Identification of Potential Driver Genes in Human Liver Carcinoma by Genomewide Screening

Hyun Goo Woo,1 Eun Sung Park,2 Ju-Seog Lee,2 Yun-Han Lee,1 Tsuyoshi Ishikawa,1 Yoon Jun Kim,1 and Snorri S. Thorgeirsson1

1Laboratory of Experimental Carcinogenesis, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, Maryland; 2Department of Systems Biology, Division of Cancer Medicine, University of Texas M. D. Anderson Cancer Center, Houston, Texas; and Department of Internal Medicine and Liver Research Institute, Seoul National University

Abstract

Genomic copy number aberrations and corresponding transcriptional deregulation in the cancer genome have been suggested to have regulatory roles in cancer development and progression. However, functional evaluation of individual genes from lengthy lists of candidate genes from genomic data sets presents a significant challenge. Here, we report effective gene selection strategies to identify potential driver genes based on systematic integration of genome scale data of DNA copy numbers and gene expression profiles. Using regional pattern recognition approaches, we discovered the most probable copy number–dependent regions and 50 potential driver genes. At each step of the gene selection process, the functional relevance of the selected genes was evaluated by estimating the prognostic significance of the selected genes. Further validation using small interference RNA–mediated knockdown experiments showed proof-of-principle evidence for the potential driver roles of the genes in hepatocellular carcinoma progression (i.e., NCSTN and SCRIB). In addition, systemic prediction of drug responses implicated the association of the 50 genes with specific signaling molecules (mTOR, AMPK, and EGFR). In conclusion, the application of an unbiased and integrative analysis of multidimensional genomic data sets can effectively screen for potential driver genes and provides novel mechanistic and clinical insights into the pathobiology of hepatocellular carcinoma. [Cancer Res 2009;69(9):4059–66]

Introduction

Microarray technologies have been used to define the global gene expression patterns in human cancer and have successfully identified gene expression signatures or novel tumor classes as well as prognostic information (1–4). However, because the gene expression profile is only a snapshot of complex genetic interactions, it is difficult to discriminate the genes driving the neoplastic process (“driver” genes) from by-stander genes (“passenger” genes). Array-based comparative genomic hybridization analyses have also identified DNA copy number changes in many cancers. Some of these copy number–altered loci have been linked to disease pathogenesis and clinical behavior (5–8), but many of the copy number–altered genes do not seem to be expressed. Thus, it seems reasonable to suggest that the combined analysis of both data sets may reveal the causally altered genes in both copy numbers and corresponding gene expression (9–12). However, dauntingly complex genomic and epigenomic interactions in cancer and the current limitations in sensitivity and specificity of genomic data may impede the identification of the regulatory genes (13). It is therefore challenging to evaluate the biological function and clinical utility of the lengthy candidate genes obtained from genomic data.

In this study, we aimed to establish a gene selection strategy to identify potential driver genes by integrating genomic data of copy numbers and gene expression profiles. Guiding our gene selection process by estimating the prognostic significance of the selected genes, we identified 50 potential driver genes. Further computational and experimental interrogations suggested that the identified genes represent potential driver genes for hepatocellular carcinomas (HCC); providing both new insights into the pathobiology of liver cancer and therapeutic opportunities.

Materials and Methods

Comparative genomic hybridization profiling. Genomic DNA of 15 HCC samples and a normal reference liver were extracted using DNeasy extraction kit (Qiagen). After labeling tumors and a reference normal liver DNA samples with Cy5 and Cy3 fluorescent dye, respectively, hybridizations and data acquisition were performed following the instructions of the manufacturer (Nimblegen). The raw data were normalized as described previously (14).

