ORIGINAL RESEARCH

Recurrent chromosomal rearrangements of *ROS1*, *FRK* and *IL6* activating JAK/STAT pathway in inflammatory hepatocellular adenomas

Quentin Bayard,[1] Stefano Caruso,[1] Gabrielle Couchy,[1] Sandra Rebouissou,[1] Paulette Bioulac Sage,[2,3] Charles Balabaud,[3] Valerie Paradis,[4,5] Nathalie Sturm,[6] Anne de Muret,[7] Catherine Guettier,[8] Benjamin Bonsang,[9] Christiane Copie,[9] Eric Letouzé,[1] Julien Calderaro,[9] Sandrine Imbeaud ![ORCID],[1] Jean-Charles Nault ![ORCID],[1,10,11] Jessica Zucman-Rossi[1,12]

Correspondence to
Dr Jean-Charles Nault, INSERM UMR 1338, INSERM, Paris 93140, France;
naultjc@gmail.com
Professor Jessica Zucman-Rossi;
jessica.zucman-rossi@inserm.fr

J-CN and JZ-R are joint senior authors.

![Check for updates]

ABSTRACT

Background Inflammatory hepatocellular adenomas (IHCAs) are benign liver tumours characterised by an activation of the janus kinase (JAK)/signal transducers and activators of transcription (STAT) pathway caused by oncogenic activating mutations. However, a subset of IHCA lacks of identified mutation explaining the inflammatory phenotype.
Methods 657 hepatocellular adenomas developed in 504 patients were analysed for gene expression of 17 genes and for mutations in seven genes by sequencing. 22 non-mutated IHCAs were analysed by whole-exome and/or RNA sequencing.
Results We identified 296 IHCA (45%), 81% of them were mutated in either *IL6ST* (61%), *FRK* (8%), *STAT3* (5%), *GNAS* (3%) or *JAK1* (2%). Among non-mutated IHCA, RNA sequencing identified recurrent chromosome rearrangement involving *ROS1*, *FRK* or *IL6* genes. *ROS1* fusions were identified in 8 IHCA, involving C-terminal part of genes highly expressed in the liver (*PLG*, *RBP4*, *APOB*) fused with exon 33–35 to 43 of *ROS1* including the tyrosine kinase domain. In two cases a truncated *ROS1* transcript from exon 36 to 43 was identified. *ROS1* rearrangements were validated by fluorescence in situ hybridisation (FISH) and led to *ROS1* overexpression. Among the 5 IHCA with *FRK* rearrangements, 5 different partners were identified (*MIA3*, *MIA2*, *LMO7*, *PLEKHA5*, *SEC16B*) fused to a common region in *FRK* that included exon 3–8. No overexpression of *FRK* transcript was detected but the predicted chimeric proteins lacked the auto-inhibitory SH2–SH3 domains. In two IHCA, we identified truncated 3'UTR of *IL6* associated with overexpression of the transcript.
Conclusion Recurrent chromosomal alterations involving *ROS1*, *FRK* or *IL6* genes lead to activation of the JAK/STAT pathway in IHCAs.

INTRODUCTION

Hepatocellular adenomas (HCAs) are rare benign tumours mainly fostered by exposure to oestrogen, particularly in women taking oral contraceptives.[1] Management of hepatocellular adenoma requires assessment of the risk of complications mainly related to tumour bleeding and malignant

Significance of this study

What is already known on this subject?
► Inflammatory hepatocellular adenomas (IHCAs) are benign liver tumours characterised by an activation of the JAK/STAT pathway caused by oncogenic activating mutations.
► A subset of IHCA lacks of identified mutation explaining the inflammatory phenotype.

What are the new findings?
► We identified recurrent chromosomal alterations involving *ROS1*, *FRK* or *IL6* in IHCAs.
► *ROS1* activation is a novel mechanism of JAK/STAT pathway activation in human benign liver tumours.

How might it impact on clinical practice in the foreseeable future?
► These new genetic alterations could be potential diagnostic biomarkers of a subtype of hepatocellular adenoma.
► Targeted therapies adapted to genetic alterations could be tested in selected patients with multiple or unresectable IHCAs.

transformation to hepatocellular carcinoma (HCC).[2]

Genotype/phenotype classification of HCA has dissected adenomas in homogeneous subgroups characterised by somatic genetic alterations dysregulating signalling pathways and linked with specific risk factors, clinical behaviours, histological and imaging features.[3] The first subgroups of HCA (HHCA, 30%–40%) are defined by bi-allelic inactivating mutations of *HNF1A* and are associated with tumour steatosis, lipogenesis and mammalian target of rapamycin (mTOR) activation.[4,5] Another group of HCA is defined by mutations in *CTNNB1*, coding for ß-catenin, either in exon 3 (b^ex3HCA, around 15% of HCA) or in exon 7 or 8 (b^ex7,8HCA, around 10% of HCA).[3,6] Mutations in *CTNNB1* exon 3 are associated with a high risk of malignant transformation in HCC.[3] A new subgroup associated

with a high risk of tumour bleeding was recently described in 4%–5% of the patients (sonic hedgegoh hepatocellular adenoma (shHCA)); it is defined by an activation of the sonic hedgehog pathway due to a *GLI1* overexpression caused by a somatic fusion with *INHBE*, its neighbouring gene.[7]

The most prevalent subgroup of HCA are inflammatory tumours (inflammatory hepatocellular adenomas (IHCAs), around 50% of HCA), characterised by the constitutive and uncontrolled activation of interleukin 6 (IL-6)/JAK/STAT pathway easily detectable by the overexpression of serum amyloid A (SAA) and C-reactive protein (CRP) at the transcriptomic level, associated with inflammatory infiltrate, sinusoidal dilatation and dystrophic arteries at histological level.[8] Overexpression of SAA and CRP in the tumour on immunohistochemistry helps to identify this subtype in routine. Interestingly, a subset of IHCA is mixed, harbouring mutation of *CTNNB1* either in exon 3 or exon 7/8 (bex³IHCA, bex⁷,⁸IHCA). Risk factors for development of IHCA are high alcohol consumption and obesity.[9] Several genetic alterations are responsible for the constitutive activation of JAK/STAT pathway in IHCA. The most frequent mutation are in *IL6ST* (coding for the gp130 receptor), *STAT3* (coding for a transcription factor of the JAK/STAT pathway), *FRK* (coding for a src kinase), *JAK1* (coding for one of the protein tyrosine kinase of the JAK/STAT pathway) and *GNAS* (coding stimulatory G-protein alpha subunit).[6 10 11] Inflammatory syndrome, anaemia or fever could be observed in patients with IHCA; these symptoms regress after tumour resection. Recently, we described an IHCA with a rearrangement of the 3'UTR of the *IL6* gene leading to a secondary amyloidosis through an autocrine/paracrine mechanism of activation of JAK/STAT pathway by IL-6 in the whole liver.[12] However, a subset of IHCA has currently no known genetic alterations explaining the constitutive activation of JAK/STAT pathway.

The aim of our study was to unravel new mechanisms of activation of inflammatory pathway in inflammatory adenomas without any mutation in known driver genes to better understand the mechanism of regulation of IL-6/JAK/STAT pathway in hepatocytes and its role in benign liver tumourigenesis.

MATERIAL AND METHODS
Selection of tumours and patients
Frozen samples from 657 HCA developed in 504 patients were analysed. This series included 124 new HCA cases apart from previously analysed HCA cases by Nault *et al*[7] . Samples were collected in 28 centres in France from 2000 to 2017. All patients gave their informed consent according to French law. All samples were frozen in nitrogen immediately after resection or biopsy and conserved at −80°C.

DNA sequencing
HNF1A (exon 1–10), *IL6ST* (exons 6 and 10), *CTNNB1* (exons 2, 3, 4, 7 and 8), *FRK* (exon 6), *STAT3* (exons 3, 6, 17 and 21), *GNAS* (exons 7, 8 and 9) and *JAK1* (all exons) were sequenced by either Sanger sequencing or Miseq Illumina PCR-based sequencing (for detailed protocol see [6 13 14]). Somatic mutations were confirmed by sequencing of the tumour and non-tumour counterpart.

