



HAL
open science

Caractérisation des réseaux d'interactions protéiques associés aux mutations oncogéniques principales retrouvées dans le cancer du poumon non à petites cellules

Benoît Beganton

► **To cite this version:**

Benoît Beganton. Caractérisation des réseaux d'interactions protéiques associés aux mutations oncogéniques principales retrouvées dans le cancer du poumon non à petites cellules. Médecine humaine et pathologie. Université Montpellier, 2017. Français. NNT : 2017MONTT157 . tel-01735769

HAL Id: tel-01735769

<https://theses.hal.science/tel-01735769>

Submitted on 16 Mar 2018

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.

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Biologie Santé

École doctorale : Sciences Chimiques et Biologiques pour la Santé – CBS2

Unité de recherche : INSERM U1194 – Institut de Recherche en Cancérologie de Montpellier (IRCM)

**Caractérisation des réseaux d'interactions protéiques
associés aux mutations oncogéniques principales retrouvées
dans le cancer du poumon non à petites cellules.**

**Présentée par Benoît BEGANTON
Le 29 Novembre 2017**

Sous la direction de Jérôme SOLASSOL

Devant le jury composé de

M. Patrick CALLIER, MD-Ph.D, INSERM U866, CHU de Dijon

M. Etienne COYAUD, Ph.D SA, Princess Margaret Cancer Center, Toronto

M. Brian RAUGHT, Ph.D PI, University of Toronto

M. Xavier QUANTIN, MD-Ph.D, CHU de Montpellier

M. Jérôme SOLASSOL, MD-Ph.D, INSERM U1194, CHU de Montpellier

Rapporteur

Rapporteur

Examineur

Examineur

Directeur de thèse



**UNIVERSITÉ
DE MONTPELLIER**

Remerciements

Je remercie les membres du jury d'avoir accepté d'évaluer mon travail, et particulièrement le Pr. Patrick Cailler et le Dr. Etienne Coyaud pour leur investissement en tant que rapporteurs de cette thèse. Merci également aux examinateurs, Dr. Brian Raught et Pr. Xavier Quantin pour le temps que vous consacrez à juger à ce travail. C'est un honneur pour moi de vous présenter ce travail qui a occupé ces trois dernières années de ma vie. J'espère que vous l'accueillerez avec le même intérêt que j'y ai porté.

Je souhaite remercier très sincèrement mon directeur de thèse, le Pr. Jérôme Solassol pour m'avoir fait confiance et permis de réaliser ce doctorat au sein de son laboratoire. L'aide précieuse qu'il m'a apporté, la pertinence de ses remarques m'ont accompagné tout au long de ce projet de recherche, et je lui en suis profondément reconnaissant. Travailler avec quelqu'un d'aussi accessible et bienveillant est plutôt rare, alors je suis heureux d'avoir eu l'opportunité de vous rencontrer.

Brian this part is for you! I guess that you are wondering what I am saying with all this French, but I would like you to know that I am truly grateful for your welcome in your lab. It was a deep honor for me to work with you. I enjoyed your vision of research and how you manage your lab. It was a great experience to me and I will keep this in memory. I know that we still have a lot of work and data to analyse, and I guess we will continue this great collaboration for a long time, and find data that could impact cancer treatment.

Estelle, merci pour tout ce que tu as fait pour moi. D'avoir pris le temps de m'encadrer, et tous ces moments et discussions que nous avons partagés. Même de l'autre côté de l'atlantique, j'avais le sentiment d'avoir comme une famille sur qui je peux compter à n'importe quel moment et ça permet d'avancer plus posément. Tous ces souvenirs resteront, c'est sûr. Merci aussi aux autres membres et anciens membres de l'équipe, et plus particulièrement à Audrey et Jeremy pour avoir pris le temps de répondre à mes 1000 questions et pour tous les bons moments que nous avons passés. Merci aussi à Paul et Valentina.

Mes remerciements vont également à Alain Mangé. Alain, je veux saluer ta disponibilité sans failles, ta gentillesse et te remercie profondément de m'avoir aidé lorsque j'en avais besoin. Merci aussi pour tes qualités humaines qui ont fait que je pouvais compter sur toi. Je tiens également à remercier tous les membres de l'équipe, Peter Coopman, Gilles Freiss et Carole Puech pour leur disponibilité et les discussions qui m'ont permis d'avancer. Merci à ma super stagiaire, Eulalie Corre. Eulalie, je ne t'ai pas laissé beaucoup de répit mais je suis reconnaissant de ton investissement et de ton implication. Tu as des qualités scientifiques et humaines qui, je pense, te permettront de réussir là où tu veux aller, et c'est ce qui compte.

Je tiens à remercier le Dr. Muriel Le Romancer et Julien Jacquemetton. Merci Muriel pour ton accueil Lyonnais si chaleureux et de m'avoir permis d'avancer dans mon travail. Julien, merci d'avoir partagé ton savoir-faire et d'avoir répondu à mes nombreuses questions.

J'adresse mes sincères remerciements à Julie Vendrell. « Impossible n'est pas Vendrell », c'est par cette phrase que j'ai appris à te connaître et ce n'était qu'un début ! Merci pour ton dynamisme, ta disponibilité, ta générosité débordante. Tu fais partie des personnes qui vont de l'avant et donne envie d'en faire de même. À tes côtés, et souvent à distance, j'ai appris des méthodes de travail et de réflexion qui permettront d'envisager les projets à venir plus sereinement. Merci aussi pour les Déferlantes, la fête des pêcheurs, les BBQs, j'en garde d'excellents souvenirs qui d'une certaine façon ont contribué à la réalisation de cette thèse.

Merci au personnel du laboratoire de Biopath et plus particulièrement à Marion Larrieux, Paul Vilquin et Charles Van Goethem. Merci de m'avoir accueilli et pris le temps de répondre à mes questions, de m'avoir aiguillé sur des aléas techniques. Merci aussi pour les à-côtés du labo, nous avons quelques soirées à rattraper et à prévoir...

Plus largement, mes remerciements vont à l'ensemble des membres de l'IRCM, qui chacun ont contribué à entretenir cette ambiance si appréciable. Merci aux autres thésards et post-docs que j'ai eu plaisir à côtoyer, et ce, pas uniquement dans le cadre du labo. Je crois qu'il existe une certaine entraide, ou du moins une compréhension mutuelle, qui porte vers l'avant. Parmi eux, je me dois de remercier plus particulièrement le gang des libanais avec Toufic, Amanda, Émile, Mona, Rita, Zahra, Olga, Moukbel, Diala. Même si je ne pourrai jamais être naturalisé (wallah c'est pas normal), j'ai apprécié découvrir votre culture et le Liban est sur ma liste de destinations :) Merci aussi à Kalil, Hélène, Sara, Leïla, Nabiya, Diane, il y'a du monde et je m'excuse de ne pouvoir tous vous citer.

Il y'a aussi des amis qui changent une vie, sans trop savoir l'expliquer. Merci à Mathieu, Richard et Guillaume pour votre soutien constant tout au long de ces années. Je crois que vous me connaissez mieux que personne, et vous côtoyer ouvre l'esprit. Regardez attentivement et vous verrez que vous êtes aussi dans ce travail.

Enfin, merci à mes parents à qui je dois à peu près tout. Merci de m'avoir permis de toujours faire ce que je voulais, d'aller là où je voulais et de m'avoir soutenu. C'est une chance, et en fait je ne sais pas comment vous remercier de tout ce que vous m'avez apporté et depuis toujours. Yann et Alice, merci à vous aussi pour votre soutien permanent. Je pense aussi à mes grands-parents qui m'encouragent de près ou de bien plus loin.

Je souhaite dédier ce travail à mes parents.

« Comme quoi il ne suffit pas de savoir. Encore faut-il savoir savoir.

Cela dit, « sait-on jamais » ? Là résiderait même notre espérance : on ne sait jamais. »

Lucien Jerphagnon

Résumé

Le cancer du poumon est la première cause de mortalité liée au cancer en France et dans le monde. Il s'agit d'un cancer de mauvais pronostic, diagnostiqué généralement aux stades tardifs III ou IV et avec une survie à 5 ans faible, respectivement de 6 et 1%. Ce cancer comprend différents types histologiques, parmi lesquels les adénocarcinomes sont les plus fréquents et comptent pour 40% des diagnostics. L'analyse génétique des tumeurs des patients en stade IV révèle des altérations récurrentes, et généralement mutuellement exclusives, au niveau des gènes *KRAS*, *EGFR*, *BRAF* et *ALK*. L'identification de ces altérations génétiques permet dans le cadre de tests compagnons de décider de l'éligibilité des patients aux thérapies ciblées lorsqu'elles sont disponibles.

Bien que ces thérapies ciblées représentent une véritable révolution dans la prise en charge des patients, une majorité de patients ne peuvent bénéficier de ces traitements car ils ne présentent pas au niveau tumoral de mutations activatrices actionnables (absence d'altération de l'*EGFR*, *BRAF* ou *ALK*, présence d'une mutation *KRAS* ...). De plus, dans l'éventualité où ils peuvent être traités par une molécule orale, le bénéfice observé reste malheureusement de courte durée et tous finiront à terme par échapper au traitement. Certains mécanismes de résistances acquises sont par ailleurs particulièrement bien décrits comme la mutation secondaire *EGFR* T790M qui affecte directement la protéine traitée par la thérapie ciblée, et compte pour environ la moitié des cas de résistances.

Ainsi, afin de mieux comprendre quels sont les mécanismes moléculaires associés aux phénomènes de résistances thérapeutiques observés chez les patients, et dans l'objectif de proposer de nouvelles voies thérapeutiques pour les patients non éligibles aux thérapies ciblées, j'ai étudié l'impact des mutations canoniques du cancer du poumon sur les réseaux d'interactions protéiques. Plus particulièrement, j'ai focalisé mon travail sur la caractérisation des réseaux d'interactions des mutations activatrices de l'*EGFR*, *KRAS*, *BRAF* et des translocations *ALK*. Les protéines de la famille RAS (*HRAS*, *NRAS*, *KRAS*) étant mutées dans plus de 25% des cancers, et étant donné le taux plus élevé de mutations de *KRAS* *versus* *HRAS* et *NRAS*, j'ai également mené une étude comparative d'interactomes entre ces trois isoformes afin de mettre en évidence les interactions protéiques spécifiques à l'isoforme *KRAS*.

Au cours de ma thèse, j'ai montré que les réseaux d'interactions identifiés par la technique du BioID (proximity-dependent biotinylation identification) sont bien plus riches que ceux obtenus par des techniques plus conventionnelles de purification par affinité, et qu'ils permettent de mettre en évidence des interacteurs spécifiquement dérégulés par l'apparition de mutations activatrices. Par ailleurs, les modèles cellulaires HEK293 (cellules humaines embryonnaires de rein) et BEAS2B (cellules pulmonaires d'origine non-cancéreuse) ont montré un fort recouvrement des interacteurs identifiés, ce qui suggère que la stratégie employée est pertinente pour identifier des interacteurs spécifiques de mutations. Ce travail de thèse a permis d'identifier une série de plusieurs interacteurs spécifiques des formes mutantes de *KRAS*, d'*EGFR*, de *BRAF*, de *NRAS* et de la fusion *EML4-ALK*. Treize interacteurs spécifiques de la forme mutée de *KRAS* ont été validés fonctionnellement dans des modèles cellulaires de cancers du poumon. Enfin, en utilisant les données de BioID, j'ai également proposé un modèle préliminaire de résistance de l'*EGFR* aux thérapies ciblées. Ce modèle suggère que *CBL* et *IGF2R* pourraient interagir avec l'*EGFR* et être impliquées dans l'échappement thérapeutique observé.

Ce travail de thèse ouvre ainsi de nouvelles perspectives quant à la détermination des mécanismes de résistance spécifiques aux thérapies ciblées, et à l'identification de nouvelles cibles thérapeutiques chez les patients qui présentent une mutation activatrice ou une translocation.

Mots-clés : Cancer du poumon, thérapies ciblées, résistance, interactome, signalisation cellulaire

Abstract

Lung cancer is the leading cause of cancer-related death in France and in the world. It is a cancer of poor prognosis, diagnosed at the late stage III or IV, with a low 5-years survival of 6% and 1%, respectively. This cancer encompasses several histological types, and among them adenocarcinoma are the most frequent and account for 40% of the diagnosis. Genetic analysis of stage IV tumors reveals redundant alterations, and generally exclusives from each other's, of *KRAS*, *EGFR*, *BRAF* and *ALK* genes. The identification of these genetic alterations enables, within companion test, to make eligible patients for targeted therapies when molecules are available.

Even though these targeted therapies represent a true revolution in patient's care, the majority of these patients cannot benefit from such due to the absence of targetable activating mutations (e.g. no *EGFR*, *BRAF* or *ALK* alterations, presence of *KRAS* mutation ...). Additionally, even if they can be treated by an oral molecule, the observed benefit unfortunately remains transient and all these patients will eventually relapse. Some acquired resistance mechanisms are particularly well described, such is the case of the *EGFR* T790M secondary mutation directly affecting the protein targeted by the therapy, and accounting for half of resistance events.

To better understand the molecular mechanisms associated with the resistance events observed in patients, and to propose new therapeutics for patient non-eligible for targeted therapies, I studied the impact of the canonical lung cancer mutations on protein interaction networks. I thus focused my work on the network characterization of *EGFR*, *KRAS*, *BRAF* activating mutations and *ALK* translocations. Considering that mutations of proteins from the RAS family (*HRAS*, *NRAS* and *KRAS*) account for more than 25% of cancers, and given the higher rate of *KRAS* mutations *versus* *HRAS* and *NRAS*, I have also conducted a comparative interactomic study between these three isoforms to uncover *KRAS* interactome specificities.

During my thesis, I showed that the protein networks characterized using BioID (proximity-dependent biotinylation identification) are much richer compared to those identified with the more conventional technique of affinity purification, and that they enable efficient identification of interactors specifically deregulated upon activating mutation. Additionally, the HEK293 cell model (Human Embryonic Kidney) and BEAS2B cell model (non-cancerous lung cell line) showed a high overlapping degree of the interactors identified, suggesting that the strategy used is relevant to identify interactome modifications related to specific lung cancer genetic alterations. This thesis allowed the identification of several interactors specific to the mutants of *KRAS*, *EGFR*, *BRAF*, *NRAS* and *EML4-ALK* fusion. Thirteen interactors specific to the mutated-*KRAS* have been functionally validated in lung-cancer cell lines models. Finally, using BioID data I have been able to propose a preliminary model of *EGFR* resistance to targeted therapies. This model suggests that *CBL* and *IGF2R* could partner to *EGFR* and play a role in the therapeutic escape.

Overall, this thesis opens new perspectives to determine mechanisms of resistance to targeted therapies, and to identify new therapeutic targets for patients harboring an activating mutation or translocation.

Key-words: Lung cancer, targeted therapies, resistance, interactome, cell signaling

Liste des abréviations

14-3-3 (ou YWHA): Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein

ALK: Anaplastic Lymphoma receptor tyrosine Kinase

AP: Purification par Affinité

APEX: Engineered Ascorbate Peroxidase

BAD: BCL2 associated Agonist of cell Death

BFRD: FDR Bayésien

BIND: Biomolecular Interaction Network Database

BioGRID: Biological General Repository for Interaction Datasets

BioID: Proximity-dependent Biotinylation Identification

BirA: Bifunctional ligase/repressor BirA

BirA*: BirA R118G

BRAF: B-Raf proto-oncogene, serine/threonine kinase

BRS: BRAF-Specific Region

CBL: CBL Proto-Oncogene

CDC42: Cell Division Cycle 42

CDK1: Cyclin Dependent Kinase 1

CDKN2A: Cyclin Dependent Kinase Inhibitor 2A

CNND1: Cyclin D1

CRD: Domaine Riche en Cystéines

CREB: cAMP Response Element-Binding protein

DH/PH: Dbl homology/Pleckstrin homology

ECE1: Endothelin Converting Enzyme 1

EGF: Epidermal Growth Factor

EGFR: Epidermal Growth Factor Receptor

eIF4F: Eukaryotique Initiation Factor 4 F

ELK1: ELK1, ETS transcription factor

EML4: Echinoderm Microtubule Associated Protein Like 4

ERBB2 (ou HER2): Erb-B2 Receptor Tyrosine Kinase 2

FDA: Food and Drug Administration

FDR: False Discovery Rate

FGFR: Fibroblast Growth Factor Receptor

FRET: Förster Resonance Energy Transfer

GAP: GTPase Activating Protein

GAPDH: Glyceraldehyde-3-phosphate déshydrogénase

GEF: Guanine nucleotide Exchange Factor

GFP: Green Fluorescent Protein

GO: Gene Ontology

GRB2: Growth Factor Receptor Bound Protein 2

HA: Human influenza hémagglutinine

HELP: Hydrophobic EML protein

HPRD: Human Protein Reference Database

HRAS: HRAS Proto-Oncogene, GTPase

HRNP: Heterogeneous Nuclear Ribonucleoprotein

HVR: Région Hypervariable

IGF2R: Insulin Like Growth Factor 2 Receptor

IL1RAP: Interleukin 1 Receptor Accessory Protein

IP: Immunoprecipitation

IPP: Interactions Protéine-Protéine

ITK: Inhibiteur de tyrosine kinase

KEGG: Kyoto Encyclopedia of Genes and Genomes

KRAS: KRAS proto-oncogene, GTPase

LASP1: LIM and SH3 Protein 1

LC: Chromatographie Liquide

LDHB: Lactate Déshydrogénase

MAP2K1 (ou MEK1): Mitogen-Activated Protein Kinase Kinase 1

MAP2K2 (ou MEK2): Mitogen-Activated Protein Kinase Kinase 2

MAP3K8: Mitogen-Activated Protein Kinase Kinase Kinase 8

MAPK: Mitogen-Activated Protein Kinases

MAPK1 (ou ERK2): Mitogen-Activated Protein Kinase 1

MAPK3 (ou ERK1): Mitogen-Activated Protein Kinase 3

MAPKK: MAP Kinase Kinase

MAPKKK: MAP Kinase Kinase Kinase

MET: MET Proto-Oncogene, Receptor Tyrosine Kinase,

MK: MAPK-activated protein Kinases

MNK: MAPK interacting Kinases

MRPL: Mitochondrial Ribosomal Protein L

MS: Spectrométrie de Masse

MSK: Mitogen- and Stress-activated Kinases

MTOR: Mechanistic Target of Rapamycin

MYC: MYC proto-oncogene, bHLH transcription factor

MYO1E: Myosin IE

NF1: Neurofibromin 1

NRAS: NRAS Proto-Oncogene, GTPase

NSCLC: Non-Small Cell Lung Cancer

NTRK1: Neurotrophic Receptor Tyrosine Kinase 1

PD1: Programmed cell Death protein 1

PDGFR: Platelet-Derived Growth Factor Receptor.

PDHB: Pyruvate Dehydrogenase E1 Beta subunit

PDL1: Programmed Death-Ligand 1

PGD: Phosphogluconate Dehydrogenase

PH: Domaine Pleckstrin Homology

PI3KCA: Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha

PKM: Pyruvate Kinase, Muscle

PLA: Proximity Ligation Assay

PLC γ : Phospholipase C Gamma

PLC ϵ : Phospholipase C Epsilon

PSMD1: Proteasome 26S Subunit, Non-ATPase 1

RA: RAS Association domain

RAB27A: RAB27A, Member RAS Oncogene Family

RAC1: Rac family small GTPase 1

RAF1: Raf-1 Proto-Oncogene, Serine/Threonine Kinase

RALGDS: Ral Guanine Nucleotide Dissociation Stimulator

RAPGEF2: Rap Guanine Nucleotide Exchange Factor 2

RAPGEF6: Rap Guanine Nucleotide Exchange Factor 6

RBD: RAS-binding domain

RBL: RBD-like

RET: Förster Resonance Energy Transfer

RIN1: RAS and Rab Interactor 1

ROS1: ROS Proto-Oncogene 1, Receptor Tyrosine Kinase

RPS: Ribosomal Protein S

RPS6KA1: Ribosomal protein S6 kinase, 90kDa, polypeptide 1

RSK: 90-kDa Ribosomal S6 Kinases

RTK: Récepteur à activité Tyrosine Kinase

SAINT: Significance Analysis of INTeractome

SCLC: Small Cell Lung Cancer

SH2: Src Homology 2

SH3: Src Homology 3

SLC2A3 (ou GLUT3): Solute Carrier Family 2 Member 3

SOS: SOS RAS/Rac Guanine Nucleotide Exchange Factor 1

SPRED2: Sprouty Related EVH1 Domain Containing 2

SRC: SRC Proto-Oncogene, Non-Receptor Tyrosine Kinase

SRF: Serum Response Factor

STAT5: Signal Transducer And Activator of Transcription 5

TGF- α : Transforming Growth Factor alpha

TIAM1: T-Cell Lymphoma Invasion and Metastasis 1

TP53: Tumor Protein P53

TRAF7: TNF Receptor Associated Factor 7

TSC2: Tuberous Sclerosis 2

USP32: Ubiquitin Specific Peptidase 32

WES: Whole Exome Sequencing

Y2H: Yeast two Hybrid

YFP: Yellow Fluorescent Protein

Liste des figures

Figure 1. Classement de l'incidence et de la mortalité des tumeurs solides par sexe en France métropolitaine.....	3
Figure 2. Classification histologique et incidences respectives des cancers du poumon.	5
Figure 3. Évolution de la caractérisation du profil mutationnel des adénocarcinomes pulmonaires....	11
Figure 4. Description générale de la voie des MAP Kinases.	12
Figure 5. Représentation de l'EGFR et de ses régions et domaines.	14
Figure 6. Représentation des sites de phosphorylation de l'EGFR et des protéines activées correspondantes.	16
Figure 7. Représentation des mutations ponctuelles et indels identifiés dans les exons 18 à 21 de l'EGFR ainsi que de leurs sensibilités respectives aux ITKs.....	17
Figure 8. Représentation du mécanisme d'inhibition compétitive des ITKs de l'EGFR et de la position des mutations les plus fréquemment identifiées.	18
Figure 9. Mécanismes de résistances aux ITKs ciblant l'EGFR.	19
Figure 10. Structure des isoformes RAS.	21
Figure 11. Boucle de régulation des protéines RAS.....	23
Figure 12. Mécanisme d'activation de RAS.....	24
Figure 13. Interacteurs et voies de signalisation activés par l'activation de RAS.....	25
Figure 14. Structure des isoformes RAF.	30
Figure 15. Cycle d'activation des protéines RAF.....	32
Figure 16. Variants de translocations EML4-ALK et incidences respectives.	35
Figure 17. Approche multidimensionnelle du protéome.	39
Figure 18. Principe généraux des purifications d'affinité.	41
Figure 19. Principe général du Yeast-2-Hybrid.....	42
Figure 20. Principe général du BioID.....	44
Figure 21. Principe général du split-BioID.	45
Figure 22. Principe général de l'APEX.....	46
Figure 23. Principe général du PLA.	47
Figure 24. Principe général du FRET.....	48
Figure 25. Méthodologie d'analyse des réseaux d'interactions protéiques.	56
Figure 26. Plan expérimental du BioID.....	56
Figure 27. Validation de l'expression du transgène KRAS WT et G12C.	113
Figure 28. Comparaison des interacteurs de KRAS identifiés à partir des modèles cellulaires HEK293 et BEAS2B..	114
Figure 29. Analyse des localisations et fonctions des interacteurs de KRAS WT.	116

Figure 30. Complexes protéiques enrichis parmi les interacteurs de KRAS WT.....	117
Figure 31. Complexes protéiques significatifs parmi les interacteurs de KRAS WT et G12C.....	118
Figure 32. Dérégulation d'interactions par la mutation KRAS G12C.....	120
Figure 33. Analyse des localisations et fonctions des interacteurs dérégulés de KRAS.....	120
Figure 34. Sélection des interacteurs dérégulés et partagés entre les modèles cellulaires HEK293 et BEAS2B.....	122
Figure 35. Validation de l'extinction des cibles sélectionnées.....	125
Figure 36. Effet sur la viabilité cellulaire de l'extinction d'expression des cibles sélectionnées.....	125
Figure 37. Effet sur la viabilité cellulaire de l'extinction d'expression des 7 dernières cibles sélectionnées.....	126
Figure 38. Effet de la délétion de GLUT3 sur la viabilité des cellules HPAFII.....	127
Figure 39. Validation de l'expression du transgène NRAS WT et Q61K dans les cellules HEK293.....	130
Figure 40. Comparaison des interacteurs de NRAS WT et NRAS Q61K identifiés dans les cellules HEK293.....	131
Figure 41. Identification des interacteurs communs dérégulés par les mutations NRAS Q61K et KRAS G12C.....	132
Figure 42. Réseau d'interactions fonctionnelles des protéines dérégulées par les mutants NRAS Q61K et KRAS G12C.....	133
Figure 43. Identification des interacteurs dérégulés de l'EGFR par les mutations primo-activatrices (L858R ou Del.19) ou de résistance (T790M).....	138
Figure 44. Identification des interacteurs affectés par la mutation EGFR T790M.....	139
Figure 45. Identification des réseaux fonctionnels et interacteurs les plus spécifiques aux mutations primaires et de résistance, ou nouvellement altérés par la mutation de résistance.....	141
Figure 46. Réseau fonctionnel d'interaction des interacteurs dérégulés par les mutations primaires et lors de l'acquisition de la mutation de résistance.....	142
Figure 47. Analyse des composants cellulaires, processus biologiques et voies cellulaires des interacteurs de BRAF WT.....	145
Figure 48. Identification des interacteurs dérégulés entre les formes BRAF WT et BRAF V600E.....	147
Figure 49. Réseau d'interaction fonctionnel des interacteurs préférentiels de BRAF V600E.....	148
Figure 50. Identification des procédés biologiques, composants cellulaires, voies de signalisation et domaines protéiques enrichis par EML4-ALK v3a.....	151
Figure 51. Réseau fonctionnel des interacteurs d'EML4-ALK v3a.....	152

Liste des tableaux

Tableau 1. Classification TNM des NSCLC.....	5
Tableau 2. Résumé des pourcentages de diagnostic et taux de survie à 1 et 5 ans selon le type de cancer et le stade.....	6
Tableau 3. Résumé des cibles recherchées par séquençage et thérapies disponibles associées ou en cours d'essai clinique.....	9
Tableau 4. Tableau récapitulatif des clonages réalisés et des modèles cellulaires correspondants.....	58
Tableau 5. Exemple de représentation des données de BioID.	62
Tableau 6. Tableau récapitulatif des lignées cellulaire utilisées pour la validation fonctionnelle des interacteurs dérégulés de KRAS.....	66
Tableau 7. Interacteurs exclusifs de KRAS WT ou G12C.....	121
Tableau 8. Tableau résumant les interacteurs attribués aux catégories KEGG Glycolysis/Gluconeogenesis et Oxidative phosphorylation.	146
Tableau 9. Sélection des catégories KEGG et InterPro enrichies.	151

Sommaire

Résumé	I
Abstract	III
Liste des abbréviations	V
Liste des figures.....	XI
Liste des tableaux	XIII
Introduction	1
A. Le cancer du poumon	3
I. Épidémiologie	3
II. Diagnostic	4
III. Classification histologique et stades.....	4
IV. Traitements.....	6
B. Biomarqueurs	11
I. La voie des ‘Mitogen-Activated Protein Kinases’ (MAPK).....	12
II. Le récepteur aux facteurs de croissance EGFR.....	14
III. Les protéines de la famille RAS.....	21
IV. La protéine BRAF	30
V. Les gènes de fusion impliquant <i>ALK</i>	34
VI. Autres cibles thérapeutiques.....	37
C. Étude des interactions protéiques	39
I. Dynamique des interactions protéiques.....	39
II. Identification des interactions protéine-protéine <i>in vitro</i>	41
III. Identification des interactions protéine-protéine <i>in vivo</i>	43
IV. Méthode de validation des interactions	46
Objectifs de thèse	49
Résultats	53
A. Introduction	55

I.	Approche expérimentale	55
II.	Expériences de BioID.....	56
III.	Traitement des réseaux d'interactions protéiques	62
IV.	Analyses <i>in vitro</i> ou <i>in vivo</i>	64
B.	Analyse de l'interactome de RAS	67
I.	Introduction de l'étude	67
II.	Manuscrit.....	67
III.	Commentaire	109
IV.	Discussion	110
C.	Réseau d'interaction de KRAS : application au NSCLC	113
I.	Introduction de l'étude	113
II.	Résultats	113
III.	Discussion	128
D.	Autres réseaux d'interactions protéiques.....	130
I.	Réseau d'interaction de NRAS	130
II.	Réseau d'interaction de l'EGFR : application au NSCLC	137
III.	Réseau d'interaction de BRAF : application au NSCLC	145
IV.	Réseau d'interaction de EML4-ALK : application au NSCLC.....	150
	Discussion	155
	Publications	159
	Annexes.....	197
	Bibliographie.....	231
	Abstract	251
	Résumé.....	252

Introduction

A. Le cancer du poumon

I. Épidémiologie

Le cancer du poumon est aujourd'hui la première cause de décès par cancer en France et dans le monde avec environ 30 000 et 1,6 millions de victimes par an, respectivement. Il s'agit également d'un des cancers le plus diagnostiqué chez les hommes et femmes avec chaque année en France environ 45 000 nouveaux cas, en 3^e position après le cancer du sein et le cancer de la prostate (Figure 1). Il y a 2 fois plus d'hommes que de femmes touchées mais, alors que l'incidence est stable chez les hommes, elle est en nette progression chez les femmes (+ 5,4% ces 7 dernières années). Au moment du diagnostic, le patient type est un fumeur, âgé en moyenne de 65 ans avec un taux de survie à 5 ans de 17% (INCa, 2017). Ce faible taux de survie amène à considérer ce cancer comme étant de mauvais pronostic.

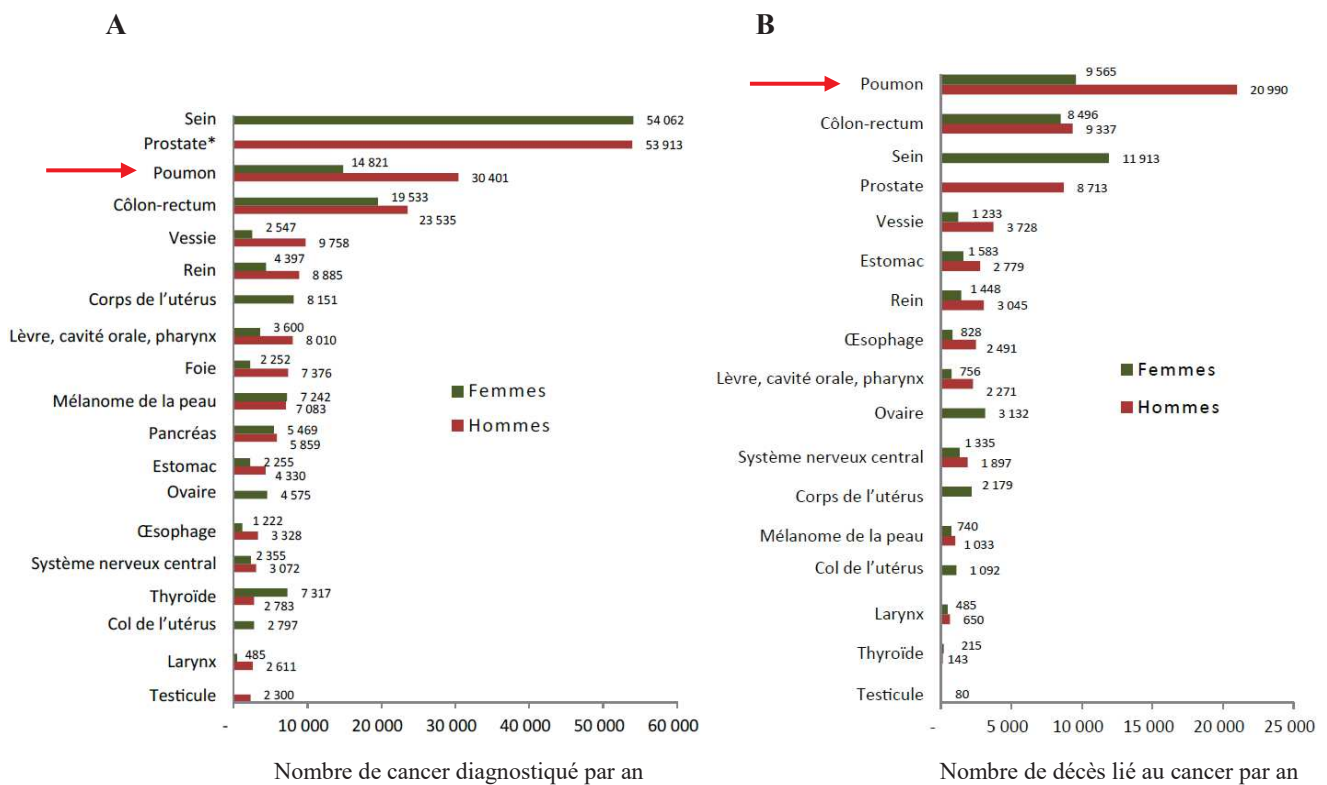


Figure 1. Classement de l'incidence et de la mortalité des tumeurs solides par sexe en France métropolitaine. (A) Incidence des tumeurs solides. Le cancer du poumon est le premier cancer touchant à la fois les hommes et les femmes. **(B)** Mortalité des tumeurs solides. Le cancer du poumon est la première cause de décès liée au cancer. Ce classement est représentatif des incidences et mortalités dans le monde. La flèche rouge indique la position du cancer du poumon dans ces classements. Source INCa, 2017.

Le tabac est le premier facteur de risque du cancer du poumon. En effet, 89% des décès dus aux cancers broncho-pulmonaires sont liés à la consommation de tabac (Bonaldi et al., 2016). Quel que soit l'âge, l'arrêt de la cigarette augmente l'espérance de vie, et ce d'autant plus que l'arrêt est précoce (Doll et al., 2004). L'arrêt du tabac est également bénéfique après diagnostic du cancer du poumon, en augmentant la survie à 5 ans, en réduisant les risques de récurrence et enfin en limitant le développement d'autres formes de cancers (Parsons et al., 2010).

Il a été démontré également que la qualité de l'environnement notamment à travers l'exposition à la pollution atmosphérique impacte le développement de ce cancer (INSERM, 2008). La présence dans l'atmosphère de carcinogènes provenant de combustibles d'origine anthropique et l'exposition au radon, gaz radioactif naturel provenant de la dégradation de l'uranium de la croûte terrestre, sont deux facteurs significativement impliqués dans la survenue du cancer du poumon (Sun et al., 2007).

II. Diagnostic

Le cancer du poumon est difficile à diagnostiquer à un stade précoce car les patients sont le plus souvent asymptomatiques (Tyczynski et al., 2003). Les motifs de consultation sont principalement une toux persistante, des dyspnées ou la présence de sang dans les crachats (hémoptysies). Ces symptômes non-exclusifs au cancer du poumon surviennent alors que la maladie s'est déjà développée et a disséminé localement ou à échelle systémique. Cette absence de signes précliniques explique la faible survie à 5 ans des patients lors du diagnostic.

Le diagnostic des cancers broncho-pulmonaires consiste en deux examens :

- a) Imagerie (radiographie ou scanner) : cet examen permet de détecter le cancer, en précisant la localisation, la taille et l'étendue des lésions.
- b) Biopsie solide : le prélèvement d'un échantillon de la tumeur permet de réaliser un examen anatomopathologique pour confirmer le diagnostic, puis établir le type histologique de la tumeur.

III. Classification histologique et stades

Les caractéristiques morphologiques et immunohistochimiques du prélèvement biologique permettent de confirmer le diagnostic puis définir le type histologique de la tumeur. Le cancer du poumon est divisé en cancer du poumon à petites cellules (ou Small Cell Lung Cancer, SCLC) et cancer du poumon non à petites cellules (ou Non-Small Cell Lung Cancer, NSCLC). Les NSCLC sont les plus fréquents et comptent pour environ 80% des cas. Celui-ci comprend trois sous-types histologiques : les adénocarcinomes (40% des cas), les carcinomes épidermoïdes (25%) et les carcinomes à grandes cellules (15%). Les 5% restants comprennent des sous-types histologiques variés (Figure 2).