Generation of the t statistic and the transcriptome correlation maps. The regional copy number alteration was estimated by a t statistic–based sliding window approach, the t statistic map (TM). Before constructing TM, considering the absolute copy number values exceeding 0.5 in at least three samples to be significant, 309,939 out of 3,080,000 probes were identified as having nontrivial changes. TM scores were calculated as the t statistic values which were obtained by applying a one-sample t test on the probe values for 15 samples within a moving window. We chose 100 Kb as the best-fit window size as it most accurately reflects the genomic copy number changes when various different window sizes (100 Kb, 1 Mb, and 10 Mb) were tested (data not shown). The TM scores generated by fewer than five probes in a moving window were regarded as missing values. The threshold for significant TM score was determined as the highest TM score from 100 random data sets which were generated by randomly ordering the probe positions. Significant TM regions were determined by a segmentation algorithm. The probes above the threshold were determined as valid probes. If a valid probe was not neighbor within 100 Kb of another valid probe, it was regarded invalid. Then, if the valid probes were neighbor over 5 Mb, the probe position between them was segmented. Segments less than 100 Kb were excluded.
Regional concordance of transcriptional regulation in 139 HCC expression profiles were constructed using the transcriptome correlation map (TCM), as described previously (13). Briefly, for each gene, the sum of pair-wise Pearson's correlation coefficients of the gene expression levels in neighboring \( n \) genes was calculated. Moving window size was determined to include the largest significant genes with the smallest neighboring genes (\( n = 24 \)). The threshold for significant TCM values and the segments were determined by the same algorithms used in the TM analysis.

**Estimation of the prognostic values of gene sets.** The prognostic value (\( P_0 \)) of a selected gene set was evaluated by using a clustering algorithm and log rank test as described previously (4). Then, a permutation test was performed to compare the prognostic value of a selected region with those from other regions. Two thousand random data sets were generated by randomly selecting the same number of genes from outside the selected regions (or genes) to be tested. For each random data set, hierarchical clustering was performed to stratify the HCCs into two groups, and the log rank test \( P \) value (\( P_m \)) between the groups was calculated. Iterative hierarchical clustering for the 2,000 random data sets was performed by a command-line version of Cluster 3.0 (15). The significance of the permutation test was taken as the frequency of events \( P_m < P_0 \) from the 2,000 trials. Prognostic impact score (PIS) was calculated as \(-\log_{10}(\text{permutation } P \text{ value})\). A permutation \( P < 0.05 \) (PIS > 1.3) was regarded as significant.

Small interference RNA–mediated gene knockdown. HepG2 and HuH-7 cells were transfected with siGENOME SMARTpool small interference RNA (siRNA; Dharmacon) for the selected target genes using LipofectAMINE 2000 (Invitrogen) according to the manufacturer’s protocol. Nontargeting siRNA (siGENOME Non-Targeting Pool no. 1 siRNA) was used as control. After 96 h of incubation with siRNAs, the cell viability was assessed by 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay.

**Screening of therapeutic drugs using in trans correlated gene signatures.** Raw data of Connectivity Map (build01) with two different Affymetrix platforms (i.e., hgu133a and HT-hgu133a) were downloaded from the authors’ web site and normalized independently using the RMA method implemented in BioConductor (16). For each of the in trans gene signatures, positive or negative signatures were assigned according to the algebraic sign of the correlation coefficient of each gene. For each treatment instance, the average connectivity score for the in trans gene signatures was calculated as the representative connectivity score (\( S_{\text{avg}} \)). The connectivity scores (\( S_i \)) for each in trans gene signature to drug treatment instance (\( i \)) were calculated based on Kolmogorov-Smirnov test and the significance of

---

**Figure 1.** Identification of genomic regional patterns of DNA copy numbers and mRNA expression in HCC. A, TM was constructed based on one-sample \( t \) statistic values of the copy numbers in 15 HCCs with a moving window size of 100 Kb. B, gene expression profiles of 139 HCC samples in the order of chromosomal location. C, TCM for the gene expression profiles of 139 HCC. Thresholds for TM (29.9) and TCM (28.7) were determined by 100 random permutation tests.

