHCC stage but also strongly indicate that the *MYC*-driven expression signature could be used successfully for early cancer diagnosis.

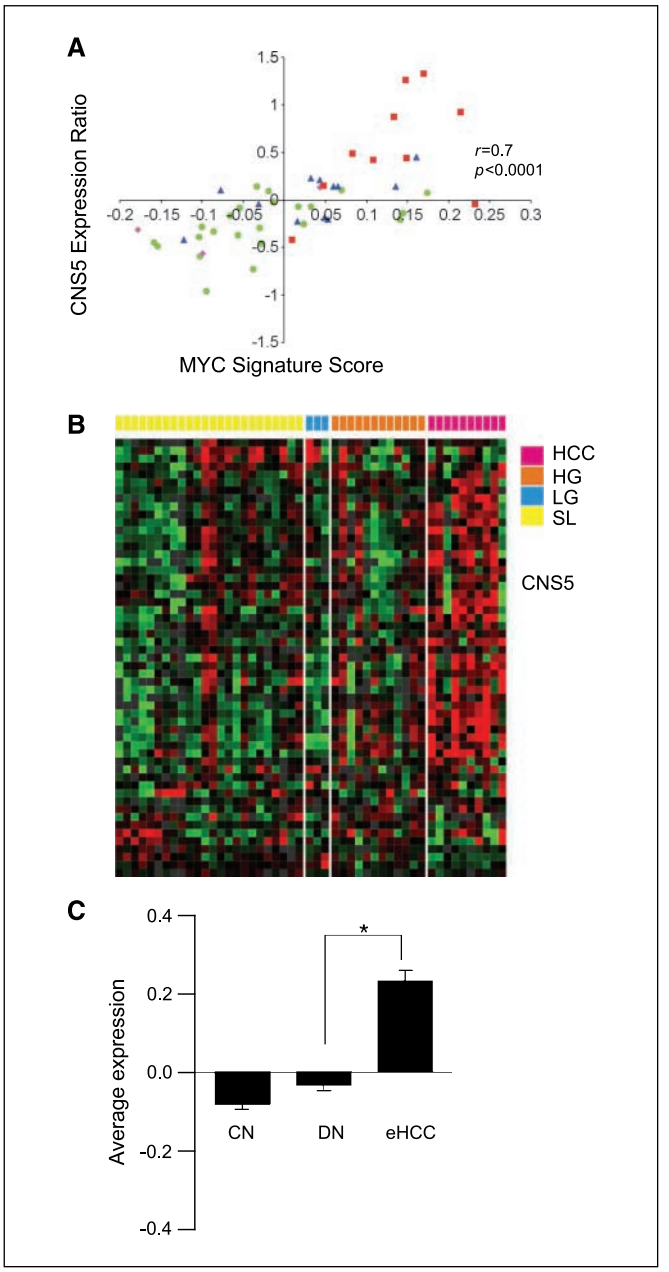

Figure 3. Correlation of *MYC* signature with *CSN5* expression and 8q amplification score. *A,* expression levels of *CSN5* in human preneoplastic liver lesions and eHCC showed a nearly linear correlation (Pearson $r = 0.7$, $P < 0.0001$) with a *MYC* signature score. A signature score for each sample was calculated from the correlation between expression of signature genes and canonical *MYC* signature. *B,* heat map shows the expression profile of the down-regulated genes ($n = 83$, $P < 0.01$) by *CSN5* knockdown experiment. The genes with missing values of >30% were excluded. Heat map colors represent relative expression levels (gene centered) across the samples. *C,* average of expression differences (gene centered) of down-regulated genes by *CSN5* knockdown treatment between CN, DN, and eHCC. *Columns,* mean; *bars,* SE (*, $P < 0.001$, two-sample Student's *t* test).

Published OnlineFirst March 10, 2009; DOI: 10.1158/0008-5472.CAN-08-3357

The exact mechanism leading to aberrant activation of *MYC* and its target genes remains undetermined. In previous studies, frequent DNA copy number aberrations were identified in many cancers, including HCC (3). Therefore, we tested whether *MYC* activation in HCC was due to the copy number gain. The effect of regional copy number changes can be inferred by calculating the regional concordance of gene expressions. In this context, we calculated the gene expression differences between DN and eHCC for each cytoband. A total of 30 differentially expressed cytobands were identified ($P < 0.001$, two-sample t test; Supplementary Table S6). As expected, many of the known regions with copy number alterations in HCCs (e.g., 1q21-22,6p21, 8q22-24, 11q11-14, 19q13, 20p12; Supplementary Table S6) were differentially expressed between DNs and eHCCs. In particular, 1q21-22 and 8q22-24 were overexpressed in eHCCs compared with DNs (Supplementary Table S6). The "early stage" copy number gains, especially 8q24, including MYC gene, may play a critical role in the malignant conversion of DN.

