MicroRNA/gene profiling unveils early molecular changes and NRF2 activation in a rat model recapitulating human HCC.

MiRNA/gene profiling unveils early molecular changes and NRF2 activation in a rat model recapitulating human HCC

Annalisa Petrelli[1]*, Andrea Perra[2]*, Davide Cora'[1]*, Pia Sulas[2], Silvia Menegon[1], Claudia Manca[2], Cristina Migliore[1], Marta Anna Kowalik[2], Giovanna Maria Ledda-Columbano[2], Silvia Giordano[§1°], Amedeo Columbano[§2°].

[1]IRCC, Institute for Cancer Research and Treatment, University of Torino School of Medicine, 10060 Candiolo (Torino), Italy; [2]Department of Biomedical Sciences, Unit of Oncology and Molecular Pathology, University of Cagliari, Cagliari, Italy;

* § these Authors equally contributed to the work.

° to whom correspondence should be addressed.

Key words: miRNome; transcriptome; multistage liver carcinogenesis; Cytokeratin-19; NRF2

Correspondence:
Amedeo Columbano, PhD
Department of Biomedical Sciences
Unit of Oncology and Molecular Pathology
University of Cagliari
Via Porcell 4, 09124 Cagliari, Italy
Phone: +39-070-6758345
Fax: +39-070-666062
e-mail: columbano@unica.it

Silvia Giordano MD, PhD
Department of Oncology
University of Torino, Medical school
Institute for Cancer Research and Treatment (IRCC)

Strada Provinciale 142
Candiolo (Torino), 10060, Italy
Phone + 39 0119933233
Fax +39 011 9933225
e-mail silvia.giordano@ircc.it

Abbreviations: 2-AAF, 2-acetylaminofluorene; Ade, adenoma; aHCC, advanced HCC; BH, Benjamini-Hochber; DENA, diethylnitrosamine; eHCC, early HCC; GCLC, glutamate-cysteine ligase; GST, glutathione S-transferase; GSTP, placental glutathione S-transferase; IHC, immunohistochemistry; IL-1, interleukin-1; IPA, Ingenuity Pathway Analysis; KEAP1, kelch-like ECH-associated protein 1; KRT-19, cytokeratin-19; LPS, lipopolysaccharide; MAF, musculoaponeurotic fibrosarcoma oncogene homolog; miRNA, microRNA; NQO1, NAD(P)H dehydrogenase, quinone 1; NRF2, nuclear factor (erythroid-derived 2)-like 2; PPARa, peroxisome proliferator activated receptor alpha; qRT-PCR, quantitative reverse transcriptase polymerase chain reaction; R-H model, Resistant-Hepatocyte model; RXR, retinoic-X-receptor; T3, thyroid hormone.

Financial Support: this work was supported by Associazione Italiana Ricerca sul Cancro (AIRC, Grants IG-11821 to AC and IG-11819 to SG), Ministero Università e Ricerca Scientifica (PRIN 2009X23L78 to SG and PRIN 2010LC747T to AC), R.A.S. 2012 to AC. CM is an AIRC fellow. MAK is a FIRC fellow.

ABSTRACT

Studies on gene and/or microRNA (miRNA) dysregulation in the early stages of hepatocarcinogenesis are hampered by the difficulty of diagnosing early lesions in humans. Experimental models recapitulating human hepatocellular carcinoma (HCC) are then entailed to perform this analysis. We performed miRNA and gene expression profiling to characterize the molecular events involved in the multistep process of hepatocarcinogenesis in the Resistant-Hepatocyte rat model. A high percentage of dysregulated miRNAs/genes in HCC were similarly altered in early preneoplastic lesions positive for the stem/progenitor cell marker cytokeratin-19, indicating that several HCC-associated alterations occur from the very beginning of the carcinogenic process. Our analysis also identified miRNA/gene-target networks aberrantly activated at the initial stage of hepatocarcinogenesis. Activation of the NRF2 pathway and up-regulation of the miR-200 family were among the most prominent changes. The relevance of these alterations in the development of HCC was confirmed by the observation that NRF2 silencing impaired while miR-200a overexpression promoted HCC cell proliferation *in vitro.* Moreover, T3-induced *in vivo* inhibition of the NRF2 pathway accompanied the regression of cytokeratin-19 positive nodules, suggesting that activation of this transcription factor contributes to the onset and progression of preneoplastic lesions towards malignancy. The finding that 78% of genes and 57% of dysregulated miRNAs in rat HCC have been previously associated to human HCC as well underlines the translational value of our results.

Conclusions: this study indicates that most of the molecular changes found in HCC occur in the very early stages of hepatocarcinogenesis. Among these, the NRF2 pathway plays a relevant role and may represent a new therapeutic target.

Hepatocellular carcinoma is the third cause of cancer-related deaths worldwide and a major health problem. Liver cirrhosis is the underlying disease in more than 80% of cases and can be due to different etiologies such as hepatitis B and C, non-alcoholic and alcoholic fatty liver disease (1).

Although many factors, including chromosomal anomalies, genetic polymorphisms, genetic and epigenetic alterations, contribute to HCC onset and progression (2), the relevant molecular mechanisms remain largely unclear. Recently, genome-wide gene expression microarray and qRT-PCR studies have tried to identify the genes abnormally expressed in HCC and to generate molecular signatures able to classify different types of liver tumors and distinguish progressive from non-progressive lesions (3). These studies indicate a general aberrant activation of signalling pathways involved in cellular proliferation, survival, differentiation and angiogenesis, but a general consensus for a signature or even a single pathway prominent and characteristic of this cancer is still missing.

Recent studies have shown that classification of tumors can also be performed by evaluating the modulation of miRNAs, small non coding RNAs acting as regulators of gene expression (4). MiRNAs are dysregulated in several cancers where they can behave as oncogenes (5) or tumor suppressor genes (6), depending on the cellular function of their targets. As far as HCC is concerned, several miRNAs have been consistently found up- or down-regulated in different study cohorts (7,8). Moreover, a putative role of miRNAs in predicting tumor outcome has been assessed on the basis of their expression (9). However, although miRNA changes between tumor specimens and the normal corresponding tissues have been investigated in

HCC, no clear-cut conclusion has been drawn from most of the various microarray-based studies. It is now becoming increasingly evident that the molecular pathogenesis of HCC cannot be understood without more detailed studies of the molecular alterations characterizing its early development. Indeed, by investigating fully developed HCC, it is difficult to discriminate changes causing cancer progression from those being the consequence of cell transformation. Comprehensive studies on gene expression patterns and/or miRNA dysregulation in early stages of liver carcinogenesis are, so far, scanty and the clinical and diagnostic value of gene expression changes occurring in human HCC remains elusive. Thus, investigations on the precancerous lesions and early stage liver cancers are likely to reveal useful insights on the molecular mechanisms underlying the multistep process of hepatocarcinogenesis.

Since studies on initial HCC stages in humans are hampered by the clinical difficulty of diagnosing early lesions, experimental models allowing to dissect the several steps of HCC are mandatory. Among the animal models used to characterize the process of HCC development, the Resistant-Hepatocyte (R-H) model allows to dissect the different steps of the carcinogenic process, offering the possibility to identify phenotypically distinct lesions at well defined timings (10). Histologically, livers sequentially develop preneoplastic nodules (1-2 months after initiation with DENA) arising in a fibrotic environment, adenomas and early HCCs (eHCC, 6-10 months) and fully advanced HCCs (aHCC, 14 months). A translational value of this model has been recently shown, since the gene expression signature associated to the rat lesions (positive for the stem/progenitor cell marker KRT-19) can successfully predict

the clinical outcome of human HCC (11). The finding that the KRT-19+ HCC subtype is characterized by the worst clinical prognosis among all human HCC subclasses (12) suggests that KRT-19 is a potential prognostic marker for HCC.

The aim of the present study was to perform an integrative analysis of global miRNA and mRNA expression profiles in the R-H model of hepatocarcinogenesis for enhanced marker and therapeutic target discovery. Specifically, we aimed at i) identifying miRNAs/genes dysregulated during the carcinogenic cascade, mainly focusing on the less investigated early steps; ii) analyzing miRNA/mRNA correlations to unveil integrated networks that are altered at the beginning of the process and maintained along tumor progression; iii) validating the translational value of this rat model also for miRNA studies, by conducting comparative analyses between miRNAs and mRNAs dysregulated in rat preneoplastic and neoplastic lesions and those identified in human HCCs.

We demonstrate here that several deregulated miRNAs/genes in fully developed rat HCC, including many miRNAs/genes altered in human HCC, are already dysregulated in the very early step of tumorigenesis. Importantly, our findings unveil the activation of the NRF2 transcription factor pathway from the very beginning and throughout the process and they also reveal the existence of regulatory networks between miRNAs and their target genes. In particular, we found up-regulation of miR-200a that controls the NRF2 pathway. Finally, we show that an extremely high number of dysregulated miRNAs/genes in rat preneoplastic and neoplastic lesions are dysregulated in

primary human HCC as well, suggesting the potential utility of this model to investigate into the critical molecular changes underlying HCC development.

EXPERIMENTAL PROCEDURES

Animals and treatment

Guidelines for Care and Use of Laboratory Animals were followed during the investigation. All animal procedures were approved by the Ethical Commission of the University of Cagliari and the Italian Ministry of Health. Male Fischer F-344 rats (100-125g) were purchased from Charles River (Wilmington, MA, USA). Preneoplastic lesions and HCCs were induced as described in Supporting Material. Histologic classification of preneoplastic nodules, adenomas, eHCCs and aHCCs was performed as previously described (13).

MiRNA and mRNA expression profiling

RNA was extracted and purified from each individual lesion after laser microdissection from the liver of four to five animals (for microdissection procedures see Supporting Material). The miRNA expression profiling was performed using TaqMan Low Density Arrays starting from 600ng of RNA and following manufacturer instructions (Applied Biosystems, Carlsbad, CA, USA). For the gene expression profile, 150ng of RNA were amplified (Illumina TotalPrep RNA Amplification Kit), labeled and hybridized on Illumina microarray (RatRef-12 V1 BeadChips, Illumina Inc., San Diego, CA, USA), including 21.791 gene specific oligonucleotide probes (for further details and data analysis see Supporting Material). MiRNAs and mRNAs validation was performed using specific TaqMan assays (Applied Biosystems).

RESULTS

MiRNome classifies the different steps of hepatocarcinogenesis and identifies miRNA changes occurring in very early phases of the tumorigenic process.

To generate miRNA expression signatures specific for the different steps of hepatocarcinogenesis, we analyzed microdissected nodules (10 weeks after initiation with DENA), adenomas and eHCCs (10 months) and aHCCs (14 months). Immunohistochemical analysis showed that less than 25% of preneoplastic nodules, equally positive for the placental form of GSTP, were also positive for KRT-19 (Supp.Table1). Notably, almost all aHCCs examined at the end of the experiment showed positivity for KRT-19, supporting our preliminary findings that, in this model, KRT-19 positive preneoplastic lesions are the progenitors of HCC, while those negative for KRT-19 are more likely to spontaneously regress (11).

In microdissected lesions, 200 out of 375 analyzed miRNAs were detectable by TaqMan Low-Density Array technology and were further considered in the present study (see Supporting Material for selection criteria). Unsupervised hierarchical clustering, based on the relative expression levels of the different miRNAs, revealed the existence of two major clusters, separating early preneoplastic lesions from more advanced stages (Fig.1A). Within the two clusters, the miRNome was able to classify the different types of lesions; moreover, a more stringent selection of differentially expressed miRNAs allowed a nearly complete separation also between concomitant adenomas and early carcinomas (Supp.Fig.1). Quantitative RT-PCR validation performed

on 10 randomly selected miRNAs in individual lesions confirmed the TaqMan array results in 80% of cases (Supp.Fig.2).

In order to identify miRNAs differentially expressed at each stage compared to the matched normal controls, we applied the Limma analysis package (14) (P<0.05; BH-corrected) (Fig.1B). Comparing each step with the previous one, we found both miRNAs dysregulated in specific transitions (Fig.1C and Supp.Table2) and miRNAs commonly altered in consecutive steps (Fig.1C and Supp.Table3A,B,C). Interestingly, 13 miRNAs already modified in KRT-19+ early lesions (e.g. miR-224, miR-122 and miR-375) were altered throughout the entire process (Fig.1C and Supp.Table3D), suggesting their essential role in cancer development. Further support for the relevance of these miRNAs in liver carcinogenesis comes from the observation that 7 out of 13 miRNAs have been found similarly dysregulated in human HCC as well (15).