---

4 http://www.broad.mit.edu/cmap/
5 http://www.bioconductor.org/
the connectivity of a perturbagen was evaluated by a permutation test for 10,000 times as described previously (16). Additional procedures are described in Supplementary Materials and Methods.

Results

Comparison of gene copy number and transcription levels in HCC. We generated fine-scale whole genome copy number profiles of 15 HCC samples using comparative genomic hybridization microarrays containing 3,080,000 isothermal oligonucleotide probes. To represent the HCC heterogeneity, we selected five tumor samples from each of the subgroups of A, B, and HB, which were previously defined as having homogeneous expression patterns with prognostic differences (Supplementary Table S1; refs. 1, 4). First, we screened for the copy number–altered regions in HCC by applying the TM method. TM was optimized to identify chromosomal regions with low stringent but high regional concordance of copy number changes, which yielded 57 TM regions. These regions were highly concordant with previous findings including the recurrent gains at 1q, 3q, 6p, 7q, 8q, 17q, 20q, 22q, and the losses at 1p, 6q, 8p, 9p, 14q, 16q, and 17p (Fig. 1; Supplementary Table 2A; refs. 7, 10, 13, 17, 18).

To assess the chromosomal regional influence of the gene expression, we next rearranged the previous 139 HCC gene expression data (4) according to a sequence map of the human genome (hg17), which revealed a strong regional concordance of gene expression patterns across the heterogeneous tumor samples (Fig. 1B). This regional pattern can be shown more clearly with a TCM. TCM was previously used to identify concordant regional coexpression which was thought to be largely due to the copy number alteration (Fig. 1C; ref. 19). With a TCM, we identified 48 regions in which gene expression levels were likely to be copy number–dependent (Supplementary Table S2B). Remarkably high TCM scores were observed in chromosome 8, implying strong influence of the copy numbers on gene expression levels. This agrees well with the previous studies indicating marked copy number aberration of chromosome 8 and its regulatory role in cancer development and progression (20, 21). We then cross-compared the TM and TCM regions, which yielded 25 overlapped regions (CR) spanning 112.8 Mb and containing 580 genes.

![Figure 2](https://example.com/figure2.png)

**Figure 2.** Prognostic effects of TM or TCM regions. A, Kaplan-Meier plot analyses (left) and PIS (right) of the genes located in the TM, TCM, and CR regions, respectively. To show the significance of the PIS, the distribution of log rank test P values generated from 2,000 random data sets (Pm) were plotted with the log rank test P values calculated from the selected gene sets (P0, red line). B, Kaplan-Meier plot analyses and log rank tests for overall survival and recurrence-free survival (C) in the HCC subgroups based on expression similarities of the genes located in the CR1, CR3, CR13, and CR20.

www.aacrjournals.org 4061 Cancer Res 2009; 69: (9). May 1, 2009

Downloaded from cancerres.aacrjournals.org on April 12, 2017. © 2009 American Association for Cancer Research.
These regions may represent the most plausible regions in which copy numbers and expressions are concomitantly deregulated.

**Prognostic effect of the copy number–dependent regions.** We hypothesized that if genomic alterations of the genes in CR are critical events during HCC development, the gene expression patterns in CR would be well associated with patient prognosis. We evaluated the prognostic relevance of the selected regions by calculating PIS, which may represent the relative prognostic significance compared with the nonselected regions. The TM and TCM regions showed significant log rank test $P$ values, but their PIS were not significant compared with the other regions (Fig. 2A). However, the genes residing in the CRs showed a significant prognostic value ($PIS = 1.62$), which may suggest that the genes in CR play regulatory roles in the HCC progression (Fig. 2A). In particular, 4 out of the 25 CR segments (CR1, 1q21.3–23.2; CR3, 1q42.11–42.12; CR13, 7q36.1; and CR20, 8q24.11–24.22) were highly predictive for patient survival ($P < 0.001$; Fig. 2B). CR13 was significant in predicting tumor recurrence as well as survival (Fig. 2C). Deregulated genes in these regions (1q21, 1q42, 7q36, and 8q24) may therefore have a high probability of functioning as potential driver genes.

