



HAL
open science

Analyse Intégrée génomique, protéomique et radiomique des Sarcomes Pléomorphes Indifférenciés : Identification et Validation de nouvelles cibles thérapeutiques

Maud Toulmonde

► To cite this version:

Maud Toulmonde. Analyse Intégrée génomique, protéomique et radiomique des Sarcomes Pléomorphes Indifférenciés : Identification et Validation de nouvelles cibles thérapeutiques. Médecine humaine et pathologie. Université de Bordeaux, 2019. Français. NNT : 2019BORD0429 . tel-02900286

HAL Id: tel-02900286

<https://theses.hal.science/tel-02900286>

Submitted on 16 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université de Bordeaux
Sciences de la Vie et de la Santé
Doctorat de Biologie Cellulaire et Physiopathologie

Maud TOULMONDE

Née le 20 Juillet 1983

**Analyse Intégrée génomique, protéomique et radiomique des
Sarcomes Pléomorphes Indifférenciés :
Identification et Validation de nouvelles cibles thérapeutiques**

Thèse dirigée par M. le Professeur Antoine ITALIANO, MD, PhD

Soutenue publiquement le 20 Décembre 2019

Jury

M. le Professeur Jean-Michel COINDRE, MD, PhD

Président du Jury

M^{me} le Professeur Florence PEDEUTOUR, MD, PhD

M^{me} Laurence BIANCHINI, PhD

M. Franck TIRODE, PhD

**Analyse Intégrée génomique, protéomique et radiomique des
Sarcomes Pléomorphes Indifférenciés :
Identification et Validation de nouvelles cibles thérapeutiques**

Les Sarcomes indifférenciés pléomorphes (UPS pour Undifferentiated Pleomorphic Sarcoma) forment un groupe hétérogène « par défaut ». Nous avons émis l'hypothèse qu'il existe un lien entre le niveau de dédifférenciation des UPS et l'infiltrat immun intra-tumoral, et que cette relation repose sur des altérations génomiques ainsi que sur l'activation de voies de signalisation spécifiques associées à un impact thérapeutique potentiel. Les objectifs de ce travail étaient de générer une classification des UPS intégrant une approche génomique, immunophénotypique, protéomique et radiomique, et d'identifier et tester de potentielles cibles thérapeutiques sur des lignées cellulaires et des xénogreffes dérivées de tumeurs de patients (PDX) caractérisées génomiquement.

Nous avons réuni 135 tumeurs de patients de type UPS dans notre institution, parmi lesquels 25 ont été sélectionnées pour séquençage de l'ARN et de l'exome entier. L'analyse hiérarchisée non supervisée des données de séquençage ARN identifia trois groupes : A, B et C. Le groupe A était principalement enrichi en gènes jouant un rôle dans le développement et le phénotype « cellule souche », notamment *LHX8*, *LRN1*, *LGR5*, *BMP5* et *FGFR2*. Le group B était fortement enrichi en gènes impliqués dans l'immunité, incluant *MARCO*, *TIMD4*, *TIGIT*, *CD27*, *IFNG*, *CD8B*, *PDCD1*, *CD3D* et *IDO1*, ainsi que *DKK1*. Le groupe C était trop petit pour être caractérisable de façon fiable. Nous avons pu confirmer cette classification sur une cohorte indépendante de 41 UPS du consortium TCGA. Nous avons trouvé une forte corrélation entre l'expression génique et la densité protéique en IHC sur les échantillons tumoraux correspondants pour CD8, PD-1 et IDO1, amenant à appeler le groupe B « immune-high » et le groupe A « immune-low ». Dans une cohorte indépendante de validation de 110 UPS, la densité de CD8 était significativement associée à une meilleure survie sans métastase ($p = 0.04$). Nous avons retrouvé une fréquence significativement plus importante de variation du nombre de copies de gènes dans le groupe « immune-low », principalement des délétions, touchant notamment *PTEN*, *RBI*, *FANCA*, *FAS*, *CDKN2A*, *TP53*, *AXINI*, *NF2* et *BRCA2*. Les analyses protéomiques permirent de détecter deux principaux groupes - PA and PB – fortement corrélés avec les deux principaux groupes identifiés en transcriptomique – A « immune low » et B « immune-high ». Le groupe PB était significativement enrichi en protéines appartenant aux voies de la réponse immunitaire, et le groupe PA en protéines appartenant aux cibles de MYC et à la transition épithélio-mésenchymateuse. Nous avons ensuite développé des lignées cellulaires et des PDX issues de tumeurs provenant de l'étude de profilage moléculaire et obtenu des modèles représentant chaque groupe A, B et C. L'inhibiteur de FGFR JNJ-42756493 avait une activité anti-tumorale *in vitro* et *in vivo* dans les lignées et les modèles de PDX du groupe A, de façon sélective. Nous avons également montré *in vitro* une sensibilité sélective des lignées cellulaires du groupe A aux trois inhibiteurs de protéines à bromodomaines BET CBP/P300 : CPI637, NEO1132 and NEO2734. Enfin, nous avons identifié un set de neuf variables de texture issues de séquences d'IRM conventionnelles hautement corrélées à notre classification transcriptomique des UPS, permettant de poser la base d'une signature radiomique pour sélectionner les UPS « immune-high » sur leur imagerie pré-thérapeutique.

Cette étude est la première à réaliser un portrait intégré d'UPS primaires non pré-traités. Nous avons pu identifier deux groupes principaux, avec des implications thérapeutiques potentielles : le groupe « immune-high », fortement inflammé et probablement le meilleur candidat à l'immunothérapie, et le groupe « immune-low », avec un rationnel pour l'utilisation d'inhibiteurs de FGFR et de BET dans ce dernier.

Mots-Clefs : Sarcomes, Génomique, Immunologie, Pré-clinique, Cible thérapeutique

Genomic Characterization and Target Functional Validation In Undifferentiated Pleomorphic Sarcomas: an Integrated Approach

Undifferentiated Pleomorphic Sarcoma (UPS) are an heterogeneous group of poorly differentiated tumors made up 'by default'. We hypothesized that there is a link between dedifferentiation state of UPS and immune infiltrate and that this relation relies on specific pathways activation and related genomics alterations with potential therapeutic impact. Objectives of this work were to generate a comprehensive Omics landscape of UPS, integrating genomic, immuno-phenotypic, proteomic and radiomic approach, and to identify and test potential targets for therapeutic approach on cell lines and patients tumor mouse xenografts (PDX).

We analyzed a cohort of 135 UPS samples from patients in our institution, of whom 25 were selected for full exome and RNA-sequencing. Unsupervised consensus and hierarchical clustering of RNA-sequencing identified 3 groups, A, B and C. Group A was mainly enriched in genes that play a crucial role in both normal development and stemcellness, notably *LHX8*, *LRRN1*, *LGR5*, *BMP5* and *FGFR2*. Group B was strongly enriched in genes involved in immunity, including *MARCO*, *TIMD4*, *TIGIT*, *CD27*, *IFNG*, *CD8B*, *PDCD1*, *CD3D* and *IDO1*, but also *DKK1*. Group C was too small to be analyzed with sufficient robustness. This classification was confirmed on an independent cohort of 41 UPS from TCGA consortium.

We found a high correlation between gene expression and protein density by IHC on related tumor sample slides for CD8, PD-1 and IDO1, leading to call group B '*immune-high*' and group A '*immune-low*'. In an independent validation cohort of 110 UPS patients, CD8 expression was significantly associated with metastase-free survival ($p = 0.04$).

Copy numbers variations were significantly more frequent in the *immune-low* group. Main recurrent events were deletions, notably in *PTEN*, *RBI*, *FANCA*, *FAS*, *CDKN2A*, *TP53*, *AXINI*, *NF2* and *BRCA2*. Proteomic analysis allowed us to detect two main proteomic groups - PA and PB – that highly correlated with the two main transcriptomic groups - A and B. Group PB was significantly enriched in immune response pathways, whereas group PA was enriched in MYC targets and epithelial-mesenchymal transition pathways.

We then further developed cell lines and PDX models from patient tumor samples included in the molecular profiling study for each class, A, B, C. We showed robust *in vitro* and *in vivo* anti-tumor activity of FGFR inhibitor JNJ-42756493 in cell lines and PDX models from group A, selectively. We also showed *in vitro* activity of three potent dual inhibitors of BET-proteins CBP/P300, CPI637, NEO1132 and NEO2734, in cell lines from group A, selectively.

Finally, we showed that a set of 9 radiomic features from basic MRI conventional sequences correlated well with our UPS molecular classification and provided the basis for a radiomics signature that could select *immune-high* UPS on their pretherapeutic imaging.

This study is the first to give a comprehensive genomic, immuno-phenotypic, proteomic and radiomic landscape of non-pretreated primary UPS. We identified two main groups of UPS with therapeutics potential: the *immune-high* group, strongly inflamed and probably the best candidate for immunotherapy, and the *immune-low* group, with a rational for FGFR and BET inhibitors activity in this one.

Key-Words: Sarcoma, Genomic, Immunology, Pre-Clinical, Therapeutic target

Ce travail a été mené au sein de l'équipe de laboratoire du Pr Antoine ITALIANO :

Equipe SARCOTARGET

Unité INSERM 1218 ACTION

Actions for onCogenesis understanding and Target Identification in Oncology

INSTITUT BERGONIE

229 cours de l'Argonne

33000 BORDEAUX

Remerciements

Nous tenons tout d'abord à remercier les membres du jury qui nous font l'honneur de juger ce travail.

M^{me} le Pr Florence Pedoutour, M^{me} Laurence Bianchini, M. Franck Tirode, experts en biologie des sarcomes, merci pour votre disponibilité et l'évaluation de ce travail.

M. le Pr Jean-Michel Coindre, merci de votre confiance. Vous êtes un mentor, votre carrière de renommée internationale faite d'expertise, d'exigence, de curiosité et de collaborations est un exemple.

M. le Pr Antoine Italiano, mon mentor numéro 1, celui qui m'a inoculé les virus du sarcome et de la recherche clinique, celui par et pour lequel j'ai fait ces choix de vie professionnelle. Tu es mon modèle et mon moteur, ton intelligence et ta passion dans ton travail me guident et me challengent au quotidien, et je t'en remercie.

Ce travail est le fruit de multiples collaborations et échanges, passés, en cours, mais aussi à venir. C'est ici l'occasion de remercier tous ceux qui y ont contribué, un peu, beaucoup, passionnément...

Carlo, par où commencer ? Merci pour tous ces fichiers excels, ces ppt, ces docs, ces old, ces new, ces V1, 2, 3, 4, 5, 6, 7, 8 ! Merci pour toutes nos discussions, et ce « great achievement » qui ne serait pas sans toi !

Stef, par où continuer ? Pertinence, vivacité, efficacité, tu es la 2^{ème} jambe de ce travail, merci d'avoir tout fait si bien, si vite.

Julien, merci d'avoir pris le temps dans ton agenda très rempli pour cette collaboration aux multiples étapes et pour nos discussions.

Amandine, merci de ton intérêt pour notre travail et pour ton apport concernant l'obscur (mais contrasté) monde de la radiomique.

Audrey, pour ta supervision sans faille, Vanessa, Jessica, toute l'équipe bio-info - notamment Damien et Frédéric - Laura pour le dernier WB, Raul pour ta contribution aux marquages FGFR2.

Valérie, toujours disponible, même aux pires moments, et experte ès IHC internationale.

Alban Bessede pour ton expertise de la voie IDO/kynurénine et ton intérêt pour les sarcomes.

Christophe Rey, tu as participé à ce projet dès le début, et pas forcément aux moments les plus simples !

Laurence Bianchini et son équipe pour les RT-qPCR.

L'équipe de Nantomics pour la protéomique.

François Bertucci pour nos échanges et cette opportunité de collaboration qu'il faudrait creuser.

Florent Petitprez pour nos échanges et cette poster discussion de l'ESMO 2019 en commun.

Tous mes collègues de l'Institut Bergonié, qui m'ont aidée à me rendre disponible pour ce travail de recherche et m'ont soutenue dans mon activité clinique, le sens numéro un de mon investissement professionnel.

Ma famille et mes amis, qui m'ont soutenue dans les moments difficiles, les doutes, les découragements et aussi offert de nombreux moments de vie joyeuse au cours de ces dernières années plutôt intenses.

Table des matières

ABREVIATIONS	12
INTRODUCTION	13
DES MFH AUX UPS, HISTORIQUE ANATOMO-PATHOLOGIQUE	13
LES SARCOMES DES TISSUS MOUS	13
LES MFH	14
CHANGEMENTS TAXONOMIQUES	15
GENOMIQUE ET PHYLOGENIE DES UPS, ETAT DES LIEUX A L'ORIGINE DE L'ETUDE	16
CARACTERISATION GENOMIQUE DES UPS	16
HYPOTHESES PHYLOGENIQUES DES UPS	25
EN RESUME	29
HYPOTHESES ET OBJECTIFS	30
METHODES	31
RESULTATS	31
ARTICLE	33
MATÉRIEL SUPPLÉMENTAIRE	59
DISCUSSION DES RESULTATS ET PERSPECTIVES	83
GENOMIQUE	83
CLASSIFICATION TRANSCRIPTOMIQUE, PROTEOMIQUE ET IMPACT FONCTIONNEL	89
IMMUNE-HIGH UPS	89
IMMUNE-LOW UPS	97
RADIOMIQUE	107
LIMITES ET REMARQUES	108
CONCLUSION	110
ANNEXES	112
REFERENCES	113

Abréviations

CGH : array-Comparative Genomic Hybridization

CMH : Complexe Majeur d’Histocompatibilité

CNS : Copy Number Signature

CNV : Copy Number Variation

DDLPS : Dedifferentiated Liposarcoma

EMA : European Medicine Agency

EMT : Epithelial-Mesenchymal Transition

FACS : Fluorescence Activated Cell Sorting

FDA : Food and Drug Administration

FISH : Fluorescent In Situ Hybridization

GIST : Gastro-Instestinal Stromal Tumor

GSF-GETO : Groupe Sarcome Français – Groupe d’Etude des Tumeurs Osseuses

IDO : Indoleamine 2,3-DiOxygenase

IHC : Immuno-Histo-Chimie

LOH : Loss Of Heterozygosity

Mb : Megabase

MFH : Malignant Fibrous Histocytoma

MMR : MisMatch Repair

MPNST : Malignant Peripheral Nerve Sheat Tumor

NK : Natural Killer

PCA : Principal Component Analysis

PDX : Patient-Derived Xenograft

TCGA : The Cancer Genome Atlas

WHO : World Health Organization

Introduction

Des MFH aux UPS, historique anatomo-pathologique

Les sarcomes des tissus mous

Les sarcomes des tissus mous sont un groupe de tumeurs mésenchymateuses hétérogènes représentant 1% des cancers de l'adulte. Ces tumeurs sont classées majoritairement en fonction de leur ligne de différenciation vers un tissu d'origine mésenchymateuse (adipocytaire, musculaire - lisse ou strié, chondroblastique, ostéoblastique, vasculaire) déterminé par la morphologie et des caractéristiques immuno-histo-chimiques (IHC) (1).

La génétique permet de classer les sarcomes des tissus mous en deux grands types :

- les sarcomes à « génomique simple », associés à des altérations génomiques simples comme des translocations spécifiques. Les synovialosarcomes, caractérisés dans plus de 80% par la translocation $t(X;18)(p11.2;q11.2)$ produisant un gène de fusion *SYT/SSX1* ou *SYT/SSX2*, ainsi que les liposarcomes myxoïdes, caractérisés de façon pathognomonique par la translocation $t(12;16)(q13;p11)$ produisant le gène de fusion *FUS/DDIT3*, appartiennent à ce groupe.

- les sarcomes à « génomique complexe », caractérisés par des aberrations chromosomiques de nombre et de structure nombreuses et non récurrentes et par des mutations fréquentes de *TP53*. Ce groupe est principalement constitué de tumeurs à cellules fusiformes et/ou pléomorphiques, c'est à dire aux noyaux irréguliers, parfois monstrueux, aux nucléoles proéminents, et de formes et de tailles variables.

Les principaux représentants de cette classe sont :

- les léiomyosarcomes, à différenciation musculaire lisse, exprimant en IHC l'actine musculaire lisse, la caldesmone, la transgeline, la calponine et inconstamment la desmine. Les léiomyosarcomes ont un mauvais pronostic, avec un risque de rechute métastatique > 50%.

- Les myxofibrosarcomes, tumeurs mésenchymateuses fibroblastiques caractérisées par un stroma myxoïde. Les myxofibrosarcomes sont les tumeurs les plus fréquentes chez

l'adulte et sont généralement de bas grade, avec un risque élevé de rechute locale mais un risque métastatique plus faible.

- Les Sarcomes indifférenciés/inclassés. Ce terme introduit dans la classification WHO de 2013 regroupe une partie des anciens histiocytomes fibreux malins (MFH, pour Malignant Fibrous Histiocytoma), et regroupe des tumeurs mésenchymateuses non rattachables aux autres catégories établies, et pouvant avoir une prédominance de cellules fusiformes, rondes, épithélioïdes ou encore pléomorphes. Dans ce dernier cas on parle de sarcomes indifférenciés pléomorphes (UPS, pour Undifferentiated Pleomorphic Sarcoma) (2).

- Les liposarcomes dédifférenciés (DDLPS pour Dedifferentiated Liposarcoma) appartiennent aux sarcomes à génomique complexe de par leurs remaniements génétiques complexes et le caractère pléomorphe et/ou fusiforme des cellules constituant leur partie dédifférenciée, mais ils forment un groupe à part, car caractérisé par une amplification constante de la région chromosomique 12q13-15 comportant les gènes *MDM2* et *CDK4*.

Les MFH

Depuis sa première description en 1964 (3), l'entité MFH a été considérée comme le type de sarcome le plus fréquent. Les principales caractéristiques anatomo-pathologiques des MFH étaient alors l'absence de différenciation et la prédominance de pléomorphisme, mais étaient rapportées aussi des variantes inflammatoires, myxoïdes, angiomatoïdes et à cellules géantes (4). Les MFH étaient ainsi un groupe hétérogène de tumeurs défini « par défaut » constituant jusqu'à 30% des sarcomes des séries publiées les plus anciennes (5), mais dont la réalité propre a été remise en question sur les trois dernières décennies (6). Les avancées des techniques utilisées en routine en anatomie pathologique et en biologie moléculaire ont en effet permis de reclasser et parfois démembrer les MFH en de nombreuses autres entités, telles que des liposarcomes, des léiomyosarcomes, des myxofibrosarcomes, mais aussi des carcinomes sarcomatoïdes ou des mélanomes.

La mise en évidence de chromosomes en anneau, de grands fragments du chromosome 12, et d'une amplification de la région 12q13-15 comportant notamment les gènes *MDM2* et *CDK4* dans les liposarcomes bien différenciés (WDLPS) et les DDLPS a permis d'en améliorer le diagnostic (7, 8). Avec la dissémination de l'utilisation en routine de l'IHC et de la technique de FISH (Fluorescence In Situ Hybridization) pour ces gènes (9), il est apparu que la grande majorité des MFH dits « inflammatoires », notamment rétro-péritonéaux, étaient en fait des DDLPS, caractérisés par cette amplification de régions du chromosome 12 contenant *MDM2* et *CDK4* (10-13).

D'autre part, avec l'utilisation plus systématique de marqueurs IHC de différenciation musculaire lisse, un certain nombre de MFH ont également pu être reclassés en léiomyosarcomes, dont l'existence est associée à un plus mauvais pronostic en termes de survie sans métastases et de survie globale (14).

En 1992, Fletcher publia une série de 159 cas « étiquetés » MFH qui, une fois relus sur le plan morphologique et immuno-histo-chimique, étaient pour 13% des tumeurs non mésenchymateuses, pour 63% n'étaient en fait pas des MFH (25% des DDLPS et 18% des léiomyosarcomes), et pour 26% étaient effectivement des sarcomes pléomorphes non classables, la moitié étant des biopsies ou du matériel trop nécrotique pour permettre un autre diagnostic en dehors d'un pléomorphisme nucléaire. Ceci a fait remettre en question l'existence même des sarcomes pléomorphes inclassés.

Changements taxonomiques

En même temps que les MFH inflammatoires étaient reconnus comme des DDLPS avec leur amplification caractéristique, les MFH myxoïdes s'individualisaient en myxofibrosarcomes, et les MFH à cellules géantes disparaissaient en carcinome anaplasique à cellules géantes, léiomyosarcomes avec réaction à cellules géantes, ou ostéosarcomes riches en cellules géantes. Le terme MFH, comportant toujours une partie hétérogène de tumeurs à composante notamment pléomorphe, a alors été dans un premier temps remplacé par « UPS » dans la nouvelle classification WHO de 2002 (15).

Le terme MFH a ensuite été totalement abandonné et le groupe de sarcomes restant indifférenciés et inclassables malgré une analyse morphologique, immuno-histo-chimique et génétique extensive a été reconnu dans la classification WHO de 2013 comme une entité spécifique sous le nom de « sarcomes indifférenciés », avec quatre sous-types distincts en fonction de la morphologie cellulaire prédominante : fusiforme, ronde, épithélioïde ou pléomorphe (16).

Ces sous-types font l'objet de nombreux travaux de recherche ayant pour but d'identifier de nouvelles entités distinctes et homogènes en leur sein.

Génomique et phylogénie des UPS, état des lieux à l'origine de l'étude

Caractérisation génomique des UPS

Les études de génomique, notamment la CGH et la transcriptomique, ont aidé au cours des deux dernières décennies à mieux caractériser et mieux comprendre les altérations distinctes sous-jacentes aux sarcomes à génomique complexe. La majorité ont porté sur les sarcomes en général, incluant MFH, léiomyosarcomes, DDLPS, et ont permis d'avancer des hypothèses en ce qui concerne les relations entre les UPS et les autres sarcomes à génomique complexe, notamment les léiomyosarcomes et les myxofibrosarcomes. Malheureusement, la grande difficulté dans l'interprétation de ces études, notamment les plus anciennes, réside dans la définition variable et la réalité discutable des histologies données comme MFH dans ces séries, empêchant d'y retrouver les « vrais » UPS avec certitudes.

En 2003, l'équipe de Lee étudia l'expression des gènes de 27 échantillons de sarcomes, neuf synoviosarcomes, neuf léiomyosarcomes, et neuf MFH à l'aide d'une plateforme de puce à ADN de 5603 sondes. Ils identifièrent avec fiabilité le groupe des synoviosarcomes, mais les MFH et les léiomyosarcomes étaient répartis en trois groupes : le premier incluait quatre léiomyosarcomes, un synoviosarcome et un MFH, le second trois MFH, et le quatrième était un mélange de léiomyosarcomes et MFH. Ils n'ont pu identifier de voie de signalisation spécifique au groupe de MFH (17).

En 2005, Baird et son équipe étudièrent 181 échantillons représentant 16 classes de sarcomes des tissus mous ou osseux par puce à ADN de 12601 sondes (18). En analyse de regroupement non supervisée à partir des 1527 gènes les plus variablement exprimés dans leur cohorte, les MFH se répartissaient largement dans les différents groupes, soulignant à nouveau la problématique de la réalité des histologies proposées (Figure 1).

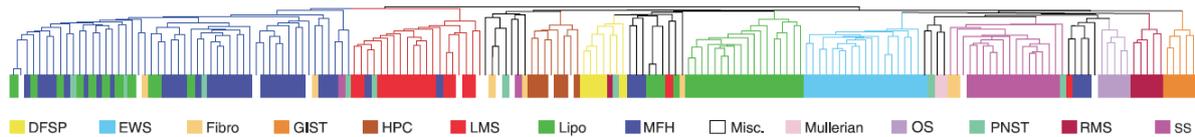


Figure 1 Dendrogramme de l'analyse hiérarchisée non supervisée des données de séquençage ARN des 181 échantillons de sarcomes de l'étude de Baird *et al.* *Cancer Res*, 2005

(DFSP : dermatofibrosarcome de Darrier Ferrand; EWS : sarcome d'Ewing; Fibro : fibrosarcome; GIST : tumeur stromale gastro-intestinale; HPC : hémangiopéricytome; léiomyosarcomes : léiomyosarcome; Lipo : liposarcome; MFH : histiocytome fibreux malin; Misc : autre sarcome inclassé, Mullerian : tumeur mixte mullérienne; OS : ostéosarcome, PNST : tumeur des gaines des nerfs périphériques; RMS : rhabdomyosarcome; SS : synovialosarcome)

L'analyse hiérarchisée de ces données leur permet de déterminer deux groupes de MFH, et l'analyse pondérée des gènes différentiellement exprimés dans ces deux groupes leur permet d'identifier 279 gènes significativement différentiels dans l'un ou l'autre des deux groupes ($p < 0.001$) (Figure 2). Le premier groupe était enrichi en gènes correspondant à des protéines impliquées dans le profil musculaire tels que la tenascine C, le sarcoglycane, la myosine X. De façon intéressante, le second groupe était très enrichi en gènes de l'inflammation tels que *HEM1*, *MX1*, *DAP10*, *PCVG2*, *FOLR3* (Annexes, Table 1).

En 2007, Nakayama et collègues évaluèrent les profils d'expression génique de 105 échantillons de sarcomes de diverses histologies, incluant 21 MFH, au moyen d'une puce à ADN incluant 22283 sondes (Gene-Chip Human Genome U133A array, Affymetrix, Santa Clara, CA, USA) (19). En analyse non supervisée, les sarcomes à cellules fusiformes et pléomorphes se regroupaient en formant des clusters plus ou moins distincts, les MFH se répartissant à nouveau dans différents autres groupes, en dehors d'un cluster propre de six échantillons. Ils étudièrent ensuite les similarités génomiques des MFH avec les cinq autres histologies à partir des profils de gènes différentiellement exprimés dans chacun de ces cinq groupes. Quinze MFH avaient des similarités moyennes à fortes avec d'autres histologies : six avec le groupe myxofibrosarcomes, cinq avec le groupe fibrosarcome, deux avec les tumeurs malignes des gaines des nerfs périphériques (MPNST), un avec les liposarcomes et un dernier avec les léiomyosarcomes.

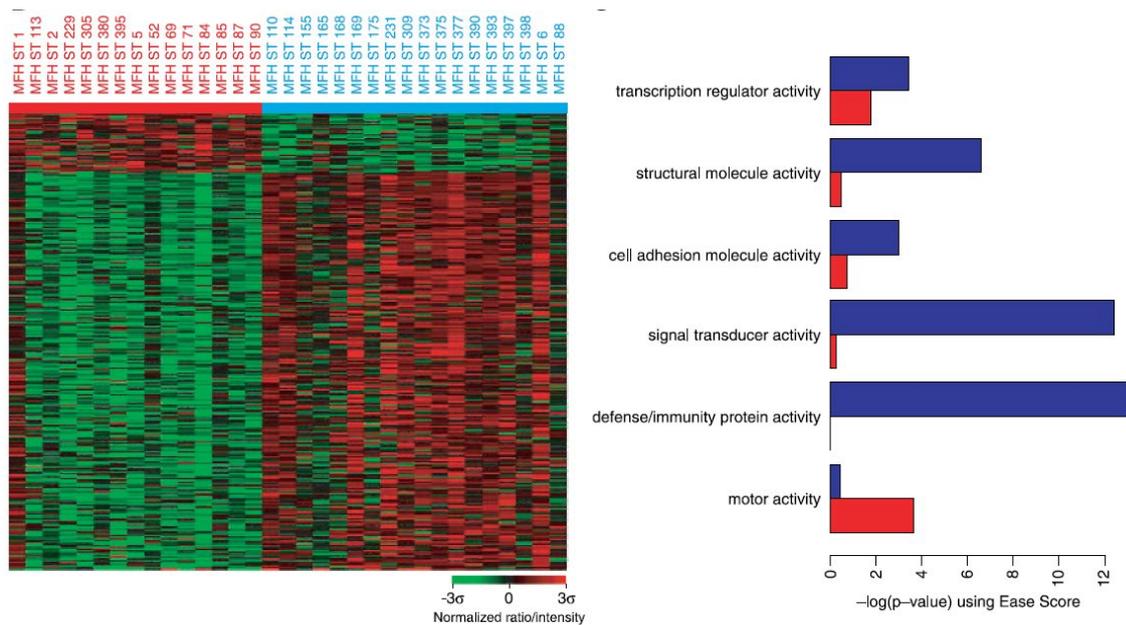


Figure 2 Deux groupes de MFH sont identifiés dans l'étude de Baird *et al. Cancer Res, 2005*

Gauche : *Heat Map* de l'analyse pondérée de l'expression génique dans les deux principaux groupes de MFH en rouge et bleu; Droite : analyses ontologiques des fonctions moléculaires les plus représentées dans chaque groupe de MFH, respectivement en rouge et bleu.

Ils concluent que la majorité des MFH pouvaient être reclassés en une variante à composante pléomorphe de plusieurs sarcomes à composante fibroblastique, mais que six parmi 21 (30%) des MFH de leur cohorte ne présentaient pas de corrélation génomique évidente avec les autres histologies et pourraient représenter une entité de réels « de novo UPS ».

En 2007, l'équipe de Francis utilisa une puce à ADN de 27649 sondes pour proposer une analyse non supervisée des profils d'expression génique de 177 sarcomes, incluant 60 MFH/UPS (20). Les MFH de leur cohorte se répartissaient à nouveau très largement entre les histologies à génomique complexe, regroupées dans le cluster C (Figure 3). Ils identifièrent une signature de 500 gènes caractérisant les MFH/UPS, incluant des gènes de la matrice extracellulaire et la réponse inflammatoire.

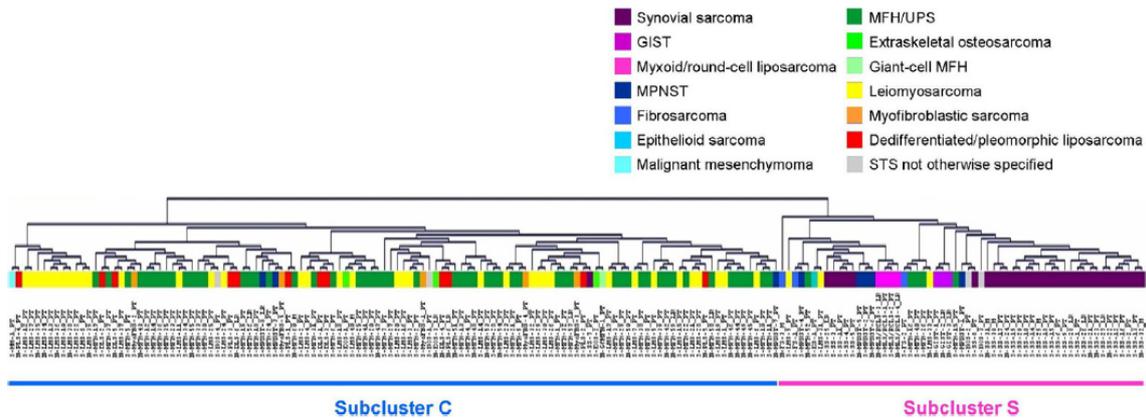


Figure 3 Dendrogramme de l'analyse hiérarchisée non supervisée des données de séquençage ARN des 177 échantillons de sarcomes de l'étude de Francis *et al.* BMC Genomics, 2007

En 2008, Carneiro et collègues ont porté leur intérêt sur les léiomyosarcomes et les UPS et remis à nouveau en question la différence entre ces deux histologies. Ils analysèrent les profils de CGH de 18 léiomyosarcomes et 31 UPS (21). Globalement, la majorité des tumeurs avaient > 30% de leur génome altéré, et 16% des pertes et des gains répartis sur > 50% du génome, soulignant le caractère complexe de la génomique de ces tumeurs. Les altérations du nombre de copies étaient réparties sur l'ensemble des chromosomes, sans différence évidente entre léiomyosarcomes et UPS. Dans cinq tumeurs (trois UPS et deux léiomyosarcomes), une amplification de *MDM2* (12q14.3-q15) a pu être retrouvée, associée pour quatre d'entre elles à une amplification de *CDK4* (12q14.1), révélant l'inclusion de cinq (10%) DDLPS dans cette cohorte. La délétion du 13q, affectant le segment 13q12.11-q31.1, était la plus commune. La plus fréquente délétion homozygote (16%) concernait le gène *RB1* sur le locus 13q14.2-14.3, et une délétion homozygote de 9p21, portant *CDKN2A/CDKN2B*, était retrouvée dans deux tumeurs. Cette délétion du chromosome 13q impliquant le gène *RB1* avait déjà été rapportée dans les UPS et les léiomyosarcomes par l'équipe du Pr Aurias à l'Institut Curie (22, 23). L'équipe de Carneiro a également pu mettre en évidence la perte de la région 4q31 (comportant le gène *SMAD1*, impliqué dans la voie du TGF- β) comme prédictive de métastase, et la perte de la région 18q22 comme favorable, ceci dans les deux groupes de MFH et léiomyosarcomes. L'analyse non supervisée des profils d'expression génique identifia quatre clusters de gènes exprimés dans les léiomyosarcomes et les UPS. Neuf gènes étaient rapportés comme différentiellement exprimés entre les léiomyosarcomes et les UPS (*TAGLN3*, *D4S234E*, *KIAA1729*, *PDLIM5*, *TEAD3*, *TPM2*, *ALDH1B1*, *TRDMT1*, *DHODH*), sans claire explication biologique, hormis le gène de la tropomyosine B β , et sans qu'aucune

méthode bioinformatique ne valide le caractère reproductible de ce set de gènes pour reclasser les échantillons correctement.

Poursuivant dans l'hypothèse d'un continuum entre léiomyosarcomes et UPS avec différents niveaux de différenciation, Larramendy et son équipe postulèrent que les UPS étaient en fait des léiomyosarcomes complètement indifférenciés (24). Ils utilisèrent à nouveau la technique de CGH pour caractériser 102 MFH et 82 léiomyosarcomes. Dans cette cohorte, le génome des MFH comportait trois fois plus de gains que de pertes, tandis que le ratio était de 1 pour les léiomyosarcomes. Ils identifièrent un groupe de 16 MFH se regroupant ensemble et un second cluster comportant le reste des MFH et des léiomyosarcomes. Le petit groupe de MFH comportait des régions très amplifiées restreintes à 1p33~p37, 17q22-q23, et 22p. Le reste des caractéristiques retrouvées en CGH ne permettait pas de différencier les deux histologies. L'ensemble de ces données a alimenté la théorie des auteurs selon laquelle le pléomorphisme est une voie oncogénique fréquente dans le développement et la progression des léiomyosarcomes.

En 2011, Les équipes du Pr Aurias et du Pr Coindre ont rapporté les résultats d'une analyse intégrée de CGH (puce de 3829 BAC-clones couvrant le génome développée dans le laboratoire) et de transcriptomique couvrant plus de 47 000 transcrits (Gene-Chip Human Genome U133 Plus 2.0 Array, Affymetrix, Santa Clara, CA, USA) portant sur 160 sarcomes à génomique complexe (25). Les analyses non supervisées et hiérarchisées de leurs données de transcriptome permirent de définir cinq groupes de tumeurs et six sets de gènes exprimés différemment entre les groupes. (Figures 4 et 5). Le groupe A, le plus homogène car composé exclusivement de léiomyosarcomes bien différenciés, était caractérisé par une surexpression de *MYOCD*, l'absence d'expression de *DKK1* et une délétion du 10q (comportant notamment *DKK1* et *PTEN*) dans 80% des cas. Les autres groupes étaient hétérogènes, incluant des UPS mais aussi des léiomyosarcomes pléomorphes et des myxofibrosarcomes. Le groupe B était caractérisé par une très faible expression des gènes de l'inflammation et du contrôle mitotique et l'absence de perte du 10q et du 13q (comportant *RBI*).

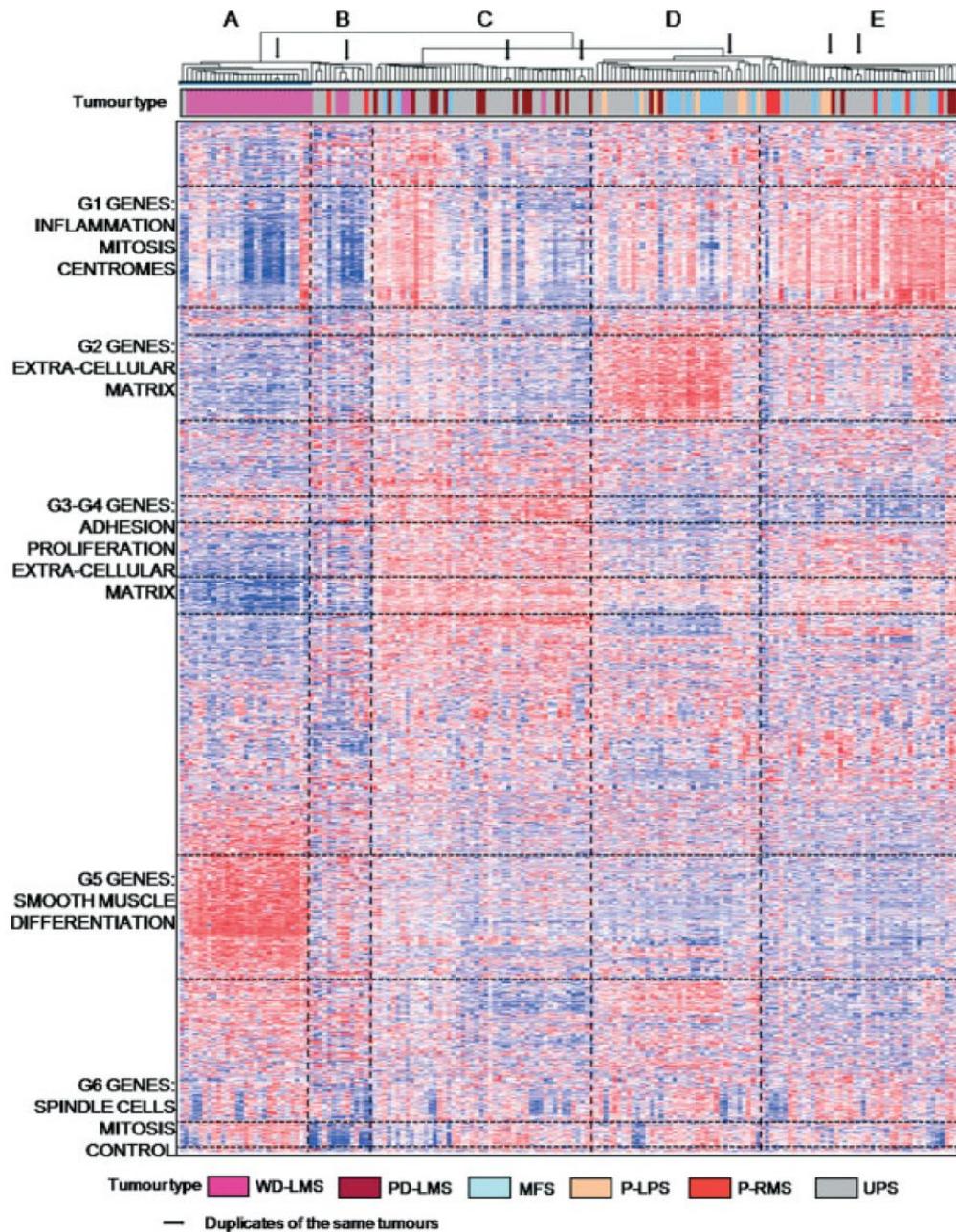


Figure 4 Cinq groupes de sarcomes à génomique complexe sont retrouvés dans l'étude de Gibault *et al. J Pathol*, 2011

Dendrogramme issu de l'analyse hiérarchisée non supervisée des données de séquençage ARN des 160 échantillons de sarcomes de l'étude ; *Heat Map* de l'analyse pondérée de l'expression génique dans les cinq principaux groupes de sarcomes à génomique complexe (WD-LMS : léiomyosarcome bien différencié ; PD-LMS : léiomyosarcome peu différencié ; MFS : myxofibrosarcome ; P-LPS : liposarcome pléomorphe ; P-RMS : Rhabdomyosarcome pléomorphe ; UPS : sarcome pléomorphe indifférencié)(rouge : surexpression, bleu : sous-expression, blanc : niveau d'expression intermédiaire).

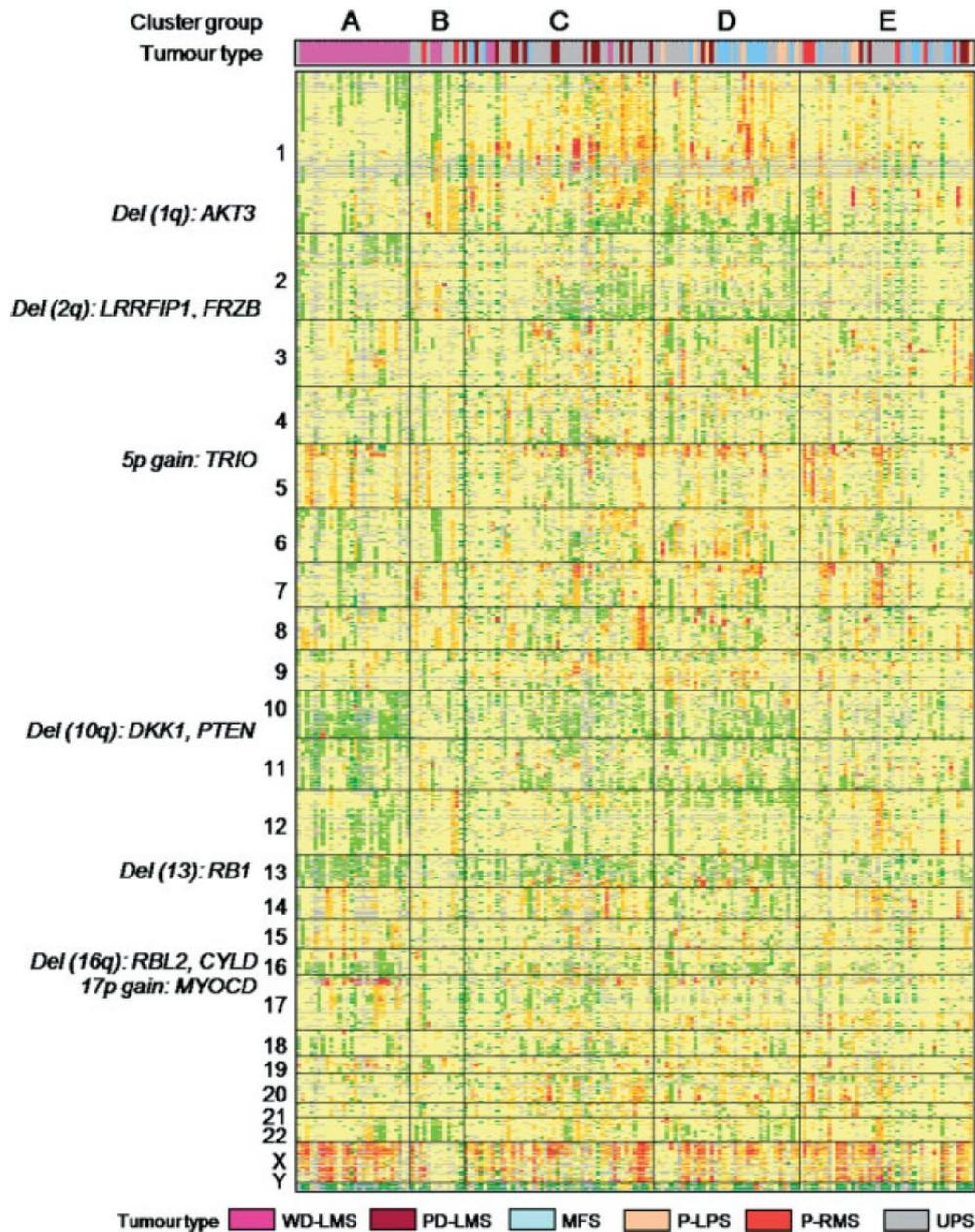


Figure 5 Profils de variation du nombre de copie des gènes retrouvés dans les cinq groupes de sarcomes à génomique complexe de l'étude de Gibault *et al. J Pathol*, 2011

Analyse en CGH des 160 échantillons de sarcomes présentés dans le même ordre que celui du dendrogramme issu des données de transcriptome de la Figure 4 (WD-LMS : léiomyosarcome bien différencié ; PD-LMS : léiomyosarcome peu différencié; MFS : myxofibrosarcome; P-LPS : liposarcome pléomorphe; P-RMS : Rhabdomyosarcome pléomorphe; UPS : sarcome pléomorphe indifférencié)(gains (ratio > 1.2) en rouge, pertes (ratio < 0.8) en vert et absence de déséquilibre (0.8-1.2) en jaune).