Quantitative RT-PCR
Gene expression of a selection of 17 genes was assessed by quantitative reverse transcriptase PCR (RT-PCR) using Fluidigm technology with Taqman probes detailed in online supplementary table 1, as previously described.[7 13] Data were normalised using RNA ribosomal 18S as calibrator. We used the 2 delta delta CT method with the level of the corresponding gene expression in tumour tissues expressed as an n-fold ratio compared with the mean level of expression in five normal livers.

As previously described,[7] HHCAs were defined by *HNF1A* bi-allelic mutations and/or downregulation of *LFABP* and *UGT2B7* (T/N<0.2); IHCAs were defined by SAA and/or CRP overexpression T/N>5 and/or activating mutations in *IL6ST*, *FRK*, *JAK1*, *STAT3*, *GNAS*. b^ex3^HCAs were defined by mutations or in frame deletion of *CTNNB1* exon 3, b^ex7,8^HCAs were defined by mutations of *CTNNB1* in exon 7 or 8, bHCAs were defined by overexpression of Wnt/catenin pathway target genes *GLUL* or *LGR5* (T/N>5) without identified gene mutation. b^ex3^IHCAs were identified by the combination of b^ex3^HCA and IHCA criteria, bIHCA by the combination of bHCA and IHCA criteria and b^ex7,8^IHCA by the combination of b^ex7,8^HCA and IHCA criteria. shHCA was defined by overexpression of the following target genes of sonic hedgehog pathway: *GLI1*, *TNNC1*, *FCRLA*, *GPR97*, *HHIP*, *PTCH1* and *PTGDS*.

Whole exome sequencing
Whole exome data were analysed in 19 HCA cases which included 4 previously published and 15 new cases. Sequence capture, enrichment and elution of genomic DNA samples from the tumour/normal pairs were performed by IntegraGen (Evry). Agilent in-solution enrichment was used with the manufacturer's biotinylated oligonucleotide probe library SureSelect Human All-Exon kit V.4 70Mb (n=4), SureSelect Human All-Exon kit V.5 +UTRs (n=8), or Twist Human Core Exome Enrichment System (n=7). Briefly, 3 µg of each genomic DNA was sonicated and purified to yield fragments of 150–200 bp. Adaptor oligonucleotides were ligated on A-tailed fragments and enriched by four to six PCR cycles. Purified libraries (500 ng) were hybridised to the SureSelect library for 24 hours. The eluted enriched DNA sample was sequenced on an Illumina HiSeq 2000 (n=12) or HiSeq 4000 (n=7) as paired-end 75 bp reads. Sequencing details for each sample are indicated in online supplementary table 2. Variant calling was processed using Mutect2 by comparing each tumour sample with its matched non-tumour counterpart, following the genone analysis toolkit (GATK) best practices.[15 16] The copy number profiles were defined using an in-house pipeline. Briefly, the logR ratio was processed using the coverage of each tumour and its non-tumour counterpart in each 1 kb genomic windows; then a segmentation was done in order to identify copy number alterations.

RNA sequencing
RNA samples from the 22 HCA tumours were sequenced comprising 1 previously published[12] and 21 new cases. RNA samples were enriched for polyadenylated RNA from 5 µg of total RNA, and the enriched samples were used to generate sequencing libraries with the NEBNext Ultra II Directional RNA (n=13) or Illumina TruSeq Stranded messenger RNA (mRNA) (n=9) kits and associated protocol as provided by the manufacturer. Libraries were sequenced by IntegraGen on an Illumina HiSeq 2000 or 4000 as paired-end 101 bp or 75 bp reads (n=8 and n=14, respectively). Full Fastq files were aligned to the reference human genome GRCh38 using TopHat2 V.2.0.14.[17] Sequencing details for each sample and the parameters used for TopHat2 are indicated in online supplementary table 2. Read mapping from multiple locations was removed and HTSeq[18] was used to obtain the number of reads associated with each gene in the Gencode V.27 database, restricting to protein-coding genes,

pseudogenes, antisense and lincRNAs (n=39 947). Bioconductor DESeq2 package[19] was used to import raw HTSeq counts for each sample into R statistical software and apply variance stabilising transformation to the raw count matrix. FPKM scores (number of fragments per kilobase of exon model and millions of mapped reads) were calculated by normalising the count matrix for the library size estimated with DESeq2 package and the coding length of each gene. In order to handle batch effect, we used an in-house method, according to our previous published method.[20] Briefly, area under the ROC curve (AUC) was used to identify and remove 2696 genes with a significant batch effect (AUC >0.95 between one sequencing project and others).

To perform gene set enrichment analysis (GSEA) analysis, genes from RNA-seq data were sorted according to fold-change from Limma analysis (genes with positive fold-changes first, then negative fold-changes) using an input of 16 387 genes (Hallmarks gene set; MSigDB V.6 database used). This gene list was obtained from Genecode database; then genes with a variance equal 0 or with a total number of reads from samples inferior of 50 were removed.

Gene fusion detection

Fusions detected by TopHat2 (--fusion-search --fusion-min-dist 2000 --fusion-anchor-length 13 --fusion-ignore-chromosomes chrM) were filtered using the TopHatFusion[20] pipeline. Fusions validated by BLAST and with at least 10 split-reads or pairs of reads spanning the fusion event were retained, whereas any fusion identified at least twice in a cohort of 36 normal liver samples was removed. For each fusions of interest, the sequence corresponding to chimaera was generated and used as a reference to realign RNAseq reads. Visual inspection of chimaera alignments on Integrative Genome Viewer (broad institute) was used to refine junction positions and determine in phase chimeric transcripts. Oncofuse V.1.0.9b2 was finally used to annotate fusions of interest.[21]

FISH experiment

Whole-tissue sections from formalin-fixed, paraffin-embedded surgical samples of HCA were deparaffinised in xylene and washed in ethanol. After rehydratation, they were subjected to heat pre-treatment (Pre-Treatment Solution, DakoCytomation, 10 min, 75°C). After digestion with pepsin (10 min, DakoCytomation) and airdrying, the SPEC ROS1 dual colour break apart probe (Zytovision Z-2144) was applied to the tissue sections. Probe mixture and tissue sections were further codenatured and hybridised overnight (Dako Hybridizer). After washing, dehydration and drying, tissue sections were finally counterstained using fluorescence mounting media containing de-amidino-phenyl-indole.

Statistical analysis

Statistical analysis was performed using the GraphPad Prism and the R V.3.4.2 software. Comparison between qualitative data was performed using Fisher's exact test. Quantitative data in two different group were compared using the non-parametric Mann-Whitney test. All tests were two-tailed and a p value<0.05 was considered as significant.

Data availability

The aligned sequencing data reported in this paper have been deposited to the EGA (European Genome-phenome Archive) database (RNA-seq accession EGAS00001003025 and

Table 1 Description of the tumours

Variables	Number (percentages) n=657
HCA	605 (92%)
Borderline lesions (HCA/HCC)	39 (6%)
HCC on HCA	13 (2%)
H HCA	226 (34%)
I HCA	222 (34%)
bex 7,8 IHCA	21 (3%)
bex3 IHCA	44 (7%)
b IHCA	9 (1%)
bex 7,8 HCA	19 (3%)
bex3 HCA	52 (8%)
b HCA	4 (1%)
Sh HCA	34 (5%)
U HCA	26 (4%)

b IHCA, IHCA with activation of B-catenin pathway without mutations. b HCA, HCA with activation of B-catenin pathway without mutations ;bex3 IHCA, IHCA with mutations of *CTNNB1* in exon 3; bex7,8 IHCA, IHCA with mutations of *CTNNB1* in exon 7 or 8;bex3 HCA, HCA with mutations of *CTNNB1* in exon 3;bex7,8 HCA, HCA with mutations of *CTNNB1* in exon 7 or 8; H HCA, *HNF1A* inactivated HCA;HCA, hepatocellular adenoma; HCC on HCA, hepatocellular carcinoma developed on hepatocellular adenoma;HCC, hepatocellular carcinoma; IHCA, inflammatory hepatocellular adenoma; Sh HCA, sonic hedgehog HCA; U HCA, unclassified HCA;

EGAS00001003685; WES accessions EGAS00001000679 and EGAS00001003686).