MiRNA expression analysis between the KRT-19- and KRT-19+ preneoplastic lesions revealed that no miRNA reached the BH correction for multiple testing at the selected threshold of P<0.05; however, using a one-sample t-test, we found 23 differentially expressed miRNAs (P<0.05; Supp.Table4), that are likely important in the progression of KRT-19+ lesions towards malignancy.

***mRNA profiling reveals that most of the HCC associated molecular changes already occur in preneoplastic lesions*.**

Gene expression profiling was performed in the same lesions using the Illumina microarray. A total of 1.144 out of 21.791 genes included in the array were selected as described in Supporting Material. Hierarchical cluster

analysis stratified the rat lesions into two major clusters: 1) normal liver and preneoplastic KRT-19- lesions; 2) preneoplastic KRT-19+ lesions, adenomas/eHCCs and aHCCs, forming 3 distinct subclusters (Fig.2A). A major difference between transcriptome and miRNome was the ability of the latter to clearly separate preneoplastic from neoplastic lesions, still maintaining the difference between KRT-19- and KRT-19+ nodules. Quantitative RT-PCR validation performed on randomly selected genes confirmed the microarray expression data for all the examined genes (Supp.Fig.3).

To identify the differentially expressed genes in each type of lesion towards its age-matched control we applied the Limma analysis package. As shown in the Venn diagram (Fig.2B, left), although KRT-19- and KRT-19+ lesions are histologically very similar, they exhibited a strikingly different number of modified genes. Besides the 64 dysregulated genes shared between KRT-19- and KRT-19+ likely involved in nodule formation, 602 genes were exclusively altered in KRT-19+ preneoplastic lesions, suggesting that they are relevant for nodule progression. Interestingly, 216 out of 234 altered genes in aHCC were altered in KRT-19+ nodules as well (Fig.2B, right). Among these, 33/39 of the most up-regulated (fold change vs. controls >5) and 12/15 of the most down-regulated (fold change vs. controls <-5) genes in aHCC were the most dysregulated in KRT-19+ nodules as well (Table1). These results suggest that the major expression changes leading to HCC occur in the very first stages of tumor progression.

Ingenuity Pathway Analysis unveils NRF2 activation in preneoplastic and neoplastic lesions.

Ingenuity Pathway Analysis (IPA) of genes altered at the final stage of the carcinogenic process (aHCC) revealed that most of the dysregulated genes are involved in metabolic pathways. Among these, there are NRF2-mediated oxidative stress response, LPS/IL-1-mediated inhibition of RXR function, aryl hydrocarbon receptor signalling and xenobiotic metabolism (Fig.3A). Strikingly, most of the altered pathways in aHCC were already modified in KRT-19+ preneoplastic lesions. Functional investigation also underlined common pathway modifications between early and late stages of hepatocarcinogenesis (Fig.3B).

Analysis of transcription factor-dependent pathways differentially activated in early lesions or in aHCC identified MYC and JUN/FOS (Fig.3C and Supp.Fig.4), already described to be involved in HCC (16). Remarkably, among the most dysregulated, there was the NRF2-dependent pathway (Fig.3C), whose role in cancer development has recently become a topic of an important controversy (17). The finding that the majority of NRF2 target genes were up-regulated indicates activation of the NRF2 pathway (Fig.3D). Notably, several of these genes were among the most up-regulated in both KRT-19+ lesions and aHCCs (see Table1). Moreover, the small MAF family of transcription factors, which operate as co-activators of NRF2 (18), was found activated at all stages of HCC development (Fig.3C and Supp.Fig.4).

Integrated miRNA:mRNA analysis uncovers differentially expressed miRNA/gene modules in early preneoplastic lesions.

To identify a possible correlation between miRNAs and genes dysregulated at the initial stage of the process, we selected conserved putative miRNA targets in rat, predicted by the TargetScanS algorithm. For each of the differentially expressed genes in KRT-19+ nodules, we extracted the annotated regulating miRNAs and performed an intersection with the differentially expressed miRNAs in the same stage. We found that a consistent percentage of modified genes are targets of dysregulated miRNAs (Fig.4). To study the relative expression of miRNAs and of their target genes, we evaluated the number of positive and negative correlations between the predicted miRNA-target gene pairs, by means of the fold-change value reported by the Limma package. Results show that 171 out of 215 miRNA/mRNA pairs (79%) shown in Fig.4 displayed anti-correlated expression, while 44 were positively correlated.

In the transition from normal liver to KRT-19+ nodules we observed at least 2 relevant nodes, where two miRNA families, namely miR-30 and miR-200, control the expression of many modified genes. While expression of the miR-30 family was mainly modified at this initial stage, miR-200 family was up-regulated in both preneoplastic lesions and aHCCs (Supp.Fig.5). Notably, one of the target genes of miR-200a is KEAP1, a negative regulator of NRF2 (19).

The NRF2 pathway promotes HCC cell growth and is controlled by miR-200a.

NRF2 is an integrated redox sensitive signaling system that regulates 1-10% of human genes and is negatively controlled by KEAP1, which promotes NRF2 proteasome-mediated degradation (19). As mentioned above, the

NRF2 pathway was already activated in early preneoplastic lesions and along tumor progression (Fig.3C,D). Accordingly, IHC performed on lesions at different stages revealed that all of them were strongly positive for NRF2 (Fig.5A); most important, IHC showed NRF2 nuclear staining in several preneoplastic and neoplastic hepatocytes, clearly demonstrating translocation and suggesting activation of this transcription factor. Activation of the NRF2-dependent pathway was confirmed by the increased expression of three examined NRF2 target genes (20) (NQO1, GSTA4 and GCLC), at all stages of hepatocarcinogenesis (Fig.5B). In parallel with NRF2 nuclear translocation, a decrease of KEAP1 was observed (Fig.5A).

To mechanistically investigate the role of this pathway, we moved to *in vitro* experiments and evaluated cell growth upon NRF2 or KEAP1 silencing in three human HCC cell lines. KEAP1 silencing was associated with increased NRF2 protein levels (Fig.6C and Supp.Fig.6C). When cell growth was assessed in the absence or in the presence of oxidative stress (H_2O_2), we found that NFR2 silencing impaired cell proliferation, while KEAP1 silencing increased growth rate (Fig.6A and Supp.Fig.6A). Analysis of NRF2 target gene expression demonstrated either inhibition or activation of the pathway upon NRF2 or KEAP1 silencing, respectively (Fig.6B and Supp.Fig.6B).

Since in humans KEAP1 is negatively controlled by miR-200a (21), which is up-regulated in KRT-19+ lesions and aHCC compared to normal liver (Supp.Table2 and Supp.Fig.5), we evaluated cell growth upon transfection of a miR-200a mimic. MiR-200a promoted cell growth, recapitulating the effect of KEAP1 silencing (Fig.6A and Supp.Fig.6A). The effect of miR-200a modulation was confirmed by its ability to down-regulate KEAP1 both in

human and rat HCC cells (Fig.6C and Supp.Fig.6C) and to promote the expression of NRF2 target genes (Fig.6B and Supp.Fig.6B). Altogether, these results suggest that miR-200a controls the NRF2 pathway, whose activation promotes liver cancer cell growth.

Inhibition of NRF2 is associated with regression of KRT-19 positive lesions.

Since the role of NRF2 in early and/or intermediate stages of HCC development is still unknown, we aimed at investigating the effect of NRF2 modulation in these stages, characterized by a remarkable NRF2 activation (Fig.3). Our previous studies have shown that a 7 days treatment with thyroid hormone (T3) is able to cause a significant reduction in the number of preneoplastic hepatic lesions in rats previously exposed to the R-H model (22). Based on the above findings, we wished to determine whether the anti-tumorigenic activity of T3 could be mediated by NRF2. To this aim, 9 weeks after DENA initiation, nodule-bearing rats were fed T3 for 4 and 7 days. While T3 treatment for 7 days caused a 50% reduction in the number of preneoplastic lesions compared to untreated animals, no difference between the two groups was observed at day 4 (Fig.7A). We evaluated the expression of NRF2 and KEAP1 in microdissected KRT-19+ lesions from 4 day-treated or untreated rats, namely at a time that preceded the loss of preneoplastic nodules. T3 treatment resulted in global reduction of NRF2 expression and in loss of its nuclear localization (Fig.7B), suggesting that T3 inactivates the NRF2-dependent pathway. This was confirmed by the down-regulation of NRF2 target genes (Fig.7C). In parallel with NRF2 reduction, we observed an

increase of KEAP1, which returned to the levels of the surrounding tissue. (Fig.7B). *In vitro* experiments confirmed the inhibitory effect of T3 on the NRF2-dependent pathway (Supp.Fig.7).

The expression signature of HCC developed by the R-H model recapitulates that of human HCC.

To assess the translational value of our results, we determined how many differentially expressed genes and miRNAs in rat HCC are dysregulated also in human HCC. After removal of features not annotated on standard rat gene symbols from the gene array list, we extracted 159 genes from the original list of 234 dysregulated ones in rat HCCs. According to the Ensembl Database (V66; http://www.ncbi.nlm.nih.gov/pubmed/22086963), 110 genes out of 159 were rat-human orthologous (see Supp.Fig.8 for experimental scheme). By integrating our data with the datasets annotated in Liverome, a curated database of liver cancer related gene signatures (23), we found that 78% (86/110) of the orthologous genes were dysregulated in HCC of both species (Supp.Table 5). Among these, 49 genes were dysregulated in a similar fashion in at least 50% of the works included in Liverome; moreover, 18 genes had the same type of dysregulation in at least 1 included work. Overall, the concordance between dysregulated genes in rat and human scored 78%, supporting the translational value of this model.

To determine which genes were already dysregulated at the beginning of the murine carcinogenic process, we intersected the 86 altered genes in HCCs in both species with those modified in KRT-19+ early nodules. Remarkably, we

found that 65 genes out of 86 were dysregulated in rat preneoplastic lesions as well (bold in Supp.Table 5).

Literature-based analysis showed that 26 out of 46 miRNAs were commonly altered in rat and human HCC (Supp.Table 6) and that 10 out of these 26 miRNAs were already dysregulated in KRT-19 positive early lesions (bold in Supp.Table 6).

Altogether, these results clearly support the potential utility of the R-H model for detecting molecular alterations occurring in human HCC and identifying those arising at the onset of the process and likely critical for cancer development.

DISCUSSION

Hepatocarcinogenesis is a slow process whereby the accumulation of genetic and epigenetic alterations eventually leads to the emergence and the expansion of clonal populations of transformed hepatocytes that evolve toward HCC. However, in humans, the precise sequence of molecular events involved in tumor initiation and progression is not well defined, due to the limited access to early stages of tumor development. Thus, most of the studies have focused on fully developed HCCs and, as a consequence, information about the molecular alterations of early preneoplastic lesions is scanty. Since the molecular pathogenesis of HCC cannot be fully understood without more comprehensive knowledge of the molecular changes occurring during its early development, the histological precursors of human HCC - foci of phenotypically altered hepatocytes and dysplastic hepatocytes - deserve particular attention (24). In the current work, we have analyzed stepwise global miRNA and mRNA expression in a rat model of HCC, with emphasis on the very early steps characterized by the occurrence of phenotypically-altered preneoplastic nodules.

One of the most relevant findings stemming from our work is that a number of miRNAs are already dysregulated in KRT-19+ preneoplastic nodules. Since these lesions are considered the HCC precursors in the carcinogenesis model used in the present study (11), it is likely that these miRNAs play a relevant role in HCC onset. The identification of miRNAs altered at the beginning of the carcinogenic process is a novel finding, since very few contributions have attempted to address the impact of miRNA dysregulation at this stage of HCC

development. Indeed, previous studies aimed at identifying miRNA alterations at the beginning of hepatocarcinogenesis have evaluated miRNA expression only in the whole liver of mice exposed to a carcinogenic regimen - characterized by hepatic fat accumulation and inflammatory response (the choline-devoid methionine deficient model) - before the appearance of preneoplastic lesions, rather than in isolated nodules (25,26).

Among the miRNAs found dysregulated in our study, some have been reported as modified in human HCC, while others have not been previously associated with liver cancer. Although further studies are warranted to better define the role of these miRNAs and of their targets, they might represent novel critical players in the development and progression of HCC. In particular, the present study identified 13 miRNAs that are dysregulated from the very early stages of the carcinogenic process throughout the progression to HCC, suggesting that they participate in the initial events leading to HCC development and that are required for neoplastic progression. Among these miRNAs, miR-224, miR-125b, miR-375 and miR-122 had already been identified as dysregulated in human HCC (7,9,27,28), whereas others, such as miR-802, miR-429 and miR-499 have not been previously described.