Le groupe C, composé d'UPS et de léiomyosarcomes peu différenciés, était caractérisé par une haute expression de gènes impliqués dans l'adhésion cellulaire, la matrice extra-cellulaire et la prolifération, ainsi qu'un gain de 5p (comportant le gène *TRIO*) dans 50% des cas, et des pertes de 1q (comportant *AKT3*), 2q (comportant *FRZB* et *LRRFIP1*, 2 régulateurs de la voie

Wnt/bêta-caténine), du chromosome 13 (*RBI*) dans 40-50% des cas, ainsi qu'une perte de 10q (*DKK1* et *PTEN*) dans 35% des cas. Le groupe D incluait des myxofibrosarcomes et des UPS associés à une haute expression des gènes de la matrice extra-cellulaire. De façon intéressante, 70% des tumeurs du groupe D surexprimaient *DKK1*, et > 55% présentaient des pertes de 13q (*RBI*) et 16q (*CYLD*, connu pour son interaction avec LRRFIP1, régulateur de la voie Wnt/bêta-caténine). Enfin, le groupe E correspondait principalement à des UPS de haut grade des membres et était caractérisé par une surexpression de *DKK1* dans 70% des cas et également une surexpression franche de gènes impliqués dans la réaction inflammatoire. Les auteurs spéculèrent que dans les léiomyosarcomes bien différenciés du groupe A, l'absence d'expression de *DKK1* autoriserait l'activation de la voie de la bêta-caténine, qui associée à l'activation de la voie Akt, notamment via la perte de *PTEN* retrouvée dans 80% des cas, autoriserait à son tour la différenciation musculaire lisse médiée par l'inhibition de *GSK3B* et la surexpression de *MYOCD*. A l'inverse, la surexpression de *DKK1* pourrait être associée à la perte de la différenciation.

En 2014, Pérot et collègues s'intéressèrent au niveau de différenciation musculaire lisse entre les léiomyosarcomes et les UPS et de l'impact de celui-ci en termes de pronostic (26). Ils évaluèrent l'expression en IHC de quatre marqueurs musculaires : actine musculaire lisse, calponine, transgeline, et caldesmone, dans quatre catégories de sarcomes : 148 UPS, 190 léiomyosarcomes, 61 myxofibrosarcomes et 13 liposarcomes pléomorphes. De façon discutable, ils associèrent les myxofibrosarcomes et les liposarcomes pléomorphes aux UPS, arguant de l'absence de différence pronostique en termes de risque métastatique entre ces tumeurs. En IHC, 46% des tumeurs étaient négatives pour les quatre marqueurs, 13% exprimaient un marqueur, 9% deux marqueurs, 10% trois marqueurs et 22% exprimaient les quatre marqueurs. Ils définirent ces dernières comme les « vrais » léiomyosarcomes, et reclassèrent les autres en UPS. Cette classification musculaire lisse avait un impact pronostique, les « vrais » léiomyosarcomes ayant le pronostic métastatique le plus sombre, les léiomyosarcomes histologiques reclassés en UPS sur la base de leur signature IHC partielle ayant une position intermédiaire, et le meilleur pronostic étant retrouvé pour les UPS définis par les deux classifications (mais comportant les myxofibrosarcomes et les liposarcomes pléomorphes). Ils s'intéressèrent ensuite aux UPS définis par leur première classification et proposèrent une seconde classification prenant en compte l'addition du nombre de marqueurs musculaires lisses exprimés et de leur intensité : le groupe 1 avait un score à 0, le groupe 2 avait un score entre 1 et 3, et le 3ème groupe un score entre 4 et 6. De façon intéressante, les sarcomes du groupe 3, avec le phénotype musculaire lisse le plus fort (mais incomplet pour

être considérés de « vrais » léiomyosarcomes) avait le même pronostic métastatique que ces derniers, tandis que les UPS des groupes 1 et 2, avec un score entre 0 et 3, avaient un pronostic métastatique plus favorable (Figure 6).

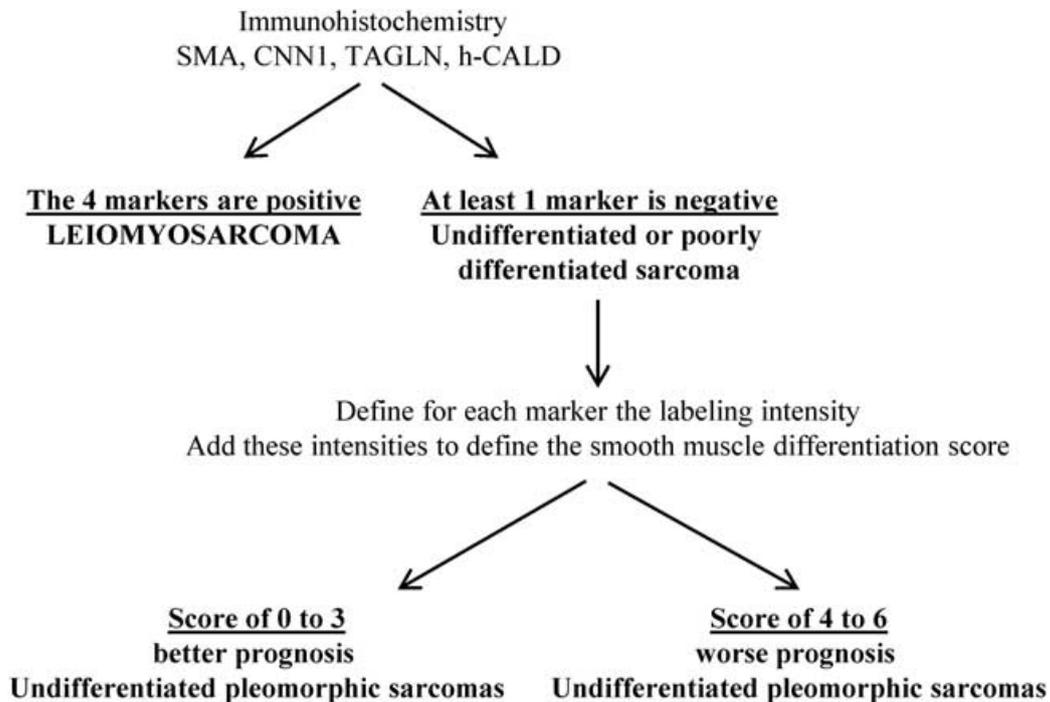


Figure 6 Schéma de la classification de différenciation musculaire lisse de Pérot *et al.* *Mod Pathol*, 2014

Enfin, l'analyse de 78 de ces tumeurs en CGH leur permit de montrer une meilleure corrélation entre le score de différenciation musculaire lisse et le type de profil génomique qu'entre la classification histologique et ce dernier. En effet, selon la classification histologique, 60% des léiomyosarcomes et 93% des UPS avaient un profil de type « complex/rearranged », caractérisé par un niveau élevé de complexité génomique. En revanche, selon leur classification par score musculaire lisse, 93% des UPS avaient un profil de type « complex/rearranged », tandis que 81% des léiomyosarcomes avaient un profil de type « complex/arm », caractérisé par des altérations moins nombreuses et impliquant des portions plus larges et des bras entiers de chromosomes. De façon intéressante, les léiomyosarcomes histologiques reclassés en UPS par leur score étaient à 95% de type « complex/rearranged ». Ces résultats montrent que les évènements génomiques sous-jacents aux sarcomes à génomique complexe semblent différents selon que la différenciation musculaire est complète ou incomplète, et suggèrent également que les léiomyosarcomes

différenciés et les UPS sont bien deux entités distinctes génomiquement, mais qu'une différenciation variable peut être associée au caractère pléomorphe dans ces derniers.

Hypothèses phylogéniques des UPS

La phylogénie des sarcomes est un sujet de débat et de controverse, et celle des UPS l'est particulièrement. Malgré leurs limites, les études de génomique et de transcriptomique des 15 dernières années ont permis de dessiner les contours des UPS et leurs relations avec les autres sarcomes à génomique complexe que sont les léiomyosarcomes et les myxofibrosarcomes. L'opinion qui reste prédominante est que les UPS sont une entité hétérogène contaminée par les sarcomes des autres lignages partiellement dédifférenciés. Cependant, les causes génétiques sous-jacentes à ces nuances de pléomorphisme/dédifférenciation, ainsi que les voies de signalisation qui leur sont associées, restent peu comprises. La perception historique est que les UPS pourraient être originaire d'une cellule souche mésenchymateuse, dans laquelle une instabilité chromosomique surviendrait, avec une capacité de pluripotence et de différenciation variable expliquant l'hétérogénéité associée au pléomorphisme (Figure 7) (27, 28).

Les critères minimaux pour définir une cellule souche mésenchymateuse en culture sont de trois ordres : les cellules doivent être « immatures », c'est-à-dire exprimer des marqueurs tels que CD73, CD105 ou CD90 et ne pas exprimer CD34, CD45, CD11b, CD14, CD19, CD79 ou HLA-DR (29); les cellules doivent être capable d'auto-renouvellement, et également être adhérentes en conditions de culture standard; et elles doivent être pluripotentes, c'est-à-dire avoir la capacité de se différencier en ostéocytes, chondrocytes ou adipocytes, en fonction de modifications apportées au milieu de culture. Les cellules souches mésenchymateuses possèdent des caractéristiques particulières ou « privilèges », notamment le fait d'induire immuno-tolérance et immuno-suppression.

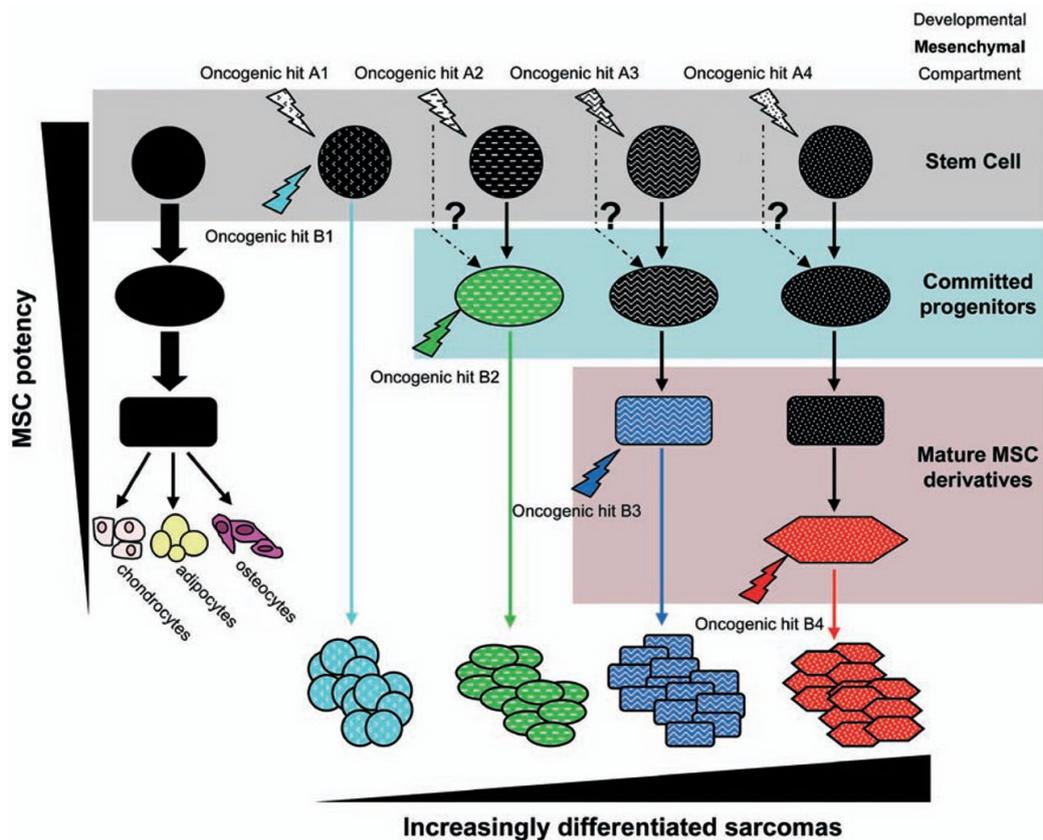


Figure 7 Proposition d'arbre phylogénique de la sarcomagénèse en fonction de la hiérarchie souche mésenchymateuse et des niveaux de différenciation par Rodriguez *et al.* Cell Res, 2012

Ce schéma illustre la coordination de l'acquisition d'altérations oncogéniques dans les cellules appropriées en termes de hiérarchie mésenchymateuse. Modèle du double hit : un hit touche la prolifération/apoptose, le second inhibe la différenciation. Ce modèle met l'accent à la fois sur l'importance des capacités transformatrices intrinsèques des mutations spécifiques et sur la nécessité de cibler la cellule appropriée en termes de différenciation/lignée.

Les cellules souches mésenchymateuses ont une expression faible à nulle des molécules de CMH de classe II (complexe majeur d'histocompatibilité), ce qui leur permet d'échapper aux lymphocytes CD4 (30). Elles expriment cependant le CMH de classe I, mais pas certaines molécules de co-stimulation nécessaires à l'activation des lymphocytes T effecteurs. Elles peuvent également avoir un effet suppresseur direct sur les lymphocytes T, B et NK, notamment par la production d'IDO1 et une activation des voies PGE2 et TGF-bêta, et induire l'expansion des lymphocytes T régulateurs (31-34).

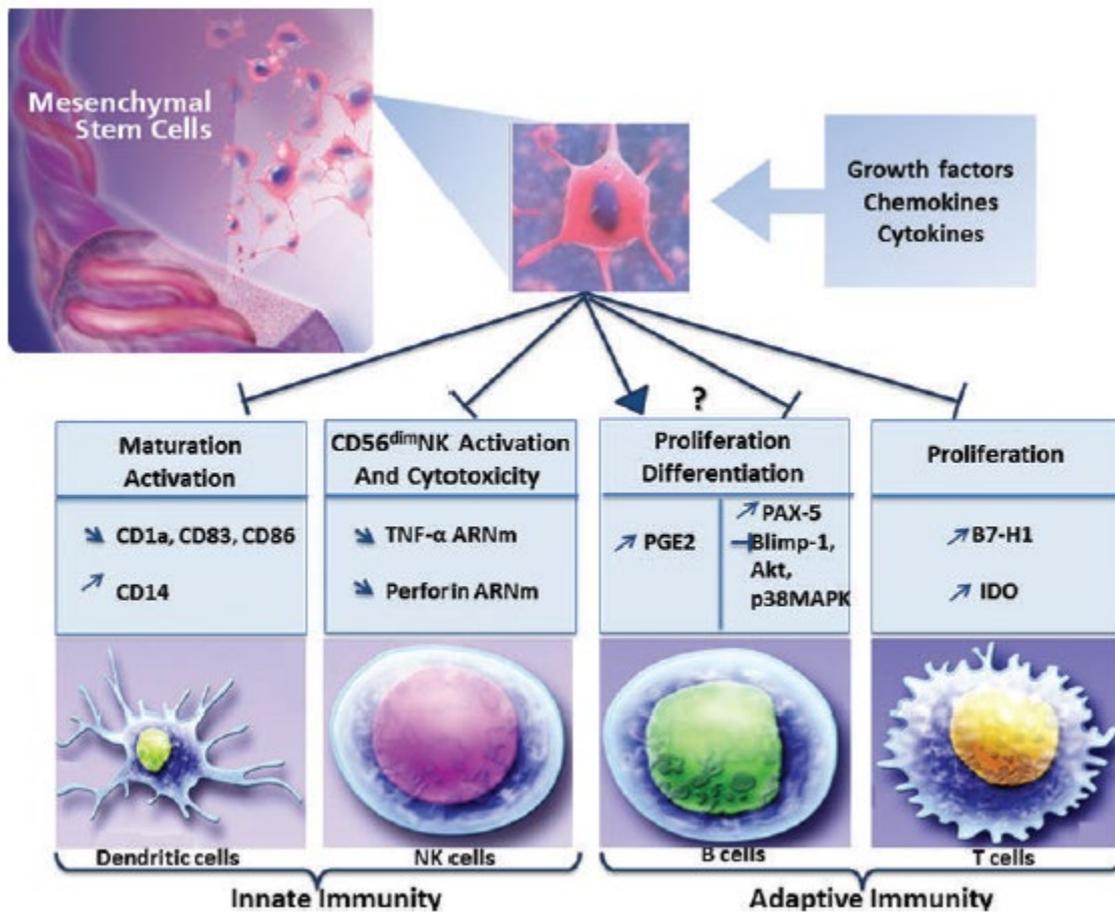


Figure 8 Schéma récapitulatif des propriétés immuno-suppressives des cellules souches mésenchymateuses sur le système immunitaire inné et adaptatif d'après El Omar *et al. Tissue Eng Part B Rev*, 2014

La transformation de cellules souches mésenchymateuses peut survenir par la perte de régulateurs du cycle cellulaire et des altérations des voies Wnt/bêta-caténine, Notch, TGF- β , et PI3K/Akt (35). La voie Wnt/bêta-caténine module la balance entre l'auto-renouveau et la différenciation dans les cellules souches et progénitrices selon un gradient d'activation dose-dépendant (36). Le rôle de la voie Wnt/bêta-caténine diffère entre les sarcomes et les carcinomes. Dans les modèles de carcinomes, il s'agit d'altérations activatrices de la voie Wnt/bêta-caténine - telle que par exemple la perte d'APC dans les cancers colorectaux (37) - alors que c'est l'inhibition de la voie Wnt/bêta-caténine qui semble associée avec la dédifférenciation des tumeurs mésenchymateuses et la sarcomagenèse des UPS. En 2007, Matushansky et collègues ont en effet montré que les cellules souches mésenchymateuses pourraient être les progéniteurs des MFH/UPS via l'inactivation de la voie Wnt/bêta-caténine. Ils ont trouvé que DKK1, un inhibiteur de Wnt et un médiateur de la prolifération des cellules souches mésenchymateuses, empêchait leur différenciation par la

voie canonique Wnt/bêta-caténine. Ils ont pu montrer dans des modèles murins que l'inhibition de la voie Wnt/bêta-caténine pouvait induire la transformation de cellules souches humaines mésenchymateuses en UPS/MFH, et qu'en rétablissant cette voie la dédifférenciation pouvait être inversée et que les cellules pouvaient se différencier en tissus mésenchymateux matures (38).

En 2012, l'équipe de Wang rapporta un rôle de l'activation des voies Hedgehog et Notch dans une fraction de cellules souches mésenchymateuses responsables du développement des UPS. Après avoir isolé une fraction de cellules tumorales dites « souches », ils comparèrent les profils d'expression des cellules souches et des cellules non souches par puce ADN (Gene-Chip Human Genome U133 Plus 2.0 Array, Affymetrix, Santa Clara, CA, USA) dans 10 tumeurs de type UPS (39). Les gènes différentiellement régulés entre les deux populations de cellules étaient ensuite confirmés par RT-PCR. L'analyse de la liste des gènes différentiellement régulés identifia les voies Hedgehog (*GLII*, *PTCH1*, *HHIP*), Notch (*HES1*, *HEY1*, *HEY2*), Wnt/bêta-caténine, TGF-bêta, interleukine, KIT, EGF, Intégrine Alpha6-Beta4 comme préférentiellement activées dans la fraction de population de cellules souches. Ils s'intéressèrent alors aux voies Hedgehog et Notch et montrèrent dans des modèles de xénogreffes de souris NOD/SCID que le blocage de ces voies faisait diminuer la fraction de ces cellules souches dans les tumeurs, diminuait la croissance des tumeurs, et supprimait la capacité d'auto-renouvellement et de prise de greffe de ces tumeurs. A l'inverse, la chimiothérapie conventionnelle augmentait la proportion de ces cellules souches et n'affectait pas leur capacité à être transplantées de façon sérielle.

En résumé

Les UPS sont un groupe de sarcomes indifférenciés pléomorphes dont la définition a beaucoup changé dans les 20 dernières années. Le diagnostic d'UPS repose sur l'absence de ligne de différenciation après un examen histologique minutieux et l'utilisation de techniques d'IHC et de biologie moléculaire adaptées (40). Les études d'omiques existantes ont principalement porté sur l'analyse du transcriptome par puces ADN et celle des altérations du génome par CGH. Ces études sont dominées par la contamination des MFH/UPS par d'autres histologies, et par le problème de situer les «vrais UPS» par rapport aux autres sarcomes à génomique complexe. Elles ont toutefois permis de dresser les contours des UPS incluant instabilité génomique et défaut d'engagement dans les voies de différenciation mésenchymateuse (41). Nous avons voulu mettre à profit la disponibilité de techniques d'analyse à haut débit du transcriptome, de l'exome, et du protéome ainsi que les avancées récentes en matière de traitement des images en radiologie par la radiomique, pour établir une classification intégrée multi-plateforme et actualisée de «vrais» UPS primaires non pré-traités.

Hypothèses et Objectifs

Nos hypothèses pour ce travail étaient :

- Les UPS sont originaires de cellules souches mésenchymateuses sans engagement de différenciation.
- Il existe un continuum entre les UPS et les sarcomes pléomorphes plus différenciés comme les myxofibrosarcomes non classiques et les léiomyosarcomes pléomorphes ou à phénotype musculaire lisse incomplet, et ce continuum est associé à une gradation de ce niveau d'engagement.
- Il y a une relation entre le niveau d'engagement dans un lignage et le niveau d'activation de l'immunité intra-tumorale.
- Les UPS peuvent être divisés en sous-groupes avec des profils génomiques et transcriptomiques spécifiques.
- Ces sous-groupes peuvent être identifiés au moyen d'une approche intégrée de données transcriptomiques, génomiques, immuno-phénotypiques et protéomiques.
- Les voies de signalisation différenciellement exprimées dans ces sous-groupes peuvent être ciblées en thérapeutique par des approches pharmacologiques de type thérapies ciblées ou immuno-modulation.

Notre objectif principal était d'utiliser les techniques de biologie moléculaire actuelles de type séquençage haut débit du transcriptome et de l'exome et l'évaluation *in situ* par immuno-histo-chimie des infiltrats tumoraux pour générer une classification intégrée de vrais UPS primaires non prétraités diagnostiqués après relecture anatomo-pathologique experte systématique et utilisation en routine des moyens diagnostiques moléculaires les plus récents.

Les objectifs secondaires étaient :

- D'identifier des gènes et des voies de signalisation spécifiques aux UPS et altérés de façon différentielle et récurrente,
- D'élaborer des lignées cellulaires et des modèles de xénogreffes dérivées de tumeur de patients (PDX) caractérisés génomiquement,
- De tester le rôle fonctionnel de cibles sélectionnées *in vitro* et *in vivo* dans ces modèles de lignées et de PDX,
- De produire des données pouvant permettre de mieux sélectionner les patients dans les approches thérapeutiques développées actuellement, notamment dans le cadre de la recherche clinique.

Méthodes

Voir les parties « **Methods** » et « **Supplementary Methods** » de l'article joint.

Résultats

Voir l'article joint, en cours de seconde relecture dans le *Journal of Clinical Investigation*.

Les tables référencées en supplément seront disponibles en ligne sur le site du journal sous la forme de fichiers excel.

The Integrated Molecular and Imaging Landscape of Undifferentiated Pleomorphic Sarcomas

Maud TOULMONDE^{1,2}, Carlo LUCCHESI^{1,3}, Stéphanie VERBEKE^{1,3}, Amandine CROMBE^{1,2}, Julien ADAM⁴, Damien GENESTE³, Vanessa CHAIRE^{1,3}, Audrey LAROCHE-CLARY^{1,3}, Raul PERRET¹, François BERTUCCI⁵, , Frederic BERTOLO³, Laurence BIANCHINI⁶, Bérengère DADONE-MONTAUDIE⁶, Todd HEMBROUGH⁷, Steeve M. SWEET⁷, Yeoun Jin KIM⁷, Fabiola CECCHI⁷, François LE LOARER¹, Antoine ITALIANO^{1,2,3}

¹Sarcoma Unit, Institut Bergonié, Bordeaux, France

²University of Bordeaux, Bordeaux, France

³INSERM U1218, Bordeaux, France

⁴Pathology Department, Gustave Roussy, Villejuif, France

⁵Centre de Recherche en Cancérologie de Marseille (CRCM), CNRS UMR725, INSERM U1068, Institut Paoli-Calmettes, Aix-Marseille Université, Marseille, France

⁶Laboratory of solid tumor genetics, Université Côte d'Azur (UCA), CNRS UMR7284, INSERM U1081, Institute for Research on Cancer and Aging, Nice (IRCAN), Nice, France

⁷Nantomics, Rockville, MD, USA

Corresponding Author: Maud TOULMONDE, MD

Institut Bergonié, Medical Oncology Dept

Early Phase Trials and Sarcoma Units

229 cours de l'Argonne

33076 Bordeaux Cedex France

Tel: + 33 5 47 30 60 88

+ 33 6 18 47 19 79

Fax: + 33 5 47 30 60 83

m.toulmonde@bordeaux.unicancer.fr

The authors declare no conflict of interest.

ABSTRACT

Undifferentiated pleomorphic sarcoma (UPS) is one of the most aggressive and least-characterized sarcoma subtypes. By using high-throughput technologies encompassing RNA-sequencing, whole exome-sequencing, mass spectrometry, immunohistochemistry as well as radiomics, we identified and confirmed in three independent cohorts of patients two main disease entities with distinct molecular features, immune phenotypes, and MRI textures, as well as differential sensitivity to specific anti-cancer agents.

INTRODUCTION

Since its first description in 1964 (1), malignant fibrous histiocytoma (MFH) has been considered the most common soft tissue sarcoma (STS) of adults. Main distinctive features of MFH are an apparent lack of specific differentiation and predominant pleomorphic patterns (2), leading some to consider it as a heterogeneous group of tumors made up ‘by default’(3). In 2002, the WHO reclassified the group under undifferentiated pleomorphic sarcoma (UPS) (4). UPS currently represents a diagnosis of exclusion based on the absence of a specific line of differentiation after careful histological examination and judicious use of ancillary techniques, including immunohistochemistry and molecular biology (5, 6). Despite adequate locoregional treatment, up to 40% of patients will develop metastatic disease, with the worse progression-free survival in palliative stage among all histological sarcoma subtypes (7). We report here the results of a multi-omics approach on centrally reviewed, high-grade, non-pretreated primary UPS, with the aims to deeply assess their molecular characteristics, better classify these entities, and identify specific therapeutic targets.

COMBINED RESULTS AND DISCUSSION

The clinical characteristics of the 135 patients included in the study are shown in Supplementary Table 1. Twenty-five cases were selected for full exome and RNA-sequencing based on availability of frozen material. Unsupervised consensus and hierarchical clustering of RNA-sequencing data identified three groups of patients with associated clusters of genes (Figure 1A and Supplementary Figures 1A and 1B). The analysis of differential gene expression between main groups A and B identified 1405 genes (Supplementary Table 2). We observed that group A was mainly enriched in genes that play a crucial role in both normal development and stemcellness (Supplementary Table 2). Such genes include *LHX8*, which encodes an important mediator of stem cell fate, *LRRN1*, implicated in pluripotency maintenance of stem cells, *LGR5*, which encodes a seven transmembrane spanning receptor

that has recently been found to be a stem cell marker, and *BMP5*, a member of the TGF-beta family involved in development and differentiation. Interestingly also, *FGFR2* was one of the top genes upregulated in group A (fold-change: 23.5). Group B was strongly enriched in genes involved in immunity, such as *MARCO*, *TIMD4*, *TIGIT*, *CD27*, *IFNG*, *CD8B*, *PDCD1*, *CD3D*, *IDO1* and numerous cytokines, but also demonstrated upregulation of *DKK1*, a key inhibitor of Wnt/beta-catenine pathway (Supplementary Table 2). Further assessing enrichment in gene-sets of the Hallmarks of Cancer, Gene Ontology Biological Processes, and Cybersort LM22 immune-cell signatures (8, 9) (<http://www.geneontology.org>), we found that genes over-expressed in group B belong to inflammatory response and INFgamma response immune pathways, CD8+ T-lymphocyte (TL) and monocyte, but also Natural Killer, CD4+ TL helper, activated dendritic cell, and memory BL, as well as regulatory TL immune cell signatures, indicating a complete immune representation *in situ* (Supplementary Table 3). In order to test the robustness of our model, we then used an independent cohort of 41 UPS from The Cancer Genome Atlas (TCGA) consortium for which up-to-date pathological review and RNA-sequencing data were publicly available (10). Using the same bioinformatics methods, we also identified three groups of patients with associated differential gene expression (Figure 1B and Supplementary Figure 1C). Interestingly, group A from TCGA cohort was enriched in genes involved in stemcellness, notably *LHX8*, *LRRN1* and transcription factors found in group A from Institut Bergonié (IB) cohort, and demonstrated a differential overexpression of *FGFR2*. Furthermore, group B from TCGA cohort was enriched in immune clusters, resembling well group B from IB cohort. Indeed, *IDO1*, *IFNG*, *TIMD4*, *PDCD1*, *CD3D*, *CD8B* as well as *TIGIT* and *CTLA4* were among the genes the most upregulated, together with *DKK1* (Supplementary Table 4). Analysis of Agreement between Differential EXpression (AGDEX) of the 25 UPS from IB cohort and the 41 UPS from TCGA cohort revealed a very high correlation between clusters (Spearman = 0.83, Pearson = 0.73) (Figure 1C).

Physiologically, precise regulation of FGFR expression is required in a wide range of processes in embryonic development, stem cell maintenance, and differentiation (11). FGFR signaling is required for sustaining self-renewal and pluripotency of human embryonic stem cells (12), and FGFR signaling has been implicated in a variety of tissue stem cell activities, including bone marrow mesenchymal stem cells (13) and hematopoietic stem cells (14). *FGFR2* is a tyrosine kinase receptor that activates several oncogenic pathways such as MAPK or PI3K/Akt pathways. Given its crucial role in driving oncogenic growth in several cancers (15) and the current clinical development of several highly specific FGFR inhibitors, we then

assessed the impact of *FGFR2* overexpression in UPS tumorigenesis. Using quantitative Reverse Transcriptase-PCR (RT-qPCR) and IHC, we first confirmed *FGFR2* overexpression in patient samples from group A compared to other groups (Supplementary Figure 2A and 2B). We then established four UPS cell lines from tumor samples with different profiles: IB106 and JR588 were representative of group A, with high *FGFR2* expression confirmed on RT-qPCR and Western Blot (WB), whereas KN473 and IB119 originated from patients samples of group B and C, respectively, with no overexpression of *FGFR2* (Supplementary Figure 2C). Using two different shRNA of *FGFR2*, we were able to inhibit *FGFR2* in IB106 and KN473 cell lines. IB106 cells were highly sensitive to *FGFR2* inhibition and demonstrated significantly higher rate of apoptosis than control for both shRNA, whereas KN473 cells demonstrated no *FGFR2* expression and no increase in apoptosis after infection with sh*FGFR2* compared to shcontrol (Supplementary Figure 2D and 2E). The specific *FGFR1-4* inhibitor JNJ-42756493 inhibited growth of IB106 and JR588 cells with IC50 values of 1.05 and 0.81 $\mu\text{mol/l}$ (Figure 2A). In contrast, the compound did not potently inhibit growth of KN473 cells ($\text{IC}_{50} > 10 \mu\text{mol/l}$) and had only intermediate efficacy on IB119 cells ($\text{IC}_{50} = 4.6 \mu\text{mol/l}$). We further assessed two other specific *FGFR* inhibitors with the same sensitivity profiles among the four tested cell lines (Supplementary Figure 2F). JNJ42756493 induced upregulation of *FGFR2* expression and relative downregulation of its phosphorylated form in IB106 and JR588 cells but not in KN473 and IB119 cells, whereas Phospho-Erk (pErk), a key downstream mediator of the *FGFR* pathway, was significantly downregulated by JNJ42756493 in IB106 and JR588 cells, as well as IB119 but not in KN473 cells (Figure 2B). After 24 hours of treatment, JNJ-42756493 also induced significant G0–G1 arrest ($P < 0.001$) in IB106 and JR588 cells while it did not affect the cell cycle of IB119 and KN473 cells (Figure 2C). We finally assessed in vivo efficacy of JNJ-42756493 using two UPS patient-derived xenograft models (PDX): JR588, overexpressing *FGFR2*, and KN473, without *FGFR2* overexpression. We observed a clear trend for delayed tumor growth in the JR588 model (Figure 2D). Collectively, our data provide preclinical evidence for *FGFR* targeting as a promising therapeutic strategy in a subset of UPS.

We then investigated the correlation between gene expression profiling and tumor immune infiltration patterns on tumor. We assessed protein expression density by IHC for several immune markers (CD8 for CD8+ LT, Program Death 1 (PD-1), CD68, CD163 and Indole 2,3-dioxygenase 1 (IDO1)) on related tumor sample slides from the 25 UPS with available gene expression profile (16). M1 macrophages were defined as CD68+ CD163- cells and M2 macrophages as CD163+ cells. All immune markers tested apart from those defining M1

macrophages were significantly more expressed in group B than group A and C ($p < 0.05$). Interestingly, IDO1 as well as CD8 and PD-1 densities were significantly higher in group B, than group A and C ($p < 0.05$) (Supplementary figure 3A, 3B, 3C, 3D, 3E). The most significant predictors of RNA-sequencing-derived group B with the highest prediction performance were CD8 and IDO1 (Figure 3A and 3B). This led us to call group B the *immune-high* group, whereas group A was named the *immune-low* group and group C, less well characterized, the group *other*. Because group C was too small to have sufficient statistical robustness, we decided to further focus on the two main groups of immune-high and immune-low UPS.

Infiltration of tumors by CD8+ TL has been associated with favorable prognosis in different malignancies (17). We therefore analyzed the prognostic impact of CD8 density and IDO1 expression by IHC in an independent larger series of 110 primary non-pretreated UPS. In this independent series, median CD8 density was 57.5 cells/mm² and 63 among 107 assessable samples (59%) had IDO1 H-score > 0 , with a median H-score of 5 (0-210). Importantly, 87% of CD8-high tumors (CD8+ TL infiltrates above the median) were IDO1-positive versus 30.2% of CD8-low tumors ($p < 0.0001$). Patients with CD8-high tumors had a significantly higher Metastasis-Free Survival (MFS) ($p = 0.04$) and a trend for better Overall Survival ($p = 0.17$) (Supplementary Figure 4A and 4B). This in line with a previous work showing that IDO1 expression is a frequent hallmark of soft-tissue sarcomas and that this pathway could be an important mechanism of primary resistance to PD-1 inhibition observed in these tumors (16). Altogether, these data suggest a role for the IDO1 pathway in UPS progression and a therapeutic potential of IDO1 inhibition in the immune-high subset of UPS.

There is growing evidence that tumor genetic events have a direct impact on tumor microenvironment. We then assessed the relationship of somatic mutation and gene copy number alteration (CNA) rates with immune-high and immune-low gene expression profiles in our UPS cohort. Overall, mutation burden in the whole series was low (median: 4.3 mutations/Mb). However, all UPS samples with mutation burden > 5 mutations/Mb belonged to the immune-high group (Figure 3C). The most frequently mutated genes in our cohort included *TP53* and *ATRX* (Supplementary Table 5), with no relevant mutation differentially recurrent in either of each group. Strikingly, the UPS samples from the immune-low group were characterized by a significantly higher rate of gene CNA in comparison with the immune-high UPS (OR = 10.8 for homozygous deletions, OR = 6.9 for amplifications) (Supplementary Figure 5A, 5B and 5C, Supplementary Table 6). There was notably a significant enrichment in deletions of Tumor Suppressor Genes (TSG) belonging to DNA

repair, cell cycle, apoptosis and beta-catenin or PI3K/mTOR pathways, such as *FANCA*, *TP53*, *CCDC6*, *RB1*, *AXIN1*, *PI3KR1* or *TSC2* in the immune-low UPS (Figure 3D). These findings are concordant with recent publications showing correlation between mutation burden and cytotoxic immune infiltrates and association between high level of CNA with low immune infiltrates and worse response to immune checkpoint blockade (18-21). Interestingly also, our results echo recent work from Steele et al. who reported significant enrichment for immune related pathways in a subset of undifferentiated sarcomas (US) characterized by higher mutational burden and lower rearrangement level, as well as recurrent deletions in genes such as *TP53*, *RB1*, *CDKN2A*, *TERT*, *ATRX* or *PTEN* in another subset of US characterized by lower mutational burden and high rearrangement level (22).

Analysis of chimeric fusion genes from RNA-sequencing did not identify any significant recurrent alteration that could discriminate between the immune-high and the immune-low groups. However, we found a significantly higher number of non-recurrent fusion genes in the immune-low group (OR = 1.96, $p < 0.01$). One sample had a *TRIO/TERT* fusion, previously described in sarcoma by Delespaul et al. (23).

We then explored whether proteomics analysis could recapitulate our UPS genetic classification. Using Data-Independent Acquisition (DIA) mass spectrometry, we assessed expression of 3215 proteins in our genomically-characterized UPS samples. Unsupervised consensus clustering and hierarchical clustering of protein expression identified three proteomics groups, that we named PA (n = 9), PB (n = 10) and PC (n = 4) (Figure 4A). Importantly, the two main proteomic groups - PA and PB - correlated well with the two main transcriptomic groups, A/immune-low and B/immune-high, with high agreement on label classification between proteomics and RNA-sequencing clusterings (precision: 82%). Overall, 565 proteins had differential expression between PA and PB ($FC \geq 2$, $p = 0.01$). When performing AGDEX between proteomics and RNA-sequencing data, we found good correlation (Spearman = 0.56, Pearson = 0.57) between 114 protein/gene couples differentially expressed in both PA vs PB, and immune-low vs immune-high groups (Supplementary Table 7 and Supplementary Figure 6A). Hallmark of Cancer gene-sets analysis revealed that the proteome of the PB group, associated to the immune-high group, was significantly enriched in immune response and metabolism pathways such as glycolysis and oxidative phosphorylation, whereas the PA group proteome, associated with the immune-low group, was enriched in epithelial mesenchymal transition (EMT) and c-Myc targets pathways (Supplementary Figure 6B). C-Myc is a master transcription factors involved in many developmental pathways, notably EMT and stemcellness, two pathways differentially

represented in our immune-low group of UPS (24, 25). Upregulation of *MYC* has also been associated with increased genomic instability (26). Interestingly, *FGFR2* is one of c-Myc target genes, whereas c-Myc is a downstream effector of the FGFR/MEK/ERK pathway (27). C-Myc is associated with histone lysine acetylation and recruitment of proteins with acetyl-lysine-binding motifs or bromodomains. Bromodomain and extraterminal (BET) proteins are a subfamily of bromodomain proteins. The suppressing effect on *FGFR2* and *MYC* transcription observed with BET inhibitor JQ1 (28), together with BET inhibitors activity reported on transcriptional activity of c-Myc-driven tumor models, and notably the role of BET proteins CBP/P300 inhibition on c-Myc driven proliferation (29), drove us to assess activity of compounds such as CPI637, NEO1132 and NEO2734, that are potent dual inhibitors of BET proteins CBP/P300, on our UPS models (28). Interestingly, when exposed to the three different potent dual inhibitors of CBP/P300, JR588 and IB106 cells (derived from the immune-low/PA group) were significantly more sensitive to all three compounds than KN473 cells (immune-high/PB group) and IB119 cells (group other/PC) (Supplementary Figure 7). These preliminary data promote further assessing BET modulation in the immune-low subset of UPS.

Radiomics is a promising field for systematizing and predicting tumor biology (30). Finally, we explored if radiomic features could correlate with our UPS molecular classification. Fourteen patients (seven immune-high, six immune-low and one from the group other) had an available pre-treatment MRI with contrast-enhanced T1-weighted imaging. After voxel resampling, signal intensities normalization and slice-by-slice whole tumor volume segmentation, 62 radiomic features quantifying shape and heterogeneity were extracted (Supplementary Table 8). Of these, nine were statistically different between the immune-high and the remaining UPS (range of p-value = 0.01-0.05). Hierarchical clustering of patients based on this 9-feature radiomic signature separated the immune-high UPS from the others with a specificity of 100% (7/7), a sensitivity of 86% (6/7) and an accuracy of 93% (13/14) (Figure 4B). Radiomics approach could be particularly relevant for soft tissue sarcomas, and especially UPS because of their marked inter-individual but also intra-tumoral heterogeneity. Here we show that nine radiomic features from basic conventional MRI sequences provide the basis for a radiomic signature that could accurately select patients with immune-high UPS on their pre-therapeutic imaging. These data warrant further validation on a larger independent cohort of UPS patients.

This study provides the most comprehensive omics characterization of true primary UPS to date, revealing two main subsets with distinct genomic, immunologic, proteomic and

radiomic features, and direct therapeutic potentials. This study improves the biological understanding of UPS and brings to light the reasons for mixed results obtained with checkpoint inhibitors in soft tissue sarcomas to date (31, 32). These data also directly help building new strategies to improve patient care. Since tumor infiltration by CD8+ TL correlates with better response to immune checkpoint inhibitors and patient outcome (33), patients with immune-high UPS may certainly be the best candidates for immune checkpoint inhibitors, including IDO1 inhibitors (16). Moreover, radiomics could be a powerful tool to identify these patients from pre-therapeutic MRI. Finally, our results provide rationale for assessing FGFR and BET inhibition as single agents or in combination in immune-low UPS.

METHODS

Tumor samples

Formalin-Fixed Paraffin-Embedded (FFPE) tumor blocks, frozen tumor and matched normal tissue samples of 25 patients with non-pretreated, surgically resected, primary, high grade UPS treated at Institut Bergonié were collected for the genomically-characterized cohort. A Tissue Micro Array (TMA) was produced from FFPE tumor blocks of 110 patients with non-pretreated surgically resected primary high grade UPS that constitute the IHC-characterized cohort. Three spots for each tumor sample were assessed.