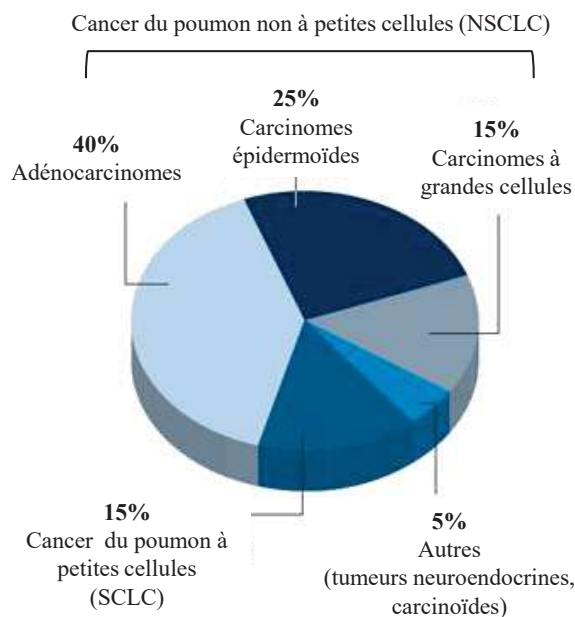


Figure 2. Classification histologique et incidences respectives des cancers du poumon.

Considérant la taille, l'étendu des lésions et le type de cancer, le stade du cancer peut être établi. Pour les NSCLC, le stade de la tumeur est évalué à partir de la classification TNM. Cette classification prend en compte trois paramètres : la taille et la propagation de la tumeur sur le site primitif (T), la propagation aux ganglions lymphatiques environnants (N) et la présence ou non de métastases (M). Chacun de ces paramètres sont cotés et leur association permet de définir le stade de la maladie. Ceux-ci s'échelonnent de I à IV, où le stade IV décrit la forme la plus avancée métastatique (Tableau 1).

Tableau 1. Classification TNM des NSCLC. Les stades I, II et III sont répartis en sous-stades A et B selon la classification TNM. La taille de la tumeur s'évalue de T1 à T4 selon son étendue. La propagation aux ganglions lymphatiques s'évalue de N0 à N3. Les métastases s'évaluent selon leur présence (M1) ou non (M0). Source Medical School of California Regents University.

Stade	Classification TNM associée
I	A: T1 N0 M0
	B: T2 N0 M0
II	A: T1 N1 M0
	B: T2 N1 M0 ou T3 N0 M3
III	A: T3 N0 M0 ou T1-3 N2 M0
	B: T1-4 N3 M0 ou T4 N0-3 M0
IV	Tout M1

Les SCLC ne sont en revanche pas inclus dans cette classification mais répartis en stade localisé (tumeur présente dans un seul poumon) ou disséminé (tumeur métastatique).

Le NSCLC est généralement diagnostiqué à un stade III ou IV, et associé à une survie à 5 ans d'environ 10 et 1%, respectivement (Goldstraw et al., 2011). Plus de la moitié des SCLC sont diagnostiqués à un stade disséminé où la survie à 5 ans est estimée à environ 1% (Jackman and Johnson, 2005) (Tableau 2).

Tableau 2. Résumé des pourcentages de diagnostic et taux de survie à 1 et 5 ans selon le type de cancer et le stade. Les données présentées sont résumées de Goldstraw et al., 2011 ; Jackman and Johnson, 2005 et Cancer Research UK (consulté Juin 2017, mis à jour Février 2017).

Cancer du poumon non à petites cellules (NSCLC)

Stade	Diagnostic (%)	Taux de survie (%)	
		1 an	5 ans
Local			
I	15	83	35
II	7	68	21
Localement avancé			
III	19	44	6
Métastatique			
IV	48	17	<1
Inconnu	11	26	6

Cancer du poumon à petites cellules (SCLC)

Stade	Diagnostic (%)	Taux de survie (%)	
		1 an	5 ans
Localisé	35	65	15
Disséminé	65	20	1

Définir le type de cancer et son stade fournit aux praticiens hospitaliers un indicateur pronostique de la maladie et jalonne la prise en charge du patient.

IV. Traitements

Les cancers broncho-pulmonaires sont traités par chirurgie, chimiothérapie, radiothérapie ou thérapies ciblées. Ces traitements peuvent être administrés seuls ou en combinaisons et leur utilisation diffère selon le type de cancer, le stade et l'état général du patient.

IV.1 Cancer du poumon à petites cellules

Du fait d'une capacité de croissance rapide et d'un développement métastatique précoce, le SCLC est le type de cancer le plus agressif. Il est alors principalement diagnostiqué à un stade disséminé où la chirurgie n'est plus applicable. Ces patients sont traités par chimiothérapie, à partir de cisplatine ou

carboplatine et etoposide en combinaison. Malgré une chimio-sensibilité initiale importante, les patients deviennent rapidement résistants et la survie médiane est de 7 à 12 mois (Arcaro, 2015). Une minorité de patients atteints de cancers à un stade limité est traitée par des combinaisons de chimiothérapies et radiothérapies. Aucune thérapie ciblée n'est aujourd'hui disponible pour traiter ces patients mais de nombreux essais cliniques ciblant notamment la protéine SRC, la protéine VEGF impliquée dans l'angiogenèse ou la voie de signalisation PI3K/AKT/mTOR sont actuellement en phase 2 (Altan and Chiang, 2015).

IV.2 Cancer du poumon non à petites cellules

Au moment du diagnostic, environ 20% des patients présentent un cancer localisé (stade I ou II), 30% des patients présentent un stade avancé local (stade IIIA et IIIB) et environ la moitié des patients présentent un stade avancé métastatique (stade IV).

IV.2.a Tumeurs résécables

La chirurgie reste le traitement curatif le plus efficace car il permet de retirer la tumeur et protéger les tissus sains. Ce traitement n'est cependant proposé qu'à des stades I ou II, voir IIIA. C'est-à-dire à des stades précoces, et ne concerne donc qu'une minorité de patients.

Le traitement consiste alors en une lobectomie, pouvant être étendue selon le cas à une pneumonectomie, et d'un curage ganglionnaire. Une chimiothérapie post-opératoire est également administrée pour les stades II et IIIA, celle-ci augmente de 5 à 10% la survie à 5 ans du patient (Dunant et al., 2005). Une radiothérapie peut-être également proposée selon l'étendue des lésions.

IV.2.b Tumeurs non résécables et non métastatiques

Ce type de tumeurs correspond aux cancers de stade III, soit un stade localement avancé non traitable par chirurgie. Ces patients sont traités par combinaison de chimiothérapie et radiothérapie, simultanée ou séquentielle. La chimiothérapie classiquement délivrée et la plus efficace est à base d'étoposide et cisplatine. Des protocoles alternatifs notamment à base de taxol et vinorelbine peuvent être proposés, ceux-ci ont cependant une efficacité réduite comparativement au traitement usuel. La radiothérapie délivrée est de 65 à 70 Gray, l'intensité du rayonnement ionisant dépendant notamment du volume de la tumeur à traiter.

IV.2.c Tumeurs métastatiques

Pour ces cancers de stade IV, les traitements proposés sont soit une thérapie ciblée lorsqu'elle est possible, soit une combinaison de chimiothérapie classique et radiothérapie. Grâce à une meilleure

compréhension des mécanismes moléculaires à l'origine du cancer du poumon, plusieurs thérapies ciblées ont récemment vu le jour. Ces thérapies ciblent des récepteurs ou protéines impliquées dans les voies de signalisation intracellulaires et jouant un rôle dans le développement tumoral. Afin d'orienter la prise en charge thérapeutique du patient vers une thérapie ciblée ou non, des tests compagnons définissent les altérations génétiques de la tumeur. Ceux-ci consistent au séquençage de la tumeur (site primitif ou métastatique) afin d'identifier principalement les mutations sur les gènes de l'*EGFR* (Epidermal Growth Factor Receptor), *KRAS* (KRAS proto-oncogene, GTPase), *BRAF* (B-Raf proto-oncogene, serine/threonine kinase), *HER2* (Erb-B2 Receptor Tyrosine Kinase 2), *MET* (MET Proto-Oncogene, Receptor Tyrosine Kinase), *PI3KCA* (Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha) et les translocations des gènes *ALK* (Anaplastic Lymphoma Receptor tyrosine kinase) et *ROSI* (ROS Proto-Oncogene 1, Receptor Tyrosine Kinase) (INCa, 2017). Ces gènes altérés sont regroupés sous le terme de biomarqueurs et classés selon le type d'altération et, dans le cas de mutation, selon la position de la mutation. Les tests compagnons sont l'élément essentiel permettant d'orienter le patient vers une thérapie ciblée ayant eu une autorisation de mise sur le marché (AMM) français. Si le patient n'est pas éligible aux thérapies ciblées ayant une AMM, il pourra être éventuellement inclus dans des essais cliniques en cours.

Les mécanismes d'altérations génétiques ainsi que les gènes altérés seront détaillés dans les prochains paragraphes. Ils sont regroupés en trois catégories :

- Mutation d'un gène codant pour un récepteur membranaire ou une protéine de signalisation cellulaire (*EGFR*, *KRAS*, *BRAF*, *PIK3CA*, *HER2*).
- Translocation des gènes *ALK* et *ROSI*.
- Amplification de *MET*.

Le gefitinib (IressaTM) et l'erlotinib (TarcevaTM) sont des inhibiteurs de tyrosine kinase (ITKs) prescrits lors de mutations spécifiques sur l'*EGFR* (délétion de l'exon 19, mutations de l'exon 20 codant pour les mutations EGFR L858R ou EGFR T790M). Le crizotinib (XalkoriTM) est un ITK prescrit lors de translocation du gène *ALK* ou *ROSI*. Enfin, le vemurafenib (ZelborafTM), est un ITK prescrit lors de la mutation de *BRAF* codant pour la mutation BRAF V600E. Ces thérapies ciblées peuvent être associées à une chimiothérapie pour une action synergique. Le bénéfice de telles thérapies est une survie sans progression de la maladie pouvant aller de 9 à 12 mois selon la traitement considéré (Hiley et al., 2016).

L'ensemble des cibles recherchées ainsi que les thérapies disponibles sont résumées dans la Tableau 3. Précisons que le statut d'autorisation de mise sur le marché d'un médicament varie selon les pays concernés.

Tableau 3. Résumé des cibles recherchées par séquençage et thérapies disponibles associées ou en cours d'essai clinique. Certains inhibiteurs comme le crizotinib ne sont pas exclusifs de leur cible. L'olmutinib marqué d'une étoile (*) est seulement approuvé en Corée du Sud. Tableau extrait de Hirsch et al., 2016.

Target and drugs	Status	Target and drugs	Status
EGFR sensitising		HER2	
Gefitinib	Approved	Trastuzumab emtansine	Phase 2
Erlotinib	Approved	Afatinib	Phase 2
Afatinib	Approved	Dacomitinib (PF-00299804)	Phase 1
Osimertinib	Approved		
Olmutinib (BI1482694/HM61713)*	Approved	BRAF	
ASP8273	Phase 3	Vemurafenib	Phase 2
EGF816	Phase 2	Dabrafenib	Phase 2
PF-06747775	Phase 2	Dabrafenib plus trametinib	Phase 2
ALK		RET	
Crizotinib	Approved	Cabozantinib	Phase 2
Alectinib	Approved	Alectinib	Phase 2
Ceritinib	Approved	Apatinib (YN968D1)	Phase 2
Lorlatinib (PF-06463922)	Phase 2	Vandetanib	Phase 2
Brigatinib (AP26113)	Phase 3	Ponatinib	Phase 2
X-396	Phase 3	Lenvatinib	Phase 2
TSR-011	Phase 2		
Entrectinib	Phase 2	NTRK1	
ROSI		Entrectinib	Phase 2
Crizotinib	Approved	LOXO-101	Phase 2
Cabozantinib	Phase 2	Cabozantinib	Phase 2
Ceritinib	Phase 2	DS-6051b	Phase 1
Lorlatinib	Phase 2	PIK3CA	
Entrectinib (RXDX-101)	Phase 2	LY3023414	Phase 2
DS-6051b	Phase 1	PQR309	Phase 1
MET		MAP2K1	
Crizotinib	Phase 2	Selumetinib	Phase 3
Cabozantinib	Phase 2	Trametinib	Phase 2
INC280	Phase 2	Cobimetinib	Phase 1
MGCD265	Phase 2		

Une telle analyse nécessite de pouvoir disposer d'un échantillon de la tumeur. Des progrès technologiques (Next Generation Sequencing, NGS) et fondamentaux récents (mise en évidence de l'ADN tumoral circulant, ctDNA) permettent désormais de définir le statut mutationnel d'une tumeur à partir d'une simple prise de sang (Thompson et al., 2016). Ces biopsies liquides sont d'un intérêt particulier pour les patients non opérables. Au cours de ma thèse, j'ai contribué à la rédaction d'une revue traitant des outils technologiques à disposition pour détecter le ctDNA et orienter la prise en charge des patients atteints d'un cancer du poumon (Vendrell et al., 2017a - voir Publication 1). Avec ces technologies, de nouveaux biomarqueurs sont également mis en évidence et ouvrent de nouvelles

perspectives de traitements (Hirsch et al., 2016). C'est le cas de la détection de la mutation activatrice du gène *PDL1* qui peut être traitée par immunothérapie.

Lorsque le patient n'est éligible à aucune thérapie ciblée ou lorsque la recherche de mutation n'est pas possible, celui-ci bénéficiera d'une chimiothérapie à base de cisplatine associée à une radiothérapie.

Face à un enjeu majeur de santé publique, la France a proposé à une échelle nationale, au travers des différents plans cancers, une organisation du séquençage des tumeurs basée autour de la création de plateformes régionales de biologie moléculaire. Celles-ci, réparties sur l'ensemble du territoire, permettent une prise en charge thérapeutique pour une médecine personnalisée optimale, et une accessibilité pour tous à ces nouveaux traitements. Le troisième et dernier plan cancer couvre la période 2014-2019. Les objectifs donnés par le gouvernement sont d'atteindre 60 000 séquençages en 2018 et de doubler le nombre d'essais cliniques en France. Pour se donner les moyens de ces enjeux, 60 millions d'euros ont été investis dans le développement et la généralisation de la médecine personnalisée au travers de différentes actions (INCa, 2014).

B. Biomarqueurs

Comme nous l'avons vu précédemment, sauf rares exceptions, l'essentiel des altérations décrites et pour lesquels il existe des thérapies ciblées concerne les NSCLC. Au cours de ma thèse, je me suis intéressé à l'étude de ce type de cancer et plus précisément à l'étude des adénocarcinomes pulmonaires.

Le séquençage des tumeurs de patients de stade IV montre une très importante variété d'altérations génétiques se concentrant sur une quinzaine de gènes. Parmi eux, les gènes *KRAS*, *EGFR* et *BRAF* sont les plus fréquemment mutés avec des incidences respectives d'environ 30%, 10% et 5%. D'autres altérations telles que le réarrangement du gène *ALK* ou *ROS1* sont également impliquées avec des fréquences variant de 1 à 7% selon les études (Soda et al., 2007; Vargas and Harris, 2016). Toutes ces altérations sont généralement mutuellement exclusives et celles-ci sont partagées en trois groupes :

- Mutations activatrices des gènes codant des récepteurs aux facteurs de croissance (*EGFR*, *ERBB2*) ou activatrices des gènes codant pour des protéines de signalisation (*KRAS*, *BRAF*, *MAP2K1*, *NRAS*, *HRAS*).
- Translocation des gènes *ALK*, *ROS1*, *RET* et *MET*.
- Amplification des gènes *ERBB2* et *MET*.

La figure 3 décrit l'évolution du génotypage des adénocarcinomes pulmonaires au cours des dernières décennies et met en évidence les sauts technologiques dans les techniques de séquençage.

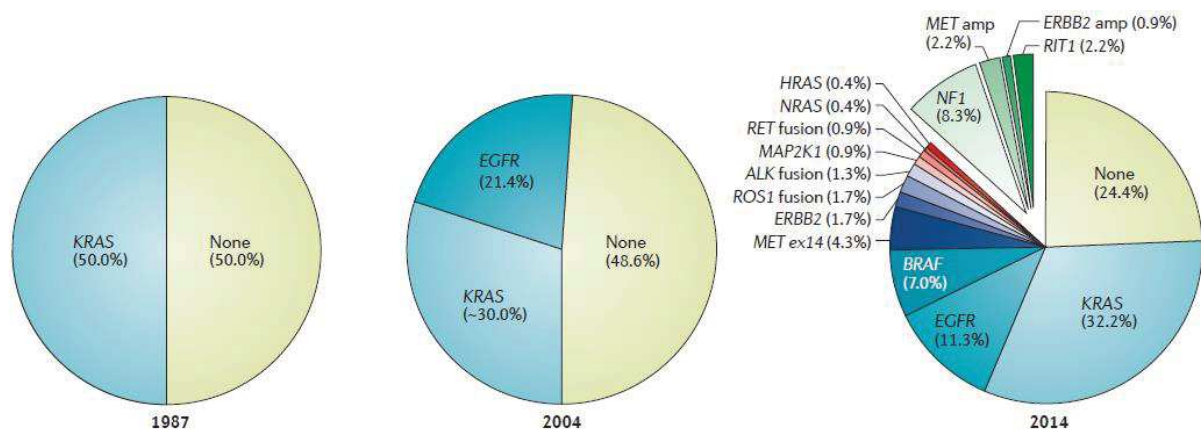


Figure 3. Évolution de la caractérisation des altérations génétiques des adénocarcinomes pulmonaires. Le développement de nouvelles techniques de séquençage associé à l'initiative de projets collaboratifs de grande échelle permet d'une part d'identifier des mutations fréquentes pouvant potentiellement être ciblées par des drogues et permet d'autre part une meilleure classification des patients selon différentes dimensions, notamment en termes d'altérations identifiées (translocations, mutations ponctuelles) ou de voies de signalisation activées. Amp., amplification. Figure extraite de Vargas and Harris, 2016.

I. La voie des ‘Mitogen-Activated Protein Kinases’ (MAPK)

Avec les mutations touchant les gènes de l'*EGFR*, *KRAS* ou *BRAF*, environ 50% des mutations décrites concernant des acteurs majeurs impliqués dans la voie des MAPK. De plus, il a été démontré plus récemment que la translocation ALK active également cette voie de signalisation (Hrustanovic et al., 2015). D'un point de vue plus global, il apparaît donc qu'une des voies majeures activée dans les adénocarcinomes pulmonaires est la voie des MAP Kinases.

Cette voie de signalisation conservée au cours de l'évolution permet aux cellules de s'adapter à leur environnement. En effet, cette voie est constituée d'une cascade de phosphorylations qui relaie l'information d'un signal extracellulaire au noyau pour contrôler *in fine* la prolifération, la différenciation, la survie et la mobilité (Arthur and Ley, 2013). Selon les stimuli extérieurs, ces processus cellulaires s'activeront spécifiquement. La figure 4 présente le schéma général de cette voie (Kim and Bar-Sagi, 2004).

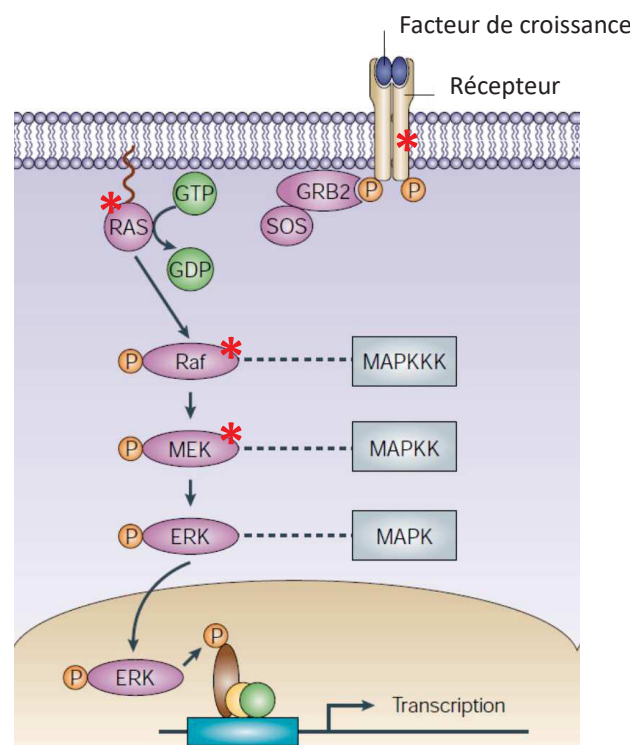


Figure 4. Description générale de la voie des MAP Kinases. Il s'agit d'une cascade de phosphorylations qui débute par l'activation par phosphorylation d'un récepteur aux facteurs de croissance. Le complexe GRB2-SOS est alors recruté à la membrane et active RAS. RAS activé active à son tour les protéines kinases RAF (MAPKKK) qui s'activent par phosphorylation. RAF activé active par phosphorylation MEK (MAPKK), puis MEK activé active par phosphorylation ERK (MAPK). ERK activé sera adressé dans le noyau et activera spécifiquement des gènes impliqués dans les fonctions notamment de prolifération et de croissance. ERK phosphorylé peut également activer des protéines présentes dans le cytoplasme (non représenté). Les protéines marquées d'une étoile (*) sont fréquemment mutées, et de façon exclusive, dans les adénocarcinomes pulmonaires. Figure extraite et adaptée de Kim and Bar-Sagi, 2004.

La voie des MAP Kinase est une cascade de phosphorylations complexe, faisant intervenir de nombreux effecteurs conventionnels ou atypiques. Il a été décrit environ 18 MAPKKK (MAP Kinase Kinase Kinase), 11 MAPKK (MAP Kinase Kinase) et 14 MAPK (MAP Kinase) avec leurs spécificités propres. Cette cascade de phosphorylation peut donc emprunter différents chemins au sein de cette même voie afin d'adapter selon le contexte la cellule à son environnement (Arthur and Ley, 2013; Cargnello and Roux, 2011).

Brièvement, l'activation de la voie des MAP Kinases débute par l'activation de récepteurs à domaine tyrosine kinase par fixation d'un ligand présent dans le milieu extra-cellulaire. Le récepteur le plus fréquemment décrit est l'EGFR, mais il en existe d'autres tels que FGFR (Fibroblast growth factor receptor) ou PDGFR (Platelet-derived growth factor receptor). Cette activation se caractérise par la dimérisation des récepteurs puis par une auto- ou transphosphorylation des domaines tyrosine kinase. Cette phosphorylation permet, *via* des protéines adaptatrices en complexe telles que GRB2 (Growth Factor Receptor Bound Protein 2) et SOS (SOS RAS/Rac Guanine Nucleotide Exchange Factor 1), l'activation des GTPases de la famille RAS. Les protéines RAS ainsi activées interagissent alors avec les MAPKKK *via* un domaine protéique RBD (RAS-binding domain) ou RBL (RBD-like). Une famille importante de MAPKKK est la famille des RAF kinases, dont *BRAF* est fréquemment muté. L'interaction avec les protéines RAS permet alors l'activation des MAPKKK par phosphorylation, puis à leur tour les MAPKKK activent les MAPKK, dont MAP2K1 (Mitogen-Activated Protein Kinase Kinase 1, ou MEK1) et MAP2K2 (Mitogen-Activated Protein Kinase Kinase 2, ou MEK2). Les MAPKK activées phosphorylent ensuite les MAP Kinase (MAPK) comme MAPK3 (Mitogen-Activated Protein Kinase 3, ou ERK1) et MAPK1 (Mitogen-Activated Protein Kinase 1, ou ERK2). Une fois activées, ces dernières vont réguler l'activité de protéines cytoplasmiques ou nucléaires après translocation dans le noyau.

Parmi les cibles nucléaires de ERK les plus souvent décrites, nous pouvons mentionner les facteurs de transcription ELK1 (ELK1, ETS transcription factor) qui régulent notamment la prolifération et survie cellulaire (Boros et al., 2009), MYC (MYC proto-oncogene, bHLH transcription factor) qui contrôle notamment la croissance cellulaire (Patel et al., 2004), et TSC2 (Tuberous Sclerosis 2) qui régule la synthèse protéique *via* la voie de signalisation PIK3/AKT/mTOR (Tchevkina and Komelkov, 2012). Les cibles cytoplasmiques de ERK1 et ERK2 sont les protéines MK (MAPK-activated protein kinases) qui comprennent les protéines MNK (MAPK interacting kinases), MSK (mitogen- and stress-activated kinases) et RSK (90-kDa ribosomal S6 kinases). Les protéines MNK régulent le métabolisme à travers l'activation du complexe eIF4F (eukaryotique initiation factor 4 F) et le contrôle de la traduction. Les protéines MSK transloquent dans le noyau et agissent comme régulateurs transcriptionnels. Enfin, l'activation par phosphorylation des protéines RSK induit leur translocation dans le noyau et régulent la transcription de gènes impliqués notamment dans la survie et le cycle

cellulaire. Les cibles des RSK sont notamment les protéines SRF (Serum Response Factor), CREB (cAMP response element-binding protein) et BAD (BCL2 associated agonist of cell death).

Cette cascade de phosphorylation est particulièrement altérée et toutes les altérations oncogéniques décrites au niveau de l'EGFR, de la famille RAS, de BRAF ou de ALK ont pour point commun de virtuellement dissocier la cellule de son environnement extérieur en fournissant à la cellule un signal d'activation de la voie MAPK, alors qu'aucun stimulus externe n'est en réalité présent. Ces gènes ainsi que la structure des protéines, et leurs mutations associées sont présentés dans les sections suivantes.

II. Le récepteur aux facteurs de croissance EGFR

II.1 Structure et fonction

Le gène de l'Epithelial Growth Factor Receptor est situé sur le brin positif du bras court du chromosome 7 (7p11.2). D'une longueur de 193 608 paires de bases, il est constitué de 28 exons. À ce jour, 9 transcrits ont été décrit et le variant 1 est considéré comme canonique. Constitué de 6369 bases, ce gène code pour une protéine d'environ 170 kDa et de 1210 acides aminés, dont 24 constituent un motif d'adressage à la membrane plasmique clivé lors de modifications post-traductionnelles.

L'EGFR est un récepteur à activité tyrosine kinase (RTK) dont la localisation principale est à la membrane plasmique. Il assure notamment l'activation de la voie des MAPK, et contrôle en particulier la prolifération, la différenciation, la survie et la motilité. La protéine est composée de 5 régions : extracellulaire, transmembranaire, juxta-membranaire, tyrosine kinase et enfin une région régulatrice C-terminale (Ferguson, 2008) (Figure 5).

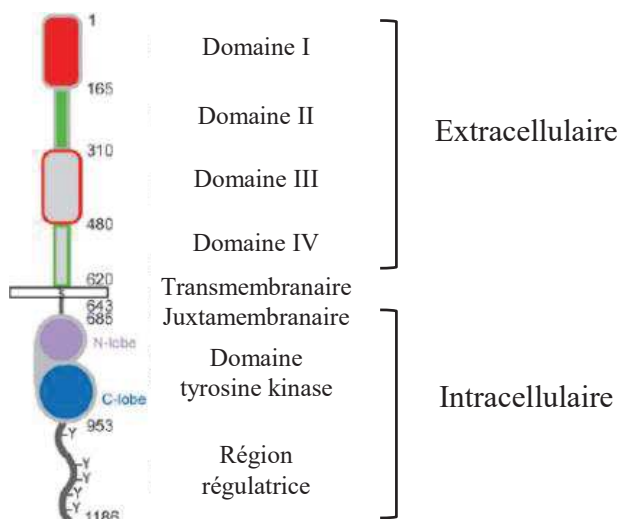


Figure 5. Représentation de l'EGFR et de ses régions et domaines. Notons qu'il est indiqué que la protéine est composée de 1186 acides aminés, et non 1210 comme mentionné. Ceci s'explique par la perte d'un peptide d'adressage à la membrane de 24 résidus. Figure extraite et adaptée de Ferguson, 2008.

- **La région extracellulaire** permet la fixation du ligand dont les principaux sont l'EGF (Epidermal Growth Factor) et TGF- α (Transforming Growth Factor alpha). Après fixation du ligand, cette région permet aussi la dimérisation du récepteur. Plus précisément, le ligand se fixe sur les domaines I et III et induit un changement conformationnel du domaine II qui est la boucle de dimérisation. Cette modification conformationnelle du domaine II permet l'homo- ou l'hétéro-dimérisation du récepteur avec ERBB2 (Erb-B2 Receptor Tyrosine Kinase 2) (Brand et al., 2011).
- **La région transmembranaire** hydrophobe de 23 acides aminés riche en leucines permet l'ancrage de la protéine à la membrane plasmique. Elle est composée d'un seul segment transmembranaire.
- **La région juxtamembranaire** serait essentielle à l'activation du récepteur après l'étape de dimérisation. En effet, son remplacement par une séquence dégénérée aboutit à l'absence totale de phosphorylation des résidus tyrosine (He and Hristova, 2012). En revanche, son rôle dans la dimérisation du récepteur est controversée (Abd Halim et al., 2015; Purba et al., 2017).
- **La région catalytique tyrosine kinase** riche en sérines et tyrosines codée par les exons 18 à 24 est le lieu de phosphorylations et d'activation du récepteur. À l'état inactif, ce domaine est auto-inhibé par le repliement structural des lobes N et C (Figure 5). Lors de la dimérisation du récepteur, les domaines kinases entrent en contact et lèvent cette contrainte d'auto-inhibition (Changeux and Christopoulos, 2016). Cette région porte deux domaines majeurs : un premier domaine de fixation à l'ATP et un second domaine d'activation stabilisant l'ATP et son substrat pour la phosphorylation. Les motifs phosphorylés permettront ensuite l'interaction avec des protéines adaptatrices portant des domaines d'interaction SH2 (Src homology 2), telles que SRC ou STAT5 (Signal Transducer And Activator Of Transcription 5) qui permettront d'activer différentes voies de signalisation, selon le motif phosphorylé (Shostak and Chariot, 2015) (Figure 6).
- **La région régulatrice C-terminale** d'autophosphorylation adjacente au domaine kinase est moins caractérisée. Il semblerait cependant qu'elle entre en contact avec le domaine kinase par recouvrement et régulerait ainsi négativement son activité (Gajiwala, 2013). Lors de l'activation du récepteur, l'autophosphorylation de ce domaine permettrait de lever cette inhibition (Linggi and Carpenter, 2006). L'activation par phosphorylation de cette région permet de recruter d'autres protéines comme CBL, GRB2, PLC γ .

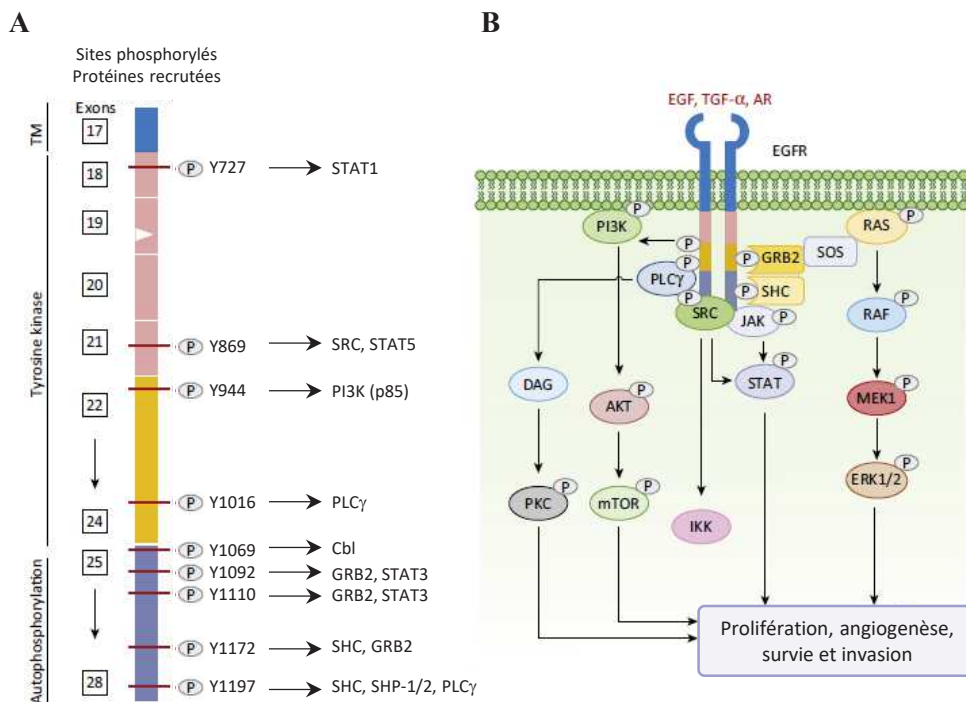


Figure 6. Représentation des sites de phosphorylation de l'EGFR et des protéines activées correspondantes. (A) Selon le contexte cellulaire, l'EGFR est phosphorylé sur certains résidus tyrosine et recrute spécifiquement des protéines de signalisation. (B) L'EGFR activé recrute des protéines impliquées dans des voies de signalisation variées qui conduisent principalement à l'activation des voies de prolifération, survie et invasion. TM, domaine transmembranaire. Figure extraite et adaptée de Shostak and Chariot, 2015.

Les phosphorylations sont les modifications post-traductionnelles les plus couramment décrites de l'EGFR. Toutefois, il existe d'autres modifications post-traductionnelles comme la N-glycosylation qui joue également un rôle central dans le repliement de la protéine, et est déterminante pour l'interaction entre la protéine et son ligand (Kaszuba et al., 2015). Parmi ses autres fonctions, l'EGFR peut être également internalisé par endocytose et transloquer dans le noyau où il agira comme facteur de transcription de gènes impliqués notamment dans la prolifération et le cycle cellulaire tels que *MYC* et la cycline D1 (*CNND1*), respectivement (Wang et al., 2010).

II.2 Les mutations de l'EGFR

Dans les adénocarcinomes pulmonaires, l'*EGFR* est très souvent muté. Ces mutations touchent plutôt les faibles consommateurs de tabac, les femmes et les patients d'origine asiatiques. Dans les stades IV de la maladie, les analyses oncosomatiques des adénocarcinomes pulmonaires ont permis de mettre en évidence principalement 5 types d'anomalies de l'*EGFR* : des mutations faux-sens, des délétions, des insertions, des amplifications et des surexpressions. La détection de ces altérations activatrices au niveau du domaine tyrosine kinase est une étape essentielle dans la mesure où elles sont potentiellement

sensibles aux thérapies ciblées par ITKs. La figure 7 ci-après décrit les principales mutations de l'*EGFR* et leur sensibilité aux ITKs de première génération, le gefitinib et l'erlotinib.

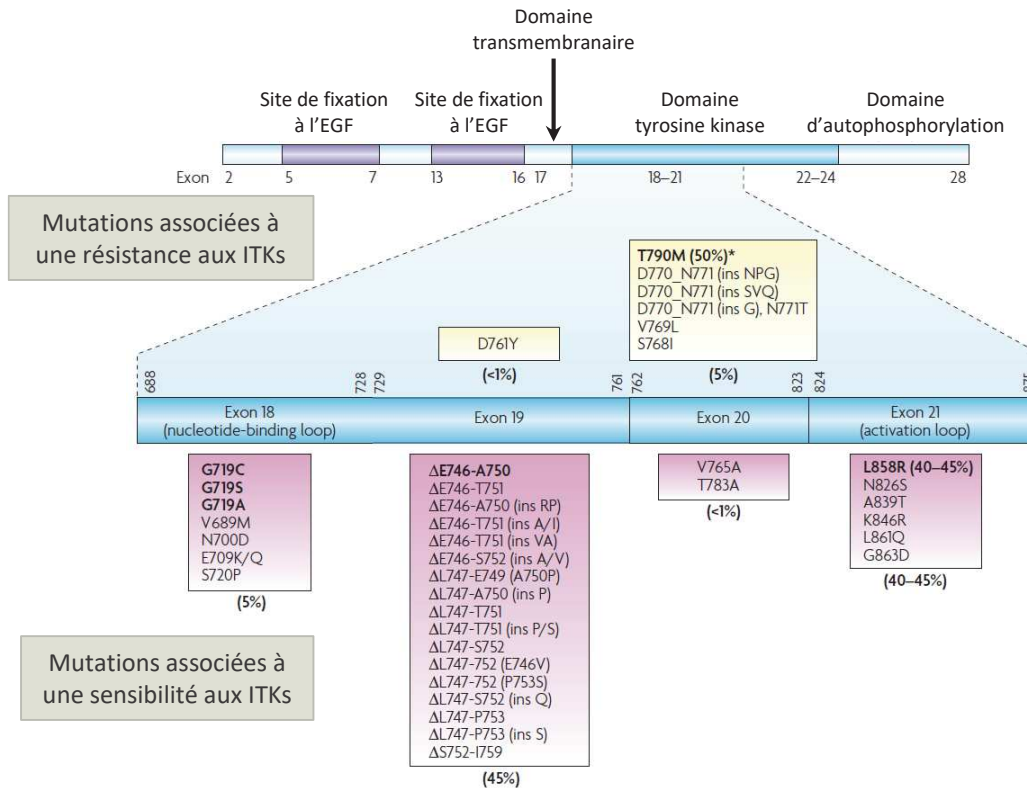


Figure 7. Représentation des mutations ponctuelles et indels identifiés dans les exons 18 à 21 de l'*EGFR* ainsi que de leurs sensibilités respectives aux ITKs. Plus d'une 20aine de mutations ont été décrites. Elles peuvent se classer en mutation de résistance aux thérapies ciblées (en jaune) et mutation de sensibilité aux thérapies ciblées (rose). La mutation EGFR T790M de l'exon 20 noté d'une étoile (*) est détectée comme mutation de résistance aux ITKs dans 50% des cas. Sous chacun des groupes de mutations est indiqué le pourcentage de diagnostic correspondant. Les ITKs mentionnées ici correspondent à l'erlotinib et au gefitinib (ITKs de première génération). EGF, Epidermal Growth Factor ; ITKs, Inhibiteurs Tyrosine Kinases. Figure extraite et adaptée de Sharma et al., 2007.