A second important finding is that 85% of the most up-regulated and 80% of the most down-regulated genes in rat HCC were already altered in early KRT-19+ preneoplastic nodules. Remarkably, an impressive number of genes involved in xenobiotic metabolism and NRF2-mediated oxidative stress signaling pathway were modified from the beginning of the tumorigenic progress. This is very relevant, as it suggests that metabolic changes are likely necessary, although not sufficient, to allow the upsurge of preneoplastic

lesions and to sustain the progression of early lesions to a malignant condition. This metabolic readjustment might be the consequence of a coordinated survival response to the DENA/2-AAF induced-damage. However, since a similar metabolic shift has been observed also in a transgenic mouse model in the absence of chemical treatment (29) and in human HCC (30), a response to oxidative stress appears to be a general feature caused by hostile conditions, such as viruses, alcohol or chemicals. Overall, these data suggest that the antioxidant response is an initial adaptive mechanism triggered to assure a better survival to liver cells, becoming irreversibly altered in some of them (initiated cells?). The finding that in the R-H model this metabolic derangement persists after the first 4 weeks of treatment, when the chemical-induced stress is no longer present, and the observation that dysregulation of the same genes is observed in advanced HCC indicate that these initial changes may likely play a far more critical role than previously recognized.

Another interesting observation is that the comparison between dysregulated genes and miRNAs in each step of progression showed the existence of networks where a consistent percentage of modified genes are indeed targeted by dysregulated miRNAs. In the transition from normal liver to KRT-19+ nodules we identified two relevant nodes where miR-30 and miR-200 families control the expression of many modified genes. While miR-30a, -30d and -30e were modified mainly at this initial stage, miR-200a, -200b and -429 were up-regulated at both initial and late stages. It is worth to note that miR-200a is known to negatively regulate the NRF2 pathway, which is already activated in early preneoplastic KRT-19+ lesions. Following exposure of cells

to electrophiles or oxidative stress, NRF2 is able to escape KEAP1-mediated degradation, translocate to the nucleus and activate the expression of a series of antioxidative and cytoprotective proteins including NQO1, GCLC and several members of the GST family (31). Notably, the role of this transcription factor has recently become the topic of an important controversy, since it is unclear whether NRF2 acts as a tumour suppressor or as an oncogene(17). Indeed, while many studies suggest that NRF2 activation mediates the beneficial effects of chemopreventive drugs, genetic analyses of human tumours indicate that this transcription factor may exert an oncogenic effect and cause resistance to chemotherapy (32). Moreover, a recent work has identified activating mutations of this gene in 6.4% of HCCs, further sustaining its oncogenic role (33). However, as recently reviewed (17), NRF2 role in early/intermediate steps of the tumorigenic process is largely unknown. Our study provides evidence of an oncogenic role of NRF2 in preneoplastic/pre-malignant stages of hepatocarcinogenesis. Indeed: i) many NRF2 target genes were among the most up-regulated, ii) many members of the small MAF family of NRF2 coactivators were found activated, iii) NRF2 silencing impaired liver cancer cell proliferation *in vitro* and iv) *in vivo* treatment of rats with T3 caused the inhibition of the NRF2 pathway followed by the regression of KRT-19+ preneoplastic lesion.

Finally, the present study shows the common dysregulation of many miRNAs and genes in both rat and human HCC. Such a striking similarity in miRNA/gene expression perturbation among HCC from different species and etiologies demonstrates that the R-H model recapitulates human HCC to an extent that allows successful mining of target genes. In this line, our results

indicating an oncogenic role of the NRF2 pathway in preneoplastic lesions should be considered for their possible translational value. Further investigation is warranted to verify NRF2 activation in human preneoplastic stages as well. If so, targeting this pathway would offer new therapeutic options in stages of progression that could dramatically change the evolution of the disease.

Acknowledgment: we thank G. Diaz for support in statistical analysis, M. Angioni and A. Follenzi for generation and characterization of RH cells, B. Martinoglio for qRT-PCR and F. Natale for editing the manuscript.

Reference List

1. Bosch FX, Ribes J, Cleries R, Diaz M. Epidemiology of hepatocellular carcinoma. Clin Liver Dis 2005;9:191-211, v.

2. Hoshida Y, Toffanin S, Lachenmayer A, Villanueva A, Minguez B, Llovet JM. Molecular classification and novel targets in hepatocellular carcinoma: recent advancements. Semin Liver Dis 2010;30:35-51.

3. Lee JS, Thorgeirsson SS. Genome-scale profiling of gene expression in hepatocellular carcinoma: classification, survival prediction, and identification of therapeutic targets. Gastroenterology 2004;127:S51-S55.

4. Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer 2006;6:857-866.

5. He L, Thomson JM, Hemann MT, Hernando-Monge E, Mu D, Goodson S, et al. A microRNA polycistron as a potential human oncogene. Nature 2005;435:828-833.

6. Johnson SM, Grosshans H, Shingara J, Byrom M, Jarvis R, Cheng A, et al. RAS is regulated by the let-7 microRNA family. Cell 2005;120:635-647.

7. Giordano S, Columbano A. MicroRNAs: New tools for diagnosis, prognosis, and therapy in hepatocellular carcinoma? Hepatology 2012.

8. Imbeaud S, Ladeiro Y, Zucman-Rossi J. Identification of novel oncogenes and tumor suppressors in hepatocellular carcinoma. Semin Liver Dis 2010;30:75-86.

9. Jiang J, Gusev Y, Aderca I, Mettler TA, Nagorney DM, Brackett DJ, et al. Association of MicroRNA expression in hepatocellular carcinomas with hepatitis infection, cirrhosis, and patient survival. Clin Cancer Res 2008;14:419-427.

10. Solt DB, Medline A, Farber E. Rapid emergence of carcinogen-induced hyperplastic lesions in a new model for the sequential analysis of liver carcinogenesis. Am J Pathol 1977;88:595-618.

11. Andersen JB, Loi R, Perra A, Factor VM, Ledda-Columbano GM, Columbano A, et al. Progenitor-derived hepatocellular carcinoma model in the rat. Hepatology 2010;51:1401-1409.

12. Lee JS, Heo J, Libbrecht L, Chu IS, Kaposi-Novak P, Calvisi DF, et al. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. Nat Med 2006;12:410-416.

13. Petrelli A, Perra A, Schernhuber K, Cargnelutti M, Salvi A, Migliore C, et al. Sequential analysis of multistage hepatocarcinogenesis reveals that miR-100 and PLK1 dysregulation is an early event maintained along tumor progression. Oncogene 2012.

14. Smyth GK. Limma: linear models for microarray data. In: R.Gentleman VCSDRIWH, ed. Bioinformatics and Computational Biology Solutions using R and Bioconductor.New York: Springer, 2005. 397-420.

15. Borel F, Konstantinova P, Jansen PL. Diagnostic and therapeutic potential of miRNA signatures in patients with hepatocellular carcinoma. J Hepatol 2012;56:1371-1383.

16. Malz M, Pinna F, Schirmacher P, Breuhahn K. Transcriptional regulators in hepatocarcinogenesis--key integrators of malignant transformation. J Hepatol 2012;57:186-195.

17. Sporn MB, Liby KT. NRF2 and cancer: the good, the bad and the importance of context. Nat Rev Cancer 2012;12:564-571.

18. Marini MG, Chan K, Casula L, Kan YW, Cao A, Moi P. hMAF, a small human transcription factor that heterodimerizes specifically with Nrf1 and Nrf2. J Biol Chem 1997;272:16490-16497.

19. Kobayashi A, Kang MI, Okawa H, Ohtsuji M, Zenke Y, Chiba T, et al. Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate proteasomal degradation of Nrf2. Mol Cell Biol 2004;24:7130-7139.

20. Yeager RL, Reisman SA, Aleksunes LM, Klaassen CD. Introducing the "TCDD-inducible AhR-Nrf2 gene battery". Toxicol Sci 2009;111:238-246.

21. Eades G, Yang M, Yao Y, Zhang Y, Zhou Q. miR-200a regulates Nrf2 activation by targeting Keap1 mRNA in breast cancer cells. J Biol Chem 2011;286:40725-40733.

22. Ledda-Columbano GM, Perra A, Loi R, Shinozuka H, Columbano A. Cell proliferation induced by triiodothyronine in rat liver is associated with nodule regression and reduction of hepatocellular carcinomas. Cancer Res 2000;60:603-609.

23. Lee L, Wang K, Li G, Xie Z, Wang Y, Xu J, et al. Liverome: a curated database of liver cancer-related gene signatures with self-contained context information. BMC Genomics 2011;12 Suppl 3:S3.

24. Thorgeirsson SS, Grisham JW. Molecular pathogenesis of human hepatocellular carcinoma. Nat Genet 2002;31:339-346.

25. Wang B, Majumder S, Nuovo G, Kutay H, Volinia S, Patel T, et al. Role of microRNA-155 at early stages of hepatocarcinogenesis induced by choline-deficient and amino acid-defined diet in C57BL/6 mice. Hepatology 2009;50:1152-1161.

26. Wang B, Hsu SH, Majumder S, Kutay H, Huang W, Jacob ST, et al. TGFbeta-mediated upregulation of hepatic miR-181b promotes hepatocarcinogenesis by targeting TIMP3. Oncogene 2010;29:1787-1797.

27. Ladeiro Y, Couchy G, Balabaud C, Bioulac-Sage P, Pelletier L, Rebouissou S, et al. MicroRNA profiling in hepatocellular tumors is associated with clinical features and oncogene/tumor suppressor gene mutations. Hepatology 2008;47:1955-1963.

28. Budhu A, Jia HL, Forgues M, Liu CG, Goldstein D, Lam A, et al. Identification of metastasis-related microRNAs in hepatocellular carcinoma. Hepatology 2008;47:897-907.

29. Coulouarn C, Factor VM, Conner EA, Thorgeirsson SS. Genomic modeling of tumor onset and progression in a mouse model of aggressive human liver cancer. Carcinogenesis 2011;32:1434-1440.

30. Inami Y, Waguri S, Sakamoto A, Kouno T, Nakada K, Hino O, et al. Persistent activation of Nrf2 through p62 in hepatocellular carcinoma cells. J Cell Biol 2011;193:275-284.

31. Baird L, nkova-Kostova AT. The cytoprotective role of the Keap1-Nrf2 pathway. Arch Toxicol 2011;85:241-272.

32. Wang XJ, Sun Z, Villeneuve NF, Zhang S, Zhao F, Li Y, et al. Nrf2 enhances resistance of cancer cells to chemotherapeutic drugs, the dark side of Nrf2. Carcinogenesis 2008;29:1235-1243.

33. Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad IB, et al. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nat Genet 2012;44:694-698.

FIGURE LEGENDS

Figure 1. ***MiRNA profiling classifies the different steps of hepatocarcinogenesis.*** A) Hierarchical clustering of miRNA expression profile in KRT-19- and KRT-19+ preneoplastic lesions, adenomas (Ade), early HCCs (eHCCs) and advanced HCCs (aHCCs) in rats exposed to the R-H model. Each row represents the expression profile of a miRNA expressed as $\Delta\Delta$ct calculated respect to age-matched controls (see Supporting Material for details) and each column represents a sample. Different types of lesions are indicated by colored bars. Red and green colors represent higher or lower expression levels of the miRNA (median centered), respectively. B) Number of differentially expressed miRNAs in each step compared to age-matched controls, defined by $P<0.05$ (BH-corrected). C) Venn diagram illustrating the overlapping subsets of significantly dysregulated miRNAs in HCC progression.

Figure 2. ***Gene profiling classifies the different steps of hepatocarcinogenesis.*** A) Hierarchical clustering of 1.144 genes in normal liver (ctr), KRT-19- and KRT-19+ preneoplastic lesions, adenomas (Ade), early HCCs (eHCCs) and advanced HCCs (aHCCs) in rats exposed to the R-H model. Each row represents the expression profile of a gene and each column represents a sample. Controls and different types of lesions are indicated by colored bars. Red and green colors represent higher or lower expression levels of the mRNA (median-centered), respectively. B) Venn diagrams illustrating overlapping subsets of significantly dysregulated genes

in KRT-19- and KRT-19+ preneoplastic nodules (*left*) and in HCC progression *(right*). The differentially expressed genes between each sample and its age-matched control were defined by $P<0.01$ (BH-corrected) and fold-change (>2 or <-2).

Figure 3. ***Functional analysis of differentially expressed genes in the progressive steps of hepatocarcinogenesis.*** Top 5 enriched pathways (A) and functions (B) in early KRT-19- and KRT-19+ preneoplastic lesions and fully developed HCC. P values were determined using the Ingenuity scoring system and were judged significant at $P<0.05$ (vertical black line). C) Differentially activated transcription factor-dependent pathways (according to IPA) in the lesions indicated in A and B. TF: transcription factor. D) Network visualization of the transcription factor NRF2 pathway. TF → target gene. Red: up-regulated genes; green: down-regulated genes. Additional networks and symbol legends are shown in Supporting Fig. 4.

