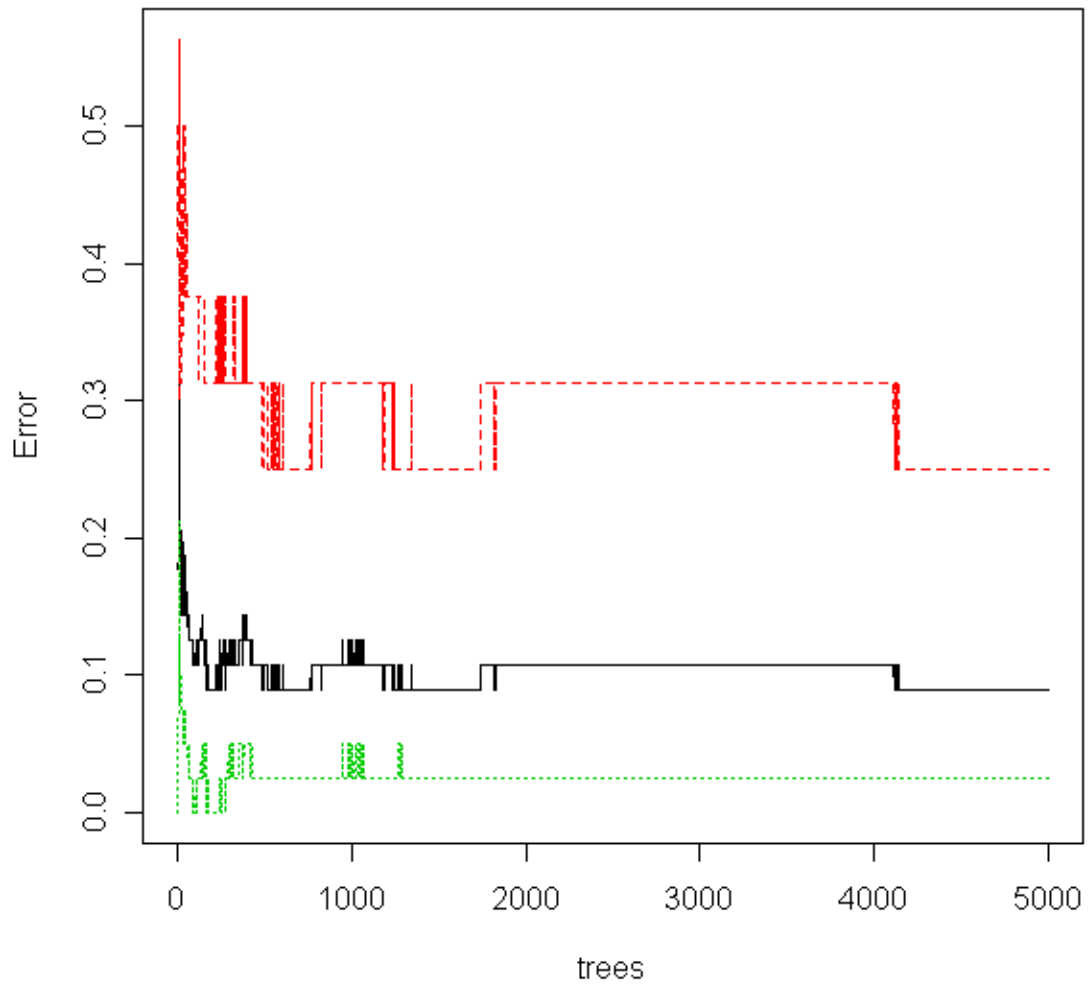


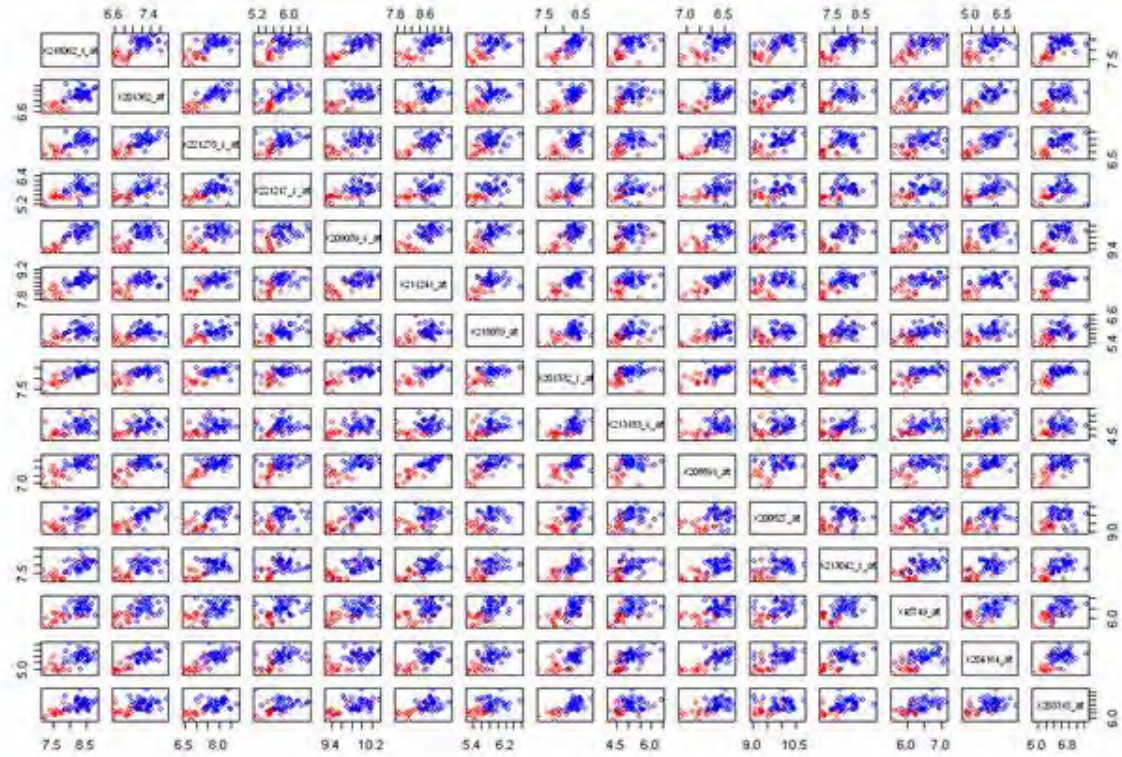
Additional file 1 – Supplemental Data

Quality Control. The quality of the total cellular RNA isolated was assessed by examining RNA purity by spectrophotometer as the absorbance at 260, 270 and 280 nm, and estimating the $A_{260}:A_{280}$ and $A_{260}:A_{270}$ ratios. In addition, RNA integrity was assessed by Agilent Bioanalyzer by estimating the *28S:18S* rRNA ratios as the sum of both rRNA peaks as a proportion of all RNA, the percent rRNA, and RNA integrity number (RIN). Quality assessment of each GeneChip was performed by assessing the average background, scaling factor, percent present calls, the 3':5' ratio for GAPDH, β -actin, and ISGF, and the overall 3':5' ratio with associated 95% confidence interval. Classical multidimensional scaling (MDS) was applied to the standardized QC variables and the resulting two-dimensional MDS plot to determine whether any chips had outlying quality.

Quality control. RNA samples met all of the quality control criteria for sample preparation. cDNA synthesis and IVT were checked in all the samples showing results in concordance with the pre-established quality control criteria. Average background intensity and percent of probe sets called Present were consistent across the set of GeneChips. The 3' to 5' prime ratios for GAPDH and β -actin were consistently close to 1.0 indicating high quality RNA. No microarray was eliminated due to quality concerns.



Supplemental Figure-1: Unbiased error estimates (black line) with 95% interval (red and green lines) obtained using the out of bootstrap observations are plotted as the number of trees included in the random forest ranged from 1 to 5,000. Final error estimate was 8.93%.



Supplemental Figure-2: Pairwise scatterplot for these 15 probe sets revealed that all probe sets were correlated, with the cirrhotic tissues with HCC having lower expression values (red) than the cirrhotic tissues without HCC (blue).

