SUPPLEMENTARY DATA

Supplementary FIGURES

Supplementary Figure 1. NMF clusterings of mRNA expression data from 35 FLCs and their validation through the Affinity propagation clustering.

Supplementary Figure 2. FLC tumors over-expressed neuroendocrine markers and their integration in the FLC molecular classification.

Supplementary Figure 3. Heatmap representation of FLC chromosomal alterations and scheme of the DNAJB1-PRKACA fusion transcript.

Supplementary Figure 4. Somatic mutations found by WES and technical validation by TES and Sanger sequencing.

Supplementary FIGURES LEGENDS

Supplementary Figure 1.

Supplementary Figure 2.

Supplementary Figure 3.

Supplementary Figure 4.

Supplementary MATERIALS and METHODS

Gene-expression microarray profiles

Immunohistochemistry

Genome-wide analysis of DNA copy number alteration

Validation of the DNAJB1-PRKACA fusion transcript

Whole-exome sequencing chromosomal rearrangements analysis

Whole-exome sequencing somatic mutation detection

Targeted-exome sequencing somatic mutation detection

Prognostic gene expression data analysis

Supplementary REFERENCES

Supplementary TABLES

Supplementary Table 1. Clinicopathological characteristics of 72 FLC patients in the training and validation cohorts

Supplementary Table 2. Leave-one-out cross validation of FLC subclasses

Supplementary Table 3. FLC gene signature

Supplementary Table 4. FLC immunohistochemistry grading

Supplementary Table 5. Chromosomal arm-level genomic copy number alterations

Supplementary Table 6. Molecular profile of the broadly chromosomal altered samples

Supplementary Table 7. Focal amplifications in FLC samples

Supplementary Table 8. Focal deletions in FLC samples

Supplementary Table 9. Chromosomal arm-level genomic copy number alterations of the Fresh Frozen cohort

Supplementary Table 10. Summary of the damaging mutations found by WES

Supplementary Table 11. Mortality Index for the different prognostic group samples

Supplementary Table 12. Enrichment of selected gene signatures in 3 liver cancer cohorts (HCC, ICC and FLC) evaluated by NTP (FDR<0.05)

Supplementary FIGURES

Supplementary FIGURE LEGENDS

Supplementary Figure 1. A) NMF consensus matrices depicting inter-sample correlation among 35 samples when 2 to 5 classes were assumed. Red color indicates highly robust co-clustering of samples. Plot of cophenetic correlation coefficient indicates that the most robust clustering is achieved when 3 classes are assumed in the dataset. **B)** Molecular classification validation by the Affinity propagation clustering represented by Principal component analysis (PCA), where PC1 means first principal component, and PC2 second principal component.

Supplementary Figure 2. A) Plot representation of neuroendocrine gene markers expression in FLC patients compared to non-tumoral samples by fold-change. FLC tumors present significant higher expression of VCAN. Vertical axis: gene expression of samples; Horizontal axis; FLC tissues (left) versus non-tumoral tissues (right). **B)** Integration of the neuroendocrine gene expression of FLC patients and non-tumoral tissues in the molecular classification, and their enrichment, when significant, in the FLC classes.

Supplementary Figure 3. A) Heatmap representation of FLC chromosomal alterations. Each tumor is displayed in separate columns, grouped by molecular class, and chromosome positions are indicated along the y axis. Chromosomal amplifications and deletions are represented in red and blue, respectively. **B)** Scheme of the DNAJB1-PRKACA fusion transcript

and example of the sanger sequencing result of the RT-PCR product from a FLC sample with the fusion transcript in the intersection point between the end of exon 1 of DNAJB1 and the start of exon 2 of PRKACA.

Supplementary Figure 4. A) Distribution of the mutations obtained using SIFT and PolyPhen2 algorithms as non-coding and coding, synonymous and non-synonymous, missense and nonsense, and finally among damaging, probably or possibly damaging and benign. **B)** Alignment visualization of TES results through Golden Helix Genome Browse software. **C)** Results of the technical validation of the damaging mutations found in *BRCA2* by Sanger sequencing.

Supplementary MATERIALS and METHODS

Gene-expression microarray profiles

Total RNA was isolated from 4 freshly cut 5-µm thick FFPE sections with the miRNA easy FFPE kit (Qiagen) using the semi-automated procedure QIAcube (Qiagen). To extract the tumoral and non-tumoral RNA, tissue sections were macrodissected in order to avoid contamination between liver tissues. Tissue macrodissection consists on defining areas representing tumoral and non-tumoral tissue by an H&E immunostaining, which is used as reference once obtaining specific tissue material in the FFPE slides with a blade. Quant-iT Ribogreen RNA assay kit (Invitrogen) was used to quantify the RNA. RNA quality was assessed by real-time quantitative reverse transcription PCR (qRT-PCR) using RPL13A Taq-Man probe (Applied Biosystems), cut-off Ct<28 cycles. Total RNA (250ng) was used for whole-genome gene-expression profiling by DASL array (Illumina). Poor quality mRNA expression profiles exhibiting outlier patterns were detected and removed as previously described[1]. Quality threshold was set based on %P-call (proportion of gene probes with a "present" signal) to ensure inter-sample similarity measured by Pearson correlation to "median" array defined as a vector of median signal intensity for each probe. We excluded 7 samples with a %P-call of <30%, resulting on 35 tumoral and 4 normal liver tissues. Raw data were normalized using cubic spline algorithm implemented in the Illumina Normalizer module of the GenePattern analysis Toolkit[2,3] and preprocessed using the PreprocessDataset module (GenePattern) with the following parameters: floor=100, fold change=3, ceiling=120.000 and min delta=100.

Non-negative matrix factorization (NMF) based consensus clustering method[4] was performed for the unsupervised clustering of the gene expression profiles, 2000 iterative

clusterings of randomly bootstrapped samples to identify highly robust molecular classes were determined (NMFconsensus module, GenePattern).

Robustness of the obtained FLC molecular classification was evaluated in a conventional leave-one-out cross validation (LOOCV) procedure, where one sample at a time was held-out and a predictive model was trained using the rest of samples. Based on the model, the sample held-out was reclassified. This process was repeated for all 35 samples, and the error number was counted by k-nearest neighbor prediction method (KNNXValidation, modules, GenePattern).

To define robust FLC molecular classes gene signature not affected by outlier gene/sample, we employed the following LOOCV signature gene refinement procedure following the same process as explained above. Genes over-expressed in each subclass compared to the rest were identified by t-test with a significance threshold $p<0.00001$. Genes selected as significant throughout the LOOCV procedure were included in the final list of signature genes. The analysis was performed by using custom R codes (www.r-project.org).

Validation of the number of molecular classes was performed using the independent method Affinity propagation clustering[5] that calculates the optimal number of clusters in which a set of samples can be divided. The input for the algorithm was the expression values of the genes which were able to differentiate two or more classes, described as those passing the Tukey Honestly Significant Difference *post hoc* test with a p-value < 0.01 and absolute log2 fold change >2. The dispersion between samples and the cluster they belong to is represented graphically (**Suppl. Figure 1**) by calculating a Principal component analysis and plotting the first two components (PC1 and PC2).

Molecular pathways and gene expression signatures associated with classes were evaluated using Gene Set Enrichment Analysis (GSEA)[6] for Molecular Signature Database gene sets (MSigDB, www.broadinstitute.org/msigdb)[7] and Nearest template prediction (NTP), both modules from GenePattern. The gene-expression signatures in liver cancer were previously reported[8]. Ingenuity® pathway analysis was also run for functional annotation.

Immunohistochemistry

Immunohistochemical staining was performed in 3-µm thick FFPE whole-section slides from 42 FLC tumours and the paired non-tumoral liver when available, after heat-induced antigen retrieval in microwave with sodium citrate (pH=6) or EDTA (pH=8) buffer. The primary antibodies and dilutions used were: rabbit EGFR (D38B1, Cell Signalling) 1:100, rabbit phospho-

RPS6 (Ser240/244) (Cell Signalling) 1:200, mouse HepPar1 (OCH1E5, Dako) ready to use, mouse Cytokeratin 7 (OV-TL 12/30, Dako) 1:50, mouse EpCAM (VU-1D9, Thermo Scientific) 1:100, and mouse Cytokeratin 19 (RCK108, Dako) 1:50. Immunoreactivity was independently graded by three liver expert pathologists (ST, LR, and SW) blinded to the expression profiling results. The variables measured were immunostaining intensity (score 0-3; 0: absent, 1: weak, 2: moderate and 3: strong), pattern (score 1-3; 1: few cells (≤1%), 2: focal (≤ 10%) and 3: diffuse), and localization of the staining (membranous, cytoplasmic or nuclear). Samples were defined positive for HepPar1, CK7 and when intensity of staining was 2 or higher, and focal o diffuse pattern, and for p-RPS6, CK19, EGFR and EpCAM by the presence of stained tumoral cells. Once patients were classified as positive or negative, Fisher exact test was performed to assess significance of the distribution between the molecular classes.

Genome-wide analysis of DNA copy number alteration

Genomic DNA was isolated from 7 macrodissected 5-µm thick FFPE sections with the QIAamp DNA FFPE Tissue Kit using QIAcube (Qiagen). To extract tumoral and non-tumoral DNA, tissue sections were macrodissected to avoid contamination between tissues. DNA quantity was assessed using Quant-It PicoGreen dsDNA Assay kit (Invitrogen) and DNA quality was determined by qRT-PCR of RNase P (Applied Biosystems), deltaCt defined as ct values(FFPE DNA) - Ct value (non-fragmented DNA), should not exceed a value of 4.

In order to depict the chromosomal alterations in FLC, copy number variations (CNVs) at over 715,000 genomic loci were analyzed in 32 primary tumors and 17 paired normal livers from FFPE samples. The frequency and magnitude of copy number gains and losses were evaluated in tumors using Genomic identification of significant targets in cancer (GISTIC) algorithm[9]. Genomic DNA (250 ng) of these 49 samples were assessed using Infinium FFPE restoration kit (Illumina) and HumanOmniExpress FFPE-12 v1.0 DNA Analysis BeadChip (Illumina).

This array provides optimized tag SNPs with high coverage of common variants in a genome-wide scale as well as high-resolution signals for DNA copy number. Two independent types of information, log R ratio (LRR) and B allele frequency (BAF), reported directly from GenomeStudio (Illumina) along with genotype calls at each SNP locus, were used for CNV analysis[10]. BAF values of tumor tissue can be transformed into the so called mirrored BAF (mBAF), by either taking advantage of paired normal tissue genotype or applying some pre-specified criteria as detailed[11], providing better signal-to-noise ratio for segmentation in CNV

analysis than LRR, which are commonly used in other cancer studies. Circular binary segmentation (CBS)[12,13] is employed to partition tumor genome into segments with constant copy number based on mBAF. Within each segment, the median LRR values at all the SNP loci was taken as a surrogate measurement of the actual copy number, and used in genomic identification of significant targets in cancer (GISTIC) analysis, as implemented in the GISTIC2 package[9]. Broad and focal regions with recurrent copy number changes at relatively high frequency in the study cohort were detected, and the significance of such findings was evaluated through a permutation test. The reported regions are those with false discovery rate (FDR)[14] less than 0.25, and those genes within these regions were considered as candidates for further investigation. Focal aberrations are defined as frequently altered regions that affect less than half size of the chromosome arm where they are located, and broad alterations are those that affect more than half size of the chromosome. All genomic coordinates were annotated based on the hg19 assembly.

For the 23 FF FLC tissue samples, DNA was extracted as previously published[15], and hybridized on the Infinium HumanHap370CNV Genotyping BeadChip SNP array. Data was extracted and evaluated as described above for the FFPE cohort.