**Screening of the potential driver genes.** Next, we tested whether the previous correlation-based gene selection method (9, 11) can identify the potential driver genes which have prognostic relevance. Before gene selection, we evaluated the genomewide influence of all copy number changes on gene expression. The distribution correlation coefficients showed significant global influence of the copy number changes on the gene expression levels (Supplementary Fig. S1A). Moreover, the copy number profile of each sample was best correlated with its corresponding expression levels, demonstrating that the overall copy number changes substantially reflected the gene expression profiles in each individual genome (Supplementary Fig. S1B).

Confident that the copy number changes were strongly reflected in the corresponding gene expression levels, we selected 379 correlated copy number alteration (corCNA) genes based on the correlation between the expression levels and the copy numbers ($P < 0.01$, $r > 0.649$, Pearson's correlation test, false discovery...
rate = 0.0019). To validate the correlation by an independent method, we measured the copy numbers of two corCNA genes (POLR2K and C1orf43) in 52 HCC samples using quantitative PCR (Supplementary Fig. S1C and D). However, the prognostic effect of the 379 genes was not significantly different from the non-correlated genes (PIS = 0.76) even though the log rank test revealed a significant difference (Fig. 3A, top). We next selected the corCNA genes residing in CR, which yielded 50 genes. Strikingly, these 50 corCNA genes showed a strong prognostic effect (P = 5.06 × 10⁻⁶; PIS = 2.03; Fig. 3A, middle). This result suggests that our gene selection strategy using the regional pattern recognition approach can effectively select the functionally relevant genes. Thus, we reasoned that these 50 genes were the most probable candidate driver genes (Table 1).

Interestingly, of the 50 corCNA genes, 30 genes resided in 1q (n = 8) and 8q (n = 22), which were previously shown to have copy number gains at the early stage of HCC development (7). The prevalence of 1q and 8q genes among 50 corCNA genes is unlikely
Cancer Res 2009; 69: (9). May 1, 2009

Published OnlineFirst April 14, 2009; DOI: 10.1158/0008-5472.CAN-09-0164

Discussion

In this study, we conducted a combined analysis of copy numbers and gene expression with the following unbiased strategies. By applying regional pattern recognition approach, TM and TCM, we could identify the most probable copy number–dependent genes. In each step of our gene selection strategy, the functional relevance of the selected genes was evaluated by estimating the prognostic effect from the expression patterns and clinical information. Although our study was not intended to classify prognostic groups using the copy numbers or the gene expression patterns, we adopted clustering algorithm and log rank test to calculate PIS which allowed the evaluation of the functional relevance of our gene selection strategy. Our analysis revealed 50 potential driver genes which have significant prognostic relevance. Indeed, the 50 genes included many cancer-related genes. For example, RRM2B, HAX1, CUTF1, HSFI, EPHB4, and MTDH were known to regulate tumorigenesis, tumor migration, or adhesion. The genes related to RNA processing/transcription (e.g., ADAR,

to be due to the gene abundance in those regions \( P = 9.42 \times 10^{-5} \), hypergeometric test). These 30 genes in the “early event” regions (1q21.3–42.12 and 8q21.13–24.13) showed the most significant prognostic relevance \( P = 1.32 \times 10^{-7} \); PIS = 3.30; Fig. 3A, bottom), indicating the importance of the early dysregulation of these genes in HCC development. The prognostic values of the 50 or 30 1q/8q corCNA genes were further validated by two independent HCC data sets (SNU; ref. 22 and GSE6764; ref. 2) using class prediction algorithms (for details see Supplementary Methods). The tumor classes defined by the expression similarity of the 50 or 30 genes could successfully predict the prognostic outcome of HCC in the independent data sets (Supplementary Fig. S2; Table S3), supporting the robustness and consistency of our finding.