Figure 4. ***MiRNA:mRNA integrated network in the early step of hepatocarcinogenesis*.** Modules represent miRNAs (orange diamonds) and genes (light blue squares). Edges connect miRNAs and putative target genes predicted by TargetScanS, both differentially expressed in KRT-19+ nodules vs. controls. Module size reflects the degree of connectivity between miRNAs and target genes in the network.

Figure 5*. NRF2 and KEAP1 expression in the different stages of hepatocarcinogenesis. A)* NRF2 (left panels) and KEAP1 (right panels)

expression analysis performed by immunohistochemistry in preneoplastic lesions, eHCC and aHCC. The insets show staining of nodules and normal surrounding tissue at lower magnification. Each panel is representative of 5 samples for each stage of the carcinogenic process (magnification X20; inset magnification X4). B) QRT-PCR analysis of NRF2 target gene expression in control livers, KRT-19+ preneoplastic lesions, eHCC and aHCC. Values are shown as fold change difference compared to normal liver (control). The histogram represents mean values ± SE of 4 to 6 animals per group *P<0.05; **P<0.01.

Figure 6. ***Modulation of the NRF2 pathway affects HCC cell growth.*** A) Malhavu and Huh7 human HCC cells were transiently transfected with NRF2, or KEAP1 or control (siC) siRNAs or miR-200a mimic and seeded in 96-well plates. Cell growth was evaluated by crystal violet staining 72 hours after seeding, in either the presence (HuH7) or the absence (Mahlavu and HuH7 NT) of 25uM hydrogen peroxide (H_2O_2). B) Expression of NRF2 target genes NQO1, GCLC and GSTA4 was evaluated by qRT-PCR on RNA extracted from the cells transfected in (A). Values are shown as fold change difference compared to control transfected cells (siC). C) Cells were transiently transfected with the indicated siRNAs or miR-200a mimic. Total cell lysates were extracted 48 hours upon transfection and analyzed by Western blot. Bactin was used as loading control. Quantification of KEAP1 related to the corresponding βactin is shown for miR-200a (control siC was set as 1).

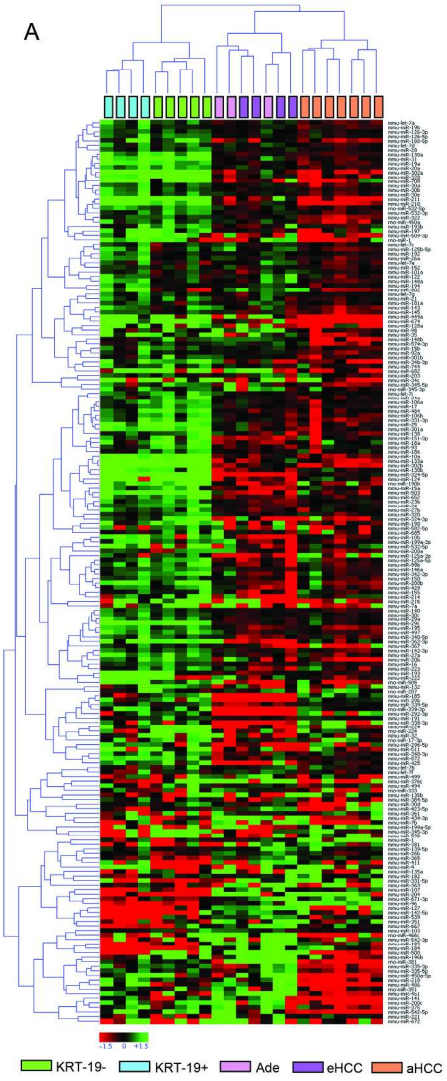
Figure 7*. In vivo inhibition of the NRF2 pathway is followed by regression of KRT-19+ preneoplastic lesions.* A) Number of hepatic GSTP-positive preneoplastic lesions after 4 or 7 days of T3 treatment in rats previously exposed to the R-H model. The histogram represents mean values ± SE of 4 to 6 animals per group. NT: not treated. NS: not significant. ****P*<0.001. B) Immunohistochemical staining of NRF2 and KEAP1 in preneoplastic lesions of untreated (NT) or 4 days T3-treated animals. The insets show staining of nodules and normal surrounding tissue at lower magnification. Each panel is representative of 5 samples for each stage of the carcinogenic process (magnification X20; inset magnification X4). C) QRT-PCR analysis of NRF2 target gene expression in normal livers (control), KRT-19+ preneoplastic lesions either not treated (NT) or after 4 days of T3 treatment. Values are shown as fold change difference compared to normal liver. **P*<0.05.

Table1. Most dysregulated genes in aHCC

Gene	FC aHCC/ctr	Gene	FC aHCC/ctr
most up-regulated genes		Ccnb2	6.39
Akr1b8	69.80	**Npdc1**	6.34
Defb1	59.73	**Bzrp**	6.12
Gstp1	51.91	**Pir**	5.93
Gstp2	46.60	***Gclc***	5.91
Gpx2	30.62	**Anxa1**	5.76
Ca2	28.32	**Ica1**	5.74
Yc2	25.61	Esm1	5.54
Nqo1	20.54	Phgdh	5.26
Ddit4l	19.51	Traf4af1	5.16
Akr7a3	18.40	***Txnrd1***	5.07
Krt1-19	17.09	**Serpinb1a**	5.00
Pcp4	13.79	**most down-regulated genes**	
Cryab	12.34	**Cyp2c**	-69.57
Abcc3	11.53	**Obp3**	-49.07
Fabp4	11.25	**Ca3**	-29.79
Slc25a4	10.78	**Dhrs7**	-22.80
Anxa2	10.05	**Hao2**	-10.86
Tubb6	9.25	**Cdh17**	-10.80
Lbp	9.08	**Olr59**	-9.86
S100a11	9.07	**Avpr1a**	-8.75
Abcb1	7.75	Sult1c1	-8.55
Scd2	7.54	**Aox3**	-8.47
Ugt1a6	7.12	**Cyp3a3**	-7.03
Ltb4dh	7.09	**Ust5r**	-6.56
Cdc2a	6.96	**Cyp2c37**	-5.81
Cd24	6.77	Akr1c18	-5.50
Rhbg	6.42	Gnmt	-5.42

Genes in bold are dysregulated in KRT-19+ preneoplastic lesions as well. NRF2 target genes are in Italic.

A



B

N° of miRNAs differentially expressed in each step compared to age-matched controls

	KRT-19-	KRT-19+	Ade/eHCC	aHCC
Total	27	41	45	46
Up	7	15	26	37
Down	20	26	19	9

C

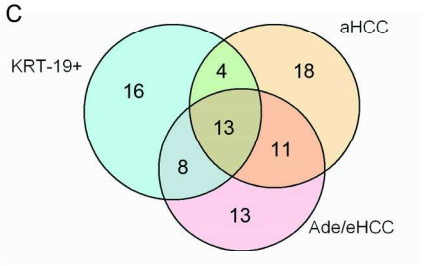
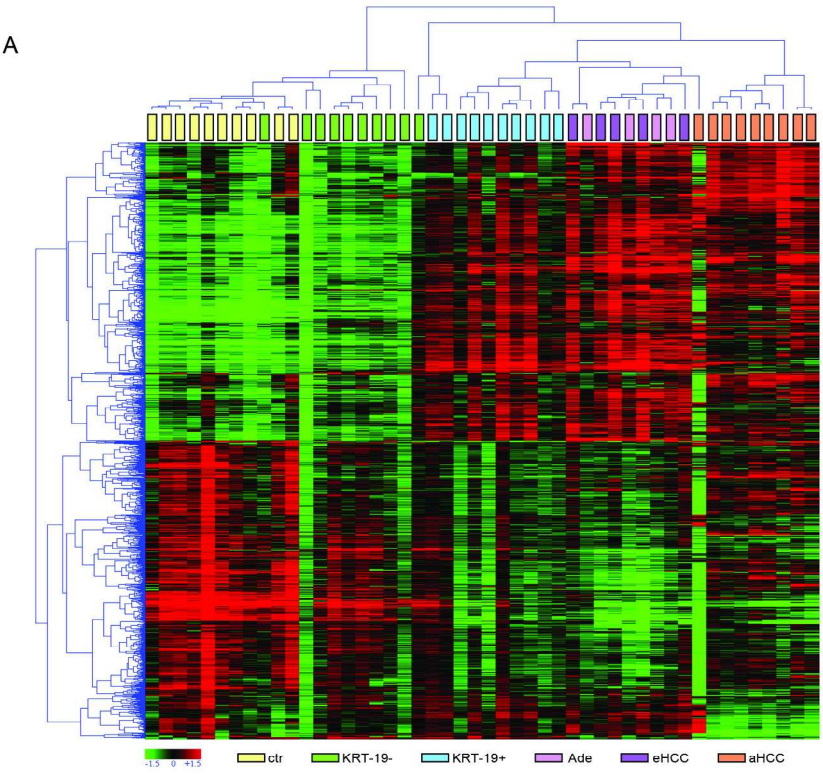


Fig. 1

190x254mm (300 x 300 DPI)



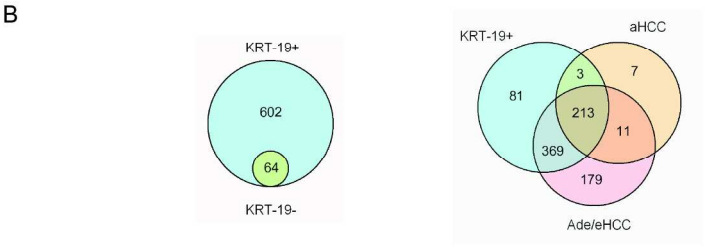
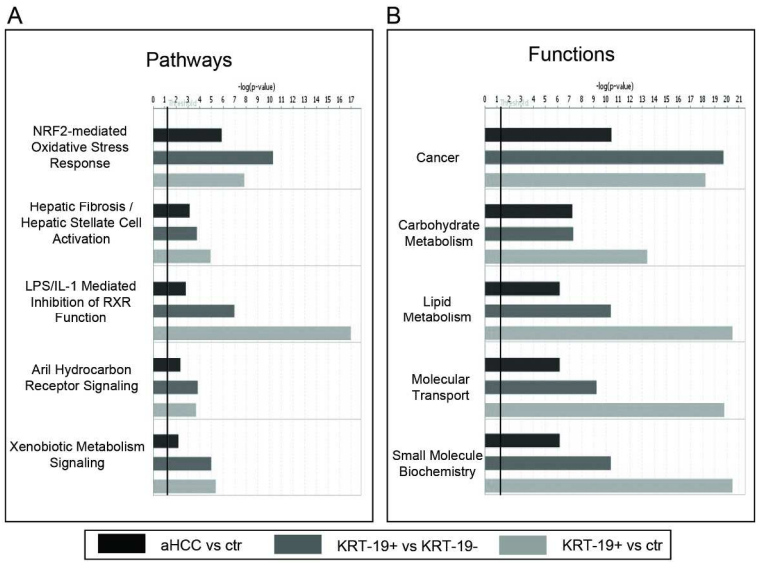


Fig. 2

190x254mm (300 x 300 DPI)



C

Transcription factor-dependent pathways differentially activated in the various steps

Step	TF	*P* value
KRT-19+ vs ctr	PPARA	8,1E-24
	RXRA	3,7E-12
	NRF2	3,3E-10
	MAFK	3,3E-06
	MAFG	1,4E-05
	MAFF	2,5E-05
KRT-19+ vs KRT-19-	PPARA	1,2E-15
	MYC	7,7E-13
	NRF2	1,2E-12
	JUN	2,9E-10
	MAFK	2,8E-08
	MAFG	1,6E-07
	FOS	3,0E-07
	MAFF	2,0E-04
aHCC vs ctr	MAFG	5,2E-10
	MAFK	1,2E-09
	NRF2	2,7E-09
	MAFF	5,0E-09
	PPARA	7,3E-09
	MYC	4,8E-08
	PPARG	9,8E-08
	FOS	1,0E-07

D

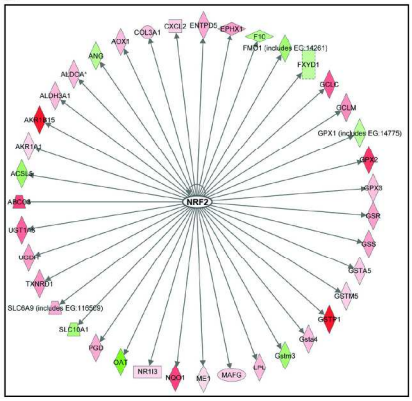


Fig. 3

190x254mm (300 x 300 DPI)