La délétion $\Delta E746-A750$ de l'exon 19 et la mutation L858R de l'exon 21 constituent à elles seules plus de 90% des cas détectés.

II.2.a Délétion de l'exon 19

Plus d'une dizaine de variants de délétions ont été décrits. Ils se caractérisent par des longueurs de délétions variables, parfois associées à des insertions d'acides aminés. La délétion la plus fréquente (45% des cas) est la perte du motif de 5 acides aminés 746-ELREA-750 codée par l'exon 19 et est communément appelée mutation EGFR Del.19. Cette délétion est adjacente au domaine de fixation à l'ATP de la région tyrosine kinase. Elle a pour conséquence un changement conformationnel

stabilisateur de la fixation de l'ATP, favorisant l'activation du domaine kinase (Imai and Takaoka, 2006; Kumar et al., 2008).

II.2.b Mutation L858R

Cette mutation de l'exon 21 code pour une mutation faux-sens de la boucle d'activation de la région catalytique tyrosine kinase. Le remplacement de la leucine par une arginine conduit à un changement conformationnel de la boucle d'activation, stabilisant la conformation active de l'EGFR en rendant moins flexible l'hélice α C du domaine kinase. Comme dans le cas de la délétion de l'exon 19, la fixation de l'ATP est ainsi stabilisée (Gazdar, 2009).

Ces deux mutations EGFR Del.19 et EGFR L858R ont été décrites pour la première fois en 2004 et leurs présences ont été corrélées à la réponse au gefitinib la même année (Lynch et al., 2004; Paez et al., 2004). D'autres essais cliniques ont permis de mettre en évidence l'efficacité de l'erlotinib, un autre inhibiteur de l'EGFR. Plus généralement, les patients porteurs d'une de ces mutations présentent un taux de réponse de plus de 70% et une survie sans progression de la maladie augmentée d'environ 6 mois comparativement aux chimiothérapies conventionnelles (Asahina et al., 2006; Pircher et al., 2010). Le gefitinib et l'erlotinib sont donc les deux ITKs prescrit lors de mutation de l'EGFR. D'un point de vue enzymatique, ces drogues sont des inhibiteurs compétitifs de la fixation de l'ATP à l'EGFR (Figure 8) (Imai and Takaoka, 2006). Cette gêne stérique induit par la fixation de ces drogues réduit donc l'activité kinase de la protéine et *a fortiori* son activation. Rappelons que d'autres mutations moins fréquentes de l'EGFR sont également sensibles à ces ITKs, c'est le cas par exemple des mutations EGFR G719C et EGFR L861Q.

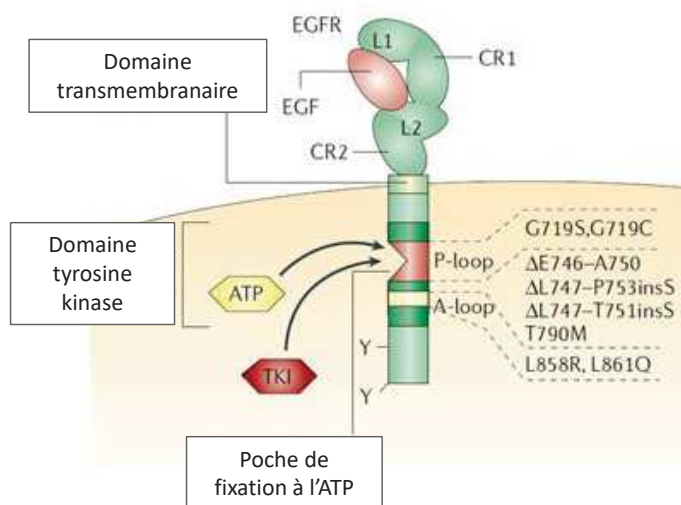


Figure 8. Représentation du mécanisme d'inhibition compétitive des ITKs de l'EGFR et de la position des mutations les plus fréquemment identifiées. Figure extraite et adaptée de Imai and Takaoka, 2006.

II.2.c Mutations de résistance

Malgré la réponse initiale importante des patients aux ITKs, la majorité des patients développeront des mécanismes de résistance environ un an après le début du traitement. Plus d'une dizaine de mécanismes ont été décrits pour expliquer cette résistance, lesquels peuvent être regroupés en deux familles : les mécanismes de résistance ciblant directement l'*EGFR*, et ceux contournant l'inhibition de son activité par les ITKs en ne ciblant pas l'*EGFR* mais par exemple ses partenaires (Figure 9) (Camidge et al., 2014).

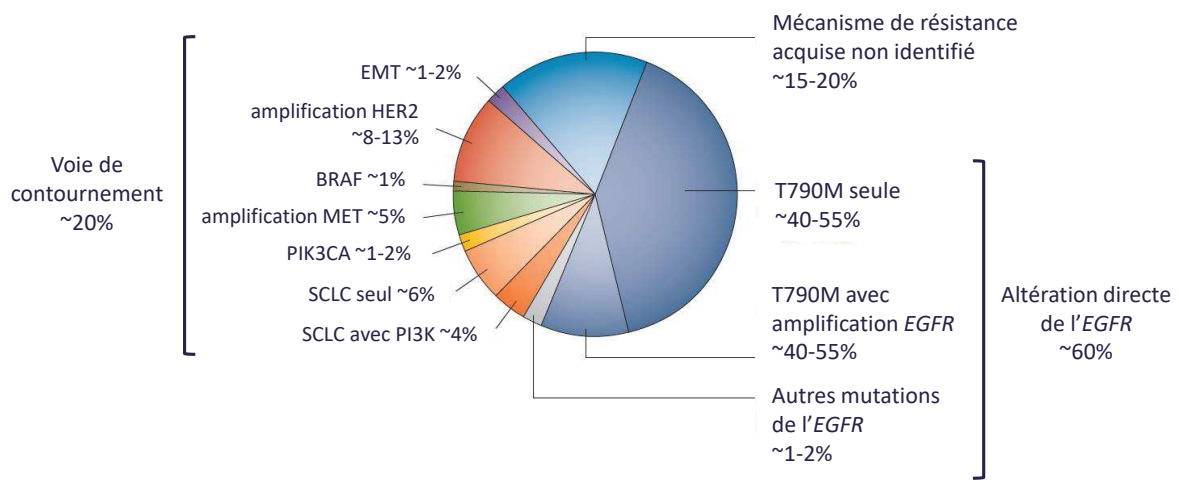


Figure 9. Mécanismes de résistances aux ITKs ciblant l'*EGFR*. Ces mécanismes sont réparties en deux catégories, ceux ciblant directement l'*EGFR* (amplifications et/ou mutations ponctuelles) et ceux contournant son inhibition (mutations ou amplification d'autres gènes, transitions histologiques). Environ 20% des mécanismes de résistance sont encore inconnus. Figure extraite et adaptée de Camidge et al., 2014

Les mécanismes ciblant l'*EGFR* comptent pour environ 60% des cas et correspondent presque exclusivement à la mutation secondaire *EGFR* T790M. Cette thréonine 790 est localisée dans la poche de fixation de l'ATP et régule l'accessibilité de la drogue à ce domaine. Plus précisément, cette mutation augmente la gêne stérique de l'inhibiteur et réduit son affinité au domaine kinase au profit de l'ATP. *In fine*, l'inhibition de l'*EGFR* s'en trouve réduite (Yun et al., 2008). La fréquence élevée de la mutation T790M a amené à développer des inhibiteurs de seconde génération tels que l'afatinib (Giotrif™) ou le dacomitinib (Nerlynx™). Cependant, un essai clinique démontrant un taux de réponse inférieur à 10% chez les patients porteur d'une mutation secondaire, et des effets secondaires importants ont conduit à reconsidérer l'efficacité de l'afatinib dans le traitement des rechutes (Miller et al., 2012). Des ITKs de troisième génération ciblant à la fois les mutations primaires et secondaires ont alors été développés, et ont montré leur efficacité. En mars 2017, l'inhibiteur osimertinib (AZD9291, Tagrisso™) développé par la société AstraZeneca a reçu l'approbation de la FDA (U.S. Food and Drug Administration) pour le traitement des patients avec tumeurs métastatiques portant une mutation de l'*EGFR* T790M. Cette molécule inhibe spécifiquement l'*EGFR* T790M mais également les mutations primaires L858R et

Délétion 19 – l'EGFR sauvage n'étant pas ciblé (Skoulidis and Papadimitrakopoulou, 2017). À l'inverse du gefitinib ou de l'erlotinib (ITKs de première génération), la fixation de cet inhibiteur est irréversible. Cependant, comme pour les inhibiteurs de première génération des mutations de résistance apparaissent environ un an après le début de traitement, notamment les mutants C797S et L718Q. Aujourd'hui des inhibiteurs de quatrième génération comme l'EAI045 sont en cours de développement pour cibler ces dernières mutations (Jia et al., 2016; Wang et al., 2017).

Les mécanismes contournant l'inhibition par les ITKs sans pour autant altérer l'EGFR comptent pour 20% des cas et sont variés. Ils peuvent correspondre à l'amplification du gène *HER2* ou *MET* ou à des mutations ciblant notamment les gènes *BRAF* et *PIK3CA*. L'amplification ou le recrutement d'autres kinases permet également de contourner l'inhibition de l'EGFR en activant d'autres voies de signalisation telles que PI3K/AKT/MTOR ou FGFR et ainsi maintenir un état prolifératif (Sequist et al., 2011; Ware et al., 2013). Les stratégies actuelles de traitement des rechutes sans mutation secondaire de l'EGFR portent sur des combinaisons de thérapies ciblant ces voies alternatives activées tout en maintenant l'inhibition de l'EGFR (Bean et al., 2007). D'autres mécanismes de résistance existent comme par exemple la différenciation histologique des tumeurs par transition épithélio-mésenchymateuse.

Pour les 20% des cas restants, les mécanismes de résistance aux ITKs ne sont pas identifiés.

III. Les protéines de la famille RAS

III.1 Structure des protéines RAS

Les protéines de la famille RAS sont de petites GTPases d'environ 21kDa ancrées à la membrane plasmique. Celles-ci sont activées par l'activation des récepteurs à activité tyrosine kinase, tels que l'EGFR. Les 4 isoformes, nommément HRAS, NRAS, KRAS4A et KRAS4B, sont codés par 3 gènes, *HRAS* (11p15.5), *NRAS* (1p13.2) et *KRAS* (12p12.1). L'exon 4 du transcrit de *KRAS* est épissé de manière alternative pour générer les isoformes KRAS4A et KRAS4B. L'isoforme KRAS4B est le plus exprimé et a longtemps été considéré comme canonique. Ainsi dans la littérature lorsqu'il est seulement mentionné KRAS, les auteurs traitent de KRAS4B.

Lorsqu'on compare ces isoformes, on remarque qu'ils partagent une très forte homologie de séquence (>80%) sur l'ensemble du domaine G qui constitue 90% de la longueur totale des RAS. Toutefois, adjacente à ce domaine G une région hypervariable (HVR) de 25 acides aminés homologues à 15% constitue la particularité de chacun de ces isoformes (Figure 10) (Bryant et al., 2014).

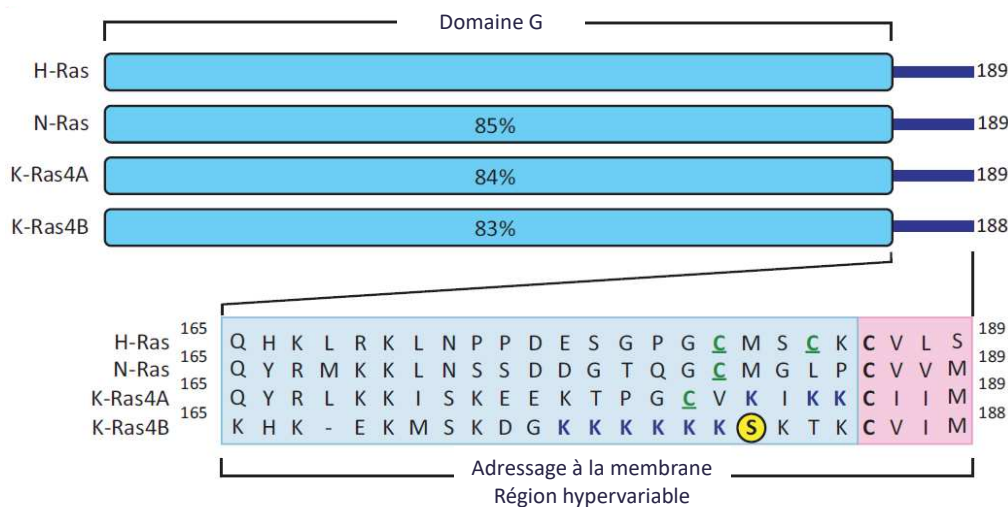


Figure 10. Structure des isoformes RAS. Le domaine G en bleu comprend les domaines de régulation des RAS et présente la plus forte homologie de séquence entre les isoformes (>80%). La séquence d'acides aminés présentée correspond à la séquence hypervariable des RAS (15% d'homologie), directement adjacente au domaine G. Cette séquence comprend les cystéines en vert et les lysines en bleu nécessaires aux modifications post-traductionnelles permettant l'adressage à la membrane plasmique. Les 4 derniers acides aminés de l'extrémité C-terminale formant la séquence -CAAX, est la séquence d'adressage au réticulum endoplasmique des RAS. C, cystéine ; A, acide aminé aliphatique ; X, acide aminé variable. Figure extraite et adaptée de Bryant et al., 2014.

Cette région HVR est le lieu de modifications post-traductionnelles spécifiques, telles que la farnésylation, la carboxyméthylation, la palmitoylation conduisant à des adressages membranaires spécifiques. Plus précisément, la protéine native est farnésylé sur la cystéine 186 de la séquence -CAAX (C, cystéine ; A, acide aminé aliphatique ; X, acide aminé variable) à l'extrémité C-terminal. Cette

modification cible RAS à la surface du réticulum endoplasmique. Il s'en suit une hydrolyse du motif –AAX puis une carboxymethylation de la cystéine restante. C'est à partir de ce point que les modifications post-traductionnelles diffèrent entre les isoformes. Alors que HRAS, NRAS et KRAS4A sont tous les 3 palmitoylés, KRAS4B est adressé directement à la membrane plasmique. Cet adressage à la membrane plasmique est facilité par une séquence basique riche en lysine. HRAS, NRAS et KRAS4A sont quant à eux palmitoylés une seconde fois et ceci permet leur adressage à l'appareil de Golgi. À la différence de NRAS et KRAS4A, HRAS sera palmitoylé une troisième fois. HRAS, NRAS et KRAS4A seront ensuite adressés à la membrane plasmique par transport vésiculaire. Le temps de demi-vie des palmitoylations est d'environ 20 min, ce qui indique une recirculation permanente entre l'appareil de Golgi et la membrane plasmique. Au final, les différentes modifications post-traductionnelles de RAS induisent des localisations membranaires différentes :

- HRAS est localisé dans les micro-domaines riches en cholestérol (radeaux lipidiques et caveolae) mais s'en dissocie lorsqu'il est activé.
- NRAS est localisé dans les radeaux lipidiques mais jamais associé aux caveolae.
- KRAS est toujours situé en dehors des radeaux lipidiques, quel que soit son statut d'activation.

Une fois adressées à la membrane, les protéines RAS sont fonctionnelles et peuvent également subir d'autres modifications telles que la phosphorylation et l'ubiquitination (Castellano and Santos, 2011).

III.2 Régulation des protéines RAS

Le domaine G est le domaine de régulation de l'activité des protéines RAS. Celui-ci comprend la boucle P, région de fixation du GTP ou GDP, et les régions Switch I et II qui contrôlent les interactions avec les régulateurs et effecteurs de RAS. Plus précisément, les protéines RAS agissent toutes comme des interrupteurs moléculaires qui pourront activer selon leur interaction avec le GTP les mécanismes de prolifération, survie ou différenciation (Ahearn et al., 2011; Bryant et al., 2014) (Figure 11).

À l'état inactif, les RAS sont liés au GDP et c'est l'interaction avec les protéines régulatrices de la famille GEF (Guanine nucleotide Exchange Factor) *via* un domaine protéique RAS-GEF qui favorise la dissociation du GDP au profit de la fixation du GTP. Une fois liées GTP, les protéines RAS sont activées et interagissent avec différents effecteurs de la signalisation cellulaire.

L'interaction de RAS avec les protéines régulatrices de la famille GAP (GTPase activating protein) *via* un domaine protéique RAS-GAP favorise l'hydrolyse du GTP en GDP et donc l'inactivation des protéines RAS. Malgré l'activité GTPase intrinsèque de RAS, sa constante catalytique est faible, de l'ordre de $2 \times 10^{-4} \text{s}^{-1}$. L'interaction avec les GAPs permet alors d'augmenter jusqu'à 10^5 fois l'hydrolyse (Chung et al., 1993). Des études de cristallographie ont permis de mettre en évidence que cette

inactivation est assurée par un résidu arginine qui se loge dans le site de fixation du GTP et favorise la conversion du GTP en GDP (Kupzig et al., 2009).

Qu'il s'agisse des GEFs ou des GAPs, en plus de leurs domaines respectifs d'interactions avec les protéines RAS, ces protéines portent des domaines protéiques qui permettent l'interaction avec les récepteurs membranaires ou des protéines associées à la membrane (les domaines SH2 et SH3), l'interaction avec les phospholipides de la membrane plasmique (domaine PH) (Vigil et al., 2010). Il a été décrit plus d'une dizaine de GEFs et de GAPs, dont les protéines SOS et NF1 (Neurofibromin 1) sont respectivement les régulateurs les plus décrits de ces deux familles (Bos et al., 2007).

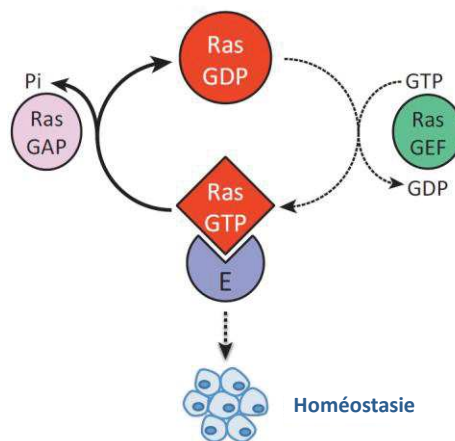


Figure 11. Boucle de régulation des protéines RAS. RAS inactivé et lié au GDP, est activé par interaction avec les protéines GEF (Guanine Exchange Factor), qui déstabilisent le GDP et favorisent la fixation de GTP. L'inactivation de RAS est assurée par l'interaction avec les protéines GAPs (GTPase activating protein) qui activent l'activité hydrolase intrinsèque des protéines RAS. E., Effecteur. Figure extraite et adaptée de Bryant et al., 2014.

D'un point de vue mécanistique, l'activation par phosphorylation d'un récepteur au domaine tyrosine kinase tel que l'EGFR permet le recrutement de protéines adaptatrices à domaine SH2, telle que GRB2. GRB2 *via* 2 domaines SH3 va recruter la protéine SOS, et c'est ce complexe GRB2-SOS qui assure la connexion entre le récepteur à activité tyrosine kinase et les protéines RAS. La protéine SOS porte un domaine DH/PH (Dbl homology/Pleckstrin homology) en tandem et un domaine catalytique RAS-GEF (ou CDC25 homology domain). C'est le domaine RAS-GEF qui interagit avec RAS et déstabilise son interaction avec le GDP au profit de la fixation du GTP, ce dernier étant 10 fois plus présent dans le cytoplasme. Le domaine DH/PH, le second domaine protéique porté par SOS, régule d'une part l'activité du domaine RAS-GEF, et d'autre part régule la translocation du complexe GRB2-SOS du cytoplasme à la membrane plasmique (Ahearn et al., 2011; Yadav and Bar-Sagi, 2010) (Figure 12).

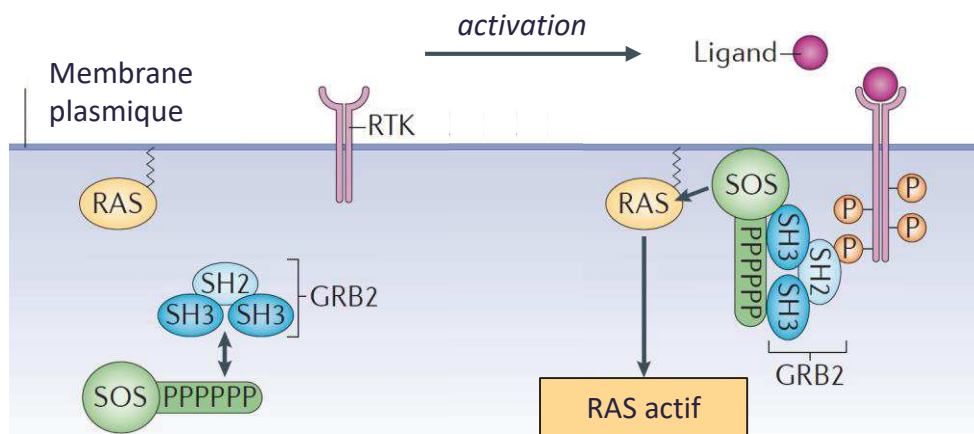


Figure 12. Mécanisme d'activation de RAS. L'activation d'un RTK par fixation d'un ligand permet le recrutement à la membrane plasmique du complexe GRB2-SOS. Ce complexe interagit d'une part avec le RTK et d'autre part avec RAS et permet ainsi son activation. Figure extraite et adaptée de Mayer, 2015

III.3 Les voies effectrices des protéines RAS

Les protéines RAS activées présentent alors une affinité plus élevée pour ses effecteurs et décrivent un nouveau réseau d'interaction avec des protéines contrôlant des processus biologiques distincts comme la transcription, l'apoptose, l'endocytose, la migration et l'organisation du cytosquelette (Figure 13) (Berndt et al., 2011; Mayer, 2015; Schubert et al., 2007).

Les voies effectrices dépendantes de RAS les plus souvent décrites sont les voies MAPK et PI3K/AKT, et leurs activations sont respectivement médiées par l'interaction avec les protéines de la famille RAF et PI3K. Comme mentionné précédemment, la voie MAPK contrôle principalement les processus de prolifération, de différenciation, de survie et de contrôle du cycle cellulaire. La voie PI3K/AKT est quant à elle impliquée dans des procédés biologiques tels que la croissance, la survie, la migration, la régulation du cycle cellulaire.

Outres ces deux voies bien décrites, il existe des effecteurs de RAS moins décrits impliqués dans l'endocytose comme RALGDS (Ral Guanine Nucleotide Dissociation Stimulator) ou RIN1 (RAS And Rab Interactor 1) ou dans la régulation du cytosquelette comme TIAM1 (T-Cell Lymphoma Invasion And Metastasis 1). Cette dernière protéine régule l'organisation du cytosquelette *via* notamment l'activation du complexe CDC42-RAC1 (Cell Division Cycle 42-Rac family small GTPase 1). Les protéines RAS sont également impliquées dans la régulation du calcium à travers ses interactions avec PLC ϵ (Phospholipase C Epsilon 1). C'est une voie de signalisation non négligeable car le calcium régule notamment l'apoptose, l'exocytose et la migration (Clapham, 2007). Plus globalement, il apparaît que les protéines RAS sont en amont de nombreuses voies de signalisation impliquées dans des processus aussi variés que la régulation de la transcription et le remodelage du cytosquelette.

Les protéines RAS interagissent avec leurs effecteurs par l'intermédiaire d'un domaine protéique RBD ou RA (RAS Association domain).

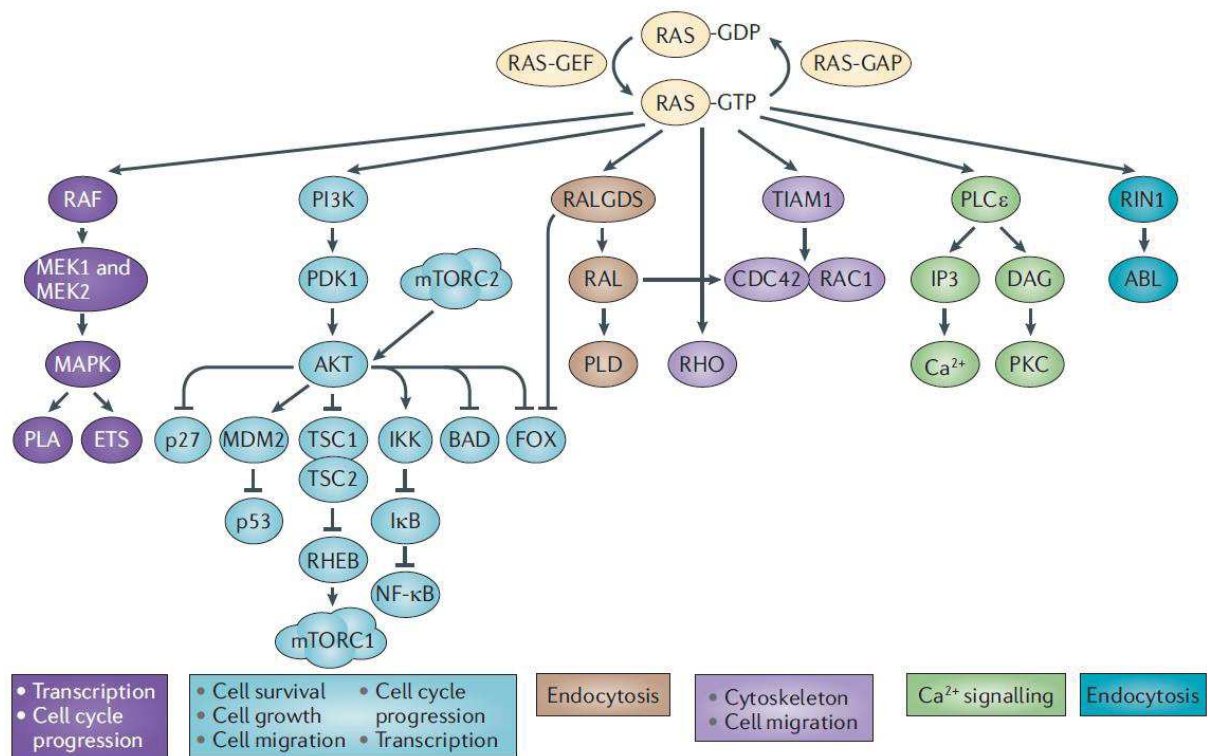


Figure 13. Interacteurs et voies de signalisation activés par l'activation de RAS. Les mécanismes cellulaires contrôlés par l'activation de RAS sont nombreux. Parmi les voies de signalisation les plus fréquemment décrites en aval de RAS mentionnons la voie MAPK (violette, à gauche) et la voie PI3K/AKT (en bleu). Figure extraite de Berndt et al., 2011.

III.4 Notions de spécificités des protéines RAS

Considérant la forte homologie du domaine de régulation (>80%), et les particularités d'adressage des isoformes de RAS, on peut s'interroger sur leurs spécificités fonctionnelles. Effectivement, peu d'articles mentionnent les spécificités fonctionnelles des isoformes de RAS, et pour cause HRAS est le premier isoforme décrit et a longtemps été considéré comme modèle d'étude des RAS. Cependant, ce modèle d'étude s'efface peu à peu pour les raisons suivantes :

- La spécificité de la région HVR dirige des spécificités d'interactions avec les protéines du réticulum endoplasmique et de l'appareil de Golgi nécessaires aux modifications post-traductionnelles et à l'adressage des protéines à la membrane.
- Les spécificités d'adressages membranaires laissent supposer des fonctions propres à chacun des isoformes, résultants de colocalisation avec des régulateurs et effecteurs spécifiques.

- Des modèles murins ont montré que la délétion de KRAS est létale pour le développement embryonnaire, à l'inverse de HRAS et NRAS (Johnson et al., 1997; Koera et al., 1997).
- Les mutations germinales de KRAS et HRAS sont retrouvées respectivement dans les syndromes Noonan et Costello, et ces RASopathies présentent des phénotypes distincts (Aoki et al., 2016).
- Indépendamment de l'expression cellulaire, certains tissus sont préférentiellement mutés pour certains isoformes. C'est le cas des mutations de NRAS dans le mélanome et KRAS dans les adénocarcinomes pulmonaires ou le cancer colorectal.

Ensemble, ces observations supportent l'hypothèse de spécificités fonctionnelles des protéines RAS. Bien que ces isoformes soient mutés dans environ 25% des cancers (Fernandez-Medarde and Santos, 2011), très peu d'études comparatives sont cependant menées.

III.5 Les mutations des protéines RAS

Les mutations activatrices des protéines RAS sont observées dans environ 25% des cancers. Il s'agit de mutations faux-sens survenant principalement aux codons 12, 13 ou 61. Toutes ces mutations affectent le domaine de régulation G, et plus précisément la boucle P (résidus 12 et 13) ou la région switch II (résidu 61). L'ensemble de ces mutations impacte l'activité GTPase intrinsèque des protéines RAS et confère une affinité d'interaction réduite aux GAPs. Ceci a pour conséquences une accumulation de ces protéines dans leur forme active, conduisant à une activation constante des voies de signalisation en aval. Les positions des mutations ainsi que l'acide aminé résultant de la mutation varie selon l'isoforme considéré ainsi que le tissu dans lequel se développe cette mutation.

III.5.a La protéine HRAS

Cet isoforme est principalement muté dans les cancers urothéliaux tous stades et grades confondus (~6% des cas) et les cancers épidermoïdes de la tête du cou (~5% des cas). Toutefois, cet isoforme est le plus faiblement muté des isoformes de RAS et également le moins impliqué dans les cancers. Il ne présente pas non plus de point de mutation préférentiel, si bien que les positions G12, G13 et Q61 sont mutées de façon équivalente (Cox et al., 2014; Goebell and Knowles, 2010).

III.5.b La protéine NRAS

NRAS est muté dans 20 à 30% des mélanomes cutanés, et est après *BRAF* le 2^e gène le plus fréquemment muté dans cette maladie. On le retrouve également muté dans certaines formes de myélomes et dans les cancers de la thyroïde (Cox et al., 2014). Les mutations de NRAS touchent principalement le résidu Q61 (~80% des cas). Les mutations Q61K and Q61R sont les plus prévalentes dans le mélanome, et constituent respectivement 34% et 68% des cas (base de données COSMIC, Forbes et al., 2017). D'autres mutations touchant les codons 12 et 13 sont également identifiées (~20% des cas). Dans le mélanome, NRAS aurait un rôle à la fois dans l'initiation et la progression tumorale. Effectivement, ce gène est muté dès les formes précoces (Platz et al., 2008), et des modèles murins ont lié le développement de métastases à la mutation de cet isoforme (Tsao et al., 2012). Plus précisément, associé à la mutation de perte de fonction de CDKN2A (Cyclin Dependent Kinase Inhibitor 2A) et/ou TP53 (Tumor Protein P53), l'activation NRAS conduit à la présence importante de métastases dans les nodules lymphatiques et d'autres organes, alors que l'activation de HRAS n'est pas métastatique. Au sein de la voie des MAPK, il a été montré par ailleurs que NRAS muté interagit préférentiellement avec CRAF (Raf-1 Proto-Oncogene, Serine/Threonine Kinase), alors que BRAF est l'interacteur préférentiel de la forme sauvage (Dumaz et al., 2006).

III.5.c La protéine KRAS

Il s'agit de l'isoforme le plus fréquemment mutée dans les cancers. La mutation de cette protéine est impliquée principalement dans les adénocarcinomes pulmonaires (~30% des cas), le cancer du côlon (~45%) et les adénocarcinomes ductal pancréatiques (~98%). Les codons mutés sont les codons 12 et 13 et l'acide aminé de substitution est généralement une cystéine (C), une aspartate (D) ou une valine (V), avec des variations selon la nature du tissu. Ainsi, les adénocarcinomes pulmonaires sont principalement mutés G12C, le cancer du côlon muté G12D, G12V ou G13D et les adénocarcinomes ductal pancréatiques mutés G12D ou G12V (Stephen et al., 2014). Il semble également que le type de mutation définit l'activation des voies de signalisation en aval. Ainsi, dans le cancer du poumon, les mutations G12C et G12V activent préférentiellement les voies RALGDS (Ral Guanine Nucleotide Dissociation Stimulator) impliquées dans l'endocytose alors que la mutation G12D active préférentiellement les voies MAPK et PI3K (Berndt et al., 2011). De plus, les tumeurs portant les mutations KRAS G12C et G12V sont plus agressives que celles présentant une mutation G12D (Ihle et al., 2012). Ceci laisse supposer des réseaux d'interactions dépendant de la nature de la mutation. Le fait que *KRAS* soit l'isoforme le plus muté dans les cancers est étonnant dans la mesure où (i) tous les isoformes sont exprimés dans les tissus adultes, (ii) tous les isoformes sont des oncogènes et (iii) la régulation des différents isoformes par le domaine G est *a priori* similaire. Le locus de *KRAS* pourrait avoir une importance dans la mesure où l'insertion de *HRAS* à la place de *KRAS* peut initier le

développement de cancer du poumon en réponse à des carcinogènes, et ceci à la même fréquence que *KRAS* (To et al., 2008). Ceci lève alors la notion de point chaud de mutation (« mutational hot-spot »). De plus, une étude de la vitesse de réparation des gènes codants pour les isoformes à la suite d'exposition à des carcinogènes a montré que le gène *KRAS* est moins efficacement réparé comparativement à *NRAS* ou *HRAS* (Feng et al., 2002).

Quel que soit le type de mutation, l'activation de *KRAS* est un facteur de mauvais pronostic associé à une non-sensibilité aux inhibiteurs de tyrosines kinase. À l'inverse des mutations de l'*EGFR* dans les adénocarcinomes pulmonaires, les mutations de *KRAS* touchent principalement les patients fumeurs et d'origine non-asiatique (Hirsch et al., 2016). Il n'existe à ce jour aucune thérapie ciblée pour traiter ces patients, si bien qu'ils seront pris en charge par chimiothérapie associée ou non à de la radiothérapie.

La mise au point de traitement ciblant les protéines RAS est particulièrement délicate car il a été démontré que l'inhibition totale des protéines RAS, indépendamment des isoformes et de leurs statuts mutationnels, est létale chez la souris adulte (Simanshu et al., 2017). Ceci amène alors à envisager des drogues exclusivement spécifiques aux protéines mutées. Du fait de leurs tailles réduites et de l'absence d'une poche de fixation à une molécule de forte affinité, les protéines RAS ont régulièrement mis en échec tout développement de nouvelles molécules (Cox et al., 2014). Toutefois, ces observations sont aujourd'hui reconsidérées car les molécules ARS-853 et SML-8-73-1 ont été identifiées très récemment comme des inhibiteurs spécifiques et irréversibles de *KRAS* G12C (Lito et al., 2016; Patricelli et al., 2016). Cette inhibition conduit à une réduction de la prolifération et l'induction de l'apoptose.