## RESULTS
Clinical features of the cohort of patients

657 HCA developed in 504 patients with a median age of 38 years and comprising of 85% of women with 87% intake of oral contraceptive pills were included in the study (see flow chart, online supplementary figure 1).

Forty-four patients harboured liver adenomatosis (features of the series are described in table 1). 296 HCA (45%) were IHCA with an overexpression of *SAA* and *CRP* at the transcriptomic level and/or a mutation in the JAK/STAT signalling pathway (see material and methods). 76% of patients with IHCA showed an increase in serum CRP. Patients with IHCA were older (median age 40 years old vs 36 years old, p value=0.039, Mann-Whitney test), with a longer cumulative uses of oral contraception (median 19 years vs 11 years, p value<0.0001, Mann-Whitney test), a higher body mass index (median 25 kg/m² vs 22 kg/m², p value=0.003, Mann-Whitney test) and more frequent chronic alcohol intake (24% vs 9%, value=0.002, Fisher's exact test) compared with patients harbouring other subgroups of HCA.

Spectrum of mutations in known oncogenes in IHCA

Among the 296 IHCA, 74 were mixed tumours with 21 bex7,8IHCA (3% of the whole series, defined by mutations of *CTNNB1* in exon 7 or 8) and 53 bex3IHCA/bIHCA (8% of the whole series) including 44 bex3IHCA (defined by a mutation of *CTNNB1* in exon 3, 7%) and 9 bIHCA (defined by an overexpression of Wnt/catenin pathway target genes without mutation of *CTNNB1*, 1%). Borderlines lesions between HCA and HCC or HCC developed on HCA were identified in 6% of all IHCA (n=18). Nine borderlines lesions between HCA and HCC or HCC developed on HCA were identified among 53 bex3IHCA/bIHCA (17%) versus 9 among 243 IHCA without activation of ß-catenin pathway (4%, p value Fisher's exact=0.0013) confirming that activation of ß-catenin pathway is associated
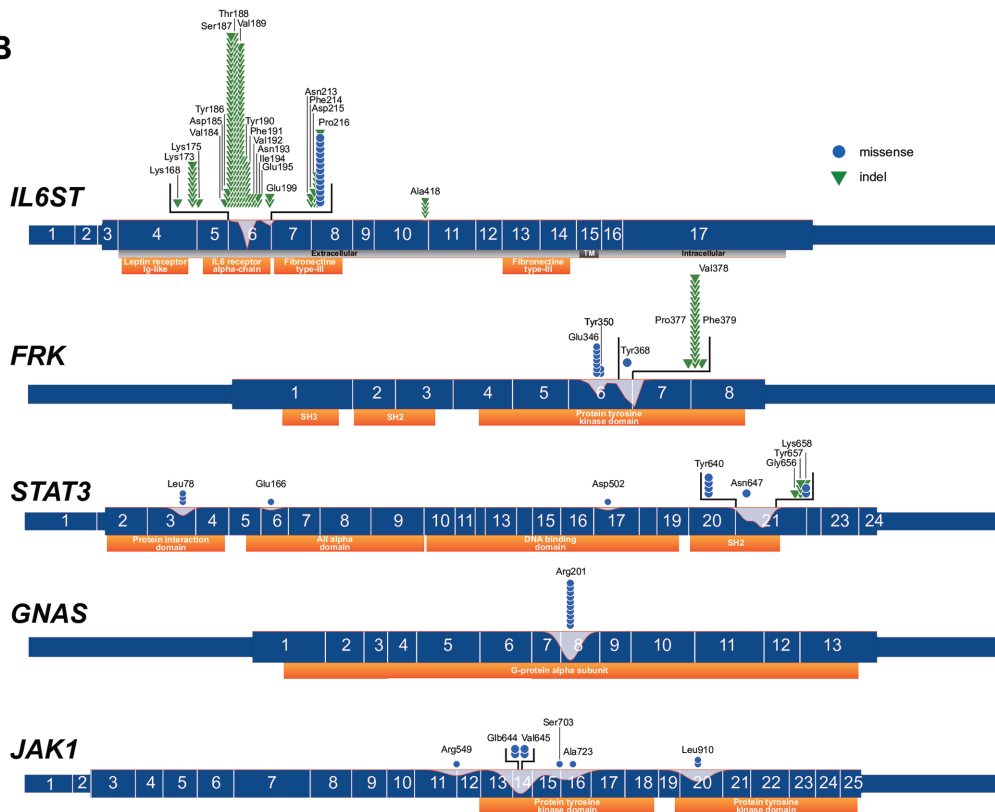
1669
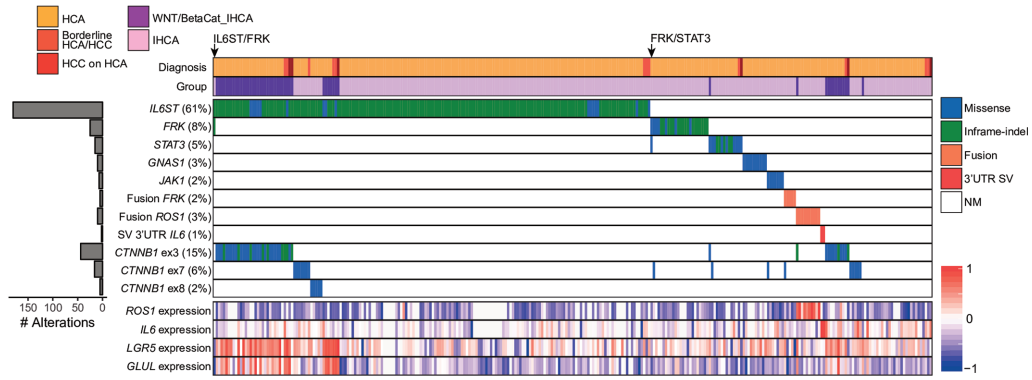
Figure 1 Distribution of mutations in inflammatory hepatocellular adenomas (IHCAs). (A) We represented a heatmap of the somatic mutations and chromosomal rearrangements in driver genes (*IL6ST, FRK, STAT3, GNAS, JAK1, ROS1, IL6* in line) in 296 IHCA (in column) (top). We included the expression level in log2 of *ROS1, IL6, GLUL, LGR5,* assessed by quantitative RT-PCR (bottom). (B) Mutational spectrum of *IL6ST, FRK, STAT3, GNAS* and *JAK1*. Red curves represent the density of mutation among genes. Orange boxes are protein domains annotations from the pfam database. HCA, hepatocellular adenoma; IHCA, inflammatory hepatocellular adenoma; HCC on HCA, hepatocellular carcinoma developed on hepatocellular adenoma; indel, inframe deletion.

with malignant transformation even in IHCA. In contrast, *HNF1A* inactivation (n=226) and sonic hedgehog activation (n=34) were almost exclusive from the inflammatory phenotype since only one mixed IHCA with bi-allelic inactivating mutations of *HNF1A* (#1355) was identified in our series.

Among the 296 IHCA, we identified an activating mutation of a gene belonging to the *JAK/STAT* pathway in 81% of the cases (figure 1A and B, online supplementary table 3):

► Mutations in *IL6ST,* identified in 180 cases, were almost all located in exon 6 in the IL-6 receptor alpha-chain domain and were mostly small in-frame deletions and rarely point mutations. Nevertheless, the remaining four in-frame

deletions clustered in exon 10 corresponding to the D4 domain of gp130 delineating this second hotspot of mutations in *IL6ST*.