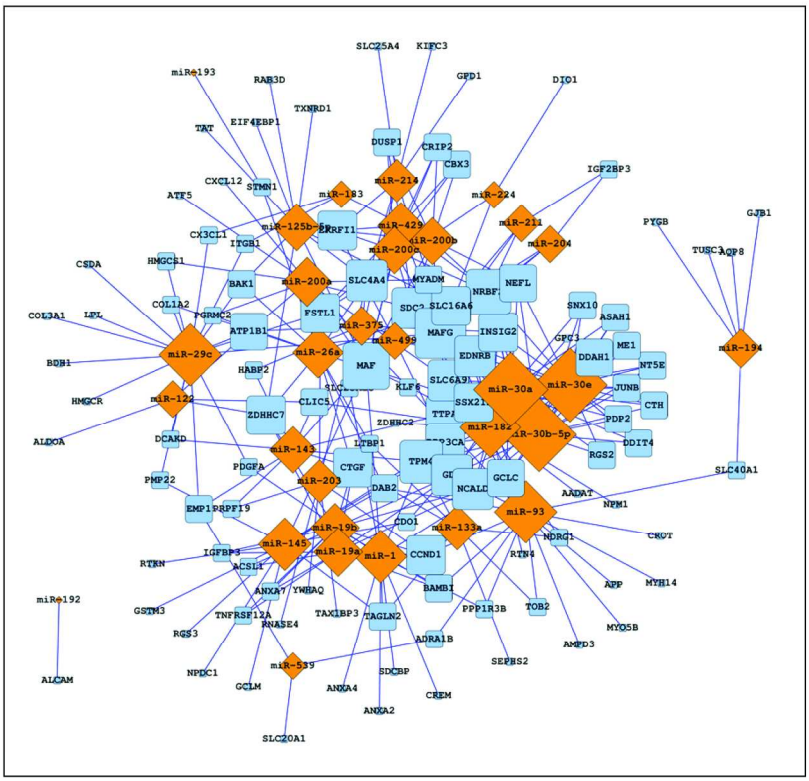


Fig. 4

190x254mm (300 x 300 DPI)

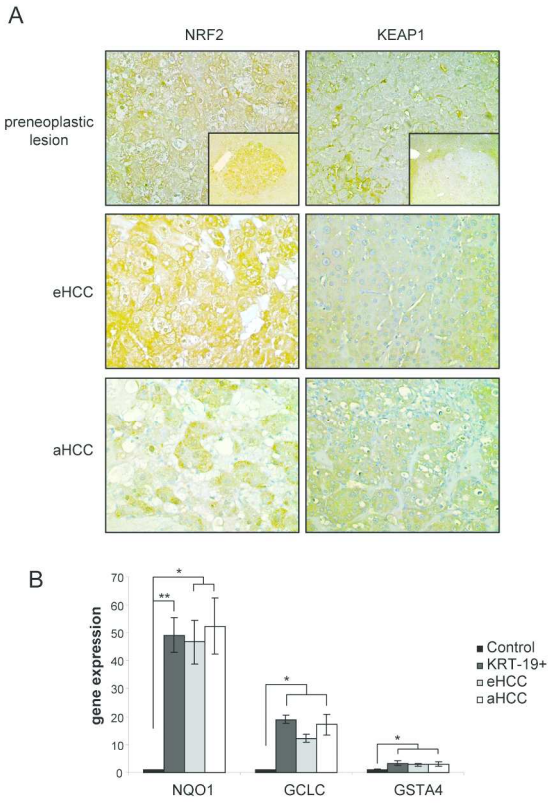


Fig. 5

190x254mm (300 x 300 DPI)

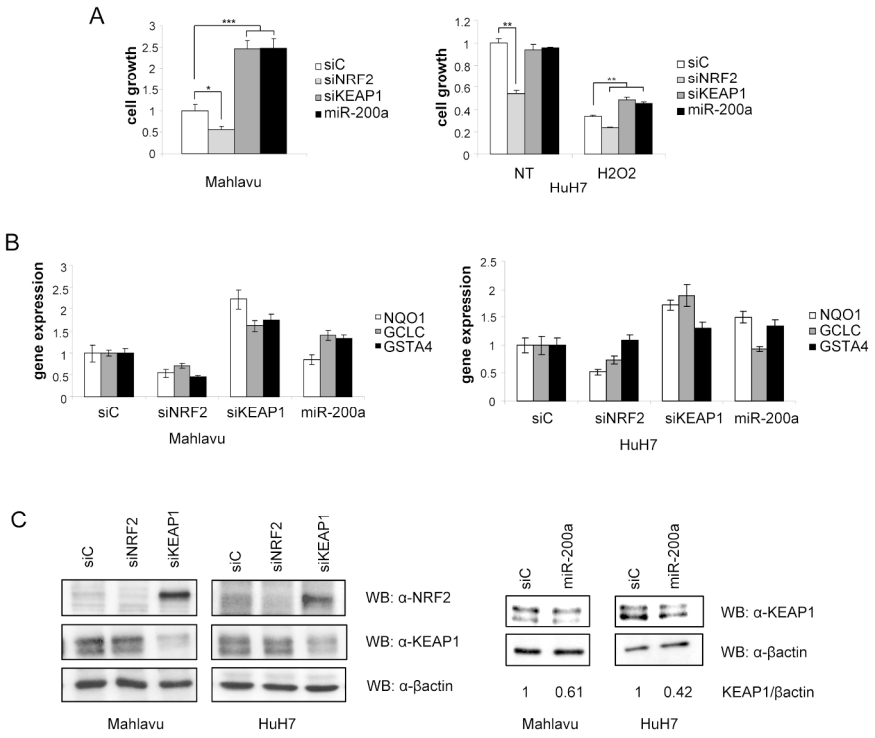


Fig. 6

193x254mm (300 x 300 DPI)

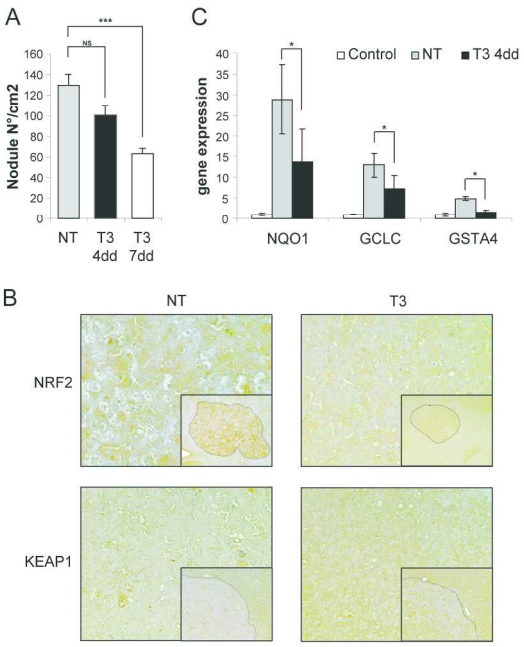


Fig. 7

190x254mm (300 x 300 DPI)

SUPPORTING MATERIAL

Animal experiments

Male F-344 rats were injected intraperitoneally with diethylnitrosamine (DENA, Sigma, St Louis, MO, USA) at a dose of 150 mg/kg body weight. After a 2-week recovery period, rats were fed a diet containing 0.02% 2-acetylaminofluorene (2-AAF, Sigma) for 1 week, followed by a two-thirds partial hepatectomy and by an additional week of 2-AAF diet. The animals were then returned to the basal diet and euthanized at 10 weeks, at 10 months and at 14 months. The number of GSTP-positive preneoplastic foci was determined as previously described (1). To investigate the effect of T3 (Sigma) on NRF2-target genes, F-344 rats were exposed to the DENA + 2-AAF protocol. Five weeks after 2-AAF release, they were divided into two groups; the first was maintained on normal diet, while the second was fed a T3-supplemented diet (4 mg/kg) for 4 and 7 days.

Immunohistochemistry

Frozen liver sections were collected and fixed in -20°C acetone for 20 minutes at room temperature. Following endogenous peroxidases block (Dako Envision, Glostrup, Denmark), aspecific sites were blocked by incubating sections for 1 hour at room temperature in PBS. Sections were incubated for 30 minutes with anti-GSTP antibody (MBL, Nagoya, Japan, cat# 311, 1:1000) and detected by anti-rabbit, HRP and 3-3’ diaminobenzidine (Dako Envision). Sections were counterstained with Harris hematoxylin. For KRT-19 staining, serial sections were treated as described (2). Anti-KRT-19 antibody

(Novocastra, Wetzlar, Germany, cat# NCL-CK19) was applied at 1:100 dilution. For NRF2 and KEAP1 staining, rabbit polyclonal antibodies (SantaCruz Biotechnology, Dallas, Texas, USA, sc-722 and sc-33569, respectively) were used at a 1:80 and 1:100 dilution.

Laser-capture Micro-dissection (LMD)

We microdissected 20 nodules (10 weeks after initiation with DENA), 4 adenomas (10 months), 5 eHCCs (10 months) and 9 aHCCs (14 months). Sixteen-µm-thick serial frozen sections of rat livers were attached to 2-µm RNase free PEN-membrane slides (Leica, Wetzlar, Germany). Microdissection (Leica, LMD6000) was followed by a 2.45 minutes H&E staining. RNA was extracted from micro-dissected samples using the PicoPure RNA Isolation Kit (Arcturus, Mountain View, CA, USA) or the mirVana miRNA Isolation Kit (Ambion, Austin, TX, USA), according to manufacturer instructions.

Supplementary bioinformatic methods

Microarray data analysis: the intensity files were loaded into the Illumina BeadStudio 3.0.19.0 software (Illumina Inc, San Diego, CA, USA) and BRB Array Tools (Version 4.2.0) for quality control and gene expression analysis. First, the quantile normalization algorithm was applied on the dataset. Only genes whose expression differed by at least 1.5 fold from the median in at least 20% of the arrays and characterized by a 50th percentile of intensities greater than 300 were retained. The FDR-adjusted p-values were calculated using the Benjamini-Hochberg procedure (3). According to these criteria,

1.144 expressed transcripts out of 21.791 showed reproducible up- or down-regulation. Custom R scripts based on the Limma package (4) were used to identify genes differentially expressed at $P<0.01$ between early nodules and advanced lesions, based on Log2ratio expression data (median centered). Normal livers were used as reference. Following this analysis, 869 genes showed reproducible up- or down-regulation in at least one comparison. Log2ratio expression data were clustered and visualized in the heat-maps using the GEDAS software (5), using Pearson correlation as distance matrix and Complete Linkage as method of calculating distance between clusters.

TLDA data analysis: for miRNA analysis, expression data for each array were first normalized to the internal control Mamm-U6 RNA (ΔCT). Subsequently, the ΔΔCT method (6) was used to evaluate the relative expression level of miRNAs in each sample vs. the respective age-matched control. As for gene expression data, R scripts based on the Limma package (4) were used to identify miRNAs differentially expressed at $P<0.05$ between the different lesions. The GEDAS software (5) was used to represent the data in heatmaps.

To investigate the statistical relevance of our results, we performed a Fisher test for each of the functional categories (controls, KRT-19-, KRT-19+ preneoplastic nodules, adenomas, eHCC and aHCC), evaluating the enrichment of the functional categories into the group of items under the same node tree. Analysis was performed using the R package for statistical computing.

The comparison between human and rat deregulated genes has been done using the annotation provided by the Liverome database. We used the

supplementary file published in the Liverome database, “combined_gene_list_file.xls”, in which, according to Liverome's authors, a complete list of all the genes that have at least one published evidence of differential expression in their database is reported. In particular, for each gene and for each experiment under consideration, Liverome already defined -where available- a value for the differential expression in a uniform representation (see Liverome's Materials and Methods for further details). We concentrated our efforts on those genes characterized by an “Up” or “Down” regulation flag in the Liverome “combined_gene_list_file.xls” file. Then we performed a comparison between the Fold-change obtained by the Limma package in our dataset and the Up/Down regulation status reported by Liverome (according to the correspondence for gene ids).

Functional analysis by means of the Ingenuity IPA Software: rat standard gene symbols (RGD ids) were submitted to the Ingenuity IPA analysis pipeline and converted to human gene ID, where possible. Analysis of the pathways was based on the number of genes significantly dysregulated (fold difference cutoff $\pm$ 2.0) with corresponding biological functions, with the restriction of at least 8 genes per function to emphasize the functions with most genes differentially expressed. The significance of each network and the connectivity was estimated in IPA. The “Transcription Factors regulators” option was used to identify Transcription Factors whose targets are enriched in our datasets. Only standard Rat gene IDs (RGD) were retained and the Ensembl Database (www.ensembl.org) was interrogated to establish human-rat orthologue genes. A number of 1000 permutations was used to assess the phenotypes

enrichments and only results corresponding to a FDR q-value<0.01 were retained for the discussion.