Samples used in this study have been selected according to the following inclusion criteria:

- Histological review performed for each case by a soft tissue sarcomas referral Pathologist of the French Sarcoma Group,
- High grade defined as grade 2-3 according to the FNCLCC grading system,
- No previous treatment done before sampling,
- Informed consent obtained for the analysis.

Cell culture and Reagents

The UPS cell lines used in this study were derived from human UPS surgical specimens after obtaining written, informed patient consent and Institut Bergonié IRB approval. Cells were grown in RPMI medium 1640 GlutaMAX™ Supplement (Life Technologies, Carlsbad USA) in the presence of 10% (v/v) fetal bovine serum and Penicillin/Streptomycin 1% (Dutscher, Merignac, France), in flasks. Cells were maintained at 37°C in a humidified atmosphere containing 5% CO₂. Cells were routinely passaged every 2 or 3 days. JNJ-42756493, a pan-FGFR inhibitor, was provided by Johnson&Johnson (New Brunswick, USA), AZD4547 and BGJ398, two other pan-FGFR inhibitors, were purchased from Selleck Chemicals (Houston, USA), CPI637, NEO1132 and NEO2734, three BET inhibitors, were provided by Epigene Therapeutics Inc. (Montreal, Canada), prepared as a 10 mmol/L stock solution in DMSO and stored at -20°C for in vitro studies. Cultured cells were treated with medium changes without antibiotics and fresh drug additions as indicated in the figure legends. Two FGFR2 shRNA were purchased from Sigma-Aldrich (Saint-Louis, USA) (TRCN0000218493 shFGFR2-1 -5'-AGCCCTGTTTGATAGAGTATA-3' and TRCN0000231051 shFGFR2-2 -5'-TTAGTTGAGGATACCACATTA-3'). Viral particles were produced by calcium phosphate transfection of 293T cells and infections were performed at a multiplicity of 1 or 5 infectious units per cell.

RNA isolation and RT-qPCR

Total RNA from UPS patients and cell lines were isolated using the RNeasy Mini Kit (Qiagen, Hilden, Germany) according to the supplier's instructions. The purity and concentration of isolated RNA were determined by spectrophotometer NanoDropTM and capillary electrophoresis using Bioanalyser 2100 (Agilent Technologies Inc., Santa Clara, USA) was carried out to determine the RNA integrity number (RIN). For RT-qPCR, cDNA was prepared from 500 ng total RNA using the High Capacity cDNA Reverse Transcription Kit (Life Technologies, Carlsbad, USA). To analyze the FGFRs expression, the following human TaqMan Gene Expression Assays were used: *RPLP0* Hs999999_m1, *FGFR1* Hs00241111_m1, *FGFR2* Hs01552918_m1, *FGFR3* Hs00179829_m1, *FGFR4* Hs00242558_m1. Each gene was analyzed in duplicate with 2.5 ng cDNA in three independent experiments with an ABI PRISM 7500 Fast Detection System instrument (Applied Biosystems, Thermo Fisher Scientific, Waltham, USA). *RPLP0* was used as a reference gene. Reaction conditions are the following: pre-incubation of 50 °C for 2 min, initial denaturation at 95 °C for 10 min and 40 cycles of: 95 °C/15 s, 60 °C/1 min. The relative comparative method was used to analyze RT-qPCR data and fold change are calculated using the Universal Human Reference RNA (Agilent Technologies Inc., Santa Clara, USA).

Western blot

Treated and control cells were harvested in 60 μ l of radio-immuno-precipitation assay (RIPA) lysis buffer. The lysate was centrifuged (13 000 rpm, 15 min, 4°C), and the supernatant was stored at -80°C. Equal amounts of total protein (30 μ g) were electrophoresed on 12% or 10% sodium dodecyl sulfate polyacrylamide gels and transferred onto polyvinylidene difluoride membranes. The blots were probed overnight at 4°C with an anti-FGFR2 (#23328, Cell Signaling Technology), anti-actin (A3853, Sigma), anti-phospho-FGFR2 (ser 782) (PA5-64796, Thermo Fisher), anti-phospho-Erk (Thr202/Tyr204) (#4370, Cell Signaling Technology) and anti-Erk (ab17942, Abcam) primary antibody diluted in PBST (DPBS 10X (GibcoTM) after 1X dilution; 0.1% Tween-20) with 5% bovine serum albumin. The horseradish peroxidase-conjugated secondary antibody (Santa Cruz Biotechnology Inc. Dallas, USA) was diluted 1:10000. Bound antibodies were visualized on Fusion Fx7 imaging system (Fisher Bioblock Scientific, Waltham, USA) using the ImmobilonTM Western enhanced chemiluminescence detection kit (Millipore Corporation, Billerica, USA). The resulting bands were analyzed and quantified using ImageJ 1.48v software (National Institutes of Health, Bethesda, USA).

Cell viability assay

Cells were seeded in triplicate at 4000 cells/well into 96-well plates, cultured with fresh growth medium for at least 24 hours and treated with a range of increasing concentrations of drugs for 72 hours. Cell viability was assessed by MTT (2-deoxyglucose (2-DG) and 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay (Sigma-Aldrich, St. Quentin Fallavier, France) at a final concentration of 0.5 mg/mL and 3h of incubation. Then, supernatant was discarded, 100 μ L of dimethyl sulfoxide (DMSO) was added and the absorbance at 570 nm was monitored using a Flexstation 3 Plate reader (Molecular Devices, Sunnyvale, USA), using 630 nm as a reference. The half maximal inhibitory concentration (IC₅₀) was calculated with GraphPad Prism software version 5.0 for Windows (GraphPad Software, La Jolla, USA). Each experiment was repeated at least 3 times.

Cell cycle analysis

The cell cycle distribution was studied by examining DNA content using fluorescence-activated cell sorting (FACS) and analyzed using Cell Quest Pro software (BD Biosciences, San Jose, USA). About 3×10^5 cells were seeded in complete medium in 6-well plates. After 6h, cells were washed once with PBS and medium without FBS were added to synchronize cells overnight. Then, the cells were treated for 24h in medium with FBS and JNJ-42756493 at the IC₅₀ concentration, centrifuged at 1500 g for 5 min, and washed with PBS. The cells were then fixed with 70% ethanol at -20°C overnight. Following ethanol removal, the cells were washed with PBS. Next, 300 μ l of a Propidium Iodide (PI) and ribonuclease-containing solution were added to the cells and then analyzed by FACS. The data were analyzed with FlowJo v.7.6.3 software, and the results were expressed in terms of percentage of cells in a given phase of the cell cycle based on 3 independent experiments.

Animal studies

Induction of tumor patient derived xenograft (PDX) was performed by subcutaneous implantation of fragments of fresh surgically resected UPS tumor of approximately 3 mm³ into the right flanks of Ragγ2C^{-/-} mice. When tumors were established (18 and 30 days after implantation of JR588 and KN473, respectively), mice were randomized according to tumor volume to either vehicle alone (10% hydroxypropyl- β -cyclodextrin) or vehicle containing JNJ-42756493, administered by oral gavage daily at 30mg/kg for 4 weeks. At the end of the experiments, mice were euthanized, and tumors were excised. Tumor progression was

analyzed with GraphPad Prism software. Experiments have been performed in agreement with Bioethics Law No 2004-800 and the Ethics Charter from the National Institute of Cancer (INCa).

Immunohistochemistry

IHC was performed on FFPE tumor samples using automated protocols on a Ventana BenchMark Ultra platform (Roche, Bâle, Switzerland). Monoclonal primary antibodies for CD8 (clone SP16, Spring Bioscience), PD-1 (clone NAT105, Abcam), IDO1 (clone D5J4E, Cell Signaling Technology), CD68 (clone PGM1, Dako), CD163 (clone 10D6, Novocastra, Leica Biosystems) and FGFR2 (clone D4L2V, Cell Signaling Technology) were used. Amplification and detection steps were performed with an Ultraview kit and 3,3'-diaminobenzidine was used as a chromogen. CD68/CD163 double staining was performed on a Ventana Discovery Ultra platform (Roche, Bâle, Switzerland) and image analysis as described previously (34). Image analysis was performed under the supervision of a pathologist (JA) to detect the density of CD8 and PD-1 positive cells in the tumor areas and the proportion of IDO1, CD68 and CD163 stained surface of tumor slides or spots, as previously described (35). IDO1 staining on TMA slides was evaluated semi-quantitatively by a trained pathologist (JA) including percentage (0-100) and intensity (0 = null, 1 = low, 2 = moderate, 3 = strong) of staining on tumor spots, and a H-score (0-100) was defined by percentage x intensity of staining. FGFR2 staining on tumor slides was evaluated semi-quantitatively by a trained pathologist (RP), and included percentage (0-100) and intensity (0 = null, 1 = low, 2 = moderate, 3 = strong) of staining on tumor cells.

Statistics

Statistical analyses were performed using R (36). Quantitative variables were described using the median and range, and qualitative variables were described using frequency and percentage. Metastasis-free survival (MFS) was defined from start of treatment to time of first metastasis or death (from any cause). Overall survival (OS) was defined from start of treatment to death (from any cause) or last patient contact. Patients alive and progression-free were censored at the date of last follow-up. Survival endpoints (OS and MFS) were analyzed using the Kaplan-Meier method. Statistical analyses of survival data were performed using SPSS software (SPSS Version 21.0, IMB Corp., Armonk, USA).

Study Approval

Institutional review board approval and patients consent have been obtained for this study. Animal care and procedures were approved by the Institutional Animal Care and Use Committee Office of Bordeaux University, France (APAFIS #8415).

Sequencing protocols, Bioinformatics Analysis, Proteomics and Radiomics

See supplementary material

AUTHOR CONTRIBUTIONS

Study design : MT, CL, SV, AL, AI

Conducting experiments: all authors

Acquiring data: all authors

Analyzing data: all authors

Writing the manuscript: MT, CL, AI

ACKNOWLEDGMENTS

This work was supported by a grant from La Ligue. The authors thank Valérie Vélasco for technical support and Pr Jean-Michel Coindre for pathological review and cohort constitution supervision.

REFERENCES

1. O'Brien JE, and Stout AP. Malignant Fibrous Xanthomas. *Cancer*. 1964;17:1445-55.
2. Weiss SW, and Enzinger FM. Malignant fibrous histiocytoma: an analysis of 200 cases. *Cancer*. 1978;41(6):2250-66.
3. Fletcher CD. Pleomorphic malignant fibrous histiocytoma: fact or fiction? A critical reappraisal based on 159 tumors diagnosed as pleomorphic sarcoma. *Am J Surg Pathol*. 1992;16(3):213-28.
4. Fletcher CD. The evolving classification of soft tissue tumours: an update based on the new WHO classification. *Histopathology*. 2006;48(1):3-12.
5. Coindre JM, Mariani O, Chibon F, Mairal A, De Saint Aubain Somerhausen N, Favre-Guillevin E, Bui NB, Stoeckle E, Hostein I, and Aurias A. Most malignant fibrous histiocytomas developed in the retroperitoneum are dedifferentiated liposarcomas: a review of 25 cases initially diagnosed as malignant fibrous histiocytoma. *Mod Pathol*. 2003;16(3):256-62.
6. Guillou L, and Aurias A. Soft tissue sarcomas with complex genomic profiles. *Virchows Arch*. 2010;456(2):201-17.
7. Savina M, Le Cesne A, Blay JY, Ray-Coquard I, Mir O, Toulmonde M, Cousin S, Terrier P, Ranchere-Vince D, Meeus P, et al. Patterns of care and outcomes of patients with METAstatic soft tissue SARcoma in a real-life setting: the METASARC observational study. *BMC Med*. 2017;15(1):78.
8. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, and Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417-25.
9. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautes-Fridman C, Fridman WH, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol*. 2016;17(1):218.
10. Cancer Genome Atlas Research Network. Electronic address edsc, and Cancer Genome Atlas Research N. Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. *Cell*. 2017;171(4):950-65 e28.
11. Dvorak P, Dvorakova D, Koskova S, Vodinska M, Najvirtova M, Krekac D, and Hampl A. Expression and potential role of fibroblast growth factor 2 and its receptors in human embryonic stem cells. *Stem Cells*. 2005;23(8):1200-11.
12. Xu C, Inokuma MS, Denham J, Golds K, Kundu P, Gold JD, and Carpenter MK. Feeder-free growth of undifferentiated human embryonic stem cells. *Nat Biotechnol*. 2001;19(10):971-4.
13. Eom YW, Oh JE, Lee JI, Baik SK, Rhee KJ, Shin HC, Kim YM, Ahn CM, Kong JH, Kim HS, et al. The role of growth factors in maintenance of stemness in bone marrow-derived mesenchymal stem cells. *Biochem Biophys Res Commun*. 2014;445(1):16-22.
14. Itkin T, Ludin A, Gradus B, Gur-Cohen S, Kalinkovich A, Schajnovitz A, Ovadya Y, Kollet O, Canaani J, Shezen E, et al. FGF-2 expands murine hematopoietic stem and progenitor cells via proliferation of stromal cells, c-Kit activation, and CXCL12 down-regulation. *Blood*. 2012;120(9):1843-55.
15. Babina IS, and Turner NC. Advances and challenges in targeting FGFR signalling in cancer. *Nat Rev Cancer*. 2017;17(5):318-32.
16. Toulmonde M, Penel N, Adam J, Chevreau C, Blay JY, Le Cesne A, Bompas E, Piperno-Neumann S, Cousin S, Grellety T, et al. Use of PD-1 Targeting, Macrophage Infiltration, and IDO Pathway Activation in Sarcomas: A Phase 2 Clinical Trial. *JAMA Oncol*. 2017.

17. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med*. 2015;21(8):938-45.
18. Rooney MS, Shukla SA, Wu CJ, Getz G, and Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*. 2015;160(1-2):48-61.
19. Chan TA, Wolchok JD, and Snyder A. Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma. *N Engl J Med*. 2015;373(20):1984.
20. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. 2015;348(6230):124-8.
21. Davoli T, Uno H, Wooten EC, and Elledge SJ. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science*. 2017;355(6322).
22. Steele CD, Tarabichi M, Oukrif D, Webster AP, Ye H, Fittall M, Lombard P, Martincorena I, Tarpey PS, Collord G, et al. Undifferentiated Sarcomas Develop through Distinct Evolutionary Pathways. *Cancer Cell*. 2019;35(3):441-56 e8.
23. Delespaul L, Lesluyes T, Perot G, Brulard C, Lartigues L, Baud J, Lagarde P, Le Guellec S, Neuville A, Terrier P, et al. Recurrent TRIO Fusion in Nontranslocation-Related Sarcomas. *Clin Cancer Res*. 2017;23(3):857-67.
24. Cho KB, Cho MK, Lee WY, and Kang KW. Overexpression of c-myc induces epithelial mesenchymal transition in mammary epithelial cells. *Cancer Lett*. 2010;293(2):230-9.
25. Yin S, Cheryan VT, Xu L, Rishi AK, and Reddy KB. Myc mediates cancer stem-like cells and EMT changes in triple negative breast cancers cells. *PLoS One*. 2017;12(8):e0183578.
26. Kuzyk A, and Mai S. c-MYC-induced genomic instability. *Cold Spring Harb Perspect Med*. 2014;4(4):a014373.
27. Liu H, Ai J, Shen A, Chen Y, Wang X, Peng X, Chen H, Shen Y, Huang M, Ding J, et al. c-Myc Alteration Determines the Therapeutic Response to FGFR Inhibitors. *Clin Cancer Res*. 2017;23(4):974-84.
28. Delmore JE, Issa GC, Lemieux ME, Rahl PB, Shi J, Jacobs HM, Kastiris E, Gilpatrick T, Paranal RM, Qi J, et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell*. 2011;146(6):904-17.
29. Garcia-Carpizo V, Ruiz-Llorente S, Sarmentero J, Grana-Castro O, Pisano DG, and Barrero MJ. CREBBP/EP300 bromodomains are critical to sustain the GATA1/MYC regulatory axis in proliferation. *Epigenetics Chromatin*. 2018;11(1):30.
30. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, Sanduleanu S, Larue R, Even AJG, Jochems A, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14(12):749-62.
31. Tawbi HA, Burgess M, Bolejack V, Van Tine BA, Schuetz SM, Hu J, D'Angelo S, Attia S, Riedel RF, Priebat DA, et al. Pembrolizumab in advanced soft-tissue sarcoma and bone sarcoma (SARC028): a multicentre, two-cohort, single-arm, open-label, phase 2 trial. *Lancet Oncol*. 2017;18(11):1493-501.
32. Toulmonde M, and Italiano A. PD-1 inhibition in sarcoma still needs investigation. *Lancet Oncol*. 2018;19(1):e6.
33. Daud AI, Loo K, Pauli ML, Sanchez-Rodriguez R, Sandoval PM, Taravati K, Tsai K, Nosrati A, Nardo L, Alvarado MD, et al. Tumor immune profiling predicts response to anti-PD-1 therapy in human melanoma. *J Clin Invest*. 2016;126(9):3447-52.

34. Ou D, Adam J, Garberis I, Blanchard P, Nguyen F, Levy A, Casiraghi O, Gorphe P, Breuskin I, Janot F, et al. Influence of tumor-associated macrophages and HLA class I expression according to HPV status in head and neck cancer patients receiving chemo/bioradiotherapy. *Radiother Oncol*. 2019;130:89-96.
35. Ou D, Adam J, Garberis I, Blanchard P, Nguyen F, Levy A, Casiraghi O, Gorphe P, Breuskin I, Janot F, et al. Clinical relevance of tumor infiltrating lymphocytes, PD-L1 expression and correlation with HPV/p16 in head and neck cancer treated with bio- or chemo-radiotherapy. *Oncoimmunology*. 2017;6(9):e1341030.
36. Team RC. R: A Language and Environment for Statistical Computing.

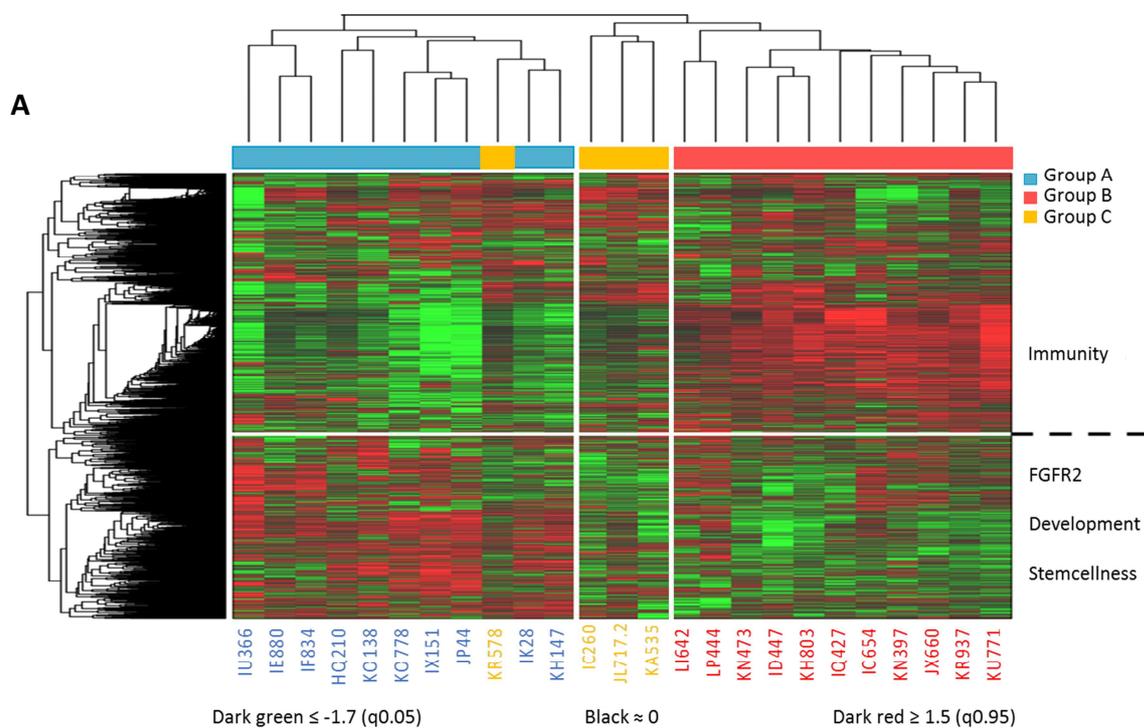
FIGURES

Figure 1. Gene expression profiling allows reproducible classification of UPS

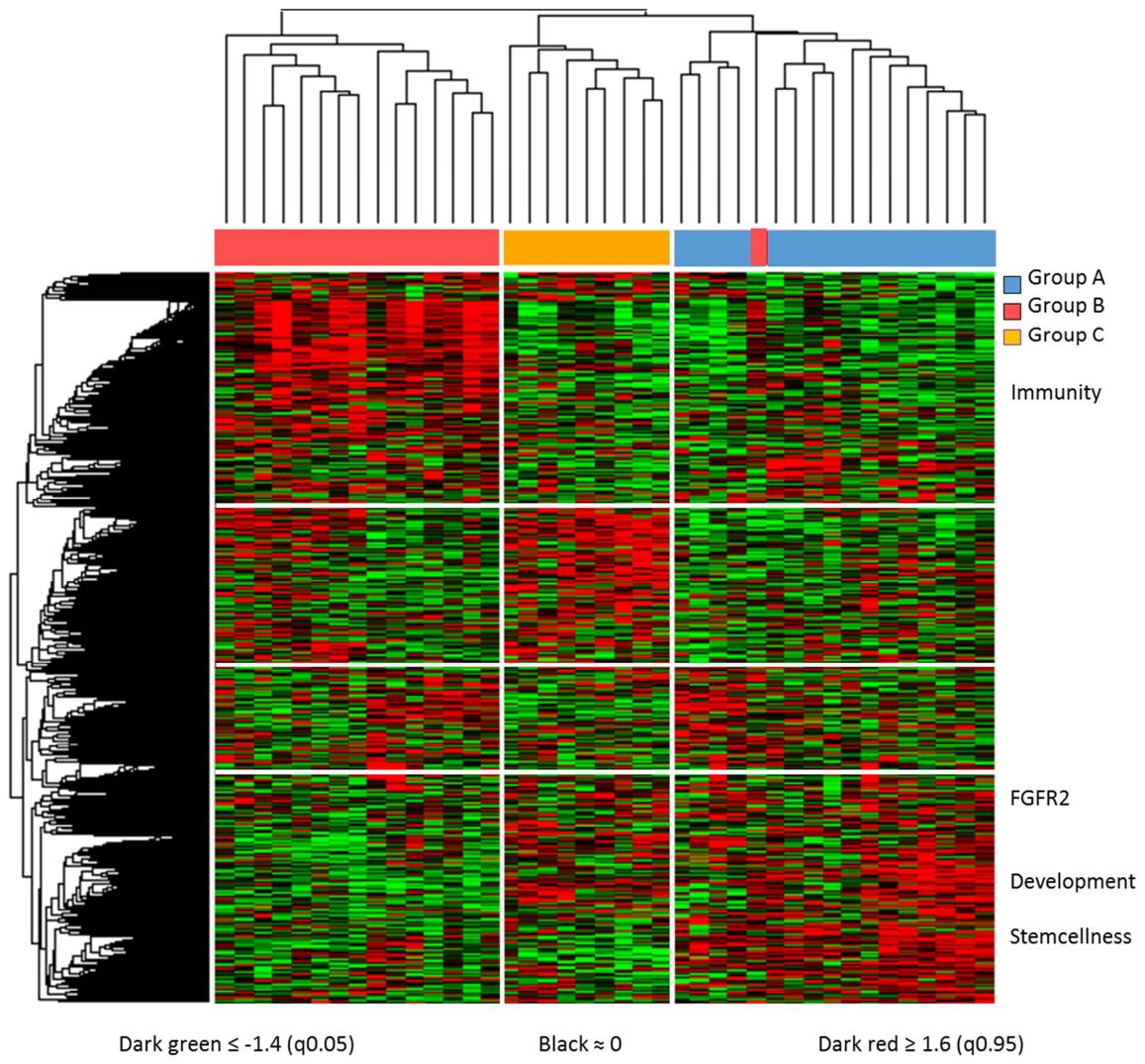
(A) Unsupervised Hierarchical Clustering of RNA-sequencing in UPS from Institut Bergonié (n = 25); 3 groups of patients are identified, with 3 associated gene-clusters.

(B) Unsupervised Hierarchical Clustering of RNA-sequencing in UPS from TCGA consortium (n = 41); 3 groups of patients are identified, with 3 associated gene-clusters.

(C) Analysis of Agreement between Differential EXpression of genes in group A vs B UPS from Institut Bergonié and group A vs B UPS from TCGA consortium reveals a very high correlation (1405 genes, Spearman = 0.83, Pearson = 0.73).



B



C

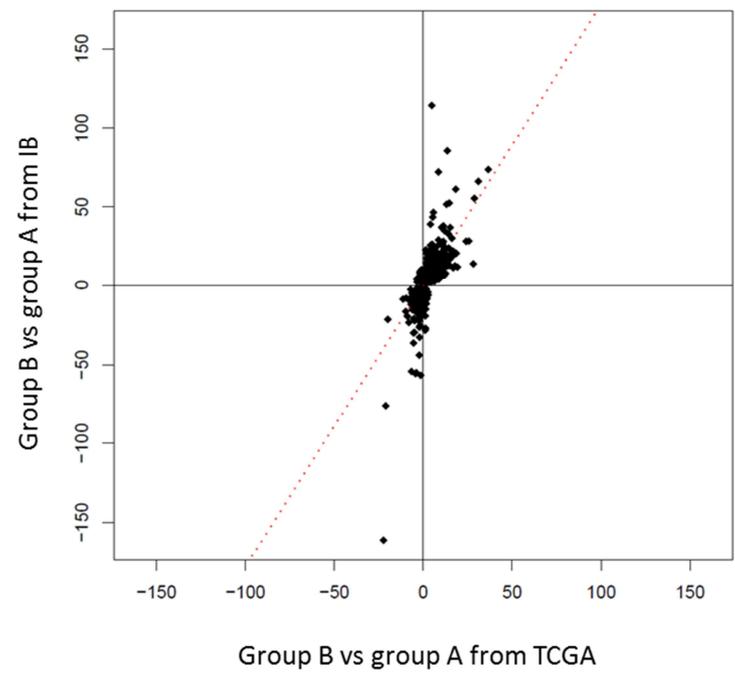


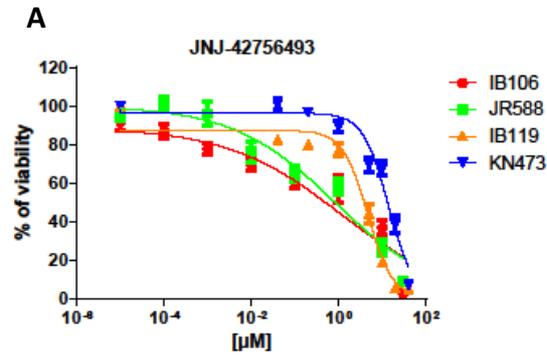
Figure 2. Therapeutic potential of FGFR inhibition in one specific subset of UPS

(A) Antiproliferative activity of pan-FGFR inhibitor JNJ-42756493 in 4 UPS cell lines; Growth curves indicate growth inhibition and IC50 of the 4 UPS cell lines after JNJ-42756493 treatment for 72 hours (n = 6) (****p < 0.0001, one-way ANOVA).

(B) FGFR-inhibitor induces MAPK pathway inhibition in FGFR2 overexpressing cell lines; Phospho-FGFR2/FGFR2 ratio and Phospho-Erk /Erk ratio decrease in FGFR2 overexpressing cell lines after 24h of treatment with JNJ-42756493 at 1 μ M (n = 3) (*p < 0.05 and **p < 0.01, two-way ANOVA) (Western-blot).

(C) Activity of pan-FGFR inhibitor (JNJ-42756493) on cell cycle of 4 UPS cell lines; Top: cell-cycle profiles after 24 hours of treatment with or without JNJ-42756493 at the IC50 analyzed by Propidium Iodide incorporation and flow cytometry; Bottom: cell cycle phase distributions were analyzed with FlowJo software and presented as mean \pm SEM of 3 independent experiments (n = 3) (**p < 0.01 and ***p < 0.001, ns: not significant, two-way ANOVA).

(D) Antitumoral effect of a pan-FGFR inhibitor (JNJ-42756493) in two Patient-Derived Xenograft (PDX) models of UPS; Mice were randomly assigned to receive 30 mg/kg of drug or vehicle; Tumor volume progression curves were drawn over 4 weeks of JNJ-42756493 treatment; The data points represent an average of 11 mice for the JR588 PDX (top) and 8 mice for the KN473 PDX (bottom) (bars, SEM) (***p < 0.001, two-way ANOVA).



Cell lines	IB106	JR588	IB119	KN473
IC ₅₀ (μM)	1,049	0,812	4,639	13,91****

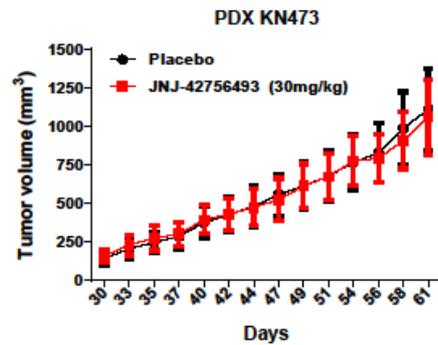
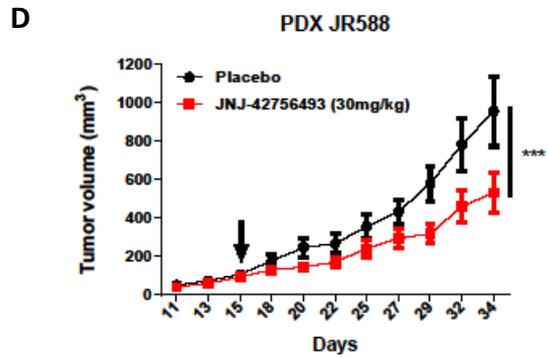
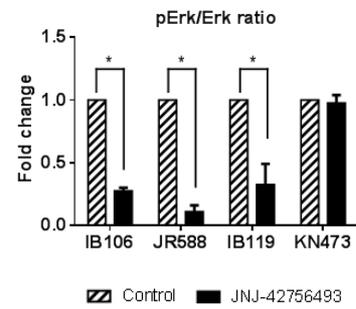
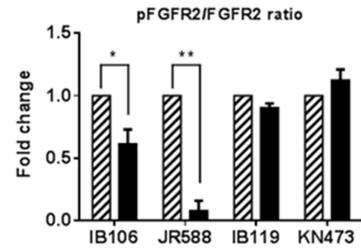
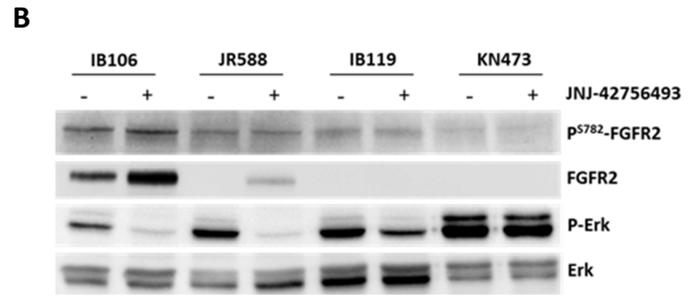
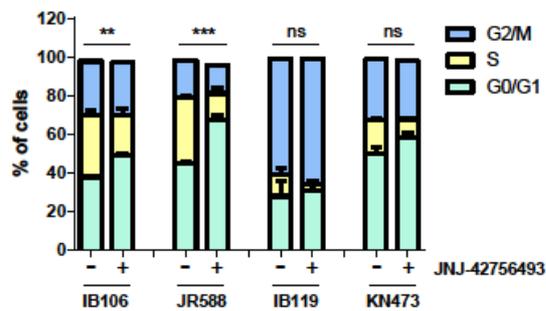
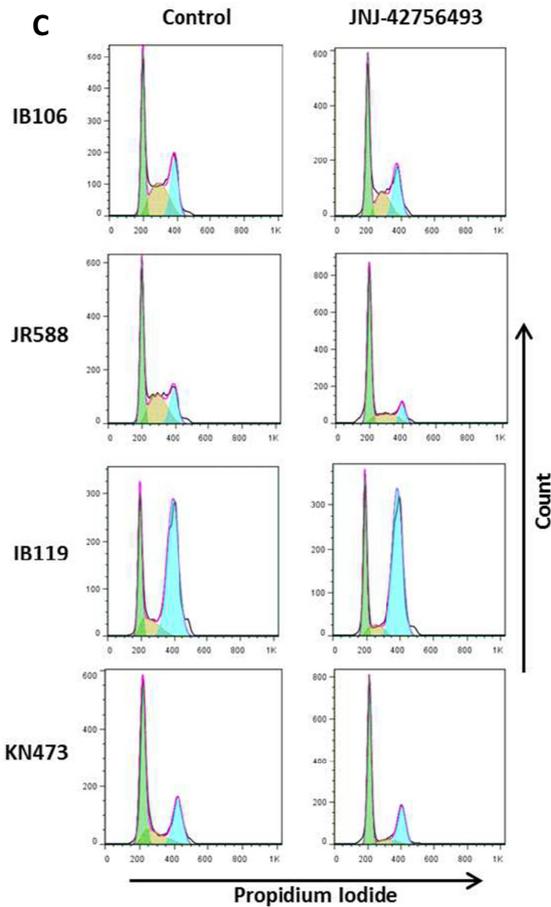


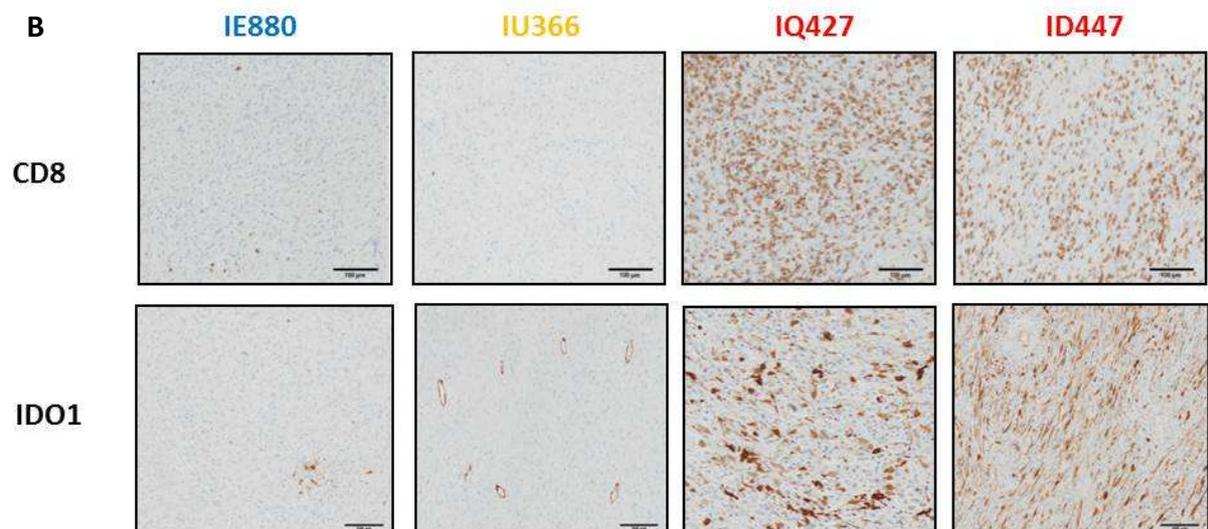
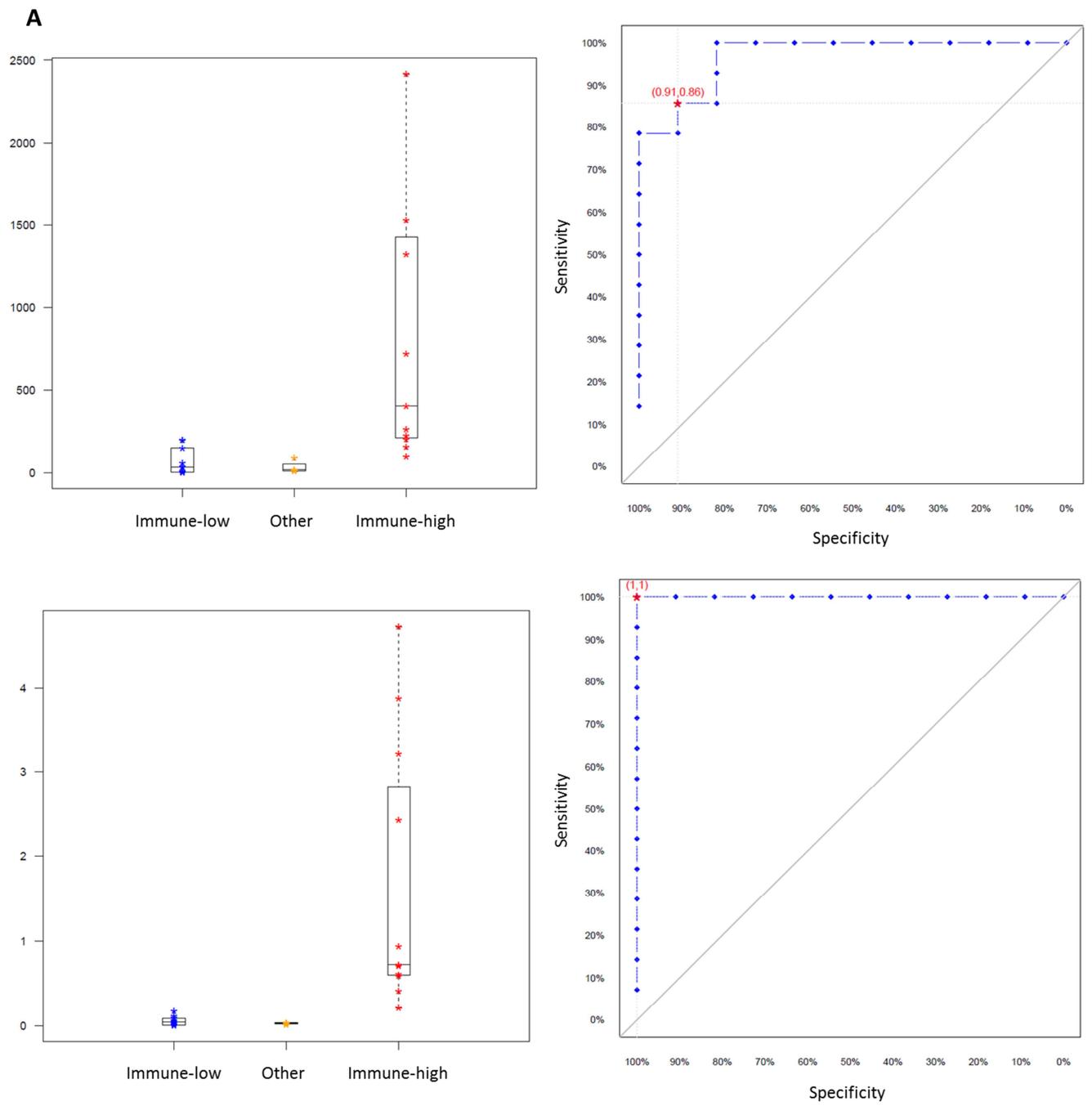
Figure 3. The UPS classification is consistent from tumor immune infiltration patterns to underlying genomic events

(A) Left: Quantification of tumor immune infiltrate density by IHC on tumor slides for CD8 (top) and IDO1 (bottom), in UPS from Institut Bergonié classified by group A = immune-low, B = immune-high, and C = other, according to gene expression clustering (n = 25); The 3 groups of UPS with different gene expression profiles have different tumor immune infiltration patterns. Right: ROC curve of predictive performance of IHC staining to predict RNA-sequencing classification for each marker (accuracy, sensitivity, specificity and AUC); Top: CD8 (cutoff = 146.2, accuracy: 0.88, sensitivity \geq 0.85, specificity \geq 0.9, AUC = 0.97, 95%CI: [0.9-1]); Bottom: IDO1 (cutoff = 0.18, accuracy: 1, sensitivity \geq 1, specificity \geq 1, AUC = 1, 95%CI: [1-1]).

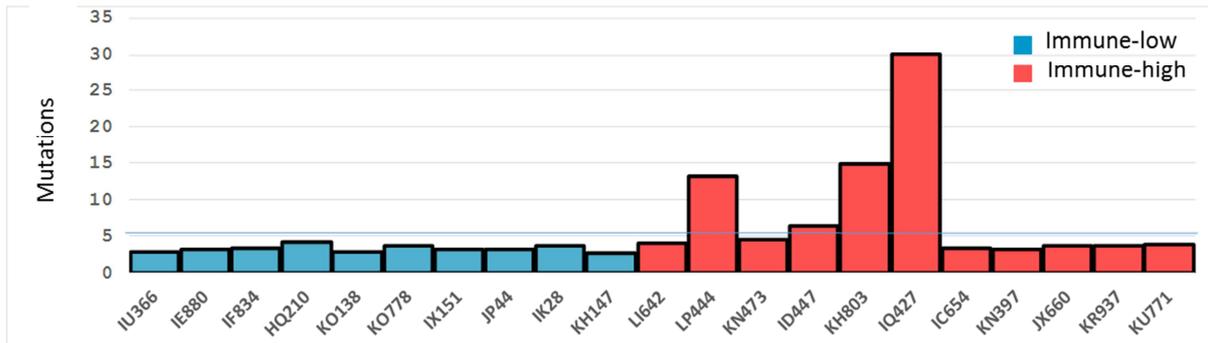
(B) Expression of CD8 (upper panel) and IDO1 (lower panel) by Immunohistochemical (IHC) staining shows concordance with RNA profiling group. Left: a sample belonging to group A negative for CD8 and IDO1 on IHC (IE880), middle: a sample belonging to group C negative for CD8 and IDO1 on IHC (IU366), right: 2 samples belonging to group B positive for CD8 and IDO1 on IHC (IQ427 and ID447) (magnification \times 100).

(C) Frequency of mutations per Megabase (MB) in immune-low and immune-high UPS from Institut Bergonié, ordered according to gene expression clustering (n = 21).

(D) Gene focal somatic copy-number alterations, focusing on deletions in tumor suppressor genes (TSG) in immune-low and immune-high UPS from Institut Bergonié, ordered according to gene expression clustering (n = 21). Enrichment in TSG deletions is found in immune-low UPS.