► *FRK* mutations, identified in 25 cases, clustered in the protein tyrosine kinase domain (exon 6) and were mainly small in-frame deletions. Interestingly, three samples displayed two *FRK* mutations on the same allele in exon 6 including two samples with p.Glu346Gly/p.Tyr350Cys mutations and one sample with p.Glu346Gly/p.Tyr368Cys mutations.

► *STAT3* mutations, identified in 15 cases, were either small in frame deletion or single nucleotide variation located in exons 3, 6, 17 and 21. Most of the mutations clustered in exons 3

and 21 coding for the protein interaction and SH2 domains, respectively. Interestingly, one sample with a mutation in *STAT3* exon 21 was also mutated in exon 17 (p.Lys658Tyr/p.Asp502Tyr) and another was mutated both in exons 21 and 3 (p.Tyr640Phe/p.Leu78Arg). Finally, one IHCA harboured a single mutation in exon 6 (p.Glu166Gln).

► All mutations in *GNAS,* identified in 10 cases, were missense identified at the Arginine 201, the known exon 8 hotspot activating the protein Gs-alpha subunit.

► *JAK1* mutations were identified in seven cases (2% of IHCA) in exons 11, 14, 15, 16 and 20 all leading to amino acid substitution activating *JAK1*. Interestingly, two samples were double mutated in *JAK1* exon 14 on the same allele at residues 644 and 645.

All oncogenic mutations identified in IHCA were mutually exclusive except for two cases: an IHCA harboured both *STAT3* (p.Leu78Arg) and *FRK* (p.Glu346Gly) mutations, another tumour showed an *IL6ST* in-frame deletion (p.Tyr186_Tyr190del) together with a *FRK* mutation (p.Val378_Lys380delinsGlu) (figure 1A). As previously described,[11] the expression level of *SAA2* and *CRP* was lower in *GNAS* activating mutated IHCA compared with IHCA with other genetic alterations (online supplementary figure 2)

Overall, 61 IHCA (21%) harboured no mutations in driver genes known to activate the JAK/STAT pathway.

Identification of recurrent *ROS1* fusion genes in IHCA
In order to search for new oncogenes in IHCA, we performed whole exome sequencing (WES) in 19 IHCA without known mutations. Comparing tumours and matched non-tumour liver samples, a median number of 2 synonymous mutations and 14 non-synonymous mutations per tumour were identified (online supplementary table 4). None of these samples harboured recurrent genetic alterations or recurrent copy number variation and no genes belonging to the *JAK/STAT* pathway was altered (online supplementary tables 4,5).

In contrast, using RNA sequencing in 22 IHCA without a known driver gene mutation, recurrent *ROS1* transcript alterations were identified in 10 cases (figure 2, online supplementary table 6); among these, 8 IHCA showed a chimeric transcript between the 3' part of *ROS1* and the 5' part of three different partner genes which are highly expressed in normal hepatocytes. *PLG-ROS1* fusions were identified in five cases as the result of a chromosome translocation t(6,10) (q22.1;q23.33). *RBP4-ROS1* fusions were identified in two cases as the result from a chromosome inversion at chromosome 6, Inv6 (q22.1-q25.3). Finally, *APOB-ROS1* fusion was identified in one case resulting from a chromosome translocation t(2,6) (q22.1;q24.1). Breakpoints in *ROS1* and the partner genes were variable leading in each case to at least one in phase transcript . However, all fusions included at least the last 9 exons of *ROS1* (exon 33–43 in one case, exon 34–43 in six cases and exon 35–43 in one case) fused in frame with the 5' of the various partner genes. One tumour with a *ROS1* fusion was also mutated for *CTNNB1* in exon 3 (figure 1A, online supplementary table 2).

Finally, in the last two cases, we identified a truncated mRNA of *ROS1* with an overexpression of exons 36–43 (#4089T) or 37–43 (#1896T) and, in these cases, we did not identify any fusion partner in RNA sequencing data; however, RNA-seq quality of sample #1896T was poor, due to a degraded RNA. In three cases, we performed a FISH experiment that confirmed the *ROS1* break in each tumour, including the two samples with only a truncated transcript identified (#4089T and #1896T)

(figure 2D). GSEA analyse of the transcriptome of HCA with ROS1 fusions confirmed an enrichment in JAK/STAT pathway (HALLMARK IL6 JAK STAT3 SIGNALING; online supplementary figure 3).

In all the cases, the tyrosine kinase domain of *ROS1* (located in exons 36–42) was retained in the rearranged genes and the *ROS1* transcript was highly overexpressed (figure 2B, median FPKM=20.4) compared with HCA without *ROS1* fusion (median FPKM=7.5×10^{-3}, p value=8.6×10^{-5}) and compared with non-tumour liver (n=4, median FPKM=5.1×10^{-2}, p value=7.0×10^{-3}).

Identification of recurrent *FRK* fusion genes in IHCA
By analysing the 22 IHCA with RNAseq, five tumours with a chimeric transcript involving *FRK* were identified (figure 3A, online supplementary table 6), with one of them also mutated for *CTNNB1* in exon 7 (online supplementary table 2). All of them harboured an enrichment in genes belonging to JAK/STAT signalling at the transcriptomic level using GSEA analysis (online supplementary figure 3). *FRK* fusions resulted from various translocations linking exons 1–25 of *MIA3* (chromosome 1q41), or exons 1–20 of *MIA2* (chromosome 14q21.1) or exons 1–25 of *LMO7* (chromosome 13q22.2) or exons 1–19 of *PLEKHA5* (chromosome 12p12.3) with exons 3–8 of *FRK* (chromosome 6q21.1). The last *FRK* fusion linked exons 1–20 of *SEC16B* (chromosome 1q25.2) to exon 1–8 of *FRK* resulting in both a chimeric transcript including all exons of *FRK* and a truncated transcript starting with an alternative TSS in the intron 2 (figure 3C). All the fusions were in phase at the nucleotide level. In all the chimeric transcript, the tyrosine kinase domain of *FRK* (exons 4–8) was retained whereas the auto-inhibitory SH2 and SH3 domains were absent in the predicted aberrant protein. The expression level of *FRK* was not increased in IHCA with *FRK* fusions compared with other HCA (figure 3B).

Recurrent *IL6* gene rearrangements in IHCA
Finally, we identified two IHCA with abnormal 3' untranslated region (UTR) of *IL6* on RNA sequencing (figure 4A) and characterised by activation of JAK/STAT pathway (GSEA analysis, online supplementary figure 3). One case was previously published (#2615T with inflammatory SAA amyloidosis); the tumour showed a deletion of 3'UTR of *IL6* due to an inversion of 18.7 megabases validated by whole genome sequencing.[12] We identified a second IHCA, #1215T, with a truncation of the 3'UTR regulatory region of *IL6* due to a eight megabases deletion identified using RNA sequencing. In contrast to the first case, #1215T was not associated with SAA amyloidosis symptoms. In these two IHCA, *IL6* was strongly overexpressed (figure 4B, median FPKM=56.4) compared with others IHCA (median FPKM=2.01×10^{-1}, p value=2.6×10^{-2}).

Finally, six IHCAs analysed by WES and/or RNA sequencing remained without identifiable driver genetic alterations. Interestingly, on pathological reviewing, 5 among these 6 IHCA (83%) showed haemorrhage on histology (vs 58% in the whole series of IHCA) with a minimal expression of *CRP* and *SAA* at the transcriptomic level.