Although the mutation status of *MYC* is involved in cancer development and progression, no mutations in *MYC* genes were found in the 10 eHCC samples by sequencing analysis (data not shown). However, a recent study has revealed that *CSN5*, a member of COP9 signalosome, is linked to *MYC* activation and wound healing signature in breast carcinomas (14). We also observed that *CSN5* expression was significantly correlated with the average expression level of the *MYC* signature ($\gamma = 0.72$, $P < 1 \times 10^{-5}$, Pearson's correlation test) implying its regulatory role in *MYC* activation (Fig. 3*A*). Thus, we examined the association of *CSN5* with the malignant conversion of DN. We generated *CSN5* regulating signature by inhibiting *CSN5* gene expression using siRNA technology. The human HCC cell line HuH-7 was transfected with the *CSN5* siRNA, and the differentially expressed genes were identified by gene expression profiling (for details, see Materials and Methods). Eighty-three genes, including *CSN5*, were identified as putative *CSN5* target genes, which were significantly down-regulated by *CSN5* knockdown ($P < 0.01$, two-sample t test). The average expression levels of these genes were significantly elevated in eHCC compared with those in DN or CN ($P < 0.001$, two-sample t test; Fig. 3*B* and *C*). Taken together, these data suggest that *CSN5* and *MYC* activation, possibly caused by copy number gains at 8q, may play a driver role in the malignant conversion of DN to eHCC.

Discussion

To obtain a comprehensive picture of the global transcriptomic alterations during early steps of human hepatocarcinogenesis, we performed expression profiling of 49 nodular liver lesions ranging from regenerative and DNs to eHCC. Our data revealed unique expression patterns associated with the consecutive stages of hepatocarcinogenesis and identified potential markers, such as *HSPA1A* and *CSN5*, which may facilitate an early diagnosis of liver cancer. By using the most advanced pathway prediction and comparative genomics analysis tools, we showed that malignant transformation of preneoplastic liver lesions coincides with the universal activation of the *MYC*-regulated expression signature. Moreover, based on a set of liver-specific *MYC* target genes, we were able to construct a genomic prediction model that could successfully differentiate between DN and eHCC.

The role of *MYC* in hepatocarcinogenesis has been extensively investigated. Chromosome gains at *MYC* locus are among the most frequently reported genetic abnormalities in advanced human HCCs (3). Amplification of 8q22-24 region is also observed in 40% to 60% of eHCCs, although it only affects a small percentage of DNs (15). A cytogenetic tumor progression model constructed by Poon and colleagues determined that gains of 8q22-24 are among the earliest genomic events associated with HCC development (16). In animal models, *MYC* also proved to be a key contributor to hepatic carcinogenesis. Overexpression of *MYC* in mouse liver results in cellular dysplasia followed by tumor formation (3). Recently, the effect of *MYC* inactivation has been evaluated in different tetracycline-inducible transgenic systems (17). Inactivation of *MYC* in invasive HCCs led to sustained tumor regression with concomitant proliferation arrest, differentiation, and apoptosis of tumor cells (18). T-cell and B-cell lymphomas, as well as carcinomas of breast, skin, and pancreas, also displayed varying degrees of *MYC* oncogene dependence (17). Previously, using a comparative functional genomic approach, we established a clear molecular relationship between human HCCs and mouse liver tumors arising on a *MYC* transgenic background. In contrast, expression patterns of ciprofibrate-induced and $Acox^{-/-}$ tumors displayed less similarities with human carcinomas (19). Thus, our current findings implicating *MYC* as a central mediator of human hepatocarcinogenesis are potentially consistent with a more universal tumorgenesis model wherein *MYC* activation is required for maintenance and expansion of transformed cells.