Integration of miRNA:mRNA data: we downloaded from the TargetScanS (7) database (http://www.targetscan.org/) predicted miRNAs targets for Rat, resulting in a miRNA-mediated post-transcriptional network composed by 210 rat unique miRNA IDs and 9.060 rat unique gene symbols. Using these data, for each of the differentially expressed genes in the various defined sets (KRT-19+ nodules, adenoma/eHCC and aHCC), we extracted the regulating miRNAs. Then, we compared the corresponding mature miRNA IDs to the ones of the differentially expressed miRNAs in the same experimental condition. In this way, we defined sets of miRNA:mRNA pairs in which both the elements are differentially expressed in the same experimental comparison and in which both the elements are annotated as interactors by TargetScanS. The resulting set of miRNA:mRNA pairs was used to define the edges of a miRNA:mRNA network, depicted by means of the Cytoscape software (8). For the miRNA-gene network, the number of positive and negative correlations between the pairs of predicted miRNA-target gene interactions was evaluated by means of the Fold-Change value reported by the Limma Package.

Cell culture and in vitro experiments

HepG2, HuH7 (ATCC, Manassas, VA, USA), Mahlavu (kindly provided by Dr. N. Atabey) human cells and FaO rat cells (Sigma) were cultured in DMEM complete medium with 10% fetal bovine serum (Lonza, Basel, Switzerland) in a 5% CO_2 atmosphere. The RH (Resistant Hepatocyte) HCC rat cells were

obtained by classical collagenase perfusion techniques from rats exposed to the R-H protocol and sacrificed at 14 months. At the time of sacrifice, the liver of the rat presented several macroscopic lesions identified as advanced HCC at histological level. Cells were cultured as human HCC cell lines in collagen-coated plates. Cells scored positive for KRT-19 and albumin and were able to grow for more than 50 passages in colture.

Cells were transiently transfected with 200 pmol of miRNA mimic (Ambion) and 100 pmol of siRNAs (Ambion) using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA). The day after, 3000 cells/well were seeded in 96well plates and treated with 25uM H_2O_2. Cell growth was evaluated after 96 hours as described in (9). For *in vitro* T3 experiments, cells were treated with 10^{-6} M triiodothyronine for 48 hours. To evaluate T3 receptor activation, the expression of G6PC and CPT1A genes was evaluated.

Protein extraction and western blot

For protein analysis, cells were lysed in 2% SDS, 0.5µM Tris–HCl, 48hours after transfection. Western blots were performed according to standard methods. The following antibodies were used: anti-NRF2 and anti-βactin (Santa Cruz Biotechnology), anti-KEAP1 (Proteintech, Chicago, USA) and anti-vinculin (Sigma). Final detection was done with the ECL system (Promega, Madison, Wisconsin, USA). Quantification of KEAP1 shown in Fig. 6 and Supp. Fig. 6 was performed using the Quantity One software and calculated as ratio between the mean value of intensity of KEAP1 and of the corresponding βactin bands.

NRF2 target gene expression analysis

RNA was extracted from transfected HepG2 cells using the Tri-Reagent solution (Ambion) and from microdissected samples, as mentioned above, according to manufacturer instructions. RNA was retro-transcribed starting from 0.5µg RNA/sample using the High Capacity Kit (Applied Biosystems, Carlsbad, CA, USA). Analysis of NQO1, GCLC and GSTA4 expression was performed using specific TaqMan probes (Applied Biosystems) and GAPDH as endogenous control.

Reference List

1. Ledda-Columbano GM, Perra A, Loi R, Shinozuka H, Columbano A. Cell proliferation induced by triiodothyronine in rat liver is associated with nodule regression and reduction of hepatocellular carcinomas. Cancer Res 2000;60:603-609.

2. Andersen JB, Loi R, Perra A, Factor VM, Ledda-Columbano GM, Columbano A, et al. Progenitor-derived hepatocellular carcinoma model in the rat. Hepatology 2010;51:1401-1409.

3. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological) 1995;57:289-300.

4. Smyth GK. Limma: linear models for microarray data. In: R.Gentleman VCSDRIWH, ed. Bioinformatics and Computational Biology Solutions using R and Bioconductor.New York: Springer, 2005. 397-420.

5. Fu L, Medico E. FLAME, a novel fuzzy clustering method for the analysis of DNA microarray data. BMC Bioinformatics 2007;8:3.

6. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001;25:402-408.

7. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 2005;120:15-20.

8. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8:new features for data integration and network visualization. Bioinformatics 2011;27:431-432.

9. Migliore C, Petrelli A, Ghiso E et al. MicroRNAs impair MET-mediated invasive growth. Cancer Res 2008;68:10128-10136

SUPPORTING FIGURE LEGENDS

Figure S1. ***MiRNA profile classifies the different steps of hepatocarcinogenesis.*** Hierarchical clustering of miRNA expression profile in KRT-19- and KRT-19+ preneoplastic lesions, adenomas (Ade), early HCCs (eHCCs) and advanced HCCs (aHCCs). Only miRNAs whose expression was dysregulated more than 4-fold in a minimum of two samples in at least one experimental group were considered. Each row represents the expression profile of a miRNA expressed as $\Delta\Delta$ct calculated respect to age-matched controls (see Supporting Material for details) and each column represents a sample. Red and green colors represent higher or lower expression levels of the miRNA (median centered), respectively.

Figure S2. ***QRT-PCR validation of randomly selected miRNAs in the stepwise development of HCC.*** MiRNA expression was determined in KRT-19- and KRT-19+ nodules, adenomas, eHCCs and aHCCs. MiRNA expression is reported as log fold-change relative to age-matched controls. TLDA: TaqMan Low Density Array.

Figure S3. ***QRT-PCR validation of randomly selected genes in the stepwise development of HCC.*** Dynamic mRNA expression was determined in KRT-19- and KRT-19+ nodules, adenomas, eHCCs and aHCCs. Gene expression is reported as log fold-change relative to age-matched controls.

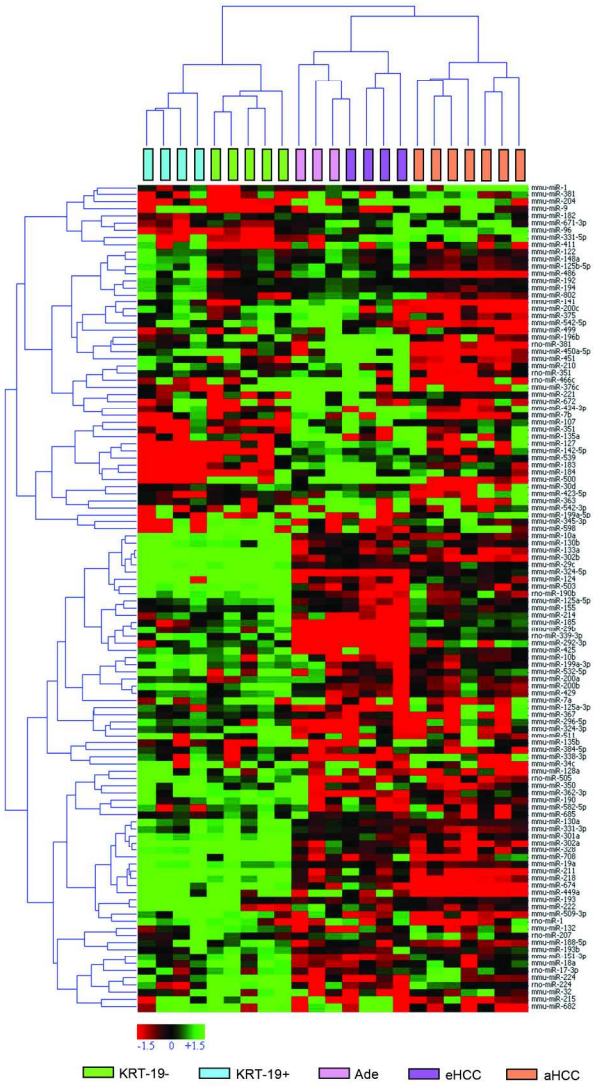
Figure S4. ***Network visualization of the indicated transcription factors*****.** Data and visualization were obtained using IPA. Red: up-regulated genes; green: downregulated genes. *TF → target gene*.

Figure S5. ***The miR-200 and miR-30 families are dysregulated during hepatocarcinogenesis.*** MiR-200 (A) and miR-30 (B) family expression evaluated by TLDA in different stages. MiRNA expression is reported as log fold-change compared to age-matched controls. TLDA: TaqMan Low Density Array.

Figure S6. ***Modulation of the NRF2 pathway affects HCC cell growth.*** A) HepG2 cells were transiently transfected with NRF2 or KEAP1 or control (siC) siRNAs or a miR-200a mimic and seeded in 96-well plates. Cell growth was evaluated by crystal violet staining 72 hours after seeding, in either the presence (H_2O_2) or the absence (NT) of 25uM hydrogen peroxide. B) Expression of the NRF2 target genes NQO1, GCLC and GSTA4 was evaluated by qRT-PCR on RNA extracted from the cells transfected in (A). Values are shown as fold change difference compared to control transfected cells (siC). C) HepG2 and rat cells (FaO and RH) were transiently transfected with the indicated siRNAs or miR-200a. Total cell lysates were extracted 48 hours upon transfection and analyzed by Western blot. Vinculin or βactin were used as loading control. Quantification of KEAP1 related to the corresponding βactin is shown for miR-200a (control siC was set as 1).

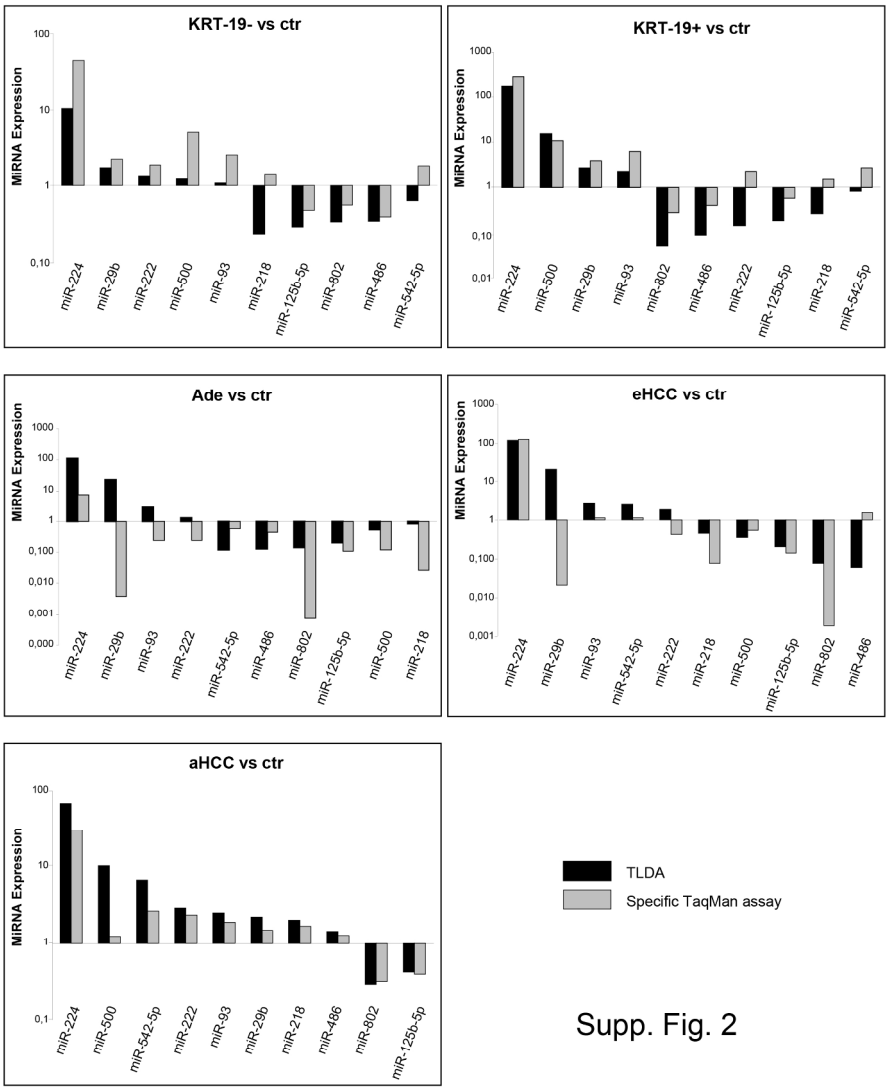
Figure S7. ***In vitro T3 treatment of HepG2 cells inhibits the NRF2 pathway.*** HepG2 cells were treated with 10^{-6} M triiodothyronine for 48 hours. Total RNA was then extracted and expression of NRF2 target genes (NQO1, GCLC, GSTA4) was evaluated. Data are indicated as fold changes compared to not treated cells (NT).