The enrichment of the 1q and 8q genes among the 50 corCNA genes raised the possibility that the gene expression alterations at 1q and 8q might concomitantly occur during the early stage of HCC development. We observed that the average copy numbers of 1q and 8q corCNA genes were correlated to each other \( (r = 0.564, P = 0.028; \text{Fig. 3B}) \). In parallel, the gene expression levels of 1q and 8q corCNA genes were also correlated in our 139 HCC data set \( (r = 0.333, P = 6.17 \times 10^{-5}, \text{permutation } P = 0.042; \text{Fig. 3C}) \), and this was validated in an independent 35 HCC data set \( (r = 0.398, P = 0.017; \text{Fig. 3D}) \). In addition, the result from GSE6764 also showed that the coexpression levels of the 1q and 8q genes corresponded with the pathologic staging of the liver (cirrhosis, dysplasia, and early HCC), and advanced HCC \( (n = 65, r = 0.609; P = 7.11 \times 10^{-8}, \text{permutation } P = 0.005; \text{Fig. 3D}) \). Permutation tests from 10,000 random trials indicated that our findings are not likely to be observed by chance. Taken together, the 30 corCNA genes might be coamplified and coexpressed during the early stage and play critical roles in HCC development and progression.

Functional assessment and identification of therapeutic targets for the 50 genes. Because the 50 genes were selected by unbiased screening, we next examined their potential driver roles. We randomly chose 11 genes from the 50 genes to include the copy number–gained genes in the distinct chromosomes (CCT3, SCRIB, PYCR2, HSFI, NCSTN, EPHB4, MTDH, CCNE1, RRM2B, TATDN1, and POLR2K) and tested the siRNA-mediated knockdown effects on the cancer cell viability. Of the 11 genes, 5 siRNAs for SCRIB, NCSTN, HSFI, TATDN1, and POLR2K significantly reduced the viability of both HepG2 and HuH-7 liver cancer cell lines \( (P < 0.01; \text{Fig. 4A}) \). In particular, the loss of NCSTN and SCRIB showed the most potent growth inhibition \( (P < 0.001) \) supporting the potential driver role of these genes in HCC development and making them attractive therapeutic targets. The knockdown of each five effective target genes was confirmed by quantitative reverse transcription-PCR, and its target selectivity was validated by two independent siRNAs (Supplementary Fig. S3). These results show that our screening strategy can successfully identify novel therapeutic targets for HCC. The 50 potential driver genes were identified by computational and experimental methods. However, the global influence of the 50 genes on cell function might be mediated by their target effector genes. We therefore sought to identify the putative effector genes whose expression levels were correlated in trans with expression of each driver gene \( (P < 0.001, \text{Pearson's correlation test}) \). Overall, the in trans correlated gene sets \( (15–1,606) \) were enriched with the cancer-dominant functions such as protein transport/translation, oxidative stress, RNA processing/DNA replication, and metabolism/immune-related functions; this may support their representative roles for the effector genes (Supplementary Fig. S4).