En alternative à une inhibition directe de RAS, de nombreuses études ont cherché à cibler ses effecteurs. Toutefois, l'inhibition de la voie MAPK à travers l'inhibition de la voie des RAF kinases, interacteur direct de RAS, est inefficace et conduit à l'inverse à une activation de la voie de signalisation (Holderfield et al., 2014). L'inhibition de la protéine MEK donne aussi de faibles résultats du fait de boucles de contrôle régulant positivement ERK et l'activation des récepteurs au domaine tyrosine kinase (Lake et al., 2016). Une alternative à ce traitement est d'inhiber spécifiquement les protéines portant un domaine RBD ou RBD-like, les domaines protéiques d'interaction avec RAS. Les interacteurs de RAS comme RAF, RALGDS et PI3K portent ce domaine et le rigosertib est une des molécules qui peut cibler spécifiquement les domaines RBD de ces protéines (Athuluri-Divakar et al., 2016). Ceci empêche une interaction directe avec RAS et ainsi l'activation de voies respectives des protéines RAF, RALGDS et PI3K. *In vivo*, cette molécule conduit à la réduction de la croissance tumorale. Cependant, comme c'est le cas pour le traitement de l'*EGFR*, des mécanismes de résistance pourraient réduire l'efficacité de ces traitements, et déjà la voie Hippo avec la protéine YAP1 (Yes Associated Protein 1) pourrait contourner l'inhibition de *KRAS* (Kapoor et al., 2014; Shao et al., 2014).

Nous voyons ici que l'identification de protéines effectrices spécifiques à des protéines mutées est motrice d'innovation dans la recherche de nouvelles cibles thérapeutiques, qu'il s'agisse de cibler une protéine en particulier ou une voie de signalisation dérégulées par l'acquisition de mutations.

IV. La protéine BRAF

IV.1 Structure et fonction

Le gène *BRAF* (B-Raf Proto-Oncogene, Serine/Threonine Kinase) est situé sur le bras long du chromosome 7 (7q34). D'une longueur d'environ 200 000 paires de bases, il est constitué de 18 exons. Le transcrit majeur de *BRAF* est constitué de 4560 nucléotides et celui-ci code pour une protéine de 766 acides aminés et 84kDa. Cette serine/thréonine kinase appartient à la famille RAF qui comporte 3 isoformes : ARAF, BRAF et CRAF. Ces isoformes partagent une même structure, avec trois régions conservées (CR) (Lavoie and Therrien, 2015):

- **La région CR1** est composée d'un domaine RBD d'interaction avec RAS, et d'un domaine riche en cystéines (CRD) essentiel au recrutement de RAF à la membrane plasmique.
- **La région CR2** est riche en serines et thréonines, résidus de phosphorylation essentiels à l'activité catalytique de RAF.
- **La région CR3** est le domaine catalytique à activité kinase. Il comprend une boucle d'activation (boucle A), et une boucle de fixation de l'ATP (boucle P). C'est ce domaine qui activera les effecteurs de RAF.

À l'extrémité N-terminale, BRAF présente un domaine BRS (BRAF-specific region) qui lui est propre (Figure 14).

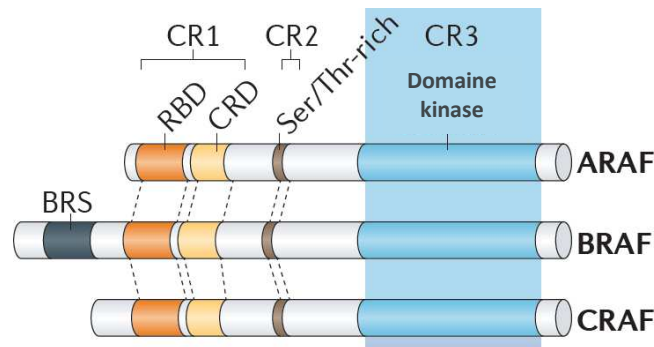


Figure 14. Structure des isoformes RAF. Les protéines RAF (ARAF, BRAF et CRAF) comportent 3 régions conservées (CR1 à 3). La région CR1 comprend le domaine RBD (RAS binding domain) et CRD (Cysteine rich domain). La région CR2 comprend un domaine riche en sérine et thréonine. Le domaine CR3 est celui de la région catalytique kinase. BRAF possède une région propre, BRS (BRAF-specific region). Figure extraite et adaptée de Lavoie and Therrien, 2015.

Bien que les protéines RAF aient été décrites au milieu des années 80, leur activation est complexe et fait intervenir de nombreuses étapes dont certaines sont encore mal comprises (Figure 15). À l'état inactif, les protéines RAF sont cytoplasmiques et dans un état replié (auto-inhibé) avec la région N-terminal qui masque le domaine kinase. Ce repliement est facilité par l'interaction avec de protéines de

régulation de la famille 14-3-3 (tyrosine 3-monooxygénase/tryptophan 5-monooxygénase activation protein). La phosphorylation de certains résidus des protéines RAF, notamment la sérine 365 pour BRAF, favorise aussi ce repliement, les interactions avec les protéines 14-3-3 et donc l'inhibition des protéines RAF (Tran et al., 2005) (Étape 1 de la figure 15).

Lorsque RAS est activé, RAF est transloqué à la membrane plasmique. Ceci est favorisé par le domaine RBD de RAF qui permet l'interaction avec RAS, et par le domaine riche en cystéine CRD qui permet l'interaction avec les phosphatidylsérines de la membrane plasmique. Ces interactions stabilisent ainsi le recrutement de RAF à la membrane. Cette interaction directe entre RAS et RAF permet de déphosphoryler les sites d'inhibition, et de lever l'inhibition par les protéines 14-3-3 (Light et al., 2002; Roy et al., 1997) (Étape 2). Ces changements structuraux permettent ensuite la dimérisation et l'activation des protéines RAF. Cette étape est peu comprise et ferait intervenir un contrôle allostérique du domaine kinase, et également d'autres protéines (Étape 3). Les RAF peuvent s'associer en hétérodimère CRAF-BRAF ou en homodimères. Les hétérodimères auraient cependant une activité supérieure. L'arginine 509 de BRAF serait à l'interface des dimères et favoriserait la transition du complexe vers une forme active (Weber et al., 2001) (Étape 4).

Un autre aspect favorable à l'activation des RAF, est l'association en nanoclusters des protéines RAS activées. Ceci a pour conséquence une augmentation de la concentration locale de RAF à la membrane, favorable à leurs dimérisations (Tian et al., 2007). Enfin, les protéines 14-3-3 interviendraient également à cette étape de dimérisation en stabilisant les dimères *via* des interactions avec de nouveaux résidus sérine phosphorylés, notamment la sérine 621 de CRAF (Noble et al., 2008) (Étape 4).

Cette dimérisation modifie alors la conformation du domaine kinase en permettant d'une part, la transactivation du dimère, et d'autre part l'orientation de certains résidus sérine et thréonine dans une conformation favorable à leur phosphorylation par d'autres protéines. La transactivation est essentielle à l'activité de RAF mais cette étape est encore peu décrite. Il semblerait qu'elle nécessite un dimère asymétrique avec un RAF donneur et un RAF receveur de phosphate (Hu et al., 2013). Cette activation peut également faire intervenir des protéines extérieures au dimère, c'est le cas des protéines kinase de la famille SRC (SRC Proto-Oncogene, Non-Receptor Tyrosine Kinase) (Girotti et al., 2015) (Étape 4).

Finalement, lorsque RAF est activé il permet principalement la phosphorylation et l'activation des protéines MEK de la voie des MAPK (Cargnello and Roux, 2011) (Étape 5). Des boucles de régulation négative interviennent cependant pour réguler l'activité de RAF (Étape 6). C'est le cas notamment de l'activation de ERK, le dernier acteur de la voie des MAPK, qui va phosphoryler RAF sur différents sites et conduire ainsi à sa dissociation de RAS et donc à son inactivation (Lake et al., 2016) (Étape 7).

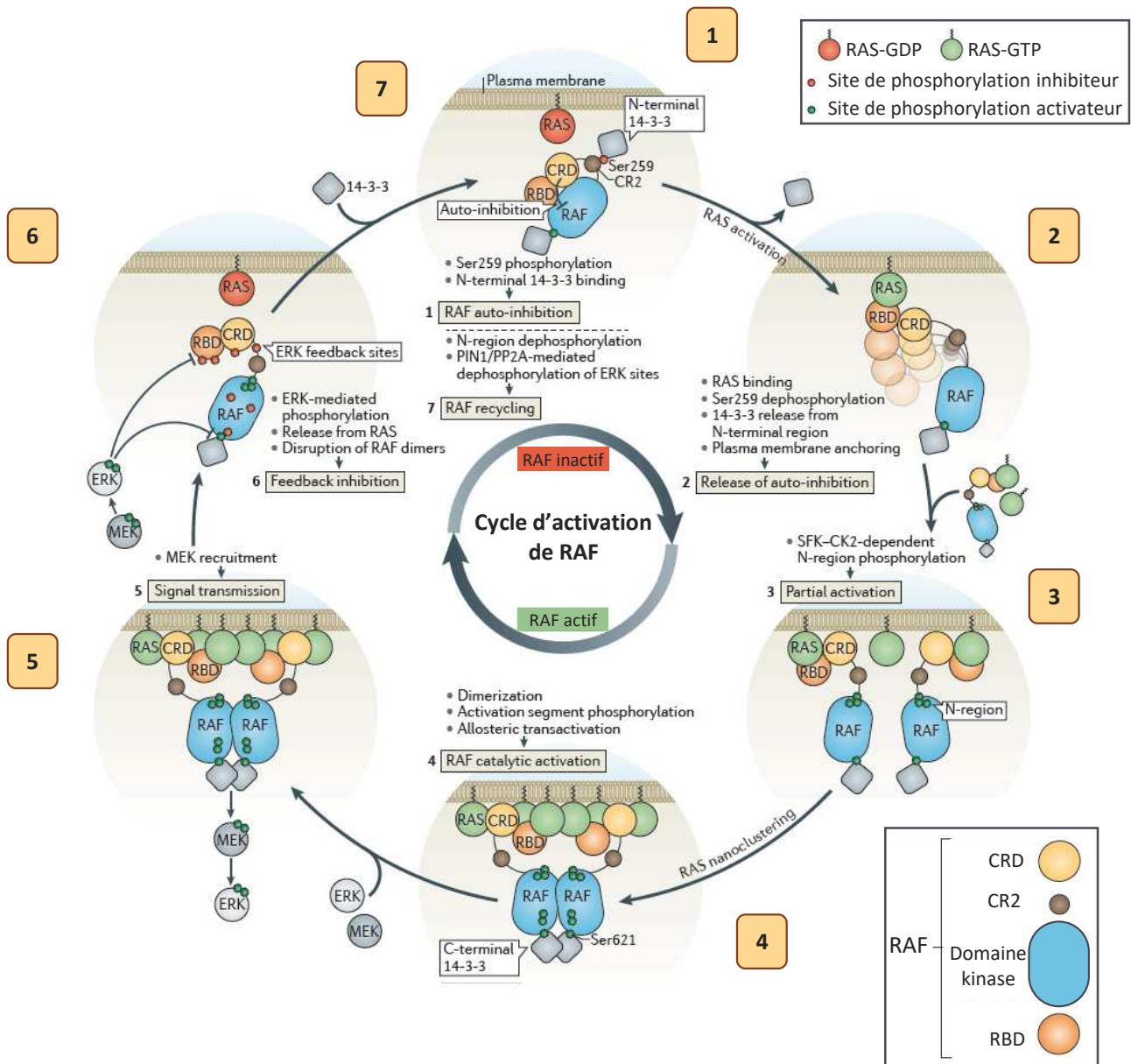


Figure 15. Cycle d'activation des protéines RAF. (Étape 1) RAF est dans un état auto-inhibé stabilisé par les protéines 14-3-3. (Étape 2) Lorsque RAS est activé, RAF interagit avec RAS par son domaine RBD. Cette interaction lève l'auto-inhibition de RAF et l'interaction avec les protéines 14-3-3. (Étape 3) L'interaction avec les kinases SFKs et CKs favorisent l'activation des protéines RAF par phosphorylation. (Étape 4) Dimérisation et transactivation de RAF. Ceci est facilité par le regroupement en nanocluster des protéines RAS et l'interaction avec les protéines 14-3-3. (Étape 5) L'activation de RAF permet le recrutement de MEK, sa phosphorylation et ainsi la transduction du signal d'activation dans la voie des MAPK. (Étape 6) Boucle de régulation négative de ERK phosphorylé. ERK phosphorylé phosphoryle RAF, dissocie les dimères de RAF, et libère RAF de RAS. (Étape 7) Recyclage de RAF par déphosphorylation des sites d'activation de RAF et déphosphorylation des sites de phosphorylation de ERK. Figure extraite et adaptée de Lavoie and Therrien, 2015.

IV.2 Les mutations de BRAF

BRAF est l'isoforme principalement muté, quel que soit le type de cancer. Avec *KRAS* et l'*EGFR* il s'agit d'un des gènes les plus mutés des adénocarcinomes pulmonaires, principalement identifié chez des patients fumeurs. La mutation majeure observée est la mutation du domaine kinase V600E (50% des cas). Ainsi, la substitution d'une valine (V) par un acide glutamique (E) mime la phosphorylation de la boucle d'activation du domaine kinase et conduit ainsi à l'activation constitutive de son activité kinase. D'autres points de mutations non négligeables sont retrouvés tels que G469A (39% des cas) et D594A (11% des cas) (Paik et al., 2011; Schubert et al., 2007). La mutation G469A quant à elle affecte le domaine de fixation à l'ATP de la boucle P. Des cas de fusions avec d'autres protéines telles que *SND1* ou *AGK* ont également été décrits, mais ceux-ci sont rares (Jang et al., 2015; Reddy et al., 2017). Chez la souris, l'expression de la mutation *BRAF* V600E dans le poumon a été montrée comme suffisante pour développer des adénomes (Dankort et al., 2007).

BRAF est le proto-oncogène principal des mélanomes où la mutation principale observée est aussi la mutation du domaine kinase V600E (90% des cas). L'efficacité des thérapies ciblées visant les mélanomes métastatiques mutés *BRAF* V600E a conduit à leur utilisation dans les adénocarcinomes pulmonaires. Le vemurafenib et le dabrafenib sont en essai clinique de phase II et les patients traités présentent une réponse positive (~42%) avec une survie sans progression de la maladie d'environ 7 mois (Planchard et al., 2016a). Du fait des boucles de rétrocontrôle, les inhibiteurs de *BRAF* induisent une augmentation de la signalisation par *RAS*, réduisant l'efficacité du traitement. C'est pourquoi, cibler *BRAF* et des protéines en aval de sa voie de signalisation constitue une stratégie de choix. La combinaison anti-*BRAF* et anti-MEK (trametinib) est en cours d'essai clinique de phase II et semble apporter des bénéfices supérieurs aux monothérapies avec une réponse globale d'environ 63% (Planchard et al., 2016b). Ces traitements sont cependant limités à la mutation V600E, et aucun essai clinique n'est en cours pour le traitement des autres mutations *BRAF* par ces mêmes thérapies (Hirsch et al., 2016).

Toutefois, comme dans le traitement de l'*EGFR*, des mécanismes de résistance aux traitements ont été décrits, notamment chez les patients atteints de mélanome, et constituent aujourd'hui la meilleure base pour appréhender et anticiper les résistances aux anti-*BRAF* dans le traitement des adénocarcinomes pulmonaires. La majorité de ces mécanismes font état d'une réactivation de la voie *MAPK*, en conséquence à une mutation activatrice de *NRAS*, *MEK1* ou *MEK2* (Wagle et al., 2014). La surexpression de la protéine *MAP3K8* (Mitogen-Activated Protein Kinase Kinase Kinase 8), régulateur de *MEK*, a également été observée, comme la surexpression de *BRAF* ainsi que son épissage alternatif (Poulikakos et al., 2011; Shi et al., 2012). Des mécanismes alternatifs, indépendant de la voie *MAPK*, ont aussi été décrits tels que l'activation de la voie *PI3K/AKT*, *HER3* ou encore *PDGFR β* (Abel et al., 2013; Villanueva et al., 2010).

V. Les gènes de fusion impliquant *ALK*

Les réarrangements de *ALK* sont courants en oncologie et il a été décrit plus d'une quinzaine de partenaires de fusion différents (Pulford et al., 2004). En plus du cancer du poumon, ceux-ci ont été décrits notamment dans les lymphomes, le cancer du côlon et le cancer du sein. La translocation EML4-*ALK* a été décrite en 2007 par une équipe Japonaise et compte pour environ 5% des mutations des adénocarcinomes pulmonaires (Soda et al., 2007).

V.1 Le récepteur *ALK*

Le gène *ALK* est situé sur le bras court du chromosome 2 (2p23). D'une longueur d'environ 730 000 paires de bases, il est constitué de 29 exons et le transcrit majeur est le variant 1. Celui-ci est constitué de 6267 nucléotides, code pour une protéine de 1620 résidus d'environ 176kDa. *ALK* est un récepteur au domaine tyrosine kinase peu décrit et peu exprimé à l'âge adulte. Il interviendrait dans le développement du système nerveux central et périphérique. *ALK* est composé de 3 domaines (Iwahara et al., 1997) :

- **Le domaine extracellulaire** permet la fixation des ligands intervenant dans la dimérisation du récepteur. Ce domaine peut être clivé et libéré dans le milieu extracellulaire, mais nous ne disposons que peu d'informations à ce sujet.
- **Le domaine transmembranaire** est en interaction avec la membrane plasmique. *ALK* ne compte qu'un seul domaine transmembranaire.
- **Le domaine tyrosine kinase** comprend une boucle P de fixation d'ATP, et une boucle A d'activation du récepteur.

L'activation d'*ALK* est sous le contrôle d'une régulation allostérique et fait intervenir des mécanismes de phosphorylation et d'autophosphorylation. Le mécanisme complet d'activation du domaine kinase est peu décrit mais, comme dans le cas d'autres récepteurs à activité tyrosine kinase, la fixation d'un ligand induirait la dimérisation et la réorientation des domaines catalytiques vers une configuration adéquate à la fixation de l'ATP et au transfert de phosphate vers une tyrosine acceptrice. La tyrosine 1278 serait essentielle à ce mécanisme (Hallberg and Palmer, 2016).

V.2 La protéine EML4

Le gène codant pour EML4 (Echinoderm Microtubule associated protein Like 4) est situé sur le bras court du chromosome 2 (2p21). D'une longueur d'environ 163 000 paires de bases, il est constitué de 23 exons et le transcrit majeur est le variant 1. Celui-ci est constitué de 5577 nucléotides et code pour une protéine de 981 résidus d'environ 109kDa. Cette protéine cytoplasmique interviendrait dans la dynamique du cytosquelette et la stabilisation des microtubules. EML4 est composé d'un domaine

hydrophobe HELP (hydrophobic EML protein), de 9 domaines WD et d'un domaine coiled-coil en N-terminal (Richards et al., 2015). Peu d'études s'intéressent à la protéine sauvage, mais elle interviendrait dans la régulation du cycle cellulaire par ses interactions avec les fuseaux mitotiques lors de la phase M. Cette interaction serait régulée par phosphorylation de résidus sérine et thréonine par l'activité de protéines kinases (Pollmann et al., 2006).

V.3 Le gène de fusion *EML4-ALK*

La littérature décrit 13 variants de fusion *EML4-ALK*. Ces variants impliquent tous des régions d'*EML4* différentes mais tous possèdent l'intégralité du domaine catalytique tyrosine kinase d'*ALK*. Au niveau de l'ADN, ceci correspond à un point de rupture intronique localisé entre les exons 19 et 20, conservant ainsi les exons 20 à 29 qui codent pour le domaine kinase (Sasaki et al., 2010). Cette translocation permet au moins 3 événements :

- Une surexpression d'*ALK* dans les tissus adultes *via* le promoteur d'*EML4*. *ALK* étant normalement peu exprimé dans les tissus adultes.
- Une relocalisation du domaine kinase d'*ALK* selon la localisation du fragment d'*EML4* transloqué.
- Une dimérisation et une activation du domaine kinase d'*ALK* *via* le domaine protéique coiled-coil d'*EML4*, en l'absence de fixation d'un ligand.

Le variant 1 comprenant les exons 1 à 13 d'*EML4* (E13:A20), et les variants 3a et 3b (E6:A20) comprenant les exons 1 à 6a ou 6b d'*EML4* sont les plus fréquents et retrouvés dans environ 33% et 29% des cas, respectivement (Figure 16) (Sasaki et al., 2010). Selon les études, ces chiffres varient jusqu'à une incidence de 58% de variants 1 et 30% de variants 3 (Maus et al., 2012).

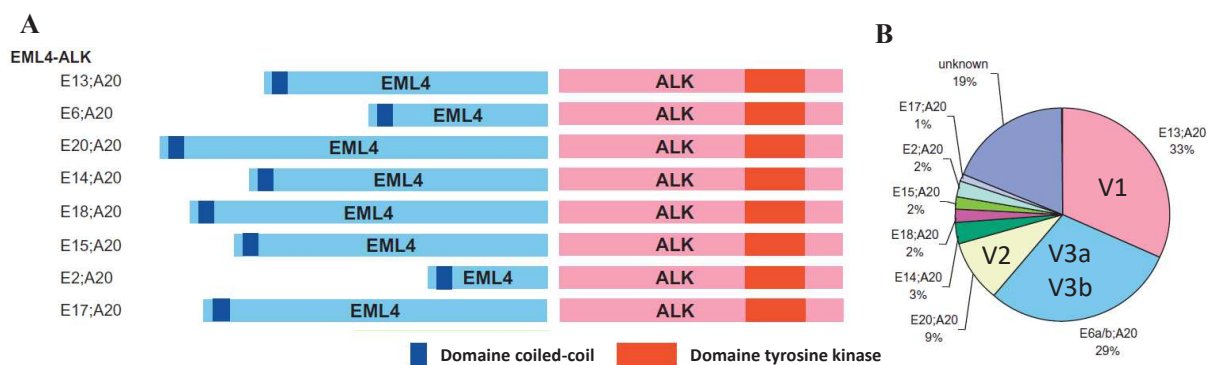


Figure 16. Variants de translocations *EML4-ALK* et incidences respectives. (A) Variants de translocation identifiés avec les exons impliqués au point de translocation. (B) Incidence respective de chacun des variants de fusion. Les variants 1 et 3a/3b sont les plus fréquents avec des incidences respectives de 33% et 29%. Figure extraite et adaptée de Sasaki et al., 2010.

C'est le fragment d'EML4 transloqué qui dirige la localisation de la protéine de fusion. Associé à l'activation du domaine kinase, ce changement de localisation permet un réseau d'activation variable selon l'environnement cellulaire. Les trois voies de signalisation les plus décrites sont les voies MAPK, PI3K/AKT et STAT (Bayliss et al., 2016). Comme décrit précédemment, celles-ci régulent notamment les mécanismes de prolifération, survie, croissance (Shaw and Solomon, 2011).

EML4-ALK est la protéine de fusion majeure des adénocarcinomes pulmonaires, mais nous pouvons mentionner également KIF5B-ALK, TFG-ALK, DCTN1-ALK, CLIP1-ALK et GCC2-ALK. Au cours de ma thèse, j'ai contribué à l'identification de ces deux derniers transcrits de fusions (Vendrell et al., 2017b - voir Publication 2).

V.4 Traitements

Le crizotinib est l'ITK prescrit lors de l'identification de cette mutation. Celui-ci cible les translocations *ALK* mais est également prescrit lors des translocations des gènes *MET* et *ROS1* qui codent tous deux pour des récepteurs à activité tyrosine kinase. Comparé aux chimiothérapies habituelles, le crizotinib augmente de 4 mois la survie sans progression de la maladie et a ainsi reçu une approbation totale de mise sur le marché par la FDA (Solomon et al., 2014). Cependant, l'ensemble des patients développent des résistances qui se caractérisent par des mutations secondaires du domaine kinase, telles que C1156Y ou G1202R (Matikas et al., 2016). Une seconde génération d'ITK ciblant *ALK* (ceritinib et alectinib) a alors été développée. Ces molécules semblent efficaces à la fois pour traiter les patients résistants au crizotinib que les patients porteurs d'une mutation *ALK* mais n'ayant encore reçu aucun traitement (Shaw et al., 2016).

Bien que l'efficacité de ces drogues contre les translocations *ALK*, autres que EML4-*ALK*, ne soient pas encore étudiée du fait de la faible prévalence de ces mutations, le crizotinib est aujourd'hui le traitement de référence pour toutes les translocations *ALK*, quel que soit le partenaire de fusion.

VI. Autres cibles thérapeutiques

Les thérapies ciblées traitant les altérations de l'EGFR, de ALK ou de BRAF ont apporté une véritable révolution dans la prise en charge des patients atteints d'adénocarcinomes pulmonaires. Cependant, quel que soit le traitement considéré, le bénéfice apporté est de l'ordre de 6 mois à 1 an de survie sans progression de la maladie, et est inévitablement marqué par un échappement thérapeutique. C'est pourquoi de nouvelles stratégies thérapeutiques sont perpétuellement recherchées. Parmi celles-ci, on peut citer d'une part les modulations de la réponse immunitaire aux cancers (e.g. PD1/PDL1) et d'autre part la recherche systématique de mutation par séquençage haut-débit. Cette dernière stratégie vise à proposer aux patients une thérapie dans le cadre d'essais cliniques basés sur les anomalies génétiques détectées dans leurs tumeurs.

VI.1 PD1 et PDL1

La suppression de l'activité anti-tumorale du système immunitaire fait partie des caractéristiques du développement cancéreux (Hanahan and Weinberg, 2011). Dans le NSCLC, il a été démontré que l'interaction entre PDL1 (Programmed death-ligand 1), surexprimé à la surface des cellules tumorales, et le récepteur PD1 (Programmed cell death protein 1) exprimé à la surface des lymphocytes T, inhibait la réponse anti-tumorale des cellules T et favorisait la progression tumorale (Pardoll, 2012). Les anticorps nivolumab, atezolizumab et pembrolizumab ciblent spécifiquement PD1 et PDL1 afin de perturber leur interaction et *in fine* activer l'activité anti-tumorale immunitaire. Ces 3 molécules sont en cours d'essai clinique et montrent toutes une efficacité significativement supérieure au traitement par chimiothérapie, et ainsi améliorent la survie sans progression de la maladie (Borghaei et al., 2015; Fehrenbacher et al., 2016; Garon et al., 2015).

VI.2 Essais cliniques

Basé sur la recherche ciblée d'anomalies génétiques fréquemment identifiées dans les cancers comme les mutations affectant par exemple les gènes *MET* ou *ROS*, les patients pourront être orientés dans des essais cliniques dits « basket ». Le nombre de patient éligible n'est toutefois que de 20% (ANSM, 2014) mais l'évolution des sensibilité de séquençage devrait permettre d'inclure plus de patient dans ces essais. De plus, on peut envisager que d'autres anomalies, si elles étaient retrouvées de façon récurrente, deviennent des cibles contre lesquelles de nouvelles molécules pourraient être développées, représentant donc une source d'innovation majeure. Une autre source d'innovation est le séquençage entier de l'exome des patients (WES, Whole Exome Sequencing) afin d'identifier des mutations récurrentes et potentiellement actionnables par des drogues seules ou en combinaison.

Plus globalement, on voit que le grand défi dans la lutte contre le cancer est d'identifier de nouvelles cibles thérapeutiques pour permettre à tout patient de bénéficier d'une médecine personnalisée et améliorer la survie sans progression de la maladie. Ceci comprend l'identification des nouvelles cibles pour traiter des cancers pour lesquels aucune thérapie ciblée n'est disponible, et identifier des nouvelles cibles pour contrer les mécanismes de résistance. De plus, considérant le coût élevé de ces thérapies un autre enjeu se situe dans l'identification de marqueurs de réponse aux thérapies ciblées car tous les patients n'y répondent pas de façon strictement identique. Définir le portrait mutationnel des tumeurs mais aussi mieux comprendre à l'échelle des protéines les partenaires engagés dans ces mécanismes sont donc des points essentiels pour une meilleure compréhension de la maladie, et être à même d'identifier des opportunités thérapeutiques.

C. Étude des interactions protéiques

I. Dynamique des interactions protéiques

Dans les sections précédentes, j'ai abordé les mécanismes d'activation des protéines RAS suite à la fixation du ligand sur l'EGFR, et comment cette activation va activer en cascade les autres partenaires de la voie de signalisation des MAPK, par exemple. Les interactions protéine-protéine (IPP), qu'elles soient transitoires ou permanentes et dépendante du compartiment subcellulaire, sont au cœur de l'activation de ces voies de signalisation et de ces effets cellulaires. Nous avons vu également que dans ce processus, toute mutation au niveau des protéines clef est susceptible de modifier la structure et la conformation des protéines en question et de modifier ainsi les réseaux d'interactions et d'activations en aval. Ces mécanismes d'activation sont des systèmes complexes avec une dynamique spatio-temporelle importante dans lesquelles les protéines s'interconnectent entre elles pour au final contribuer au phénotype de la cellule (Figure 17). Dans son ouvrage *General system theory* publié en 1967, Ludwig von Bertalanffy explicitait déjà l'importance d'une approche systémique des réseaux biologiques: « The properties and modes of action of higher levels are not explicable by the summation of the properties and modes of action of their components taken in isolation. If, however, we know the ensemble of the components and the relations existing between them, then the higher levels are derivable from the components ».

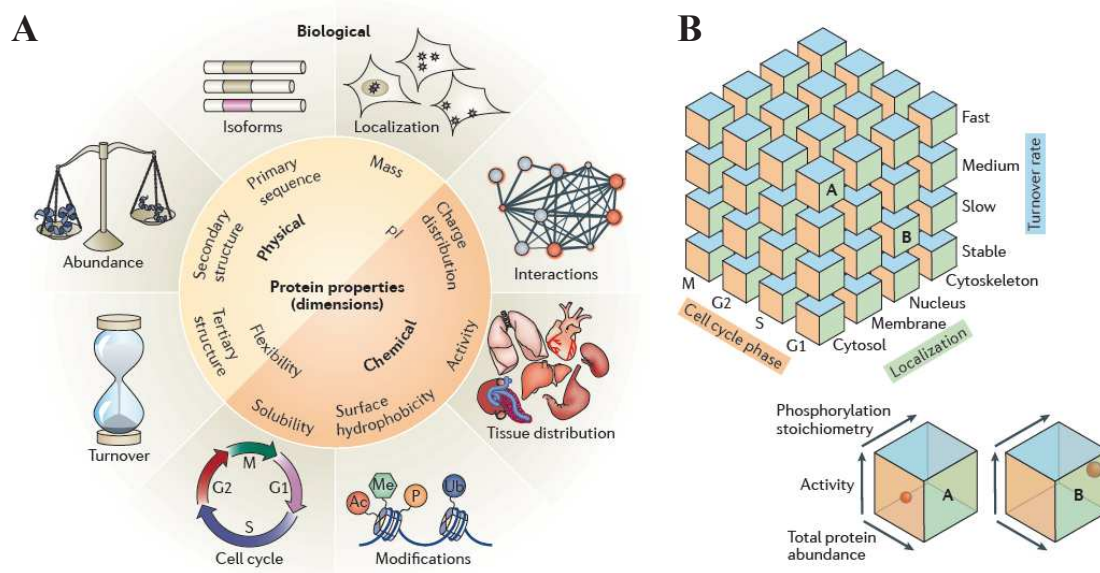


Figure 17. Approche multidimensionnelle du protéome. (A) Le protéome se définit par la complémentarité des propriétés de chacune des protéines le composant, notamment abondance, localisation, réseaux d'interactions. (B) Série de cubes empilés représentant la dynamique du protéome. Les cubes sont orientés selon les dimensions de phase du cycle cellulaire, localisation et taux de renouvellement. En phase G1, une protéine (représentée par la sphère rouge) peut être localisée dans le cytosol et le noyau, et présenter un taux de renouvellement différent dans ces compartiments respectifs. Des propriétés supplémentaires telles que l'activité, l'abondance et l'état de phosphorylation participent également à ces différences. Figure extraite de Larance and Lamond, 2015.

La construction de ces réseaux d'interactions est complexe et représente un véritable défi à la croisée entre différentes disciplines, notamment protéomique, biologie cellulaire et bio-informatique. Cette complexité s'explique par la nature et la diversité des IPP. Ces dernières peuvent être :

- Obligatoires (l'interaction est nécessaire à la stabilité des protéines et/ou à la fonction), ou non obligatoires (bien que fonctionnelle indépendamment, l'interaction des protéines sera responsable d'un effet cellulaire).
- Transitoires (par exemple en s'associant ou se dissociant en fonction du contexte cellulaire, d'une modification post-traductionnelle, de la localisation intra-cellulaire des protéines) ou permanentes comme dans le cas du complexe protéique constituant le protéasome et dans lequel la délétion d'une des protéines du complexe conduit à son inactivation complète et la mort cellulaire.
- Directes ou indirectes dans un complexe.

Cet aspect multidimensionnel des réseaux d'interactions conduit à des défis technologiques important dans l'identification des IPPs et la construction des interactomes. Pour répondre à ce challenge plusieurs techniques ont été développées. Parmi elles, les techniques classiques et conventionnelles, comme les purifications d'affinité (AP) et les doubles hybride chez la levure (*Yeast two hybrid*, Y2H), ne donnent qu'un reflet partiel des IPPs. En effet, ces techniques *in vitro* n'identifient que les interactions fortes en dehors de tout contexte cellulaire. Il s'agit d'approches à bas ou moyen débit, capables de mettre en évidence des interactions binaires ou de complexe à forte affinité, sans tenir compte (ou difficilement) de la dynamique spatio-temporelle. Cependant, ces techniques ont été la base de l'identification de nombreuses IPPs dont beaucoup sont encore référencées dans les bases de données. Ces dernières années ont été marquées par le développement de techniques d'identification des interactions *in vivo* de haute résolution permettant de tenir compte de la dimension spatio-temporelles de ces interactions. Ces techniques, comme le BioID (proximity-dependent biotinylation identification) et l'APEX (Engineered ascorbate peroxidase), sont regroupées sous le terme de marquage de proximité (ou *proximity-dependent labeling*) et ouvrent de nouvelles perspectives dans l'identification des interacteurs.

Notons que dans les études d'IPP, la protéine dont on veut définir le réseau d'interaction est parfois appelée protéine appât et, ses partenaires ou interacteurs sont appelées protéines proies.

II. Identification des interactions protéine-protéine *in vitro*

II.1 Purification par affinité (AP)

L'approche de purification par affinité est la technique la plus populaire pour purifier une protéine d'intérêt, et avec elle, ses partenaires. Dans cette technique, une protéine endogène ou étiquetée (tag HA, GFP, FLAG), est isolée en utilisant un anticorps fixé à la protéine A ou G et couplées à des billes de sépharose. L'utilisation de conditions de lyse non dénaturantes permet d'isoler la protéine d'intérêt tout en préservant les interactions avec ses partenaires protéiques. À la suite de lavages, l'identification des complexes protéiques se fait par analyse en spectrométrie de masse après séparation éventuelle en gel d'électrophorèse (Figure 18).

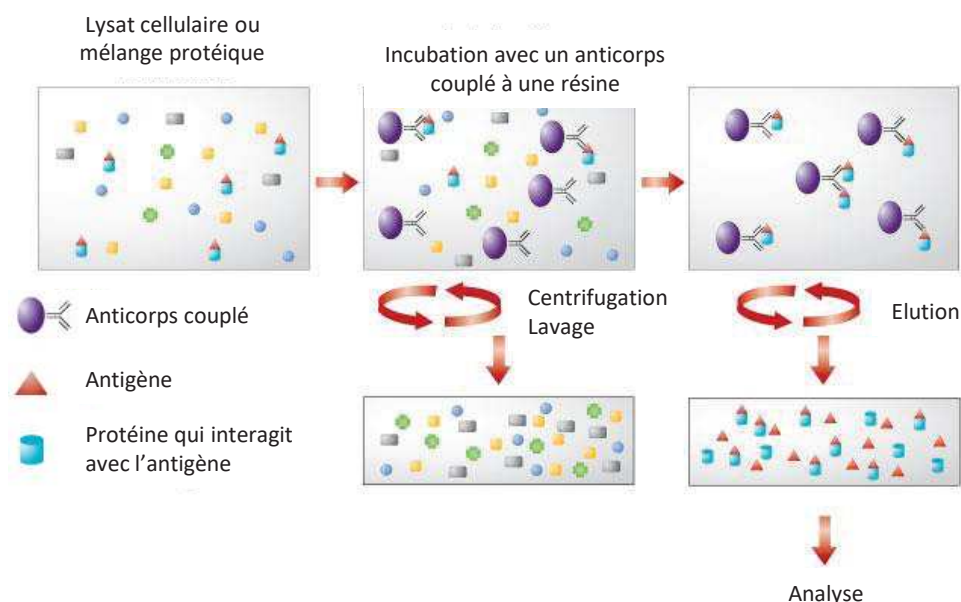


Figure 18. Principe généraux des purifications d'affinité. Un lysat protéique ou un mélange de protéine est mis au contact de bille couplées à des anticorps. Les anticorps permettent d'isoler la protéine d'intérêt et avec elle ses partenaires. Les protéines non retenues sur les billes sont éliminées par série de centrifugations et lavages. Les protéines isolées par les anticorps sont ensuite dénaturées puis éluée. L'extrait protéique peut être analysé par spectrométrie de masse. Figure extraite et adaptée de Thermofisher.com (Juillet 2017).