Figure S8. ***Flowchart for the intersection with the Liverome database.*** We first downloaded gene signatures from the Liverome database, for a total of 6.927 human genes. Then we focused on the dataset of 234 rat genes dysregulated in our model (aHCC vs ctr). For these genes, we looked for reliable human orthologous by means of the Ensembl database (version 66). Commonly dysregulated genes were then identified according to matching gene ids. Up/Down common modulations were assessed by means of the Fold-Change provided by the Limma package for the rat case and by the internal Liverome annotation for the human case.



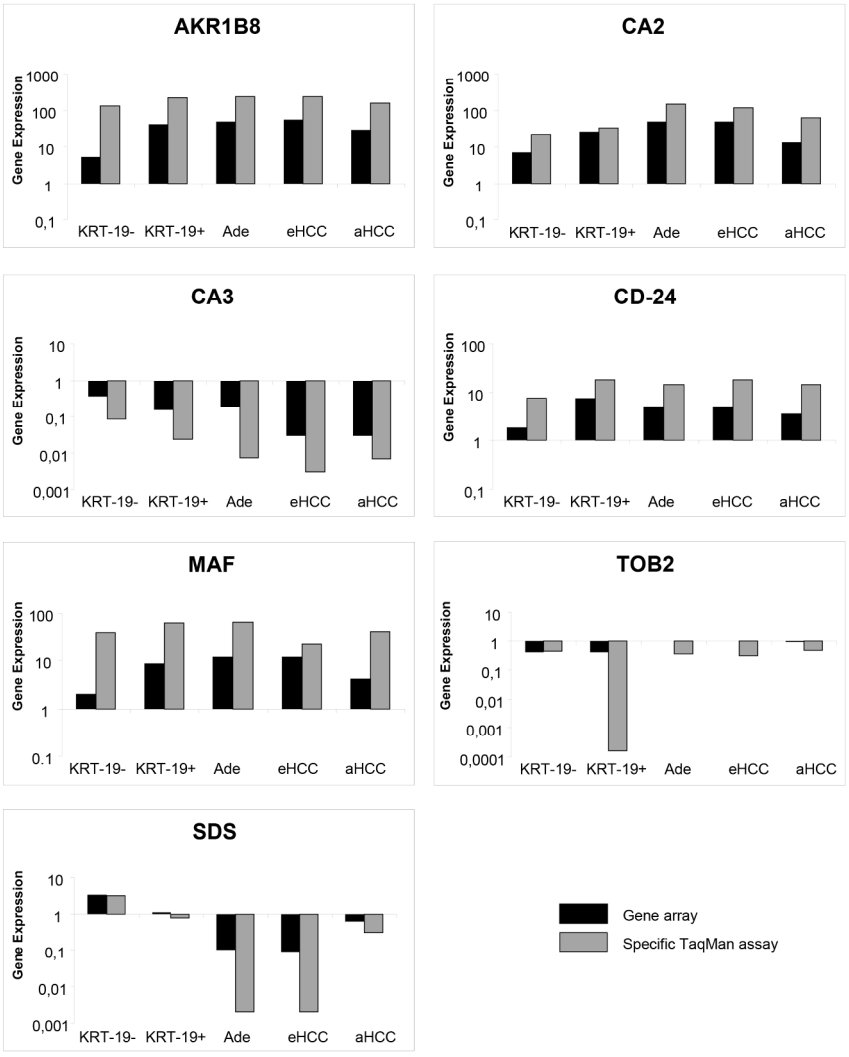
Supp. Fig. 1

190x254mm (300 x 300 DPI)



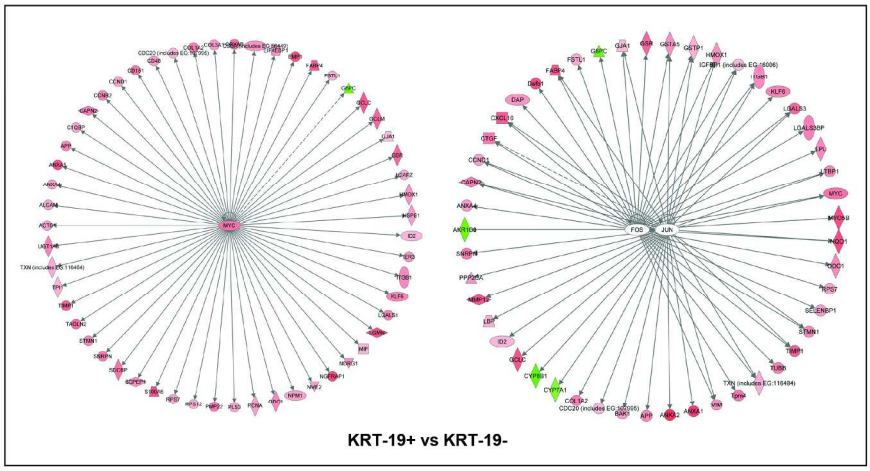
Supp. Fig. 2

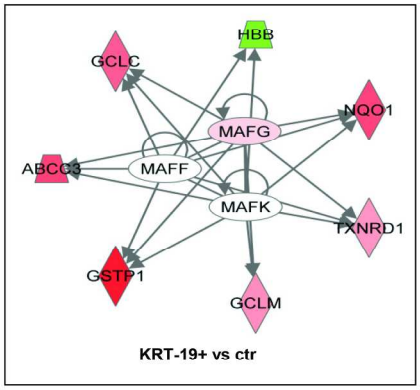
190x254mm (300 x 300 DPI)

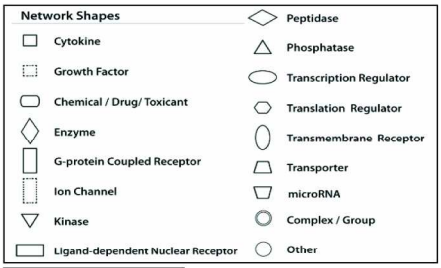


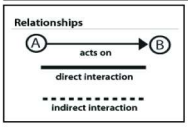
Supp. Fig. 3

190x254mm (300 x 300 DPI)



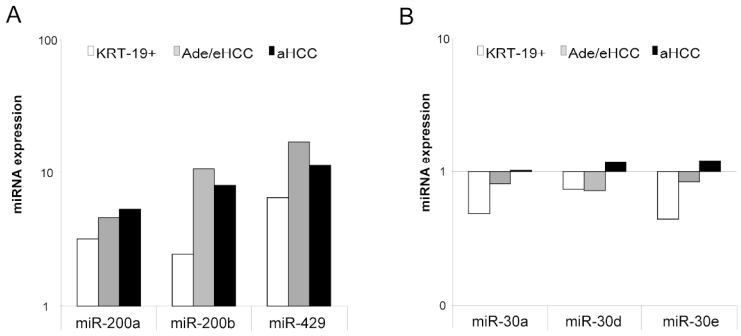






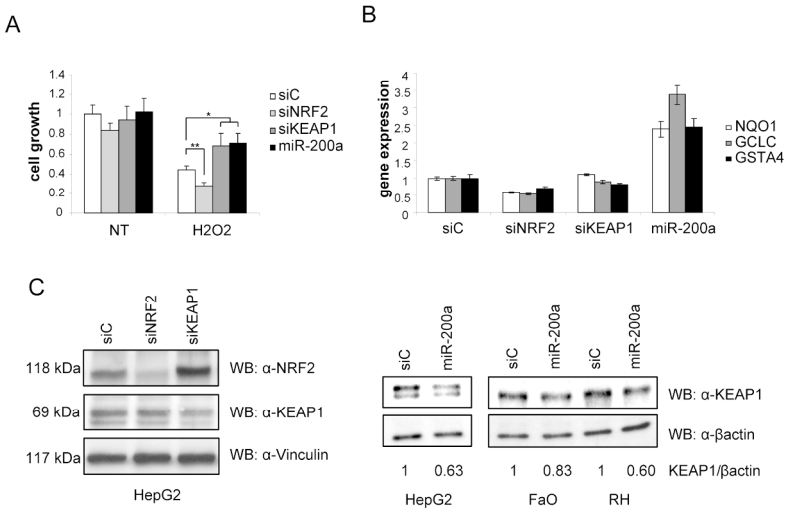
Supp. Fig. 4

190x254mm (300 x 300 DPI)



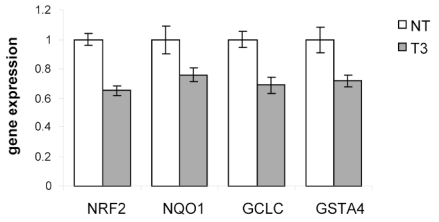
Supp Fig. 5

190x254mm (300 x 300 DPI)



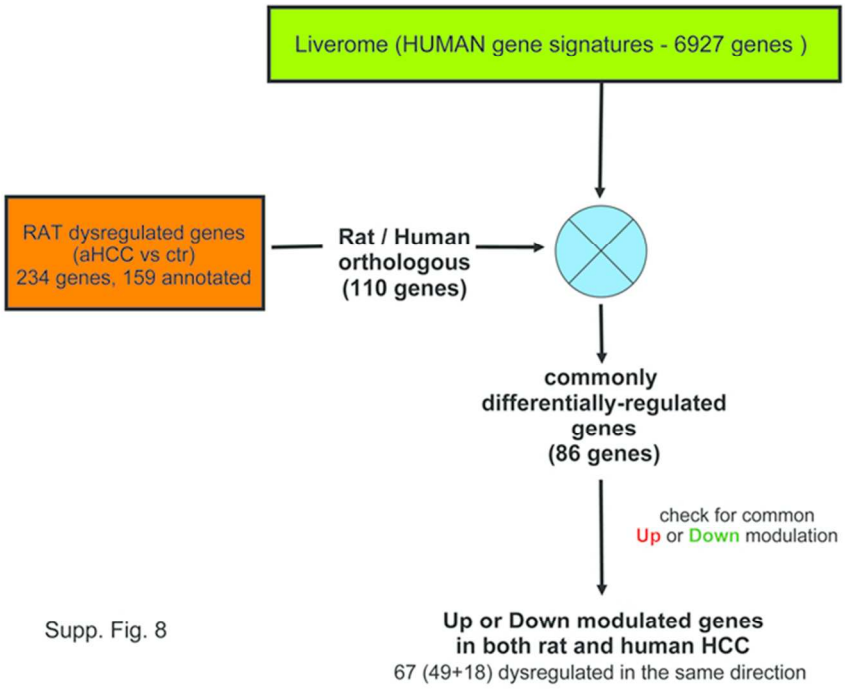
Supp. Fig. 6

190x254mm (300 x 300 DPI)



Supp. Fig. 7

190x254mm (300 x 300 DPI)



Supp. Fig. 8

67x50mm (300 x 300 DPI)

Supp. Table1. Percentage of early preneoplastic nodules (10 weeks after initiation with DENA) and aHCCs (14 months after initiation) positive for KRT-19.

Early Preneoplastic Nodules			aHCCs		
Total number of GSTP+ nodules	KRT-19+	%	Total number	KRT-19+	%
539	132	24.5	10	9	90

GSTP+ nodules and aHCCs were examined in the liver of 11 and 10 animals, respectively.