C



D

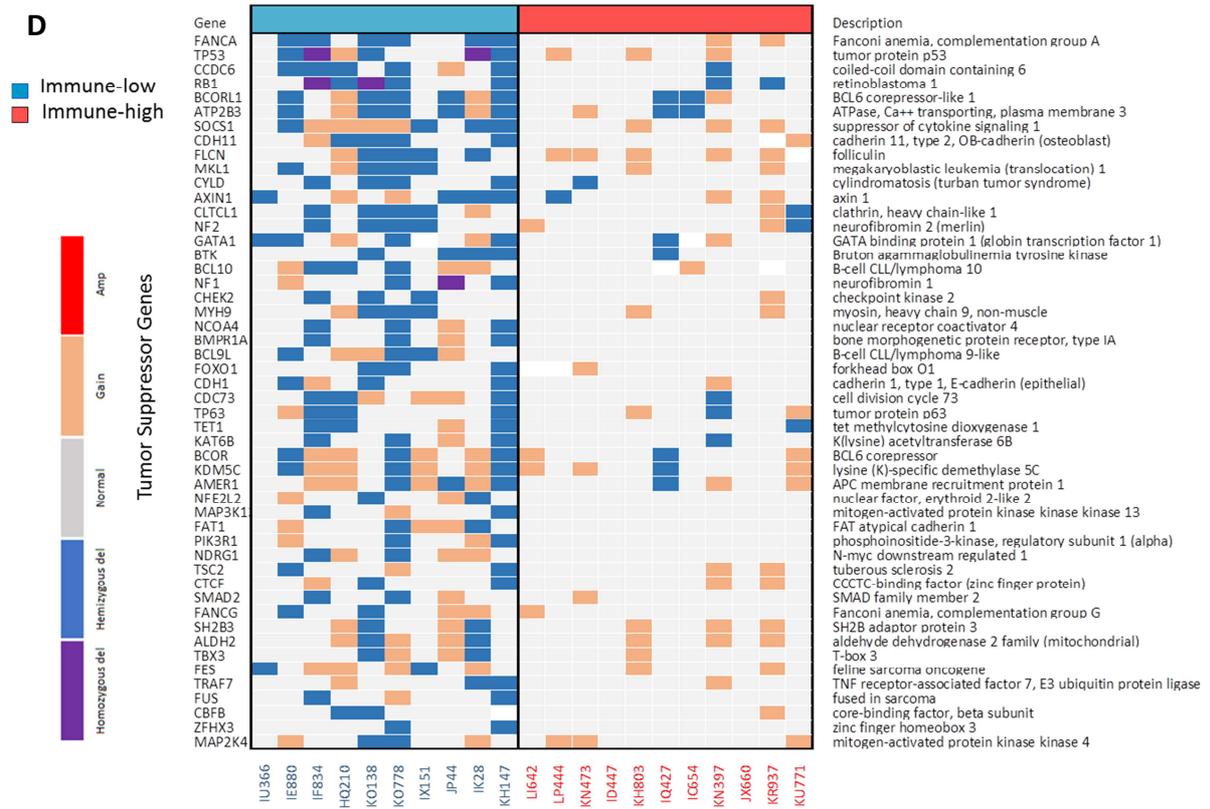
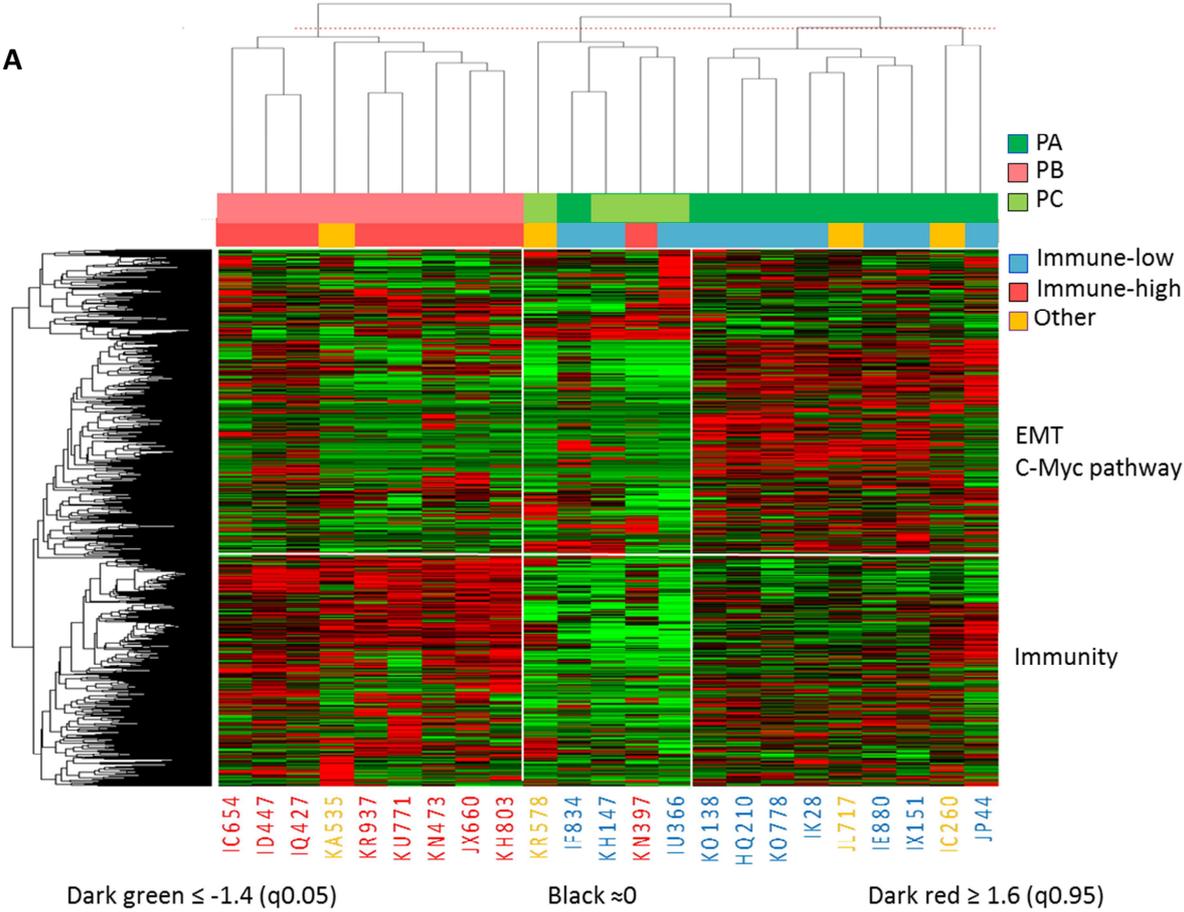


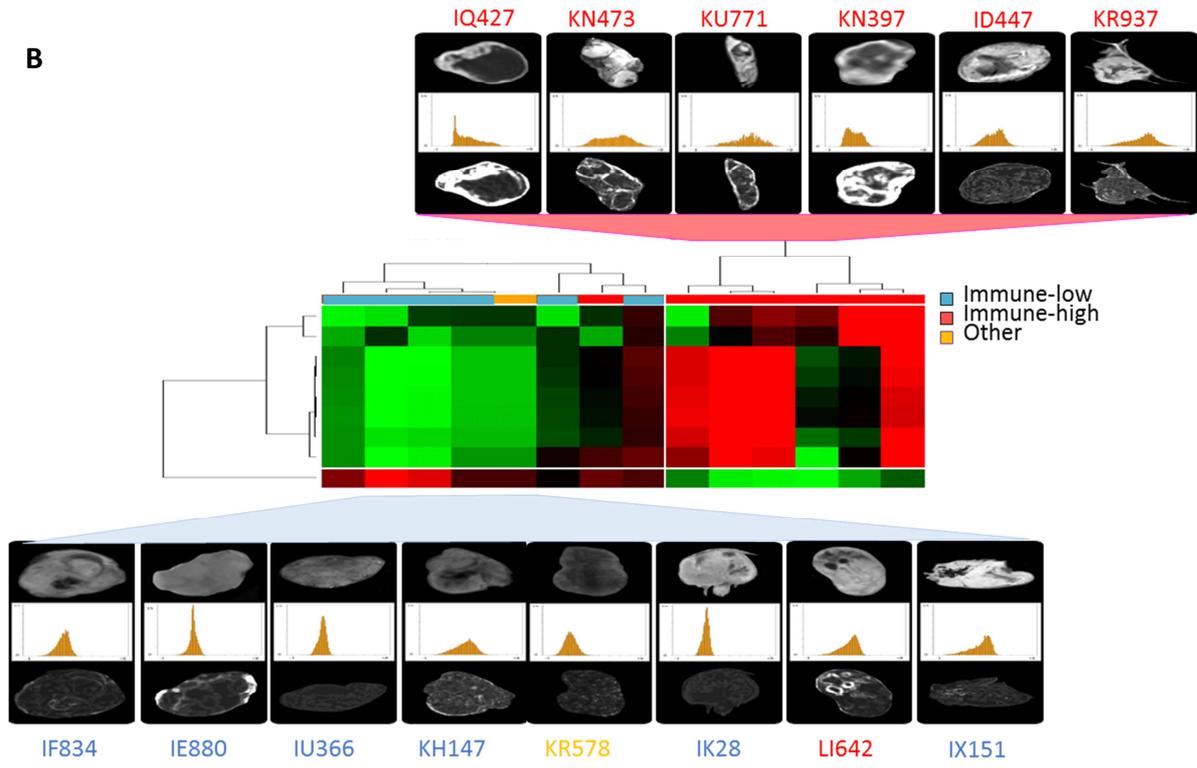
Figure 4. The UPS classification includes specific proteomic profiles and predictive radiomic features

(A) Unsupervised Hierarchical Clustering of Protein Expression in UPS from Institut Bergonié (n = 23); 3 groups of patients are identified, PA, PB, PC, with 3 associated clusters of proteins and 565 proteins with differential expression (FC ≥ 2, p = 0.01). Corresponding color labels of samples according to RNA-sequencing profiling are shown. There is a high agreement on classification labels (precision: 82%) between proteomics and RNA-sequencing clusterings.

(B) Hierarchical Clustering of UPS patients according to the 9-feature radiomic signature allows discriminating immune-high UPS from non-immune-high UPS (specificity: 100%, sensitivity: 86% and accuracy: 93%) (n = 14). Corresponding color labels of samples according to RNA-sequencing profiling are shown.



B



SUPPLEMENTARY MATERIAL

Supplementary Methods

Supplementary Tables 1-7

Supplementary Figures 1-5

SUPPLEMENTARY METHODS

Sequencing protocols

DNA extraction was performed with a QIAamp DNA mini kit (Qiagen, Courtaboeuf, France) as described by the manufacturer. The purity of the genomic DNA was measured with a NanoDrop 1000 apparatus (NanoDrop Products), and the quantity was estimated by a fluorescence-based method using a Qubit double-stranded DNA BR assay kit and a Qubit fluorometer (Life Technologies, Carlsbad, CA, USA) according to the manufacturer's instructions. For Exome-sequencing, capture of DNA exome regions to sequence was performed via the Agilent SureSelect XT All exon Human V5 kit, and sequencing was performed on a HiSeq 2500 platform (Illumina Inc. San Diego, CA, USA) via the Illumina TrueSeq paired-end 125x2 reads protocol. Exome-sequencing coverage produced by the sequencer was at least fixed at 60X and 30X for Tumor DNA and Constitutional DNA, respectively. For RNA extraction, quantity and purity of double-stranded cDNA from total RNA were verified by GATC Biotech using an Agilent 2100 Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA, USA). Extractions passed strict control criteria concerning quantity (1 µg per sample), concentration (> 20 ng/µl), purity (OD 260/280: 1.8 - 2.2) and RIN value (≥ 8). Nucleic Acids were dissolved in RNase-, DNase- and protease-free molecular grade water or Tris bufferRNAFor RnaSeq. RNA-sequencing was performed on a HiSeq 2500 platform (Illumina Inc. San Diego, CA, USA) via the Illumina TrueSeq PolyA Strand-specific cDNA library, paired-end 125x2 reads protocol. Number of paired-end reads produced by the sequencer was at least 30M per sample. NGS sequencing was performed by GATC Biotech (Konstanz, Germany).

Bioinformatics Primary Analysis

For both ExomeSeq and RnaSeq, bioinformatics analysis was performed based on the Hg19/Gh37 version of the human genome. Briefly, pre-alignment and quality control of sequences (.fastQ.gz) was done via FastqPairedEndValidator to check that R1 and R2 sequences were correctly paired (<https://github.com/orionzhou/luffy/blob/master/archive/FastqPairedEndValidator.pl>), Clumpify to remove reads having identical sequence (<https://github.com/BioInfoTools/BBMap>) and Sickle to trim low quality sequences at 5' and 3' ends (Sickle - <https://github.com/najoshi/sickle>). SeqPrep package (SeqPrep - <https://github.com/jstjohn/SeqPrep>) was used to remove sequencing adaptors from raw reads. This package also detected the proportion of nucleotide fragments whose R1 and R2 paired-end reads were overlapping and merged them into single-end reads. To prevent double coverage bias due to overlapping R1 and R2 sequences and keep exploiting those fragments, a home-made python script was developed to split those merged reads into new non-overlapping R1 and R2 paired-end reads. Alignment of quality controlled DNA sequences to .bam files was performed via Bowtie2 (1) with the “-very-sensitive” alignment strategy parameter. Alignment of RNA sequences was performed using Tophat2 (2) and Bowtie2 (1) on both the UCSC

hg19 reference genome and transcriptome. Remaining PCR duplicates reads in Post-alignment .bam files were removed via the PicardTools suite MarkDuplicates module (PicardTools - <https://broadinstitute.github.io/picard/>).

Bioinformatics Secondary Analysis

Production of raw genetic alteration files (.vcf/.bcf) was performed via Samtools/Bcftools (3) with a quality cut-off per base of Phred 20. Calculation of the number of aligned reads covering each exon was performed via HtSeq (4). Quality reports of raw reads were built with FastQC (FastQC - <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Annotation of SNV and InDel genetic variants was reported by Annovar (5), producing a wide variety of genetic, genomic, somatic and polymorphic information about the variants, including the type, localization and standard protein annotation information. The likelihood of strand bias affecting each variant proposal was estimated via a proportion test of the number of alternative bases measured in each strand compared to the reference bases using prop.test in R (6). A position was called variant if at least an alternative allele was found in Constit and/or in Tumor. The likelihood of a variant being heterozygous was estimated via a binomial test comparing the alternative allele ratio against an error rate of 5 bases over 100. If the test was rejected, the genotype of the variant was called AB (heterozygous), and AA (homozygous) otherwise (BB if the alternative allele ratio was higher than 0.5). Genotyping calling was challenged using DeepVariant as alternative method (7). Briefly, to gain in specificity without losing sensitivity for InDels detection, an InDel detected by Samtools was only kept if validated by DeepVariant.

For the analysis of DNA and RNA mutations, a variant was annotated as somatic in case the genotypes of the Constit:Tumor samples of the same patient were respectively AA:AB, AA:BB, BB:AB, BB:AA or AB:BB. In order to report high-quality somatic mutations, variants were filtered by depth of coverage ≥ 10 reads covering a position; absence of strand bias; location in exonic or in splicing sites; non-polymorphic status based on Exac (8) and 1000 genome (9) Minor Allele Frequency $\leq 1 \times 10^{-3}$; not falling into repeating genomic regions according to Repeat Masker (10); and non-synonymous. Mutations associated to somatic LOH regions were detected in first instance by their characteristic genotype (AB in Constit and BB in Tumor). Then, LOH regions were estimated by analyzing the genotype status of common exonic polymorphisms (1000 genome DB v.2017_01_17) in a region of ± 2 MB around the somatic mutation characterized by the BB genotype. The mutation was deemed to be in a LOH region if the mean alternative allele ratio of the polymorphisms in that region was ≥ 0.8 (genotype inferred: BB).

Mutational Burden

The overall tumor mutational burden included all mutations having a joint Constit and tumor genotype of AA:AB, AA:BB or BB:AB and characterized by depth of coverage ≥ 10 reads in both Constit and Tumor; absence of strand bias in either Constit or Tumor; location in exonic or splicing sites; not falling into repeating regions according to Repeat Masker; and non-synonymous. LOH mutations were excluded. Counts were expressed as number of mutations per coding mega-base (shortened as MB coding) and based on a 35Mb exonic reference size.

Detection of Copy Number Alterations and Ploidy and Group Enrichment by Genomic Alterations

DNA Copy number variation was estimated via an in-house method, Copy_Number by Exon, written in R, which calculates copy number ratios at the exon level between the Tumor DNA and Normal DNA of the same patient and reports them at gene and genomic region level. The genomic amplification threshold was set at ≥ 6 tumor alleles

(Tumor vs Normal DNA ratios ≥ 3 , log Ratios ≥ 2.58), and focal amplifications were retained in regions of ≤ 10 Mb. We estimated the enrichment in genomic deletions between 2 groups of samples via hypergeometric test and paired t-test. For the hypergeometric test, we counted the number of deletions (either homozygous or hemizygous) or non-deletions (normal, gain, amplification) events per gene and sample in the 2 groups of samples. Using the resulting 2x2 contingency table of counts, the null hypothesis of homogeneity of deletions between the 2 groups was tested via hypergeometric test implemented via the `fisher.exact` test in R (6). This also estimated the Odds Ratio (OR) between the 2 groups and the related confidence interval (CI). For the paired t-test, the number of deletions events for each gene in the 2 groups was scaled in order to take into account the difference in sample size by group. The normalization factor was the ratio between the sample sizes of the 2 groups. The null hypothesis that the mean (paired) difference of the gene deletion counts between the 2 groups was 0 was tested with the function `t.test` in R (6) with the option set to paired test. P-value and confidence interval of the mean difference were reported. After adjustment of p-values for the 2 tests, a group was reported significantly enriched in genomic deletions if both tests resulted in significant p-values. The enrichment in deletions between the 2 groups was tested on all genes.

Common list of Oncogenes and Tumor suppressor genes

We used for that purpose the set of genes published by Vogelstein B et al. (11) and in the COSMIC database (COSMIC v81 (09-05-2017) Cancer Gene Census (12) available at: <http://cancer.sanger.ac.uk/cosmic/download>). The list joint totaled 401 genes.

Detection, Annotation and Interpretation of Genetic variants

Resulting mutations and copy numbers were further annotated and interpreted via direct access to genetic, genomic and clinical trial public data portals. Briefly, we checked whether rare somatic mutations were already described in COSMIC database and/or if mutations affected known TSG/ONCO genes. Then we proceeded with the annotation of mutations of unknown origin. First we looked at mutations that could impair genes with copy number deficiencies, looking for a second mutational hit that could completely inactivate the gene product. Inactivation can come from missense mutations predicted to affect protein functionality or from truncating mutations preventing the complete conformation of the protein. We finally looked at mutations of any nature (missense or truncating) that could affect the protein of genomically normal, gained or amplified genes. We applied a similar procedure to mutations located in LOH regions.

Expression Normalization

For RNA-Sequencing analysis of quantitative expression of genes, gene count data were normalized according to the VOOM method (13), which transforms raw count values to log₂-counts per million (logCPM), estimates the mean-variance relationship and uses this relationship to compute appropriate observational-level weights.

Unsupervised Clustering

Unsupervised clustering analysis of RNA-Sequencing expression data was performed using agglomerative hierarchical clustering with distance criteria (1-Pearson_Correlation) and linkage criteria equal to average via the function `hclust` available in R (6). We performed iteratively this procedure on subset of genes with increasing variability based on their standard deviation. The choice of the number of genes to use for freezing the clustering configuration was determined empirically as follows. For each clustering performed at a given standard deviation cut-off, we measured the anti-

correlation between the 2 groups placed at the highest level in the clustering hierarchy. We assumed that clustering with anti-correlations close to 0 were associated with grouping by chance between samples. We selected the SD cutoff for which sample anti-correlation was at the lowest negative level whilst the number of genes left was no less than 3000. This procedure was meant to remove genes with weak variability that could bias clustering via spurious random correlations and at the same time assuring that sample grouping was determined by the largest as possible number of genes. To assess the robustness of clustering results at a given SD threshold, we performed consensus clustering via the package ConsensusClusteringPlus available in BioConductor (14). Consensus was established after 10,000 iterations on the subset of samples obtained by leave-n-out of 40% of samples.

Differential Gene Expression

RNA-Sequencing differential gene expression between groups of samples was performed using the statistical t-test from the R package LIMMA (15), which calculates fold changes and nominal p-values related to each gene starting from raw expression values and the normalization weights produced by VROOM (13). The set of nominal p-values from each test were adjusted according to the Benjamini-Hochberg adjustment via the package multiple.test available in R (6). We defined the significantly up- or down-regulated transcripts using an FDR threshold of 0.01. The fold-change used to further filter the differential gene expression was set to a minimum value of 2.

Gene-Set Enrichment Analysis

We assessed enrichment in biological pathways between RNA-Sequencing groups based on Broad Institute's Molecular Signature Data Base gene sets (MSigDB)(16) and the CYBERSORT LM22 immuno gene sets (17). Gene-set enrichment was performed as follows. The total count of genes (UNIVERSE) was set to the total number of genes for which the Differential Gene Expression (DGE) calculation was performed (18,399 annotated genes in our study). For each biological pathway we counted the genes that participated in the pathway (IN.GENSET) and its complement with respect to the UNIVERSE (NOT.IN.GENSET). Also the UNIVERSE was split in differentially expressed genes (SIG.DGE) and their complement (NOT.SIG.DGE). Then we built a 2x2 contingency matrix holding the count of differentially expressed genes in the pathway (SIG.DGE.IN.GENSET) and its complement (SIG.DGE.NOT.IN.GENSET) and the count of not differentially expressed genes in the pathway (NOT.SIG.DGE.IN.GENSET) and its complement (NOT.SIG.DGE.NOT.IN.GENSET). We performed the hypergeometric test on this table via fisher test in R (6). The raw p-values of all tests for each pathway were adjusted by Benjamini Hochberg procedure via the multiple.test package in R (6). We additionally performed the calculation of the OR based on the above contingency table and its 95% CI (results of the fisher.test function R base package). We tagged as significantly ENRICHED all pathways with a significant hypergeometric test at adjusted p-value ≤ 0.01 and the lower bound of the CI of their OR not lower than 2. We ranked all ENRICHED-tagged pathways by the count of differentially expressed genes in the pathway (SIG.DGE.IN.GENSET). Then we wanted to qualitatively assess pathways significantly over-expressed or under-expressed in sample groups. To that extent, and for each pathway and each group, we calculate the ratio between the numbers of genes that were over-expressed in the pathway and the numbers of DGE genes in the pathway. A ratio close to 0 suggests that the pathway was under-expressed while a ratio close to 1 suggested that the pathway was over-expressed in the reference group. Ratios that were close to 0.5 were considered informative of no enrichment in either group.

Detection of fusion genes

Three independent algorithms, FusionMap version 10.0.1.29 (18), defuse version 0.6.1 (19) and Starfusion version 1.6.0 (20), were used for the detection of fusion genes. The results of StarFusion were further in-silico validated by the application of Fusion Inspector (21). The same FastQ curated reads were given in input to the 3 algorithms for each sample. Default parameters were used for the 3 algorithms. In order to assess the reproducibility of the fusion genes among algorithms and the presence of recurrent fusion genes, we proceeded as follow: Using a R dplyr (6) in-house script we pulled together in a single table the information about the fusion gene partners (5' and 3'), their chromosome and 5' starting position and the total number of reads supporting the fusion proposal, plus the information of the tool generating the results and the sample in which the fusion was detected. We then sorted the table by fusion gene, number of tools detecting the fusion and mean number of reads for each. The resulting table was reduced to the only fusion gene proposals that were detected by 3 tools. This final table was analyzed to assess the presence of recurrent fusion genes among study samples. We assessed the enrichment in fusion genes between the two main groups of samples using a hypergeometric test. We counted the presence and the absence of fusion gene events in each of the 2 groups of samples. Using the resulting 2x2 contingency table of counts, the null hypothesis of homogeneity of fusion gene counts between the 2 groups was tested via hypergeometric test implemented via the fisher.exact test in R (6) that also estimated the OR between the 2 groups and the related CI.

AGDEX algorithm

Agreement between Differential Gene Expression detected in 2 independent experiments was assessed using the AGDEX algorithm (21). Briefly, the subset of differentially expressed genes reported in both experiments were retained and their fold change created. Agreement in the pairwise fold changes in both experiments was assessed by the Spearman coefficients of non-parametric regression, the Pearson linear correlation corr function and the Robust Linear Model regression rlm function in R (6). A coefficient greater than 0.5 was associated to high AGDEX.

Proteomics

Proteomic analysis sample preparation

A batch of 24 samples slides were selected by pathologists and sent to Nantomics Proteomics facility, for sample preparation as previously described (22). Two consecutive tissue sections per sample were prepared. The first sections were stained with hematoxylin and eosin (H&E), and were used to guide tumor area selection. The second sections were mounted on DIRECTOR® slides (Expression Pathology, Rockville, MD), deparaffinized with xylene, stained with hematoxylin, and subjected to microdissection based on the tumor-specific markup by a board-certified pathologist using MMI CellCut laser microdissection system (Molecular Machines and Industries, Eching, Germany) equipped with UV laser. Collected cells after laser ablation were heated in Liquid Tissue® buffer (Expression Pathology, Rockville, MD) at 95 °C for 1.5 h, and subsequently incubated with trypsin (Promega, Madison, WI) for 16h at 37 °C. A Micro BCA assay (Thermo Fisher Scientific, Rockford, IL, USA) was performed to determine total protein concentration. Samples were received, processed and analyzed by mass spectrometry in one single batch.

Data-independent acquisition (DIA) Liquid Chromatography – Mass Spectrometry (LC-MS/MS)

All LC-MS analyses were performed with an Ultimate 3000 UHPLC system coupled to a Q-Exactive HF mass spectrometer, as previously described (23, 24). For all samples a minimum of 2 µg tumor protein digest was injected per sample. Samples were loaded onto the trap column (Acclaim PepMap, 0.10 mm × 20 mm, 5 µm C18 particles), prior to

separation over the analytical column (Acclaim PepMap, 0.75 mm × 250mm, 2 μm C18 particles) at a flowrate of 300 nL/min. Mobile phase A was 0.1% formic acid in water and mobile phase B was 0.1% formic acid in acetonitrile. The mobile phase gradient was composed of the following steps (time, %B): 0, 1%; 10, 1%; 15, 6%; 90, 24%; 100, 36%; 100.1, 90%; 107, 90%, 113, 1%. For DIA analyses, the MS1 scan covered precursors between 380 and 1200 m/z at 60k resolution. MS2 scans were acquired with a precursor isolation window of 13 Th, HCD fragmentation (normalized collision energy = 27) and at 30k resolution. A total of 45 MS2 windows were used to cover a range from 380 to 920 m/z.

DIA data analysis

DIA data analysis was performed using Spectronaut™ Pulsar 12.0 (Biognosys, Schlieren, Switzerland) software (23), querying an in-house spectral library. Upon obtainment of proteins raw precursors, each of them was assigned a false-discovery rate Q-value expressing precursor protein specificity, a Q-value > 0.01 indicating non-specific protein precursor. Precursors with Q-value above that threshold were flagged as Not Evaluated (NA). Proteins were assigned the value of the most frequent precursor passing the Q-value threshold or NA otherwise. Missing values were imputed by the median value of the protein in the analyzed batch of samples. For the specific study, a total of 5091 proteins were detected. Among them, 3215 proteins were used for quantitative analysis. Unsupervised cluster consensus, hierarchical clustering and differential expression analyses were executed as for RNA-Sequencing. A protein was considered differentially expressed when the Benjamini-Hochberg adjusted p-value was lower than 0.01 and the absolute value of the Linear fold change was equal to or greater than 2. Agreement of analysis results between differential gene expression and differential protein expression were assessed using the AGDEX algorithm (21). For gene-set enrichment analysis, we applied the same procedure as for RNA-Sequencing, taking into account the different proportion of genes covered by RNA-Sequencing and Proteomics technologies (respectively 18K and 5K). Hence, we first removed all non-detectable proteins from the annotation files of each Msig (16) and CyberSort (17) gene-sets and all gene-sets which were covered ≤ 30% of genes were not included in analysis. Then, gene-set by gene-set, we extracted the proteins belonging uniquely to that gene-set (and not to any other gene-set) and retained only the differentially expressed proteins with linear absolute fold change above or equal 1.5 and Benjamini adjusted p-value below 0.05. For each gene-set we empirically assigned a score of -1 (immune-low-associated) if more than 75% of fold change values were negative, a score of +1 (immune-high-associated) if more than 75% of fold change values were positive and score of +2 (again immune-high-associated) if 100% of values were positive. Pathways not satisfying these scoring criteria were excluded from analysis because their enrichment was considered not significant.

Radiomics

MRIs were all performed on 1.5-Tesla magnets. Ranges of slices thickness, repetition time and echo time were: 3-5 mm, 500-920 msec and 8.7-15 msec, respectively. Contrast-enhanced spin-echo T1 weighted-imaging with a static Gadolinium-chelates injection of good quality (CE-T1-WI) was used. Each tumor volume was manually segmented, slice-by-slice, by a senior radiologist with expertise in STS imaging (AC). Radiomic features extraction was done with OleaSphere® Software (v3, sp14, La Ciotat, France). All images were re-sampled using a Bspline interpolation in order to obtain common voxels of 1x1x4mm³ (25). The signal intensities (SIs) of each sequence were standard scaled. SIs were then discretized into 64 fixed discrete values (bins). A total of 62 3D-radiomic features were extracted: 16 from shape analysis, 17 from first-order (i.e. histogram-based) analysis, and 29 from grey-level matrices (i.e. second-order analysis). The list of features is provided in Supplementary Table 8. Differences in means of each the features were

assessed with student t-test in R (6) (range of raw p-value: 0.01-0.05). Unsupervised clustering based on this selection of features was performed using agglomerative hierarchical clustering with distance criteria (1-Pearson_Correlation) and linkage criteria equal to Average via the function hclust available in R (6). The discriminant power of the discriminative radiomic features was characterized by specificity, sensitivity and accuracy.

SUPPLEMENTARY TABLES

Supplementary Table 1. Patients characteristics (n = 135)

	Genomically - characterized cohort (n = 25)	IHC- characterized cohort (n = 110)
	n (%)	n (%)
Gender		
Male	11 (44)	56 (50.9)
Female	14 (56)	54 (49.1)
Median age (year) (min-max)	66 (24-88)	62 (19-95)
Site		
Head and neck	1 (4)	2 (1.8)
Internal trunk	1 (4)	2 (1.8)
Trunk wall	9 (36)	18 (16.4)
limbs	14 (56)	88 (80)
Median size (mm) (min-max)	70 (24-184)	90 (20-250)
Grade		
2	12 (48)	24 (21.8)
3	13 (52)	86 (78.2)
Metastatic at diagnostic		
No	24 (96)	107 (97.3)
Yes	1 (4)	3 (2.7)
Surgery		
No	0 (0)	0 (0)
Yes	25 (100)	110 (100)
Surgical margins		
R0	13 (52)	51 (46.4)
R1	12 (48)	39 (35.5)
R2	0 (0)	3 (2.7)
Not assessable	0 (0)	17 (15.4)
Chemotherapy		
No	22 (88)	68 (61.8)
Adjuvant	3 (12)	39 (35.5)
Palliative	0 (0)	3 (2.7)
Radiation therapy		
No	11 (44)	31 (28.2)
Yes	14 (56)	79 (71.8)
Locoregional Relapse		
No	21 (84)	84 (76.4)
Yes	4 (16)	26 (23.6)
Metastatic Relapse		
No	20 (80)	72 (65.5)
Yes	5 (20)	38 (34.5)

IHC: immunohistochemistry

Supplementary Table 2. Genes differentially expressed in group B versus A in UPS from Institut Bergonié (n = 21)

Genes are listed by decreasing linear fold change and differential expression significance (UP, EQ and DOWN, according to fold change), genes overexpressed in group B are at the top of the list and genes overexpressed in group A are at the bottom.

Supplementary Table 3. Enrichment in Hallmark of Cancer, Gene Ontology and CyberSort LM22 gene-sets in group B versus A in UPS from Institut Bergonié (n = 21)

Genesets are sorted by hypergeometric test status (ENRICHED on top) and by lower bound of the OR Confidence Interval (decreasing).

Supplementary Table 4. Genes differentially expressed in group B versus A in UPS from TCGA cohort (n = 32)

Genes are listed by decreasing linear fold change and differential expression significance (UP, EQ and DOWN, according to fold change), genes overexpressed in group B are at the top of the list and genes overexpressed in group A are at the bottom.

Supplementary Table 5. Mutated genes in UPS from Institut Bergonié (n = 24)

Mutations are characterized by location on chromosome and gene, polymorphism, prediction of functional effect, code and type of base modification and frequency in each of the 24 samples.

Supplementary Table 6. Gene copy number variations in group A and B in UPS from Institut Bergonié (n = 21)

Most Frequent Losses: genes are ordered according to the frequency of deletions in the corresponding gene in the whole cohort (column “%_DELL(ALL)”) ; Most frequent gains: genes are ordered according to the frequency of gains in the corresponding gene in the whole cohort (column “%_GAIN_TOT”).

Supplementary Table 7. AGDEX identifies 114 protein/gene couples differentially expressed in both PA (n = 9) vs PB (n = 10), and immune-low (n = 10) vs immune-high (n = 11) groups in UPS from Institut Bergonié

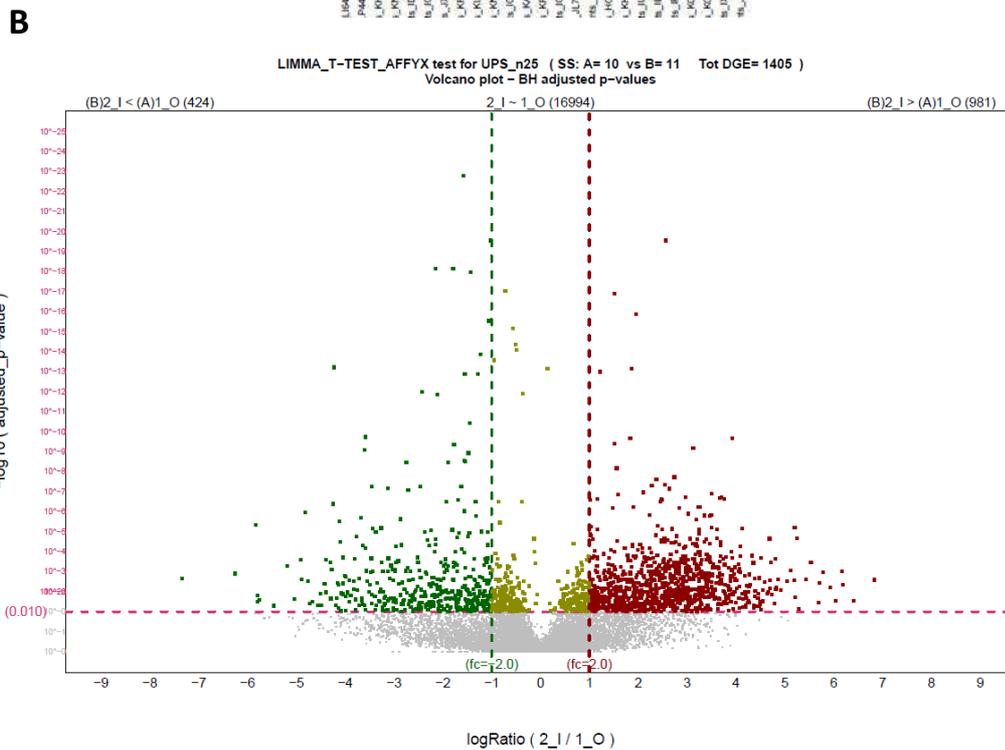
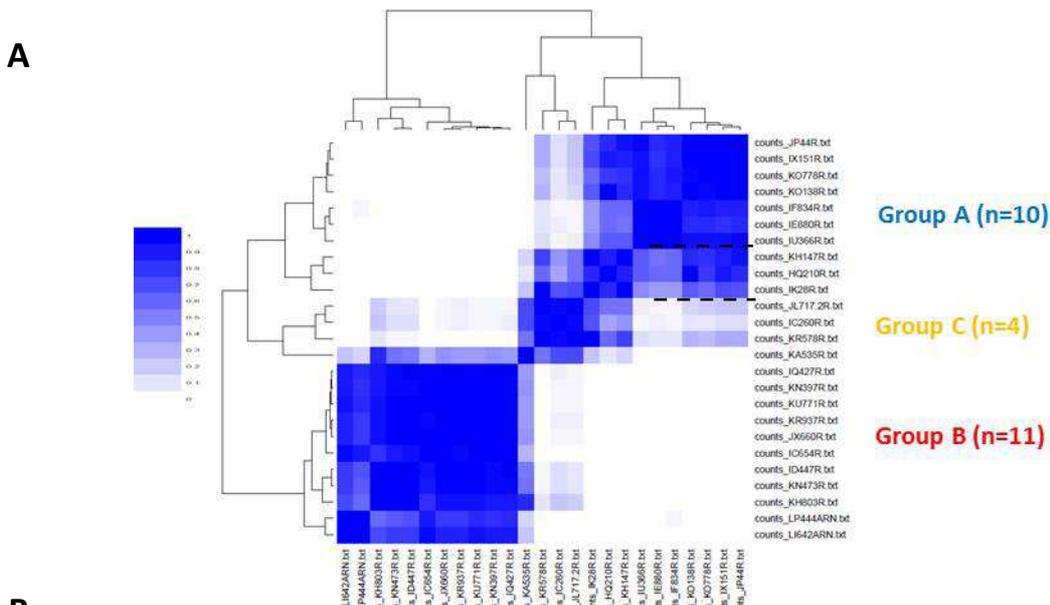
Supplementary Table 8. Radiomic features analyzed in the available pre-treatment contrast-enhanced T1-weighted MRI of the UPS patients from Institut Bergonié (n = 14)

The nine textural features that were found statistically significant in discriminating immune high from non-immune-high UPS patients (T-test, $p < 0.05$) that were retained for the radiomics signature are underlined in red. These textural features are related to the dispersion of signal intensities, lower kurtosis (i.e. increased flatness of the histogram of signal intensities), and higher short run emphasis, due to increased short-run high grey-level emphasis. This last variable indicates finer and more heterogeneous texture (i.e. contrast-enhanced areas inside the tumor volume) in immune-high UPS on contrast-enhanced MRI.

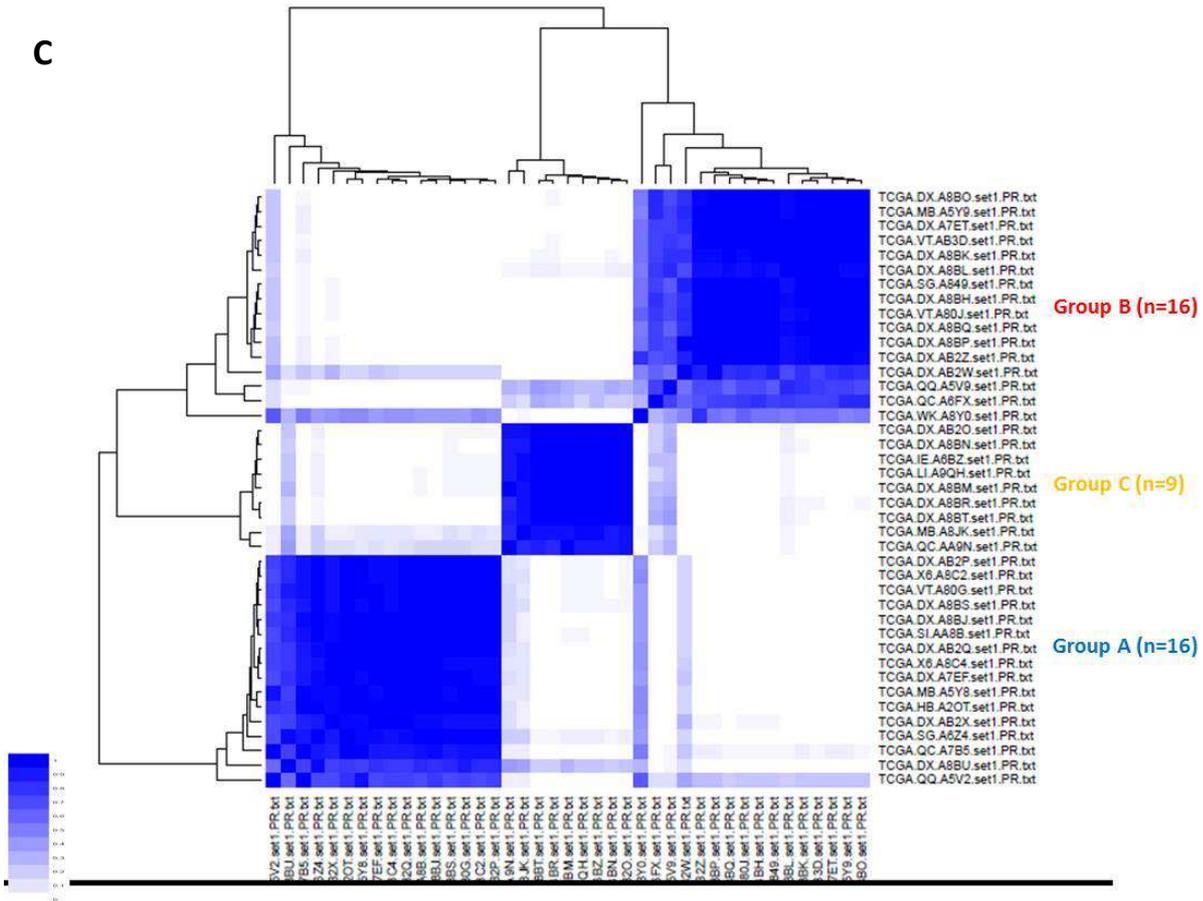
SUPPLEMENTARY FIGURES

Supplementary Figure 1. Analyses of RNA-sequencing data in 25 UPS from Institut Bergonié and 41 UPS from TCGA consortium allow reproducible classification of UPS

(A) Unsupervised consensus clustering of RNA-sequencing data in UPS from Institut Bergonié (n = 25); Blue gradient shows probability that patients cluster in 3 groups by resampling 13 out of 25 samples with 10,000 iterations; (B) Volcano plot of differentially expressed genes between group A (n = 10) and group B (n = 11) identifies 1405 differentially expressed genes; (C) Unsupervised consensus clustering of RNA-sequencing data in UPS from TCGA consortium (n = 41); Blue gradient shows probability that patients cluster within 3 groups by resampling 35 out of 41 samples with 10,000 iterations.