DISCUSSION
IHCA serves as a unique model to study the inflammatory pathways in mature hepatocytes and to identify new drivers of oncogene-induced inflammation in benign liver tumourigenesis. Based on the analysis of a large number of IHCA (see flow chart,
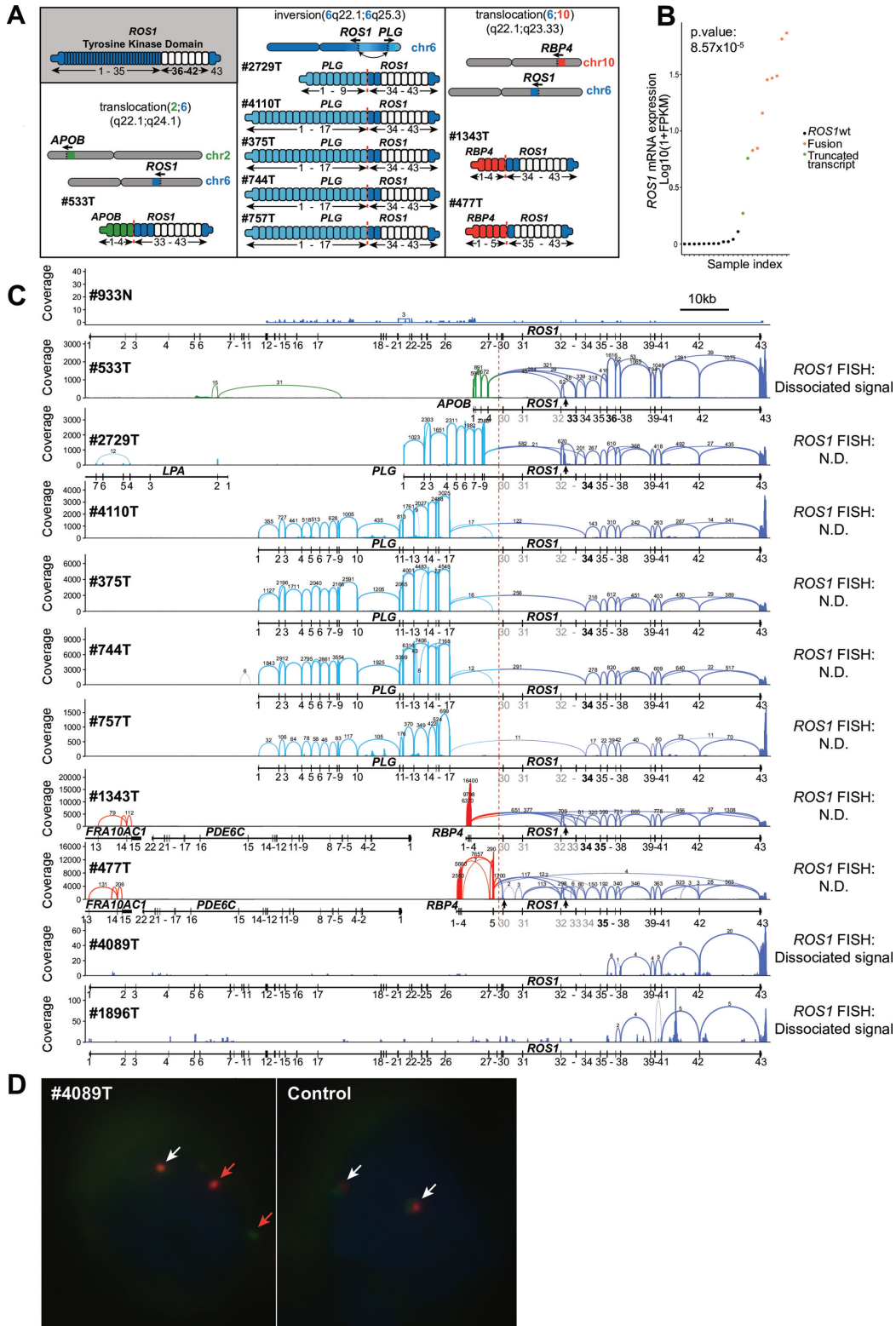
Figure 2 Recurrent *ROS1* fusions in inflammatory hepatocellular adenomas (IHCAs). (A) Representation of the *ROS1* fusion in 8 IHCAs whereas 2 IHCAs with *ROS1* truncated transcript were not represented. IHCA, inflammatory hepatocellular adenoma. (B) Expression level of *ROS1* in FPKM (log 10) in 22 hepatocellular adenoma analysed by RNA sequencing (10 IHCA with *ROS1* fusion/truncated transcripts and 12 IHCA with other genetic alterations). The p value compare the level of expression of *ROS1* between IHCA with *ROS1* alterations versus other IHCA using the Mann-Whitney non-parametric test. (C) Sashimi plots showing RNA-seq alignments for the wild-type *ROS1* in a normal liver (#933N) (top) and 10 cases of IHCA with *ROS1* fusion (bottom). Splice junction reads are represented as arcs connecting a pair of exons. Arc width is proportional to the number of reads aligning to the junction. Grey exon number shows exons not involved in chimeric transcripts and bold exon number represents the *ROS1* exons involved in the in-frame junction of the chimeric transcript. Black arrow indicates stop of transcription following abnormal junction. (D) FISH using *ROS1* dual-colour break-apart probe in 1 IHCA with known *ROS1* fusion and one control. White arrows indicate wild-type *ROS1* with combined red and green signals. Red arrows indicate dissociated signals validating *ROS1* rearrangement.
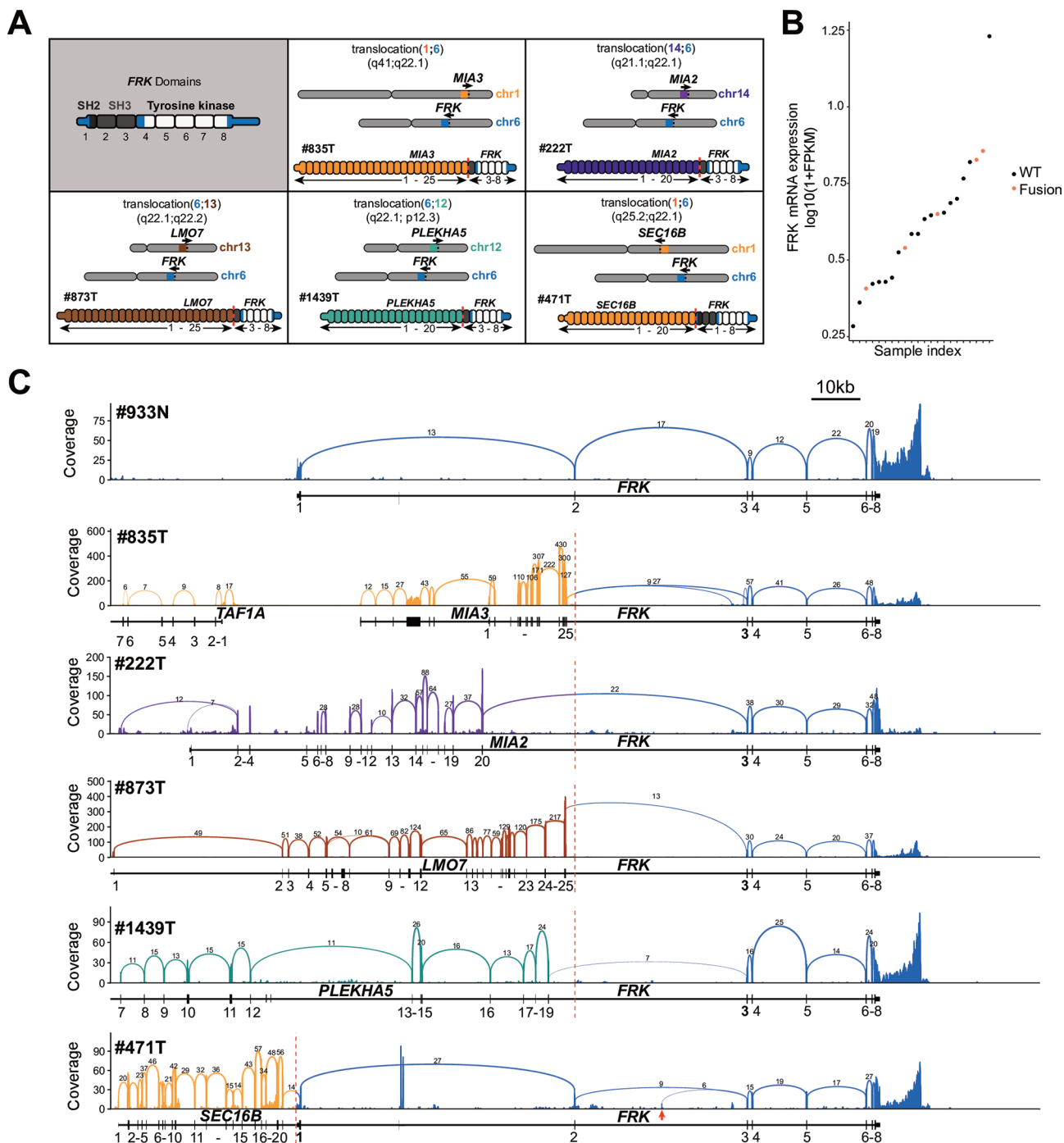
Figure 3 Recurrent *FRK* fusions in inflammatory hepatocellular adenomas (IHCAs). (A) Representation of the *FRK* fusions in five IHCAs. (B) Expression level of *FRK* in FPKM (log 10) in 22 HCA analysed by RNA sequencing. (C) Sashimi plots showing RNA-seq alignments for the wild-type *FRK* in a normal liver (#933N) (top) and five cases of IHCA with *FRK* fusion (bottom). Splice junction reads are represented as arcs connecting a pair of exons. Arc width is proportional to the number of reads aligning to the junction. Bold exon number represents the *ROS1* exons involved in the in-frame junction of the chimeric transcript. Red arrow indicates alternative transcription start site in the *FRK-SEC16B* fusion.