The average expression values of genes located in the 8q region were elevated in eHCCs and displayed a strong correlation with the presence of the *MYC* up-regulated signature. Intriguingly, neither we nor other investigators were able to detect a concomitant overexpression of the *MYC* gene itself (20). Alternatively, *MYC* activity is affected by posttranscriptional modifications affecting the half-life of the protein. Phosphorylation of the Thr-58 residue in the Myc box I domain targets *MYC* for ubiquitination and consequent proteasosome-mediated degradation (21). Mutations of this region are frequently observed in Burkitt lymphomas and lead to stabilization of *MYC*, as well as to disruption of its proapoptotic function (22). However, it is unlikely that direct *MYC* mutations have a significant role in the early stages of hepatocarcinogenesis, as sequencing of *MYC* transcript in HCC samples did not reveal any genetic alteration. A member of the COP9 signalosome, *CSN5* (alternative symbols are *JAB1* or *COPS5*), located at 8q13 locus, regulates activity of the ubiquitin ligase complex (21). Surprisingly, *CSN5* proved to be an essential activator of *MYC* transcriptional activity in nontransformed breast epithelial cells by increasing its turnover (14). In advanced HCC, high expression levels of *CSN5* significantly correlated with an increase in gene copy numbers (23). Importantly, using comparative genomics analysis, we showed that *CSN5* overexpression occurred at the early stages of hepatocarcinogenesis and showed a significant association with the presence of the *MYC*-regulated expression signature (Fig. 3). These results are consistent with the notion that *CSN5* plays an important role in liver cancer progression by a mechanism involving enhanced transcriptional activation of *MYC* targets. Fast-growing tumors were often characterized by tissue hypoxia, resulting in the activation of hypoxia-inducible factors *HIF1* and *HIF2*. The interaction between the *HIF* and *MYC* transcription factors is well established and may alternatively explain aberrant activation of the *MYC* signature (24). Thus, under hypoxic conditions, expression of the *MYC*-regulated metabolic genes may provide tumor cells with a significant growth advantage. However, considering the relatively small size of the eHCCs investigated in this study, tissue hypoxia

is more likely to be a contributing, rather than a major, cause of *MYC* activation.

Multiple attempts have been made to identify reliable tissue and serum markers that are able to differentiate between eHCC and other nodular lesions. Immunohistochemical studies validated *HSPA1A* (25) and *CSN5* (26) as potential biomarkers of eHCC. Our results confirm these findings and indicate a regulatory connection with *MYC*. Thus, *HSPA1A* and *CSN5* exemplify elements of the *MYC* signature that have already been detected in previous studies conducted on early hepatocarcinogenesis. We also constructed a 36–*MYC* gene–regulated classifier, which could predict the eHCC with 75% to 85% accuracy. Previously, other groups described that, in HBV-infected livers, a set of "grade-associated" genes could differentiate between LGDN and HGDN, as well as eHCC, with 100% accuracy (27). The discrepancy between the findings could be, in part, attributed to the smaller number and the different etiology of the LGDN analyzed in the current study. Still, during the validation of our model using independent samples sets, we showed that the 36 gene signature could be successfully applied to lesions with diverse etiologic background. However, the novelty of our approach goes beyond the description of a novel genomic classifier for eHCC. By using a comparative genomic approach, we provide comprehensive evidence for the central role of *MYC* in the malignant transformation of preneoplastic liver lesions.

In addition to the *MYC* signature, we also assessed the distribution of other oncogene-specific target gene sets, such as *RAS* and *β-catenin.* Surprisingly, neither *RAS*-induced nor *β-catenin*[7]–induced genes showed a significant enrichment in the eHCC samples. Although considerable evidence supports the importance of RAS pathway in promoting liver carcinogenesis (28), *RAS* mutations, which frequently occur in other epithelial malignancies, are almost never detected in HCCs (29). Deregulation of the *WNT/β-catenin* pathway, however, is clearly a hallmark of HCCs (30), although not detected at the initial stages (31). It is possible that these pathways do not take part in the earliest phase of liver malignant transformation. Alternatively, the lack of cross-species or cross-tissue conservation might be responsible for the lack of *RAS* and *β-catenin* signatures in the eHCC data set. We are also cognizant of the fact that transcriptional activity of *MYC* itself could be regulated by multiple pathways, including *RAS/RAF/MPAK* (32), *JAK/STAT*, and *WNT/β-catenin* (33–35) signaling, which may result in a significant overlap between the *MYC* and other oncogene-induced signatures. Similar observations have been made by Sansom and colleagues, who found that the inducible deletion of *MYC* in $APC^{-/-}$ intestinal epithelial cells leads to the loss of β-catenin expression signature together with the reversal of APC-deficient phenotype (36). Thus, the *MYC* signature might, in part, represent common effectors of proliferative stimuli with different origins. Indeed, a common induction of several cell cycle–related genes was observed in eHCCs in an independent report (37).