Considering the in trans correlated genes as putative effectors for the 50 genes, we hypothesized that the disruption of the expression of the in trans correlated genes would interrupt HCC development. To test this hypothesis, we used Connectivity Map, an analytic resource of gene expression profiles for drug responses \( (16) \). For each of the 50 in trans gene signatures, we calculated the connectivities to the drug instances in the Connectivity Map. The connectivity score profiles of the 50 in trans gene sets showed overall similar connectivities against each treatment instance, suggesting the commonality of their connectivities to the drug responses \( (\text{Fig. 4B}) \). Therefore, we used the averaged connectivity score of the 50 in trans gene sets \( (S_{\text{avg}}) \) as a representing connectivity to each treatment instance. We found only two perturbagens, metformin and sirolimus (rapamycin), to be significantly connected to the \( S_{\text{avg}} \) \( (\text{permutation } P = 0.0025 \text{ and } 0.0082, \text{respectively; Fig. 4B}) \). In addition, gefitinib, an epidermal growth factor receptor (EGFR) selective tyrosine kinase inhibitor, was also identified as having significant connectivity \( (S_{\text{avg}} = -0.335 \text{ and ranked to 7th out of 453 instances}) \). Because only one instance for gefitinib was available in Connectivity Map (build01), we performed permutation test by treating all the EGFR/IR selective tyrosine kinase inhibitors as same perturbagen \( (n = 7) \). This revealed a strong connectivity of the EGFR tyrosine kinase inhibitors to \( S_{\text{avg}} \) \( (\text{permutation } P = 0.0022; \text{Fig. 4B}) \). These results may suggest that the dysregulated potential driver genes can be interrupted by rapamycin, metformin, or gefitinib treatment.

Interestingly, the cross-talk among the target molecules of rapamycin, metformin, and gefitinib has been reported \( (\text{i.e., mTOR, AMPK, and EGFR; ref. 23}) \). This supports the idea that a combination of these drugs may provide clinical benefit for patients with HCC \( (24) \). We found that the combined treatment with these drugs could potentiate the growth inhibition of cancer cells as compared with the expected potency calculated by Bliss independence model \( (\text{for a review, see ref. 25; Fig. 4C}) \).

In addition, our data suggests a novel link between the 50 genes and the EGFR, AMPK, and mTOR signaling pathways. To verify this notion, we examined the knockdown effects of NCSTN and SCRIB on the mTOR signaling which is thought to be a common target for the three signaling pathways. Western blot analysis showed that the knockdown of NCSTN and SCRIB could indeed inhibit the phosphorylation of p70S6 kinase, one of mTOR's targets \( (\text{Fig. 4D}) \).
ELP3, HMBOX1, POLR2K, ZHX1, and KARS) and cell cycle regulation (e.g., BRD4, CCNC, and CCNE1) were also identified.

The functional significance of the 50 genes was further evaluated in part by siRNA-mediated knockdown experiments. Particularly, NCSTN, encoding a γ-secretase component, showed the most potent effect on cancer cell growth in our screening. Supporting our finding, NCSTN was previously identified as one of the copy number–dependent genes in HCC (11). Moreover, inhibition of γ-secretase has been reported to reduce leukemia cell growth via Notch and mTOR signaling pathways (26). SCRIB was also known to associate with human cancers (27), but its oncogenic role is not yet fully understood. Loss of HSF1 inhibited cancer cell growth (Fig. 4) in agreement with the recent identification of HSF1 as an oncogene (28). Taken together, our data successfully show the potential driver roles of the 50 genes, and suggest that our strategy can efficiently identify novel therapeutic targets.

The application of the in trans correlated genes to mine the Connectivity Map allowed us to identify therapeutic targets for the 50 genes. In the past, the Connectivity Map has been used to predict the therapeutic targets or diseases from gene expression signatures (16, 29, 30). In contrast with these studies, we used in trans correlated gene signatures to predict the molecular targets which were functionally associated with each of the 50 genes. In support of our approach, the coexpressed genes in microarray data