Supplemental Table I. Characteristics of the studied tissues

A) Tumor samples

Sample ID	Tumor size, number,		TNM Stage
	location	Tumor grade	
1-HCCTS	4.3-R	moderately differentiated	T4N0M0
2-HCCTS	2.2-R	moderately differentiated	T2N0M0
3-HCCTS	4.2-R	moderately differentiated	T3N0M0
4-HCCTS	2-R,1.5-R	moderately differentiated	T2N0M0
5-HCCTS	4R	moderately differentiated	T2N0M0
6-HCCTS	4.3L	well differentiated	T3N0M0
7-HCCTS	4-L, 6.5-R, 3.8-R	moderately differentiated	T4N0M0
8-HCCTS	2.3-R	well differentiated	T2N0M0
9-HCCTS	2.3 -R	moderately differentiated	T2N0M0
10-HCCTS	3.3-R	moderately differentiated	T2N0M0
11-HCCTS	3-L	moderately differentiated	T2N0M0
12-HCCTS	6.2-R	well differentiated	T3N0M0
13-HCCTS	4.3-R	moderately differentiated	T2N0M0
14-HCCTS	1.0-R	moderately differentiated	T1N0M0
15-HCCTS	3.3-L	well/moderately differentiated	T2N0M0
16-HCCTS	5.0- L	moderately diffused	T3N0M0
17-HCCTS	3.0-L	moderately differentiated	T3N0M0
18-HCCTS	3.0-R	low grade	T2N0M0
19-HCCTS	2.0-L	well differentiated	T1N0M0
20-HCCTS	5.3-L	moderately differentiated	T4N0M1
21-HCCTS	3.0-L	moderately differentiated	T3N0M0
22-HCCTS	2.7-R,1.7-R,0.9-L	moderately differentiated	T3N0M0
23-HCCTS	5.2-R,1.0-R	moderately differentiated	T3N0M0

24-HCCTS	3.8-R	moderately differentiated	T2N0M0
25-HCCTS	4.5-R	moderately differentiated	T2N0M0
26-HCCTS	2.3-R	moderately diffused	T2N0M0
27-HCCTS	4.8-R,1.9-L	modertely differentiated	T4N0M1
28-HCCTS	1.5-R	well differentiated	T4N0M0
29-HCCTS	2.2-R	moderately differentiated	T2N0M0
		well to moderately	
30-HCCTS	2.2-R	differentiated	T2N0M0
31-HCCTS	1.0-R,2.0-L	moderately differentiated	T2N0M0
32-HCCTS	1.5-R,3-9R	moderately differentiated	T2N0M0
33-HCCTS	1.3-L,1.5-R, 1.3-R	moderately differentiated	T4N0M0
34-HCCTS	5.5-L, 1.2-R	well differentiated	T4NOM1
35-HCCTS	1.0 R	moderately differentiated	T1N0M0
36-HCCTS	1.0-R	well differentiated	T1N0M0
37-HCCTS	2.6-R	moderately differentiated	T2N0M0
38-HCCTS	2.0-R	well differentiated	T2N0M0
		well to moderately	
39-HCCTS	1.6-R,1.4-L	differentiated	T2N0M0
40-HCCTS	5.2-L	moderately diffuse	T3N0M0
41-HCCTS	1.9-R,1.7-R	well differentiated	T2N0M0
		well to moderately	
42-HCCTS	4.1-R,1.6-L	differentiated	T4N0M1
43-HCCTS	1.7/1.7R/1.2L	well differentiated	T4N0M1
44-HCCTS	4.0-R,1.0-R,1.0-R	high grade HCC	T4N0M1
45-HCCTS	1.4/1.7L	moderately differentiated	T4N0M1
46-HCCTS	2.0-R	high grade HCC	T2N0M0
47-HCCTS	2.2-R	moderately differentiated	T2N0M0

B) HCV Cirrhotic tissues

Sample ID	HCV (+)	HCC	Matched tumor tissue	Ishak grade	Knodell score
1-CT	Yes	Yes	yes	6	11
2-CT	Yes	Yes	yes	6	14
3-CT	Yes	Yes	yes	6	12
4-CT	Yes	No	N/A	6	15
5-CT	Yes	Yes	no	6	13
6-CT	Yes	No	N/A	6	14
7-CT	Yes	Yes	yes	6	11
8-CT	Yes	No	N/A	6	12
9-CT	Yes	No	N/A	6	13
10-CT	Yes	No	N/A	6	9
11-CT	Yes	No	N/A	6	9
12-CT	Yes	No	N/A	6	10
13-CT	Yes	No	N/A	6	13
14-CT	Yes	No	N/A	6	11
15-CT	Yes	No	N/A	6	10
16-CT	Yes	No	N/A	6	12
17-CT	Yes	No	N/A	6	11
18-CT	Yes	No	N/A	6	13
19-CT	Yes	Yes	yes	6	11
20-CT	Yes	Yes	yes	6	11
21-CT	Yes	No	N/A	6	11
22-CT	Yes	No	N/A	6	9
23-CT	Yes	Yes	yes	6	11
24-CT	Yes	Yes	no	6	9