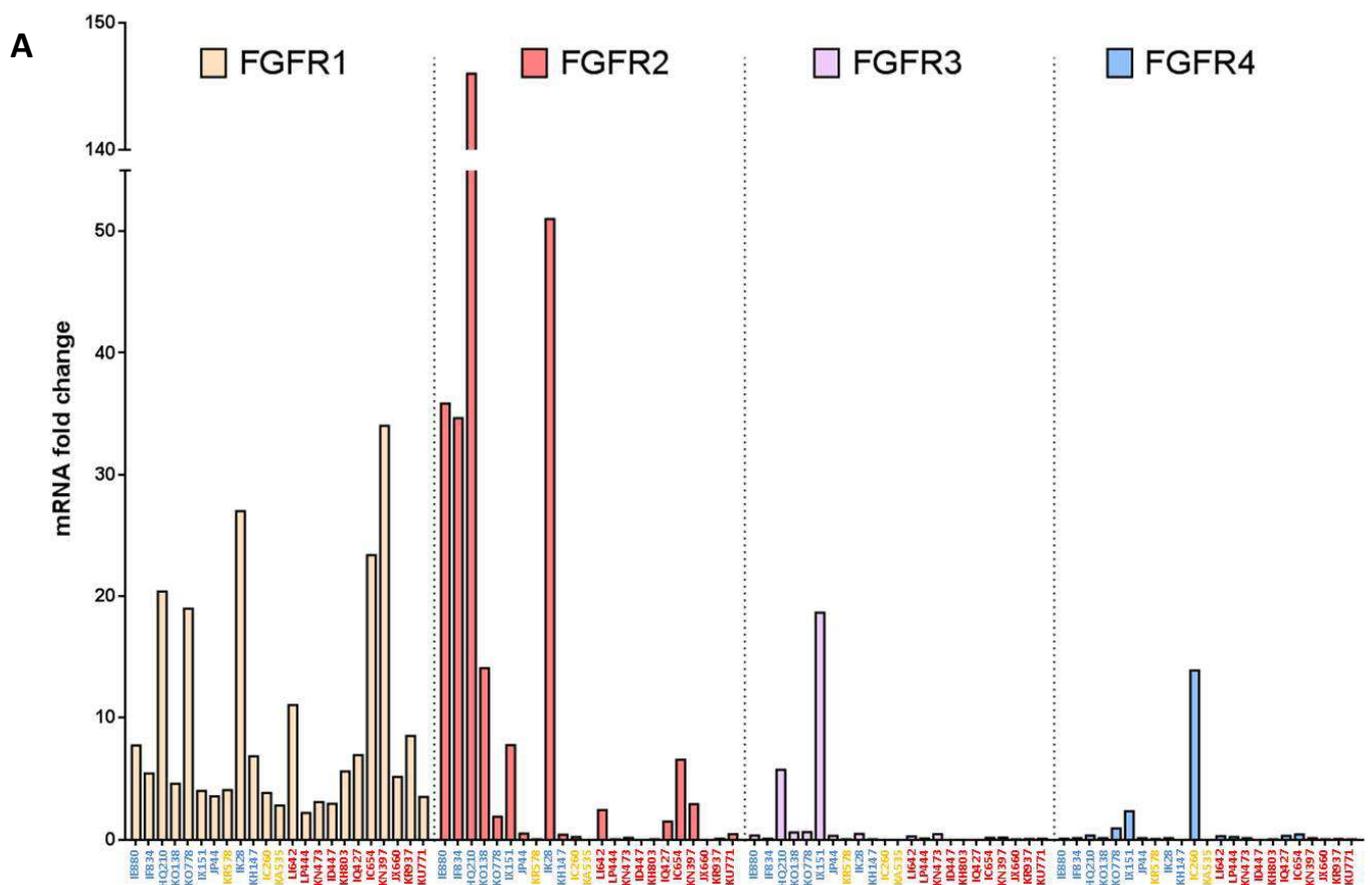


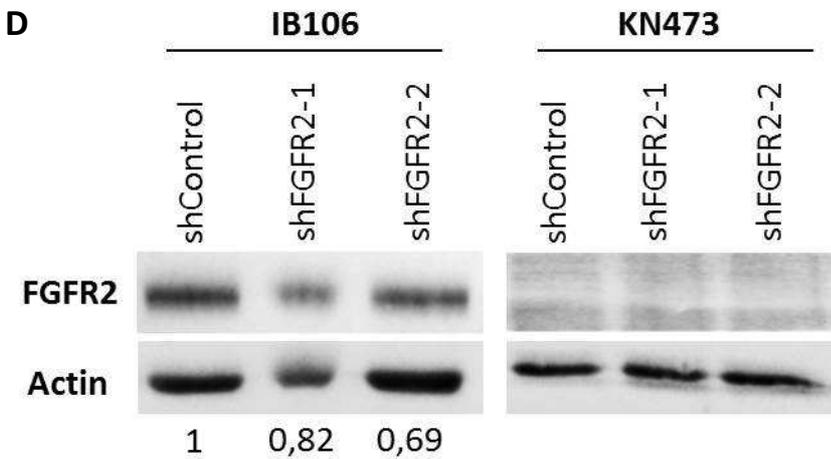
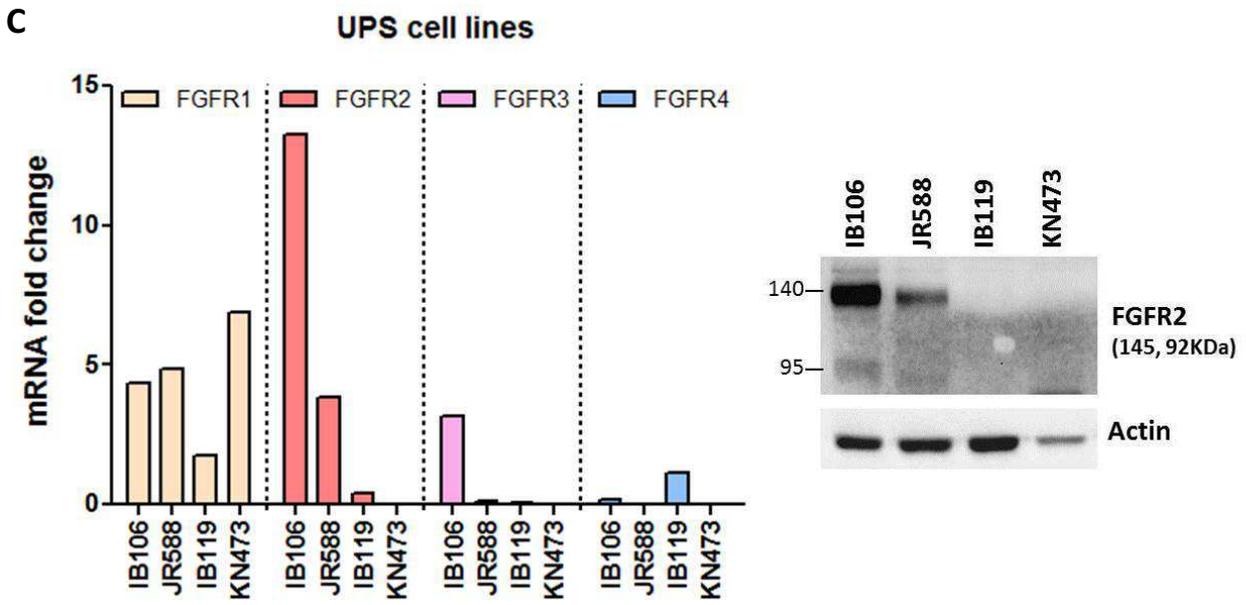
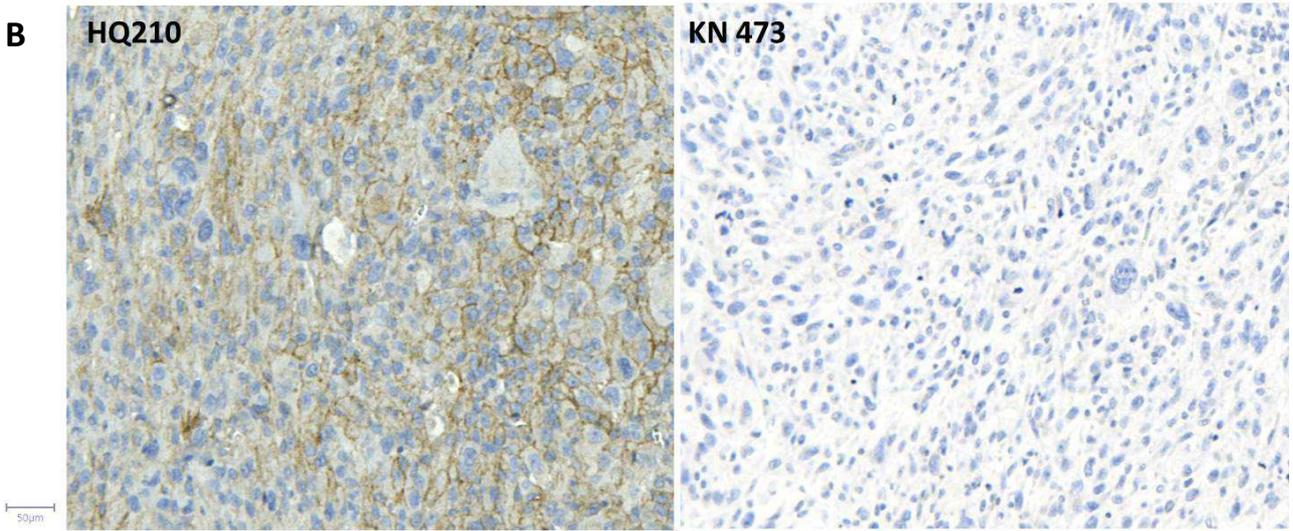
C



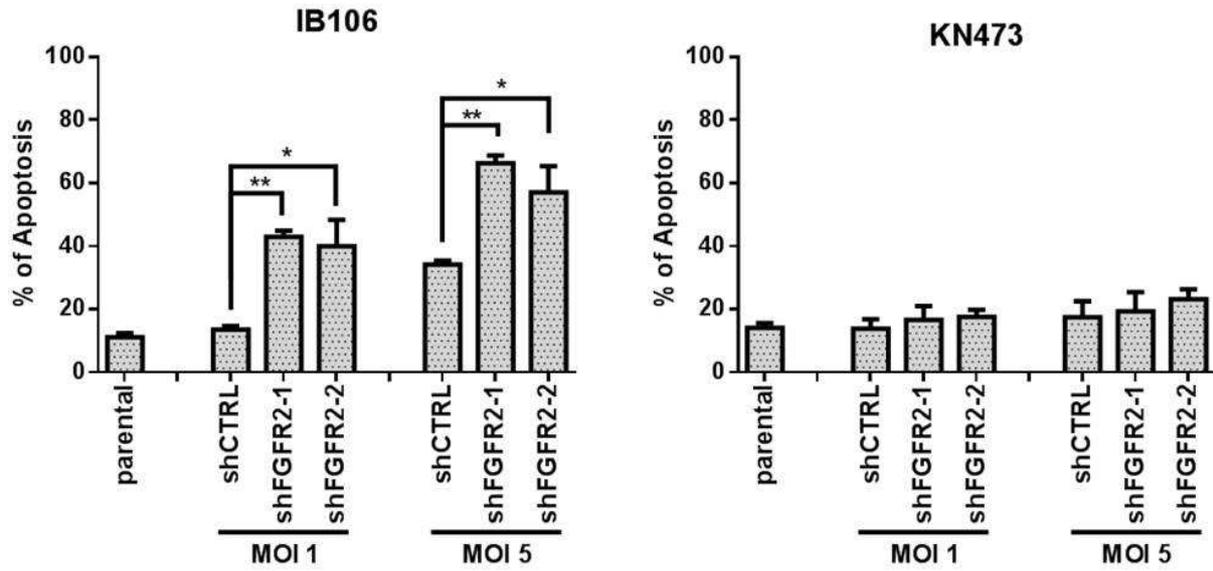
Supplementary Figure 2. Expression and targeting of FGFR2 in UPS

(A) mRNA expression of FGFR1,2,3,4 by RT-qPCR in UPS from Institut Bergonié ordered by group A, B and C according to gene expression clustering (n = 23); FGFR1 is randomly expressed in the 3 groups, FGFR2 is particularly overexpressed in group A whereas FGFR3 and 4 are weakly expressed in the whole cohort; (B) Expression of FGFR2 by Immunohistochemistry (IHC) staining shows concordance with mRNA expression by RT-qPCR. Left: sample belonging to group A positive for FGFR2 on IHC; right: sample belonging to group B negative for FGFR2 on IHC (magnification $\times 100$); (C) Left: mRNA expression of FGFR1,2,3,4 by RT-qPCR in 4 UPS cell lines; IB106 is derived from patient HQ210, IB119 from patient IC260, KN473 from patient KN473 and JR588 from a patient not included in the NGS analysis. Right: Protein expression of FGFR2 by Western blot in the same 4 UPS cell lines shows concordance with mRNA expression by RT-qPCR; (D) Protein expression of FGFR2 by Western blot in IB106 and KN473 cell lines, infected with either a shcontrol or 2 different shFGFR2; (E) Apoptosis quantification in IB106 and KN473 cell lines, infected with either a shcontrol or 2 different shFGFR2 at two multiplicity of infection (MOI: 1 and 5). Cells were analyzed by flow cytometry after Annexin V/Propidium Iodide incorporation 72h after infection (n = 3) (*p < 0.05 and **p < 0.01, one-way ANOVA); (F) Antiproliferative activity of pan-FGFR inhibitors AZD4547 (left) and BGJ398 (right) in 4 UPS cell lines; Growth curves indicate growth inhibition and IC50 of the 4 UPS cell lines after AZD4547 and BGJ398 treatment for 72 hours (n = 3) (**p < 0.01; ***p < 0.001, one-way ANOVA).

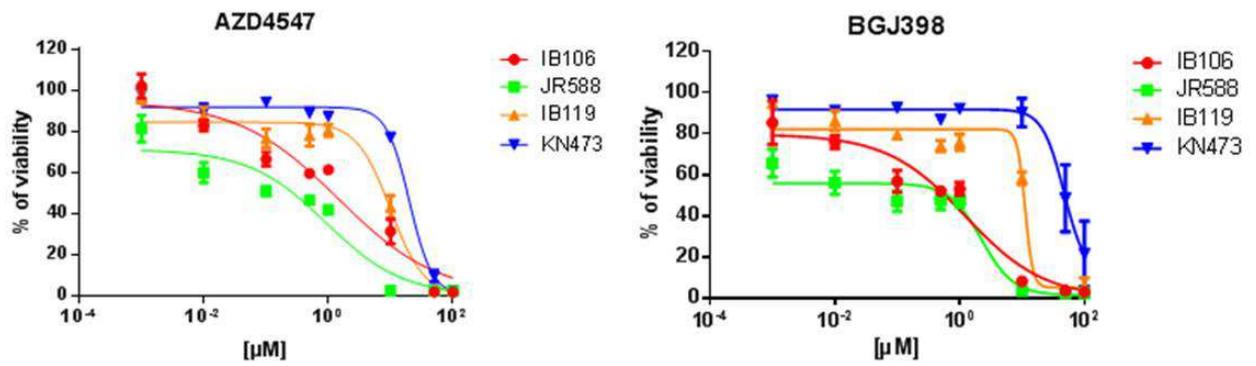




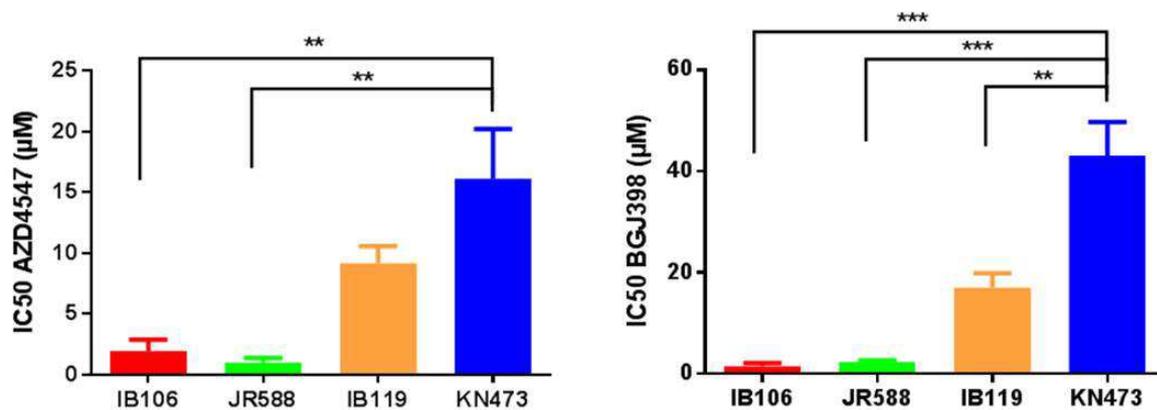
E



F

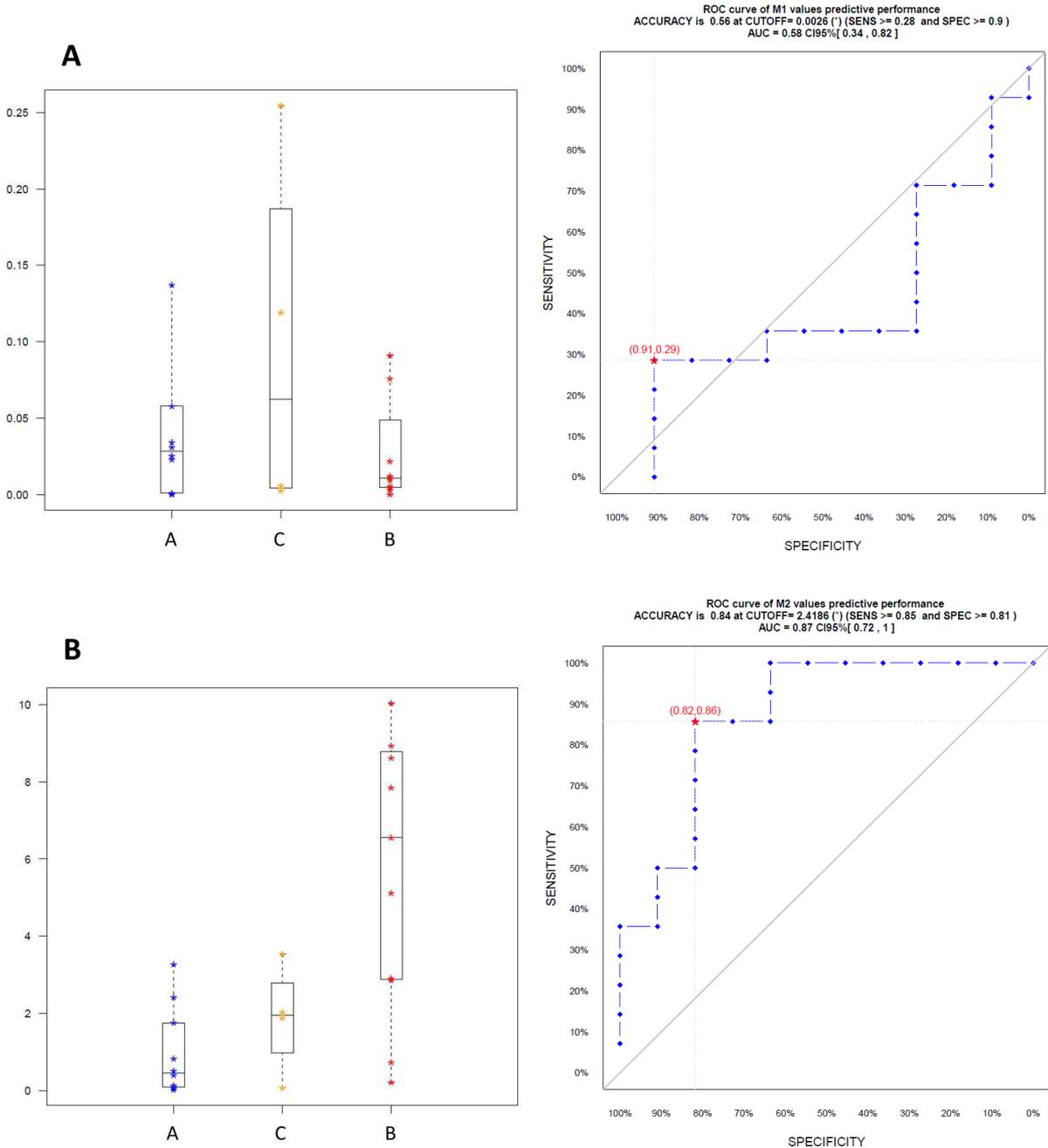


	IB106	JR588	IB119	KN473
IC50 AZD4547 (μM)	1,44	0,91	9,94	20,02**
IC50 BGJ398 (μM)	0,44	2,25	11,23	47,45***

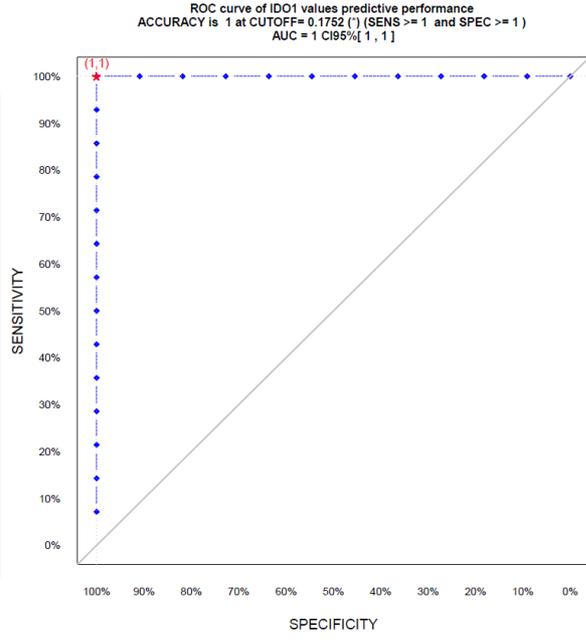
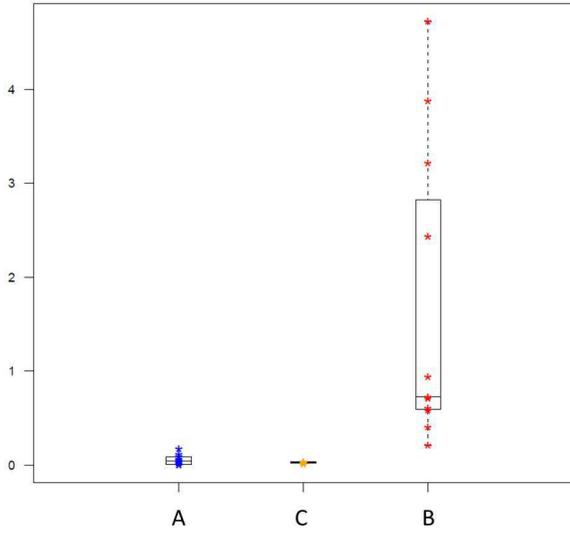


Supplementary Figure 3. Tumor immune infiltrates patterns reflect gene expression clustering of UPS (n = 25)

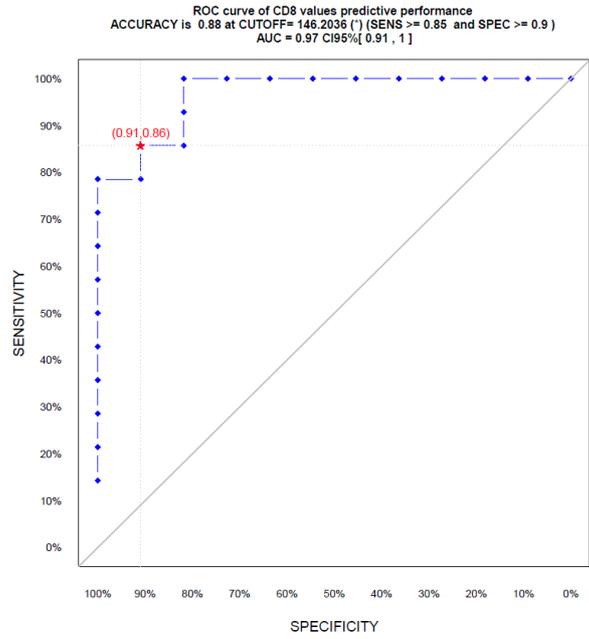
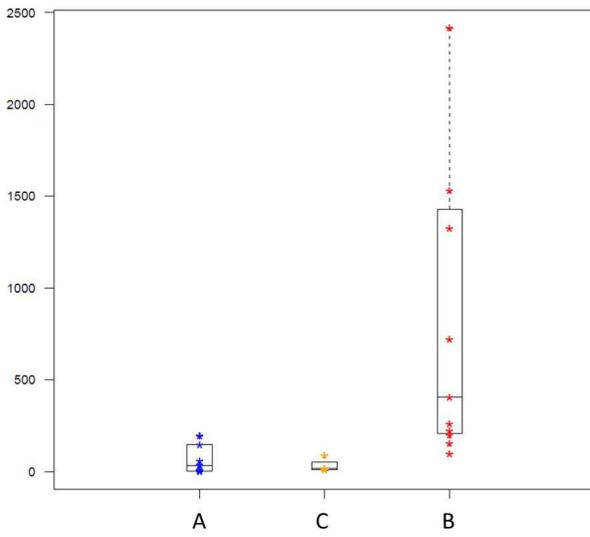
M1 (A) and M2 (B) macrophages density and IDO1 (C) expression by IHC staining in UPS from Institut Bergonié classified by group A, B and C according to gene expression clustering (n = 25); ROC curves of predictive performance of IHC staining to predict RNA sequencing classification (accuracy, sensitivity, specificity and AUC); CD8 (D) and PD-1 (E) density by IHC staining in UPS from Institut Bergonié classified by group A, B and C according to gene expression clustering (n = 25); ROC curves of predictive performance of IHC staining to predict RNA sequencing classification (accuracy, sensitivity, specificity and AUC).



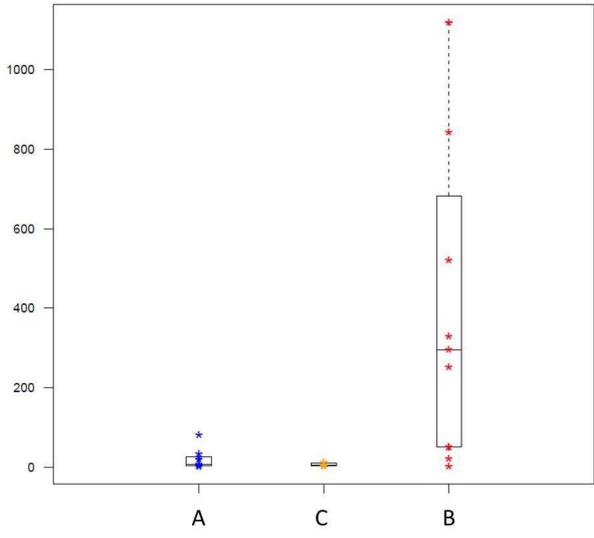
C



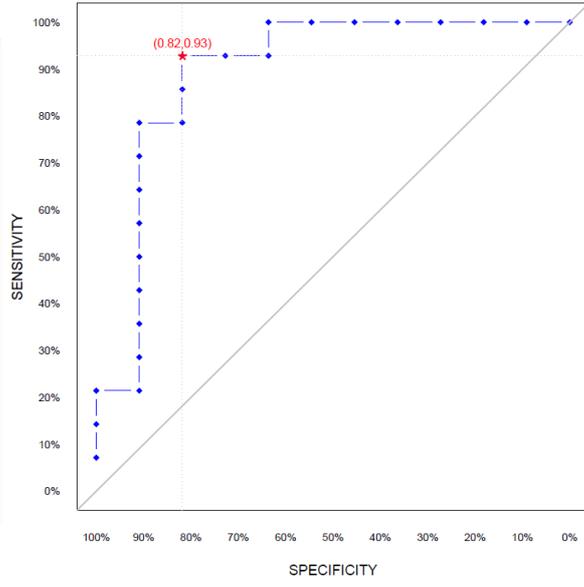
D



E

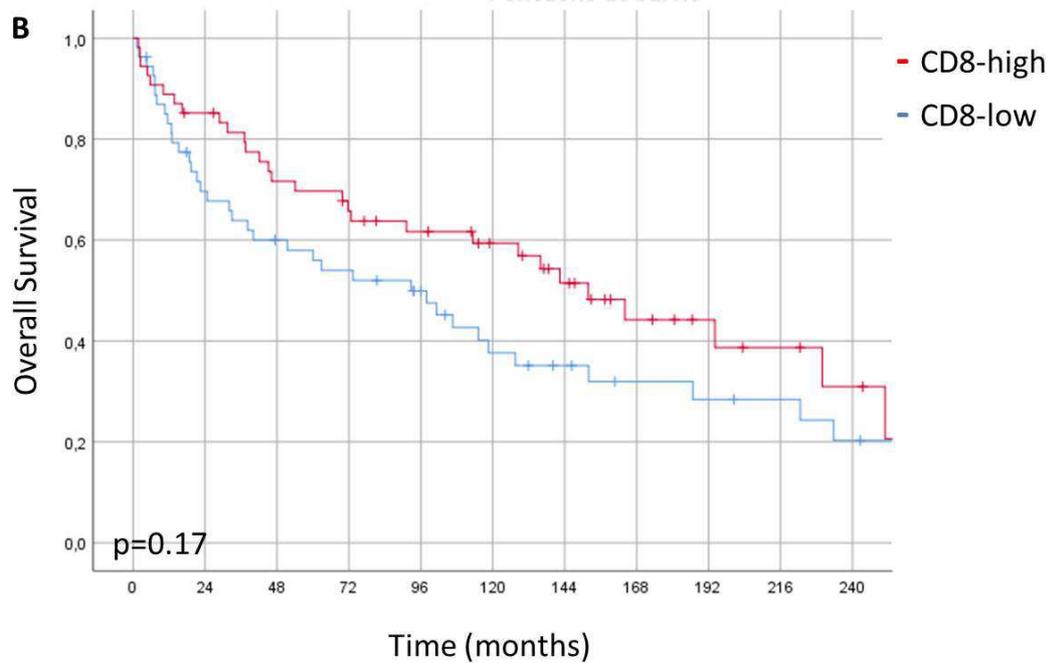
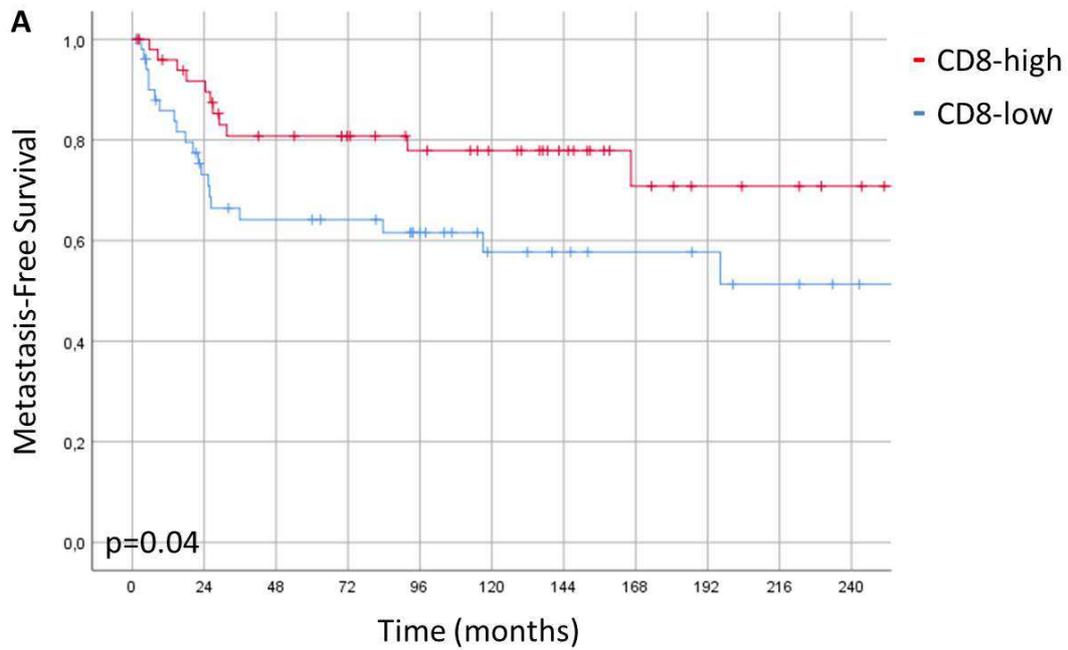


ROC curve of PD_1 values predictive performance
ACCURACY is 0.88 at CUTOFF= 33.632 (*) (SENS >= 0.92 and SPEC >= 0.81)
AUC = 0.9 CI95% [0.75 , 1]



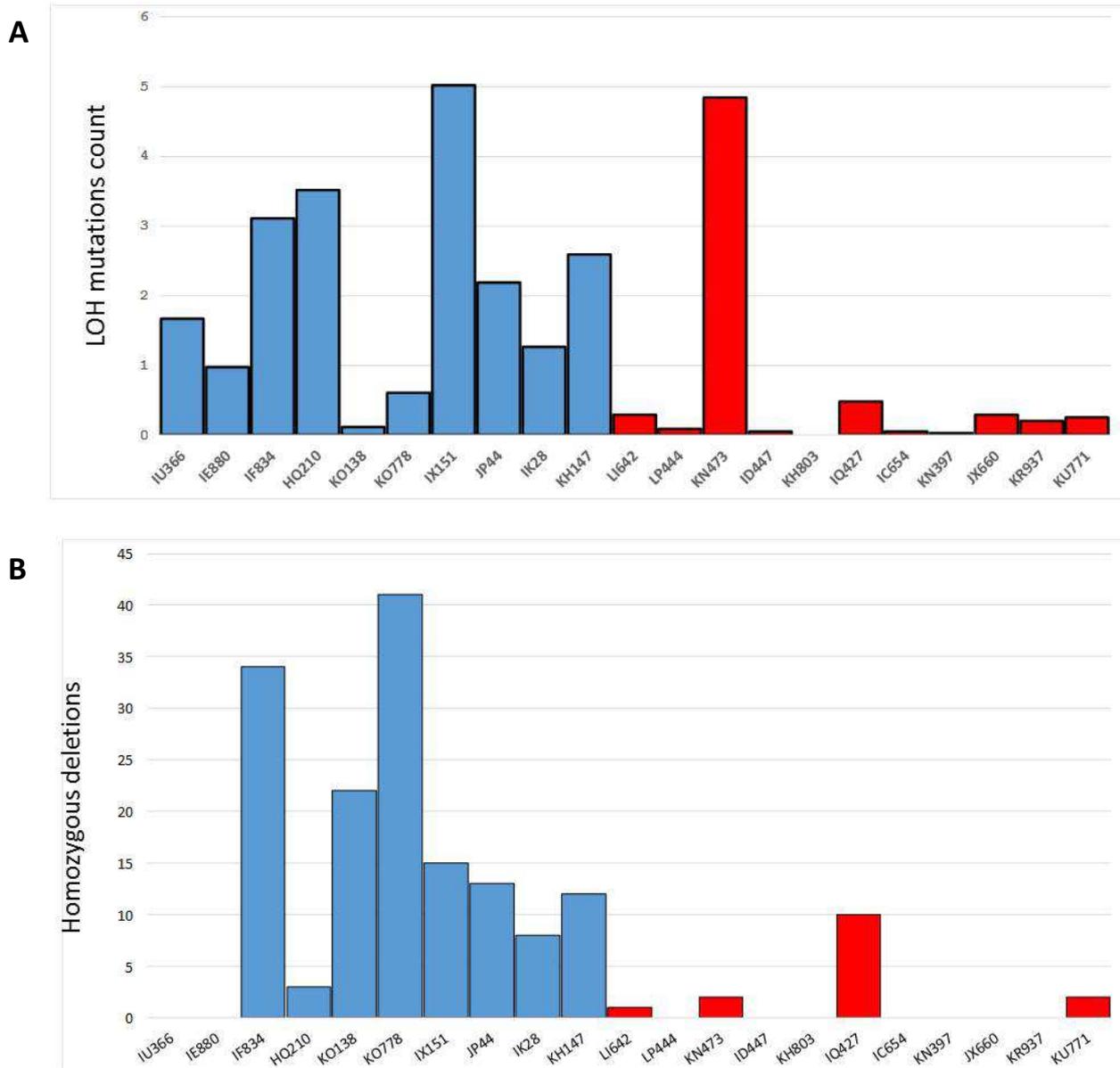
Supplementary Figure 4. Survival analysis according to tumor immune infiltration patterns in an independent group of UPS patients (n = 110)

(A) Metastasis free survival (MFS) ($p = 0.04$) and (B) overall survival (OS) ($p = 0.17$) of UPS patients according to CD8+ cells density in their respective tumor sample (n = 110).

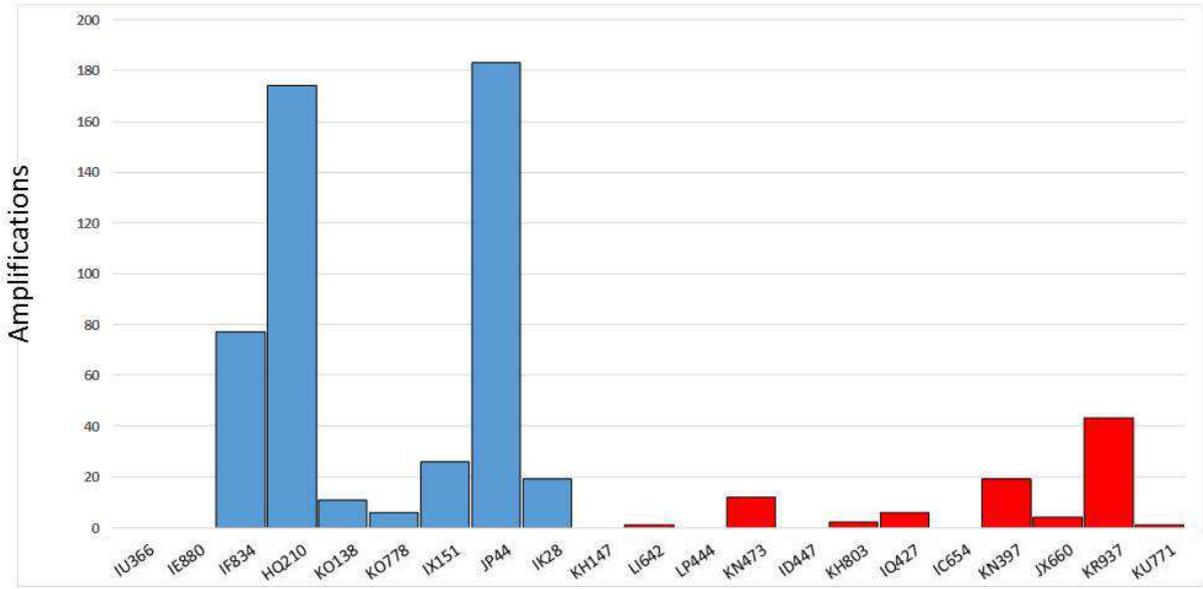


Supplementary Figure 5. Analysis of mutational burden and variation in gene copy number alterations in immune-low and immune-high UPS (n = 21)

(A) Loss of heterozygosity mutations in immune-low (n= 10, blue bars) and immune-high UPS (n = 11, red bars); (B) Homozygous gene deletions in immune-low (n = 10, blue bars) and immune-high UPS (n = 11, red bars) (OR= 10.8); (C) Genes amplification in immune-low (n = 10, blue bars) and immune-high UPS (n = 11, red bars) (OR= 6.9).

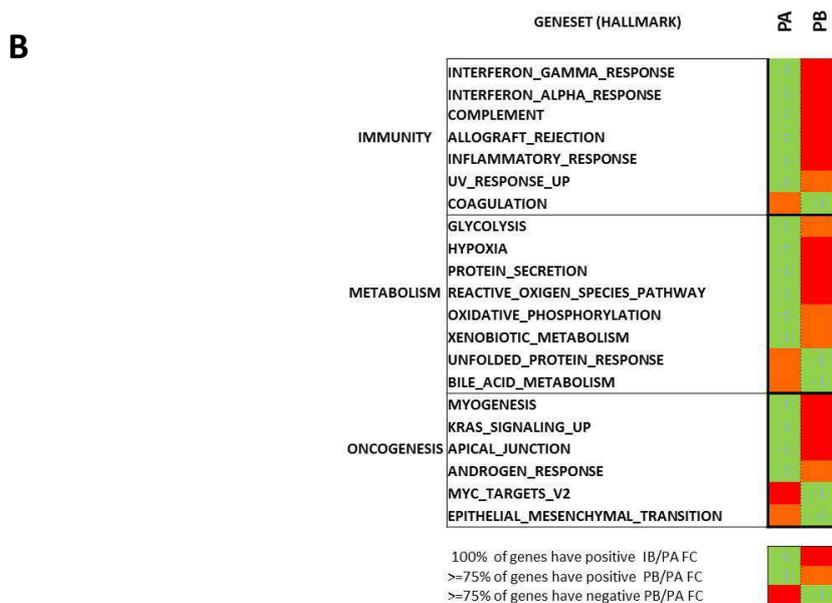
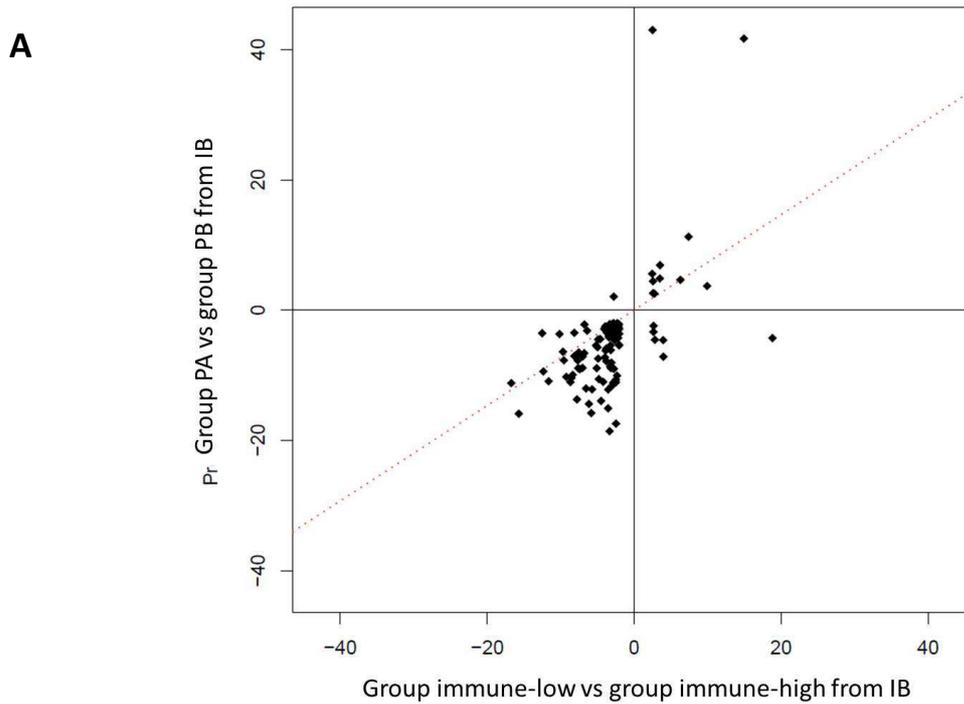


C



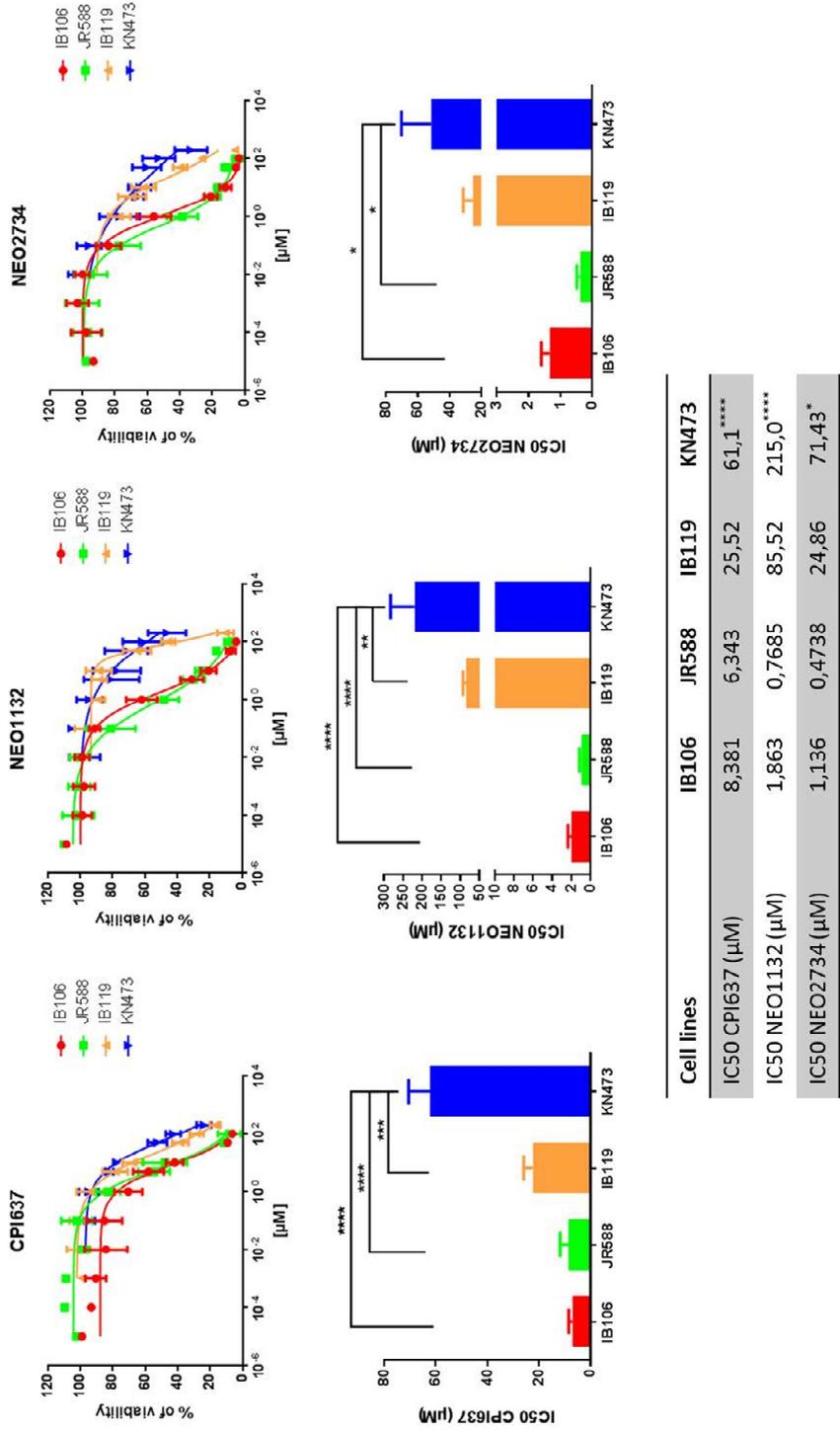
Supplementary Figure 6. Proteomics profiling and represented pathways in UPS (n = 19)

(A) Analysis of Agreement between Differential EXpression (AGDEX) of proteins in group PA vs PB and genes in group immune-low vs immune-high in UPS from Institut Bergonié; 114 couples of protein and related gene that were both differentially expressed were identified. Their analysis of agreement showed good correlation (Spearman = 0.56, Pearson = 0.57, Regression slope = 0.73); (B) Pathways differentially enriched in the PA and the PB groups from unsupervised hierarchical clustering of protein expression (n = 19). Immune response and metabolism pathways are represented in the PB/immune-high group (n = 9), whereas epithelial mesenchymal transition (EMT) and Myc targets pathways appear in the PA/immune-low group (n = 10).