Cette méthode présente un certain nombre d'avantages :

- Identification des interactions directes, en condition quasi-physiologiques.
- Possibilité d'étude des interactions au cours du temps, en réponse par exemple à un stimulus.

Parmi les inconvénients, on peut noter que cette méthode ne permet d'identifier que des interactions fortes et qu'elle ne permet pas d'analyser des protéines faiblement solubles comme les protéines membranaires. Cette méthode nécessite aussi des anticorps spécifiques. De plus, devoir préserver les

interactions protéiques tout au long de la purification, fait qu'il n'est pas possible de laver intensément les échantillons et est donc source de faux-positifs. Pour pallier à ce dernier inconvénient, des méthodes alternatives par des systèmes de purifications successives ont été développées comme l'approche TAP-TAG. D'autres méthodes comme le pull-down permettent de purifier les complexes protéiques par affinité en l'absence d'anticorps.

II.2 Méthode de double hybride chez la levure

Aussi appelée Yeast-2-Hybrid, cette méthode utilise la complémentarité fonctionnelle basée sur la modularité des facteurs de transcription, pour activer un gène rapporteur et produire un signal quantifiable de l'interaction protéique. La protéine appât est fusionnée avec un domaine de fixation à l'ADN alors que la protéine proie est fusionnée à un domaine d'activation de la transcription. Si les deux protéines interagissent entre elles, alors le complexe interagit avec le promoteur du gène rapporteur et déclenche sa transcription. Le facteur de transcription est généralement Gal4 et le gène rapporteur LacZ (Mehta and Trinkle-Mulcahy, 2016) (Figure 19).

Cette technique présente l'avantage d'être peu coûteuse et relativement facile à mettre en œuvre. Elle met en évidence exclusivement des interactions permanentes et/ou transitoires. Elle présente néanmoins plusieurs inconvénients liés à la technique elle-même comme un taux de faux-positifs important, notamment lorsque la protéine appât possède un domaine d'activation de la transcription, ou quand cette protéine interagit de manière non spécifique avec un grand nombre d'interacteurs. De plus, les protéines étudiées sont exprimées dans un système biologique ne permettant pas de mettre en place toutes les modifications post-traductionnelles qui sont parfois nécessaires à l'interaction. Comme pour les techniques de purification par affinité, l'étude des interactions des protéines membranaires n'est pas envisageable.

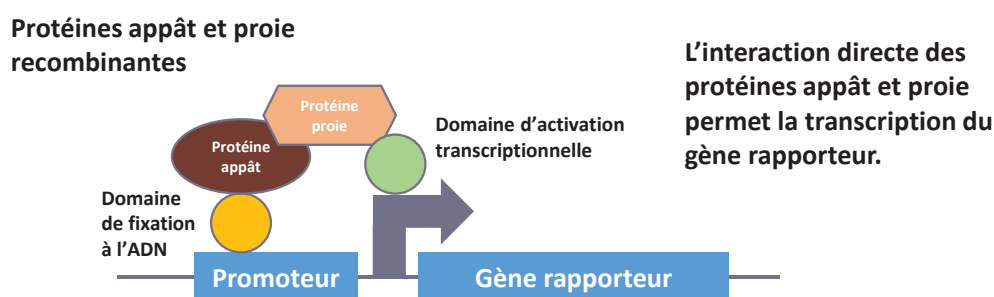


Figure 19. Principe général du Yeast-2-Hybrid. Une protéine appât (marron) est fusionnée à un domaine de fixation à l'ADN (orange) et une protéine proie (beige) est fusionnée à un domaine activateur de transcription (vert). Si les deux protéines interagissent entre elles, alors le complexe interagit avec le promoteur du gène rapporteur et déclenche sa transcription. Figure reproduite de Mehta and Trinkle-Mulcahy, 2016.

III. Identification des interactions protéine-protéine *in vivo*

Alors qu'un grand nombre de méthodes ont été développées pour identifier les IPPs, la plupart sont limitées à l'identification d'interactions fortes, à un instant donné et dans un contexte cellulaire qui n'est pas toujours le plus pertinent. De plus, les protéines membranaires sont essentielles aux fonctions cellulaires et cette localisation rend difficile la caractérisation de leur interactome par les techniques classiques. Pour répondre à ces limitations, deux techniques *in vivo* de marquage de proximité ont été développées, nommément BioID et APEX.

Le marquage de proximité est basé sur la fusion d'une protéine d'intérêt à une enzyme, BirA pour le BioID et l'ascorbate peroxidase pour l'APEX. Cette enzyme génère un produit réactif qui diffusera autour de la protéine d'intérêt et réagira avec les protéines environnantes. Ceci conduit à un marquage localisé, covalent et permettra de purifier facilement les partenaires par purification d'affinité avant de les identifier par spectrométrie de masse (MS). De plus, le marquage covalent par la biotine ($K_d \sim 10^{-15}$) permet d'utiliser des conditions agressives de lyse et de lavage afin d'augmenter le taux de pureté des échantillons. Ces techniques *in vivo* sont donc particulièrement adaptées à l'étude des protéines membranaires et limitent les faux-positifs.

III.1 BioID: Proximity-dependant Biotin Identification

III.1.a BioID conventionnel

Développée en 2012 (Roux et al., 2012), cette technique utilise la capacité naturelle des biotines ligases à biotinyler les protéines. BirA est la biotine ligase la plus étudiée et celle-ci est exprimée par *Escherichia coli*. Cette enzyme est un co-facteur de l'acétyl-CoA carboxylase et a pour fonction de biotinyler une de ses sous-unités, et de réguler transcriptionnellement l'opéron biotine. En présence d'ATP et de biotine, BirA génère l'espèce réactive biotinoyl-5'-AMP (biotine-AMP). BirA retient la biotin-AMP en son site actif et facilitera son transfert sur un résidu lysine d'une protéine cible. Afin de permettre un marquage plus large, BirA a été modifié et le variant R118G, noté BirA*, a une affinité plus faible pour la biotine-AMP, ce qui permet son relargage et la biotinylation des protéines environnantes dans un rayon de 10 nm (Kim et al., 2014). Lorsqu'une protéine cible est fusionnée à BirA*, l'expression de cette protéine chimère dans un milieu complétement avec de la biotine permet un marquage des interacteurs sur ~24 h, ceci quelle que soit sa localisation. En ce sens, le BioID est une technique qui donne un historique des interactions protéiques dans la cellule et dans des conditions proches des conditions physiologiques (Figure 20).

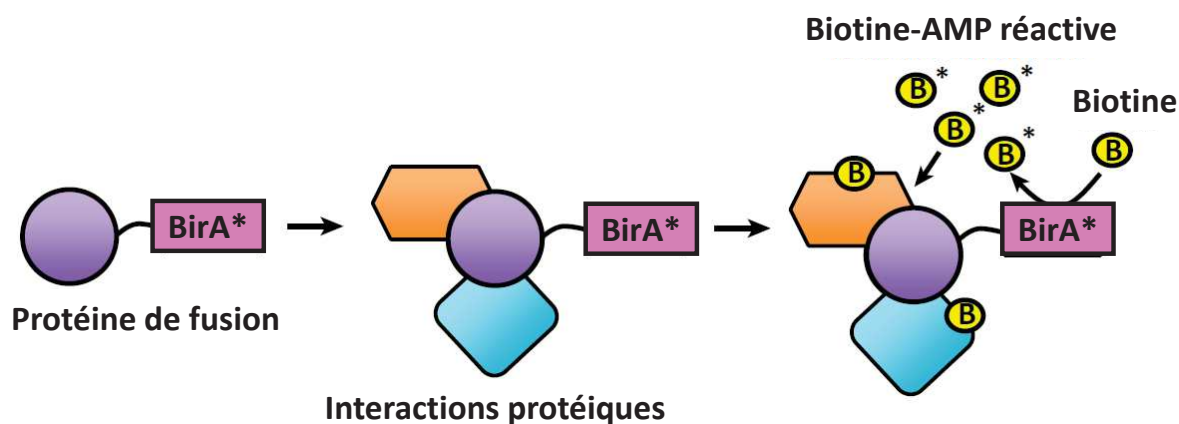


Figure 20. Principe général du BioID. Une protéine d'intérêt est fusionnée à l'enzyme BirA*. Lorsque cette protéine de fusion est exprimée et lorsque le milieu est supplémenté en biotine, BioID active la biotine en biotine-AMP. La biotine activée diffuse alors autour de BioID et autour de la protéine d'intérêt pour marquer de façon covalente les interacteurs proches (<10 nm). Figure extraite et adaptée de Lönn and Landegren, 2017.

Comme toutes techniques, celle-ci présente néanmoins des défauts. L'absence de lysine ou plus généralement d'amine primaire à la surface des protéines rend impossible le marquage à la biotine et par conséquent réduit le nombre de protéines identifiées. De plus, BirA* est une protéine relativement grosse par rapport à des étiquettes HA par exemple (32kDa *versus* 1 kDa), ce qui peut modifier le comportement de la protéine d'intérêt. Ce dernier point motive d'ailleurs la recherche actuelle de biotine ligase de taille réduite et à activité similaire ou supérieur. Un autre inconvénient est que la biotinylation peut aussi modifier le comportement de la protéine d'intérêt et de ses interacteurs, ce qui peut éloigner la cellule de ses conditions physiologiques. Précisons également qu'il s'agit d'un marquage de proximité à 10 nm qui va révéler des interactions directes ou indirectes.

Comparé aux techniques conventionnelles, le BioID offre de nombreux avantages mais ne rend compte de l'historique d'interaction que d'une seule protéine. Le split-BioID a alors été développé à partir du même principe pour permettre l'identification des interacteurs d'un complexe protéique binaire au cours du temps (Schopp et al., 2017).

III.1.b Split-BioID

Dans cette approche du split-BioID, BirA* est partagé en deux fragments, N- et C-terminal. Une première protéine est fusionnée au fragment N-terminal et la seconde est fusionnée au fragment C-terminal. Lorsque les deux protéines interagissent ensemble, le fragment N-terminal s'associe au C-terminal et ensemble génèrent une protéine BirA* fonctionnelle. Ainsi ce n'est que lorsque le complexe protéique est formé que la biotine-AMP est produite et marque les interacteurs du complexe. Ceci permet un marquage temporel et spatial plus restreint mais cette méthode garde néanmoins les désavantages du BioID conventionnel cités plus haut (Figure 21) (Schopp et al., 2017).

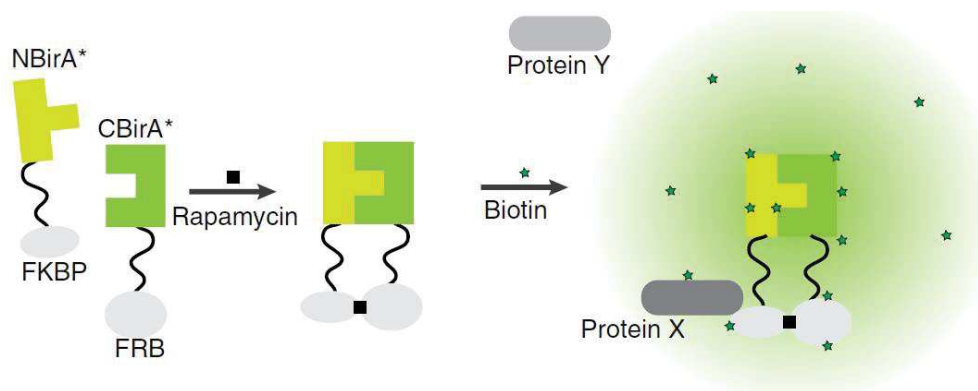


Figure 21. Principe général du split-BioID. Une première protéine (ici FKBP) est fusionnée au fragment N-terminal de BirA* et une seconde protéine (ici FRB) est fusionnée au fragment C-terminal de BirA*. Lorsque les deux protéines interagissent ensemble, les fragments N- et C-terminal s'associe et BirA* est fonctionnel. BirA* active la biotine présente dans le milieu qui diffusera autour du complexe protéique et marquera de façon covalente les protéines à proximité (<10 nm). Dans cet exemple, la rapamycine permet l'interaction de FKBP avec FRP. Figure extraite de Schopp et al., 2017.

III.2 APEX: Engineered ascorbate peroxidase

Développée en 2013 (Rhee et al., 2013), cette technique est sur le principe très proche du BioID. Elle utilise une autre enzyme, l'ascorbate peroxydase monomérique, dérivée du soja pour marquer avec de la biotine activée les protéines présentes à proximité d'une protéine d'intérêt. En présence de peroxyde d'hydrogène, cette enzyme oxyde le biotine-phénol en biotine-phénoxy. Ce dernier est un radical à demi-vie courte (<2,5 ms) qui réagira de façon covalente avec les acides aminés environnants riches en électrons tels que les tyrosines, les tryptophanes, les histidines et les cystéines, et ceci dans une région localisée (<20 nm). La différence principale avec le BioID est la durée d'incubation. Alors qu'elle est de l'ordre de 24h pour le BioID, l'incubation avec le peroxyde d'hydrogène est de l'ordre de la minute voire de quelques heures. L'APEX2 (APEX A134P), a été développée pour augmenter l'efficacité catalytique de l'enzyme et à des niveaux d'expression de la protéine plus faibles. Cette méthode permet donc d'avoir une image *in vivo* et à un instant donné de l'environnement protéique d'une protéine d'intérêt ou, lorsque l'APEX est fusionné à un motif d'adressage, d'un compartiment donné (Figure 22).

À l'inverse du BioID, l'APEX fournit un historique des interactions moins large et n'est donc pas adapté à l'identification d'interactions transitoires. De plus, la biotine-phénol et l'exposition au peroxyde d'hydrogène peut modifier la physiologie de la cellule par l'induction de voie de signalisation en réponse au stress oxydatif ou modifier la dynamique des organelles et les interactions protéiques dans leur ensemble. Comme pour le BioID, les acides aminés qui vont être biotinylés doivent être à la surface des complexes protéiques, si bien que certains pourront ne pas être marqués et donc identifiés.

Notons que le BioID et l'APEX sont des techniques transposables chez des organismes multicellulaires comme la souris et la drosophile (Chen et al., 2015; Dingar et al., 2015).

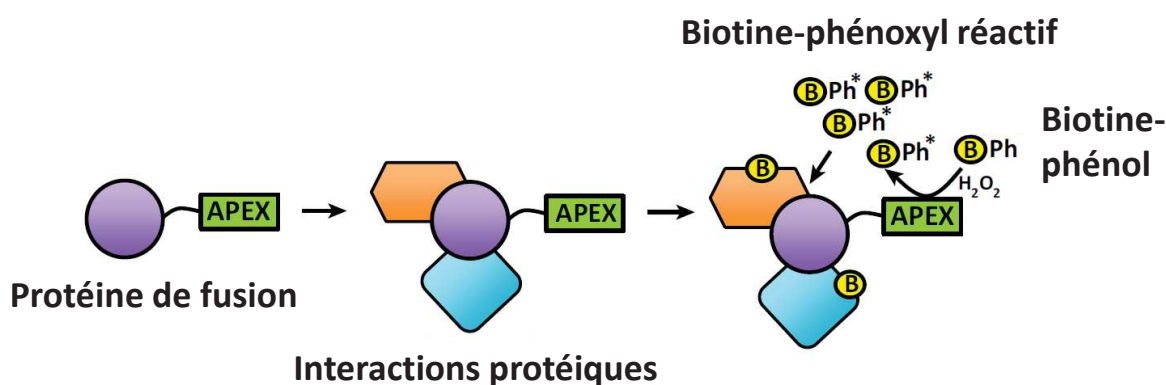


Figure 22. Principe général de l'APEX. Une protéine d'intérêt ou un motif d'adressage est fusionné à l'ascorbate peroxydase. Lorsque cette fusion de protéine est exprimée, et lorsque le milieu est supplémenté en biotine-phénol, l'APEX active la biotine-phénol en biotine-phénoxy. La biotine activée diffuse alors autour de l'APEX et autour de la protéine d'intérêt ou région cellulaire ciblée pour marquer de façon covalente les interacteurs proches (<20nm). Le temps d'incubation est de l'ordre de la minute. Figure extraite et adaptée de Lönn and Landegren, 2017.

IV. Méthode de validation des interactions

Une fois les interactions caractérisées *in vivo* ou *in vitro*, il est parfois nécessaire de les valider en utilisant d'autres techniques qui sont généralement à bas-débit. Le PLA (Proximity Ligation Assay) et le FRET (Förster Resonance Energy Transfer) font partie de ces techniques à bas-débit.

IV.1 PLA : Proximity Ligation Assay

Cette technique, parfois appelée *in situ* Proximity Ligation Assay (isPLA), est une méthode de détection par fluorescence d'interaction entre deux protéines sur des coupes tissulaires de cellules fixées ou congelées. Dans cette méthode, les deux protéines, dont on suppose l'interaction, sont ciblées par des anticorps primaires. Puis l'échantillon est mis au contact d'anticorps secondaires conjugués à des oligonucléotides. Si les deux anticorps primaires sont à proximité (<40 nm), alors les deux anticorps secondaires le seront aussi et leurs oligonucléotides respectifs pourront s'hybrider l'un sur l'autre pour former un brin d'ADN circulaire. Après ligation, cet ADN circulaire est amplifié par PCR (Polymerase Chain Reaction) en cercle roulant générant un ADN simple brin composé de copies répétées du brin complémentaire de l'ADN circulaire. Les nucléotides étant marqués par des fluorophores, ceci permet leur intégration au brin d'ADN généré par PCR. Finalement, la taille du produit réactionnel final est de l'ordre du μm , ce qui permet sa visualisation par microscopie à fluorescence. Cette technique est adaptée à la détection d'interactions protéiques impliquant des protéines membranaires, nucléaires ou

cytosoliques (Figure 23) (Bobrich et al., 2013; Söderberg et al., 2006). Comme mentionné, le PLA convient à la validation d'interactions protéiques sur coupes tissulaires fixées ou congelées, et échantillons de culture cellulaire. Précisons toutefois que l'utilisation de tissus nécessite de disposer d'anticorps primaires ciblant la protéine endogène et compatibles en immunohistochimie. En revanche, l'utilisation d'échantillons de culture cellulaire rend possible l'utilisation d'anticorps ciblant des étiquettes HA ou FLAG, par exemple.



Figure 23. Principe général du PLA. Des anticorps primaires produit chez des espèces différentes ciblent des protéines. Puis des anticorps secondaires couplés à des oligonucléotides ciblent spécifiquement les anticorps primaires. Si les anticorps secondaires sont à proximité (<40 nm), les oligonucléotides s'hybrident. Après ligation, les oligonucléotides sont utilisés en tant qu'amorce de polymérisation en cercle roulant. Des fluorophores s'associent au brin d'ADN généré et permettent la détection des interactions protéiques par microscopie à fluorescence. Figure extraite de Bobrich et al., 2013.

IV.2 FRET : Förster Resonance Energy Transfer

Le FRET est une technique qui permet de détecter l'interaction directe entre deux protéines par émission de fluorescence *in vivo*. Plus précisément, dans des modèles de culture cellulaire la protéine appât est fusionnée à un fluorophore donneur et la protéine proie à un fluorophore accepteur. Lorsque ces deux protéines sont exprimées dans les cellules et interagissent entre elles, le fluorophore donneur sous stimulus lumineux va être excité et va transférer son énergie par résonance au fluorophore accepteur. Ce dernier va alors émettre à son tour un signal fluorescent à une longueur d'onde spécifique qui pourra être lu par l'expérimentateur (Lönn and Landegren, 2017) (Figure 24).

L'avantage de cette technique est qu'elle se réalise *in vivo* et qu'elle permet une approche dynamique des études des PPIs. En effet, elle permet par lectures successives de suivre l'association et la dissociation d'un complexe selon les conditions environnementales. Par la spécificité du transfert d'énergie cette technique réduit également les faux-positifs. L'inconvénient est cependant de devoir exprimer des protéines de fusion dont l'association et le niveau d'expression peut différer des conditions endogènes. Un autre inconvénient est qu'elle n'est pas applicable à des coupes tissulaires.

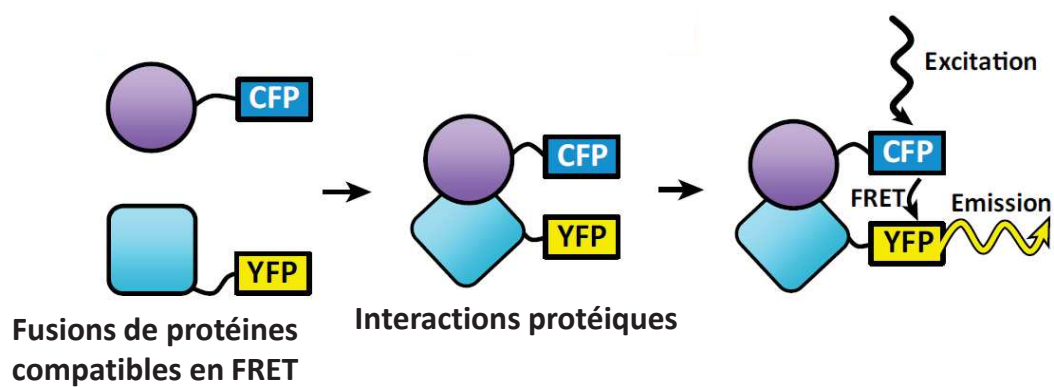


Figure 24. Principe général du FRET. Une première protéine est couplée à un fluorophore donneur (ici CFP, Cyan fluorescent protein) et une seconde protéine couplée à un fluorophore accepteur (ici YFP, Yellow fluorescent protein). Lorsque les deux protéines interagissent ensemble, et sous stimulus lumineux à longueur d'onde définie, le fluorophore donneur est excité et transfère une partie de son énergie au fluorophore accepteur. L'émission de lumière par le fluorophore accepteur sera acquise par l'expérimentateur à longueur d'onde définie. Figure extraite et adaptée de Lönn and Landegren, 2017.

Objectifs de thèse

Le cancer du poumon est la première cause de mortalité liée au cancer en France et dans le monde. Il s'agit d'un cancer de mauvais pronostic, diagnostiqué généralement aux stades tardifs III ou IV et dont les opportunités thérapeutiques sont peu nombreuses avec une survie à 5 ans faible, respectivement de 6 et 1%. Ce cancer comprend différents types histologiques, parmi lesquels les adénocarcinomes sont les plus fréquents et comptent pour 40% des diagnostics.

L'analyse génétique par séquençage des tumeurs des patients en stade IV permet de mettre en évidence des altérations génétiques récurrentes, et généralement mutuellement exclusives, au niveau des gènes *KRAS*, *EGFR*, *BRAF* et *ALK*. L'identification de mutations permet dans le cadre de tests compagnons de décider de l'éligibilité du patient aux thérapies ciblées lorsqu'elles sont disponibles.

Bien que ces thérapies ciblées représentent une véritable révolution dans la prise en charge des patients, une majorité de patients ne peuvent bénéficier de ces traitements car ils ne présentent pas au niveau tumoral de mutation activatrices actionnables (absence d'altérations *EGFR*, *BRAF* et *ALK*, présence d'une mutation *KRAS*...). De plus, dans l'éventualité où ils peuvent être traités par une molécule orale, le bénéfice observé reste malheureusement de courte durée et tous finiront à terme par échapper au traitement. Ces échappements ont plusieurs explications, et dans plus de la moitié des cas, ils sont liés à l'apparition de nouvelles mutations. C'est dans ce cas de figure que de nouveaux inhibiteurs ont été spécialement développés, comme par exemple les ITKs de 3^e génération qui ciblent spécifiquement la mutation de résistance T790M de l'EGFR. Malgré ces traitements ciblés, un échappement thérapeutique apparaît à nouveau, plus ou moins rapidement, cette fois encore caractérisé principalement par l'apparition de mutations spécifiques.

La prise en charge des patients présentant un adénocarcinome pulmonaire reste encore dépendante, (i) pour les sujets traités par ITKs, de la mise en évidence exhaustive des mécanismes cellulaires de résistances et (ii) pour les patients non éligibles aux thérapies ciblées, de la recherche de nouvelles solutions thérapeutiques. C'est dans ce double contexte que s'inscrit mon travail de thèse d'université.

Ainsi, afin de mieux comprendre quels sont les mécanismes moléculaires associés aux phénomènes de résistances thérapeutiques observées en clinique, et dans l'objectif de proposer de nouvelles voies thérapeutiques pour les patients non éligibles aux thérapies ciblées, nous avons étudié, au niveau protéique, l'impact des mutations sur les réseaux d'interactions protéiques. La caractérisation des réseaux d'interactions des protéines cibles en fonction du contexte mutationnel ouvre la voie à la mise en évidence des bouleversements cellulaires induits par ces mutations et l'identification de nouvelles stratégies thérapeutiques ciblant spécifiquement les interacteurs identifiés, seuls ou en association avec des thérapies actuellement utilisées.

Le premier objectif de ma thèse a donc porté sur la caractérisation des réseaux d'interactions des protéines de la famille RAS. Sachant que cette famille est constituée principalement de 3 isoformes (HRAS, NRAS et KRAS) qui partagent un degré important d'homologie, nous nous sommes tout d'abord intéressés à l'étude des réseaux d'interactions protéiques portés par ces trois isoformes et la mise en évidence des interactions protéiques spécifiques à l'isoforme KRAS. Le second objectif de ma thèse a consisté à étudier la dynamique des réseaux d'interactions protéiques dans un contexte mutationnel fréquemment retrouvé dans les adénocarcinomes pulmonaires, à savoir la mutation *KRAS* G12C, et pour lequel les stratégies par traitement oral sont absentes.

Nous avons par la suite appliqué cette même démarche à d'autres oncodrivers, à savoir NRAS, EGFR, BRAF et ALK, en contexte sauvage et en contexte muté. Plus précisément, je me suis intéressé à l'effet des mutations suivantes :

- *NRAS* : Q61K
- *EGFR* : L858R, Del.19, T790M, T790M-L858R, T790M-Del.19
- *BRAF* : V600E
- *EML4-ALK* : variant 1, variant 3a.

Enfin, nous avons étudiés dans une série d'analyses préliminaires l'effet thérapeutique potentiel lié à l'inhibition de l'expression des interacteurs nouvellement identifiés.

Résultats

A. Introduction

I. Approche expérimentale

L'approche expérimentale employée est résumée dans la figure 25. Elle est divisée en trois parties: (i) générer le réseau d'interaction protéique, (ii) analyser ce réseau *in silico* et par des expériences *in vivo* et *in vitro* et réciproquement, (iii) enfin générer de nouvelles connaissances et identifier des cibles thérapeutiques potentielles dans le traitement des adénocarcinomes pulmonaires. Plus précisément, l'analyse comprend deux volets complémentaires. Le premier volet est *in silico* afin d'annoter les réseaux d'interactions en termes de voies de signalisation, mécanismes biologiques, localisations, domaines protéiques significativement identifiés. Le second volet concerne l'analyse *in vitro* ou *in vivo* et comprend toutes les expériences de validation des interactions par des approches complémentaires (immunoprécipitation, PLA) et les expériences d'analyse fonctionnelle de ces interacteurs. L'aller-retour entre ces analyses *in silico* et *in vivo* ou *in vitro* permet d'affiner par itération la compréhension des mécanismes biologiques impliqués, qu'ils soient fondamentaux ou plus appliqués.

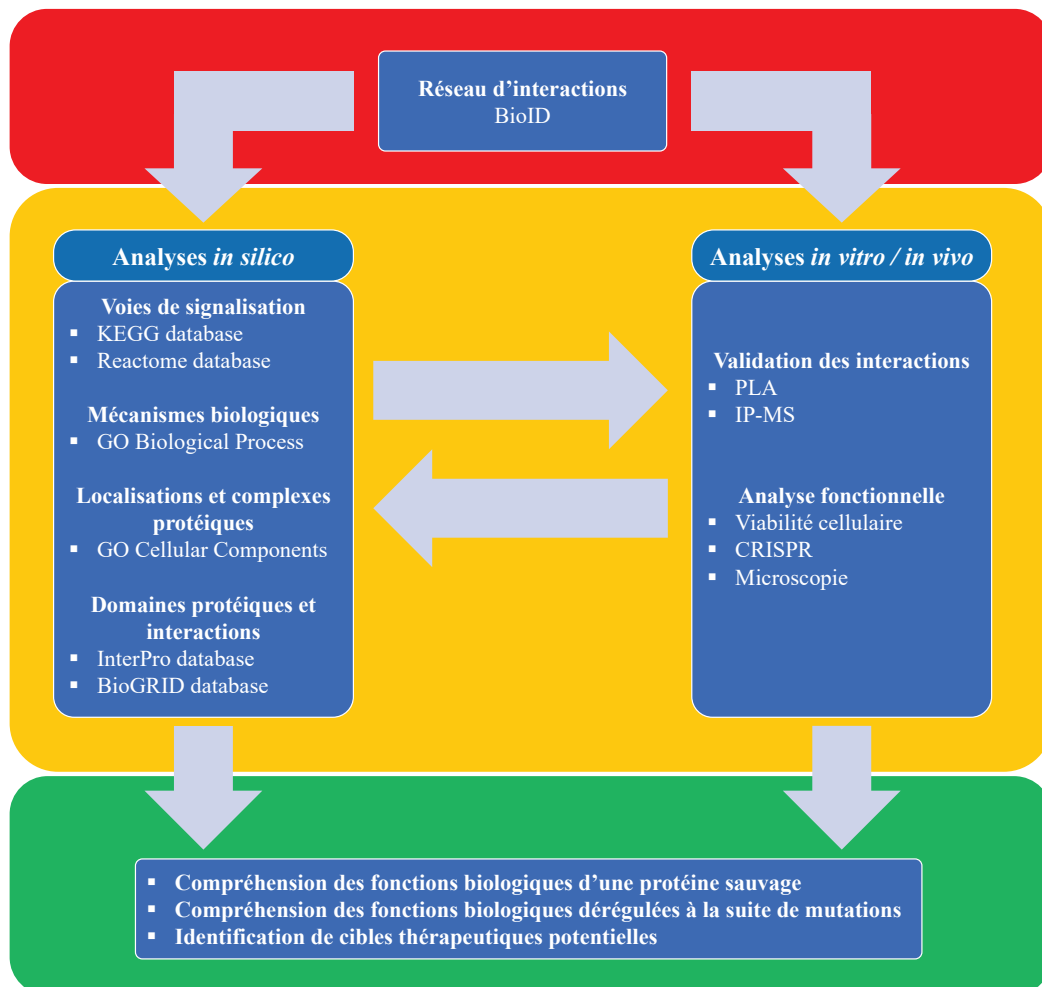


Figure 25. Méthodologie d'analyse des réseaux d'interactions protéiques. (Section rouge) Génération du réseau d'interaction protéique par la méthode du BioID. **(Section orange)** L'analyse du réseau d'interaction comprend une analyse *in silico* et une analyse *in vitro* ou *in vivo*. L'analyse *in silico* utilise notamment les bases de données KEGG, GO et InterPro afin d'annoter les interacteurs identifiés. L'analyse *in vitro* ou *in vivo* permet de valider les interactions identifiées en BioID et de les analyser en termes de fonctions. Ces deux types d'analyses sont complémentaires. **(Section verte)** Finalement, l'aller-retour entre ces analyses permet de mieux comprendre les fonctions biologiques d'une protéine et ses dérégulations lors de mutation, et finalement l'identification de cibles thérapeutiques potentielles. BioID, Proximity-dependant Biotinylation Identification; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology; BioGRID, Biological General Repository for Interaction Datasets; PLA, Proximity Ligation Assay; IP-MS, Immunoprecipitation-Mass Spectrometry; CRISPR, Clustered Regularly Interspaced Short Palindromic Repeats.

II. Expériences de BioID

Le BioID est une méthode *in vivo* d'identification des interactions protéiques. Pour cela, elle nécessite donc d'établir des lignées cellulaires exprimant stablement et de façon inductible la protéine d'intérêt. Après induction et marquage avec la biotine, les cellules sont lysées et les protéines marquées sont purifiées et analysées par MS. Enfin, les données de MS sont traitées statistiquement par une suite de logiciels afin d'identifier les interacteurs significatifs.

Le plan expérimental du BioID avec un temps estimé par étape et pour une construction est le suivant (Figure 26) :

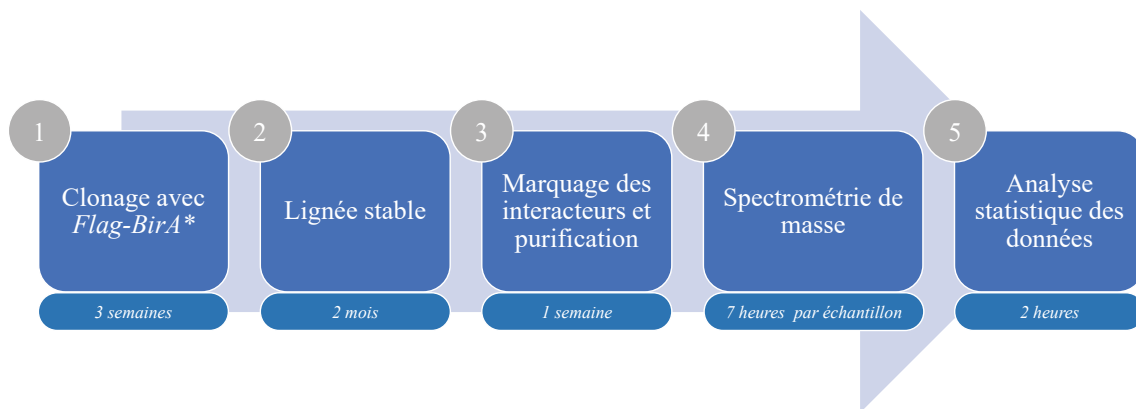


Figure 26. Plan expérimental du BioID. (1) Cette méthode débute par le clonage du gène d'intérêt pour exprimer la protéine d'intérêt fusionnée avec l'enzyme BirA*, elle-même étiquetée par la séquence Flag. (2) Les lignées stables sont ensuite générées et amplifiées. (3) La protéine de fusion est alors exprimée pendant 24h dans un milieu supplémenté en biotine, puis les cellules récoltées, les protéines extraites et les interacteurs marqués à la biotine purifiés. (4) Ces interacteurs sont analysés par spectrométrie de masse et (5) ensuite analysés statistiquement afin d'identifier les interacteurs significatifs. Cette méthode comporte de nombreuses étapes et il faut compter au minimum 3 mois pour caractériser le réseau d'interaction d'une protéine d'intérêt.

II.1 Clonage des gènes d'intérêts

Les gènes d'intérêt ont été amplifiés et clonés en amont ou en aval de la séquence codant pour Flag-BirA*. Ceci permet d'orienter la localisation de l'enzyme BirA* en N- ou C-terminale. Dans le cas des protéines membranaires EGFR et ALK, seule l'insertion en C-terminale a été réalisée. Pour les protéines cytoplasmiques et non ancrées à la membrane (BRAF, EML4, EML4-ALK, Δ EML4, et Δ ALK), les deux insertions en N- et C-terminale ont été faites pour étendre les identifications d'interactions. Pour les protéines RAS (HRAS, NRAS et KRAS), BirA* a été positionné seulement en N-terminale car l'extrémité C-terminale est essentielle à l'adressage membranaire de ces protéines.

Selon le modèle cellulaire utilisé, le vecteur de destination des gènes d'intérêt est différent. L'ensemble des clonages réalisés ainsi que le modèle cellulaire associé est résumé dans le tableau 4.

II.1.a Vecteur pcDNA5

Dans un premier temps, tous les réseaux d'interactions ont été établis dans un modèle cellulaire de choix pour ce type d'étude : la lignée cellulaire HEK293 (lignée embryonnaire de rein humaine) et plus particulièrement la lignée HEK293 T-REx flp-in (notées HEK293 ci-après). Cette lignée possède une séquence de recombinaison homologue FRT (Flippase Recognition Target) permettant l'intégration des gènes cibles en un site unique d'intégration. Le vecteur possède la séquence homologue FRT qui assurera l'intégration dans les HEK293 et place le gène sous contrôle d'un promoteur CMV (CytoMegalovirus) inductible à la tétracycline. Ce plasmide porte un gène de sélection à l'hygromycine permettant de sélectionner les cellules ayant intégrées le plasmide. Afin d'intégrer le vecteur pcDNA5 dans les HEK293, ce vecteur est co-transfecté avec le vecteur pOG44 codant pour la Flp recombinase.