25-CT	Yes	No	N/A	6	13
26-CT	Yes	Yes	yes	6	9
27-CT	Yes	No	N/A	6	9
28-CT	Yes	No	N/A	6	13
29-CT	Yes	No	N/A	6	11
30-CT	Yes	Yes	yes	6	11
31-CT	Yes	No	N/A	6	12
32-CT	Yes	No	N/A	6	9
33-CT	Yes	No	N/A	6	12
34-CT	Yes	No	N/A	6	11
35-CT	Yes	No	N/A	6	7
36-CT	Yes	No	N/A	6	7
37-CT	Yes	No	N/A	6	12
38-CT	Yes	No	N/A	6	7
39-CT	Yes	No	N/A	6	11
40-CT	Yes	No	N/A	6	11
41-CT	Yes	No	N/A	6	11
42-CT	Yes	No	yes	6	9
43-CT	Yes	No	N/A	6	9
44-CT	Yes	No	N/A	6	11
45-CT	Yes	Yes	yes	6	9
46-CT	Yes	Yes	yes	6	9
47-CT	Yes	No	N/A	6	14
47-CT	Yes	No	N/A	6	12
48-CT	Yes	No	N/A	6	12
49-CT	Yes	Yes	yes	6	11
50-CT	Yes	Yes	yes	6	13
51-CT	Yes	Yes	yes	6	12

52-CT	Yes	No	N/A	6	11
53-CT	Yes	Yes	no	6	12
54-CT	Yes	No	N/A	6	12
55-CT	Yes	No	N/A	6	11
56-CT	Yes	No	N/A	6	11
57-CT	Yes	No	N/A	6	14
58-CT	Yes	Yes	yes	6	12

Supplemental Table II. Seventeen unique probe sets differentially expressed in HCV-HCC samples when compared with HCV cirrhotic tissues from patients with HCC.

AffyID	Symbol	Gene Name	HCV-Cirrhosis	HCV-HCC	Fold change
200055_at	TAF10	TAF10 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 30kDa	8.83	9.46	1.55
201917_s_at	SLC25A36	solute carrier family 25, member 36	7.16	5.96	-2.30
202600_s_at	NRIP1	nuclear receptor interacting protein 1	9.62	7.89	-3.32
202984_s_at	BAG5	BCL2-associated athanogene 5	5.40	4.76	-1.56
204408_at	APEX2	APEX nuclease (apurinic/apyrimidinic endonuclease) 2	5.92	6.29	1.29
204713_s_at	F5	coagulation factor V (proaccelerin, labile factor)	11.05	9.71	-2.53
205171_at	PTPN4	protein tyrosine phosphatase, non-receptor type 4 (megakaryocyte)	6.37	5.60	-1.71

208994_s_at	PPIG	peptidylprolyl isomerase G (cyclophilin G)	8.02	7.37	-1.57
212106_at	UBXD8	UBX domain containing 8	7.31	6.64	-1.59
212634_at	KIAA0776	KIAA0776	7.76	6.66	-2.14
212758_s_at	ZEB1	zinc finger E-box binding homeobox 1	7.64	6.69	-1.93
214679_x_at	GNA11	guanine nucleotide binding protein (G protein), alpha 11 (Gq class)	6.41	6.91	1.41
214723_x_at	KIAA1641	KIAA1641	5.61	4.55	-2.08
215447_at	TFPI	tissue factor pathway inhibitor (lipoprotein- associated coagulation inhibitor)	7.02	5.92	-2.14
T216022_at		NA	5.01	4.21	-1.74
216775_at	USP53	ubiquitin specific peptidase 53	4.22	3.85	-1.29
218213_s_at	C11orf10	chromosome 11 open reading frame 10	9.65	10.26	1.53