Due to the small number of FF samples, GISTIC2 was not applicable to this dataset. Thus, we applied different criteria to define for each genomic segments the high confident events, such as allelic imbalance, gain and loss: 1) allelic imbalance=mBAF.median > 0.6 AND |LRR.median| < 0.5 * sigma; 2) gain=mBAF.median > 0.57 AND LRR.median > sigma, and 3) loss: mBAF.median > 0.57 AND LRR.median < -sigma, where mBAF=|BAF-0.5| + 0.5 at heterozygous sites. The parameters used stand for mBAF.median the median of mBAF values at heterozygous sites within the segment; LRR.median the median of LRR values at all sites within the segment and sigma the estimated standard deviation from all LRR values of a sample, reflecting the noise level in LRR of the sample.

Validation of the DNAJB1-PRKACA fusion transcript

500 micrograms of RNA were retro-transcribed into cDNA using the High-Capacity cDNA Reverse Transcription Kit from Applied Biosystems following manufacturer's instructions. The resulting cDNA was used as template for semiquantitative polymerase chain reaction (PCR) amplification using the primers: Fw_DNAJB1-Exon1 5'-GTTCAAGGAGATCGCTGAGG-3' and Rv_PRKACA-Exon3 5'-TTCCCGGTCTCCTTGTGTTT-3'. To detect the presence of the fusion product, the PCR amplifications on human tissues were

performed using the following protocol: 95°C for 5 minutes, 34 cycles of 95°C denaturation for 15 seconds, 56°C annealing for 30 seconds, and 68°C extension for 45 seconds, followed by a 5 minute final extension at 68°C. The PCR amplifications were performed in a volume of 35µL reaction mixture containing 1.4mM MgCl2, 0.2mM of each dNTP, 0.125mM of each primer, 0.7µL DMSO, 1U of Platinum Taq DNA Polymerase (Invitrogen) and 2µL (100ng) of cDNA. PCR products were purified using the Qiaquick PCR purification kit (Qiagen) and sequenced using an Applied Biosystems 3700 DNA sequencer (ABI PRISM® 3730XL; Applied Biosystems).

Whole-exome sequencing chromosomal rearrangements analysis

One FF FLC-normal liver pair underwent whole-exome sequencing (WES) on HiSeq2000 sequencer (Illumina) with 50X of coverage. Results achieved good quality metrics: high read quality, high mapping rate, low duplication rate and adequate coverage of the target regions. PCR and optical duplicates, as well as low-quality (Q<20) and non-uniquely mapped reads were removed.

In the setting of paired normal-tumor DNA sequencing, somatic copy number variants (sCNV) can be detected from two sources of evidences. (1) The ratio of read depth of tumor vs normal is different in CNV regions from the unchanged regions. Briefly, ratio of a given locus was defined as ratio=(read depth in tumor)/(read depth in normal). Therefore a value of log2ratio greater than 0 indicates copy number gain in tumor and less than 0 copy number loss. (2) The nucleotide composition might be altered at germline polymorphic loci (including SNPs and indels) within sCNV regions. At loci with two alternative alleles, the reference allele is referred as “A” allele and non-reference allele as “B” allele. In analogy to CNV data generated from Illumina SNP array, BAF was defined as (number of reads supporting B allele)/(total number of reads at this locus). In normal tissue, at a locus with heterozygous genotype, the number of reads supporting the reference allele and non-reference allele are roughly equal; therefore, BAF would fluctuate closely around 0.5. If the region, where this locus is located, undergoes copy number loss or gain in tumor, the tumor BAF would deviate from 0.5. A simple transformation, defined as mBAF =|BAF - 0.5| + 0.5 as in CNV analysis of Illumina array data, can dissolve the arbitrary direction of such deviation with respect to haplotype, so that segmentation algorithm can be readily applicable to this signal[11]. To make noise distribution more adhere to Gaussian, log2 of the ratio of mBAF in tumor vs. normal, denoted as log2mBAF was calculated.

Both log2ratio and log2mBAF have spatial correlation with copy number change along chromosomes and existing CNV segmentation methods (e.g. circular binary segmentation,

CBS[12]) can be readily applied. It should be noted that log2ratio and log2mBAF reflect distinct aspects of CNV and provide orthogonal information in CNV calling. To our experience, log2mBAF is more sensitive than log2ratio in detecting unbalanced copy number gain or loss. However, log2mBAF cannot directly infer the magnitude of copy number change and is not informative in balanced copy number events.

Standard GATK pipeline was applied for raw data processing, read alignment and variants calling. On each observed SNP/indel locus, the total number of reads and number of reads supporting the non-reference allele in normal and tumor tissues respectively was quantified, and used to calculate the log2ratio and log2mBAF metrics. CBS algorithm (implemented in R package "DNAcopy"[13]) was applied to perform segmentation separately on log2ratio and log2mBAF spaces. Afterwards, the union of change points called from log2ratio and log2mBAF were taken, and summarized the segment boundaries. The copy number status is decided for each segment based on the evidences from both log2ratio and log2mBAF.

Whole-exome sequencing somatic mutation detection

Remaining reads from the quality control (as detailed in the paragraph above: **Whole-exome sequencing chromosomal rearrangements analysis**) were aligned to the human genome reference 19th version using BWA[16]. Afterwards, somatic SNV were detected using Mutect software[17] and small indels were identified through VarScan2[18]. We calculated p-value using Fisher's exact test for all putative mutation sites based on the distribution of read support for different alleles in tumor and matched normal samples. The Mutect and VarScan2 software were employed in above analyses because of their desirable feature in detecting mutations in low purity or heterogeneous cancer samples. Purity is defined as the percentage of tumoral cells in the tumor, and of hepatocytes in the normal tissue versus fibroblasts, endothelial and immunological cells, and portal veins in the non-tumoral tissues. FLC tumors presented relatively low purity in both normal and tumor tissues. Among the normal tissues, the purity ranges between 80% and 95%; and 35% to 70% among tumor tissues. Purity information has been factored in mutant allele frequency estimations.

Somatic SNVs were further filtered by Mutect internal threshold, and those SNVs that failed the threshold were excluded. Somatic indels inferred by VarScan2 software were filtered following the criteria: 1) Read depth ≥ 20 in both tumor and normal samples, 2) Read support of mutant allele in tumor tissue not as a result of sequencing error (Binomial test, $p > 0.01$), 3) Quality score not significantly lower than other alleles (Wilcoxon rank sum test, $p > 0.01$), 4) Mutant allele frequency change between tumor and adjacent normal ≥ 20% and Fisher's Exact

Test p-value < 0.01, 5) Mutant allele not significantly enriched in repeatedly aligned reads, 6) Mutant allele not significantly enriched within 10 bps of 5' or 3' ends of reads (Fisher's exact test, p > 0.01) and 7) Mutant allele were observed in both forward and reverse strand of the tumor DNA. Lastly, the resulting mutation were annotated by SNPEff pipeline[19] and scored by SIFT and PolyPhen2, in terms of mutation location, impact on gene product (coding or non-coding, as well as if synonymous or non-synonymous, and missense or non-sense) and the likelihood of the mutation to be functional (damaging, probably damaging, possibly damaging or benign).

BRCA2 damaging mutation (Y2789C) was validated by Sanger sequencing once PCR amplifications on human tissues were performed using the primers: Fw_BRCA2 5'-GGCAGTTCTAGAAGAATGAAAACTC-3' and Rv_BRCA2 5'- CGATGATAAGGGCAGAGGAA-3', and following the PCR amplification protocol: 95°C for 2 minutes, 40 cycles of denaturation at 95°C for 30 seconds, 56°C annealing for 30 seconds, 72°C extension for 45 seconds, followed by a 2 minute final extension at 72°C. The PCR amplifications were performed in a volume of 35µL reaction mixture containing 1.5mM MgCl2, 0.2mM of each dNTP, 0.125mM of each primer and 1U of Platinum Taq DNA Polymerase (Invitrogen).

Targeted-exome sequencing somatic mutation detection

Forty-seven FLCs, 20 FF from the validation-French cohort and 27 FFPE with 14 paired non-tumoral samples from the training cohort, underwent targeted-exome sequencing (TES) by GeneRead DNAseq Targeted Panels V2 (Qiagen) following manufacturer's instructions. TES was applied in the whole *BRCA2* gene, and in 13 exons including the hotspots for *CSMD2, ARMCX1, COL6A6, TERT* promoter, *TP53*, *CTNNB1*, *NFE2L2*, *EGFR*, *BRAF*, *KRAS*, *NRAS*, *IDH1* and *IDH2* genes. However, the sequencing reaction was not successful for *TERT* promoter, *TP53* and *IDH2*. Then, we got good quality reads for *CSMD2, ARMCX1, COL6A6,* from which we were able to validate the mutations found by WES, and also good quality data for *CTNNB1*, *NFE2L2*, *EGFR*, *BRAF*, *KRAS*, *NRAS*, and *IDH1* genes. However, those genes were not damagingly mutated in any of their hotspots (*EGFR*-L858R, *CTNNB1*-exon 3, *BRAF*-V600X, *NFE2L2*-exon 2, *KRAS*-G12, *NRAS*-Q61, *IDH1*-R132) in any of the 47 FLCs. Those results correlated with the WES data, which also demonstrated no damaging mutations in any of the hotspots for those genes, neither for *TP53*-R249S*, IDH2*-R172, *ARID1A, TTN* or *AXIN1*.

Results were filtered by quality control (Q>20) that was performed per base and read, removing and trimming reads with low quality[20]. Reads were aligned to GRCh37/hg19 with

Burrows-Wheeler Aligner algorithm, paired-end mode, and duplicated reads were also removed[16]. Realignment and variant calling was performed with Genome Analysis Toolkit[2]. Variants associated with background polymorphisms (dbSNP b138) have been removed. Genotype quality scores (GQ>30, Read Depth > 30, and ratio AlternativeReads/ReadDepth > 20%) have been used to filter false positive gene mutations events. Variants have also been filtered according to their predicted function (only non-synonymous coding or stop-gaining mutations were considered), assessed by SnpEff pipeline[19].

Prognostic gene expression data analysis

The 8-gene prognostic signature was obtained using the random survival forest technique (RSF)[21] in those FLC patients that have been resected, excluding those transplanted. This is a method for prediction and variable selection using right-censored survival and competing risk data by growing survival trees to estimate a cumulative hazard function (CHF), which derives from each tree of the RSF. As input, we used the 15,169 Illumina probes which passed the quality filters[1,3]. First, probes were randomly split in 8 sets to fit a model using the RSF method by growing 10,000 survival trees. Variable importance scores (VIMP) were computed for all probes used to grow the trees, and those with VIMP > 0.01 were selected to fit a new model using the same RSF method. Once the model was built, redundant variables were removed by generating incremental RSF models adding one more probe at each step, ranked by the VIMP calculated in the previous model. This allowed us to get an estimation of the error as a measure of how correctly the ensemble classified two random individuals in terms of survival. As an extra step for removing the redundancy and reduce the signature to the minimum number of probes, a new model was constructed with those probes able to improve the prediction error better than the expected by adding a random probe. Therefore, the expression signature was based upon 8 non-redundant genes.

From the last RSF model, we generated a risk score (i.e. mortality index (MI), range 6.7-42) based on the 8-gene signature for each individual, computed as a sum over the CHF for each individual evaluated at distinct time points weighed by the number of individuals at risk at that moment. Using the reported MI, our cohort was divided into 2 risk groups, the upper quartile (Poor-prognosis, MI ≥ 21) and the 3 remaining quartiles (Non poor-prognosis, MI < 21). We were able to show that the risk obtained from the signature, correlated with the survival through the Kaplan-Meier estimator. This signature was assessed to calculate the MI related to

recurrence in the training cohort, being able to stratify those patients based on the upper MI quartile as threshold. The prognostic 8-gene signature was used to calculate the MI in an independent cohort to validate its capacity to predict survival. The expression values of the validation French-cohort samples were obtained from the hybridization of the mRNA extracted as previously published[15] in the HumanHT-12_V4.0 whole genome expression array from Illumina.