II.1.b Vecteur pSTV1

La lignée cellulaire humaine de poumon d'origine non-cancéreuse BEAS2B a été utilisée pour constituer les réseaux d'interactions de KRAS. Dans ce cas, les intégrations des gènes cibles sont réalisées par des particules virales et ceci a nécessité d'utiliser le vecteur lentiviral pSTV1 (laboratoire dirigé par le Dr. AC Gingras). Le gène d'intérêt est également sous le contrôle du promoteur CMV mais inductible avec de la doxycycline. Il possède également une séquence codant pour la protéine YFP (Yellow Fluorescent Protein) permettant de trier les cellules ayant intégrée le transgène. À noter qu'à la différence des HEK293, le site d'insertion n'est ici pas maîtrisé.

Pour produire les lentivirus, les cellules 293T sont co-transfectées avec le vecteur pVSVG (vecteur d'enveloppe), le vecteur pPAX2 (vecteur de packaging) et le vecteur pSTV1. Les virus sont ensuite produits sur 24h, puis récoltés à partir du surnageant de culture cellulaire.

Tableau 4. Tableau récapitulatif des clonages réalisés et des modèles cellulaires correspondants. Les cases notées d'une croix (x) indiquent le clonage réalisé. N-ter, position de l'étiquette Flag-BirA* en N-terminale de la protéine ; C-ter, position de l'étiquette Flag-BirA* en C-terminale de la protéine. Les fragments tronqués des fusions EML4-ALK v1 et EML4-ALK v3a sont notés d'un delta (Δ).

	Vecteur Etiquette Flag-BirA*	pcDNA5		pSTV1		Modèle cellulaire
		N-ter	C-ter	N-ter	C-ter	
EGFR	WT		x			HEK293
	L858R		x			HEK293
	Del.19		x			HEK293
	T790M		x			HEK293
	T790M-L858R		x			HEK293
	T790M-Del.19		x			HEK293
KRAS	WT	x		x		HEK293 (pcDNA5) BEAS2B (pSTV1)
	G12C	x		x		HEK293 (pcDNA5) BEAS2B (pSTV1)
NRAS	WT	x				HEK293
	Q61K	x				HEK293
HRAS	WT	x				HEK293
BRAF	WT	x	x			HEK293
	V600E	x	x			HEK293
ALK	WT		x			HEK293
EML4	WT	x	x			HEK293
EML4-ALK v1	-	x	x			HEK293
EML4-ALK v3a	-	x	x			HEK293
Δ ALK	-	x	x			HEK293
Δ EML4 v1	-	x	x			HEK293
Δ EML4 v3a	-	x	x			HEK293

II.2 Établissement des lignées stables et sélection

II.2.a La lignée HEK293

Comme indiqué, la lignée cellulaire HEK293 a été utilisée dans un premier temps pour établir les réseaux d'interactions protéiques. En effet, ces cellules présentent plusieurs intérêts comme :

- Une croissance rapide avec un temps de génération d'environ 24 heures.
- Une efficacité de transfection élevée (un test de transfection de la GFP (Green Fluorescent Protein) indique un taux de transfection de 90%).
- Des outils technologiques nombreux.
- Un site d'intégration de pcDNA5 maîtrisé *via* le site FRT, et donc un niveau d'expression similaire d'une lignée transfectée stablement à une autre.
- L'expression d'un large panel de gènes.

La facilité d'utilisation des HEK293 associée à l'expression d'un large panel de gènes placent ces cellules comme une plateforme de caractérisation des interactions protéiques. Il s'agit par ailleurs de la lignée de référence dans le laboratoire du Dr. B. Raught pour l'analyse en BioID.

Après transfection de pcDNA5 et pOG44, les cellules ayant intégrées stablement pcDNA5 sont sélectionnées en présence d'hygromycine (200µg/mL). Cet évènement est rare (entre 20 et 50 clones pour 10 millions de cellules) mais tous les clones sont identiques dans la mesure où toutes les cellules possèdent une seule séquence FRT qui est localisée au même endroit dans le génome.

II.2.b La lignée BEAS2B

M'intéressant au cancer du poumon, et pour comparer le réseau d'interaction obtenu dans les HEK293 à celui obtenu dans un modèle cellulaire plus proche de mon sujet d'étude, j'ai décidé d'utiliser la lignée humaine de poumon d'origine non-cancéreuse BEAS2B. Le fond génétique sauvage de cette lignée est particulièrement intéressant pour identifier des dérégulations d'interactions du seul fait de la mutation d'un transgène. En effet, l'hétérogénéité génétique des lignées cancéreuses montre que malgré la mutation du même oncogène, par exemple *KRAS*, chacune de ces lignées présentent un profil mutationnel unique avec des propriétés uniques. Réaliser alors le BioID dans une lignée cancéreuse n'assure donc pas d'identifier des dérégulations d'interactions du seul fait de la mutation de l'oncogène d'intérêt, quand bien même celui-ci serait le principal. Cette approche perd d'autant plus de sens lorsqu'il s'agit de définir le réseau d'interaction de la protéine sauvage, alors qu'il s'agit précisément de la première étape vers l'identifier des dérégulations oncogènes.

Les virus portant le transgène *KRAS* WT ou *KRAS* G12C sont utilisés pour infecter les cellules BEAS2B à une MOI (Multiplicité d'Infection) inférieure à 1, c'est-à-dire moins d'un virus par cellule. Les cellules positives exprimant la protéine YFP sont ensuite triées par cytométrie en flux. Cette étape a permis d'enrichir jusqu'à 90% les cellules positives. Les cellules triées comptaient donc environ 10% des cellules négatives, et c'est l'inconvénient des sélections par fluorescence car il n'y a aucune pression de sélection. Toutefois, j'ai évalué le pourcentage de fluorescence des cellules triées après une quinzaine de divisions et les cellules positives étaient toujours enrichies à ~90%.

II.3 Marquage des interacteurs et purification

Pour les HEK293, l'expression de la protéine d'intérêt fusionnée à BirA* est induite par l'ajout de tétracycline dans le milieu de culture. La tétracycline se complexe à son répresseur et lève l'inhibition du promoteur CMV. En plus de la tétracycline, le milieu est enrichi en D-biotine à une concentration finale de 50 µM. La biotine entre dans les cellules par diffusion facilitée et est utilisée en tant que substrat par BirA*. Comme expliqué précédemment, la biotine activée diffusera alors autour de BirA* et de la

protéine d'intérêt et marquera les interacteurs dans un rayon d'environ 10 nm, et ceci tout au long de l'induction de la protéine pendant 24 heures.

Pour la lignée BEAS2B, l'expérience est sensiblement identique à la différence que la protéine d'intérêt est induite par l'ajout de doxycycline.

Quelle que soit la lignée considérée, le nombre de cellules nécessaire à la préparation des échantillons à l'analyse en spectrométrie de masse est élevée (~100 millions de cellules). Deux réplicas biologiques (Pool A et B) sont réalisés par protéine d'intérêt pour sélectionner les interacteurs les plus pertinents.

Un jour après l'induction, les cellules sont récoltées puis lysées à l'aide d'un tampon de lyse stringent qui permet notamment de solubiliser les protéines membranaires, sans considérer la préservation des interactions protéiques (RIPA). En effet, le marquage à la biotine n'est pas sensible à ce tampon de lyse. Les protéines marquées à la biotine seront ensuite purifiées à l'aide de billes de streptavidine-sepharose, puis intensément lavées pour réduire les faux positifs. Les protéines purifiées et fixées aux billes sont ensuite digérées par la trypsine. Les peptides générés sont récoltés puis lyophilisés et ensuite analysés par MS.

II.4 Spectrométrie de masse

Les échantillons lyophilisés sont repris dans un tampon contenant 0,1% d'acide formique, puis 1/6^e de l'échantillon est analysé. Afin de minimiser les faux-positifs de l'analyse de MS, chaque échantillon est analysé deux fois (réplicas techniques). Au final, pour chaque construction ce sont donc deux réplicas biologiques et deux réplicas techniques qui sont réalisés par protéine d'intérêt. En plus des échantillons d'intérêts, des échantillons contrôles sont préparés à partir des cellules exprimant seulement Flag-BirA* pour évaluer les faux-positifs. Flag-BirA* est une protéine qui peut générer un réseau d'interactions qui représenterait ainsi de faux-positifs pour la protéine d'intérêt fusionnée à Flag-BirA*. Les échantillons d'intérêt sont donc comparés à 8 analyses de l'échantillon contrôle.

Brièvement, les échantillons sont analysés par chromatographie liquide couplée à la spectrométrie de masse en tandem (LC-MS/MS). Les peptides sont séparés sur une colonne C18 par chromatographie liquide avant d'être analysés sur un spectromètre de masse de type quadripôle (Q Exactive™ HF, Thermo Fisher). Après un scan MS, un scan MS/MS est réalisé sur les 20 ions les plus intenses (acquisition DDA pour Data-Dependent Acquisition). Les peptides et protéines sont ensuite identifiés en utilisant la base de données humaine RefSeq v45 avec la suite de logiciel TPP/iProphet (Pedrioli, 2010).

II.5 Analyse statistique des données

Une fois les peptides présents dans l'échantillon identifiés et rapportés à la protéine correspondante, il faut déterminer si l'interacteur identifié est spécifique. Pour cela, l'outil statistique SAINT (Significance analysis of interactome) a spécialement été développé (Choi et al., 2011) pour comparer statistiquement les interacteurs détectés avec les protéines cibles à ceux détectés avec la protéine Flag-BirA* seule.

SAINTE utilise la statistique de Bayes. Il s'agit de convertir le nombre de peptides obtenus pour une protéine i (interacteur), purifiée à partir d'une expérience cherchant à définir les interacteurs de la protéine j (cible), en une probabilité d'interaction vraie ou fausse entre i et j et en prenant en compte les peptides obtenus pour les expériences contrôles (Flag-BirA* seul). Au final, on obtient une probabilité d'interaction, la valeur SAINT ou SAINT score. Le nombre de peptides étant dépendant de la taille de la protéine, l'algorithme SAINT normalise le nombre de peptides obtenus par la longueur des protéines et le nombre total de peptides obtenus pour l'échantillon. Enfin, cette valeur SAINT peut être utilisée pour calculer un taux de faux-positifs ou FDR (False Discovery Rate) et un FDR Bayesian (BFRD). Un SAINT score de 0,9 est équivalent à un FDR de 2%. Afin d'augmenter la robustesse des résultats, les interacteurs d'une cible sont comparés à 8 échantillons contrôle.

Une fois l'analyse SAINT réalisée, un autre paramètre supplémentaire pris en compte pour éliminer les faux-positifs est le nombre de peptides uniques. En effet, l'analyse SAINT prend en compte le nombre de peptides totaux obtenus pour une protéine et non le nombre de peptides uniques. Étant donné que l'identification de plusieurs peptides uniques augmente la confiance dans l'interaction identifiée, les interacteurs pour lesquels seulement 1 peptide unique a été identifié par injection, et ceci pour les 4 répliques, sont retirés de la liste des interacteurs significatifs.

Enfin, de façon empirique, certaines classes de protéines sont identifiées de façon significative quelles que soient les protéines d'intérêts par rapport aux échantillons contrôles. Afin de disposer de données représentatives de la protéine étudiée, ces classes de protéines sont retirées manuellement. Il s'agit de protéines capables d'interagir avec une multitude de protéines comme les protéines ribosomales MRPL (Mitochondrial Ribosomal Protein L), RPS (Ribosomal Protein S) et les HRNP (Heterogeneous Nuclear Ribonucleoprotein).

Le tableau 5 ci-après montre un exemple de la représentation des données finales de BioID. Le tableau obtenu correspond au réseau d'interactions protéiques décrit par une protéine d'intérêt pendant 24 heures.

Tableau 5. Exemple de représentation des données de BioID. Les lignes jaunes indiquent la protéine d'intérêt, ici KRAS, et l'enzyme BirA*. Les autres lignes correspondent aux protéines identifiées dans l'échantillon. Les colonnes comprennent le nom des protéines (Gene Name), leur nom complet (Gene full name) et les identifiants génomiques (Gene ID) et protéiques (Protein ID). Les colonnes Top 2 controls présentent les deux plus hautes valeurs de peptides identifiés parmi les 8 échantillons contrôles (Flag-BirA* seul) pour chacune des protéines. Les colonnes Pool A et Pool B comprennent le nombre total de peptides identifiés pour les réplicats biologiques A et B. Les colonnes tech#1 et tech#2 indiquent le nombre total de peptides identifiés pour les réplicats techniques de chacun des réplicats biologiques. La colonne Total correspond à la somme des peptides identifiés pour les 4 réplicats. La colonne SAINT correspond au score SAINT (probabilité d'interaction) calculé pour chacun des interacteurs. Si le score SAINT n'est pas significatif, aucune valeur n'est indiquée. La colonne Total Unique Peptide totalise le nombre de peptides uniques identifiés pour l'ensemble des 4 analyses. Les cases grisées sont un nombre total de peptides identifiés correspondant à un seul peptide unique.

Gene Name	Gene full name	Gene ID	Protein ID	Top 2 controls	KRAS WT BioID				Total	SAINT	Total Unique Peptides	
					pool A		pool B					
					tech#1	tech#2	tech#1	tech#2				
BirA*	Bifunctional ligase/repressor BirA	948469	16131807	2526	2538	1618	1839	1689	1808	6954	-	100
KRAS	Kirsten rat sarcoma viral oncogene homolog	3845	15718761	0	1	199	237	283	269	988	-	48
IRS4	insulin receptor substrate 4	8471	4504733	29	32	380	460	566	558	1964	1,00	161
EPB41L3	erythrocyte membrane protein band 4.1-like 3	23136	32490572	78	84	238	329	430	408	1405	1,00	139
DSG2	desmoglein 2	1829	116534898	16	23	210	242	271	285	1008	1,00	124
ADD3	adducin 3 (gamma)	120	62912453	13	14	186	245	238	244	913	1,00	124
CTNND1	catenin (cadherin-associated protein), delta 1	1500	146231938	11	13	184	220	257	252	913	1,00	132
DST	dystonin	667	34577049	21	23	159	191	224	211	785	1,00	227
ENO1	enolase 1, (alpha)	2023	4503571	66	75	150	169	195	188	702		63
SSR1	signal sequence receptor, alpha	6745	169404009	0	0	6	7	4	4	21	1,00	5

III. Traitement des réseaux d'interactions protéiques

Une fois les réseaux d'interactions obtenus, leurs analyses sont partagées en deux catégories complémentaires (Figure 25). D'une part, une analyse *in silico* qui permet d'annoter les interacteurs identifiés selon les voies de signalisations, fonctions biologiques, localisations, domaines protéiques significativement identifiés et les interactions protéiques déjà rapportées dans la littérature. D'autre part, une analyse *in vivo* ou *in vitro* qui permet, à partir d'expériences complémentaires, de valider les interactions identifiées en BioID et de les analyser fonctionnellement.

III.1 Analyses *in silico*

Ces analyses ont principalement été réalisées en utilisant l'outil d'annotation DAVID 6.8 (Database for Annotation, Visualization and Integrated Discovery, mis à jour 2017) (Huang et al., 2009). Cet outil informatique est une plateforme d'analyse intégrant les bases de données Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) et InterPro (ressource d'analyse et de classification des séquences protéiques). Les données sont également comparées à des bases de données référençant les interactions connues (BioGRID, HPRD, BIND), disponibles sur le site de NCBI (National Center for Biotechnology Information).

III.1.a Gene Ontology

Les protéines identifiées ont été annotées en utilisant l'outil informatique Gene Ontology (GO). Il permet d'annoter les gènes et protéines correspondantes selon leurs fonctions moléculaires (Molecular Functions, MF), leur processus biologiques (Biological Process, BP) et les composants cellulaires dans lesquels ils sont impliqués (Cellular Components, CC). Les termes GO sont hiérarchisés en niveaux, des plus généraux (niveau 1) aux plus spécifiques (niveau 5). Pour avoir une vision précise des mécanismes et localisation biologiques impliquées, je me suis situé à un niveau 5 en sélectionnant les termes enrichis avec une P-value associée de $\leq 5\%$. Pour chacune des analyses le génome a été utilisé en tant que liste contrôle. Les listes de termes enrichis comportant souvent des termes redondants, celles-ci sont filtrées à l'aide de Revigo afin de disposer de listes de termes concises et représentatives des données (Supek et al., 2011).

Dans le cas de la comparaison des isoformes de RAS, j'ai cartographié la localisation globale des interacteurs identifiés en cherchant spécifiquement les termes GO CC suivants :

- Plasma membrane (GO:0005886)
- Cell junction (GO:0030054)
- Cytosol (GO:0005829)
- Cytoskeleton (GO:0005856)
- Endoplasmic reticulum (GO:0005783)
- Golgi apparatus (GO:0005794)
- Mitochondrion (GO:0005739)
- Lysosome (GO:0005764)
- Endosome (GO:0005768)
- Nucleus (GO:0005634)

III.1.b Kyoto Encyclopedia of Genes and Genomes

Les voies de signalisation et voies métaboliques enrichies ont été identifiées en utilisant la base de données KEGG (Kyoto Encyclopedia of Genes and Genomes). Comme dans le cas des analyses GO, les termes enrichis sélectionnés sont ceux associés à une P-value $\leq 5\%$. Les termes faisant référence à des maladies comme Non-small cell lung cancer (hsa05223), Pancreatic cancer (hsa05212) ou Melanoma (hsa05218) n'ont pas été sélectionnées car ceux-ci ne sont pas assez informatifs sur les voies de signalisation ou voie métaboliques.

III.1.c InterPro

InterPro est une ressource d'analyse et de classification des séquences protéiques. Les domaines protéiques enrichis parmi les interacteurs identifiés ont été sélectionnés en utilisant une P-value seuil de 5%. Ma thèse étant centré autour de protéines impliquées dans des voies de signalisation, j'ai particulièrement été attentif aux domaines protéiques tyrosine kinase, PH, RBD, SH2 et SH3.

III.1.d National Center for Biotechnology Information

NCBI est un site ressource sur le génome qui condense notamment les connaissances actuelles sur le génome, le transcriptome, et le protéome. Ce site a été utilisé pour comparer les interacteurs identifiés en BioID à ceux déjà référencés dans les bases de données BioGRID (Biological General Repository for Interaction Dataset), HPRD (Human Protein Reference Database) et BIND (Biomolecular Interaction Network Database).

L'identification combinée des voies de signalisation, des procédés biologiques, des localisations et de domaines protéiques enrichis donne une image à plus large échelle des fonctions des réseaux d'interactions. Ces analyses *in silico* ont été complétées par des analyses *in vitro* et *in vivo* afin de valider les interactions identifiées et leur rôle fonctionnel.

IV. Analyses *in vitro* ou *in vivo*

Ces expériences visent d'une part à confirmer les interactions identifiées en BioID par des approches complémentaires (IP-MS, PLA), et d'autre part à confirmer leur rôle fonctionnel au niveau cellulaire.

IV.1 Validation des interactions

IV.1.a Validation d'interactions par IP-MS

L'approche par immunoprécipitation suivie de la spectrométrie de masse a été utilisée car il s'agit de la technique historique dans l'étude des interactions protéiques, et la plupart des bases de données d'interactions protéiques sont alimentées par ce type de technique.

Dans cette méthode, la préparation des cellules est identique au BioID, seule la lyse des cellules et la purification diffèrent. Les cellules sont récoltées puis lysées par avec un tampon moins stringent que le RIPA, notamment composé de sel, d'inhibiteurs de protéases et de Nonidet P-40 (détergeant non ionique et non-dénaturant qui remplace le SDS présent dans le RIPA). Les protéines sont ensuite mises au contact de billes d'agarose conjuguées à des anticorps anti-Flag afin de purifier la protéine cible conjuguée à Flag-BirA*, et avec elle ses interacteurs. Les billes sont ensuite lavées dans des conditions moins stringentes que pour le BioID afin de préserver les interactions protéiques, mais avec le risque d'augmenter le taux de faux positifs. Les protéines purifiées sont ensuite éluées, puis lyophilisées et digérées par la trypsine. Les peptides sont repris dans un tampon contenant 0,1% d'acide formique et 1/6^e de l'échantillon est analysé par MS.

Les résultats de MS sont ensuite analysés comme pour le BioID, avec les mêmes seuils statistiques, comparativement aux cellules contrôle exprimant Flag-BirA* seul. Les interacteurs identifiés par cette approche sont ensuite comparés aux interacteurs identifiés en BioID et à ceux référencés dans les bases de données. Les différents isoformes de RAS (HRAS, NRAS et KRAS) ont été analysés par cette approche.

IV.1.b Validation d'interactions par PLA

En parallèle de l'approche *in vitro* par IP-MS j'ai choisi de valider les interactions identifiées *in vivo* par la technique PLA.

Cette validation d'interaction a été faite dans les cellules HEK293 par l'induction de la protéine cible fusionnée à Flag-BirA* conjointement à l'expression ectopique de l'interacteur d'intérêt fusionné avec un étiquette HA (Human influenza hémagglutinine). Ceci permet l'utilisation d'anticorps très spécifiques, ciblant les motifs Flag et HA, et permet de s'affranchir d'anticorps primaires dirigés contre des protéines peu étudiées ou d'anticorps non développées pour les techniques d'immunofluorescence ou d'immunohistochimie.

Cette approche a donc nécessité de cloner dans un vecteur pcDNA3 les interacteurs cibles en fusionnant avec l'étiquette HA. Ces vecteurs ont ensuite été transfectés dans les cellules, puis à 24 heures post-transfection les cellules sont induites par la tétracycline pour exprimer la protéine cible. Après 24 heures supplémentaires, les cellules sont fixées et perméabilisées au méthanol. Les cellules fixées sont ensuite hybridées avec les anticorps primaires dirigés contre les étiquettes Flag et HA. Les anticorps secondaires couplés aux oligonucléotides reconnaissent ensuite les anticorps primaires, permettant l'hybridation des oligonucléotides dans le cas où les deux protéines interagissent. Après ligation, ces oligonucléotides servent d'amorce à l'amplification en cercle roulant (Figure 23). L'étape d'amplification dure 1 heure à 37°C au cours de laquelle des fluorophores sont intégrés dans le produit d'amplification.

Cette technique a été appliquée à l'étude des interactions des isoformes RAS afin de valider la spécificité de la technique de BioID et sa sensibilité par rapport aux IP-MS. L'interaction bien décrite entre les isoformes de RAS et BRAF est utilisée comme contrôle positif d'interaction, alors que le vecteur vide pcDNA3 est utilisé comme contrôle négatif.

IV.2 Études fonctionnelles par knock-down

Une fois les interactions validées, je me suis intéressé au rôle fonctionnel des interacteurs identifiés en BioID dans un contexte muté ou non et à leur impact sur la viabilité cellulaire. La stratégie que j'ai

choisie est d'invalider l'expression d'un interacteur donné en présence de la protéine cible mutée ou non.

À partir des analyses d'interaction de la protéine KRAS sauvage et KRAS muté G12C, j'ai sélectionné 12 protéines qui interagissent préférentiellement avec la forme mutée, et 1 protéine qui interagit préférentiellement avec la forme sauvage. L'expression de ces interacteurs a été invalidée par des ARN interférants (siRNA) et ceci dans 4 lignées d'adénocarcinomes pulmonaires présentant un génotype KRAS G12C ou KRAS WT. Le génotype des lignées utilisées et leur origine est précisé sur le tableau 6.

Tableau 6. Tableau récapitulatif des lignées cellulaire utilisées pour la validation fonctionnelle des interacteurs dérégulés de KRAS. Les lignées H23 et H358 présentent un génotype KRAS G12C, EGFR WT et ALK WT. Les lignées H1650 et H1975 présentent un génotype KRAS WT, EGFR Del. 19 ou EGFR L858R et T790M, ALK WT.

Lignée cellulaire	Type de cancer	Age	Sexe	Fumeur	Nombre de paquets de cigarettes/an	Statut KRAS	Statut EGFR	Statut fusion EML4-ALK
H23	Adénocarcinoma; NSCLC	51	M	Oui	40	G12C	WT	WT
H358	Bronchioalveolar Carcinoma; NSCLC	N/A	N/A	N/A	N/A	G12C	N/A	N/A
H1650	Bronchoalveolar Carcinoma (Stage 3b); NSCLC	27	M	Oui	10	WT	Del. 19	WT
H1975	Adénocarcinoma; NSCLC	N/A	F	Non	N/A	WT	L858R ; T790M	WT

Concrètement, pour chacune des lignées cellulaires, l'extinction d'expression des protéines cibles a été réalisée par reverse-transfection en utilisant un mélange de siRNA. Après transfection, les cellules ont immédiatement étéensemencées et la viabilité cellulaire directement mesurée, puis à 96h post transfection, le temps final. La viabilité a été mesurée en utilisant de l'Alamar Blue® et l'extinction des protéines évaluée par western-blot.

B. Analyse de l'interactome de RAS

I. Introduction de l'étude

Notre étude vise à identifier en utilisant la technique du BioID les propriétés partagées et spécifiques à chacun des isoformes de RAS en termes de fonctions biologiques, localisation des interactions protéiques et domaines protéiques engagés.

La famille des protéines RAS (HRAS, KRAS et NRAS) est impliquée dans plusieurs pathologies comme les cancers, avec des mutations somatiques, et des d'autres pathologies comme les syndromes de Noonan et syndromes de Costello, avec des mutations germinales de RAS (Rauen, 2007; Schubbert et al., 2006). Dans le cas des cancers, les protéines RAS sont mutées dans plus de 25% des cas et ces mutations affectent préférentiellement certains isoformes en fonction du type de cancer, alors que tous ces cancers expriment la totalité des isoformes. Par exemple, les cancers colorectaux, pulmonaires et pancréatiques sont principalement mutés KRAS alors que les mélanomes et les cancers de la thyroïde sont principalement mutés NRAS. Bien qu'homologues à plus de 80%, l'analyse comparée de la séquence primaire de ces protéines met en évidence une région hypervariable (15% d'homologie), siège de modifications post traductionnelles et responsable d'adressages spécifiques à la membrane spécifiques. Le fait que des mutations spécifiques de certains isoformes se retrouvent dans des pathologies différentes, tout en exprimant l'ensemble des isoformes, suggère fortement que les différents isoformes de RAS possèdent des fonctions cellulaires spécifiques (Castellano and Santos, 2011).

L'établissement des réseaux d'interactions protéiques de chacun des isoformes et leur comparaison permettra ainsi de mieux caractériser les fonctions biologiques communes de ces isoformes mais surtout leurs fonctions biologiques spécifiques. Les données de MS sont disponibles en Annexe 1 et 2.

II. Manuscrit

BioID uncovers the proximal protein-protein interactions landscape of RAS isoforms

Running title: BioID-based RAS interactome identifies isoform specificities

ABBREVIATIONS

BioID, proximity-dependant biotinylation-identification

PPI, protein-protein interaction

GAP, GTPase-activating proteins

GEF, guanine nucleotide exchange factors

HVR, hyper-variable-region

PTM, post-translational modifications

IP-MS, immunopurification-mass spectrometry

SAINT, significance analysis of interactome

FDR, false discovery rate

RTK, receptor tyrosine kinase

PLA, proximity ligation assay

SUMMARY

RAS proteins, namely KRAS, NRAS and HRAS, are isoforms frequently activated in various cancers such as non-small cell lung cancer, colorectal cancer, melanoma and bladder cancer. For many years, the high degree of sequence homology (80% identity) associated with the common share of upstream and downstream interacting partners led to consider their activities as redundant. However, (i) the high conservation of the Hyper-Variable-Region across mammalian species, (ii) the preferential activation of different RAS proteins in specific types of tumors and, (iii) the specific post-translational modifications and plasma membrane localization of each isoform indicate that these proteins may play a specific role. To decipher RAS proteins specificities, we analysed by mass spectrometry their proximal protein-protein interaction landscape using the proximity-dependent biotin identification technology (BioID). Performed in the HEK293 T-REx flp-in cell system stably expressing the wild type KRAS, NRAS or HRAS protein, we characterized more than 800 high-confidence proximal interactors specific to RAS proteins. Here, we depict the first comparative large-scale network of RAS protein-protein interactions. This resource shed light on less described function of RAS like its particular proximity with proteins involved in metabolism, and depict shared but also specific RAS protein-networks that together contribute to the complexity of RAS functions. This network of protein interactions therefore opens a new vision to RAS research.

INTRODUCTION

Characterized during the eighties, RAS genes code for four highly related and nearly ubiquitously expressed proteins, namely HRAS, NRAS, KRAS4A and KRAS4B. Placed at the crossroads of a broad range of cellular signaling networks, these small GTPases are essential for cell fate as they control cell proliferation, differentiation and survival (Cox and Der, 2010; Downward, 2003). RAS proteins act as a binary switch oscillating through an active GTP- and inactive GDP-bound state, triggered by the direct interaction with Guanine nucleotide Exchange Factors (GEFs) or GTPase-activating proteins (GAPs) (Bos et al., 2007). While GEFs trigger the GDP to GTP exchange and turn on RAS signaling, GAPs stimulates GTP hydrolysis and thus RAS inactivation (Vigil et al., 2010). RAS regulation is responsive to multiple signal inputs and this cycling enable the transduction of extra-cellular stimuli to intra-cellular signaling cascade, and therefore adapt the cell to its environment (Cox et al., 2014). Understanding RAS functions represents a major challenge as they are proto-oncogenes frequently activated in human cancers (e.g. Non-Small Cell Lung Cancer (NSCLC), melanoma, bladder cancer) and developmental disorders (i.e. Costello and Noonan syndrome). KRAS mutations are solely considered to be responsible of around one million deaths per year world-wide (Schubbert et al., 2007; Simanshu et al., 2017). Over the years, RAS have been extensively studied and considerable efforts have been devoted to decipher RAS functions and identify specific vulnerabilities. However, only few strides have been made, and there is still no targeted therapies available, nor clinical trials launched (Aoki et al., 2016; Patricelli et al., 2016)

For many years, the high degree of sequence homology (~80% identity) encompassing the regulatory and effector binding region, led to consider RAS isoforms activities as redundant (Karnoub and Weinberg, 2008; Lu et al., 2016). Indeed, RAS interactors like NF1, RASAL2, RAF proteins, PI3KCA are part of the common set of upstream and downstream RAS interacting partners well described (Castellano and Downward, 2011; Downward, 2003; Drosten et al., 2010). However, the perfect overlap of protein-interactions (PPIs) network over RAS isoforms is less clear, and accumulating pieces of evidence tend to depict clear distinct biological properties. First, the overall high sequence homology is misleading since the Hyper-Variable-Region (HVR) encompass 19-20 amino acids of 15% identity responsible for different post-translational modifications (PTMs) and *in fine* plasma membrane localizations (Ahearn et al., 2011). Proteins behave in a specific environment, and given that the HVR sequence is well-conserved across mammalian species, this suggest distinct and preserved functions (Hancock, 2003). Second, RAS-activated tumors are preferentially mutated for one of the isoforms, while

they are all expressed in the corresponding tissue (Furth et al., 1987). We can mention, KRAS mutation in colorectal cancer, NRAS mutation in melanoma and HRAS mutation in bladder cancer. This suggest distinct environments that requires the activity of one isoform over the others (Cox et al., 2014). Last, KRAS4B has solely a major role in embryogenic development. Indeed, KRAS4B, is essential for mice embryogenesis while this is not true for KRAS4A, HRAS nor NRAS (Johnson et al., 1997; Koera et al., 1997). Altogether, (i) the specific post-translational modifications and localizations of each isoforms, (ii) the preferential activation of different RAS proteins in specific types of tumors, and (iii) the specific requirement of KRAS4B isoform for mouse development are accumulative evidence that support the notion of isoform-specific RAS functions.

Within cells, proteins behave together and in specific locations to wield their functions. According to the extracellular environment or PTMs, the network of protein-interactions evolves over the time and cellular space, and it is this diversity and dynamic nature of protein interactions that depict protein biological properties (Larance and Lamond, 2015). Hence, the characterization of protein proximal-landscape of interactions is an essential step to expand our knowledge on biological functions. Proximity-dependant biotinylation identification (BioID) has emerge as a powerful tool to gain new perspectives on protein functions, especially for proteins from a poorly soluble compartment like membranes (Roux et al., 2012). Indeed, classical co-IP Mass Spectrometry (MS) techniques are laborious to perform for this cellular compartment because Protein-Protein Interactions (PPIs) are often lost during the cell membrane lysis process (Lönn and Landegren, 2017). This is of particular interest for proteins like RAS that are anchored to the plasma membrane and trigger the translocation of proteins to the plasma membrane once activated (Lavoie and Therrien, 2015). Moreover, other available PPIs identification approaches (e.g. Yeast-two hybrid system, in vitro binding assays) are not performed in cells, and are thus subjected to caution. BioID circumvents many of these caveats and enables *in vivo* the identification of transient, weak and strong PPIs (Kim and Roux, 2016). BioID has already been successfully employed to reveal cilia architecture and specific signaling from oncogenic fusion proteins, and therefore shows the strength of this technology to uncover unknown PPIs and protein functions (Cutler et al., 2017; Gupta et al., 2015).

Understanding RAS function in the context of plasma membrane association is at the core of deciphering RAS regulation and activity. Over the years, whereas RAS research has been mainly focused on the pathological-involvement of these proteins, poor is known regarding RAS-wild-type activities and specificities, while this is the first step to clearly

decipher pathogenic to non-pathogenic behavior (Castellano and Downward, 2011; Simanshu et al., 2017). Additionally, high-throughput *in-vivo* proteomic analysis of RAS isoforms PPIs are today lagging behind while technological tools to identify *in vivo* protein PPIs are greatly evolving (Kim and Roux, 2016; Lönn and Landegren, 2017). These considerations prompted us to perform BioID on HRAS, NRAS and KRAS4B, the major KRAS transcript - therefore name solely KRAS. Here, we provide a resource of high-confident PPIs and depict a comprehensive picture of RAS proximal landscape of interactions in a more physiological context than what has been already published. As expect the overlapping degree of PPIs over the isoforms is high but all of them exhibit network specificities. Most of the high-confident interactors reported here are not previously reported, and thereby complete RAS' PPIs portrait and open a new vision of RAS network and isoforms particularities.

EXPERIMENTAL PROCEDURES

Cloning

All PCR reactions were performed using Q5 DNA polymerase (NEB) according to manufacturer's instructions. Primers, destination plasmids, cloning sites and DNA templates used for BioID, IP-MS and PLA experiments are listed in supplemental Table S1. PCR products were digested with *AscI* and *NotI* (NEB) and ligated into the destination plasmid using T4 DNA ligase (NEB) following manufacturer's instructions.

BioID experiment

Full length wild type human HRAS (NM_005343.3), NRAS (NM_002524.4) and KRAS (NM_004985.4) coding sequences were amplified by PCR and cloned into the pcDNA5 FRT/TO FLAG-BirA* expression vector, downstream to and in-frame with BirA* (BirA R118G) (Table S1). Using the Flp-In system (Invitrogen), 293 T-Rex Flp-In cells stably expressing FLAGBirA*-HRAS WT, FLAGBirA*-NRAS WT and FLAGBirA*-KRAS WT were generated. After selection (DMEM + 10% FBS + 200µg/mL Hygromycin B), 10 X 150 cm² plates of subconfluent (60%) cells were incubated for 24h in complete media supplemented with 1 µg/mL tetracycline (Sigma) and 50µM biotin (BioShop, Burlington, ON, Canada). Cells were collected and pelleted (2000 rpm, 3 min), the pellet was washed twice with PBS, and dried pellets were snap frozen. Pellets were lysed in 10 ml of modified RIPA lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, 0.1% SDS, 1:500 protease inhibitor mixture (Sigma-Aldrich, Saint-Louis, MO), 250U Turbonuclease (Accelagen, San Diego, CA)) at 4°C for 1 h, then sonicated (30 s at 35% power, Sonic Dismembrator 500; Fisher Scientific) to disrupt visible aggregates. The lysate was centrifuged at 16000 rpm (35000 X g) for 30 min. Clarified supernatants were incubated with 30 µl packed, preequilibrated streptavidin-Sepharose beads (GE) at 4 °C for 3 h. Beads were collected by centrifugation (2000 rpm, 2 min), washed six times with 50 mM ammonium bicarbonate pH 8.3, and treated with TPCK trypsin (Promega, Madison, WI, 16 h at 37 °C). The supernatant containing the tryptic peptides was collected and lyophilized. Peptides were resuspended in 0.1% formic acid and 1/6th of the sample was analyzed per MS run.