Supplementary Figure 7. BET inhibitors in vitro activity in UPS

Antiproliferative activity of BET-CBP/P300 inhibitors (left: CPI637, middle: NEO1132 and right: NEO2734) in 4 UPS cell lines; Growth curves indicate growth inhibition and IC50 of the 4 UPS cell lines representative of each group after treatment for 72 hours (n = 3) (**p < 0.01, ***p < 0.001, ****p < 0.0001, one-way ANOVA).



SUPPLEMENTARY REFERENCES

1. Langmead B, and Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357-9.
2. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, and Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36.
3. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and Genome Project Data Processing S. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-9.
4. Anders S, Pyl PT, and Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166-9.
5. Wang K, Li M, and Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164.
6. Team RC. R: A Language and Environment for Statistical Computing.
7. Poplin R, Chang PC, Alexander D, Schwartz S, Colthurst T, Ku A, Newburger D, Dijamco J, Nguyen N, Afshar PT, et al. A universal SNP and small-indel variant caller using deep neural networks. *Nat Biotechnol*. 2018;36(10):983-7.
8. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285-91.
9. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74.
10. Kuhn RM, Haussler D, and Kent WJ. The UCSC genome browser and associated tools. *Brief Bioinform*. 2013;14(2):144-61.
11. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., and Kinzler KW. Cancer genome landscapes. *Science*. 2013;339(6127):1546-58.
12. Sondka Z, Bamford S, Cole CG, Ward SA, Dunham I, and Forbes SA. The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nat Rev Cancer*. 2018;18(11):696-705.
13. Law CW, Chen Y, Shi W, and Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15(2):R29.
14. Wilkerson MD, and Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010;26(12):1572-3.
15. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, and Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
16. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, 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(43):15545-50.
17. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, and Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453-7.
18. Ge H, Liu K, Juan T, Fang F, Newman M, and Hoek W. FusionMap: detecting fusion genes from next-generation sequencing data at base-pair resolution. *Bioinformatics*. 2011;27(14):1922-8.
19. McPherson A, Hormozdiari F, Zayed A, Giuliany R, Ha G, Sun MG, Griffith M, Heravi Moussavi A, Senz J, Melnyk N, et al. deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. *PLoS Comput Biol*. 2011;7(5):e1001138.
20. Haas BJ, Dobin A, Li B, Stransky N, Pochet N, and Regev A. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol*. 2019;20(1):213.
21. Johnson RA, Wright KD, Poppleton H, Mohankumar KM, Finkelstein D, Pounds SB, Rand V, Leary SE, White E, Eden C, et al. Cross-species genomics matches driver mutations and cell compartments to model ependymoma. *Nature*. 2010;466(7306):632-6.
22. Kim YJ, Sweet SMM, Egertson JD, Sedgewick AJ, Woo S, Liao WL, Merrihew GE, Searle BC, Vaske C, Heaton R, et al. Data-Independent Acquisition Mass Spectrometry To Quantify Protein Levels in FFPE Tumor Biopsies for Molecular Diagnostics. *J Proteome Res*. 2019;18(1):426-35.
23. Kim YJ, Chambers AG, Cecchi F, and Hembrough T. Targeted data-independent acquisition for mass spectrometric detection of RAS mutations in formalin-fixed, paraffin-embedded tumor biopsies. *J Proteomics*. 2018;189(91-6).
24. Bruderer R, Bernhardt OM, Gandhi T, Xuan Y, Sondermann J, Schmidt M, Gomez-Varela D, and Reiter L. Optimization of Experimental Parameters in Data-Independent Mass Spectrometry

- Significantly Increases Depth and Reproducibility of Results. *Mol Cell Proteomics*. 2017;16(12):2296-309.
25. Thevenaz P, Blu T, and Unser M. Interpolation revisited. *IEEE Trans Med Imaging*. 2000;19(7):739-58.

Discussion des résultats et Perspectives

Ce travail est le premier à utiliser une plateforme « multi-omique » pour tenter d'identifier des sous-groupes d'UPS distincts avec un potentiel de ciblage thérapeutique.

Génomique

Sur le plan génomique, nos données corroborent les deux principaux travaux d'omics publiés dernièrement sur les UPS. Dans notre série, la charge mutationnelle était faible, les seules tumeurs avec une charge > 5 mutations par Mb appartenant au groupe immune-high. Les pertes les plus fréquentes se retrouvaient au niveau de segments incluant *PTEN* (38%), *RBI* (33%), *FANCA* (29%), *FAS* (29%), *CDKN2A* (29%), *TP53* (24%), *AXIN1* (24%), *NF2* (24%), *BRCA2* (24%), mais aussi de façon intéressante *CD274* (24%). Nous avons retrouvé des gains de plusieurs gènes d'intérêt dans l'oncogenèse, notamment *CCND1* (33%), *CDK6* (38%), impliqués dans le cycle cellulaire, ainsi que *MAP2K2* (43%), impliqué dans la voie des MAP kinases ou *AKT* (33%), effecteur de la voie PTEN/PI3K, mais aussi *EZH2* (38%) impliqué dans la régulation épigénétique. Nous avons ensuite étudié les altérations les plus différenciellement représentées entre les groupes immune-low et immune-high en nous intéressant spécifiquement aux gènes suppresseurs de tumeurs, et retrouvé les mêmes voies différenciellement altérées dans les UPS immune-low. Dans l'étude du TCGA portant sur 206 sarcomes, la charge mutationnelle globale était faible, en moyenne 1.06 par Mb. L'équipe a rapporté dans leurs UPS/myxofibrosarcomes des amplifications de *VGLL3* dans 11% des cas, *CCNE1* dans 10% des cas et *YAP1* dans 3% des cas mais aussi de *MYOCD*, ainsi que des pertes de *RBI*, *TP53*, *ATRX*, *CDKN2A*, *NF1* et *PTEN* (Figure 9)(42). L'analyse des 44 UPS et des 17 myxofibrosarcomes inclus dans l'étude du TCGA a révélé une grande similitude entre les deux histologies à travers toutes les plateformes étudiées : aussi bien sur le plan mutationnel, avec des mutations de *TP53*, *RBI* et *ATRX*, qu'au niveau des variations du nombre de copies. Il est cependant clairement expliqué dans le papier que ce manque de distinction entre les deux histologies peut être expliqué par la sur-représentation de composantes de haut grade de type UPS et la sous-représentation de la forme classique à composante myxoïde dans les échantillons de myxofibrosarcomes sélectionnés. En effet, la relecture des échantillons en mentionnait 65% comme étant « non-classiques » et 29% comme « épithélioïdes de haut grade ».

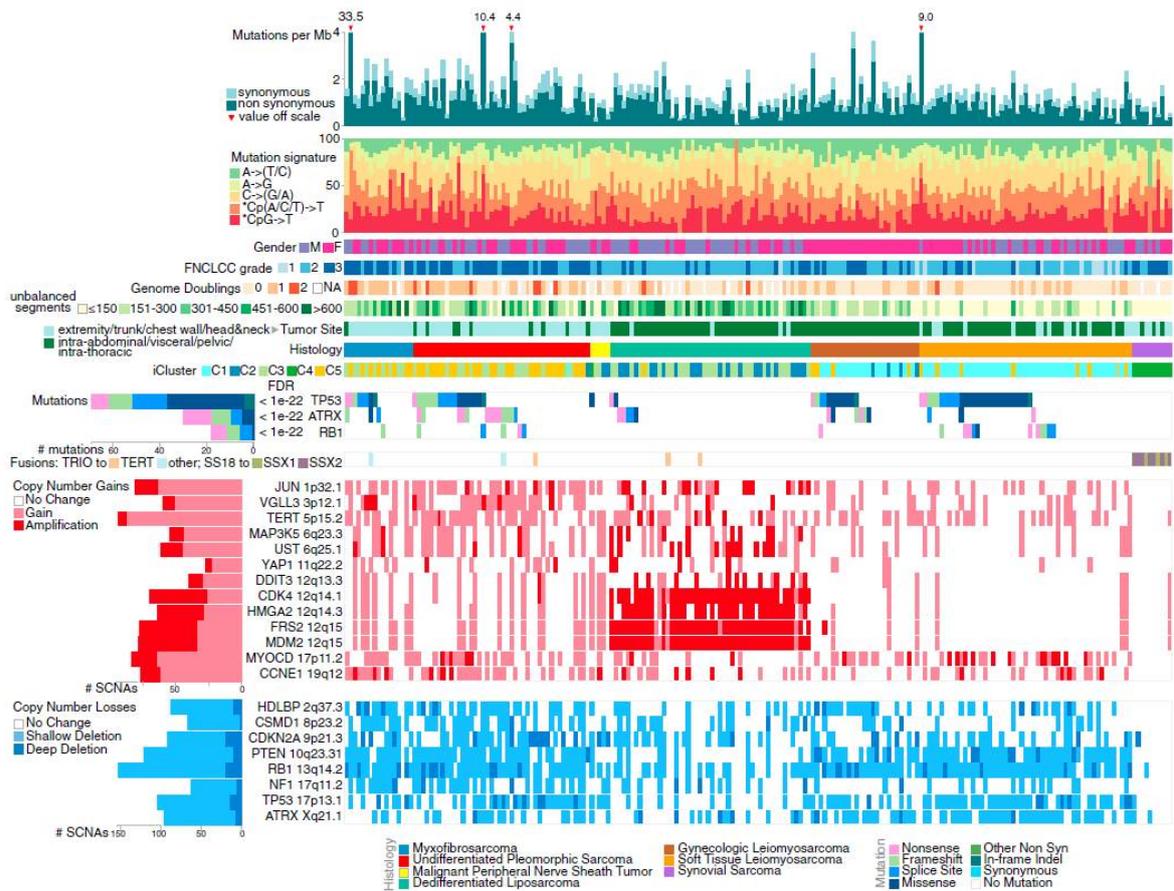


Figure 9 Caractéristiques moléculaires des 206 sarcomes de l'étude du TCGA consortium organisées par histologies (the Cancer Genome Atlas Research Network, *Cell*, 2017)

De haut en bas : Fréquence mutationnelle (/Mb), signatures mutationnelles avec les types de substitution; Genre ; Grade FNLCC; Nombre de doublements du génome; Nombre de déséquilibres de segments génomiques; Site; Type histologique - les UPS sont en rouge; Classification immunologique; Gènes les plus significativement mutés : *TP53*, *ATRX*, *RB*; Principales fusions; Altérations du nombre de copies : gains (rose), amplifications (rouge), délétions hétérozygotes (bleu clair), délétions homozygotes (bleu foncé).

En analysant à nouveau leurs données de transcriptome en fonction de la proportion de composante myxoïde dans leurs tumeurs, ils purent mieux classer les UPS d'un côté et les myxofibrosarcomes de l'autre, sans donner cependant de détails sur les clusters de gènes différentiellement exprimés, en dehors de gènes de la matrice extracellulaire dans les tumeurs fortement myxoïdes (Figure 10).

Un autre papier très important dans la thématique des sarcomes indifférenciés est l'étude de Steele publiée en mars 2019 (43). Les auteurs analysèrent les caractéristiques moléculaires d'une cohorte de 70 sarcomes indifférenciés classés selon leur variant morphologique prédominant : 25 fusiformes et pléomorphes, 18 fusiformes, 14 pléomorphes, 10 épithélioïdes et 3 fusiformes et épithélioïdes.

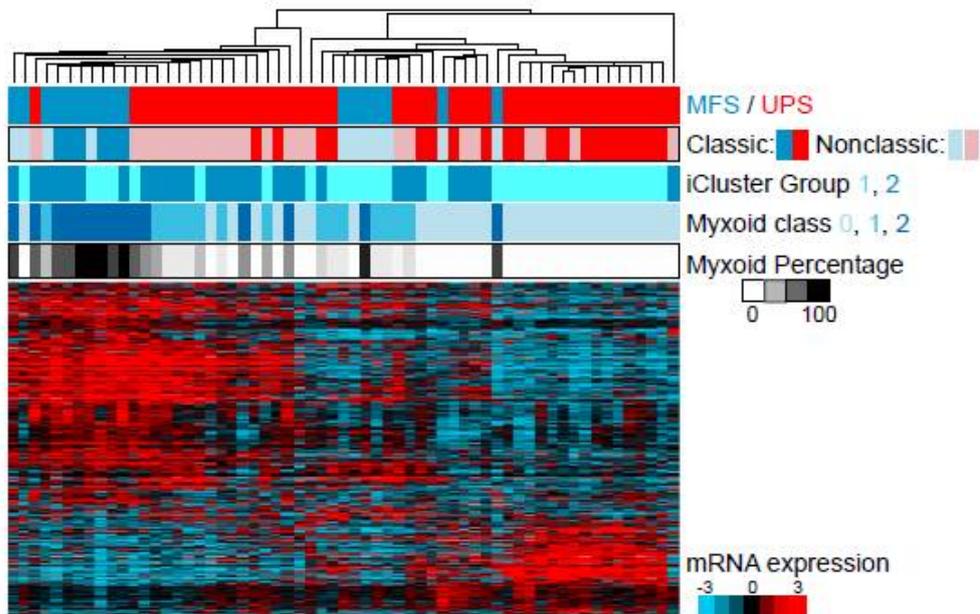


Figure 10 Classification moléculaire des UPS et myxofibrosarcomes de l'étude du TCGA consortium en fonction de la composante myxoïde (the Cancer Genome Atlas Research Network, *Cell*, 2017)

De haut en bas : Type histologique; Sous-type : classique (foncé) /non classique (clair); classification immunologique; Classification myxoïde; Pourcentage de la composante myxoïde; Profil d'expression des gènes différentiellement exprimés entre les groupes.

L'analyse du méthylome montra une relative hypo-méthylation du génome, prédominant dans des régions inter-géniques. L'analyse en composante principale des données de méthylation et d'expression génique ne permit pas de déterminer de sous-groupe reproductible. Ils analysèrent ensuite le génome complet de 52 de leurs échantillons et s'intéressèrent aux réarrangements. Ils identifièrent ainsi trois sous-groupes moléculaires distincts : un groupe mutation-high/ réarrangement-low (mutHi-rearrLo), un groupe mutation-low/ réarrangement-high (MutLo-rearrHi) et un groupe mutation-low/ réarrangement-low (mutLo-rearrLo). Ils purent mettre en évidence des altérations de gènes en rapport avec la réparation des mésappariements « MMR, mismatch repair » de type mutation inactivatrice associée à la perte de l'allèle normal, méthylation de promoteur ou mutation pathologique constitutionnelle pour *MSH2* et *MSH6* dans le groupe des sept patients mutHi-rearrLo. Ils identifièrent également des réarrangements récurrents pour 51 gènes, dont neuf étaient des gènes suppresseurs de tumeurs connus. Les réarrangements de type mutation ou délétion les plus fréquents intéressaient *TP53*, *RB1*, *ATRX*, *CDKN2A*, *CBFA2T3*, *STK11*, *TCF3*, *CYLD*. A noter que jusqu'à 50% des événements retrouvés pour *TP53*, *RB1* et *ATRX* auraient été omis en analyse de l'exome uniquement. Ils ont de plus pu mettre en évidence une altération de la voie

PI3K/Akt dans près de 33% des échantillons, représentée par des délétions de *PTEN*, *TSC1*, *TSC2* et des mutations *hotspot* de *PI3KCA*. Le groupe de tumeurs avec une charge mutationnelle plus élevée que la moyenne (mutHi-rearrLo) était associé à une activation des voies de l'immunité et à un meilleur pronostic. Ces données concordent avec les nôtres. Ils étudièrent ensuite plus spécifiquement les duplications du génome (WGD pour whole genome duplication) et les variations du nombre de copies des gènes, et identifièrent sept signatures distinctes, sans toutefois pouvoir obtenir de corrélation avec les données de transcriptomique, ni mettre en évidence de voie de signalisation préférentielle associée à l'une de ces signatures. La CNS1 (copy number signature) était associée à une perte de l'hétérozygotie (LOH pour loss of heterozygosity) amplifiée, indiquant deux ou plus duplications du génome et associée significativement à la perte de *TP53*; la CNS2 était caractérisée par des duplications de la LOH, indiquant une duplication du génome. La CNS3 était hypo-diploïde, avec de larges délétions et quelques gains/amplifications, tandis que la CNS4 se caractérisait par une LOH neutre, évocatrice d'haploïdisation. La CNS5 conservait l'hétérozygotie et quelques amplifications, la CNS6 associait de larges segments dupliqués hétérozygotes et des nombres de copies très variables. Enfin la CNS7, la plus fréquente (67% des échantillons), représentait une simple duplication du génome. Les auteurs conclurent que l'haploïdization et la duplication du génome sont deux événements fréquents précoces et distincts dans l'évolution d'une partie des sarcomes indifférenciés et proposèrent un schéma évolutionniste de ces sarcomes indifférenciés (Figure 11).

De nombreux gènes de fusion ont été décrits dans les sarcomes des tissus mous (44). Beaucoup de ces fusions peuvent permettre d'une part le diagnostic d'une entité spécifique, mais également servir de facteurs de stratification, et sont l'objet de nombreux travaux de recherche visant à mieux comprendre les mécanismes moléculaires impliqués dans la tumorigénèse, et développer des approches thérapeutiques. Les travaux portant spécifiquement sur la recherche de fusions récurrentes dans les UPS sont peu nombreux et leurs résultats limités. En 2015, deux fusions impliquant *PRDM10* (*MED12/PRDM10* et *CITED2/PRDM10*) ont été retrouvées dans trois UPS (45). Cependant les deux UPS dans lesquels ces fusions ont été observées pour la première fois étaient des UPS de bas grade sélectionnés sur la base de leur caryotype simple. De plus, les trois tumeurs avaient un faible compte mitotique, un faible grade et comportaient des cellules géantes et de la matrice myxoïde, suggérant qu'il ne s'agissait pas de vrais UPS.

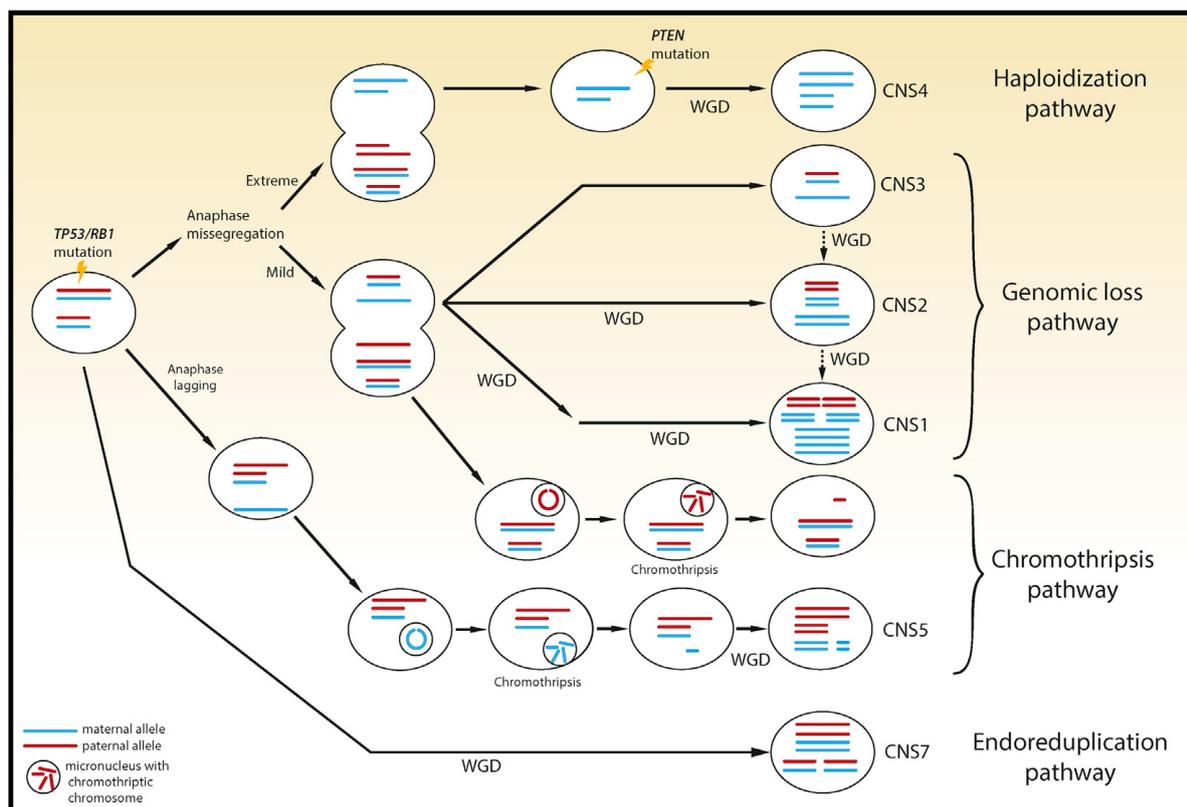


Figure 11 Proposition d'arbre évolutionniste de la tumorigénèse des sarcomes indifférenciés issue du travail de Steele *et al.* Cancer Cell, 2019

Les événements conducteurs de type mutations (*TP53*, *RB1*) sont précoces. La voie de l'haploïdisation (CNS4) se caractérise par une mal-ségrégation extrême lors de l'anaphase, générant de larges zones de perte d'hétérozygotie ; Trois signatures (CNS3, CNS2 en CNS1) sont des variations de ce type de pertes mais différentes par les duplications de génome ultérieures associées. La voie de la chromothripsis est une mal-ségrégation en anaphase ou un décalage d'anaphase amenant à des micro-nucléations de chromosome. La CNS5 reprend ce procédé associé à une duplication du génome. Dans la voie de l'endoreduplication, une cellule subit un doublement du génome avec relativement peu d'autres altérations : CNS7.

En 2016, Delespaul et collègues ont mis en évidence un type de fusion impliquant le gène *TRIO* dans des sarcomes dits à génomique complexe (46). Après séquençage ARN de 112 échantillons, ils appliquèrent une méthodologie informatique dédiée à la détection des transcrits de fusion (trois algorithmes: dDeFuse, TopHat-Fusion et ChimerScan), qu'ils validèrent ensuite en RT-PCR et séquençage par la méthode Sanger. Ils identifièrent ainsi trois tumeurs porteuses d'une fusion *TRIO*_{33/34}-*TERT*_{2/3}. Deux des trois cas présentaient également un réarrangement de *TRIO* (gain/amplification) en CGH. En élargissant leur cohorte avec 24 patients supplémentaires, ils retrouvèrent une autre fusion *TRIO*-*TERT* et deux nouvelles fusions impliquant *TRIO*₃₃ et *LINC01504*, et *TRIO*₃₄ et *ZNF558*. L'analyse supervisée des gènes différenciellement exprimés entre le groupe des tumeurs avec une fusion de *TRIO* (TRIO+) et celui des tumeurs sans fusion de *TRIO* (TRIO-) permet de mettre en évidence une surexpression des sets de gènes de l'immunité, du cycle cellulaire et de la

prolifération dans le groupe de tumeurs TRIO+, tandis que les sets de gènes principalement représentés dans les tumeurs TRIO- appartenaient à la différenciation myogénique ou adipocytaire. Au total, parmi les sept tumeurs identifiées porteuses d'une fusion impliquant *TRIO*, trois étaient des UPS, deux des DDLPS, une un myxofibrosarcomes et une un rhabdomyosarcome pléomorphe. Récemment, Zheng et collègues évaluèrent le transcriptome de 19 UPS à la recherche de transcrits de fusion récurrents et retrouvèrent deux fusions impliquant le gène *RBI* avec *RNASEH2B* et *FGF14-AS1*, associées à une perte d'expression de Rb en IHC, suggérant une perte de fonction du gène *RBI* associée à cette fusion (47). L'étude du TCGA n'a pas mis en évidence de fusion récurrente dans sa série d'UPS en dehors de deux fusions impliquant *TRIO*, dont une *TRIO/TERT*, également rapportée par l'équipe de Delespaul (42, 46). Enfin, l'étude de Steele et collègues n'a pas non plus identifié de fusion oncogénique récurrente malgré une analyse de séquençage de grande profondeur (43). Nous avons retrouvé une fusion impliquant *TRIO/TERT* dans un UPS de notre série, mais pas d'autre fusion récurrente malgré les trois algorithmes utilisés. L'ensemble de ces résultats confirme le fait que la sarcomagenèse primitive des UPS n'est pas basée sur des phénomènes génomiques primitifs *driver* de type translocation récurrente et que ce type d'altération est plutôt à considérer comme un événement de type « passager » dans ces tumeurs à la génomique très instable et hétérogène.

Classification transcriptomique, protéomique et impact fonctionnel

Immune-high UPS

Un des résultats importants de ce travail est l'identification d'un sous-groupe d'UPS appelé « immune-high », caractérisé par un fort infiltrat inflammatoire, et ceci de façon reproductible à travers deux cohortes indépendantes. Ce groupe de patients est caractérisé par l'expression de multiples clusters de gènes impliqués dans diverses réactions immunitaires, notamment les signatures monocytes, lymphocytes CD8, lymphocytes CD4 mémoires, macrophages M2, NK activés, ainsi que la réponse interféron. Ce groupe se caractérise également par l'expression de marqueurs immunitaires en IHC tels que CD8 et IDO1 au niveau tumoral, et semble avoir un profil de radiomique spécifique, ce qui suggère la possibilité d'une sélection de ces patients en routine pour des approches d'immunothérapie.

L'implication de l'immunité dans les UPS a été soulevée dès le début des années 2000 lors des premières analyses de transcriptome des UPS. L'équipe de Baird avait déjà identifié en 2005 un sous-groupe d'UPS caractérisé par une forte expression des gènes de l'inflammation (18). En 2011, l'équipe de Gibault rapportait également un sous-type de sarcomes à génomique complexe constitué principalement d'UPS de haut grade sur-exprimant fortement un cluster de gènes de l'inflammation (25). Les études récentes du TCGA et de Steele et al. ont également confirmé cette caractéristique (42, 43).

Les avancées dans la compréhension du rôle de l'immunité dans la cancérogenèse, et l'arrivée de l'immunothérapie de nouvelle génération en cancérologie, ont amené de grands espoirs et de notables améliorations dans la prise en charge de certains cancers sur les dix dernières années. Malheureusement, l'utilisation de cette approche dans les sarcomes a donné jusqu'ici des résultats préliminaires modestes. Nous avons rapporté l'étude PEMBROSARC évaluant l'association du cyclophosphamide per os à posologie métronomique avec l'anti-PD-1 pembrolizumab dans les sarcomes des tissus mous (48). Cette étude est négative sur son objectif primaire, mais lors de l'analyse des échantillons tumoraux des patients inclus nous avons pu montrer que le seul patient ayant eu une réponse partielle était le seul pour lequel on retrouvait une expression de PD-L1 au niveau des cellules immunitaires intra-tumorales > 10%, et celle-ci était associée à un fort infiltrat CD8 au sein de la tumeur. Nous avons également pu mettre en évidence une infiltration CD8 faible et variable, associée à la présence de macrophages pro-tumoraux de types M2, ainsi que l'expression d'IDO1 dans une majorité des échantillons, pouvant expliquer la résistance primaire à cette approche thérapeutique.

IDO1 est une enzyme réalisant la dégradation du tryptophane en kynurénine, principalement exprimée par les cellules immunitaires au cours de l'inflammation, et impliquée dans l'immuno-suppression (Figure 12) (49, 50).

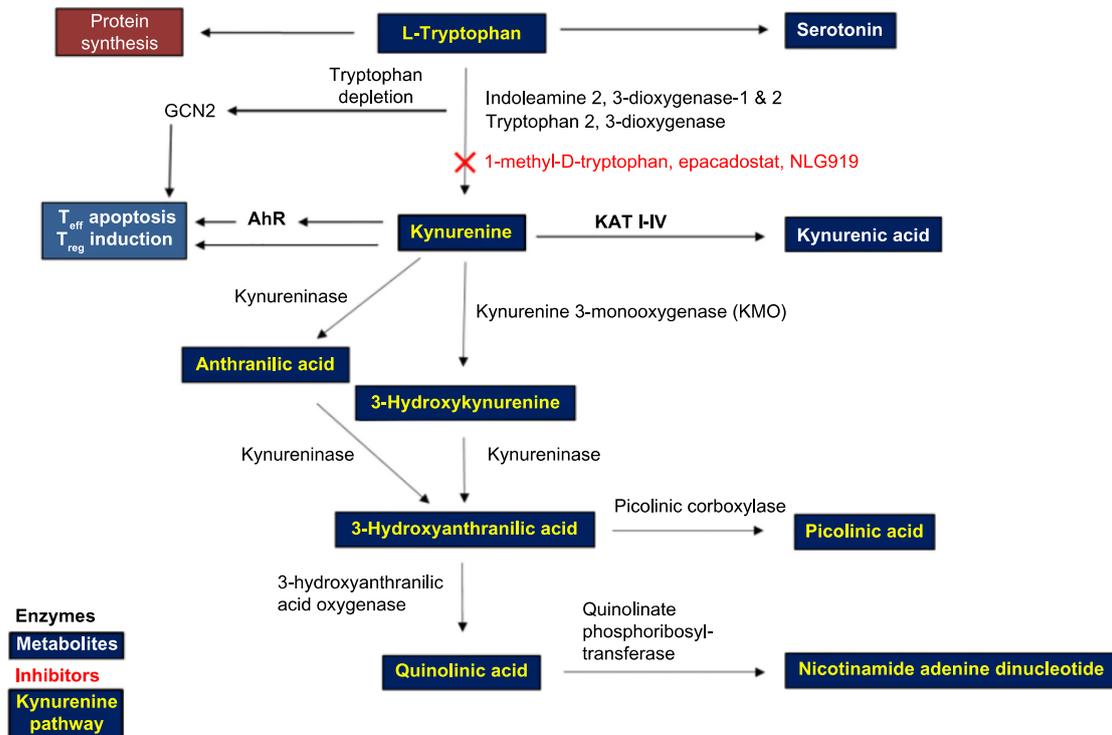


Figure 12 Représentation schématique des enzymes clés et des métabolites de la voie de la Kynurénine par Routy *et al. Int J Tryptophan Res, 2016*

En plus de son rôle important dans la synthèse protéique et la production de sérotonine, l'acide aminé essentiel Tryptophane joue un rôle majeur dans l'immuno-suppression par la voie des kynurénines. Les enzymes IDO1, IDO2 et TDO sont les enzymes limitantes de cette voie, induisant la production de kynurénines. La kynurénine se lie au récepteur AhR présent dans différents types de cellules immunitaires, notamment macrophages, cellules dendritiques et lymphocytes, et induit la transcription de gènes impliqués dans les voies immuno-suppressives telles que la voie du TGF- β ou de l'IL10 (AhR : aryl hydrocarbon receptor; GCN2 : general control nonderepressible 2; IDO : indoleamine 2,3-dioxygénase; KAT : kynurenine aminotransférase; T_{eff} : effector T cell; T_{reg} : regulatory T cell; TDO : tryptophan 2,3-dioxygénase).

De façon intéressante, dans l'étude SARC028, quatre sur dix des patients porteurs d'un UPS traités par pembrolizumab seul ont eu une réponse objective. Malheureusement, les études ancillaires rapportées dans la publication n'ont porté que sur l'analyse de l'expression de PD-L1 au niveau des cellules tumorales, faible et non corrélée (51).

L'ensemble de ces résultats très préliminaires a amené beaucoup de questions sur la manière de sélectionner les patients porteurs d'un sarcome pouvant bénéficier de l'immunothérapie, ainsi que sur les autres voies immuno-modulatrices que PD-1/PD-L1 qui pourraient être préférentiellement ciblées dans ces tumeurs (52).

Dans le papier du TCGA consortium publié en novembre 2017, on pouvait trouver 44 UPS parmi les 206 sarcomes étudiés. Parmi les gènes différentiellement exprimés retenus dans l'analyse transcriptomique non supervisée globale, 203 gènes étaient reliés à l'inflammation et ont servi de base à un score immunologique reprenant les signatures géniques des différents types de cellules immunitaires. Les UPS/myxofibrosarcomes et les DDLPS avaient le plus haut score macrophagique, et la présence des signatures des cellules dendritiques matures et immatures dans ces tumeurs était associée à un bon pronostic, soulignant le rôle important de la présentation d'antigènes dans la réponse immunitaire de ces tumeurs. Une plus forte expression de *B7H3*, *TGFBI* et *TIM3* était retrouvée dans les UPS, myxofibrosarcomes et DDLPS (42).

En parallèle au présent travail, notre équipe a eu l'occasion de nouer une collaboration avec l'équipe du Pr Fridman portant sur l'élaboration d'une classification immunologique des sarcomes. Dans ce travail, les données de transcriptome de 608 échantillons de sarcomes de type léiomyosarcomes, UPS et DDLPS provenant de quatre cohortes de patients annotées et disponibles publiquement ont été analysées (53). La composition du micro-environnement tumoral a été analysée selon la méthode « MCP-counter » décrite précédemment (54). Cette méthode produit un score d'abondance pour huit populations de cellules de l'immunité à partir de signatures d'expression génique associées (Lymphocytes T, CD8, cytotoxiques, NK, B, lignée monocyttaire, dendritique, et neutrophile, ainsi que deux populations stromales : endothéliale et fibroblastique). En développant ensuite une analyse basée sur les corrélations des centroïdes de chaque set de gènes relatifs à une population de cellules dédiée, une classification immunologique fut élaborée, basée sur cinq classes : A, B, C, D et E. La classe A était enrichie en léiomyosarcomes, et la classe E enrichie en UPS tandis que les DDLPS étaient surreprésentés dans la classe C. De façon intéressante, la classe E était extrêmement enrichie en signatures immunologiques représentatives de toutes les populations étudiées. La signature du lignage lymphocytaire B était la plus prépondérante, avec notamment le gène de *CXCL13*, chimiokine impliquée dans la formation de structures lymphoïdes tertiaires (TLS), définies par un follicule de lymphocytes B CD20+, associé à un agrégat de cellules de type lymphocytes CD3+, contenant au moins une cellule dendritique mature DC-LAMP+. A l'inverse, la classe A était qualifiée de « immune-déserte », car exprimant très faiblement les différentes signatures immunologiques, tandis que la classe C était caractérisée par une forte représentation de cellules endothéliales, associée à une représentation faible à moyenne des cellules immunitaires. D'autres gènes d'intérêt tels que *PDCD1*, *PDCD1LG2*, *CTLA4*, *TIM3*, *LAG3*, *IDO1* étaient également différentiellement représentés dans la classe E. Les patients de

la classe A avaient le moins bon pronostic, et globalement, les patients de la classe A et B avaient un pronostic significativement moins bon que ceux des classes D et E ($p = 0.015$). Parmi les signatures immunitaires, seule la signature « lymphocytes B » était associée à une meilleure survie tandis que les autres - « T », « CD8 » et « cytotoxiques » incluses - ne l'étaient pas. L'analyse de la charge mutationnelle ne retrouva pas de différence entre les classes immunologiques, mais les tumeurs avec la charge la plus haute étaient dans les groupe D et E, notamment deux tumeurs MSI-high, tandis que neuf gènes étaient retrouvés mutés dans plus de 5% des tumeurs : *TP53* (35%), *ATRX* (16%), *TTN* (10%), *RBI* (9%), *MUC16* (8%), *PCLO* (6%), *DNAH5*, *MUC17* et *USH2A* (5% chaque). La fréquence de mutation de *TP53*, *TTN* et *MUC16* était plus élevée dans les classes D et E que les autres. Enfin, en termes de variations de nombre de copies, l'analyse de la cohorte du TCGA consortium permit de mettre en évidence des pertes de 1q et de 15q fréquentes dans les UPS, mais sans différences évidentes entre les classes immunologiques, en dehors d'une tendance pour un profil plus remanié pour les A, proches des léiomyosarcomes, et une plus faible amplitude pour les classes D et E (données non publiées). Ces données concordent avec les nôtres.

De façon intéressante, l'analyse du méthylome de la cohorte issue du TCGA permit de retrouver des profils différents selon les classes immunes : globalement les gènes de l'immunité étaient hyper-méthylés dans la classe A et hypo-méthylés dans la classe E. Parmi les gènes hyper-méthylés dans la classe A on pouvait retrouver *CD8A*, *CD27*, *CD19*, *BANK1*, *CD160*, *PDCD1*, *CD274*, *LAG3*, *PRF1*, *GNLY*, *IL12*, *IL32*, *IL6*, *IL13*, et parmi les groupes de gènes hypo-méthylés dans les classes A et B se trouvaient *IL4*, *CCDC80*, en rapport avec la voie Wnt/bêta-caténine, ou *TGFB3* (données non publiées). L'équipe a ensuite étudié la corrélation entre la classification immunologique obtenue en séquençage ARN et l'expression des principaux marqueurs de l'immunité en IHC (CD3, DC-LAMP, CD20, CD8, CD34 et PD-L1) dans une cohorte indépendante de 73 tumeurs. Une très bonne corrélation a pu être mise en évidence entre ces deux techniques pour plusieurs marqueurs tels que CD3, CD20, CD8 ou CD34 (Figure). Le marqueur CD20 était le plus discriminant. Une méthode de sélection des tumeurs de la classe E à partir de la détection de TLS fut alors développée. Selon le degré de maturation du TLS s'y associait la présence de lymphocytes T « épuisés » CD3+ PD1+ LAG3+.

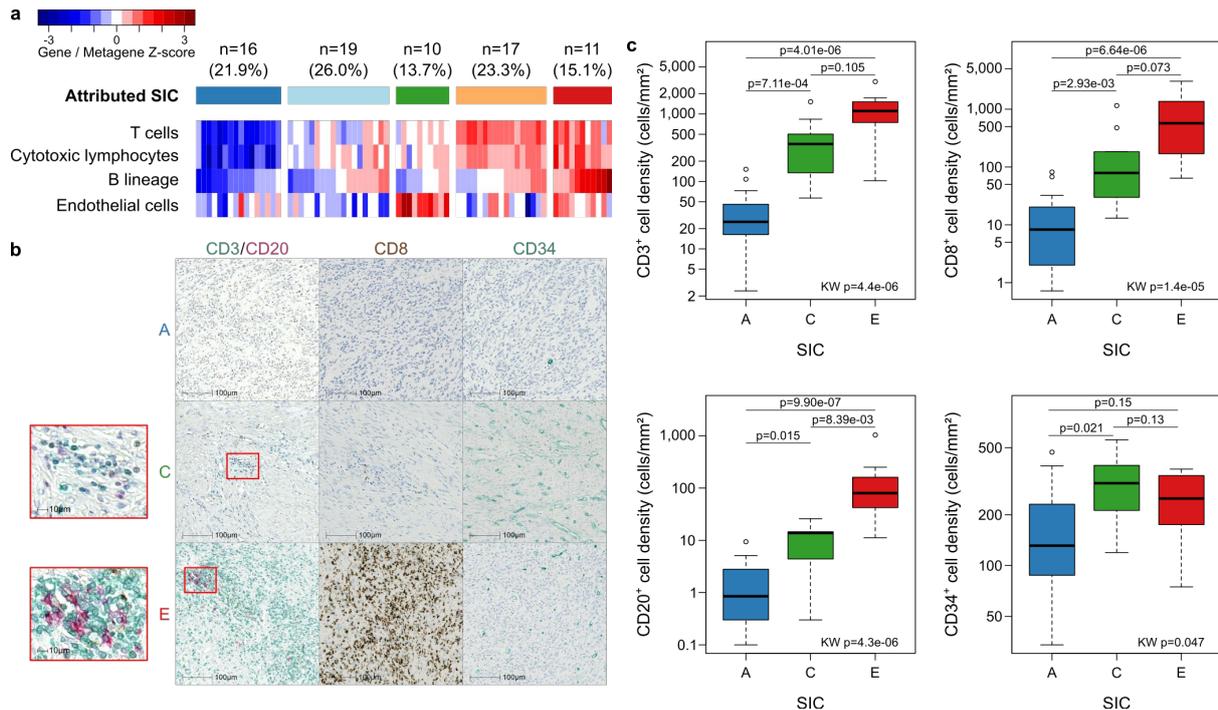


Figure 13 Validation de la classification génomique en IHC dans la cohorte de 73 sarcomes de l'étude de Petitprez et al. *Nature (accepted)*

(a) Classification génomique selon les Z-scores du MCP-counter ; (b) Images représentatives des classe A, C et E marquées par CD3 (vert)/ CD29 (rose), CD8 (brun) et CD35 (vert) en IHC ; (c) Densité cellulaire de CD3, CD8, CD20 et CD34 en IHC en fonction des 3 principales classes, A, C, E.