online supplementary figure 1), we identified novel recurrent somatic chromosomal rearrangements of *ROS1*, *FRK* and *IL6*.

In our study, IHCA is the most frequent molecular subgroup with 45% of HCA harbouring an inflammatory phenotype. Overall, we now described in IHCA a total of seven genes that are able to activate the JAK/STAT pathway in hepatocytes when mutated and therefore are associated with benign liver tumourigenesis.[22] Some of these oncogenes belong directly to the JAK/STAT pathway such as *IL6ST* (61% of mutations), *STAT3* (5%

of mutations), *JAK1* (2% of mutations) or *IL6* (1% of alterations).[6 10] However, the remaining three genes, *FRK* (8% of mutations, 2% of fusions, mutually exclusive), *GNAS* (3% of mutations) and *ROS1* (3% of fusions), belong to other signalling pathways but are also known to activate the IL-6/JAK/STAT pathway in cellulo (figures 1A and 5).[6 11 23] Though in most of the IHCA, gene mutations were exclusive, in two cases, we observed a concomitant mutation of *STAT3* L78R with either a classical *FRK* or *STAT3* mutations. We had previously shown,
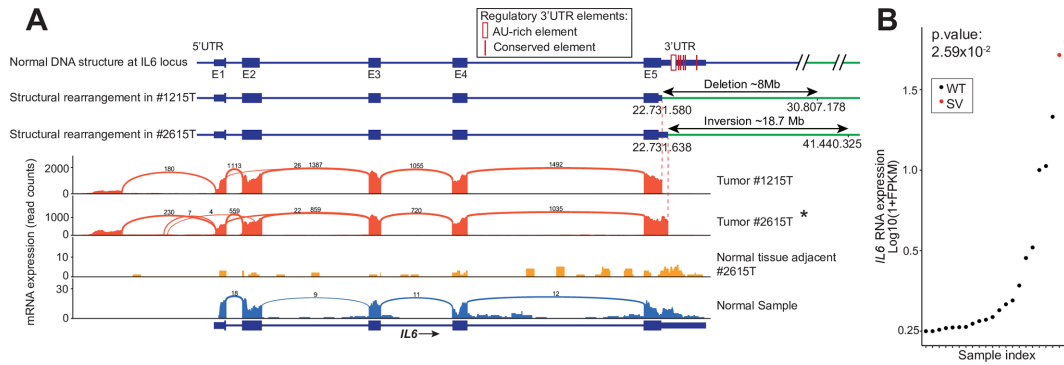
Figure 4 *IL6* rearrangements in inflammatory hepatocellular adenomas (IHCAs). (A) Sashimi plots showing RNA-seq alignments for the wild-type *IL6* in a normal liver and a corresponding non-tumour liver (bottom) and two cases of IHCA with truncation of 3'UTR (top). Splice junction reads are represented as arcs connecting a pair of exons. Arc width is proportional to the number of reads aligning to the junction. *already published[12]. (B) Expression level of *IL6* in FPKM (log 10) in 22 HCA analysed by RNA sequencing. The p value compared the level of expression of *IL6* between IHCA with *IL6* alterations versus other IHCA using the Mann-Whitney non-parametric test. Coordinates in GRCh38. FPKM, fragments per kilobase of exon model per million reads mapped; IHCA, inflammatory hepatocellular adenoma; WT, wild type; SV, structural variants.

in cellulo, that *STAT3* L78R mutations alone can only weakly activate the JAK/STAT pathway.[24] Another study suggested that the L78R mutation, located in the N-terminal domain, could prolong the activation of *STAT3*.[25] We could hypothesise that *STAT3* L78R mutation could cooperate with additional mutation in the JAK/STAT pathway to induce an inflammatory response in mature hepatocytes.

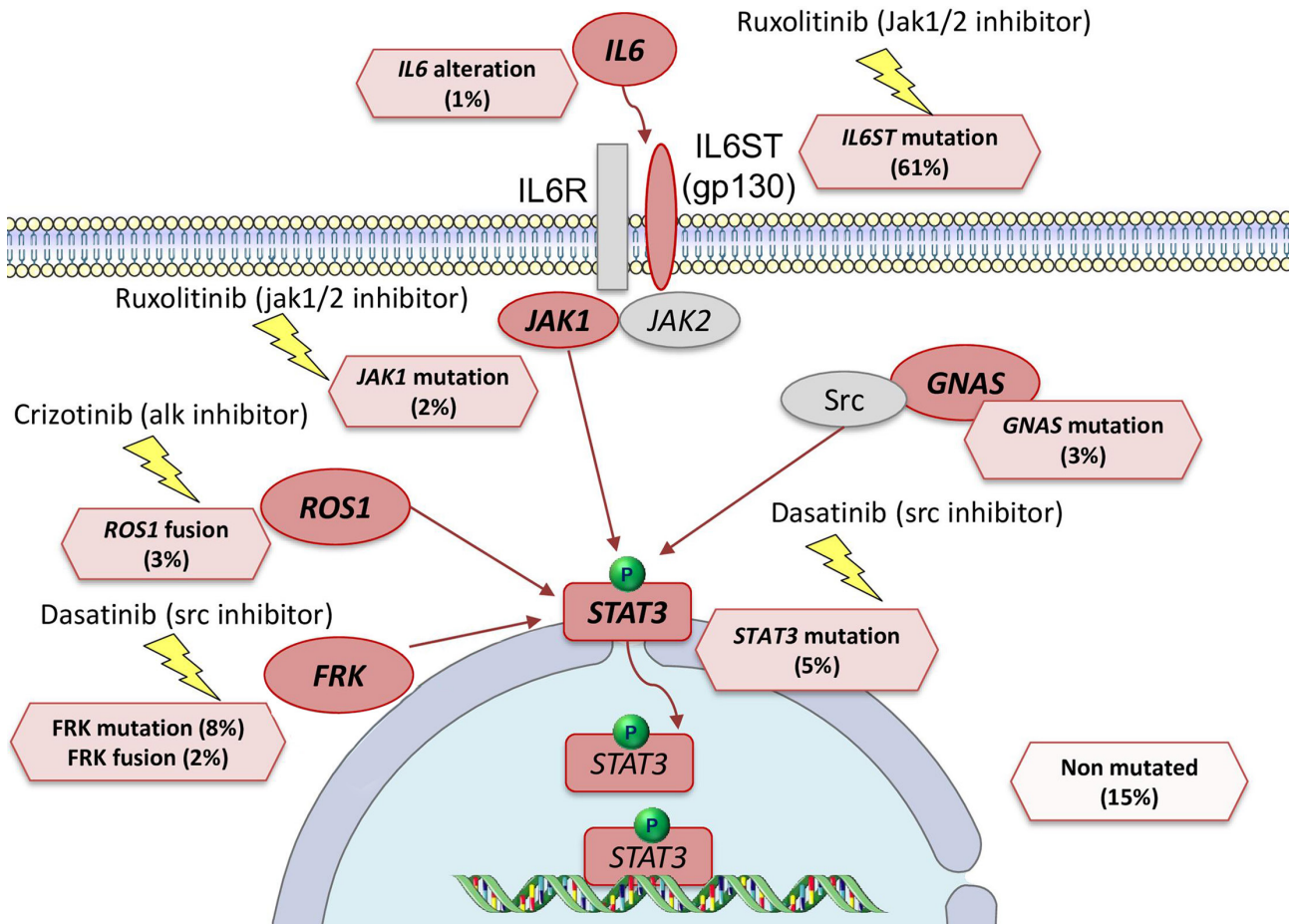# Driver genetic alterations in inflammatory hepatocellular adenomas