The integration of the two datasets (i.e. training and validation) was performed using the virtualArray[22] package from Bioconductor[23], which merges independent sets of samples by gene name and normalizes for batch effect by using surrogate variable analysis[24]. All analyses were performed using the R statistical package (www.r-project.org).

Clinical data and Statistical Analysis

Comparisons of continuous variables between 2 classes were performed with the Mann–Whitney U test. Spearman coefficient was used to assess correlations between continuous variables. Fisher exact test was used to compare proportions. Bonferroni correction for multiple hypothesis testing was applied when appropriate. Kaplan-Meier method and log-rank test were performed to analyze the association of molecular and clinical variables with overall survival and tumor recurrence.

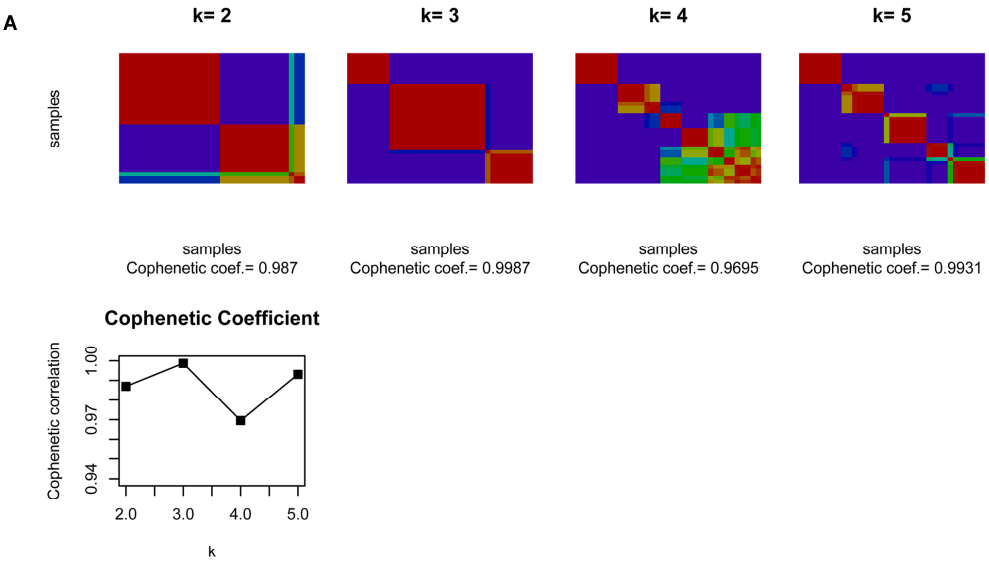
Supplementary REFERENCES

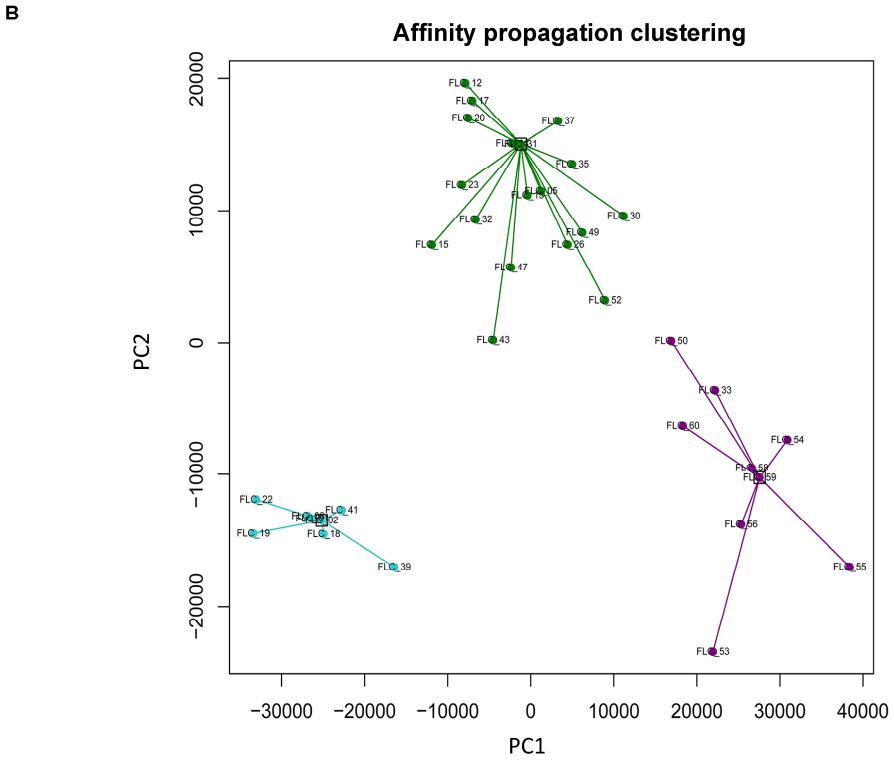
1. Hoshida Y, Villanueva A, Kobayashi M, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. N Engl J Med 2008;359:1995-2004.
2. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011;43:491-8.
3. Reich M, Liefeld T, Gould J, et al. GenePattern 2.0. Nat Genet 2006;38:500-1.
4. Brunet JP, Tamayo P, Golub TR, et al. Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci U S A 2004;101:4164-9.
5. Frey BJ, Dueck D. Clustering by passing messages between data points. Science 2007;315:972-6.
6. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545-50.
7. Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures database (MSigDB) 3.0. Bioinformatics 2011;27:1739-40.
8. Hoshida Y, Toffanin S, Lachenmayer A, et al. Molecular classification and novel targets in hepatocellular carcinoma: recent advancements. Semin Liver Dis 2010;30:35-51.

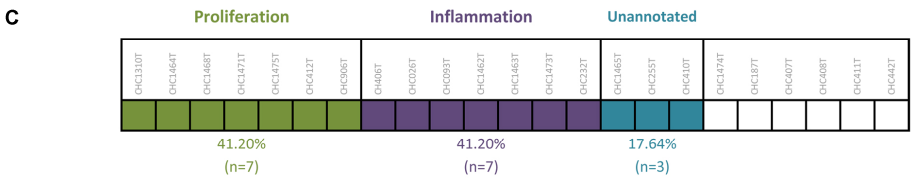
9. Mermel CH, Schumacher SE, Hill B, et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol 2011;12:R41.
10. Peiffer DA, Le JM, Steemers FJ, et al. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Res 2006;16:1136-48.
11. Staaf J, Lindgren D, Vallon-Christersson J, et al. Segmentation-based detection of allelic imbalance and loss-of-heterozygosity in cancer cells using whole genome SNP arrays. Genome Biol 2008;9:R136.
12. Olshen AB, Venkatraman ES, Lucito R, et al. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 2004;5:557-72.
13. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. Bioinformatics 2007;23:657-63.
14. Benjamini YH, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B, 1995;57:289–300.
15. Wurmbach E, Chen YB, Khitrov G, et al. Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma. Hepatology 2007;45:938-47.
16. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-60.
17. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol;31:213-9.
18. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 2012;22:568-76.
19. Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012;6:80-92.
20. http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
21. Ishwaran H KU, Chen X, Minn AJ. 2011; 4: 115–32. Random survival forests for high-dimensional data. Statistical Analy Data Mining 2011;4 115–32.
22. Heider A, Alt R. virtualArray: a R/bioconductor package to merge raw data from different microarray platforms. BMC Bioinformatics 2013;14:75.
23. Gentleman RC CV, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004;5:R80.
24. Leek JT, Johnson WE, Parker HS, et al. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 2012;28:882-3.