L'utilisation de cette signature TLS sur les échantillons tumoraux collectés lors de l'étude SARC028 évaluant le pembrolizumab dans une cohorte de sarcomes permet de montrer que les taux de réponse étaient significativement associés à la classification immunologique : 0% pour les patients des classes A et B, 22% et 25% dans les classes C et D et 50% dans la classe E ($p = 0,026$). Ces données d'un grand intérêt ont amené à proposer une sélection des patients sur la base de la présence de TLS sur du matériel tumoral à l'inclusion dans une nouvelle cohorte de l'essai clinique PEMBROSARC actuellement en cours en France au sein du GSF-GETO et promu par l'Institut Bergonié (NCT02406781). Nous n'avons pas évalué la présence de TLS dans les échantillons de notre étude. Cependant, nous avons appliqué la méthode de classification de Petitprez et collègues sur les données de transcriptome de notre cohorte de patients UPS et avons trouvé une corrélation quasi-parfaite entre nos trois groupes immune-high, immune-low et other et les cinq classes de l'étude (données non publiées).

Au total, en plus d'être des marqueurs potentiels de sélection pour l'immunothérapie, la mise en évidence de ces différentes signatures immunologiques transcriptomiques et immuno-histo-chimiques dans un sous-groupe d'UPS spécifique fait émerger de nombreuses questions en parallèle et encourage à développer d'avantage d'approches immuno-modulatrices. Tout

d'abord, comment développer des outils reproductibles permettant de sélectionner les patients ? Dans le cas de l'expression de marqueurs tels que la densité de CD8 et IDO1, le seuil permettant de classer les échantillons « immune-low » ou « immune-high » est variable en fonction des cohortes analysées et ne peut être facilement normalisé. La présence ou l'absence de TLS semble un marqueur plus facilement utilisable, au moins dans un premier temps à visée exploratoire. Cependant, 25% de patients TLS-négatifs appartenant aux classes C et D et à l'inverse seulement 50% des patients TLS-positifs de la classe E ont présenté une réponse objective dans l'étude SARC028. Les TLS sont le site d'instruction et de sélection des cellules anti-tumorales (55). Si les TLS sont absents, les cellules lymphocytaires intra-tumorales ne peuvent pas être éduquées à reconnaître les antigènes tumoraux, dont le répertoire est très large, et même si revigorées par les inhibiteurs de checkpoints tels que les antiPD-L1, elles ne peuvent être capables d'exercer une activité anti-tumorale efficace. Comment améliorer la formation de TLS dans les sarcomes ? Les analyses réalisées dans notre étude soulignent la présence de nombreux effecteurs lymphocytaires mais de nombreux freins sont également exprimés tels que LAG3, TIGIT, TIMD4, IDO1, en plus - et peut être de façon plus importante - que PD-L1. Quelles approches immuno-modulatrices pourraient être plus efficaces dans ces tumeurs ? L'inhibition des différentes voies immunosuppressives identifiées pourraient être une piste. La co-stimulation de co-facteurs exprimés tels que CD27 pourrait également être envisagée (56). Les infiltrats macrophagiques pro-tumoraux de type M2, l'angiogenèse pro-tumorale sont aussi des cibles complémentaires potentielles actuellement explorées dans des études cliniques de phase précoce, mais dans des populations de patients atteints de sarcomes non sélectionnés, avec le risque de diluer un signal déjà faible. Le rôle de ces voies immuno-suppressives et l'effet de leur ciblage dans les UPS restent de plus à définir.

Les études d'immunothérapie dans les sarcomes sont également limitées par le peu de modèles précliniques disponibles. Le modèle le plus utilisé est celui de la lignée de fibrosarcome murin MCA205. La lignée MCA205 est dérivée d'un fibrosarcome induit par le méthylcholanthrène dans la lignée de souris immunocompétente C57BL/6. Nous avons pu mener une collaboration avec le laboratoire ImmuSmol (Alban Bessède) visant à caractériser l'impact de l'inhibition de la voie de signalisation de la kynurénine/IDO1 dans les sarcomes (57). Le niveau d'expression de kynurénine en intra-tumoral était élevé dans ce modèle de souris C57BL/6 greffées avec des tumeurs de type MCA205. Le traitement avec le GDC-0919, un inhibiteur d'IDO1, seul ou en association à un anti-PD-L1 (clone10F9-G2) induisait une diminution significative du ratio kynurénine/tryptophane dans le plasma des souris, et de

la kynurénine en micro-dialyse intra-tumorale. Le GDC-0919 seul n'avait pas d'effet anti-tumoral, toutefois l'ajout de l'anti-IDO1 à l'anti-PD-L1 permettait une amélioration de l'effet anti-tumoral non significative (1/15 rejet de greffe tumorale avec l'anti-PD-L1, 3/15 rejets avec la combinaison) (Figure 14). L'analyse par FACS des tumeurs permit de montrer que l'anti-PD-L1 induisait un afflux de cellules CD45+, largement composées de LT-CD4+ et LT-CD8+ activés, associé à une production d'INF-gamma et à une limitation des macrophages immunosuppresseurs de type M2 (CD38-/Egr2+) et une augmentation des macrophages de type M1 (CD38+/Egr2-), tandis que le GDC-0919 seul ou associé à l'anti-PD-L1 n'apportait pas de changement significatif. L'analyse du transcriptome par Nanostring (PanIO360 panel) de tumeurs non traitées et traitées par l'anti-PD-L1 a permis de mettre en évidence une augmentation de l'expression d'*IDO1* et d'*IDO2* sous traitement, corrélée à l'expression d'*IL6* ($R^2 = 0.69$) et d'*INFg* ($R^2 = 0.36$), ainsi que des gènes d'enzymes d'aval de la voie, telles que *KYNU* (kynuréninase). De façon intéressante l'anti-PD-L1 induisait une régulation négative des voies Wnt/bêta-caténine, Hedgehog et de l'autophagie. En dehors d'une augmentation de *INH*, un inhibiteur de la voie du TGF-bêta, et d'une diminution de l'expression de *pvr*, partenaire de TIGIT, un inhibiteur T présent notamment sur les Lymphocytes NK, et de *Gzma*, codant pour le granzyme A, secrété notamment par les Lymphocytes NK, le GDC-0919 seul ou ajouté à l'anti-PD-L1 n'apportait en revanche pas de modulation significative supplémentaire des clusters de l'immunité. Ces éléments suggèrent d'autres voies impliquées dans la résistance primaire à l'inhibition de PD-L1 et d'IDO1, comme potentiellement l'inhibition des lymphocytes NK. Ce type de modèle a également ses limites et la transposition à d'autres modèles de sarcomes plus pléomorphes et immunogéniques reste à évaluer. Des études récentes ont montré dans des modèles de cellules souches embryonnaires que l'inhibition d'IDO1 réduisait la prolifération, mais qu'une quantité de kynurénine persistait dans le milieu, permettant le maintien du signal paracrine de maintenance de l'état indifférencié dans ces modèles (58). L'effet immuno-suppresseur de la kynurénine dans le milieu tumoral est directement lié à l'action de son récepteur intracellulaire AhR, exprimés par les cellules immunitaires mais également les cellules tumorales. AhR est activé par de nombreux ligands et est également impliqué dans plusieurs voies oncogéniques (59). Plusieurs inhibiteurs de AhR sont en développement, avec des résultats précliniques encourageants (60, 61), et pourraient être testés dans les modèles disponibles de sarcomes murins immuno-compétents.

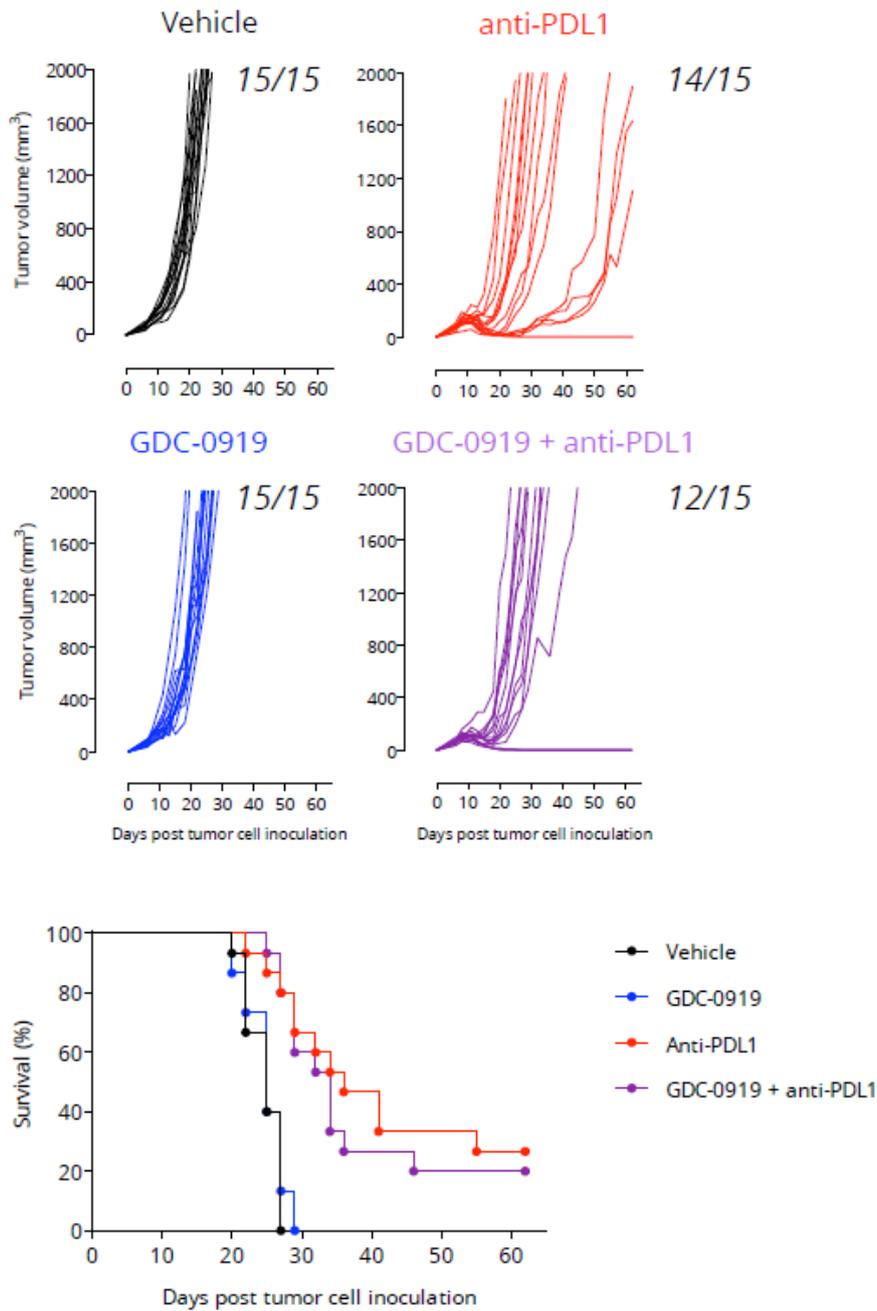


Figure 14 Effet anti-tumoral in vivo de l'inhibition de PD-L1 et IDO1 dans des modèles de souris C57BL/6 greffées avec des tumeurs de type MCA205 (Nafia et al. *Frontiers in Oncology* (submitted))

Les souris étaient randomisées entre un traitement par l'anti-PD-L1 murin (clone 10F9-G2, BioXCell) en intra-péritonéal à 5mg/kg à J6, 9, 12, et 15, et/ou l'anti-IDO1 GDC-0919, administré 2 fois par jour à 100mg/kg de J6 à J26. Le volume tumoral était évalué 3 fois par semaine selon la formule $V = (\text{Length} * \text{Width}^2) / 2$.

Immune-low UPS

Un autre résultat intéressant de cette étude est le fait que le groupe des UPS immune-low soit associé au développement et au caractère souche. Dans notre étude, nous avons retrouvé *FGFR2* surexprimé parmi plusieurs gènes codant pour des facteurs de transcription et des gènes du développement et du caractère souche, tels que *LHX8*, *LRRN1*, *LGR5*, et ceci de façon reproductible aussi bien dans notre cohorte de 25 UPS que dans la cohorte indépendante de validation de 41 UPS du TCGA consortium.

Les récepteurs FGFR sont une famille de quatre récepteurs activés par 22 ligands différents, qui régulent de nombreux processus biologiques incluant le développement embryogénique, la maintenance des cellules souches, la prolifération, la différenciation, la migration, l'apoptose ou le métabolisme (62). Après fixation d'un ligand, la dimérisation des récepteurs active plusieurs voies d'aval : la voie des MAP kinases, la voie PI3K/Akt et la voie STAT (63). Nous avons pu mettre en évidence *in vitro* et *in vivo* une sensibilité différentielle de nos modèles au JNJ-42756493 ou Erdafitinib (64), ainsi qu'au BGJ298 (65) et à l'AZD4547 (66), trois inhibiteurs spécifiques des récepteurs FGFR. Les UPS immune-low, caractérisés par une surexpression de *FGFR2*, sont apparus significativement plus sensibles à cette approche pharmacologique que les UPS immune-high et les other. Nous avons pu mettre en évidence une diminution de la phosphorylation de ERK, situé en aval de la voie des MAP kinases et un blocage en phase G0/G1 dans les lignées sensibles à l'inhibiteur JNJ-42756493. Plusieurs remarques peuvent être faites :

Tout d'abord, le fait que l'inhibition de *FGFR2* soit spécifiquement responsable de l'effet anti-tumoral dans nos modèles peut être remis en question. En effet, il n'existe pas d'inhibiteur spécifique de *FGFR2* et l'effet anti-tumoral observé pourrait également passer par l'inhibition des autres récepteurs. Malgré la double contrainte de la disponibilité de tumeur fraîche et du succès aléatoire des greffes tumorales sur modèles murins, nous avons réussi à obtenir une lignée et un PDX représentatifs de chacune de nos trois classes, immune-low, immune-high et other. IB119, dérivée de l'échantillon IC260, est la lignée représentative du groupe other (le plus petit groupe pour lequel nous n'avons que trois échantillons disponibles pour la caractérisation du statut *FGFR1,2,3,4* en qRT-PCR). Malgré une expression de *FGFR4* dans cette lignée et l'échantillon correspondant en qRT-PCR, elle était moins sensible à l'inhibiteur JNJ42756493 que les deux lignées du groupe immune-low, qui avaient la plus forte expression de *FGFR2* en qRT-PCR. La lignée représentative du groupe immune-high, KN473, dérivée de l'échantillon éponyme, était caractérisée par une expression de *FGFR1* en

qRT-PCR. Malgré cela, elle était la plus résistante au JNJ-42756493. Ces données sont spéculatives, mais vont dans le sens d'un effet de l'inhibiteur de FGFR obtenu par le ciblage de FGFR2 plutôt que FGFR4 ou FGFR1 dans ces modèles. Pour essayer d'aller plus loin, nous avons développé une approche de shRNA spécifique de FGFR2. Nous avons pu mettre en évidence une grande sensibilité à l'inhibition de FGFR2 dans le modèle IB106, surexprimant FGFR2, traduite par un fort taux d'apoptose. A l'inverse, l'inhibition de FGFR2 dans la lignée KN473 n'entraînait pas de modification en termes d'apoptose. Pour être plus complet, une approche séquentielle de sh spécifiques des différents FGFR1, 2, 3, 4 pourrait être proposée pour évaluer l'effet relatif de l'inhibition de chaque récepteur dans les lignées respectives.

D'autre part, le fait que l'événement génomique touchant FGFR2 retrouvé dans le groupe immune-low soit une surexpression relative questionne l'impact thérapeutique en clinique. Des aberrations des différents récepteurs à tyrosine kinase de la famille des FGFR sont retrouvées dans 7% des cancers, intéressant FGFR2 dans 1,5% des cas. La majorité sont des amplifications (66%), suivies des mutations (26%) et des réarrangements (8%). Les cancers les plus représentés sont les carcinomes urothéliaux (32% d'aberrations), mammaires (18%), de l'endomètre (13%) ou bronchiques (13%) (67). Des surexpressions et des amplifications de différents FGFR et/ou de leurs ligands ont également été rapportées dans d'autres sarcomes tels que les liposarcomes myxoïdes (68), les rhabdomyosarcomes, les léiomyosarcomes, les DDLPS, les myxofibrosarcomes ou les UPS (69), avec une validation *in vitro* et parfois *in vivo* de l'activité anti-proliférative d'inhibiteurs de FGFR sur des modèles dédiés. En 2005, Ishibe et collègues ont rapporté l'expression de plusieurs FGF et FGFR en RT-PCR dans des tumeurs et lignées de synoviosarcomes (70). Ils ont pu montrer dans leurs modèles *in vitro* que l'activation de la voie FGFR par différents ligands FGF induisait la prolifération cellulaire et l'induction de la voie MAPK/ERK. A l'inverse, l'utilisation d'un inhibiteur de FGFR (SU5402) était associée à une régulation négative de la phosphorylation de ERK1/2, et à une diminution de la prolifération cellulaire et de la croissance tumorale, *in vitro et in vivo*. En 2017, Chudasama et collègues ont rapporté l'existence d'hyper-expressions et d'amplifications de FGFR1 dans plusieurs types de sarcomes (69). Ils sont partis du cas d'une patiente avec un léiomyosarcome ayant une amplification de *FGFR1* qui a présenté une réponse dissociée et retardée sous un inhibiteur de FGFR (BJ398), suivi d'une très bonne réponse prolongée sous double inhibition FGFR/VEGF/PDGF et mTOR avec l'association nintedanib - everolimus. Ils évaluèrent alors l'expression de *FGFR1* en FISH sur une cohorte de 190 patients avec différents types de sarcomes, et retrouvèrent un gain de *FGFR1* dans

32% et une amplification dans 7% des cas, respectivement. Dans une seconde cohorte de 79 patients, l'analyse des variations du nombre de copies des gènes par CGH leur permit d'identifier 13 cas (16.5%) avec gain ou amplification de *FGFR1*. Parmi ces 79 patients, 16 (20%) avaient une surexpression de *FGFR1* en séquençage ARN, associée à un gain/amplification en CGH pour cinq d'entre eux. L'analyse des données de séquençage ARN de 254 sarcomes du TCGA leur permit de retrouver à nouveau une surexpression de *FGFR1* dans 15.4% des cas. Parmi ces derniers, cinq et 17 avaient une amplification ou un gain de *FGFR1*, tandis que 14 et trois étaient soit diploïdes, soit avaient une perte du locus de *FGFR1*, illustrant qu'une surexpression de *FGFR1* peut être retrouvée en l'absence d'augmentation du nombre de copies. Ils ont ensuite utilisé six lignées caractérisées par trois niveaux d'expression de *FGFR1* en RT-PCR et de nombre de copies de *FGFR1* en FISH (amplification, haut niveau d'expression et niveau modéré d'expression). Ils ont également retrouvé la voie MAPK/ERK comme voie effectrice préférentielle suite à l'activation de FGFR1 dans ces lignées, et démontré que l'effet des siRNA et des trois inhibiteurs de FGFR utilisés (PD173074, AZD4547 et BGJ398) était lié au niveau différentiel d'expression de *FGFR1* dans ces modèles, la lignée amplifiée *FGFR1* étant la plus sensible à l'inhibition. Au total, il existe une faible corrélation entre le niveau d'expression, les altérations du nombre de copies, ou la surexpression protéique des FGFR, et la valeur prédictive de ces différentes altérations en clinique est variable selon les molécules évaluées. Une étude de phase I testant le Rogaratinib chez des patients atteints de différents cancers a récemment été rapportée (71). Dans la cohorte d'extension, les patients étaient sélectionnés sur la base de l'expression tumorale de FGFR, évaluée par RNAscope et Nanostring. Une surexpression était définie par un signal en Nanostring > 800 ou un score RNAscope de 3 ou 4. Au total, 813 patients ont eu un test FGFR réalisé et valide, et 103 patients ont été inclus dans la cohorte d'extension. Parmi les 100 patients évaluables pour la réponse en imagerie, 15 (15%) ont eu une réponse objective, et 56 (56%) une stabilité. Douze des 15 patients répondeurs étaient porteurs d'un carcinome urothélial, un avec une surexpression de *FGFR1* et 11 de *FGFR3*, dont cinq avec également une anomalie de type mutation activatrice ou fusion au niveau de *FGFR3*. A noter que parmi les 19 patients atteints d'un sarcome enrôlés dans cette étude, sept (37%) avaient une surexpression de *FGFR* et ont pu être inclus dans la cohorte d'extension, mais qu'aucun n'a présenté de réponse objective. Plus que la surexpression ou les amplifications, ce sont les anomalies de FGFR de type fusion ou certaines mutations activatrices qui ont montré l'impact thérapeutique le plus significatif. L'erdafinitinib a par exemple récemment montré une efficacité clinique dans le cancer urothélial métastatique uniquement chez des patients chez

qui une mutation de *FGFR3* ou une fusion impliquant *FGFR2* ou *FGFR3* avait été retrouvée sur un échantillon tumoral (72). Dans cette étude, la détermination de l'anomalie génétique devait être centralisée, et l'autorisation par la FDA et l'EMA de la molécule dans cette indication spécifique s'accompagne de la recherche de l'anomalie par un test compagnon.

Une troisième question qui peut être posée est celle du rôle *driver* de cette surexpression de *FGFR2* dans ces tumeurs immune-low, ou bien le fait que celle-ci soit la conséquence d'activation de programmes transcriptionnels plus spécifiquement impliqués dans le développement, le caractère souche et la différenciation mésenchymateuse.

Plusieurs protéines appartenant à la voie MYC étaient différentiellement représentées dans nos analyses de protéomique. Cependant, c-Myc lui-même n'était pas inclus dans les protéines détectées par la plateforme de protéomique. *MYC* n'était pas non plus différentiellement exprimé dans les données de séquençage ARN (immune-low = 5.7, immune-high = 6.8, ns), ni amplifié dans les données de séquençage de l'exome.

C-Myc est un facteur de transcription aux effets pléiotropes impliqué dans de nombreuses voies de signalisation, notamment l'EMT et le renouvellement et l'entretien du caractère « souche » (73, 74). De façon intéressante, ces deux voies de signalisation étaient également significativement différentiellement représentées dans nos analyses de séquençage ARN et de protéomique, en faveur du groupe immune-low. De plus, *FGFR2* est un gène cible de c-Myc, tandis que c-Myc est un effecteur d'aval de la voie *FGFR/MEK/ERK* (75). C-myc est aussi fortement impliqué dans l'instabilité génomique (76). C-Myc agit comme un activateur transcriptionnel en se fixant à l'ADN après hétéro-dimérisation avec un autre facteur de transcription, Max (77, 78). L'activité de c-Myc est liée à une augmentation de l'acétylation des lysines situées au niveau des histones, qui permet le recrutement de protéines adaptatrices avec des motifs de liaison à ces lysines acétylées ou bromodomaines. BRD2, 3, 4 sont des membres de la sous-famille des protéines à bromodomaine extra-terminal, qui s'associent à la chromatine acétylée et participent d'un signal de transduction qui facilite l'activation transcriptionnelle (79). La mise en évidence d'une interaction de BRD4 avec le facteur P-TEFb (positive transcription elongation factor complex) et de c-Myc avec ce même facteur pour le contrôle de l'enzyme Pol II dans les régions promotrices qu'il gouverne a établi le rationnel pour évaluer l'impact des inhibiteurs de bromodomaines sur la transcription médiée par c-Myc. Delmore et collègues ont mis en évidence un effet inhibiteur du JQ1, un inhibiteur de BET bromodomaines tels que BRD4 notamment, sur la transcription de *FGFR2* et *MYC* dans un modèle de myélome (80). Ils ont pu montrer que l'expression de *MYC* lui-même et ses fonctions transcriptionnelles pouvaient être modulées pharmacologiquement par

l'inhibiteur de BET bromodomaines. Garcia-Carpizo et son équipe ont également montré dans des modèles de lymphomes et leucémies que les protéines co-activatrices à bromodomaines CBP (CREB (cAMP responsive element binding protein) binding protein)/P300 sont nécessaires à la prolifération médiée par c-Myc, et que le CBP30, un inhibiteur de CBP/P300, interférait avec la transcription médiée par c-Myc en diminuant le taux d'histones acétylés dans les régions régulatrices de liaison de c-Myc et réduisait ainsi l'expression des gènes contrôlés par c-Myc (81). Ceci provoquait une accumulation des cellules en G0/G1. Nous nous sommes basés sur les travaux de Delmore et de Garcia-Carpizo pour explorer l'activité de nouveaux inhibiteurs spécifiques de CBP/P300 dans nos modèles de sarcomes. Nous avons pu mettre en évidence une sensibilité différentielle aux trois composés inhibiteurs de CBP/P300 CPI637, NEO1132, et NEO 2734 entre nos lignées d'UPS immune-low et immune-high. Toutefois, ces résultats sont très préliminaires et les mécanismes fonctionnels sous-jacents à ce différentiel de sensibilité restent à explorer.

La transformation de cellules souches mésenchymateuses en tumeur de type sarcome indifférencié est obtenue d'une part par la perte de régulateurs du cycle cellulaire et l'induction d'une instabilité chromosomique (27), mais également par des altérations des voies Wnt/bêta-caténine, Notch et PI3K/Akt (28, 35). Une surexpression de gènes impliqués dans les voies Hedgehog et Notch a été rapportée dans une sous-population de cellules souches d'UPS par Wang et al. (39). Dans notre étude, nous avons pu également retrouver une surexpression de ces gènes dans le groupe d'UPS immune-low : *HEY2* (FC = 8.17), *HEYL* (FC = 2.88), *HEY1* (FC = 1.9) et *HES1* (FC = 2.82) pour la voie Notch et *GLI1* (FC = 13.13), *PTCH1* (FC = 2.13) ; *HHIP* (FC = 7.83) la voie Hedgehog. Des inhibiteurs de ces voies complexes sont développés actuellement. Une perspective de ce travail pourrait être d'explorer plus loin le rôle de ces voies dans les UPS.

DKK1, fortement impliqué dans la régulation de la voie de la bêta-caténine, était parmi les gènes les plus surexprimés dans le sous-groupe immune-high (FC = 55.28 dans notre cohorte), et ceci de façon reproductible dans une cohorte indépendante d'UPS du TCGA consortium. Ces données corroborent les travaux de Gibault et de Matushansky dans lesquels l'hyper-expression de *DKK1* était aussi associée d'une part à l'expression de gènes de l'inflammation, d'autre part au caractère indifférencié (25, 38). Dans l'étude de Gibault, les deux groupes de sarcomes les plus indifférenciés comportaient une forte surexpression de *DKK1* et une sur-représentation des gènes de l'inflammation y était associée dans le groupe E, suggérant que ce groupe soit notre groupe d'UPS immune-high. Mathushansky et collègues

ont également rapporté que l'inhibition de la voie de la bêta-caténine, notamment par la surexpression de *DKK1*, inhibait la différenciation tissulaire de cellules souches mésenchymateuses et concourrait à leur transformation en UPS. Ils ont notamment montré que l'engagement dans la différenciation était médié par *Wnt2* et la voie canonique de la bêta-caténine. Ils ont également montré une haute expression de *MYC*, *CCND1* et *WNT2* dans leurs UPS. La modulation de la voie de la bêta-caténine régule finement la balance entre cellules souches et cellules différenciées à travers de multiples effecteurs qui diffèrent selon la force de cette modulation. Ces nuances de modulation ont été largement décrites dans les modèles de cellules souches épithéliales, notamment de follicules pileux (82). Dans ces modèles, un faible niveau d'activation de la voie de la bêta-caténine stimule la prolifération et la mobilisation des cellules souches, permettant la croissance du follicule pileux sans affecter la morphologie de la niche, alors que des niveaux plus élevés d'activation de la voie stimulent des programmes transcriptionnels de différenciation, tandis que la perte de l'activation de cette voie résulte en la perte de l'expression des marqueurs souches par ces cellules. Cette activation différentielle est modulée par l'expression différentielle d'antagonistes comme *sFRP1*, *DKK3* et de facteurs de transcription promoteurs (*Lef1*) et inhibiteurs (*Tcf3*).

Des études récentes ont démontré un lien entre l'inactivation de la voie Wnt/bêta-caténine et le recrutement des cellules immunitaires dans de nombreux types de tumeurs (83). L'équipe de Spranger a notamment analysé un set de 266 mélanomes métastatiques issus du TCGA qu'ils classèrent selon l'expression ou non de gènes immunitaires en séquençage ARN (84). De façon intéressante, 48% des tumeurs non infiltrées (immune-low) présentaient une activation de la voie de la bêta-caténine, se traduisant par une surexpression de gènes tels que *APC2*, *SOX2*, *SOX11*, *WNT7B*. En analysant 197 tumeurs supplémentaires par séquençage de l'exome, 11% comportaient une délétion d'un régulateur négatif de la voie de la bêta-caténine tel que *APC*, *AXIN1*, *TCF1*, et 8% une mutation gain de fonction de *CTNNB1*. Ils créèrent un modèle de mélanome murin conditionnel *Braf*^{V600E}/*PTEN*^{-/-} +/- *CTNNB1*-stabilisée (*CAT-STA*). L'induction d'une stabilisation de la bêta-caténine dans ce modèle s'associait à une absence de lymphocytes intra-tumoraux ainsi qu'à une disparition de l'expression de *CCR5* par les cellules dendritiques CD103/CD8a, et à une inhibition de l'expression de *CCL4* par les tumeurs du fait d'un facteur transcriptionnel répresseur *ATF3* induit par l'activation de la voie de la bêta-caténine, *CCL4* étant une chimiokine nécessaire à l'activation des cellules dendritiques, qui à leur tour activent les lymphocytes (Figure 15). Nous avons également pu mettre en évidence une surexpression relative de gènes de la voie de la bêta-caténine tels que *ARHGEF4* (FC = -13.90), *HEY2* (FC = -8.17) *LEF1* (FC = -5.44) et *FZD1* (FC = -2.01),

associée à une sous-expression relative de *CCR5* (FC = 11.86) et *CCL4* (FC = 15.78) dans les UPS immune-low par rapport aux UPS immune-high, ainsi qu'une délétion d'*AXINI* dans 4 (40%) de ces UPS immune-low. Pour aller plus loin dans la caractérisation de la voie de la bêta-caténine dans nos UPS, nous avons eu l'occasion d'une collaboration avec l'équipe du Pr Bertucci à Marseille qui a développé une signature d'activation de la voie de la bêta-caténine dérivée d'un métagène-set publié par Gatza et collègues (85). L'équipe du Pr Bertucci a récemment démontré le caractère pronostique négatif de cette signature d'activation de la voie de la bêta-caténine en termes de survie sans métastase dans une cohorte de 610 patients atteints de sarcomes. Ils ont pu également montrer dans une cohorte de 1379 patients atteints de sarcomes issus de bases publiques que des valeurs élevées de leur signature étaient associées au site « extrémité », au profil « génomique complexe », au grade « élevé », ainsi qu'à l'histologie « léiomyosarcomes » (86).

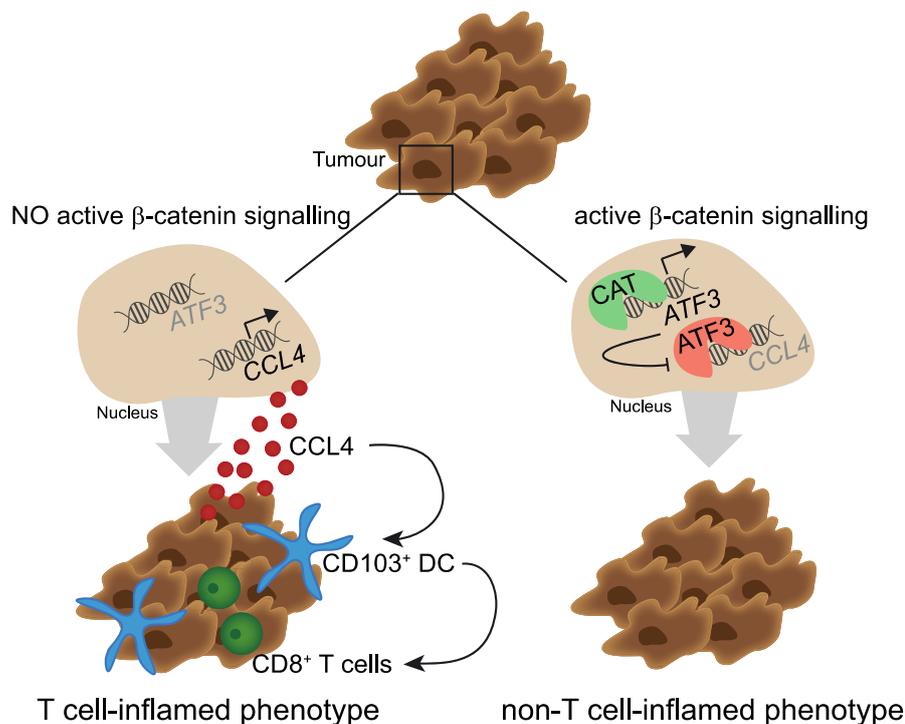


Figure 15 Schéma récapitulatif de l'activité immuno-suppressive de la voie intrinsèque de la bêta-caténine dans le mélanome proposé par Spranger *et al.* Nature, 2015

A gauche, une tumeur sans activation de la voie de la bêta-caténine, dans laquelle ATF3 n'est pas transcrit. CCL4 est sécrété, ce qui permet le recrutement des cellules dendritiques (DC) CD103+ et l'activation des lymphocytes CD8+ ; A droite, une tumeur avec activation de la voie de la bêta-caténine, qui induit la transcription d'ATF3, qui lui-même inhibe la transcription de CCL4 et empêche le recrutement des cellules dendritiques.

Nous avons testé cette signature sur notre cohorte de 21 patients immune-high et immune-low. La grande majorité des tumeurs immune-low avaient une valeur élevée de cette signature, tandis que cette valeur était plus hétérogène dans les tumeurs immune-high (Figure 16).

Ces données amènent à questionner la teneur exacte et l'effet de l'activation de la voie de la bêta-caténine dans les UPS, et le rôle différentiel des autres voies altérées. Martinez-Font et collègues ont récemment montré un effet apoptotique in vitro de plusieurs inhibiteurs de la voie de la bêta-caténine tels que le PKF118-310, un disrupteur de la liaison TCF/bêta-caténine, et IWR-1 et XAV939, deux stabilisateurs de l'inhibiteur AXIN1, sur différentes lignées de léiomyosarcomes, fibrosarcomes et liposarcomes, associé à une diminution de l'activité transcriptionnelle de cette voie (87, 88).

Il serait intéressant d'évaluer l'effet différentiel de ces composés dans nos modèles d'UPS.

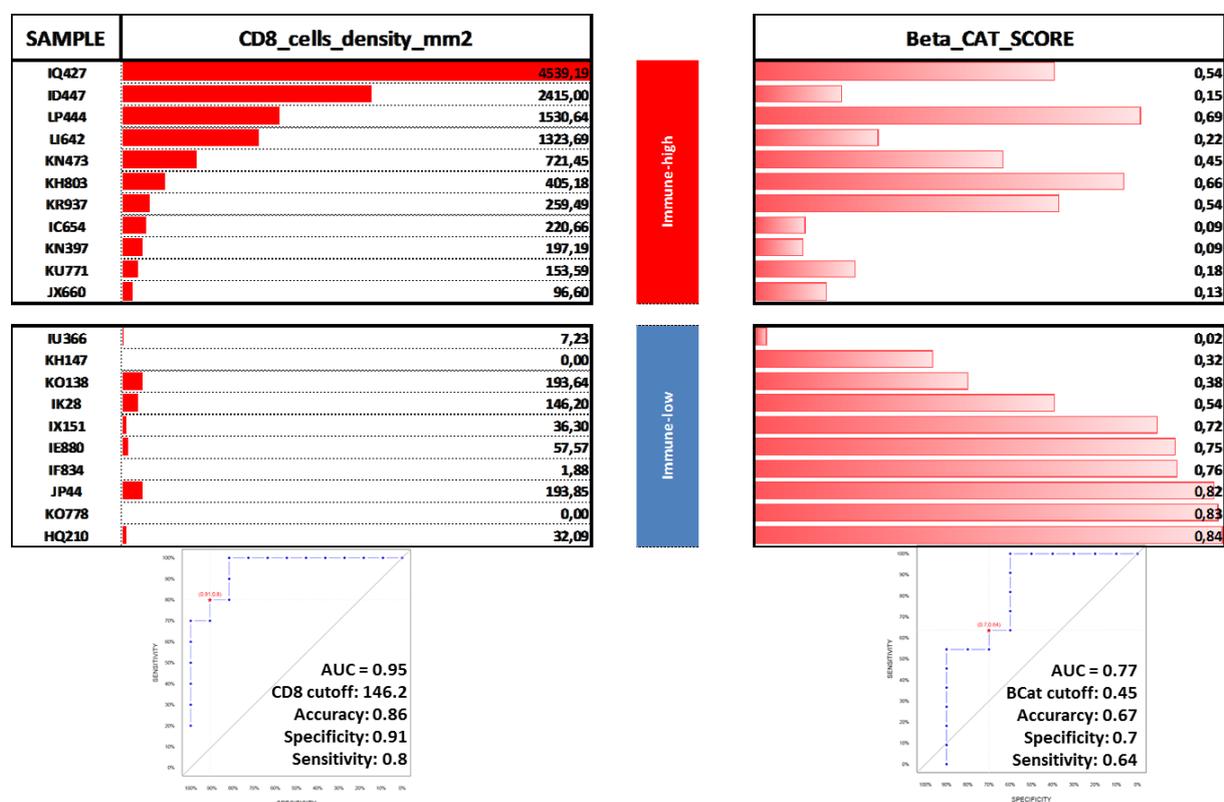


Figure 16 Relation entre la densité de CD8 et le score d'activation de la bêta-caténine développé par Bertucci *et al.* Mol Cancer Ther, 2017, dans les 21 UPS immune-high et immune-low de notre étude

Les échantillons sont classés en deux groupes : immune-high (rouge) en haut, et immune-low (bleu) en bas, puis répartis en fonction de la densité intra-tumorale de cellules CD8+ dans la colonne de gauche ; La valeur correspondante du score d'activation de la voie de la bêta-caténine (BCat) est indiquée dans la colonne de droite ; En dessous se trouvent les courbes ROC donnant la meilleure valeur en termes de précision, spécificité et sensibilité pour CD8 (gauche) et le score Bcat (droite).

De façon intéressante, 50% des UPS immune-low de notre étude comportaient une perte de *PTEN*, contre 27% pour les UPS immune-high. George et collègues ont récemment publié un papier élégant montrant à partir d'un cas de patiente porteuse d'un léiomyosarcome que la perte de *PTEN* était associée à la résistance aux inhibiteurs de PD-1 (89). Chez cette patiente avec une réponse complète de plus de 2 ans à l'anti-PD-1 pembrolizumab ayant présenté une ré-évolution tardive d'une unique métastase, l'analyse de la tumeur primaire et de la métastase résistante par séquençages ARN et exome complets permit d'identifier une amplification focale de *MYC* et une mutation mis-sense de *PTEN* avec perte de l'hétérozygotie dans la métastase résistante, en plus d'une mutation non-sens de *TP53* partagée par les deux échantillons. Ceci poussa l'équipe à analyser ensuite les données de séquençage ARN disponibles d'une cohorte de 241 sarcomes du TCGA consortium. Les tumeurs avec perte bi-allélique de *PTEN* (délétion hétérozygote et mutation oncogénique) étaient associées à des niveaux d'expression significativement plus élevés de STAT3 - marqueur de l'activation de la voie JAK2/STAT3 en cas de perte de *PTEN* - et plus bas de *PDCD*, *CD8A*, *INFG*, *PRF1*, *GZMA*, que les tumeurs n'ayant aucune ou qu'une des deux altérations.

Ceci fait écho au travail de Peng et collègues qui ont démontré dans des modèles murins de mélanome BRAF-muté que la perte de *PTEN* par shPTEN diminuait l'infiltration des lymphocytes T dans la tumeur et inhibait leur cytotoxicité *in vitro* et *in vivo* (90). En analysant les échantillons tumoraux de 39 patients traités par anti-PD-1, Peng et collègues ont également pu montrer que le statut PTEN en IHC corrélait significativement à la réponse tumorale (Figure 17A et B). En reprenant cette fois les données d'expression et de variations de nombre de copies des gènes dans la cohorte de mélanomes du TCGA consortium disponible publiquement, Peng et al. retrouvèrent également des scores d'infiltration lymphocytaire inversement corrélés à l'expression d'*AKT* et à la perte de *PTEN*. Les altérations des voies bêta-caténine et PTEN/Akt étaient retrouvées significativement plus fréquentes dans les tumeurs non-inflammées de façon non chevauchante (Figure 17C et D), suggérant que ces deux voies soient des voies indépendantes d'exclusion du système immunitaire au niveau tumoral. Enfin, le traitement de ces modèles avec un inhibiteur sélectif de PI3K-bêta améliorerait l'efficacité des anti-PD-1 et anti-CTLA-4.

Les données de protéomique relatives aux sarcomes sont rares. Une étude de 2006 a porté sur la protéomique de 80 sarcomes en général, incluant 18 MFH pléomorphes pouvant être considérés comme des UPS (91). Les auteurs identifièrent des profils d'expression protéique par électrophorèse sur gel en 2-D (2-D difference gel electrophoresis, 2-D-DIGE). Par une

méthode d'évaluation des ratios d'intensité de saturation de la coloration par la cyanine Cy5 (spécifique de chaque échantillon) par rapport à la Cy3 (contrôle interne fait d'une mixture de l'ensemble des échantillons protéiques de la cohorte), ils analysèrent l'expression des spots protéiques les plus différentiels entre les différents histologies de la cohorte (sarcomes à cellules claires, liposarcomes myxoïdes, GIST, MPNST, synoviosarcomes, MFH (10 myxoides et 18 pléomorphes) et léiomyosarcomes (7 conventionnels et 3 pléomorphes). Les 10 spots les plus différentiellement exprimés furent retenus pour chaque histologie. La classification des tumeurs à partir de ces 67 protéines permet de retrouver les léiomyosarcomes et les MFH ensemble et les autres histologies dans un autre cluster.