![Figure 4. Evaluation of functional and clinical utility of the 50 in trans correlated genes. A, the siRNAs (15 nmol/L) targeting 11 driver genes were transfected to HepG2 and HuH-7 cells for 96 h, and the cell viabilities were assessed by MTT assay. Nontargeting control siRNA (NT-CTL) was used as a control. Columns, mean percentage cell viability of three replicates compared with NT-CTL; bars, SD. Significance of growth inhibition compared with NT-CTL for each cell line was evaluated with a two-sample f test (*, P < 0.05; **, P < 0.01; and ***, P < 0.001). B, connectivity scores for each of the in trans correlated gene signatures to Connectivity Map are shown in a heatmap ordered by the average connectivity scores for individual instances (S_avg). Bar views with the instances of metformin (M, n = 5), rapamycin (R, n = 10), and EGFR/IR selective tyrosine kinase inhibitors [i.e., gefitinib, 4,5-dianilinophthalimide, butein, tyrophostin AG-1478, and HNMPA-(AM)3; G, n = 7]. Ranked distribution of S_avg for the interesting perturbagen in the ordered 453 instances (black line), and other instances with positive (green) and negative (red) connectivity scores. C, HuH-7 cells were treated with rapamycin (R, 10 nmol/L), metformin (M, 1 mmol/L), or gefitinib (G, 1 μmol/L) for 48 h in serum-free media, and the cell viability was assessed by MTT assay. The percentage of cell viability in three replicates compared with the control group (red columns) and the expected viability of combination treatment (blue columns); bars, SD. Significance of the growth inhibition effects compared with the control or expected value E were evaluated by two-sample t tests (*, P < 0.05; **, P < 0.01; and ***, P < 0.001). D, HuH-7 cells were transfected with siRNAs (15 nmol/L) targeting NCSTN or SCRIB for 48 h, and Western blot analysis was performed using antibodies for p70 S6 kinase (p70S6K), phospho-p70 S6 kinase (pp70S6K), and β-actin.](https://www.aacrjournals.org/doi/figure/10.1158/0008-5472.CAN-09-0164)
have been previously shown to be functionally related (31). Accordingly, our in silico approach could predict relevant effector targets without generating experimental signatures for each of the driver genes. Identification of common connectivities (S\text{avg}) from 50 different signatures might also be helpful in reducing chance findings which can be generated by applying a single signature. Indeed, our approach was able not only to predict well-known anticancer drugs with recognized therapeutic potential for HCC, such as rapamycin and gefitinib (8, 32–36), but also identified metformin as a therapeutic drug for targeting the 50 in trans genes signatures. Metformin is a widely used hypoglycemic drug for patients with diabetes. Recently, several studies have shown its anticancer effect, suggesting a novel utility of the antimetabolic drugs for cancer treatment (15, 37, 38). Supporting these studies, the oncogenic function of HSFI has been reported to be mediated by the antidiabetic drug metformin selective in cultured cells in vitro (Fig. 4C).

Our analysis also predicted novel links between the 50 genes and mTOR, AMPK, and EGFR pathways. Supporting this notion, the oncogenic function of HSFI has been reported to be mediated by the antidiabetic drug metformin selective in cultured cells in vitro (Fig. 4C).

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Acknowledgments
Received 1/15/09; revised 2/19/09; accepted 2/20/09; published OnlineFirst 4/14/09.

Grant support: Intramural Research Program of the Center for Cancer Research/National Cancer Institute, and in part by FG-4-2 of the 21C Frontier Functional Human Genome Project from the Ministry of Science & Technology in Korea.

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 the staff of the Laboratory of Experimental Carcinogenesis for critical reading of the manuscript.

References
39. O’Callaghan-Sunol C, Sherman MY. Heat shock proteins are possibly mediated through these pathways. Although further stringent validations for the mechanism(s) might be required, our findings open new opportunities for studying the multifaceted association of these 50 genes with the three identified pathways.

In conclusion, we suggest that our unbiased gene selection strategy of integrating multidimensional genomic data can effectively select the potential driver genes and provide new biological and clinical insights into the copy number–dependent genes in HCC. Undoubtedly, our strategy can be applied to other cancer types.

Cancer Res 2009;69 (9): May 1, 2009
4066 www.aacrjournals.org

Downloaded from cancerres.aacrjournals.org on April 12, 2017. © 2009 American Association for Cancer Research.
Identification of Potential Driver Genes in Human Liver Carcinoma by Genomewide Screening

Hyun Goo Woo, Eun Sung Park, Ju-Seog Lee, et al.