Supplementary TABLES

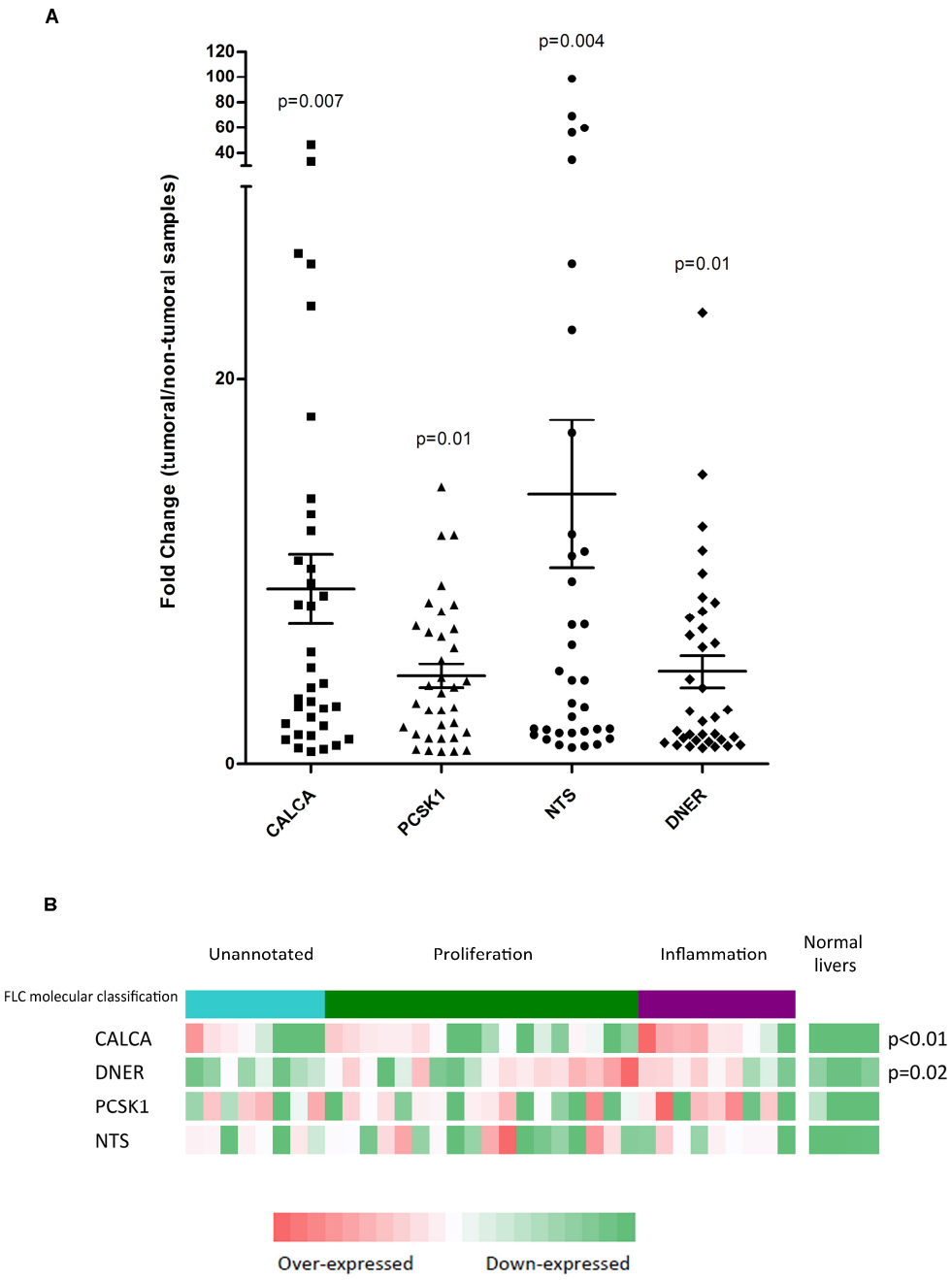
Suppl. Figure 1



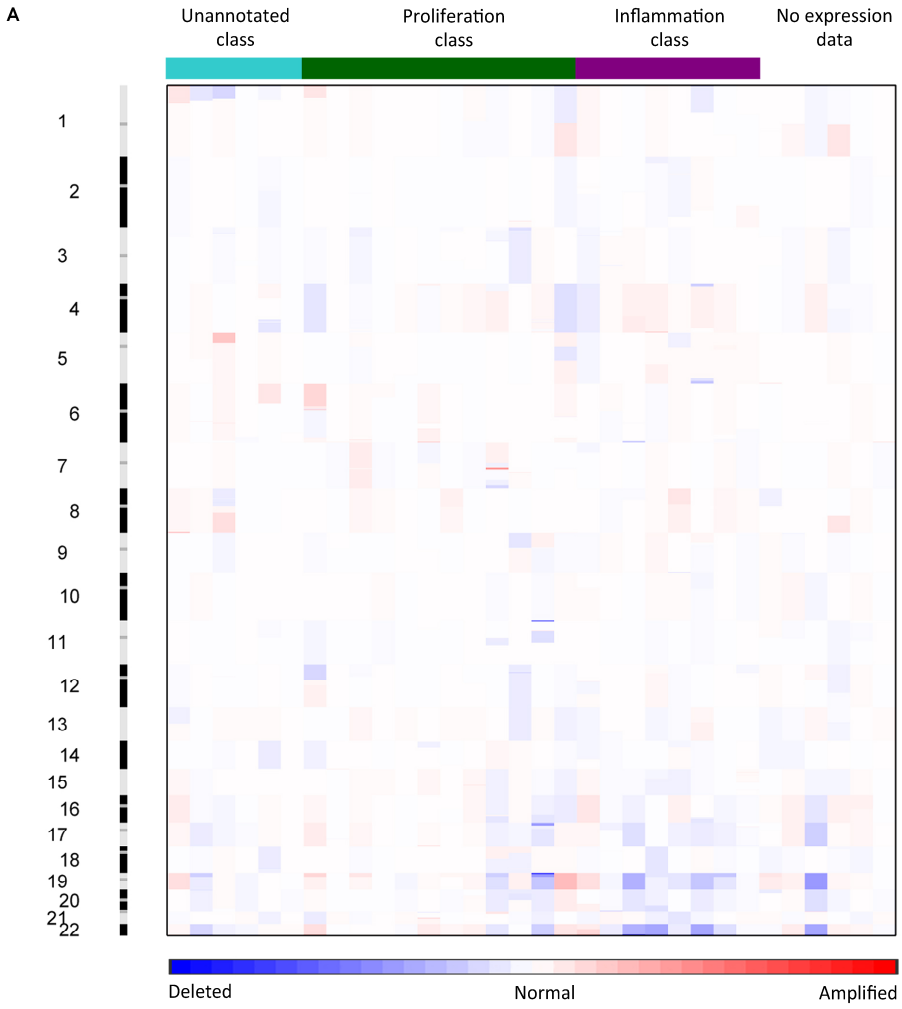


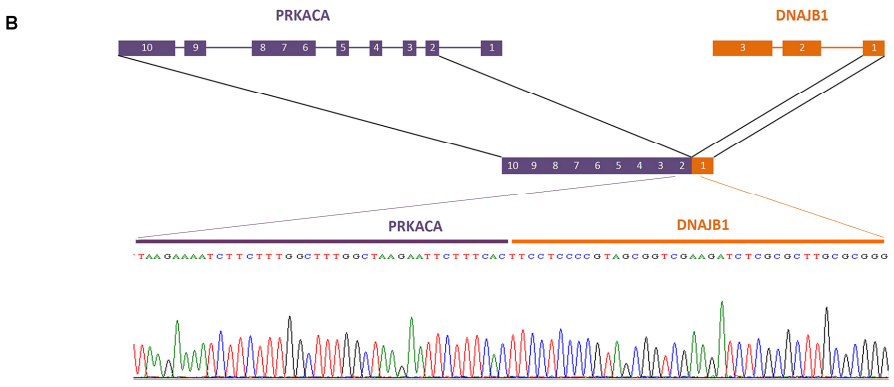


Suppl. Figure 2



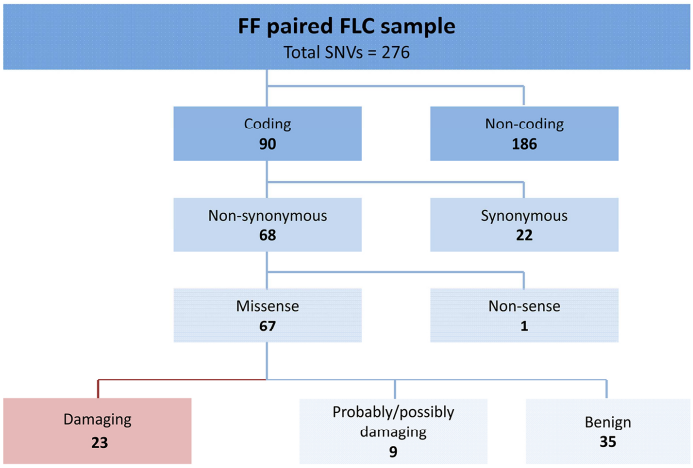
Suppl. Figure 3



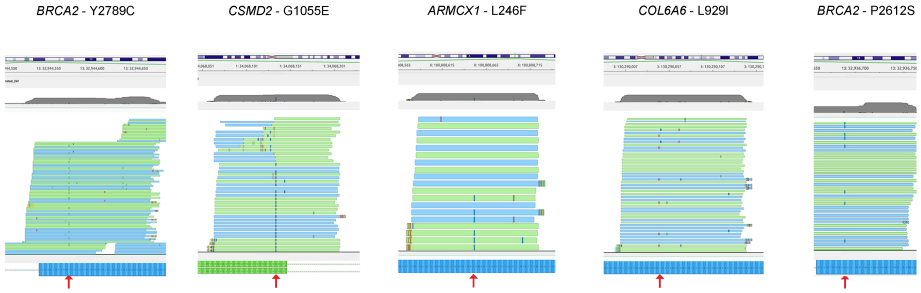


Suppl. Figure 4

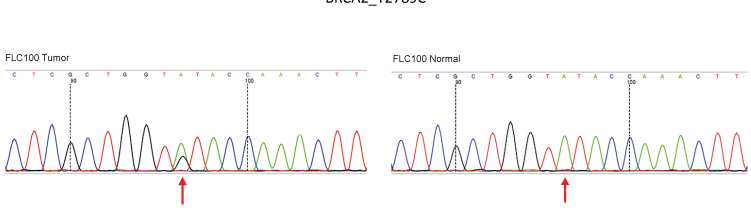
A



B



C



Supplementary Table 11. Mortality Index for the different prognostic group samples

Training Set - Survival

Sample ID	Survival status	Mortality Index	Prognosis group
FLC56	1	42.06	Poor-prognosis
FLC59	1	38.22	Poor-prognosis
FLC17	1	33.05	Poor-prognosis
FLC60	1	31.10	Poor-prognosis
FLC55	1	29.37	Poor-prognosis
FLC31	1	23.46	Poor-prognosis
FLC32	1	21.08	Poor-prognosis
FLC22	1	20.68	Non poor-prognosis
FLC47	1	20.23	Non poor-prognosis
FLC14	1	18.37	Non poor-prognosis
FLC50	1	17.36	Non poor-prognosis
FLC37	0	14.56	Non poor-prognosis
FLC12	0	13.67	Non poor-prognosis
FLC23	0	12.22	Non poor-prognosis
FLC53	1	11.76	Non poor-prognosis
FLC58	0	11.51	Non poor-prognosis
FLC35	0	10.89	Non poor-prognosis
FLC19	0	10.74	Non poor-prognosis
FLC41	1	10.19	Non poor-prognosis
FLC33	0	10.09	Non poor-prognosis
FLC13	1	8.71	Non poor-prognosis
FLC54	0	8.21	Non poor-prognosis
FLC05	0	7.28	Non poor-prognosis
FLC08	0	6.95	Non poor-prognosis
FLC43	0	6.87	Non poor-prognosis
FLC49	0	6.78	Non poor-prognosis
FLC52	0	6.74	Non poor-prognosis
FLC39	0	6.73	Non poor-prognosis
FLC01	0	6.71	Non poor-prognosis

Training Set - Recurrence

Sample ID	Survival status	Mortality Index	Prognosis group
FLC_56	1	36.72	Poor-prognosis
FLC_17	1	35.06	Poor-prognosis
FLC_32	1	31.32	Poor-prognosis
FLC_47	1	28.71	Poor-prognosis
FLC_55	1	27.87	Poor-prognosis
FLC_23	1	25.45	Poor-prognosis
FLC_14	1	24.08	Non poor-prognosis
FLC_31	1	23.82	Non poor-prognosis
FLC_22	1	23.10	Non poor-prognosis
FLC_50	1	22.17	Non poor-prognosis
FLC_13	1	21.85	Non poor-prognosis
FLC_05	1	20.39	Non poor-prognosis

FLC_12	0	19.68	Non poor-prognosis
FLC_43	1	19.03	Non poor-prognosis
FLC_41	1	18.02	Non poor-prognosis
FLC_35	0	16.58	Non poor-prognosis
FLC_33	1	15.55	Non poor-prognosis
FLC_19	1	14.49	Non poor-prognosis
FLC_53	1	14.11	Non poor-prognosis
FLC_49	1	13.93	Non poor-prognosis
FLC_58	0	13.43	Non poor-prognosis
FLC_54	0	11.10	Non poor-prognosis
FLC_52	0	10.94	Non poor-prognosis
FLC_01	0	10.91	Non poor-prognosis
FLC_39	1	10.10	Non poor-prognosis
FLC_08	0	9.06	Non poor-prognosis

Validation Set - Survival

Sample ID	Survival status	Mortality Index	Prognosis group
CHC412T	1	8.49	Poor-prognosis
CHC187T	1	8.44	Poor-prognosis
CHC411T	1	8.32	Poor-prognosis
CHC906T	1	8.31	Poor-prognosis
CHC255T	1	8.23	Poor-prognosis
CHC1471T	1	8.00	Poor-prognosis
CHC026T	1	7.93	Non poor-prognosis
CHC1468T	0	7.90	Non poor-prognosis
CHC1463T	0	7.89	Non poor-prognosis
CHC442T	0	7.83	Non poor-prognosis
CHC408T	0	7.69	Non poor-prognosis
CHC093T	1	7.65	Non poor-prognosis
CHC1464T	0	7.50	Non poor-prognosis
CHC1465T	0	7.38	Non poor-prognosis
CHC1474T	1	7.30	Non poor-prognosis
CHC1310T	0	7.26	Non poor-prognosis
CHC232T	0	7.03	Non poor-prognosis
CHC410T	0	6.69	Non poor-prognosis
CHC407T	0	6.67	Non poor-prognosis
CHC406T	0	6.26	Non poor-prognosis
CHC1462T	0	5.86	Non poor-prognosis
CHC1473T	0	5.38	Non poor-prognosis