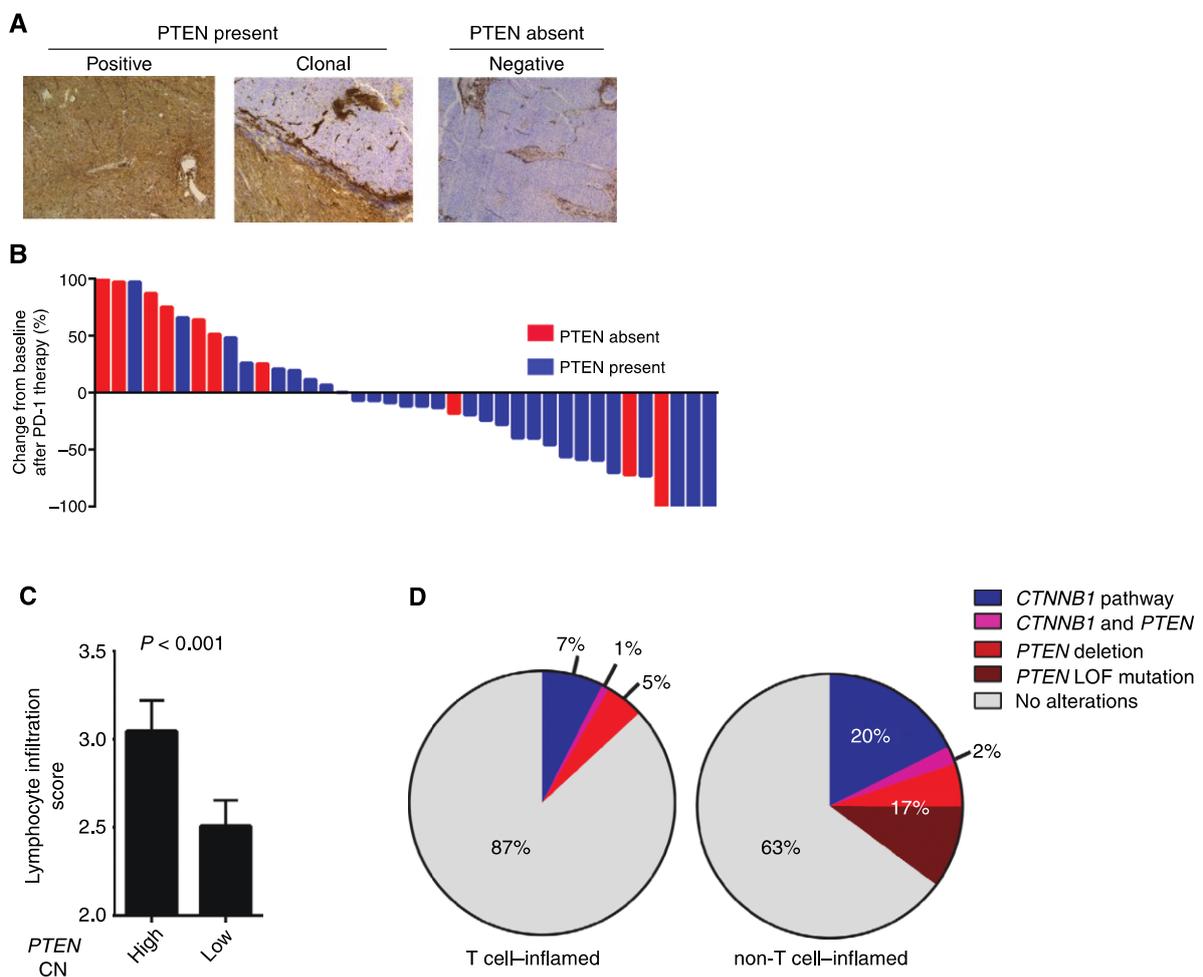


Figure 17 La perte de *PTEN* est associée à une moins bonne réponse aux anti-PD-1 et une moindre infiltration tumorale lymphocytaire T (Peng et al. Cancer Discovery, 2016).

A : Expression de *PTEN* en IHC dans les tumeurs de patients porteurs de mélanomes; B : Modifications de la taille tumorale en fonction de la présence ou non de *PTEN* en IHC dans les tumeurs de 39 patients atteints de mélanome et traités par anti-PD-1 (nivolumab, pembrolizumab); C : Comparaison des scores d'infiltration lymphocytaire entre les patients de la cohorte du TCGA en fonction du nombre de copies de *PTEN*; D : Fréquence des altérations génétiques touchant les voies de la bêta-caténine et *PTEN* entre les tumeurs infiltrées et non infiltrées par les lymphocytes T (mutation activatrice de *CTNNB1*, mutation perte de fonction des régulateurs *APC*, *PAC2*, *AXIN1*, *AXIN2*; perte et mutation-délétion de *PTEN*).

Les protéines « signatures » du groupe MFH étaient LDHB, CAPG, ANXA1, ACTB, FUBP2, HNRPK, LDHB, A1AT, LASP1. Les léiomyosarcomes pléomorphes et les léiomyosarcomes conventionnels se regroupaient séparément et exprimaient des protéines différentes : les léiomyosarcomes pléomorphes avaient en effet des taux d'expression de Tropomyosin 3 et 4, de Beta-actin, de CLIC1, et de T-plastin plus élevés, tandis que les léiomyosarcomes conventionnels exprimaient plus fortement la Tropomyosin 1 et 2, l'Alpha-actin 2 et l'alpha-actinin 1. Le gène de la Tropomyosin 3, *TPM3*, était le plus réarrangé dans notre cohorte d'UPS : 10 de nos 21 (48%) tumeurs analysées en exome avaient un gain de *TPM3*, et notamment 70% des tumeurs du groupe immune-low, versus 27% des immune high.

Radiomique

Au cours de ce travail, nous avons eu l'opportunité d'une collaboration avec le Dr Crombé, radiologue à l'Institut Bergonié et actuellement en thèse de mathématiques dans le domaine de la radiomique. Après contourage manuel des coupes IRM pré-thérapeutiques des 14 patients pour lesquels une imagerie IRM de bonne qualité était disponible, l'analyse de 62 composantes de la texture des images IRM de nos patients permit de définir neuf variables radiomiques significativement différentielles entre les UPS immune-high et les autres. Cette signature permettait d'identifier les patients du groupe transcriptomique immune-high avec de très bonnes sensibilité et spécificité. Ces données sont très préliminaires et demandent à être vérifiées et validées sur une plus large cohorte de patients atteints de sarcomes, ce qui est en cours sur l'institut Bergonié. Par ailleurs, la reproductibilité de ce type d'approche est questionnable : d'une part les algorithmes utilisés pour quantifier les multiples variables de radiomique peuvent être hétérogènes, d'autre part le contourage lui-même des zones d'intérêt sur chaque coupe de tumeur en imagerie est un point limitant crucial dans la validité de l'interprétation, la reproductibilité et la dissémination/interopérabilité de ce type de travail en multicentrique. Plusieurs projets de standardisation des techniques de contourage tumoral ainsi que des algorithmes de calcul des multiples données de texture ainsi générées sont en cours en ce sens, dans le cadre notamment de partenariat industriel utilisant l'intelligence artificielle.

Limites et remarques

Cette étude est principalement descriptive. L'analyse du transcriptome et de l'exome de ces 25 échantillons d'UPS a permis de générer une très grande quantité de données qui pourront être analysées de manière approfondie au travers de travaux de recherche dédiés à caractériser les mécanismes précis de la tumorigenèse des UPS.

Seuls les deux principaux groupes ont pu être caractérisés, le 3^{ème} groupe étant trop petit dans notre première cohorte pour permettre une analyse de cluster de gènes. L'analyse plus détaillée des données générées concernant le 3^{ème} groupe de la cohorte du TCGA pourrait permettre de mieux caractériser ce groupe intermédiaire, ni immune-low, ni immune-high.

Les résultats de transcriptome obtenus, ainsi que la recherche de mutations et de fusions, sont directement dépendants des étapes de filtrage et de traitement bio-informatique des données générées, d'autres algorithmes pourraient être développés et testés. Les analyses de protéomique sont également limitées par le nombre de protéines recherchées et détectées dans la bibliothèque utilisée (pour note, FGFR2 et MYC n'étaient pas dans la bibliothèque utilisée dans cette étude).

D'autre part, le rationnel à l'origine des analyses fonctionnelles réalisées est limitée, portant d'une part sur la surexpression différentielle d'un oncogène en séquençage ARN, FGFR2, et d'autre part la représentation différentielle de protéines impliquées dans des voies de signalisation très ubiquitaires, telle que celles de c-MYC. Il apparaît nécessaire d'aller plus loin dans la caractérisation du rôle fonctionnel de ces cibles potentielles dans nos modèles d'UPS. Nous n'avons pas réalisé de caractérisation épigénétique de notre cohorte. Il serait intéressant d'associer l'analyse du méthylome et l'évaluation de l'acétylation des histones. L'environnement métabolique n'a pas été exploré non plus. Les analyses protéomiques ont permis de mettre en évidence des voies de signalisation métabolique telles que la glycolyse, l'hypoxie, la phosphorylation oxydative, surreprésentées dans le groupe immune-high qui pourraient être explorées plus en détails dans une collaboration future avec une équipe s'intéressant à cette spécialité.

Enfin, la caractérisation de nos deux groupes repose principalement sur l'absence d'infiltration immunitaire dans un groupe de tumeurs, associée à la surexpression différentielle de voies de signalisation et à des variations différentielles du nombre de copies de gènes dans l'autre, mais la causalité de cette immuno-exclusion reste à explorer. L'existence de cette forte expression de gènes de l'immunité caractérisant le groupe immune-high a limité l'analyse plus détaillée des autres voies de signalisation représentées dans ce

groupe pouvant être responsable de cet enrichissement inflammatoire. Ces résultats amènent de nouvelles questions : Quelles altérations génomiques sont responsables du phénotype immunogène des UPS immune-high ? Quelles sont les voies de signalisation et les altérations sous-jacentes qui sont effectivement reconnues par le système immunitaire ? Sur le plan structurel, une caractérisation de l'antigénicité des mutations retrouvées dans les deux groupes pourrait apporter des informations. Quels mécanismes sont mis en œuvre pour contrôler et inhiber cette réponse immunitaire présente ? D'un autre côté, quel impact les inhibiteurs de FGFR ou de BET ont sur le système immunitaire *in vivo*, dans les UPS immune-low ? Le ciblage de ces voies pourrait-il permettre une meilleure reconnaissance par le système immunitaire ?

Au total, ce travail est descriptif et générateur de nombreuses données mais également de nombreuses nouvelles questions. Ces résultats apportent toutefois une certaine lumière sur l'hétérogénéité des UPS et permettent de nouvelles hypothèses pour avancer dans la compréhension des mécanismes sous-jacents à cette hétérogénéité.

Conclusion

Le développement d'un sarcome à génomique complexe est la manifestation d'altérations au niveau des points de contrôle du cycle cellulaire et d'une instabilité chromosomique résultante dans une cellule souche mésenchymateuse. Les événements et mécanismes primordiaux responsables de cette instabilité sont mal connus. La forte fréquence de mutations de *TP53*, *RBI* et d'autres gènes impliqués dans les voies de contrôle du cycle cellulaire, et le fait que les maladies héréditaires associées à une mutation germinale de ces gènes induisent un risque élevé de développer un sarcome à génomique complexe, suggèrent un rôle important de ces altérations dans la sarcomagenèse. Ces événements peuvent survenir dans une cellule souche mésenchymateuse 1/ à différentes étapes de sa différenciation, en fonction du niveau d'activation de plusieurs voies de signalisation impliquées dans cette différenciation, 2/ et à différents moments du cycle cellulaire et notamment de la mitose, entraînant des altérations génomiques structurelles différentes, principalement des altérations de la ploïdie.

La phylogénie des sarcomes à génomique complexe est un sujet de controverse et de recherche active. La complexité de la problématique est illustrée par les différents changements de classification des UPS au cours des 20 dernières années, et encore actuellement par les relectures et modifications faites dans les bases publiques, incluant celle du TCGA. Comme mentionné par Steele, les sarcomes indifférenciés comportent plusieurs sous-types, allant du vrai pléomorphisme au phénotype fusiforme, ou épithélioïde (43). Les sarcomes indifférenciés fusiformes peuvent ainsi comporter une proportion variable de pléomorphisme, tout en exprimant des composantes myogéniques, ce qui a fait proposer une forme frontière entre UPS et léiomyosarcomes sous le nom de sarcomes à différenciation musculaire lisse incomplète (26). La sur-représentation de formes frontières « non classiques » possédant un pléomorphisme dominant sur la composante myxoïde a également pu faire rapprocher les myxofibrosarcomes non classiques des UPS à travers différentes plateformes d'analyses génétiques (42).

Notre volonté dans la sélection d'UPS les plus indifférenciés était de comprendre les voies de signalisation derrière le phénotype « pléomorphe ». Nos résultats ont permis d'identifier un groupe immune-high, ou « inflammé », caractérisé par un fort infiltrat immunitaire, et un groupe immune-low, ou « non-inflammé », caractérisé par des remaniements chromosomiques plus complexes, notamment des pertes, et des altérations ou des sur-représentations des principales voies de signalisation en rapport avec le caractère souche,

telles que FGFR, PTEN/PI3K/Akt, Wnt/bêta-caténine, Notch, Hedgehog. Ces résultats amènent la proposition suivante : 1/ Les sarcomes pléomorphes immune-low dériveraient de cellules souches mésenchymateuses ayant amorcé une étape de différenciation intermédiaire, génomiquement instables, et conservant des mécanismes de privilège immunitaire, potentiellement par le biais des voies de signalisation responsables de la différenciation mésenchymateuse, qui appellent à être mieux caractérisées ; 2/ Les sarcomes pléomorphes immune-high pourraient être issus des cellules souches mésenchymateuses les plus indifférenciées, moins remaniées génomiquement et porteuses de caractéristiques pléomorphes reconnaissables par le système immunitaire, mais également dotées de mécanismes d'échappement/d'épuisement immunitaire qui restent à étudier.

Cette étude est la première à réaliser un portrait intégré transcriptomique, génomique, immuno-phénotypique, protéomique et radiomique d'UPS primaires non pré-traités. Cette étude améliore la compréhension biologique des UPS et amène un éclairage sur les résultats mitigés obtenus en recherche clinique avec les anti-PD-1 et anti CTLA-4 dans les sarcomes à ce jour. Ces données aident également à développer de nouvelles stratégies pour améliorer la prise en charge des patients. L'infiltration LT-CD8, et plus récemment LB, ont été corrélées à une meilleure réponse à l'immunothérapie, ainsi les patients du groupe immune-high devraient certainement être les meilleurs candidats à l'immunothérapie, sous réserve d'une meilleure compréhension des voies immuno-suppressives en place. De plus, la radiomique pourrait être un outil intéressant pour identifier ces patients dès l'IRM pré-thérapeutique. Enfin, nos résultats procurent un rationnel pour l'utilisation des inhibiteurs de FGFR et de BET, mais également d'autres voies comme les voie de la beta-caténine, Notch ou PTEN/PI3K, seuls ou en combinaison, dans les UPS immune-low.

Annexes

Tableau 1 Listes des 20 principaux gènes discriminants dans les deux groupes de MFH identifiés par Baird *et al.* Cancer Res, 2005

Table D MFH discriminator weighted gene list

Gene Name	Clone	Weight	p value
TMEPAI	transmembrane, prostate androgen induced RNA	4.894034	0.000042
NUDT4P2	nudix (nucleoside diphosphate linked moiety X)-type motif 4	4.864848	0.000042
MYO10	myosin X	4.859004	0.000042
DD96	epithelial protein up-regulated in carcinoma, membrane assoc. protein 17	4.693524	0.000075
SGCB	sarcoglycan, beta (43kD dystrophin-associated glycoprotein)	4.48026	0.000134
SPHAR	S-phase response (cyclin-related)	4.39871	0.000179
FLJ12815	Homo sapiens cDNA FLJ12815 fis, clone NT2RP2002546	4.394127	0.000179
TNC	hexabrachion (tenascin C, cytotoxin)	4.374351	0.000179
IGSF4	immunoglobulin superfamily, member 4	4.368213	0.000179
MSTP028	MSTP028 protein	4.305659	0.000179
ITPR1	inositol 1,4,5-triphosphate receptor, type 1	4.287211	0.00024
SSNA1	Sjogren's syndrome nuclear autoantigen 1	4.254941	0.00024
PIN	dynein, cytoplasmic, light polypeptide	4.241266	0.00024
GGH	gamma-glutamyl hydrolase (conjugase, folylpolyglutamyl hydrolase)	4.210993	0.00024
CHSY1	KIAA0990 protein	4.201867	0.00024
DKFZp434M229	Homo sapiens mRNA; cDNA DKFZp434M229 (from clone DKFZp434M229)	4.187824	0.000319
LOC51186	pp21 homolog	4.152696	0.000319
FLJ00012	hypothetical protein FLJ00012	4.126175	0.000319
GEM	GIP-binding protein overexpressed in skeletal muscle	4.117367	0.000319
GS3955	GS3955 protein	4.098805	0.000424

Gene Name	Clone	Weight	p value
HEM1	hematopoietic protein 1	8.045924	0
MX1	myxovirus (influenza) resistance 1	7.793819	0
DAP10	DNAX-activation protein 10	7.390443	0
DAP10	DNAX-activation protein 10	7.177107	0
PLCG2	phospholipase C, gamma 2 (phosphatidylinositol-specific)	7.104073	0
FOLR3	folate receptor 3 (gamma)	7.044833	0
TNXB	cytochrome P450, subfamily XXIA (steroid 21-hydroxylase, CAH), polypeptide 2	7.020555	0
CCL8	small inducible cytokine subfamily A (Cys-Cys), member 8	6.88491	0
LGALS9	lectin, galactoside-binding, soluble, 9 (galectin 9)	6.880873	0
IAN4L1	likely homolog of mouse immunity-associated nucleotide 4	6.804767	0
OAS1	2',5'-oligoadenylate synthetase 1 (40-46 kD)	6.672744	0
RNASE1	ribonuclease, RNase A family, 1 (pancreatic)	6.672215	0
RABL2B	RAB, member of RAS oncogene family-like 2A	6.518627	0
HML2	macrophage lectin 2 (calcium dependent)	6.356948	0.000001
VAMP8	vesicle-associated membrane protein 8 (endobrevin)	6.19316	0.000001
IFIT1	interferon-induced protein with tetratricopeptide repeats 1	6.131028	0.000001
BF	B-factor, properdin	6.130303	0.000001
CSF1R	colony stimulating factor 1 receptor	6.106375	0.000001
FLT1	fms-related tyrosine kinase 1 (VEGF/vascular permeability factor receptor)	6.078917	0.000001
TCN2	transcobalamin II; macrocytic anemia	6.073769	0.000001

Table D: The top 20 discriminating probes associated with MFH tumors visualized in figure 2b & 2c. present. The top list exhibits a preponderance of muscle related genes whereas the bottom group shows an abundance of inflammatory probes.

Références

1. Doyle LA. Sarcoma classification: an update based on the 2013 World Health Organization Classification of Tumors of Soft Tissue and Bone. *Cancer*. 2014;120(12):1763-74.
2. Fletcher CDM, Bridge JA, Hogendoorn PCW, and Mertens F. In: Kleihues P ed. *World Health Organisation Classification of Tumours*. Lyon, France: IARC Press; 2013.
3. O'Brien JE, and Stout AP. Malignant Fibrous Xanthomas. *Cancer*. 1964;17:1445-55.
4. Weiss SW, and Enzinger FM. Malignant fibrous histiocytoma: an analysis of 200 cases. *Cancer*. 1978;41(6):2250-66.
5. Kilkenny JW, 3rd, Bland KI, and Copeland EM, 3rd. Retroperitoneal sarcoma: the University of Florida experience. *J Am Coll Surg*. 1996;182(4):329-39.
6. Fletcher CD. Pleomorphic malignant fibrous histiocytoma: fact or fiction? A critical reappraisal based on 159 tumors diagnosed as pleomorphic sarcoma. *Am J Surg Pathol*. 1992;16(3):213-28.
7. Pedeutour F, Suijkerbuijk RF, Van Gaal J, Van de Klundert W, Coindre JM, Van Haelst A, Collin F, Huffermann K, and Turc-Carel C. Chromosome 12 origin in rings and giant markers in well-differentiated liposarcoma. *Cancer Genet Cytogenet*. 1993;66(2):133-4.
8. Coindre JM, Pedeutour F, and Aurias A. Well-differentiated and dedifferentiated liposarcomas. *Virchows Arch*. 2010;456(2):167-79.
9. Sirvent N, Coindre JM, Maire G, Hostein I, Keslair F, Guillou L, Ranchere-Vince D, Terrier P, and Pedeutour F. Detection of MDM2-CDK4 amplification by fluorescence in situ hybridization in 200 paraffin-embedded tumor samples: utility in diagnosing adipocytic lesions and comparison with immunohistochemistry and real-time PCR. *Am J Surg Pathol*. 2007;31(10):1476-89.
10. Binh MB, Sastre-Garau X, Guillou L, de Pinieux G, Terrier P, Lagace R, Aurias A, Hostein I, and Coindre JM. MDM2 and CDK4 immunostainings are useful adjuncts in diagnosing well-differentiated and dedifferentiated liposarcoma subtypes: a comparative analysis of 559 soft tissue neoplasms with genetic data. *Am J Surg Pathol*. 2005;29(10):1340-7.
11. Chibon F, Mariani O, Derre J, Malinge S, Coindre JM, Guillou L, Lagace R, and Aurias A. A subgroup of malignant fibrous histiocytomas is associated with genetic changes similar to those of well-differentiated liposarcomas. *Cancer Genet Cytogenet*. 2002;139(1):24-9.
12. Coindre JM, Mariani O, Chibon F, Mairal A, De Saint Aubain Somerhausen N, Favre-Guillevin E, Bui NB, Stoeckle E, Hostein I, and Aurias A. Most malignant fibrous histiocytomas developed in the retroperitoneum are dedifferentiated liposarcomas: a review of 25 cases initially diagnosed as malignant fibrous histiocytoma. *Mod Pathol*. 2003;16(3):256-62.
13. Coindre JM, Hostein I, Maire G, Derre J, Guillou L, Leroux A, Ghnassia JP, Collin F, Pedeutour F, and Aurias A. Inflammatory malignant fibrous histiocytomas and dedifferentiated liposarcomas: histological review, genomic profile, and MDM2 and CDK4 status favour a single entity. *J Pathol*. 2004;203(3):822-30.
14. Fletcher CD, Gustafson P, Rydholm A, Willen H, and Akerman M. Clinicopathologic re-evaluation of 100 malignant fibrous histiocytomas: prognostic relevance of subclassification. *J Clin Oncol*. 2001;19(12):3045-50.
15. Fletcher CD. The evolving classification of soft tissue tumours: an update based on the new WHO classification. *Histopathology*. 2006;48(1):3-12.
16. Fletcher CD. The evolving classification of soft tissue tumours - an update based on the new 2013 WHO classification. *Histopathology*. 2014;64(1):2-11.
17. Lee YF, John M, Edwards S, Clark J, Flohr P, Maillard K, Edema M, Baker L, Mangham DC, Grimer R, et al. Molecular classification of synovial sarcomas, leiomyosarcomas and malignant fibrous histiocytomas by gene expression profiling. *Br J Cancer*. 2003;88(4):510-5.
18. Baird K, Davis S, Antonescu CR, Harper UL, Walker RL, Chen Y, Glatfelter AA, Duray PH, and Meltzer PS. Gene expression profiling of human sarcomas: insights into sarcoma biology. *Cancer Res*. 2005;65(20):9226-35.
19. Nakayama R, Nemoto T, Takahashi H, Ohta T, Kawai A, Seki K, Yoshida T, Toyama Y, Ichikawa H, and Hasegawa T. Gene expression analysis of soft tissue sarcomas: characterization and reclassification of malignant fibrous histiocytoma. *Mod Pathol*. 2007;20(7):749-59.
20. Francis P, Namlos HM, Muller C, Eden P, Fernebro J, Berner JM, Bjerkehagen B, Akerman M, Bendahl PO, Isinger A, et al. Diagnostic and prognostic gene expression signatures in 177 soft tissue sarcomas: hypoxia-induced transcription profile signifies metastatic potential. *BMC Genomics*. 2007;8:73.

21. Carneiro A, Francis P, Bendahl PO, Fernebro J, Akerman M, Fletcher C, Rydholm A, Borg A, and Nilbert M. Indistinguishable genomic profiles and shared prognostic markers in undifferentiated pleomorphic sarcoma and leiomyosarcoma: different sides of a single coin? *Lab Invest.* 2009;89(6):668-75.
22. Mairal A, Terrier P, Chibon F, Sastre X, Lecesne A, and Aurias A. Loss of chromosome 13 is the most frequent genomic imbalance in malignant fibrous histiocytomas. A comparative genomic hybridization analysis of a series of 30 cases. *Cancer Genet Cytogenet.* 1999;111(2):134-8.
23. Chibon F, Mairal A, Freneaux P, Terrier P, Coindre JM, Sastre X, and Aurias A. The RB1 gene is the target of chromosome 13 deletions in malignant fibrous histiocytoma. *Cancer Res.* 2000;60(22):6339-45.
24. Larramendy ML, Gentile M, Soloneski S, Knuutila S, and Bohling T. Does comparative genomic hybridization reveal distinct differences in DNA copy number sequence patterns between leiomyosarcoma and malignant fibrous histiocytoma? *Cancer Genet Cytogenet.* 2008;187(1):1-11.
25. Gibault L, Perot G, Chibon F, Bonnin S, Lagarde P, Terrier P, Coindre JM, and Aurias A. New insights in sarcoma oncogenesis: a comprehensive analysis of a large series of 160 soft tissue sarcomas with complex genomics. *J Pathol.* 2011;223(1):64-71.
26. Perot G, Mendiboure J, Brouste V, Velasco V, Terrier P, Bonvalot S, Guillou L, Ranchere-Vince D, Aurias A, Coindre JM, et al. Smooth muscle differentiation identifies two classes of poorly differentiated pleomorphic sarcomas with distinct outcome. *Mod Pathol.* 2014;27(6):840-50.
27. Helman LJ, and Meltzer P. Mechanisms of sarcoma development. *Nat Rev Cancer.* 2003;3(9):685-94.
28. Rodriguez R, Rubio R, and Menendez P. Modeling sarcomagenesis using multipotent mesenchymal stem cells. *Cell Res.* 2012;22(1):62-77.
29. El Omar R, Beroud J, Stoltz JF, Menu P, Velot E, and Decot V. Umbilical cord mesenchymal stem cells: the new gold standard for mesenchymal stem cell-based therapies? *Tissue Eng Part B Rev.* 2014;20(5):523-44.
30. Ryan JM, Barry FP, Murphy JM, and Mahon BP. Mesenchymal stem cells avoid allogeneic rejection. *J Inflamm (Lond).* 2005;2(8).
31. Ghannam S, Pene J, Moquet-Torcy G, Jorgensen C, and Yssel H. Mesenchymal stem cells inhibit human Th17 cell differentiation and function and induce a T regulatory cell phenotype. *J Immunol.* 2010;185(1):302-12.
32. Spaggiari GM, Capobianco A, Abdelrazik H, Becchetti F, Mingari MC, and Moretta L. Mesenchymal stem cells inhibit natural killer-cell proliferation, cytotoxicity, and cytokine production: role of indoleamine 2,3-dioxygenase and prostaglandin E2. *Blood.* 2008;111(3):1327-33.
33. Tipnis S, Viswanathan C, and Majumdar AS. Immunosuppressive properties of human umbilical cord-derived mesenchymal stem cells: role of B7-H1 and IDO. *Immunol Cell Biol.* 2010;88(8):795-806.
34. Yoo KH, Jang IK, Lee MW, Kim HE, Yang MS, Eom Y, Lee JE, Kim YJ, Yang SK, Jung HL, et al. Comparison of immunomodulatory properties of mesenchymal stem cells derived from adult human tissues. *Cell Immunol.* 2009;259(2):150-6.
35. Larue L, and Bellacosa A. Epithelial-mesenchymal transition in development and cancer: role of phosphatidylinositol 3' kinase/AKT pathways. *Oncogene.* 2005;24(50):7443-54.
36. Fodde R, and Brabletz T. Wnt/beta-catenin signaling in cancer stemness and malignant behavior. *Curr Opin Cell Biol.* 2007;19(2):150-8.
37. Janssen KP, Alberici P, Fsihi H, Gaspar C, Breukel C, Franken P, Rosty C, Abal M, El Marjou F, Smits R, et al. APC and oncogenic KRAS are synergistic in enhancing Wnt signaling in intestinal tumor formation and progression. *Gastroenterology.* 2006;131(4):1096-109.
38. Matushansky I, Hernando E, Succi ND, Mills JE, Matos TA, Edgar MA, Singer S, Maki RG, and Cordon-Cardo C. Derivation of sarcomas from mesenchymal stem cells via inactivation of the Wnt pathway. *J Clin Invest.* 2007;117(11):3248-57.
39. Wang CY, Wei Q, Han I, Sato S, Ghanbari-Azarnier R, Whetstone H, Poon R, Hu J, Zheng F, Zhang P, et al. Hedgehog and Notch signaling regulate self-renewal of undifferentiated pleomorphic sarcomas. *Cancer Res.* 2012;72(4):1013-22.
40. Goldblum JR. An approach to pleomorphic sarcomas: can we subclassify, and does it matter? *Mod Pathol.* 2014;27 Suppl 1(S39-46).
41. Kelleher FC, and Viterbo A. Histologic and genetic advances in refining the diagnosis of "undifferentiated pleomorphic sarcoma". *Cancers (Basel).* 2013;5(1):218-33.
42. Cancer Genome Atlas Research Network. Electronic address edsc, and Cancer Genome Atlas Research N. Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. *Cell.* 2017;171(4):950-65 e28.

43. Steele CD, Tarabichi M, Oukrif D, Webster AP, Ye H, Fittall M, Lombard P, Martincorena I, Tarpey PS, Collord G, et al. Undifferentiated Sarcomas Develop through Distinct Evolutionary Pathways. *Cancer Cell*. 2019;35(3):441-56 e8.
44. Nakano K, and Takahashi S. Translocation-Related Sarcomas. *Int J Mol Sci*. 2018;19(12).
45. Hofvander J, Tayebwa J, Nilsson J, Magnusson L, Brosjo O, Larsson O, Vult von Steyern F, Mandahl N, Fletcher CD, and Mertens F. Recurrent PRDM10 gene fusions in undifferentiated pleomorphic sarcoma. *Clin Cancer Res*. 2015;21(4):864-9.
46. Delespaul L, Lesluyes T, Perot G, Brulard C, Lartigue L, Baud J, Lagarde P, Le Guellec S, Neuville A, Terrier P, et al. Recurrent TRIO Fusion in Nontranslocation-Related Sarcomas. *Clin Cancer Res*. 2017;23(3):857-67.
47. Zheng B, Zhang S, Cai W, Wang J, Wang T, Tang N, Shi Y, Luo X, and Yan W. Identification of Novel Fusion Transcripts in Undifferentiated Pleomorphic Sarcomas by Transcriptome Sequencing. *Cancer Genomics Proteomics*. 2019;16(5):399-408.
48. Toulmonde M, Penel N, Adam J, Chevreau C, Blay JY, Le Cesne A, Bompas E, Piperno-Neumann S, Cousin S, Grellety T, et al. Use of PD-1 Targeting, Macrophage Infiltration, and IDO Pathway Activation in Sarcomas: A Phase 2 Clinical Trial. *JAMA Oncol*. 2018;4(1):93-7.
49. van Baren N, and Van den Eynde BJ. Tumoral Immune Resistance Mediated by Enzymes That Degrade Tryptophan. *Cancer Immunol Res*. 2015;3(9):978-85.
50. Routy JP, Routy B, Graziani GM, and Mehraj V. The Kynurenine Pathway Is a Double-Edged Sword in Immune-Privileged Sites and in Cancer: Implications for Immunotherapy. *Int J Tryptophan Res*. 2016;9:67-77.
51. Tawbi HA, Burgess M, Bolejack V, Van Tine BA, Schuetz SM, Hu J, D'Angelo S, Attia S, Riedel RF, Priebat DA, et al. Pembrolizumab in advanced soft-tissue sarcoma and bone sarcoma (SARC028): a multicentre, two-cohort, single-arm, open-label, phase 2 trial. *Lancet Oncol*. 2017;18(11):1493-501.
52. Toulmonde M, and Italiano A. PD-1 inhibition in sarcoma still needs investigation. *Lancet Oncol*. 2018;19(1):e6.
53. Petitprez F, de Reyniès A, Keung EZ, Chen TW, Sun CM, Calderaro J, Jeng YM, Hsiao LP, Lacroix L, Bougouïn A, et al. B cell signatures are associated with improved sarcoma survival and immunotherapy response *Nature*. 2019(accepted).
54. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautes-Fridman C, Fridman WH, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol*. 2016;17(1):218.
55. Sautes-Fridman C, Petitprez F, Calderaro J, and Fridman WH. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat Rev Cancer*. 2019;19(6):307-25.
56. Buchan SL, Fallatah M, Thirdborough SM, Taraban VY, Rogel A, Thomas LJ, Penfold CA, He LZ, Curran MA, Keler T, et al. PD-1 Blockade and CD27 Stimulation Activate Distinct Transcriptional Programs That Synergize for CD8(+) T-Cell-Driven Antitumor Immunity. *Clin Cancer Res*. 2018;24(10):2383-94.
57. Nafia I, Toulmonde M, Bortolotto D, Chaibi A, Bodet D, Rey C, Velasco V, Larmonier C, Cerf L, Adam J, et al. IDO targeting in sarcoma: biological and clinical implications. *Front Immunol*. 2019(submitted).
58. Yamamoto T, Hatabayashi K, Arita M, Yajima N, Takenaka C, Suzuki T, Takahashi M, Oshima Y, Hara K, Kagawa K, et al. Kynurenine signaling through the aryl hydrocarbon receptor maintains the undifferentiated state of human embryonic stem cells. *Sci Signal*. 2019;12(587).
59. Murray IA, Patterson AD, and Perdew GH. Aryl hydrocarbon receptor ligands in cancer: friend and foe. *Nat Rev Cancer*. 2014;14(12):801-14.
60. Campesato LF, Budhu S, Tchaicha J, Pourpe S, Liu C, Manfredi MG, McGovern K, Wolchok JD, Merghoub T, and Ibrahim F. Interaction between Kynurenine and the AhR is an effector mechanism of tumor immunosuppression and represents a potential immunotherapy target *J Immunol*. 2018;200(1 Suppl):177.5.
61. Pinto S, Steeneck C, Albers M, Anderhub S, Birkel M, Buselic-Wölfel L, Eisenhardt G, Kremoser C, Hoffmann T, and Deuschle U. Targeting the IDO1-Kynurenine-AhR pathway for cancer immunotherapy. *Cancer Res*. 2019;79(13 Suppl):Abstract 1210.
62. Turner N, and Grose R. Fibroblast growth factor signalling: from development to cancer. *Nat Rev Cancer*. 2010;10(2):116-29.
63. Touat M, Ileana E, Postel-Vinay S, Andre F, and Soria JC. Targeting FGFR Signaling in Cancer. *Clin Cancer Res*. 2015;21(12):2684-94.
64. Perera TPS, Jovcheva E, Mevellec L, Vialard J, De Lange D, Verhulst T, Paulussen C, Van De Ven K, King P, Freyne E, et al. Discovery and Pharmacological Characterization of JNJ-42756493

- (Erdafitinib), a Functionally Selective Small-Molecule FGFR Family Inhibitor. *Mol Cancer Ther.* 2017;16(6):1010-20.
65. Nogova L, Sequist LV, Perez Garcia JM, Andre F, Delord JP, Hidalgo M, Schellens JH, Cassier PA, Camidge DR, Schuler M, et al. Evaluation of BGJ398, a Fibroblast Growth Factor Receptor 1-3 Kinase Inhibitor, in Patients With Advanced Solid Tumors Harboring Genetic Alterations in Fibroblast Growth Factor Receptors: Results of a Global Phase I, Dose-Escalation and Dose-Expansion Study. *J Clin Oncol.* 2017;35(2):157-65.
 66. Gavine PR, Mooney L, Kilgour E, Thomas AP, Al-Kadhimi K, Beck S, Rooney C, Coleman T, Baker D, Mellor MJ, et al. AZD4547: an orally bioavailable, potent, and selective inhibitor of the fibroblast growth factor receptor tyrosine kinase family. *Cancer Res.* 2012;72(8):2045-56.
 67. Helsten T, Elkin S, Arthur E, Tomson BN, Carter J, and Kurzrock R. The FGFR Landscape in Cancer: Analysis of 4,853 Tumors by Next-Generation Sequencing. *Clin Cancer Res.* 2016;22(1):259-67.
 68. Kunstlinger H, Fassunke J, Schildhaus HU, Brors B, Heydt C, Ihle MA, Mechtersheimer G, Wardelmann E, Buttner R, and Merkelbach-Bruse S. FGFR2 is overexpressed in myxoid liposarcoma and inhibition of FGFR signaling impairs tumor growth in vitro. *Oncotarget.* 2015;6(24):20215-30.
 69. Chudasama P, Renner M, Straub M, Mughal SS, Hutter B, Kosaloglu Z, Schwessinger R, Scheffler M, Alldinger I, Schimmack S, et al. Targeting Fibroblast Growth Factor Receptor 1 for Treatment of Soft-Tissue Sarcoma. *Clin Cancer Res.* 2017;23(4):962-73.
 70. Ishibe T, Nakayama T, Okamoto T, Aoyama T, Nishijo K, Shibata KR, Shima Y, Nagayama S, Katagiri T, Nakamura Y, et al. Disruption of fibroblast growth factor signal pathway inhibits the growth of synovial sarcomas: potential application of signal inhibitors to molecular target therapy. *Clin Cancer Res.* 2005;11(7):2702-12.
 71. Schuler M, Cho BC, Sayehli CM, Navarro A, Soo RA, Richly H, Cassier PA, Tai D, Penel N, Nogova L, et al. Rogaratinib in patients with advanced cancers selected by FGFR mRNA expression: a phase 1 dose-escalation and dose-expansion study. *Lancet Oncol.* 2019;20(10):1454-66.
 72. Loriot Y, Necchi A, Park SH, Garcia-Donas J, Huddart R, Burgess E, Fleming M, Rezaqadeh A, Mellado B, Varlamov S, et al. Erdafitinib in Locally Advanced or Metastatic Urothelial Carcinoma. *N Engl J Med.* 2019;381(4):338-48.
 73. Cho KB, Cho MK, Lee WY, and Kang KW. Overexpression of c-myc induces epithelial mesenchymal transition in mammary epithelial cells. *Cancer Lett.* 2010;293(2):230-9.
 74. Yin S, Cheryan VT, Xu L, Rishi AK, and Reddy KB. Myc mediates cancer stem-like cells and EMT changes in triple negative breast cancers cells. *PLoS One.* 2017;12(8):e0183578.
 75. Liu H, Ai J, Shen A, Chen Y, Wang X, Peng X, Chen H, Shen Y, Huang M, Ding J, et al. c-Myc Alteration Determines the Therapeutic Response to FGFR Inhibitors. *Clin Cancer Res.* 2017;23(4):974-84.
 76. Kuzyk A, and Mai S. c-MYC-induced genomic instability. *Cold Spring Harb Perspect Med.* 2014;4(4):a014373.
 77. Amati B, Brooks MW, Levy N, Littlewood TD, Evan GI, and Land H. Oncogenic activity of the c-Myc protein requires dimerization with Max. *Cell.* 1993;72(2):233-45.
 78. Blackwood EM, and Eisenman RN. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science.* 1991;251(4998):1211-7.
 79. Schreiber SL, and Bernstein BE. Signaling network model of chromatin. *Cell.* 2002;111(6):771-8.
 80. Delmore JE, Issa GC, Lemieux ME, Rahl PB, Shi J, Jacobs HM, Kastiris E, Gilpatrick T, Paranal RM, Qi J, et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell.* 2011;146(6):904-17.
 81. Garcia-Carpizo V, Ruiz-Llorente S, Sarmentero J, Grana-Castro O, Pisano DG, and Barrero MJ. CREBBP/EP300 bromodomains are critical to sustain the GATA1/MYC regulatory axis in proliferation. *Epigenetics Chromatin.* 2018;11(1):30.
 82. Niemann C. Controlling the stem cell niche: right time, right place, right strength. *Bioessays.* 2006;28(1):1-5.
 83. Luke JJ, Bao R, Spranger S, Sweis RF, and Gajewski T. Correlation of WNT/ β -catenin pathway activation with immune exclusion across most human cancers. *Journal of Clinical Oncology.* 2016;34(suppl; abstr 3004).
 84. Spranger S, Bao R, and Gajewski TF. Melanoma-intrinsic beta-catenin signalling prevents anti-tumour immunity. *Nature.* 2015;523(7559):231-5.
 85. Gatz ML, Lucas JE, Barry WT, Kim JW, Wang Q, Crawford MD, Datto MB, Kelley M, Mathey-Prevot B, Potti A, et al. A pathway-based classification of human breast cancer. *Proc Natl Acad Sci U S A.* 2010;107(15):6994-9.
 86. Bertucci F, Finetti P, and Birnbaum D. Wnt Signaling Inhibition Promotes Apoptosis in Sarcomas-Letter. *Mol Cancer Ther.* 2017;16(10):2324.

87. Martinez-Font E, Vogler O, Alemany R, and Obrador-Hevia A. Wnt Signaling Inhibition Promotes Apoptosis in Sarcomas-Response. *Mol Cancer Ther.* 2017;16(10):2325.
88. Martinez-Font E, Felipe-Abrio I, Calabuig-Farinas S, Ramos R, Terrasa J, Vogler O, Alemany R, Martin-Broto J, and Obrador-Hevia A. Disruption of TCF/beta-Catenin Binding Impairs Wnt Signaling and Induces Apoptosis in Soft Tissue Sarcoma Cells. *Mol Cancer Ther.* 2017;16(6):1166-76.
89. George S, Miao D, Demetri GD, Adeegbe D, Rodig SJ, Shukla S, Lipschitz M, Amin-Mansour A, Raut CP, Carter SL, et al. Loss of PTEN Is Associated with Resistance to Anti-PD-1 Checkpoint Blockade Therapy in Metastatic Uterine Leiomyosarcoma. *Immunity.* 2017;46(2):197-204.
90. Peng W, Chen JQ, Liu C, Malu S, Creasy C, Tetzlaff MT, Xu C, McKenzie JA, Zhang C, Liang X, et al. Loss of PTEN Promotes Resistance to T Cell-Mediated Immunotherapy. *Cancer Discov.* 2016;6(2):202-16.
91. Suehara Y, Kondo T, Fujii K, Hasegawa T, Kawai A, Seki K, Beppu Y, Nishimura T, Kurosawa H, and Hirohashi S. Proteomic signatures corresponding to histological classification and grading of soft-tissue sarcomas. *Proteomics.* 2006;6(15):4402-9.