Previously, we successfully applied comparative functional genomic tools to identify subclasses of human HCC with progenitor cell origin (38) and with dominant activation of *MET* gene (39). Similar analysis conducted on atypical liver nodules confirms the power of this approach and shows a universal transcriptional activation of *MYC* target genes at the eHCC stage. In conclusion, the combination of genome-wide expression profiling with a comparative genomics approach not only produced a prediction model that could significantly facilitate early diagnosis of liver cancer but also permitted the identification of *MYC* as a central regulator of malignant transformation in early hepatocarcinogenesis. In the future, this strategy may provide a molecular classification system embracing the full spectra of preneoplastic and neoplastic liver lesions, the cellular origin of HCC, and may contribute to the identification of novel diagnostic and therapeutic targets.

[7] Own unpublished observations.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Received 9/8/08; revised 12/10/08; accepted 1/2/09; published OnlineFirst 3/10/09.

Grant support: NIH intramural research program.

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

We thank Dr. Joe W. Grisham for the thoughtful discussion of the manuscript.

References

1. De Marzo AM, Meeker AK, Zha S, et al. Human prostate cancer precursors and pathobiology. Urology 2003;62:55–62.
2. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61:759–67.
3. Thorgeirsson SS, Grisham JW. Molecular pathogenesis of human hepatocellular carcinoma. Nat Genet 2002;31:339–46.
4. Grisham JW. Interspecies comparison of liver carcinogenesis: implications for cancer risk assessment. Carcinogenesis 1997;18:59–81.
5. Terminology of nodular hepatocellular lesions. International Working Party. Hepatology 1995;22:983–93.
6. Kojiro M, Roskams T. Early hepatocellular carcinoma and dysplastic nodules. Semin Liver Dis 2005;25:133–42.
7. Lencioni R, Cioni D, Della Pina C, Crocetti L, Bartolozzi C. Imaging diagnosis. Semin Liver Dis 2005;25:162–70.
8. Bruix J, Hessheimer AJ, Forner A, Boix L, Vilana R, Llovet JM. New aspects of diagnosis and therapy of hepatocellular carcinoma. Oncogene 2006;25:3848–56.
9. Pfaffl MW, Horgan GW, Dempfle L. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 2002;30:e36.
10. Coulouarn C, Gomez-Quiroz LE, Lee JS, et al. Oncogene-specific gene expression signatures at preneoplastic stage in mice define distinct mechanisms of hepatocarcinogenesis. Hepatology 2006;44:1003–11.
11. Menssen A, Hermeking H. Characterization of the c-MYC-regulated transcriptome by SAGE: identification and analysis of c-MYC target genes. Proc Natl Acad Sci U S A 2002;99:6274–9.
12. Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature 2006;439:353–7.
13. Sweet-Cordero A, Mukherjee S, Subramanian A, et al. An oncogenic KRAS2 expression signature identified by cross-species gene-expression analysis. Nat Genet 2005;37:48–55.
14. Adler AS, Lin M, Horlings H, Nuyten DS, van de Vijver MJ, Chang HY. Genetic regulators of large-scale transcriptional signatures in cancer. Nat Genet 2006;38:421–30.
15. Zondervan PE, Wink J, Alers JC, et al. Molecular cytogenetic evaluation of virus-associated and non-viral hepatocellular carcinoma: analysis of 26 carcinomas and 12 concurrent dysplasias. J Pathol 2000;192:207–15.
16. Poon TC, Wong N, Lai PB, Rattray M, Johnson PJ, Sung JJ. A tumor progression model for hepatocellular carcinoma: bioinformatic analysis of genomic data. Gastroenterology 2006;131:1262–70.
17. Arvanitis C, Felsher DW. Conditional transgenic models define how MYC initiates and maintains tumorigenesis. Semin Cancer Biol 2006;16:313–7.
18. Shachaf CM, Kopelman AM, Arvanitis C, et al. MYC inactivation uncovers pluripotent differentiation and tumour dormancy in hepatocellular cancer. Nature 2004;431:1112–7.
19. Lee JS, Chu IS, Mikaelyan A, et al. Application of comparative functional genomics to identify best-fit mouse models to study human cancer. Nat Genet 2004;36:1306–11.
20. Chan KL, Guan XY, Ng IO. High-throughput tissue microarray analysis of c-myc activation in chronic liver diseases and hepatocellular carcinoma. Hum Pathol 2004;35:1324–31.