Supplementary Table 2. Leave-one-out cross validation of FLC subclasses

Prediction Method	# Error	% Correct Prediction
k-nearest neighbor	2	94%

Supplementary Table 3. FLC gene signature

Gene symbol	Molecular class	tstat
RAB40C	Unannotated	14.237
GCN1L1	Unannotated	13.242
DAP3	Unannotated	13.072
BRWD2	Unannotated	11.640
MRPL36	Unannotated	11.616
NUP188	Unannotated	11.392
ASXL1	Unannotated	10.963
DPAGT1	Unannotated	10.959
ZNF16	Unannotated	10.390
NSDHL	Unannotated	10.193
ASAP3	Unannotated	10.089
NEURL1B	Unannotated	9.845
KCNJ5	Unannotated	9.759
USP16	Unannotated	9.682
C6ORF47	Unannotated	9.659
ATP6V0E1	Unannotated	9.470
NCOA4	Unannotated	9.427
UBL4A	Unannotated	9.335
BUB3	Unannotated	9.295
UBR7	Unannotated	9.267
LASP1	Unannotated	9.220
TUG1	Unannotated	9.182
ZC3HC1	Unannotated	9.169
IRF3	Unannotated	9.113
INTS4	Unannotated	9.054
UROS	Unannotated	9.040
ZNHIT1	Unannotated	8.977
YTHDC2	Unannotated	8.975
TMEM177	Unannotated	8.880
EIF3K	Unannotated	8.800
SF3B3	Unannotated	8.791
C20ORF20	Unannotated	8.731
KIAA1310	Unannotated	8.653
BRMS1	Unannotated	8.628
PRPF8	Unannotated	8.604
METTL13	Unannotated	8.548
FAM105B	Unannotated	8.526
RAB1A	Unannotated	8.444
RPAIN	Unannotated	8.438
PSMC6	Unannotated	8.314
TBC1D3H	Unannotated	8.300
PRRC1	Unannotated	8.296
PRKCDBP	Unannotated	8.242
TRIM39	Unannotated	8.199
CLUAP1	Unannotated	8.167
ATG4B	Unannotated	8.134
IPO13	Unannotated	8.105

KARS	Unannotated	8.060
RP9	Unannotated	8.001
UBN1	Unannotated	7.997
TSPYL1	Unannotated	7.991
MPHOSPH10	Unannotated	7.966
SETD1A	Unannotated	7.850
SLC35B1	Unannotated	7.816
C19ORF22	Unannotated	7.803
INTS9	Unannotated	7.785
RIT1	Unannotated	7.766
ERGIC3	Unannotated	7.765
ZNF586	Unannotated	7.738
GHITM	Unannotated	7.714
RANBP1	Unannotated	7.669
CNOT6	Unannotated	7.661
LPHN1	Unannotated	7.644
VPS45	Unannotated	7.638
ROD1	Unannotated	7.637
MAPK6	Unannotated	7.598
ZDHHC7	Unannotated	7.591
LYST	Unannotated	7.562
UBB	Unannotated	7.514
CUGBP1	Unannotated	7.503
ROBLD3	Unannotated	7.463
ALDH9A1	Unannotated	7.403
WRNIP1	Unannotated	7.381
IFITM3	Unannotated	7.376
MCM3AP	Unannotated	7.339
AP2S1	Unannotated	7.316
ATP5SL	Unannotated	7.314
TMEM170B	Unannotated	7.314
SAP30L	Unannotated	7.313
C20ORF96	Unannotated	7.312
FEN1	Unannotated	7.309
UBAP2	Unannotated	7.302
SNRNP48	Unannotated	7.297
RNASEK	Unannotated	7.266
DNAJB11	Unannotated	7.225
GRINA	Unannotated	7.201
L3MBTL2	Unannotated	7.198
TSFM	Unannotated	7.179
HRB	Unannotated	7.149
VCL	Unannotated	7.139
MIA	Unannotated	7.115
ABCA1	Unannotated	7.109
IMPDH2	Unannotated	7.103
TOR1A	Unannotated	7.102
ZNF8	Unannotated	7.093
RBP4	Unannotated	7.085
GOLGA5	Unannotated	7.084

DDHD2	Unannotated	7.082
CCDC21	Unannotated	7.077
PHB2	Unannotated	7.071
URM1	Unannotated	7.067
SERPINB9	Unannotated	7.055
FLJ45966	Unannotated	7.050
ASCC1	Unannotated	7.036
SNX14	Unannotated	7.025
TRIM11	Unannotated	7.019
PI4KA	Unannotated	7.018
VPS11	Unannotated	6.979
FAM82A2	Unannotated	6.939
CEP170	Unannotated	6.897
MMS19	Unannotated	6.883
ESD	Unannotated	6.878
XPOT	Unannotated	6.864
TULP4	Unannotated	6.855
AQR	Unannotated	6.854
LTBR	Unannotated	6.843
ZZZ3	Unannotated	6.826
TSTD2	Unannotated	6.826
NME4	Unannotated	6.813
SP140L	Unannotated	6.773
NOTCH3	Unannotated	6.763
ARID4B	Unannotated	6.759
PRAGMIN	Unannotated	6.756
RHOC	Unannotated	6.750
CSTF3	Unannotated	6.744
LRCH3	Unannotated	6.743
NCOA6	Unannotated	6.741
IRF9	Unannotated	6.726
GTF2I	Unannotated	6.708
HARS2	Unannotated	6.689
RBM33	Unannotated	6.650
PDPN	Unannotated	6.648
RXRB	Unannotated	6.644
PRKRIR	Unannotated	6.637
DKC1	Unannotated	6.634
RPUSD3	Unannotated	6.627
CMAS	Unannotated	6.622
EIF3G	Unannotated	6.616
FXC1	Unannotated	6.597
CALR	Unannotated	6.596
YY1AP1	Unannotated	6.564
PIP4K2A	Unannotated	6.562
PDCD2	Unannotated	6.550
MRFAP1L1	Unannotated	6.546
SLC6A16	Unannotated	6.496
TOMM34	Unannotated	6.489
FCHSD2	Unannotated	6.470

C19ORF2	Unannotated	6.469
MRPL17	Unannotated	6.458
DOLK	Unannotated	6.457
ETF1	Unannotated	6.452
C7ORF55	Unannotated	6.451
ILK	Unannotated	6.446
RECQL	Unannotated	6.444
ZFYVE16	Unannotated	6.443
UBE3A	Unannotated	6.425
ZNF234	Unannotated	6.420
LYAR	Unannotated	6.415
SKP1A	Unannotated	6.410
SLC35E3	Unannotated	6.402
CSK	Unannotated	6.352
CYBASC3	Unannotated	6.349
SMAD2	Unannotated	6.328
IQCK	Unannotated	6.323
NUDCD3	Unannotated	6.317
DCAKD	Unannotated	6.316
ZSWIM1	Unannotated	6.316
ZNHIT6	Unannotated	6.315
CCS	Unannotated	6.308
PIK3R2	Unannotated	6.301
MORF4L2	Unannotated	6.288
KCNC3	Unannotated	6.278
SLC35D2	Unannotated	6.245
CHMP4B	Unannotated	6.228
NECAP2	Unannotated	6.223
AGAP8	Unannotated	6.217
DCTN2	Unannotated	6.186
FAM171A1	Unannotated	6.184
VEZT	Unannotated	6.184
UTP14C	Unannotated	6.183
DPH3	Unannotated	6.179
ERH	Unannotated	6.173
SNRPA	Unannotated	6.171
TSC2	Unannotated	6.168
CDC42BPB	Unannotated	6.163
COPZ1	Unannotated	6.157
MESDC2	Unannotated	6.153
SFT2D1	Unannotated	6.152
SLFN12	Unannotated	6.142
RFP	Unannotated	6.137
QKI	Unannotated	6.134
SP100	Unannotated	6.133
HEATR6	Unannotated	6.119
USP7	Unannotated	6.118
CPT1C	Unannotated	6.112
ARMCX6	Unannotated	6.106
ATG9A	Unannotated	6.094

CCNH	Unannotated	6.090
NFIX	Unannotated	6.086
C8ORF33	Unannotated	6.074
ACTG1	Unannotated	6.065
SFRS6	Unannotated	6.045
FAM122A	Unannotated	6.043
HDAC3	Unannotated	6.030
GNB1	Unannotated	6.022
C2ORF56	Unannotated	6.021
CYB561D1	Unannotated	6.020
SLC25A3	Unannotated	6.019
CCDC53	Unannotated	6.017
GPATCH8	Unannotated	6.000
C19ORF53	Unannotated	5.998
C17ORF95	Unannotated	5.993
PNPO	Unannotated	5.990
FBXL10	Unannotated	5.987
ZCWPW1	Unannotated	5.976
C1QBP	Unannotated	5.974
SH3GL1	Unannotated	5.969
DBNL	Unannotated	5.968
PSMC1	Unannotated	5.965
ZNF860	Unannotated	5.960
ZNF613	Unannotated	5.946
DYNLT1	Unannotated	5.943
UBE2V2	Unannotated	5.921
OTUD4	Unannotated	5.920
SEC22C	Unannotated	5.914
RPUSD2	Unannotated	5.914
ZNF395	Unannotated	5.907
GIPC1	Unannotated	5.904
NDFIP1	Proliferation	9.920
TMEM205	Proliferation	9.777
NANP	Proliferation	9.346
PPP6C	Proliferation	9.329
RNF181	Proliferation	9.254
KLHDC3	Proliferation	9.109
NPC2	Proliferation	9.005
N6AMT2	Proliferation	8.988
REV1	Proliferation	8.767
FNTA	Proliferation	8.731
SEPT9	Proliferation	8.715
SIN3A	Proliferation	8.649
BRAF	Proliferation	8.643
WAC	Proliferation	8.642
YIF1A	Proliferation	8.498
MRPL18	Proliferation	8.322
PPP2CA	Proliferation	8.303
CCT8	Proliferation	8.190
FOXN2	Proliferation	8.136

ABHD5	Proliferation	8.076
SEC31A	Proliferation	8.026
DNAJC8	Proliferation	8.000
BTAF1	Proliferation	7.996
SIKE	Proliferation	7.895
MED6	Proliferation	7.867
GTF2H2	Proliferation	7.730
KLF3	Proliferation	7.717
GAK	Proliferation	7.589
DARS	Proliferation	7.569
EIF3L	Proliferation	7.502
MRFAP1	Proliferation	7.473
P76	Proliferation	7.440
LAMA5	Proliferation	7.419
SNORD71	Proliferation	7.386
ATL2	Proliferation	7.368
RBM18	Proliferation	7.360
HSPA4	Proliferation	7.356
BRP44L	Proliferation	7.351
SRPK1	Proliferation	7.342
MITD1	Proliferation	7.326
FAM125A	Proliferation	7.309
RPUSD4	Proliferation	7.230
NDUFA4	Proliferation	7.199
NCKAP1	Proliferation	7.190
ZFAND6	Proliferation	7.186
CTGLF3	Proliferation	7.180
SAPS3	Proliferation	7.176
EIF3F	Proliferation	7.115
LOC644907	Proliferation	7.112
HP1BP3	Proliferation	7.078
EEA1	Proliferation	7.071
MAP1LC3A	Proliferation	7.023
MRPS31	Proliferation	7.020
PIGY	Proliferation	7.008
GNG2	Proliferation	6.999
TPST2	Proliferation	6.977
ECD	Proliferation	6.960
ELF1	Proliferation	6.958
H1F0	Proliferation	6.941
BAMBI	Proliferation	6.934
OSBP2	Proliferation	6.931
UTX	Proliferation	6.931
ANKRD11	Proliferation	6.909
RANGAP1	Proliferation	6.901
SIPA1	Proliferation	6.896
TMED2	Proliferation	6.883
PIN1	Proliferation	6.878
GNG12	Proliferation	6.859
EIF4G2	Proliferation	6.814