Figure 5 Mutations in oncogenes in inflammatory hepatocellular adenomas (IHCAs). Frequencies of somatic genetic alterations in driver genes observed in IHCA are indicated. Targeted therapies potentially active against each genetic alteration are indicated by a yellow flag. IHCA, inflammatory hepatocellular adenoma.

In this work, we identified recurrent *ROS1* rearrangements in 10 IHCA without other mutations in driver genes. *ROS1* is a tyrosine kinase receptor composed of a N-terminal extracellular domain, a transmembrane domain and a C-terminal intracellular tyrosine kinase domain, known to activate several downstream pathways such as the MAP kinase and AKT/mTOR pathways.[26 27] Recurrent *ROS1* rearrangements are identified in 2% of non-small cell lung carcinoma and in other cancer types including cholangiocarcinoma, lymphoma, glioblastoma, colorectal and gastric carcinoma.[26] In anaplastic lymphoma and in inflammatory myofibroblastic tumours, *ROS1* fusions induce a strong activation of the JAK/STAT pathway.[23 28] To our knowledge IHCA is the second type of benign tumour showing *ROS1* fusion after Spitz nevi.[29] This observation reinforces the hypothesis that *ROS1* fusion is an early event during the mechanism of carcinogenesis in solid tumours but it is not sufficient alone to promote a full malignant transformation of hepatocyte or melanocyte.

In IHCA, exons 33–35 of *ROS1* are fused in-phase with different partners including *PLG*, *APOB* and *RBP4*. All these partners are highly expressed in adult liver whereas *ROS1* is almost not expressed in normal hepatocytes. Resulting from the gene fusion, *ROS1* is under the control of the promoter of the different partner genes and consequently overexpressed specifically in all IHCA harbouring the *ROS1* rearrangement. Interestingly, all *ROS1* fusions retained the intracellular tyrosine kinase domain (exon 36–42) and lost the transmembrane domain. These findings are similar to all the *ROS1* fusion proteins already identified in lung cancer, lymphoma or cholangiocarcinoma that all retained the tyrosine kinase domain, essential for its oncogenic activity.[23 30 31] In cellulo model we have also demonstrated that *ROS1* fusions were able to activate *STAT3* in haematological cancer.[23] Moreover, most of *ROS1* fusions described in cancer lack their transmembrane domain leading to the relocalisation of the fusion proteins in the cytoplasm that could interact with different proteins and modified downstream signalling pathway.[27 32] In two of the IHCA cases, though we identified a truncated overexpressed *ROS1* mRNA, its partner genes could not be identified owing to the poor quality of RNA. However, *ROS1* breaks were confirmed by FISH in both the cases.

We previously identified in IHCA recurrent mutations of *FRK* able to activate JAK/STAT pathway in cellulo.[6] In this study, recurrent *FRK* fusions were identified in five cases of IHCA, all showed an in-frame fusion gene. The fusion transcript was predicted to code for a chimeric protein retaining the tyrosine kinase domain, encoded by exons 4–8, but lacking the SH2 and SH3 auto-inhibitory domains, encoded by exons 1–3.[33] Interestingly, *FRK* fusions have been described in acute myeloid leukaemia with *ETV6-FRK* and in anaplastic large cell lymphoma with *CAPRIN-FRK*, all with a similar junctions at exon 3 of *FRK*[34 35]. *In cellulo* model showed that *ETV6-FRK* fusion protein has a constitutive autophosphorylation on its tyrosine residues with an elevated kinase activity.[35] All these data suggested that in hepatocyte like in haematological diseases, the deletion of the SH2/SH3 domain together with an intact tyrosine kinase domain is critical for the oncogenic mechanism of *FRK* fusions.

We also described two cases of somatic alterations of *IL6* with truncation of the 3'UTR due to large chromosome inversion or deletion at the genomic level. These two IHCA harboured a strong overexpression of *IL6* suggesting that the lacks of the 3'UTR regulatory elements induced an overexpression of the transcript. Interestingly, the IL-6 mRNA has a short half-life of around 30 min that is regulated by the 3'UTR.[36] Consequently, the truncation of the 3'UTR in these two IHCA could lead to an increased stability and accumulation of the transcript. We previously described one of the cases that was associated with occurrence of inflammatory amyloidosis,[12] but the second case did not.

In our series, we identified *IL6* alterations, *ROS1* and *FRK* fusions in 3%, 2% and 1% of all HCA, respectively. In all these HCA, no other driver mutations were identified except one *CTNNB1* mutation exon 3 in one case and one *CTNNB1* mutation exon 7 in another case. As *ROS1* and *IL6* rearrangements induced an mRNA overexpression, we did not identify an additional candidate IHCA for these genetic alterations after screening of the whole series of IHCA by quantitative RT-PCR. In contrast, the exact frequency of *FRK* rearrangement could be under-estimated.

In six IHCAs, no functional genetic alterations were identified using next generation sequencing. However, in 5 of these 6 cases, we observed haemorrhage at histological level which is known to induce inflammation and thus a slight increase in SAA and CRP expression in these adenomas without a genomic driven activation of the *IL6/JAK/STAT3* pathway.

Finally, IHCA harboured several driver genes that could be used as a target for drug therapy in the future in clinical practice. *ROS1* fusion could be targeted by crizotinib, a small molecule currently approved by Food and Drug Administration (FDA) for the treatment of lung cancer with *ROS1* fusion.[37] Moreover, we previously showed that dasatinib, a src inhibitor, shutdown JAK/STAT activation due to *FRK* mutations *in cellulo*.[6] We can predict that the same results could be obtained in IHCA with *FRK* fusions. These data suggested that targeted therapies adapted to genetic alterations (ruxolitinib in *IL6ST* mutation, dasatinib for *FRK* mutations/fusions and crizotinib for *ROS1* fusion) could be tested in selected patients with multiple or unresectable IHCA (figure 5).[38]

In conclusion, we described recurrent chromosomal alterations of *ROS1*, *FRK* and *IL6* as new mechanisms of JAK/STAT pathway activation driving benign tumourigenesis in IHCAs that could be tested as therapeutic targets in future clinical trials.

Short summary

We identified recurrent chromosomal alterations involving *ROS1*, *FRK* or *IL6* in IHCAs. *ROS1* activation is a novel mechanism of JAK/STAT pathway activation in human benign liver tumours.