FLAG Affinity Purification experiment

For FLAG pulldowns, 10 X 150 cm² dishes of FLAGBirA*-HRAS WT, FLAGBirA*-NRAS WT and FLAGBirA*-KRAS WT were treated with tetracycline and biotin as above, scraped

into PBS, pooled, washed twice in 25mL PBS and collected by centrifugation at 1000 X g for 5 min at 4°C. Cell pellets were stored at -80°C until lysis. The cell pellet was weighed and 1:4 pellet weight/lysis buffer (by volume) was added. Lysis buffer consisted of 50 mM HEPES-NaOH (pH 8.0), 100 mM KCl, 2 mM EDTA, 0.1% Nonidet P-40, 10% glycerol, 1 mM PMSF, 1 mM DTT and 1:500 protease inhibitor mixture (Sigma-Aldrich, St. Louis, MO). Upon resuspension, cells were incubated on ice for 10 min, subjected to one additional freeze-thaw cycle, then centrifugated at 27000 X g for 20 min at 4°C. Supernatant was transferred to a fresh 15mL conical tube and 250U Turbonuclease (Accelagen) plus 30µL packed, pre-equilibrated FLAG-M2 agarose beads (Sigma-Aldrich) were added. The mixture was incubated for 2 h at 4 °C with end-over-end rotation. Beads were pelleted by centrifugation at 1000 rpm (100 X g) for 1 min and transferred with 1 ml of lysis buffer to a fresh centrifuge tube. Beads were washed once with 1 ml lysis buffer and twice with 1 ml ammonium bicarbonate (ammbic) rinsing buffer (50 mM ammbic pH 8.0, 75 mM KCl). Elution was performed by incubating the beads with 150 µl of 125 mM ammonium hydroxide (pH >11). The elution step was repeated twice more, and the combined eluate centrifuged at 15000 X g for 10 min, transferred to a fresh centrifuge tube and lyophilized. Following overnight trypsin digestion (as above), peptides were resuspended in 0.1% formic acid and 1/6th of the sample was analyzed per MS run.

Liquid Chromatography-Mass Spectrometry

Liquid chromatography (LC) analytical columns (75 nm inner diameter) and precolumns (150 nm ID) were made in house from fused silica capillary tubing from InnovaQuartz (Phoenix, AZ) and packed with 100Å C18-coated silica particles (Magic, Michrom Bioresources, Auburn, CA). LC-MS/MS was conducted using a 120 min reversed-phase buffer gradient running at 250 nl/min (column heated to 40 °C) on a Proxeon EASY-nLC pump in line with a Q Exactive™ HF Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo Fisher Scientific). A parent ion scan was performed in the Orbitrap, using a resolving power of 60000. For protein identification, .raw files were converted to the .mzXML format using Proteowizard (Kessner et al., 2008), then searched using X!Tandem (Craig and Beavis, 2004) against Human RefSeq Version 45 (containing 36113 entries). Search parameters specified a parent MS tolerance of 15 ppm and an MS/MS fragment ion tolerance of 0.4 Da, with up to two missed cleavages allowed for trypsin. Oxidation of methionine was allowed as a variable modification. Data were analyzed using the trans-proteomic pipeline (Pedrioli, 2010) via the ProHits 2.0.0 software suite (Liu et al., 2010). Proteins identified with a ProteinProphet cut-off of 0.85 (corresponding to ≤1% FDR) were analyzed with SAINT Express v. 3.3 (Choi et al., 2011; Teo et al., 2014). For

BioID and IP-MS, respective eight control runs were used for comparative purposes and four runs were conducted on 293 T-REx cells expressing the FLAGBirA*-HRAS WT, FLAGBirA*-NRAS WT and FLAGBirA*-KRAS WT. The eight controls were collapsed to the highest two spectral counts for each hit. Proteins identified with two or more unique peptides and scoring above a Bayesian False Discovery Rate of 1% (corresponding to a SAINT score ≥ 0.75) were considered to be high confidence proximity interactors.

Network representation

The networks of protein-protein interactions were made using Cytoscape v3.2.1. Only the proteins identified with a significant SAINT score were considered as interactor to the corresponding bait. The interactors were mapped to each other according to the total spectral count of the four MS runs, and using the edge-weighted spring embedded layout algorithm.

Protein-protein interaction database

All the known RAS interactors were downloaded from the National Center for Biotechnology Information (NCBI) at <https://www.ncbi.nlm.nih.gov/>. It encompasses protein interactions from MINT, BioGRID and HPRD databases.

Dotplot layout and functional genomic analysis

The dot plot of BioID and IP-MS interactors were performed using the Prohits Visualization tool with the SAINT output and the standard parameter (Knight et al., 2017).

Enrichment in Gene Ontology (GO) biological process, GO cellular components, KEGG pathways, Reactome pathways and InterPro protein domains were performed using the DAVID v6.8 server at <https://david.ncifcrf.gov/> (Huang et al., 2009). GO_BP_5 category were used and the output were filtered to keep non-redundant ontologies using REVIGO analysis tool at <http://revigo.irb.hr/> (Supek et al., 2011). For the cellular mapping the interactors identified were tested for GO_CC_ALL category and the ontologies relative to each compartment were searched manually. With the exception of the cellular mapping, GO_CC_Direct were used to identify significant cell compartments. To identify preferential interaction with one specific isoform, the fold change was calculated as: $\log_2(1 + \text{spectral count sum Bait A}) / (1 + \text{spectral count sum Bait B})$. If a \log_2 fold change ≥ 1 or ≤ -1 was measured, the interactor was considered as preferential to the corresponding bait.

Proximity Ligation Assay

293 T-Rex cells (1.75×10^5) were grown on coverslips into 6-well plates, incubated for 48 hours and subjected to transfection with 500ng of the pcDNA3 3xHA-tagged RAS substrate candidate using PolyJet transfection reagent (Signagen, Rockville, MD), according to manufacturer's instructions. On day one, transfection media was removed and replaced by fresh media with tetracycline ($1\mu\text{g}/\text{mL}$). On day two, cells were fixed with ice-cold methanol for 2 min, washed twice in PBS. Then, the slides were treated according to the manufacturer's instructions (Duolink In Situ Red starter kit Mouse/Rabbit, Sigma DUO92101). Firstly, the samples were saturated using the blocking solution, then mouse anti-FLAG antibody 1:500 (Euromedex EL-1B11) and rabbit anti-HA antibody 1:3000 (Cell signaling C29F4) were incubated 1 h at 37°C . After two washes in PBS, the PLA minus and plus probes (containing the secondary antibodies conjugated with complementary oligonucleotides) were added and incubated 1 h at 37°C . The next step allows the ligation of oligonucleotides if the two proteins are in close proximity ($\leq 30\text{nm}$) thanks to the ligase during an incubation of 30 min at 37°C . Then, after two washes, the addition of nucleotides and polymerase, allowed amplification by rolling-circle amplification (RCA) reaction using the ligated circle as a template during an incubation of 100 min at 37°C . The amplification solution also contains fluorescently labelled oligonucleotides that hybridize to the RCA product. Then, the samples were dried at room temperature in the dark and were mounted with Duolink II Mounting Medium containing DAPI, then analysed on fluorescence microscope.

Image acquisition and analysis

The hybridized fluorescent slides were viewed under a Leica DM6000B microscope. Images were acquired under identical conditions at objective X63. For each sample, 8 fields were acquired with, for each field, a z-stack of 11 images spaced from $1\mu\text{m}$ to each other's. Analyses and quantifications of these samples were performed using Image J software (<https://imagej.nih.gov/ij/>). Each field were merged and dots counted on 8 bits image. The plugin "Counter cells" allowed to count the number of cells. The ratio Dots/Cell were calculated and a Student's t-test was performed to calculate the significance of interaction. Quantifications are represented as Mean \pm SEM.

SDS-PAGE and western blotting

Cells were lysed using a commercial cell lysis buffer (CST #9803), incubated 10 minutes on ice and centrifuged at 15,000 rpm for 15 minutes. The supernatant was collected and the protein concentration

measured with a micro-Bradford assay (Thermo Scientific). 10µg of proteins were loaded on 10% Bis-Tris gels and transferred on activated PVDF membranes at 300mA for 2 hours. Subsequently, proteins were detected using anti-FLAG-M2 antibody (1:5000, Sigma F1804), anti-HA antibody (1:5000, CST #C29F4) and anti-beta-actin (1:5000, Santa Cruz sc-47778). The antibody-bound proteins were revealed using an ECL western blotting substrate (Thermo Scientific) and visualized using a chemiluminescent imaging system.

RESULTS AND DISCUSSION

A RAS BioID-based interactome

To gain a better understanding of RAS interaction landscape, BioID analysis was conducted on the three RAS isoforms HRAS, NRAS and KRAS. We used the 293 T-REx cells, a model system that has been successfully used to identify proximal interactors for a large number of bait proteins, including nuclear proteins, fusion proteins, and centriolar proteins (Cutler et al., 2017; Gupta et al., 2015; Mehus et al., 2016). To wield their functions, RAS proteins require post-translational modifications. The cleavage of a targeting motif at the C-terminus followed by prenylation enables RAS proteins to be targeted to the plasma membrane (Wang and Casey, 2016). Considering these essential PTMs, the FLAGBirA* tag (therefore name FB) was fused at the N-terminus of RAS proteins. Two independent inducible and stable cell pools were created for each RAS protein. Cells were subjected to intensive lysis, sonication and nuclease treatment to maximize the solubilisation of proteins localized at poorly soluble intracellular locations, or associated with membranes, RNA or chromatin. A streptavidin-Sepharose matrix was then used to capture biotin-labeled proteins, and after extensive washes the matrix was treated with trypsin. The eluted peptides were subsequently analysed using nanoflow liquid chromatography-electrospray ionization-tandem mass spectrometry (nLC-ESI-MS/MS). Cell pools expressing the FB tag alone were subjected to the same analysis to control for endogenously biotinylated proteins that interact non-specifically with streptavidin. Using the ProHits (Liu et al., 2010) system, MS data were analyzed using the Trans-proteomic pipeline (Deutsch et al., 2010). Finally, the data were subjected to Significance Analysis of INteractomes (SAINT) (Choi et al., 2011; Teo et al., 2014) to identify *bona fide* RAS interacting partners. Proteins identified with two or more unique peptides and scoring above a SAINT score of 0.75 (equivalent to 1% FDR) were considered to be high-confidence proximity interactors.

Altogether, 985 high confidence RAS-interactors were identified. This represent more than 4 times the total RAS-interactors referenced in BIND, BioGRID and HPRD databases. As expected, this large number of interacting partners encompass shared and specific RAS-interactors (Fig. 1, Table S2). Based only on the SAINT score, 734 proteins were identified to interact with HRAS, NRAS interacted with 609 proteins, and 630 proteins interacted with KRAS. Most of the interactors identified are found in the proximity of every isoforms, this represent a total of 376 proteins. However, each isoform appeared to depict a specific interactome and HRAS has a larger network of protein-interactions than KRAS or NRAS. Indeed, 197 partners were found to specifically interact with HRAS, whereas 95 and 81 were found to specifically interact with KRAS or NRAS, respectively (Fig. 1). Overall, this list includes the well-known RAS partners like BRAF, CRAF, PI3KCA, NF1 and RASAL2.

To determine how BioID method performs compared to a conventional IP approach, a standard FLAG tag-based immunopurification (IP-MS) was also conducted on the same cells, following the same induction and biotin-labelling treatment. In this case, 982 high-confidence RAS interactors were identified (Table S3). As expected, proteins identified with IP-MS shown only a partial overlap with the BioID-identified proteins, whatever the considered isoform. Indeed, around 15% of HRAS BioID dataset was identified using IP-MS, and this percentage is even lower for KRAS and NRAS, with 9% and 7% respectively (Fig. 2). Considering each isoform, a closer analysis shown that both BioID and IP-MS identify RAF kinases and PIK3 proteins. These proteins family are members of the mitogen-activated protein kinase (MAPK) pathway and PIK3/Akt signaling pathway, both considered as the major signaling pathway triggered by RAS activation. In line with this observation, compared to the databases BioGRID, BIND and HPRD, members of the MAPK and PIK3 signaling were also identified. However, the overlap with thesedatabases is surprisingly low, but once should be careful considering database. Indeed, GAPs and GEFs are well-described regulating RAS partners, but only a subset of them are referenced as RAS partners. More strikingly, the BRAF kinase is not referenced as a NRAS partner, while described as one of the RAS major effector.

BioID and IP-MS depict a different view of RAS functions

The low overlap observed from the comparison of BioID and IP-MS experiments prompted us to analyse the functions and protein domains identified from these respective networks of interactions. Indeed, the low overlap indicated two sub-populations of RAS interactors, likely driven by the technical specificities of these approaches. More precisely, 167 proteins were identified and shared between BioID and IP-MS, while 818 and 815 partners

were specifically identified from the respective techniques (Fig. 3A). Independent Gene Ontology (GO) Biological Process (BP) analysis of BioID and IP-MS RAS-interactors were performed and the comparison of the TOP5 ontologies shown clear distinct functions. Consistent with RAS localisation, the interactors identified with BioID appear to mostly wield their functions at the plasma membrane or to be involved in membrane architecture (Fig. 3B). Indeed, all of the TOP5 GO BP ontologies identified deal with plasma membrane organization, transport, or actin remodeling. With the exception of “Cell morphogenesis”, all the ontologies identified here are also identified using IP-MS dataset but at a much lower significance. This is particularly striking for the “Plasma membrane organization” ontology identified with a P-value of $5.6 \cdot 10^{-35}$ with BioID, while identified with a P-value of $4.4 \cdot 10^{-2}$ with IP-MS experiments. Given the fitness of BioID to identify protein from this particular location, this result was expected. IP-MS experiment enabled also to identify protein involved in transport but the membrane close interaction is less evident. On the contrary, IP-MS appeared to be suitable to identify proteins in interaction with the Golgi apparatus, an organelle required for RAS PTMs, or involved in the nuclear trafficking or mitotic cell cycle process. Of note, this last category was not significant using BioID dataset. Considering the plasma membrane localisation of RAS, the high significance of nuclear trafficking is surprising but RAS trigger a variety of signals that end inside the nucleus and could explain why RAS interactors could fit to this ontology, as well as to the mitotic cell cycle process. Overall, the TOP5 GO BP identified from IP-MS experiment appeared less corroborant to the ontologies identified from BioID and more generally, appeared less relevant when compared to RAS known functions and localizations.

Protein domains are at the core of PPIs and enable proteins to act together. RAS activation and signaling through proteins harboring kinase, SH2, SH3 or PH domain is well described and thus, we were questioning about the prevalence of these particular domains in BioID and IP-MS datasets. Using the InterPro database, the analysis of protein domains of the BioID-identified RAS-interactors highlighted the major prevalence of domains that enable proteins to be targeted to the plasma membrane (PH domain-like), domains involved in membrane trafficking (SNARE coiled-coil domain) or domain related to cell signaling (i.e. SH3-domain, kinase-like domain, small GTP-binding protein domain) (Fig. 3C). Altogether, the protein domains identified from RAS-interactors identified with BioID depict a more comprehensive picture of RAS mechanism of activation and signaling with domain involved in membrane interaction, signal transduction and vesicle trafficking. Thereby, these significant specificities greatly fit with RAS known-properties. Interestingly, these domains were poorly

or not identified from IP-MS experiments and this was most particularly the case for PH domain-like, SNARE coiled-coil domain and SH3 domains that were not identified. Additionally, the protein-kinase like domain was not significantly identified from the IP-MS experiment while it is one of the major domain that trigger RAS activity. Overall, with the exception of the small GTP-binding protein domain, IP-MS shown a different view of RAS interactions. Indeed, in line with the GO BP previously identified from IP-MS experiments, the domains identified are more related to nuclear trafficking (Importin-beta, N-terminal), cell cycle regulation and transcriptional control (Tetratricopeptide-like helical domain) or intracellular transport (HEAT type2 domain). However, the Armadillo-type fold domain is also identified and harbored by signaling proteins such as PIK3 proteins and the small Rho GTPases, members of the Ras superfamily and involved in cytoskeleton remodeling (Ridley, 2015).

Overall, BioID depicted a strong involvement of RAS interactors in membrane architecture, protein transport and cell signaling from the plasma membrane. Consistent with these functions, the protein domains that enable signal transduction and interaction with the plasma membrane were recovered, as well as protein domains necessary for protein transport. Interestingly, IP-MS depicted a strong involvement of RAS in protein transport, nuclear trafficking and cell cycle regulation. Consistent with these observations, the corresponding protein domains were also identified. However, as expected, the functions depicted solely by IP-MS were much less reliable to RAS known signaling, and thus appeared less relevant to study RAS interactome in the context of its essential interaction with the plasma membrane. Taken together, this comparison demonstrated the reliability of BioID over IP-MS to study RAS proteins in the context of membrane interaction.

RAS BioID strongly overlap RAS known pathway

Literature describe RAS proteins at the core of a signaling network that involve different families of proteins (e.g. Receptor Tyrosine Kinases (RTKs), GAPs, GEFs,) and different signaling pathways (i.e. MAPK, PIK3/Akt, RAP1, calcium signaling, endocytosis and actin cytoskeleton regulation). Considering the current knowledge referenced in the KEGG database, we mapped the interactors identified with BioID on the RAS-known pathways (Fig. 4). This mapping shown proteins implicated in every aspect of RAS-known activities, encompassing regulatory proteins and effector pathways. More precisely, we identified 17 RTKs, 4 GEFs and 7 GAPs that represent 25%, 15% and 33% of the members of these respective families. Although RTKs generally require adaptor proteins and thus do not necessarily physically interact with RAS proteins to promote their functions, this family is essential for RAS activity

and were identified in a close proximity with RAS. This is notably the case for EGFR that was identified but is not referenced as direct RAS interactor. Other members of the ERBB family were identified like ERBB2 and ERBB4. GAPs and GEFs proteins identified included the well-described GAB1 and NF1 proteins, respectively. RAS proteins are known to signal through MAPK, PIK3/Akt, RAP1 and Calcium pathways, and are known to be implicated in actin cytoskeleton and endocytosis regulation. As expected, we identified proteins from all these pathways, with an identification's percentage ranging from 6% to 26% of the total referenced members of these pathways. These percentages may appear low but it corresponds mainly to protein localised at the plasma membrane, while pathways such as MAPK or PIK3/Akt depict cascades of signaling that only start at the plasma membrane to finally spread to other compartments like nucleus. Additionally, these pathways include proteins that may not be expressed in every cell model. Only 6% and 13% of the MAPK and PIK3/Akt pathway were recovered, respectively. However, proteins related to the MAPK pathway included the RAF proteins (i.e. BRAF, RAF1 – ARAF identified in BioID is not referenced as MAPK signaling component) and the MEK proteins (i.e. MAP2K1, MAP2K2). The MAPK pathway encompasses a large amount of proteins but the RAF and MEK proteins are among the major members describing the canonical pathway. ERK proteins (i.e. MAPK1, MAPK3) are not described in the direct proximity of RAS and accordingly were not identified. The PIK3/Akt signaling proteins identified included proteins from the MAPK signaling (i.e. RAF1, MAP2K1, MAP2K2) but comprised also specific PIK3 proteins members. Indeed, the catalytic subunit alpha and beta of PI3K (PIK3CA, PIK3CB) as well as its regulatory subunits 1, 2 and 3 (PIK3R1, PIK3R2, PIK3R3) were identified and PIK3 is the first component of the PIK3/Akt cascade, described in a direct interaction with RAS. Interestingly, MTOR was identified in the close proximity of RAS. It was unexpected since MTOR is described as a downstream component of the PIK3/Akt pathway, activated by successive steps comprising activation of PI3K, AKT, TSC and RHEB proteins. The 14-3-3 protein members (YWHA proteins) were also highly represented with 5 out of the 7 family members, and notably none of them were previously reported as RAS interactors. This family of proteins bind to serine/threonine-phosphorylated residues and are known to be involved in various cellular processes such as cell signaling or cell cycle regulation (Hermeking, 2003). This is another line of evidence that RAS proteins are involved in a broad range and unexpected signaling networks. Actin cytoskeleton and endocytosis regulation were the two processes the most represented, with respectively 26% and 22% of the proteins belonging to these pathways identified with BioID (Fig. 4, Table S4). Actin cytoskeleton regulation is governed by the inter-connection of signaling and

cytoskeleton-related proteins. This two-component system was identified with the RAF proteins, 11 integrins, the regulating and catalytic subunits of PI3K and proteins from the cytoskeleton as actin or radixin. Finally, consistent with the biological process previously described, endocytosis is also represented with proteins belonging to RAB, SNX, CHMP and VPS families, as well as proteins from the ubiquitin-system. Other important RAS pathway regulating proteins not shown here were also identified like SHOC2.

Although not every target depicted here necessarily physically interact with RAS, taken together, this analysis supports BioID as a valid method to detect proteins relevant to RAS functions.

RAS interactors wire most of the cellular compartments

To determine the cell connection of the interactors identified, we conducted a GO analysis focused on cellular components (CC). Interestingly, every cellular compartment was identified for all the isoforms, with the exception of mitochondrion for KRAS (Fig. 5). This last observation has to be mitigated since the more specific ontology “mitochondrial outer membrane” was found significant, with 14 annotated proteins. Additionally, cancer research highlighted the link between mitochondria functions deregulations and KRAS mutations, as well as its direct interaction with the mitochondrial outer-membrane following PKC phosphorylation (Bivona et al., 2006; Neuzil et al., 2012; Weinberg et al., 2010). From a more global point of view, this analysis shown a strong connection with proteins that are localised at the plasma membrane or involved in cell junctions. Indeed, the strongest P-value ($<10^{-55}$) were observed for these cellular compartments. For every RAS isoform, the less represented but still significant compartment identified was the lysosome. This observation has to be balanced by the nature of this organelle and BioID's principles. Indeed, the aim of this cellular compartment is to degrade proteins and is therefore rich in ATPase. Moreover, BirA* requires ATP to activate biotin and label proteins. Considering these two points, the identification of interacting partners from this particular compartment is therefore particularly difficult. Overall, this analysis shown that RAS partners are able to target most of the cellular compartments, creating a dense and dynamic network of interactions. This complexity is consistent with the cross-road particular place of RAS proteins, and emphasizes the difficulties to depict clear RAS properties.

RAS isoforms exhibit a strong connection with metabolism

To decipher less known functions of RAS proteins, we performed a second run of pathway analysis without the RTKs, GAPs, GEFs proteins, the components of the MAPK,

PIK3/Akt, Rap1, Calcium signaling and without the proteins involved in endocytosis and actin cytoskeleton regulation previously identified. Only four significant pathways were identified, encompassing SNARE interactions, tight junctions, cell adhesion molecules (CAMs) and oxidative phosphorylation (Fig. 6A). SNARE interactions are involved in membrane fusions and mostly in exocytotic pathways (Chen and Scheller, 2001). With the tight junctions and CAMs, these three ontologies are located at the plasma membrane and correspond to RAS localisation. Cellular junctions are strongly connected to the actin cytoskeleton and therefore correlate with RAS known pathways. Exocytosis deal about vesicles trafficking and both HRAS and NRAS reach the plasma membrane from the Golgi apparatus through a vesicular transport (Apolloni et al., 2000). Additionally, with endocytosis, exocytosis relates to the protein transport process previously identified. More strikingly, 21 proteins were related to the oxidative phosphorylation pathway with at least 75% of these proteins identified for every isoform. Predominantly, proteins from the ATP synthase complex (5 proteins) and proteins from the NADH-ubiquinone oxidoreductase (9 proteins) were identified. This highlights a strong connection between RAS and metabolism. This observation was confirmed using the Reactome database (Fig. 6B, Table S5). This last analysis significantly identified the Respiratory electron transport pathway, which corresponds to the KEGG Oxidative phosphorylation pathway, but also three other energy-related ontologies. Indeed, out of the 7 significant pathways identified, the “Amino acid transport across the plasma membrane”, “Regulation of insulin secretion” and “Energy dependent regulation of mTOR by LKB1-AMPK” categories were also identified, and depict along with the “Respiratory electron transport” ontology an intimate interplay between cellular signaling and metabolism. This interplay encompasses vesicle trafficking, membrane transport, MTOR cell signaling and mitochondrial activity. Moreover, SNARE proteins were shown recently to be particularly involved in insulin vesicle secretion (Xiong et al., 2017) and thus could coordinate this interconnection. None of these RAS partners were previously reported and this close proximity between RAS and cellular energetics complete our view on RAS cellular environment.

RAS isoforms exhibit specific functions

Based on the spectral counts, we took advantage of the semi-quantitative aspect of this technic to decipher RAS isoform specificities. Therefore, we calculated the spectral count ratio of FB-KRAS to FB-HRAS, FB-NRAS to FB-HRAS, and FB-KRAS to FB-NRAS. If a difference of two-fold more peptides was observed for one of the two baits, this interactor was considered as preferential to the corresponding bait (Fig 7A, Table S6). Overall, when

compared to HRAS, KRAS and NRAS appeared both less linked to cellular membranes and to actin cytoskeleton (Fig 7B). Indeed, the significant GO BP ontologies referring to membranes (e.g. plasma membrane organization, protein localization to cell periphery, vesicle fusion, actin cytoskeleton organization) included a majority of interactors that were found closer to HRAS than to KRAS or NRAS. Consistently, the significant GO CC, protein domains and pathways corresponded to these functions, and most notably the GO CC plasma membrane, cell-cell adherent junction, cytoskeleton and the protein domains t-SNARE, Syntaxin/epimorphin and Chaperonin TCP-1. However, a specificity of KRAS over HRAS is its higher proximity with the regulating RAS proteins DAB2IP, IQGAP1, IQGAP2, NF1 and PLXNB2. Indeed, all these proteins harbor a GAP protein domain and were identified closer to KRAS than HRAS. Interestingly, this ontology was not significant when HRAS was compared to NRAS, suggesting a close connection of KRAS to these RAS regulating proteins. The PH-like domain harbored by membrane-interacting proteins and particular RAS interacting partners (e.g. GAB1, TRIO) were also significant and more proteins were closer to KRAS than HRAS. A shared specificity of KRAS and NRAS over HRAS is their connection to the pathway “central carbon metabolism in cancer”. This KEGG pathway includes notably proteins from the PIK3/Akt and MAPK pathways, glucose transporters, metabolic enzymes and tyrosine kinase receptors frequently activated in cancers. Both KRAS and NRAS were closer to these proteins, and more specifically to proteins from the PIK3/Akt enzymes such as AKT1, PIK3CA, and MTOR. This proximity corroborates to the connection of RAS with metabolism previously identified, and emphasizes the specific proximity of NRAS and KRAS in this interplay. Additionally, the specific identification of this pathway for this two particular baits corroborates with the predominant mutation of *KRAS* and *NRAS* isoforms in cancer. Differences between KRAS and NRAS are less evident but KRAS depicted a stronger link with vesicle trafficking than NRAS. Indeed, GO BP endosomal transport, endocytosis pathway and GO CC lysosomal membrane were significant for this specific isoform with proteins from the VPS and SNX families. Only the significant GO BP transmembrane transport was more specific to NRAS than KRAS. This feature was also identified when NRAS was compared to HRAS. In both case, this ontology included membrane transporters from the protein families SLC25 (phosphate mitochondrial transport), SLC30 (zinc balance) and SLC35 (nucleotide sugar transport) (Huang and Tepasamorndech, 2013; Palmieri, 2013; Song, 2013).

To identify properties exclusively specific to one RAS isoform, we used the ratios previously calculated (Fig. 7A) and compared all the proximal interactions found preferential

for one or the other isoform (Fig 8). The PPIs identified as preferential for one bait, when independently compared to the two other baits, were then considered to have a specific spectral count profile for this particular isoform (Table S7). Following the previous observations, the TOP5 GO BP for HRAS specific proximal interactors were related to membrane processes such as protein transport, membrane fusions, endocytosis, or exocytosis. Accordingly, the SNARE pathway was significant (not shown) and the t-SNARE, Synaptobrevin and Syntaxin/epimorphin, conserved site protein domains were significantly identified (Table 1, 2). Few significant ontologies were identified for KRAS or NRAS. However, corroborant to the previous observations, NRAS depicted a strong connection with the transmembrane transport process through the proximity with protein transporters from the SLC25, SLC30, SLC35 families, as well as the glucose transporter 1 (SLC2A1) (Table 1). The CRAL-TRIO protein domain was the only significant ontology identified for KRAS specific interactors (Table 2). This protein domain is harbored by the proteins NF1, MOSPD2, KALRN and TRIO and is necessary for lipid interactions. MOSPD2 is poorly described but KALRN and TRIO proteins are RhoGEFs essential for Rho GTPases activation. Along with NF1, they interact with the membrane-anchored prenyl group of the targeted GTPase to activate or inhibit its activity (Chiapparino et al., 2016). In addition, KALRN and TRIO proteins harbor also a SH3 and kinase domain and thus are involved in cellular signaling and regulation.

Overall, HRAS is closer to membrane and vesicle trafficking, whereas NRAS appears link to membrane transporters regulating metabolites and inorganic molecules fluxes. Finally, KRAS has a strong proximity with signaling regulating proteins and emphasizes the hub position of this particular RAS isoform.

Dataset validation

To validate RAS isoform specific interactions, we used the proximity ligation assay (PLA) that enables to detect protein-protein interactions within 30nm. Based on the spectral counts and SAINT score, we selected 8 prey proteins and BRAF as positive control of interaction. Among these 8 prey proteins, 3 were identified to specifically interact with FB-HRAS, 2 proteins with FB-NRAS and 3 proteins with FB-KRAS (Table 3). These proteins are involved in various cellular processes such as metabolism, cell signaling or vesicular transport. We performed PLA on induced FB-RAS and transfected HA-tagged prey proteins. Interaction with BRAF was observed with every RAS isoform, and dots were dependent on the presence of the ectopic BRAF protein (Fig. 9A). Specific interactions with RAS isoforms were observed and followed BioID identifications. More precisely, FB-HRAS interacted specifically with RAP2A and

S100A11, FB-NRAS interacted specifically with EBP and SHKBP1 and FB-KRAS interacted specifically with HA-PGD (Fig. 9B, Fig. S1). Only 3 interactors were discordant with BioID experiment. Notably, LASP1 was supposed to interact with HRAS, and TRIM13 and VPS35 were expected to specifically interact with KRAS. Overall, based on this dataset, more than 60% of the interactions identified with BioID were validated by PLA, and therefore support the analysis of RAS shared functions and specificities.

REFERENCES

1. Cox, A. D., and Der, C. J. (2010) Ras history: The saga continues. *Small GTPases* 1, 2–27
2. Downward, J. (2003) Targeting RAS signalling pathways in cancer therapy. *Nat. Rev. Cancer* 3, 11–22
3. Bos, J. L., Rehmann, H., and Wittinghofer, A. (2007) GEFs and GAPs: Critical Elements in the Control of Small G Proteins. *Cell* 129, 865–877
4. Vigil, D., Cherfils, J., Rossman, K. L., and Der, C. J. (2010) Ras superfamily GEFs and GAPs: validated and tractable targets for cancer therapy? *Nat. Rev. Cancer* 10, 842–857
5. Cox, A. D., Fesik, S. W., Kimmelman, A. C., Luo, J., and Der, C. J. (2014) Drugging the undruggable RAS: Mission Possible? *Nat. Rev. Drug Discov.* 13, 828–851
6. Schubert, S., Shannon, K., and Bollag, G. (2007) Hyperactive Ras in developmental disorders and cancer. *Nat. Rev. Cancer* 7, 295–308
7. Simanshu, D. K., Nissley, D. V., and McCormick, F. (2017) RAS Proteins and Their Regulators in Human Disease. *Cell* 170, 17–33
8. Aoki, Y., Niihori, T., Inoue, S., and Matsubara, Y. (2016) Recent advances in RASopathies. *J. Hum. Genet.* 61, 33–39
9. Patricelli, M. P., Janes, M. R., Li, L.-S., Hansen, R., Peters, U., Kessler, L. V., Chen, Y., Kucharski, J. M., Feng, J., Ely, T., Chen, J. H., Firdaus, S. J., Babbar, A., Ren, P., and Liu, Y. (2016) Selective Inhibition of Oncogenic KRAS Output with Small Molecules Targeting the Inactive State. *Cancer Discov.* 6, 316–329
10. Karnoub, A. E., and Weinberg, R. A. (2008) Ras oncogenes: split personalities. *Nat. Rev. Mol. Cell Biol.* 9, 517–531
11. Lu, S., Jang, H., Muratcioglu, S., Gursoy, A., Keskin, O., Nussinov, R., and Zhang, J. (2016) Ras Conformational Ensembles, Allostery, and Signaling. *Chem. Rev.* 116, 6607–6665

12. Castellano, E., and Downward, J. (2011) RAS Interaction with PI3K: More Than Just Another Effector Pathway. *Genes Cancer* 2, 261–274
13. Drosten, M., Dhawahir, A., Sum, E. Y. M., Urosevic, J., Lechuga, C. G., Esteban, L. M., Castellano, E., Guerra, C., Santos, E., and Barbacid, M. (2010) Genetic analysis of Ras signalling pathways in cell proliferation, migration and survival. *EMBO J.* 29, 1091–1104
14. Ahearn, I. M., Haigis, K., Bar-Sagi, D., and Philips, M. R. (2011) Regulating the regulator: post-translational modification of RAS. *Nat. Rev. Mol. Cell Biol.* 13, 39–51
15. Hancock, J. F. (2003) Ras proteins: different signals from different locations. *Nat. Rev. Mol. Cell Biol.* 4, 373–384
16. Furth, M. E., Aldrich, T. H., and Cordon-Cardo, C. (1987) Expression of ras proto-oncogene proteins in normal human tissues. *Oncogene* 1, 47–58
17. Johnson, L., Greenbaum, D., Cichowski, K., Mercer, K., Murphy, E., Schmitt, E., Bronson, R. T., Umanoff, H., Edelman, W., Kucherlapati, R., and Jacks, T. (1997) K-ras is an essential gene in the mouse with partial functional overlap with N-ras. *Genes Dev.* 11, 2468–2481
18. Koera, K., Nakamura, K., Nakao, K., Miyoshi, J., Toyoshima, K., Hatta, T., Otani, H., Aiba, A., and Katsuki, M. (1997) K-ras is essential for the development of the mouse embryo. *Oncogene* 15, 1151–1159
19. Larance, M., and Lamond, A. I. (2015) Multidimensional proteomics for cell biology. *Nat. Rev. Mol. Cell Biol.* 16, 269–280
20. Roux, K. J., Kim, D. I., Raida, M., and Burke, B. (2012) A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* 196, 801–810
21. Lönn, P., and Landegren, U. (2017) Close Encounters - Probing Proximal Proteins in Live or Fixed Cells. *Trends Biochem. Sci.* 42, 504–515
22. Lavoie, H., and Therrien, M. (2015) Regulation of RAF protein kinases in ERK signalling. *Nat. Rev. Mol. Cell Biol.* 16, 281–298
23. Kim, D. I., and Roux, K. J. (2016) Filling the Void: Proximity-Based Labeling of Proteins in Living Cells. *Trends Cell Biol.* 26, 804–817
24. Gupta, G. D., Coyaud, É., Gonçalves, J., Mojarad, B. A., Liu, Y., Wu, Q., Gheiratmand, L., Comartin, D., Tkach, J. M., Cheung, S. W. T., Bashkurov, M., Hasegan, M., Knight, J. D., Lin, Z.-Y., Schueler, M., Hildebrandt, F., Moffat, J., Gingras, A.-C., Raught, B., and Pelletier, L. (2015) A Dynamic Protein Interaction Landscape of the Human Centrosome-Cilium Interface. *Cell* 163, 1484–1499