UBE2D2	Proliferation	6.786
UTP6	Proliferation	6.768
SEPHS1	Proliferation	6.768
RUFY2	Proliferation	6.760
AASDH	Proliferation	6.749
CHD2	Proliferation	6.716
C12ORF41	Proliferation	6.695
PAWR	Proliferation	6.636
U2AF2	Proliferation	6.627
HDAC1	Proliferation	6.613
ANAPC1	Proliferation	6.604
MTMR6	Proliferation	6.586
CSNK1E	Proliferation	6.586
COX4I1	Proliferation	6.568
AP1S2	Proliferation	6.560
MYH9	Proliferation	6.549
FARP1	Proliferation	6.546
KIAA0776	Proliferation	6.544
RNF160	Proliferation	6.533
RPL26	Proliferation	6.504
SNRPN	Proliferation	6.490
NSUN2	Proliferation	6.479
NDUFB10	Proliferation	6.470
CHCHD4	Proliferation	6.468
NENF	Proliferation	6.457
EMD	Proliferation	6.448
CAPZA2	Proliferation	6.448
NDUFB8	Proliferation	6.446
NUDT3	Proliferation	6.424
C11ORF58	Proliferation	6.395
FAM53C	Proliferation	6.385
PPM1M	Proliferation	6.381
DMTF1	Proliferation	6.380
RNF170	Proliferation	6.379
CASP3	Proliferation	6.374
HERC4	Proliferation	6.330
BMPR2	Proliferation	6.313
GFPT1	Proliferation	6.298
CCDC90B	Proliferation	6.277
SFRS18	Proliferation	6.268
PDCD6IP	Proliferation	6.261
LOC374395	Proliferation	6.259
SNAP29	Proliferation	6.222
EIF4E2	Proliferation	6.200
C1ORF52	Proliferation	6.197
SPRYD3	Proliferation	6.195
CDKN2AIP	Proliferation	6.166
HADHA	Proliferation	6.161
MED13	Proliferation	6.156
SPOP	Proliferation	6.153

ACAP2	Proliferation	6.150
SLC30A7	Proliferation	6.148
TAOK3	Proliferation	6.137
VBP1	Proliferation	6.119
TMEM218	Proliferation	6.101
ZNF704	Proliferation	6.094
UBA1	Proliferation	6.082
C12ORF43	Proliferation	6.065
KIF1B	Proliferation	6.064
C1ORF50	Proliferation	6.064
PUM2	Proliferation	6.059
EIF1AX	Proliferation	6.046
SDHB	Proliferation	6.044
EIF3A	Proliferation	6.039
KRAS	Proliferation	6.031
PTPRK	Proliferation	6.021
SERPINB6	Proliferation	6.018
CCDC88A	Proliferation	5.986
RAD21	Proliferation	5.982
JMJD1C	Proliferation	5.977
DRAP1	Proliferation	5.976
MTRF1L	Proliferation	5.974
ST13	Proliferation	5.968
TTC39C	Proliferation	5.956
TIMM23	Proliferation	5.952
MRPS7	Proliferation	5.941
MSRB2	Proliferation	5.934
MTCH1	Proliferation	5.925
PRKAR1B	Proliferation	5.923
CD58	Proliferation	5.923
CMIP	Proliferation	5.921
DNTTIP2	Proliferation	5.920
RBM14	Proliferation	5.875
CFL1	Proliferation	5.872
MED20	Proliferation	5.870
RPL15	Proliferation	5.865
TSPAN12	Proliferation	5.859
PHACTR2	Proliferation	5.853
C17ORF70	Proliferation	5.843
GOSR2	Proliferation	5.833
PITPNB	Proliferation	5.827
CTDSP2	Proliferation	5.827
ANKHD1	Proliferation	5.827
C17ORF61	Proliferation	5.816
NFAT5	Proliferation	5.812
PTK2	Proliferation	5.808
TLN1	Proliferation	5.803
HSZFP36	Proliferation	5.787
TERF2IP	Proliferation	5.767
TMEM14A	Proliferation	5.762

UBQLN4	Proliferation	5.749
CISD2	Proliferation	5.745
NIN	Proliferation	5.728
C14ORF32	Proliferation	5.721
ZNF226	Proliferation	5.718
LRP3	Proliferation	5.714
CNTNAP1	Proliferation	5.703
KIAA1539	Proliferation	5.699
STAT6	Proliferation	5.689
RPS6	Proliferation	5.687
SNORA8	Proliferation	5.674
TFAM	Proliferation	5.666
FUBP3	Proliferation	5.666
HNRPUL1	Proliferation	5.657
BTN2A1	Proliferation	5.657
FAF1	Proliferation	5.653
FBXO22	Proliferation	5.645
BAT2D1	Proliferation	5.643
PBRM1	Proliferation	5.643
C2ORF79	Proliferation	5.642
MRPS6	Proliferation	5.641
HES1	Proliferation	5.637
SSU72	Proliferation	5.620
MUTED	Proliferation	5.613
DTX2	Proliferation	5.613
FAM18B	Proliferation	5.610
RPL23	Proliferation	5.605
GOLGA8B	Proliferation	5.592
STAT2	Proliferation	5.587
CCND3	Proliferation	5.584
CYB5R3	Proliferation	5.579
LSM1	Proliferation	5.579
HADHB	Proliferation	5.574
C10ORF116	Proliferation	5.574
MRS2	Proliferation	5.559
HEBP2	Proliferation	5.552
NDUFV3	Proliferation	5.551
RBX1	Proliferation	5.544
HIST1H2BC	Proliferation	5.544
ID2	Proliferation	5.543
C10ORF78	Proliferation	5.529
CCDC58	Proliferation	5.523
MASP1	Proliferation	5.514
UBE2E1	Proliferation	5.512
ACACB	Proliferation	5.509
RQCD1	Proliferation	5.506
TSR1	Proliferation	5.503
ZNF35	Proliferation	5.497
STX8	Proliferation	5.474
HBXIP	Proliferation	5.464

ZNF669	Inflammation	13.075
PTPLAD2	Inflammation	12.636
ORC6L	Inflammation	12.551
LILRB3	Inflammation	12.460
NAG18	Inflammation	11.844
MEX3D	Inflammation	11.473
ZNF674	Inflammation	11.405
AIRE	Inflammation	10.945
RAD51	Inflammation	10.905
MCM8	Inflammation	10.876
RNU1A3	Inflammation	10.854
ZNF223	Inflammation	10.628
LAIR1	Inflammation	10.442
ABCB8	Inflammation	10.440
SLC4A5	Inflammation	10.319
TUBA1A	Inflammation	10.277
LOC100190938	Inflammation	10.212
FAM119A	Inflammation	10.191
HFE	Inflammation	10.131
CYCSL1	Inflammation	10.107
DAPP1	Inflammation	10.062
C14ORF85	Inflammation	9.995
ZNF773	Inflammation	9.872
SFRS13A	Inflammation	9.863
PCDHB9	Inflammation	9.847
NUP93	Inflammation	9.809
FAM175A	Inflammation	9.791
MOV10	Inflammation	9.693
ZNF652	Inflammation	9.640
PIGX	Inflammation	9.584
KLHL28	Inflammation	9.572
USP49	Inflammation	9.543
EID2B	Inflammation	9.502
ZNF430	Inflammation	9.471
DENR	Inflammation	9.395
PPA2	Inflammation	9.393
ZNF273	Inflammation	9.312
MGC16703	Inflammation	9.296
MFSD6L	Inflammation	9.267
SLC5A8	Inflammation	9.245
PLA2G2D	Inflammation	9.238
ZNF557	Inflammation	9.160
C8ORF37	Inflammation	9.154
RNU5D	Inflammation	9.119
CNGB1	Inflammation	9.113
NME3	Inflammation	9.078
BMS1P5	Inflammation	9.046
CREB1	Inflammation	8.934
SHCBP1	Inflammation	8.934
PACRGL	Inflammation	8.908

LILRA6	Inflammation	8.776
FLJ45256	Inflammation	8.721
RNU6-15	Inflammation	8.691
FAM40B	Inflammation	8.682
QRFPR	Inflammation	8.668
DUSP19	Inflammation	8.666
CDAN1	Inflammation	8.619
C2ORF69	Inflammation	8.580
TMEM106A	Inflammation	8.559
CLCC1	Inflammation	8.436
TERF1	Inflammation	8.418
KIAA1143	Inflammation	8.345
SNTN	Inflammation	8.341
PNPT1	Inflammation	8.313
SETD4	Inflammation	8.294
NLRP8	Inflammation	8.291
SLC4A2	Inflammation	8.285
RPS2	Inflammation	8.280
GGA1	Inflammation	8.266
RAXL1	Inflammation	8.243
TRIM16L	Inflammation	8.234
PPP1R15B	Inflammation	8.230
PPM1K	Inflammation	8.178
IL10	Inflammation	8.177
SMCR5	Inflammation	8.174
C1ORF150	Inflammation	8.169
C8ORF45	Inflammation	8.144
CUX1	Inflammation	8.121
CEP27	Inflammation	8.072
STAG3L1	Inflammation	8.045
RAB4B	Inflammation	8.021
RPS27L	Inflammation	7.980
ZNF638	Inflammation	7.930
CDC123	Inflammation	7.913
MIA3	Inflammation	7.878
C19ORF60	Inflammation	7.808
HSPC268	Inflammation	7.764
ERGIC1	Inflammation	7.728
FAM39DP	Inflammation	7.705
TSC22D4	Inflammation	7.695
TMEM110	Inflammation	7.661
MGAT1	Inflammation	7.656
ATP6V0D1	Inflammation	7.627
NUP214	Inflammation	7.614
ZNHIT3	Inflammation	7.605
CRAT	Inflammation	7.592
NUBPL	Inflammation	7.587
SLC35C2	Inflammation	7.558
ZNF461	Inflammation	7.554
NBPF15	Inflammation	7.537

MALAT1	Inflammation	7.517
RAB11B	Inflammation	7.510
PDRG1	Inflammation	7.476
MSTO2P	Inflammation	7.469
C11ORF63	Inflammation	7.456
MLL4	Inflammation	7.454
GSTTP2	Inflammation	7.417
C11ORF83	Inflammation	7.414
CABC1	Inflammation	7.399
STAR	Inflammation	7.380
RNU6-1	Inflammation	7.378
ZNF595	Inflammation	7.372
STAP2	Inflammation	7.319
TRIM34	Inflammation	7.315
LOC646996	Inflammation	7.268
PRRG4	Inflammation	7.265
AGXT2L2	Inflammation	7.260
TRPM7	Inflammation	7.249
RBL2	Inflammation	7.244
PCTK1	Inflammation	7.231
LOC100128288	Inflammation	7.216
NOTCH2NL	Inflammation	7.214
CHERP	Inflammation	7.207
RNU1G2	Inflammation	7.206
SPIN2B	Inflammation	7.199
SIGLEC9	Inflammation	7.197
EHMT1	Inflammation	7.176
RPLP0P2	Inflammation	7.167
LOC255167	Inflammation	7.167
FAM108A3	Inflammation	7.146
DCLRE1C	Inflammation	7.143
IER3	Inflammation	7.142
C9ORF80	Inflammation	7.116
SC4MOL	Inflammation	7.087
PIGW	Inflammation	7.085
FXR1	Inflammation	7.024
RNF126	Inflammation	6.995
SUMO1	Inflammation	6.988
SSB	Inflammation	6.954
FLJ35390	Inflammation	6.911
SDC1	Inflammation	6.909
ZNF682	Inflammation	6.894
KIAA1632	Inflammation	6.868
LEP	Inflammation	6.861
CCDC94	Inflammation	6.853
ADRA2B	Inflammation	6.820
RNY4	Inflammation	6.805
RNY1	Inflammation	6.786
ALPP	Inflammation	6.786
KIAA0101	Inflammation	6.782