Author affiliations
[1]Centre de Recherche des Cordeliers, Sorbonne Université, Inserm, Université de Paris, Université Paris 13, Functional Genomics of Solid Tumors laboratory, F-75006 Paris, France
[2]Service de Pathologie, Hôpital Pellegrin, CHU de Bordeaux, F 33076 Bordeaux, France
[3]Université Bordeaux, UMR1053 Bordeaux Research in Translational Oncology, BaRITOn, F-33076 Bordeaux, France
[4]Service d'anatomopathologie, Hôpital Beaujon, Assistance-Publique Hôpitaux de Paris, Clichy, France
[5]INSERM U1149, Clichy, France
[6]Service d'anatomopathologie, CHU, Grenoble, France
[7]Service d'anatomopathologie, CHU, Tours, France
[8]Service d'anatomopathologie, CHU Bicêtre, Assistance-Publique Hôpitaux de Paris, Bicêtre, France, Bicêtre, France
[9]Service d'anatomopathologie, Hôpital Henri Mondor; Université Paris Est, Inserm U955, Team 18, Institut Mondor de Recherche Biomédicale, France, Créteil, France
[10]Service d'hépatologie, Hôpital Jean Verdier, Hôpitaux Universitaires Paris-Seine-Saint-Denis, Assistance-Publique Hôpitaux de Paris, Bondy, France
[11]Unité de Formation et de Recherche Santé Médecine et Biologie Humaine, Université Paris 13, Communauté d'Universités et Etablissements Sorbonne Paris Cité, Paris, France, Paris, France
[12]Hôpital Européen Georges Pompidou, F-75015, Assistance Publique-Hôpitaux de Paris, Paris, France

ORCID iDs
Sandrine Imbeaud http://orcid.org/0000-0001-8439-6732
Jean-Charles Nault http://orcid.org/0000-0002-4875-9353

REFERENCES
1. European Association for the Study of the Liver (EASL). EASL clinical practice guidelines on the management of benign liver tumours. *J Hepatol* 2016;65:386–98.
2. Nault J-C, Paradis V, Cherqui D, *et al*. Molecular classification of hepatocellular adenoma in clinical practice. *J Hepatol* 2017;67:1074–83.
3. Zucman-Rossi J, Jeannot E, Nhieu JTV, *et al*. Genotype-Phenotype correlation in hepatocellular adenoma: new classification and relationship with HCC. *Hepatology* 2006;43:515–24.
4. Bluteau O, Jeannot E, Bioulac-Sage P, *et al*. Bi-Allelic inactivation of TCF1 in hepatic adenomas. *Nat Genet* 2002;32:312–5.
5. Pelletier L, Rebouissou S, Paris A, *et al*. Loss of hepatocyte nuclear factor 1alpha function in human hepatocellular adenomas leads to aberrant activation of signaling pathways involved in tumorigenesis. *Hepatology* 2010;51:557–66.
6. Pilati C, Letouzé E, Nault J-C, *et al*. Genomic profiling of hepatocellular adenomas reveals recurrent FRK-activating mutations and the mechanisms of malignant transformation. *Cancer Cell* 2014;25:428–41.
7. Nault J-C, Couchy G, Balabaud C, *et al*. Molecular Classification of Hepatocellular Adenoma Associates With Risk Factors, Bleeding, and Malignant Transformation. *Gastroenterology* 2017;152:880–94.
8. Bioulac-Sage P, Rebouissou S, Thomas C, *et al*. Hepatocellular adenoma subtype classification using molecular markers and immunohistochemistry. *Hepatology* 2007;46:740–8.
9. Paradis V, Champault A, Ronot M, *et al*. Telangiectatic adenoma: an entity associated with increased body mass index and inflammation. *Hepatology* 2007;46:140–6.
10. Rebouissou S, Amessou M, Couchy G, *et al*. Frequent in-frame somatic deletions activate gp130 in inflammatory hepatocellular tumours. *Nature* 2009;457:200–4.
11. Nault JC, Fabre M, Couchy G, *et al*. GNAS-activating mutations define a rare subgroup of inflammatory liver tumors characterized by STAT3 activation. *J Hepatol* 2012;56:184–91.
12. Calderaro J, Letouzé E, Bayard Q, *et al*. Systemic AA amyloidosis caused by inflammatory hepatocellular adenoma. *N Engl J Med* 2018;379:1178–80. 20.
13. Rebouissou S, Franconi A, Calderaro J, *et al*. Genotype-Phenotype correlation of CTNNB1 mutations reveals different ß-catenin activity associated with liver tumor progression. *Hepatology* 2016;64:2047-2061.
14. Nault JC, Mallet M, Pilati C, *et al*. High frequency of telomerase reverse-transcriptase promoter somatic mutations in hepatocellular carcinoma and preneoplastic lesions. *Nat Commun* 2013;4:2218.
15. 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.
16. Van der Auwera GA, Carneiro MO, Hartl C, *et al*. From FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics* 2013;43.
17. Kim D, Pertea G, Trapnell C, *et al*. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 2013;14:R36.
18. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;31:166–9.
19. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol* 2014;15:550.
20. Bayard Q, Meunier L, Peneau C, *et al*. Cyclin A2/E1 activation defines a hepatocellular carcinoma subclass with a rearrangement signature of replication stress. *Nat Commun* 2018;9:5235.
21. Shugay M, Ortiz de Mendíbil I, Vizmanos JL, *et al*. Oncofuse: a computational framework for the prediction of the oncogenic potential of gene fusions. *Bioinformatics* 2013;29:2539–46.
22. Bromberg JF, Wrzeszczynska MH, Devgan G, *et al*. Stat3 as an oncogene. *Cell* 1999;98:295–303.
23. Crescenzo R, Abate F, Lasorsa E, *et al*. Convergent mutations and kinase fusions lead to oncogenic STAT3 activation in anaplastic large cell lymphoma. *Cancer Cell* 2015;27:516–32.
24. Pilati C, Amessou M, Bihl MP, *et al*. Somatic mutations activating STAT3 in human inflammatory hepatocellular adenomas. *J Exp Med* 2011;208:1359–66.
25. Domoszlai T, Martincuks A, Fahrenkamp D, *et al*. Consequences of the disease-related L78R mutation for dimerization and activity of STAT3. *J Cell Sci* 2014;127:1899–910.
26. Davies KD, Doebele RC. Molecular pathways: ROS1 fusion proteins in cancer. *Clin Cancer Res* 2013;19:4040–5.
27. Uguen A, De Braekeleer M. Ros1 fusions in cancer: a review. *Future Oncol* 2016;12:1911–28.
28. Lovly CM, Gupta A, Lipson D, *et al*. Inflammatory myofibroblastic tumors harbor multiple potentially actionable kinase fusions. *Cancer Discov* 2014;4:889–95.
29. Wiesner T, He J, Yelensky R, *et al*. Kinase fusions are frequent in spitz tumours and spitzoid melanomas. *Nat Commun* 2014;5:3116.
30. Takeuchi K, Soda M, Togashi Y, *et al*. Ret, ROS1 and ALK fusions in lung cancer. *Nat Med* 2012;18:378–81.
31. Gu T-L, Deng X, Huang F, *et al*. Survey of tyrosine kinase signaling reveals ROS kinase fusions in human cholangiocarcinoma. *PLoS One* 2011;6:e15640.
32. Pal P, Khan Z. ROS1 [corrected]. *J Clin Pathol* 2017;70:1001–9.
33. Goel RK, Lukong KE. Understanding the cellular roles of fyn-related kinase (FRK): implications in cancer biology. *Cancer Metastasis Rev* 2016;35:179–99.
34. Hu G, Dasari S, Asmann YW, *et al*. Targetable fusions of the FRK tyrosine kinase in ALK-negative anaplastic large cell lymphoma. *Leukemia* 2018;32:565–9.
35. Hosoya N, Qiao Y, Hangaishi A, *et al*. Identification of a SRC-like tyrosine kinase gene, FRK, fused with ETV6 in a patient with acute myelogenous leukemia carrying a t(6;12)(q21;p13) translocation. *Genes Chromosomes Cancer* 2005;42:269–79.
36. Paschoud S, Dogar AM, Kuntz C, *et al*. Destabilization of interleukin-6 mRNA requires a putative RNA stem-loop structure, an AU-rich element, and the RNA-binding protein AUF1. *Mol Cell Biol* 2006;26:8228–41.
37. Shaw AT, Kim D-W, Nakagawa K, *et al*. Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. *N Engl J Med* 2013;368:2385–94.
38. Poussin K, Pilati C, Couchy G, *et al*. Biochemical and functional analyses of gp130 mutants unveil JAK1 as a novel therapeutic target in human inflammatory hepatocellular adenoma. *Oncoimmunology* 2013;2:e27090.