Supp. Table2A. MiRNAs differentially expressed in KRT-19+ vs ctr

miRNA	FC	miRNA	FC
up-regulated		miR-802	0,04
rno-miR-224	137,62	miR-486	0,06
miR-183	74,50	miR-375	0,10
miR-184	53,58	miR-125b-5p	0,16
miR-127	51,16	miR-210	0,18
miR-500	33,70	miR-19a	0,21
miR-539	30,07	miR-296-5p	0,21
miR-499	27,73	miR-192	0,22
miR-224	26,90	miR-122	0,23
miR-204	18,01	miR-194	0,23
miR-182	9,32	miR-29c	0,23
miR-214	6,45	miR-148a	0,25
miR-429	6,35	miR-200c	0,29
miR-185	5,82	miR-451	0,29
miR-200a	3,01	miR-26a	0,31
miR-200b	2,49	miR-19b	0,32
miR-93	2,19	miR-30b	0,37
down-regulated		miR-143	0,40
miR-133a	0,00	miR-145	0,40
rno-miR-1	0,01	miR-30e	0,44
miR-211	0,03	miR-203	0,49
miR-193	0,04	miR-30a	0,49

Supp. Table2B. MiRNAs differentially expressed in aHCC vs ctr

miRNA	FC	miRNA	FC
up-regulated		miR-185	2,62
rno-miR-224	69,07	miR-125a-5p	2,62
miR-499	44,55	miR-484	2,59
rno-miR-466c	36,74	miR-93	2,52
miR-224	36,39	miR-425	2,46
miR-449a	32,98	miR-99b	2,38
miR-183	14,86	miR-106b	2,30
miR-429	11,08	miR-29b	2,21
miR-200b	8,14	miR-322	2,13
miR-130b	7,53	miR-15b	2,12
miR-214	6,13	miR-497	2,00
miR-674	6,05	miR-181a	1,97
miR-130a	5,96	let-7i	1,89
miR-200a	5,34	miR-652	1,88
miR-182	4,22	**down-regulated**	
miR-324-5p	4,06	miR-292-3p	0,04
miR-301a	3,57	miR-1	0,06
miR-335-3p	3,37	miR-375	0,28
miR-199a-3p	3,23	miR-802	0,29
miR-138	3,17	miR-192	0,40
miR-331-3p	3,16	miR-125b-5p	0,42
miR-222	2,93	miR-194	0,46
miR-151-3p	2,92	miR-122	0,47
miR-193b	2,87	miR-193	0,49
miR-98	2,83		

MiRNAs with $P<0.05$ BH-corrected are reported.

Supp. Table3A. Commonly dysregulated miRNAs in aHCC and KRT-19+ nodules

miR-183-4395380
miR-224-4373187
miR-214-4395417
miR-192-4373108
miR-499-4381047
miR-200a-4378069
miR-429-4373355
miR-200b-4395362
miR-224-4395683
miR-125b-5p-4373148
miR-182-4395729
miR-194-4373106
miR-193-4395361
miR-93-4373302
miR-185-4395382
miR-802-4395566
miR-122-4395356
miR-375-4373027

Supp. Table3B. Commonly dysregulated miRNAs in Ade/eHCC and KRT-19+ nodules

miR-224-4395683
miR-802-4395566
miR-224-4373187
miR-451-4373360
miR-125b-5p-4373148
miR-194-4373106
miR-192-4373108
miR-185-4395382
miR-486-4378096
miR-200b-4395362
miR-429-4373355
miR-499-4381047
miR-93-4373302
miR-19b-4373098
miR-127-4373147
miR-214-4395417
miR-375-4373027
miR-122-4395356
miR-539-4378103
miR-148a-4373130
miR-26a-4395166
miR-1-4395765

Supp. Table3C. Commonly dysregulated miRNAs in aHCC and Ade/eHCC

miR-324-5p-4373052
miR-130a-4373145
miR-224-4373187
miR-214-4395417
miR-192-4373108
miR-499-4381047
miR-429-4373355
miR-200b-4395362
miR-1-4395333
miR-224-4395683
miR-125b-5p-4373148
miR-194-4373106
miR-93-4373302
miR-185-4395382
miR-802-4395566
let-7i-4395332
miR-301a-4373064
miR-122-4395356
miR-99b-4373007
miR-375-4373027
miR-125a-5p-4395309
miR-331-3p-4373046
miR-151-3p-4373304
miR-29b-4373288
miR-484-4381032

Supp. Table3D. Commonly dysregulated miRNAs in KRT-19+, Ade/eHCC and aHCC

miR-802-4395566
miR-224-4373187
miR-125b-5p-4373148
miR-194-4373106
miR-192-4373108
miR-185-4395382
miR-200b-4395362
miR-429-4373355
miR-499-4381047
miR-93-4373302
miR-224-4395683
miR-214-4395417
miR-375-4373027
miR-122-4395356

In this table, miRNAs in bold are dysregulated in human HCC as well.

Supp. Table4. MiRNAs differentially expressed in KRT-19+ vs. KRT-19-

miRNA	FC
up-regulated	
miR-224	16,11
miR-500	12,82
miR-539	6,52
miR-185	2,92
miR-34a	2,38
miR-93	2,09
miR-183	1,96
miR-182	1,74
miR-17	1,61
miR-320	1,52
miR-24	1,35
down-regulated	
miR-9	0,09
miR-222	0,11
miR-802	0,15
miR-486	0,25
miR-193	0,28
miR-148a	0,56
miR-26b	0,57
miR-194	0,57
miR-122	0,61
miR-125b-5p	0,61
miR-26a	0,63
miR-192	0,66

Supp. Table5. List of the 86 rat/human orthologous genes involved in HCC progression

Gene symbol	Gene name	Frequency
ABCC3	ATP-binding cassette, sub-family C (CFTR/MRP), member 3	1 list(s)
ABCC6	ATP-binding cassette, sub-family C (CFTR/MRP), member 6	4 list(s)
AGXT	alanine-glyoxylate aminotransferase	10 list(s)
AKR1B10	aldo-keto reductase family 1, member B10	10 list(s)
AKR1C1	aldo-keto reductase family 1, member C1	3 list(s)
AKR1C2	aldo-keto reductase family 1, member C2	7 list(s)
AKR1C3	aldo-keto reductase family 1, member C3	6 list(s)
AKR1C4	aldo-keto reductase family 1, member C4	4 list(s)
ALDOA	aldolase A, fructose-bisphosphate	6 list(s)
ANGPTL4	angiopoietin-like 4	3 list(s)
ANXA1	annexin A1	6 list(s)
ANXA2	annexin A2	8 list(s)
ANXA4	annexin A4	6 list(s)
ANXA5	annexin A5	4 list(s)
ANXA7	annexin A7	2 list(s)
ASAH1	N-acylsphingosine amidohydrolase 1	1 list(s)
AVPR1A	arginine vasopressin receptor 1A	1 list(s)
BAMBI	BMP and activin membrane-bound inhibitor homolog (Xenopus laevis)	3 list(s)
BOK	BCL2-related ovarian killer	2 list(s)
C15orf23	chromosome 15 open reading frame 23	2 list(s)
C5orf13	chromosome 5 open reading frame 13	6 list(s)
CCNB2	cyclin B2	4 list(s)
CD63	CD63 molecule	2 list(s)
CDH17	cadherin 17, LI cadherin (liver-intestine)	2 list(s)
COL1A2	collagen, type I, alpha 2	9 list(s)
CRYAB	crystallin, alpha B	2 list(s)
CSDA	cold shock domain protein A	3 list(s)
CTGF	connective tissue growth factor	7 list(s)
CX3CL1	chemokine (C-X3-C motif) ligand 1	1 list(s)
CXCL16	chemokine (C-X-C motif) ligand 16	2 list(s)
DEFB1	defensin, beta 1	1 list(s)
DHRS7	dehydrogenase/reductase (SDR family) member 7	1 list(s)
EGF	epidermal growth factor	2 list(s)
ENTPD5	ectonucleoside triphosphate diphosphohydrolase 5	1 list(s)
ESM1	endothelial cell-specific molecule 1	1 list(s)
FABP4	fatty acid binding protein 4, adipocyte	1 list(s)
GCLC	glutamate-cysteine ligase, catalytic subunit	4 list(s)
GCLM	glutamate-cysteine ligase, modifier subunit	5 list(s)
GNMT	glycine N-methyltransferase	7 list(s)
GPX2	glutathione peroxidase 2 (gastrointestinal)	4 list(s)
HAGH	hydroxyacylglutathione hydrolase	6 list(s)
HAL	histidine ammonia-lyase	3 list(s)
HAO2	hydroxyacid oxidase 2 (long chain)	2 list(s)
HES6	hairy and enhancer of split 6 (Drosophila)	1 list(s)
ICA1	islet cell autoantigen 1, 69kDa	1 list(s)
IER3	immediate early response 3	3 list(s)
IGFBP3	insulin-like growth factor binding protein 3	9 list(s)
EHHADH	enoyl-CoA, hydratase/3-hydroxyacyl CoA dehydrogenase	6 list(s)
RPSA	ribosomal protein SA	12 list(s)
LBP	lipopolysaccharide binding protein	1 list(s)
LGALS3BP	lectin, galactoside-binding, soluble, 3 binding protein	4 list(s)
LCP1	lymphocyte cytosolic protein 1 (L-plastin)	2 list(s)

LPL	lipoprotein lipase	2 list(s)
MYC	v-myc myelocytomatosis viral oncogene homolog (avian)	3 list(s)
NDRG2	NDRG family member 2	3 list(s)
NQO1	NAD(P)H dehydrogenase, quinone 1	3 list(s)
PCYOX1	prenylcysteine oxidase 1	1 list(s)
PIK3C3	phosphoinositide-3-kinase, class 3	2 list(s)
PIR	pirin (iron-binding nuclear protein)	3 list(s)
PLA2G7	phospholipase A2, group VII (platelet-activating factor acetylhydrolase)	1 list(s)
PLSCR1	phospholipid scramblase 1	2 list(s)
PMM1	phosphomannomutase 1	1 list(s)
PPAP2C	phosphatidic acid phosphatase type 2C	1 list(s)
PQLC3	PQ loop repeat containing 3	1 list(s)
PYGB	phosphorylase, glycogen; brain	7 list(s)
RND1	Rho family GTPase 1	1 list(s)
MRPL13	mitochondrial ribosomal protein L13	1 list(s)
RTKN	rhotekin	1 list(s)
S100A11	S100 calcium binding protein A11	3 list(s)
SERPINB1	serpin peptidase inhibitor, clade B (ovalbumin), member 1	2 list(s)
SLC22A25	solute carrier family 22, member 25	1 list(s)
SLC40A1	solute carrier family 40 (iron-regulated transporter), member 1	1 list(s)
SNX10	sorting nexin 10	1 list(s)
SPINT1	serine peptidase inhibitor, Kunitz type 1	3 list(s)
SYNJ2	synaptojanin 2	2 list(s)
TALDO1	transaldolase 1	3 list(s)
TAX1BP3	Tax1 (human T-cell leukemia virus type I) binding protein 3	3 list(s)
TMEM55A	transmembrane protein 55A	1 list(s)
TNFRSF12A	tumor necrosis factor receptor superfamily, member 12A	2 list(s)
TSPAN3	tetraspanin 3	1 list(s)
TSPAN8	tetraspanin 8	5 list(s)
TUBB	tubulin, beta	6 list(s)
TUBB2A	tubulin, beta 2A	1 list(s)
TUBB6	tubulin, beta 6	1 list(s)
TXNRD1	thioredoxin reductase 1	2 list(s)
VIM	vimentin	11 list(s)

Genes in bold are dysregulated in KRT-19+ preneoplastic lesions as well.

Supp. Table6. List of the 26 miRNAs involved in both rat and human HCC progression

miRNA	References
miR-224	1, 4, 6, 7, 8
miR-183	4, 6, 8
miR-130b	7
miR-130a	15
miR-182	4, 6, 8, 18
miR-324-5p	4, 7, 8
miR-301a	4, 5, 6, 8
miR-331-3p	6
miR-222	3, 4, 6, 8
miR-151-3p	4, 8
miR-193b	6
miR-98	14
miR-185	2, 6
miR-93	5, 7, 8
miR-425	9
miR-99b	13
miR-106b	5, 7, 8
miR-15b	12
miR-181a	10
let-7i	14
miR-292-3p	11
miR-1	2, 5
miR-375	16, 17
miR-125b-5p	2, 5, 6, 7, 8
miR-194	2, 8
miR-122	3, 5, 7, 8

MiRNAs in bold are dysregulated in KRT-19+ preneoplastic lesions as well.

References

1 Murakami Y. et al, Oncogene, 2006
2 Budhu A. et al, Hepatology, 2008
3 Cairo S. et al, Proc Natl Acad Sci U S A, 2010
4 Wang YU. et al, J Biol Chem, 2008
5 Wong CM. et al, Hepatology, 2012
6 Wong QW. et al, Clin Cancer Res, 2010
7 Connolly E. et al, Am J Pathol, 2008
8 Borel F. et al, J Hepatol, 2012
9 Pineau P. et al, Proc Natl Acad Sci U S A, 2010
10 Junfang J. et al, Hepatology, 2009
11 Meng F. et al, Gastroenterology, 2007
12 Huang YH. et al, Plos One, 2012
13 Kutay H. et al, J Cell Biol, 2006
14 Huang YS. et al, J Gastr Hep, 2008
15 Huang XH. et al, Hepatol Res, 2009
16 Liu AM. et al, Biochem Biophys Res Commun, 2010
17 He XX. et al, Oncogene, 2011
18 Li S. et al, Hepatology, 2009