MAP3K7IP2	Inflammation	6.775
KILLIN	Inflammation	6.774
RNU5A	Inflammation	6.760
HNRPH1	Inflammation	6.758
COX7A2L	Inflammation	6.758
PPP4R1	Inflammation	6.721
MTFMT	Inflammation	6.720
C2ORF14	Inflammation	6.698
KEAP1	Inflammation	6.681
SLC30A6	Inflammation	6.672
SNORD11B	Inflammation	6.654
C7ORF64	Inflammation	6.654
NBPF1	Inflammation	6.651
EIF3B	Inflammation	6.633
SNAPC1	Inflammation	6.627
UBTF	Inflammation	6.610
ZNF787	Inflammation	6.598
UPLP	Inflammation	6.593
HMGN3	Inflammation	6.583
SKAP2	Inflammation	6.558
HNRNPA1L2	Inflammation	6.535
RGS10	Inflammation	6.533
LOC100190986	Inflammation	6.529
SGPL1	Inflammation	6.528
CCT6A	Inflammation	6.517
ZC3HAV1L	Inflammation	6.502
AFF1	Inflammation	6.488
UTP20	Inflammation	6.460
LRP10	Inflammation	6.460
PARP10	Inflammation	6.456
FAM84A	Inflammation	6.455
KCNH6	Inflammation	6.448
HPR	Inflammation	6.437
N4BP2	Inflammation	6.433
TBCD	Inflammation	6.428
ZNF48	Inflammation	6.410
RNU4-1	Inflammation	6.403
PSIP1	Inflammation	6.382
GCHFR	Inflammation	6.377
MFSD2	Inflammation	6.367
GYG2	Inflammation	6.349
TXNDC11	Inflammation	6.321
FADS2	Inflammation	6.307
ADAM15	Inflammation	6.289
LOC441150	Inflammation	6.275
LOC260339	Inflammation	6.269
POGZ	Inflammation	6.256
BRIP1	Inflammation	6.251
LOC440926	Inflammation	6.251
PEAR1	Inflammation	6.234

NUPR1	Inflammation	6.227
THAP4	Inflammation	6.225
FAM3C	Inflammation	6.217
RABAC1	Inflammation	6.201
DPP9	Inflammation	6.180
SPTLC1	Inflammation	6.177
RNU1F1	Inflammation	6.174
TBC1D2B	Inflammation	6.174
TFAMP1	Inflammation	6.173
SNORD26	Inflammation	6.157
LYL1	Inflammation	6.137
SLC25A16	Inflammation	6.136
HNRPK	Inflammation	6.129
C19ORF24	Inflammation	6.127
CCDC85B	Inflammation	6.126
CHRM2	Inflammation	6.125
ATG3	Inflammation	6.118
HIST2H3A	Inflammation	6.113
PDCD7	Inflammation	6.108
MCPH1	Inflammation	6.098
FBXO18	Inflammation	6.095
CFLAR	Inflammation	6.088
KLK7	Inflammation	6.069
GIYD1	Inflammation	6.068
KRCC1	Inflammation	6.059
PHF1	Inflammation	6.048
SDCCAG3	Inflammation	6.048
FTSJ1	Inflammation	6.047
MED15	Inflammation	6.045
PPP2R5D	Inflammation	6.029
ZNF358	Inflammation	6.024
RNU1-5	Inflammation	6.019
BMP4	Inflammation	6.011
METRN	Inflammation	6.010
ZFP64	Inflammation	6.008
TRADD	Inflammation	6.008
NAT12	Inflammation	6.007
PHC2	Inflammation	5.998
LOC399744	Inflammation	5.991
GLI2	Inflammation	5.990
MAGED2	Inflammation	5.981
GLRX2	Inflammation	5.980
MTDH	Inflammation	5.966
C1RL	Inflammation	5.950
SURF2	Inflammation	5.941
PSMB5	Inflammation	5.920
EPS8L2	Inflammation	5.919
NUDT4P1	Inflammation	5.912
GLTSCR1	Inflammation	5.907
FAM176B	Inflammation	5.899

UBE2T	Inflammation	5.871
SNORA5B	Inflammation	5.847
ZNF205	Inflammation	5.836
DOPEY2	Inflammation	5.836
SNORD3D	Inflammation	5.813
PODXL	Inflammation	5.795
RNU1-3	Inflammation	5.758
TIGD2	Inflammation	5.757
NFE2L1	Inflammation	5.751
RBM9	Inflammation	5.741
FKTN	Inflammation	5.707
TEX11	Inflammation	5.648
MLLT1	Inflammation	5.546

Supplementary Table 4. FLC immunohistochemistry grading (n=42)

Marker	pRPS6		EGFR		HepPar1		CK7		CK19		EpCAM	
Difuse	11	26%	29	69%	39	93%	35	83%	1	2%	2	5%
Focal	11	26%	3	7%	2	5%	3	7%	7	17%	8	19%
Few cels	3	7%	5	12%	0	0%	0	0%	1	2%	0	0%
Positivity	25	60%	37	88%	41	98%	38	90%	9	21%	10	24%

Supplementary Table 5. Chromosomal arm-level genomic copy number alterations

Chr.Arm	FLC01	FLC02	FLC08	FLC22	FLC39	FLC41	FLC05	FLC12	FLC13	FLC14	FLC17	FLC20	FLC23	FLC32	FLC35	FLC43	FLC47	FLC49	FLC50	FLC53	FLC54	FLC56	FLC58	FLC59	FLC60	FLC33	FLC03	FLC11	FLC45	FLC48	FLC51	FLC57
1p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.171	0	0	0	0	0	0	0	0	0	0	0	0.161	0	0
2p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4p	0	0	0	0	0	0	-0.159	0	0	0	0	0	0	0	0	0	0	-0.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4q	0	0	0	0	0	0	-0.159	0	0	0	0	0	0	0	0	0	0	-0.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5p	0	0	0.204	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6p	0	0	0	0	0.17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.168	0	0
9p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12p	0	0	0	0	0	0	-0.201	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.153	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.17	0	0	0	0	0	0	0	0	0	-0.161	0	0	0
16q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.17	0	0	0	0	0	0	0	0	0	-0.161	0	0	0
17p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.193	0	0	-0.152	0	0	0	0	-0.268	0	0	0
17q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.193	0	0	-0.152	0	0	0	0	-0.268	0	0	0
18p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.16	0	0	0	0	0	0	0	0	0	0
18q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.16	0	0	0	0	0	0	0	0	0	0
19p	0.203	-0.153	0	0	0	0	0	0	0	0	0	0	0	0	-0.277	0	-0.302	0.32	0.214	0	-0.344	0	-0.168	-0.306	-0.181	0	0	0	-0.427	0	0	0
19q	0.203	-0.153	0	0	0	0	0	0	0	0	0	0	0	0	-0.277	0	-0.302	0.32	0.214	0	-0.344	0	-0.168	-0.306	-0.181	0	0	0	-0.427	0	0	0
20p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.172	0	0	0	0	0	0	0	-0.195	0	0	0
20q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.172	0	0	0	0	0	0	0	-0.195	0	0	0
21q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.183	0	0	0	0	0	0	0	0	0
22q	0	-0.243	0	0	0	0	0.213	0	0	0	0	0	0	0	-0.169	0	-0.278	0.168	0	0	-0.351	-0.386	0	-0.389	-0.197	0	0	0	-0.394	0	0	0

Supplementary Table 6. FLC molecular subclass of chromosomal altered samples

Alteration	Chr.	Patients involved	Molecular Subclass (%)				p value
		#, (%)	Proliferation	Inflammation	Unannotated	No transcriptomic data	
Broad							
Gains							
	1q	2, (6)	50			50	
	19	3, (9)	33	33	33		
	22q	2, (6)	100				
Losses							
	4	2, (6)	100				
	17	3, (9)		66		33	
	19	8, (25)	25	50	12.5	12.5	
	20	2, (6)		50		50	
	22q	8, (25)	25	50	12.5	12.5	
Focal							
Gains							
	6q27	2, (6)	100				
	8q24.3	4, (12.5)	25		50	25	
	17q25.3	2, (6)	50	50			
Losses							
	1p36.33	3, (9)	33		66		
	8p21.3	2, (6)		50	50		
	11p15.5	1, (3)	100				
	19p13.3	9, (28)	22	56	11	11	p=0.0151
	19p13.12	9, (28)	22	56	11	11	p=0.0151
	22q13.32	8, (5)	25	50	12.5	12.5	

Supplementary Table 7. Focal amplifications in FLC samples

cytoband	17q25.3	6q27	8q24.3
q value	0.004019	0.018965	0.22437
residual q value	0.004019	0.018965	0.22437
wide peak boundaries	chr17:77364582-77395624	chr6:168352106-168552064	chr8:145919127-146364022
total number of affected genes	1	4	12
genes in wide peak	RBFOX3	KIF25	RPL8
		MLLT4	ZNF7
		FRMD1	ZNF16
		HGC6.3	COMMD5
			ZNF250
			C8orf33
			ZNF34
			ZNF251
			ZNF252
			TMED10P1
			C8orf77
			ZNF517

Supplementary Table 8. Focal deletions in FLC samples

cytoband	1p36.33	8p21.3	11p15.5	19p13.12	19p13.3	22q13.32
q value	0.22421	0.22421	0.22421	0.22421	0.22421	0.22421
residual q value	0.22421	0.22421	0.22421	0.22421	0.22421	0.22421
wide peak boundaries	chr1:1-1388267	chr8:21888571-22058121	chr11:1-3009437	chr19:14276712-14623757	chr19:1-4110491	chr22:48929172-49004394
Total number of affected genes	56	10	104	7	150	2
genes in wide peak	hsa-mir-429	BMP1	hsa-mir-483	CD97	hsa-mir-637	FAM19A5
	hsa-mir-1302-2	EPB49	hsa-mir-675	PKN1	hsa-mir-4321	LOC284933
	DVL1	SFTPC	hsa-mir-4298	PTGER1	hsa-mir-1909	
	SCNN1D	FGF17	hsa-mir-210	DDX39A	hsa-mir-3187	
	TNFRSF4	NPM2	AP2A2	GIPC1	hsa-mir-1302-11	
	TNFRSF18	HR	ASCL2	LPHN1	AES	
	ISG15	FAM160B2	CD81	LOC100507373	AMH	
	NOC2L	NUDT18	CD151		ATP5D	
	OR4F3	REEP4	CDKN1C		AZU1	
	SDF4	LGI3	CTSD		HCN2	
	MXRA8		DRD4		BSG	
	CPSF3L		DUSP8		CDC34	
	C1orf159		HRAS		CIRBP	
	AURKAIP1		IGF2		CNN2	
	MRPL20		INS		CSNK1G2	
	HES4		IRF7		DAPK3	
	VWA1		KCNQ1		CFD	
	OR4F5		LSP1		ARID3A	
	LINC00115		MUC2		EEF2	
	GLTPD1		MUC6		EFNA2	
	OR4F16		NAP1L4		ELANE	
	CCNL2		SLC22A18		GAMT	
	TAS1R3		SLC22A18AS		GNA11	
	PLEKHN1		POLR2L		GNA15	
	C1orf170		PSMD13		GNG7	
	ACAP3		RNH1		MKNK2	
	UBE2J2		MRPL23		GPX4	
	PUSL1		RPLP2		GZMM	
	B3GALT6		SCT		MATK	
	SAMD11		TALDO1		GADD45B	
	LOC148413		TH		NFIC	
	ATAD3C		TSPAN4		OAZ1	
	LOC254099		TNNI2		PALM	
	TTLL10		TNNT3		POLR2E	
	FAM41C		PHLDA2		POLRMT	
	KLHL17		RASSF7		MAP2K2	
	AGRN		IFITM1		PRTN3	
	FAM132A		BRSK2		PTBP1	
	RNF223		TSPAN32		RPS15	
	MIR200A		TSSC4		SGTA	
	MIR200B		IFITM3		STK11	
	ANKRD65		DEAF1		TBXA2R	
	MIR429		IFITM2		TCF3	
	FAM138F		KCNQ1OT1		THOP1	
	LOC643837		PKP3		TLE2	
	TMEM88B		SIRT3		MADCAM1	
	FAM138A		C11orf21		SF3A2	
	WASH7P		TRPM5		PPAP2C	
	LOC729737		IGF2-AS1		S1PR4	
	OR4F29		BET1L		AP3D1	
	LOC100130417		CEND1		APBA3	
	LOC100132062		CDHR5		MED16	
	LOC100132287		TOLLIP		FSTL3	
	LOC100133331		PIDD		APC2	