Published OnlineFirst March 10, 2009; DOI: 10.1158/0008-5472.CAN-08-3357

21. Wei N, Deng XW. The COP9 signalosome. Annu Rev Cell Dev Biol 2003;19:261–86.
22. Hemann MT, Bric A, Teruya-Feldstein J, et al. Evasion of the p53 tumour surveillance network by tumour-derived MYC mutants. Nature 2005;436:807–11.
23. Patil MA, Gutgemann I, Zhang J, et al. Array-based comparative genomic hybridization reveals recurrent chromosomal aberrations and Jab1 as a potential target for 8q gain in hepatocellular carcinoma. Carcinogenesis 2005;26:2050–7.
24. Dang CV, Kim JW, Gao P, Yustein J. The interplay between MYC and HIF in cancer. Nat Rev 2008;8:51–6.
25. Di Tommaso L, Franchi G, Park YN, et al. Diagnostic value of HSP70, glypican 3, and glutamine synthetase in hepatocellular nodules in cirrhosis. Hepatology 2007;45:725–34.
26. Takami T, Terai S, Yokoyama Y, et al. Human homologue of maid is a useful marker protein in hepatocarcinogenesis. Gastroenterology 2005;128:1369–80.
27. Nam SW, Park JY, Ramasamy A, et al. Molecular changes from dysplastic nodule to hepatocellular carcinoma through gene expression profiling. Hepatology 2005;42:809–18.
28. Calvisi DF, Ladu S, Gorden A, et al. Ubiquitous activation of Ras and Jak/Stat pathways in human HCC. Gastroenterology 2006;130:1117–28.
29. Challen C, Guo K, Collier JD, Cavanagh D, Bassendine MF. Infrequent point mutations in codons 12 and 61 of ras oncogenes in human hepatocellular carcinomas. J Hepatol 1992;14:342–6.
30. Villanueva A, Newell P, Chiang DY, Friedman SL, Llovet JM. Genomics and signaling pathways in hepatocellular carcinoma. Semin Liver Dis 2007;27:55–76.
31. Park JY, Park WS, Nam SW, et al. Mutations of β-catenin and AXIN I genes are a late event in human hepatocellular carcinogenesis. Liver Int 2005;25:70–6.
32. Sears R, Leone G, DeGregori J, Nevins JR. Ras enhances Myc protein stability. Mol Cell 1999;3:169–79.
33. Liu J, Levens D. Making myc. Curr Top Microbiol Immunol 2006;302:1–32.
34. He TC, Sparks AB, Rago C, et al. Identification of c-MYC as a target of the APC pathway. Science 1998;281:1509–12.
35. Kolligs FT, Bommer G, Goke B. Wnt/β-catenin/tcf signaling: a critical pathway in gastrointestinal tumorigenesis. Digestion 2002;66:131–44.
36. Sansom OJ, Meniel VS, Muncan V, et al. Myc deletion rescues Apc deficiency in the small intestine. Nature 2007;446:676–9.
37. Wurmbach E, Chen YB, Khitrov G, et al. Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma. Hepatology 2007;45:938–47.
38. Lee JS, Heo J, Libbrecht L, et al. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. Nat Med 2006;12:410–6.
39. Kaposi-Novak P, Lee JS, Gomez-Quiroz L, Coulouarn C, Factor VM, Thorgeirsson SS. Met-regulated expression signature defines a subset of human hepatocellular carcinomas with poor prognosis and aggressive phenotype. J Clin Invest 2006;116:1582–95.

Correction: Article on Activation of the MYC Signature in Early HCC

In the article on activation of the MYC signature in early HCC in the April 1, 2009 issue of *Cancer Research* (1), Dr. Cedric Coulouarn should have been included as the sixth author.

1. Kaposi-Novak P, Libbrecht L, Woo HG, Lee Y-H, Sears NC, Coulouarn C, Conner EA, Factor VM, Roskams T, Thorgeirsson SS. Central role of c-Myc during malignant conversion in human hepatocarcinogenesis. Cancer Res 2009;69:2775–82.

Published OnlineFirst 4/7/09.

doi:10.1158/0008-5472.CAN-69-8-COR3

Cancer Research
The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

AACR American Association for Cancer Research

Central Role of c-Myc during Malignant Conversion in Human Hepatocarcinogenesis

Pal Kaposi-Novak, Louis Libbrecht, Hyun Goo Woo, et al.

Cancer Res 2009;69:2775-2782. Published OnlineFirst March 10, 2009.