	DDX11L1		KCNQ1DN		ABCA7	
	LOC100288069		PNPLA2		HMG20B	
			PHRF1		UQCR11	
			SIGIRR		SBNO2	
			RIC8A		ZFR2	
			EPS8L2		PIP5K1C	
			CHID1		HMHA1	
			SLC25A22		SHC2	
			ATHL1		TIMM13	
			PTDSS2		DAZAP1	
			MOB2		SNORD37	
			SYT8		FGF22	
			ODF3		TJP3	
			LRRC56		ITGB1BP3	
			LOC143666		SLC39A3	
			SCGB1C1		THEG	
			NLRP6		ZBTB7A	
			NS3BP		FZR1	
			LOC255512		PIAS4	
			C11orf35		LSM7	
			H19		MBD3	
			EFCAB4A		MIER2	
			TMEM80		PCSK4	
			ANO9		C19orf24	
			LOC338651		PLEKHJ1	
			B4GALNT4		BTBD2	
			PDDC1		RNF126	
			KRTAP5-1		NCLN	
			KRTAP5-3		SPPL2B	
			KRTAP5-4		WDR18	
			IFITM5		REXO1	
			FAM99A		ZNF77	
			IFITM10		C19orf29	
			MIR210		CELF5	
			KRTAP5-5		TLE6	
			KRTAP5-2		LPPR3	
			KRTAP5-6		ZNF556	
			MIR483		OR4F17	
			SNORA52		FAM108A1	
			LOC653486		DOHH	
			SNORA54		KLF16	
			INS-IGF2		DOT1L	
			MUC5B		KISS1R	
			MIR675		LMNB2	
			FAM99B		RAX2	
			LOC100133161		MUM1	
			MRPL23-AS1		ATCAY	
			MIR4298		MIDN	
			MIR210HG		R3HDM4	
			MIR4686		C19orf6	
					TPGS1	
					REEP6	
					IZUMO4	
					SCAMP4	
					ADAT3	
					ZNF554	
					GRIN3B	
					MRPL54	
					ZNF57	
					JSRP1	
					MOB3A	
					MFSD12	
					GIPC3	

					C19orf21	
					PLK5	
					C2CD4C	
					CIRBP-AS1	
					C19orf25	
					ATP8B3	
					DIRAS1	
					ZNF555	
					C19orf26	
					CSNK1G2-AS1	
					C19orf77	
					ODF3L2	
					ADAMTSL5	
					TMPRSS9	
					NDUFS7	
					C19orf35	
					WASH5P	
					ONECUT3	
					MEX3D	
					FLJ45445	
					PRSS57	
					C19orf29-AS1	
					FAM138F	
					LINGO3	
					FAM138A	
					MIR637	
					C19orf71	
					LOC100288123	
					MIR1909	
					MIR1227	
					MIR3187	
					MIR4321	
					MIR4745	

Supplementary Table 9. Chromosomal arm-level genomic copy number alterations of the Fresh Frozen cohort

Alteration	chr.arm	92T	187T	232T	255T	334T	406T	407T	410T	1310T	total	%	93T*	337T*
Gains	8q	1	0	0	1	0	0	0	0	0	**2**	**22**	0	1
	15q	1	0	0	0	0	0	0	0	0	**1**	**11**	0	0
	20q	1	0	0	0	0	0	0	0	0	**1**	**11**	0	0
Losses	18p	1	0	1	1	0	0	0	0	0	**3**	**33**	1	1
	18q	1	0	1	1	0	0	0	0	0	**3**	**33**	1	1
	8p	1	0	0	1	0	0	0	0	0	**2**	**22**	1	1
	21q	1	0	0	0	0	0	0	0	1	**2**	**22**	0	0
	22q	1	0	1	0	0	0	0	0	0	**2**	**22**	0	0
	13q	1	0	0	0	0	0	0	0	0	**1**	**11**	1	0
	20p	1	0	0	0	0	0	0	0	0	**1**	**11**	1	0
	4q	1	0	0	0	0	0	0	0	0	**1**	**11**	0	0
	9q	0	0	0	0	0	0	0	0	1	**1**	**11**	0	0
Allelic imbalance	22q	0	0	0	0	0	1	0	0	0	**1**	**11**	1	0
	1p	0	1	0	0	0	0	0	0	0	**1**	**11**	0	0
	3p	0	1	0	0	0	0	0	0	0	**1**	**11**	0	0
	4q	0	1	0	0	0	0	0	0	0	**1**	**11**	0	0
	9p	1	0	0	0	0	0	0	0	0	**1**	**11**	0	0
	9q	1	0	0	0	0	0	0	0	0	**1**	**11**	0	0
	15q	0	0	0	0	0	1	0	0	0	**1**	**11**	0	0
	18p	0	0	0	0	0	1	0	0	0	**1**	**11**	0	0
	18q	0	0	0	0	0	1	0	0	0	**1**	**11**	0	0
	21q	0	0	0	0	1	0	0	0	0	**1**	**11**	0	0

* 93T is a tumoral sample from a different part of the same tumor as T92, and 337T is the relapse from the primary tumor 255T

Supplememntary Table 10: Summary of the damaging mutations found by WES

Chr.	Sample	Gene	Position	Ref. Allele	Alt. Allele	Mutation Profile	AminoAcid Change	Cosmic HCC
chr1	FLC100	CSMD2	34068134	C	T	NON_SYNONYMOUS_CODING	G1055E	2.7
chr1	FLC100	SASS6	100575999	T	G	NON_SYNONYMOUS_CODING	K210T	
chr2	FLC100	RTN4	55214684	C	G	NON_SYNONYMOUS_CODING	R1055T	
chr3	FLC100	COL6A6	130290045	C	A	NON_SYNONYMOUS_CODING	L929I	
chr4	FLC100	GABRA2	46252405	G	A	NON_SYNONYMOUS_CODING	P426S	1.4
chr4	FLC100	UGT2B10	69879855	T	A	NON_SYNONYMOUS_CODING	Y201F	
chr7	FLC100	WIPI2	5256247	G	A	NON_SYNONYMOUS_CODING	M127I	
chr7	FLC100	CDK13	40133813	A	G	NON_SYNONYMOUS_CODING	D1198G	
chr9	FLC100	CD72	35618023	C	A	NON_SYNONYMOUS_CODING	G60W	1.4
chr10	FLC100	TUBB8	94579	A	G	NON_SYNONYMOUS_CODING	F13L	1.4
chr10	FLC100	JMJD1C	64936109	A	C	NON_SYNONYMOUS_CODING	L2213R	1.4
chr10	FLC100	MMS19	99229890	G	A	NON_SYNONYMOUS_CODING	T238I	1.4
chr13	**FLC100**	**BRCA2**	**32944573**	**A**	**G**	**NON_SYNONYMOUS_CODING**	**Y2789C**	**1.5**
chr14	FLC100	RNASE7	21511552	G	C	NON_SYNONYMOUS_CODING	C134S	
chr14	FLC100	ACIN1	23564195	G	A	NON_SYNONYMOUS_CODING	R101W	1.4
chr16	FLC100	POLR3K	97543	A	G	NON_SYNONYMOUS_CODING	C72R	
chr16	FLC100	ITGAM	31308917	C	T	NON_SYNONYMOUS_CODING	P480L	1.4
chr16	FLC100	HAS3	69148458	C	G	NON_SYNONYMOUS_CODING	D317E	
chr18	FLC100	IMPA2	12028977	G	A	NON_SYNONYMOUS_CODING	V246M	
chr18	FLC100	ZNF236	74583728	G	T	NON_SYNONYMOUS_CODING	G203V	
chr19	FLC100	CRX	48364098	C	T	NON_SYNONYMOUS_CODING	R104W	1.4
chr20	FLC100	PRND	4705610	A	C	NON_SYNONYMOUS_CODING	E138A	
chr20	FLC100	ADNP	49518604	T	C	NON_SYNONYMOUS_CODING	T51A	
chrX	FLC100	ARMCX1	100808651	G	T	NON_SYNONYMOUS_CODING	L246F	1.4

Supplementary Table 1. Clinicopathological characteristics of 72 FLC patients in the training and validation cohorts

Variable	Training (n=38)	Validation-French (n=22)	Validation-Brazilian (n=12)	p-value
Origin				
USA	19 (50%)	0	0	
Europe	19 (50%)	22 (100%)	0	
South-America		0	12 (100%)	
Median age	25.5 (11-65)	26 (17-65)	21 (17-33)	0.26
Gender				0.51
Male	16 (42%)	6 (28%)	4 (33%)	
Female	22 (58%)	16 (72%)	8 (67%)	
Race				0.24
Caucasian	23 (61%)		11 (92%)	
Non-caucasian	8 (21%)		1 (8%)	
Virology				-
HCV	1 (3%)		0	
HBV	1 (3%)		0	
Fibrosis				0.06
F0	27 (71%)	21 (95%)	12 (100%)	
F1-2	8 (21%)	1 (5%)	0	
Cirrhosis	1 (3%)	1 (5%)	0	1.00
Median tumor size (cm)	11 (7-13)	9 (2.5-17)	13 (6-20)	**0.03**

Multiple nodules				0.66
Absent	27 (69%)		11 (92%)	
Present	7 (18%)		1 (8%)	
Satellites	6 (16%)		2 (17%)	1.00
Macro-vascular invasion	5 (13%)			-
Treatment				0.28
Resection	34 (89%)	22 (100%)	12 (100%)	
Transplantation	4 (11%)	0	0	
Adjuvant Therapy	14 (37%)		0	1.00
Chemotherapy	11 (29%)			-
Radiotherapy	2 (5%)			-
Chemoembolization	1 (3%)			-
Events				
Recurrence	21 (54%)		7 (58%)	1.00
Death	16 (42%)	9 (41%)	7 (58%)	0.65
Median survival (month)	58	68	48	0.37

Fisher's exact test for categorical data.
Anova test for parametric non-categorical data

Supplementary Table 12. Enrichment of selected gene signatures in 3 liver cancer cohorts (HCC, ICC and FLC) evaluated by NTP (FDR<0.05).

Gene signatures	Liver cancer cohorts		
	HCC n=228, (%)	**ICC n=149, (%)**	**FLC n=35, (%)**
FLC molecular subclasses			
FLC_Proliferation	54,(24)	73,(49)	18,(51)
FLC_Inflammation	8,(4)	43,(29)	8,(23)
Tumor aggressiveness/ poor-prognosis			
Solid tumors			
Breast, Sarrio_EMT[36]	55,(24)*	21,(14)*	0,(0)
Thyroid, Montero_poor prognosis[37]	37,(16)*	11,(7)	0,(0)
Melanoma, Winnepenninckx_poor prognosis[38]	59,(26)*	32,(21)*	0,(0)
HCC			
Coulouarn_TGFb[39]	32,(14)	28,(19)*	1,(3)
Boyault_G3 poor prognosis[16]	53,(23)*	4,(3)	3,(9)

* significant (p<0.05) vs FLC tumors