25. Cutler, J. A., Tahir, R., Sreenivasamurthy, S. K., Mitchell, C., Renuse, S., Nirujogi, R. S., Patil, A. H., Heydarian, M., Wong, X., Wu, X., Huang, T.-C., Kim, M.-S., Reddy, K. L., and Pandey, A. (2017) Differential signaling through p190 and p210 BCR-ABL fusion proteins revealed by interactome and phosphoproteome analysis. *Leukemia* 31, 1513–1524
26. Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. (2008) ProteoWizard: open source software for rapid proteomics tools development. *Bioinforma. Oxf. Engl.* 24, 2534–2536
27. Craig, R., and Beavis, R. C. (2004) TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* 20, 1466–1467
28. Pedrioli, P. G. A. (2010) Trans-proteomic pipeline: a pipeline for proteomic analysis. *Methods Mol. Biol. Clifton NJ* 604, 213–238
29. Liu, G., Zhang, J., Larsen, B., Stark, C., Breitkreutz, A., Lin, Z.-Y., Breitkreutz, B.-J., Ding, Y., Colwill, K., Pasculescu, A., Pawson, T., Wrana, J. L., Nesvizhskii, A. I., Raught, B., Tyers, M., and Gingras, A.-C. (2010) ProHits: integrated software for mass spectrometry-based interaction proteomics. *Nat. Biotechnol.* 28, 1015–1017
30. Choi, H., Larsen, B., Lin, Z.-Y., Breitkreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z. S., Tyers, M., Gingras, A.-C., and Nesvizhskii, A. I. (2011) SAINT: probabilistic scoring of affinity purification–mass spectrometry data. *Nat. Methods* 8, 70–73
31. Teo, G., Liu, G., Zhang, J., Nesvizhskii, A. I., Gingras, A.-C., and Choi, H. (2014) SAINTexpress: improvements and additional features in Significance Analysis of INteractome software. *J. Proteomics* 100, 37–43
32. Knight, J. D. R., Choi, H., Gupta, G. D., Pelletier, L., Raught, B., Nesvizhskii, A. I., and Gingras, A.-C. (2017) ProHits-viz: a suite of web tools for visualizing interaction proteomics data. *Nat. Methods* 14, 645–646
33. Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57
34. Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011) REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS ONE* 6, e21800
35. Mehus, A. A., Anderson, R. H., and Roux, K. J. (2016) in *Methods in Enzymology* (Elsevier), pp 3–22.
36. Wang, M., and Casey, P. J. (2016) Protein prenylation: unique fats make their mark on biology. *Nat. Rev. Mol. Cell Biol.* 17, 110–122

37. Deutsch, E. W., Mendoza, L., Shteynberg, D., Farrah, T., Lam, H., Tasman, N., Sun, Z., Nilsson, E., Pratt, B., Prazen, B., Eng, J. K., Martin, D. B., Nesvizhskii, A. I., and Aebersold, R. (2010) A guided tour of the Trans-Proteomic Pipeline. *Proteomics* 10, 1150–1159
38. Ridley, A. J. (2015) Rho GTPase signalling in cell migration. *Curr. Opin. Cell Biol.* 36, 103–112
39. Hermeking, H. (2003) The 14-3-3 cancer connection. *Nat. Rev. Cancer* 3, 931–943
40. Weinberg, F., Hamanaka, R., Wheaton, W. W., Weinberg, S., Joseph, J., Lopez, M., Kalyanaraman, B., Mutlu, G. M., Budinger, G. R. S., and Chandel, N. S. (2010) Mitochondrial metabolism and ROS generation are essential for Kras-mediated tumorigenicity. *Proc. Natl. Acad. Sci.* 107, 8788–8793
41. Neuzil, J., Rohlena, J., and Dong, L.-F. (2012) K-Ras and mitochondria: Dangerous liaisons. *Cell Res.* 22, 285–287
42. Bivona, T. G., Quatela, S. E., Bodemann, B. O., Ahearn, I. M., Soskis, M. J., Mor, A., Miura, J., Wiener, H. H., Wright, L., Saba, S. G., Yim, D., Fein, A., Pérez de Castro, I., Li, C., Thompson, C. B., Cox, A. D., and Philips, M. R. (2006) PKC Regulates a Farnesyl-Electrostatic Switch on K-Ras that Promotes its Association with Bcl-Xl on Mitochondria and Induces Apoptosis. *Mol. Cell* 21, 481–493
43. Chen, Y. A., and Scheller, R. H. (2001) SNARE-mediated membrane fusion. *Nat. Rev. Mol. Cell Biol.* 2, 98–106
44. Apolloni, A., Prior, I. A., Lindsay, M., Parton, R. G., and Hancock, J. F. (2000) H-ras but not K-ras traffics to the plasma membrane through the exocytic pathway. *Mol. Cell. Biol.* 20, 2475–2487
45. Xiong, Q.-Y., Yu, C., Zhang, Y., Ling, L., Wang, L., and Gao, J.-L. (2017) Key proteins involved in insulin vesicle exocytosis and secretion. *Biomed. Rep.* 6, 134–139
46. Palmieri, F. (2013) The mitochondrial transporter family SLC25: Identification, properties and physiopathology. *Mol. Aspects Med.* 34, 465–484
47. Song, Z. (2013) Roles of the nucleotide sugar transporters (SLC35 family) in health and disease. *Mol. Aspects Med.* 34, 590–600
48. Huang, L., and Tepaamorndech, S. (2013) The SLC30 family of zinc transporters – A review of current understanding of their biological and pathophysiological roles. *Mol. Aspects Med.* 34, 548–560
49. Chiapparino, A., Maeda, K., Turei, D., Saez-Rodriguez, J., and Gavin, A.-C. (2016) The orchestra of lipid-transfer proteins at the crossroads between metabolism and signaling. *Prog. Lipid Res.* 61, 30–39

FIGURE LEGENDS

Fig 1. The RAS BioID interactome. (A) The network of RAS protein-protein interactions describes shared and specific prey proteins (round nodes) to each RAS isoform (square nodes). Prey proteins are placed from one to each other's according to their respective total spectral count, using the edge-weighted spring embedded layout algorithm (Cytoscape v3.2.1) (B) The Venn diagram of RAS protein network specifies the number of shared or specific proteins identified.

Fig 2. Comparison of BioID to IP-MS and database. The Venn diagram depicts the shared and specific interactors identified from BioID and IP-MS experiment, and the overlap with the known RAS interactors referenced in databases (BioGRID, MINT, HPRD from NCBI). The dot plot shows the average spectral count and SAINT score measured for every selected interactor.

Fig 3. Comparison of the biological processes and protein domains identified from BioID and IP-MS. (A) Venn diagram of the total RAS interacting partners identified from BioID, AP-MS experiments and those reported in databases (BIND, BioGRID, HPRD from NCBI). (B) Comparison of the TOP 5 Gene Ontology Biological Process (level 5) of the RAS interacting partners identified with BioID to those identified with IP-MS. (C) Comparison of the TOP 5 protein domains (InterPro database) identified with BioID to those identified with IP-MS.

Fig 4. BioID-identified interactors involved in RAS-known signaling. The listed proteins are mapped according to their belonging to protein families or RAS-known cellular pathways. The protein families involved are the receptor tyrosine kinases (RTK), Guanine nucleotide Exchange Factors (GEFs), GTPase-activating proteins (GAPs). The RTK, GEFs and GAPs used for identification are compiled from literature. The KEGG annotation was used to identify proteins from the mitogen-activated protein kinase (MAPK) pathway (hsa04010), PIK3/Akt pathway (hsa04151), Rap1 pathway (hsa04015), Calcium signaling pathway (hsa04020), Regulation of actin cytoskeleton pathway (hsa04810) and endocytosis pathway (hsa04144). Sig., signaling; Cytosk., cytoskeleton. Proteins annotated with an asterisk (*) were previously identified as RAS-known physical interactor (BIND, BioGRID, HPRD from NCBI).

Fig 5. Cellular localization of each RAS isoform interactor. (left) Cellular mapping of each RAS interactor. The annotation was determined by Gene Ontology (GO) analysis focused on Cellular components. (right) Associated P-values calculated for every cellular components and RAS isoforms. NI, not identified.

Fig 6. Under-represented functions of RAS. (A) KEGG pathway analysis performed without the proteins related to the RAS-known pathways identified SNARE interaction in vesicular transport (hsa04130), tight junctions (hsa04530), oxidative phosphorylation (hsa00190) and Cell Adhesion Molecules (CAMs, hsa04514) as the most significant RAS pathways. (B) Reactome pathway analysis identified a close proximity of RAS proteins with the metabolism-related pathways Amino-acid transport across the plasma membrane (HSA-352230), Regulation of insulin secretion (HSA-422356), Respiratory electron transport (HSA-611105) and Energy dependent regulation of mTOR by LKB1-AMPK (HSA-380972). AA., amino acids; dep., dependent.

Fig 7. RAS isoforms specificities. (A) Based on the spectral count ratio of each interactor, a protein was considered as preferential to one or the other isoform when at least two-fold more peptides was measured for the corresponding bait. (B) The respective significant Gene Ontology (GO) Biological Process (BP) level 5, GO Cellular Component (CC) Direct, InterPro protein domain and KEGG pathway were identified. The blue rectangle point-out ontologies that depict the most important differences between the isoforms. The letter in each square indicates to which RAS isoform the ontology is more specific.

Fig. 8. Identification of highly preferential interactor to each RAS isoform. The comparison of the deregulated interactors when compared to two baits were compared and enabled to cluster highly preferential interactor to each isoform.

Fig. 9. PLA validation of specific RAS isoform interactions. (A) PLA detection of FB-RAS / HA-BRAF interaction as technical positive control. PLA was performed with Flag- and HA-specific antibodies. The detected complexes are represented by red dots and the nuclei were counterstained with DAPI (blue). The ectopic expression of HA-BRAF and the tetracycline-dependant expression of FB-RAS have been confirmed by western blotting. The negative control of PLA was performed with the FB-RAS-induced cells transfected with the HA-empty vector, and probed with Flag- and HA-specific antibodies. (B) PLA detection of specific FB-HRAS, FB-NRAS or FB-KRAS interaction with HA-tagged RAP2A, SHKBP1 and PGD, respectively. PLA was performed as described and the ectopic expression of the HA-tagged prey protein was validated by western blotting (Fig S1A). Quantification of the dots per cell was performed by computer-assisted analysis. The mean \pm SEM is shown and the corresponding P-value was determined by Student's t-test. The expression of the HA-tagged proteins was validated by western blotting. Tet., Tetracycline; Neg., Negative; Pos., Positive; ***, P-value \leq 0.0001.

SUPPLEMENTARY FIGURE LEGEND

Fig. S1. PLA validation of specific RAS interactions. (A) Validation of the ectopic expression of HA-RAP2A, -SHKBP1 and -PGD. (B) PLA detection of specific FB-HRAS and FB-NRAS interaction with HA-tagged S100A11 and EBP, respectively. PLA was performed with Flag- and HA-specific antibodies. The detected complexes are represented by red dots and the nuclei were counterstained with DAPI (blue). Quantification of the dots per cell was performed by computer-assisted analysis. The mean±SEM is shown and the corresponding P-value was determined by Student's t-test. The expression of the HA-tagged proteins was validated by western blotting. ***, P-value ≤ 0.0001.

TABLES LEGENDS

Table 1. Specific functions to HRAS and NRAS. The significant functions strictly specific to HRAS and NRAS were determined by Gene Ontology (GO) Biological Process (BP) level 5 analysis.

Table 2. Specific protein domains to HRAS and KRAS. The significant protein domains strictly specific to HRAS and KRAS were determined by InterPro analysis.

Table 3. Selected interactors for proximity ligation assay (PLA). Based on the spectral count and Significance Analysis of INteractomes (SAINT) score, specific interactors of each RAS isoform were selected for interaction validation by PLA. BRAF (green) was selected as positive control of interaction for every RAS isoform.

SUPPLEMENTARY TABLE LEGEND

Table S1. Cloning material. Primers, destination plasmids, cloning sites and DNA templates used for BioID, IP-MS and PLA experiments.

Table S2. RAS BioID-based interactome. RAS interactors identified with BioID. For each bait, the proteins identified are from two biological replicates and two technical replicates. Row in orange correspond to HRAS, NRAS, KRAS and BirA*. For each prey protein, the total spectral count (sum of the spectral identified from the four MS runs), total unique peptides and the corresponding SAINT score is specified. Grey boxes correspond to spectral counts with only one unique peptide.

Table S3. RAS IP-MS-based interactome. RAS interactors identified with IP-MS. For each bait, the proteins identified are from two biological replicates and two technical replicates. Row

in orange correspond to HRAS, NRAS, KRAS and BirA*. For each prey protein, the total spectral count (sum of the spectral identified from the four MS run per bait), total unique spectral count and the corresponding SAINT score is specified. Grey boxes refer to spectral counts with only one unique peptide.

Table S4. BioID-interactors mapped to the regulation of actin cytoskeleton and endocytosis pathways. The BioID interactors that correspond to the RAS-known pathways regulation of actin cytoskeleton (hsa04810) and endocytosis (hsa04144) were identified using the KEGG annotation.

Table S5. Reactome analysis. To complement the KEGG pathway analysis, a second analysis of the BioID interactors without the proteins related to the RAS-known pathways were performed using the Reactome database. The seven significant ontologies are listed, as well as the corresponding p-values and proteins identified.

Table S6. Preferential RAS interactors. Based on the spectral counts, the \log_2 ratio of each prey protein was calculated to identify a preferential interaction with the bait. A preferential interaction was defined considering at least two-fold more peptides for the corresponding bait.

Table S7. Highly preferential RAS interactors. To identify the highly preferential interactors, the \log_2 ratios previously calculated were compared. A prey protein identified as preferential to one bait when independently compared to the two-other baits was considered as highly preferential to the corresponding bait.

Fig. 1

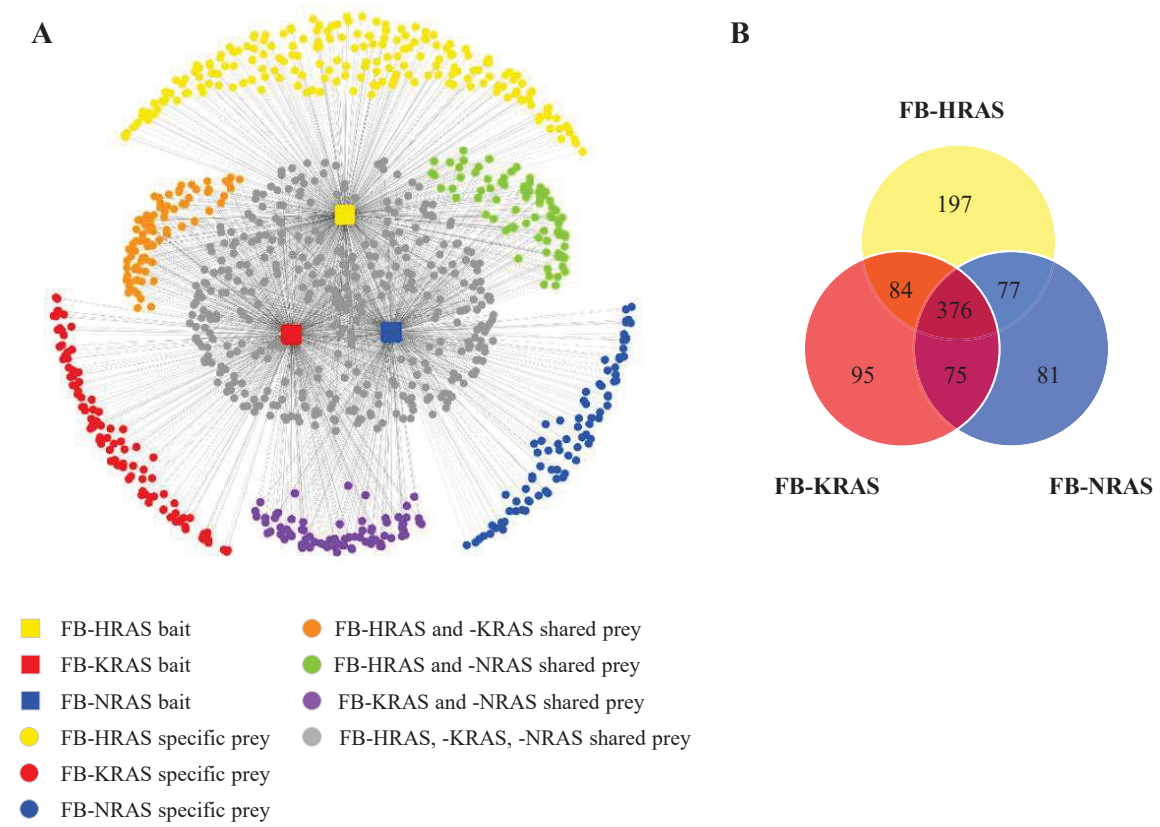


Fig. 2

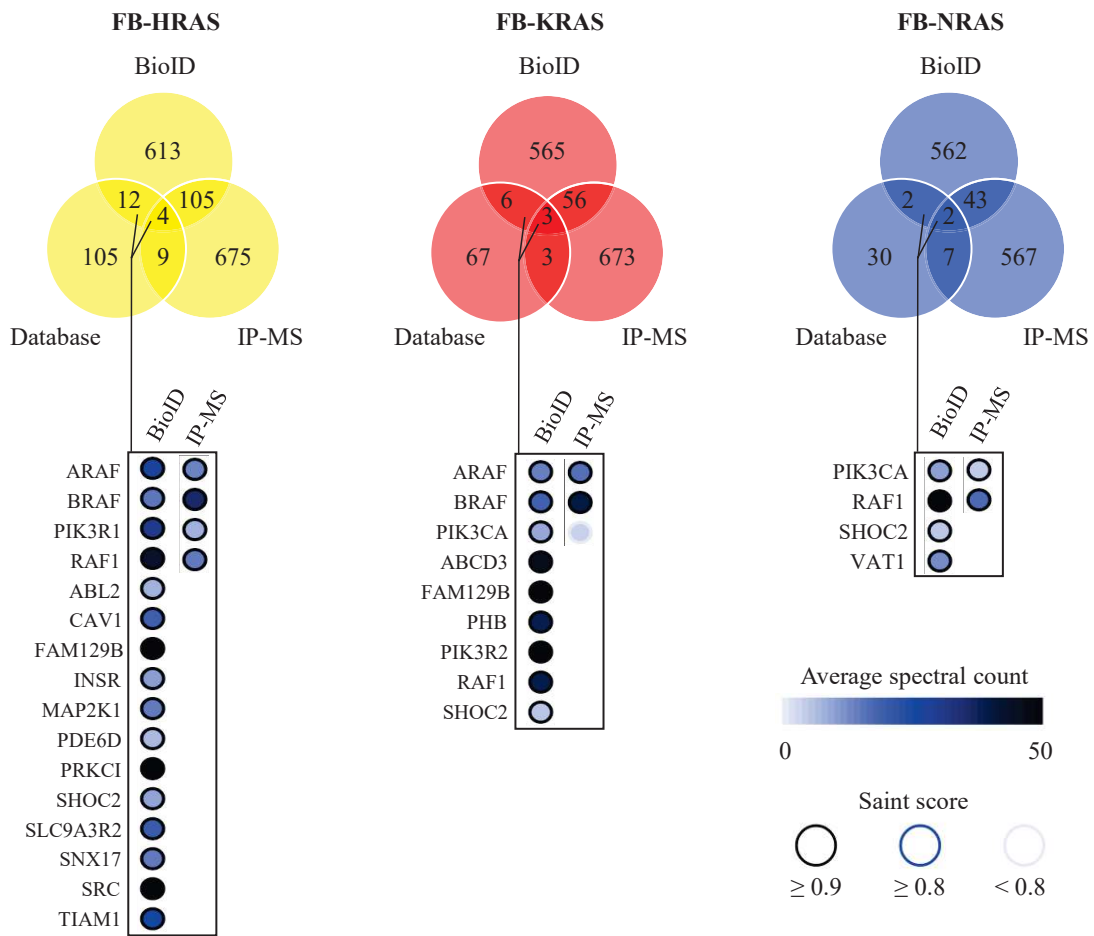


Fig. 3

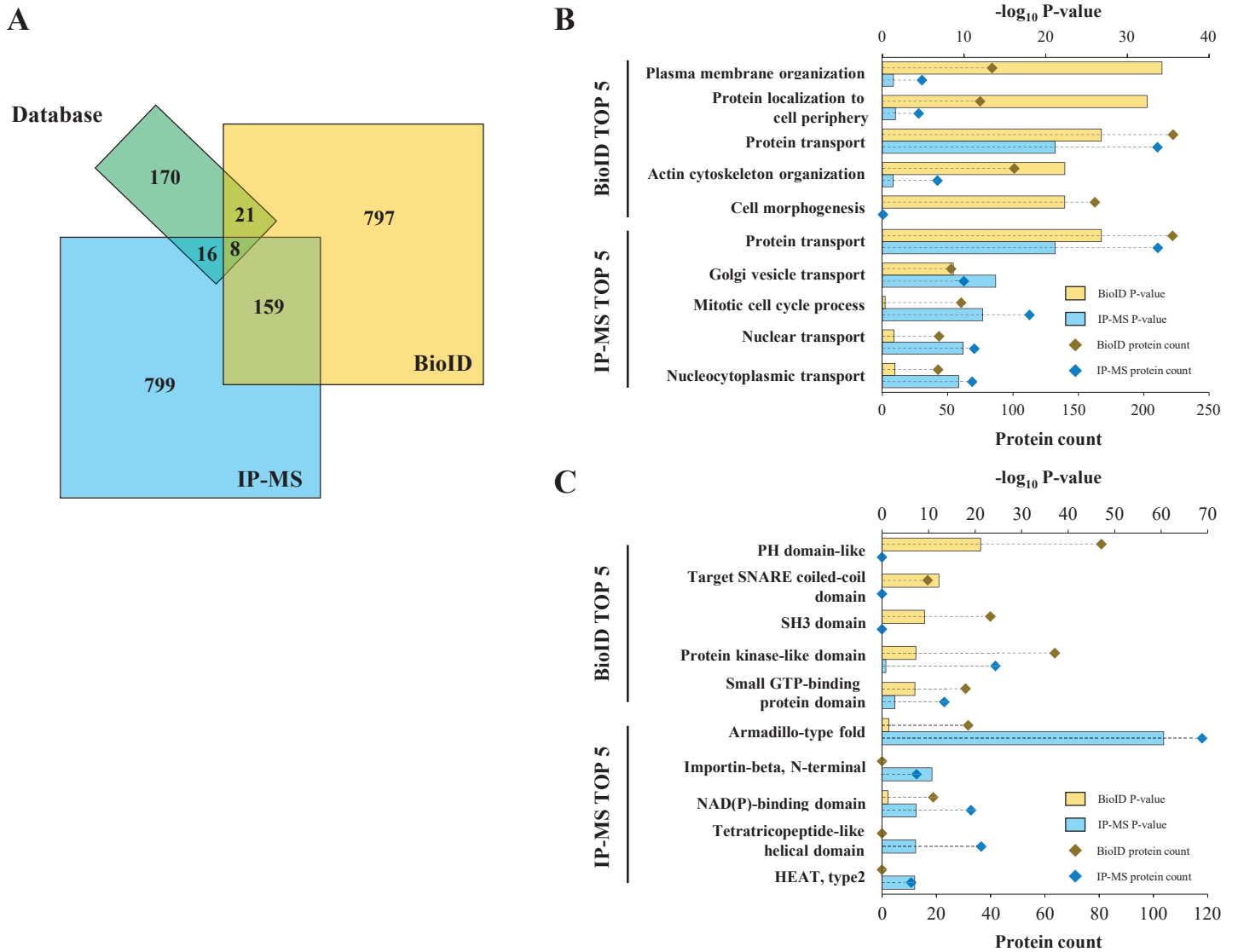


Fig. 4

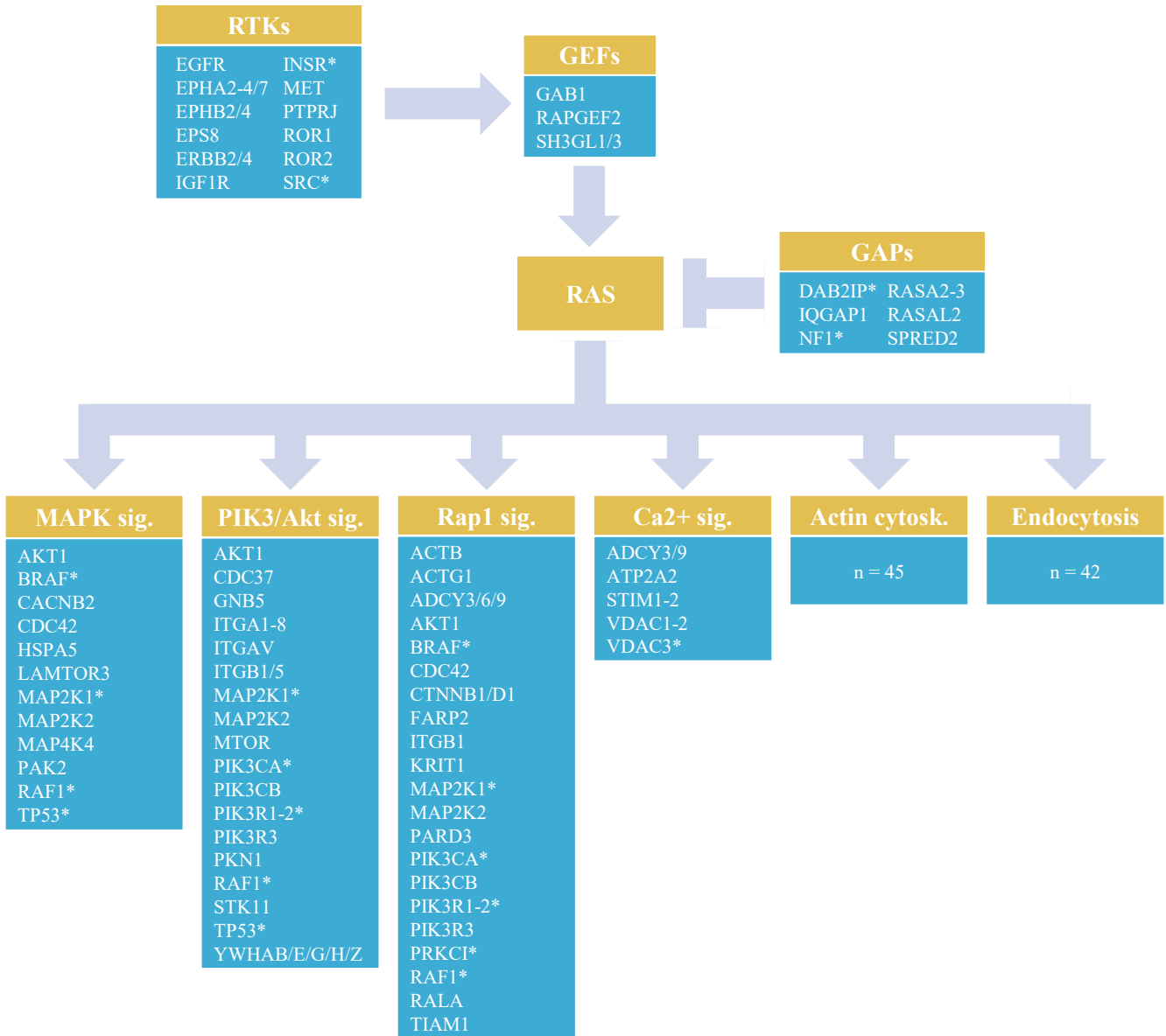


Fig. 5

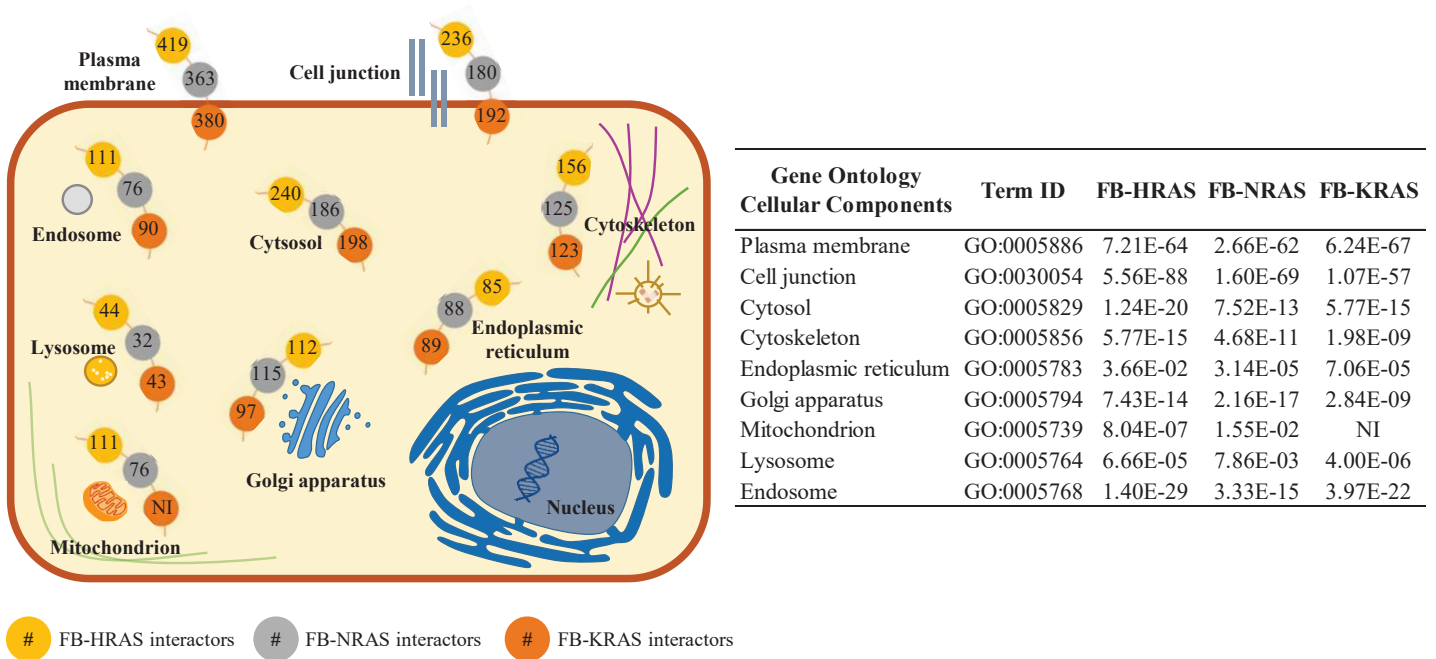


Fig. 6

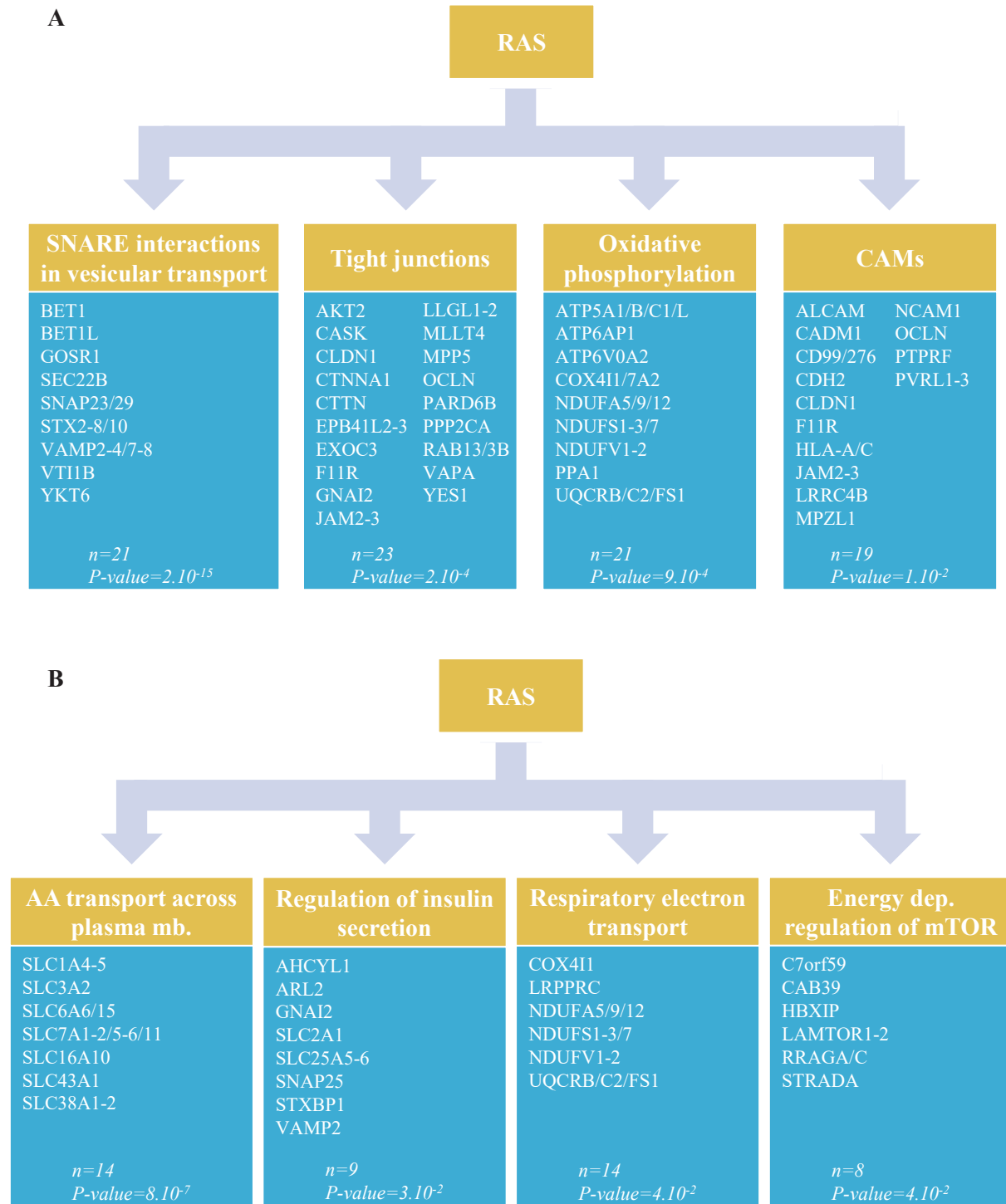
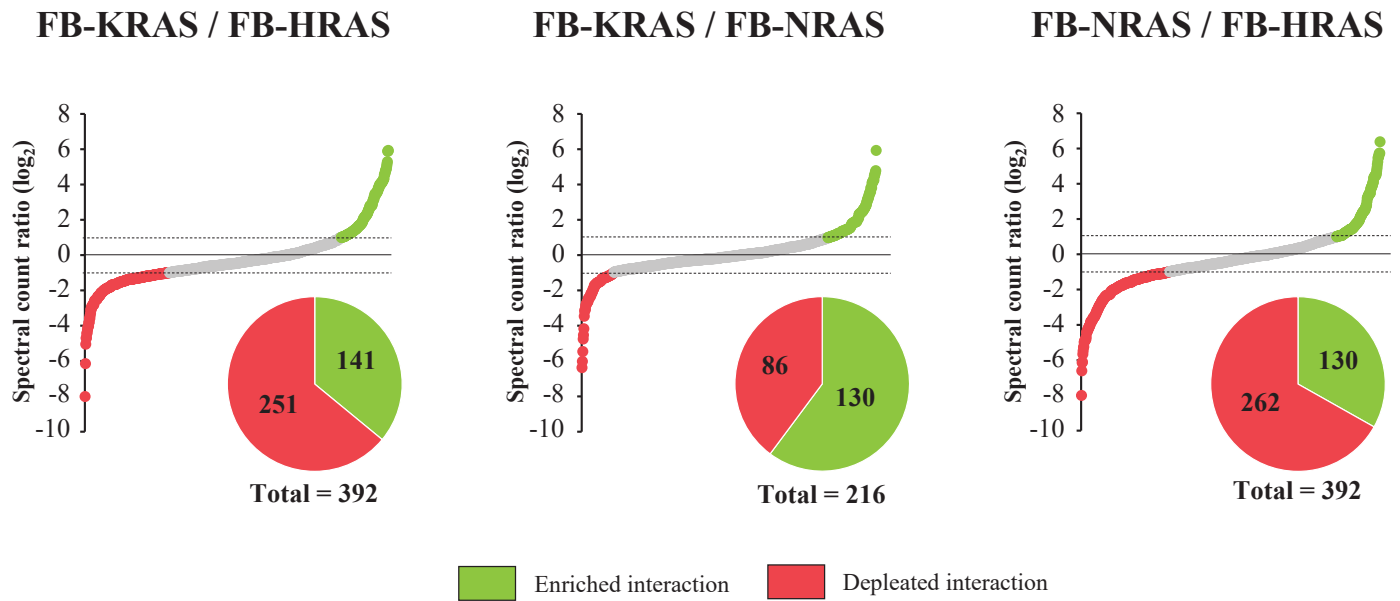


Fig. 7

A



B

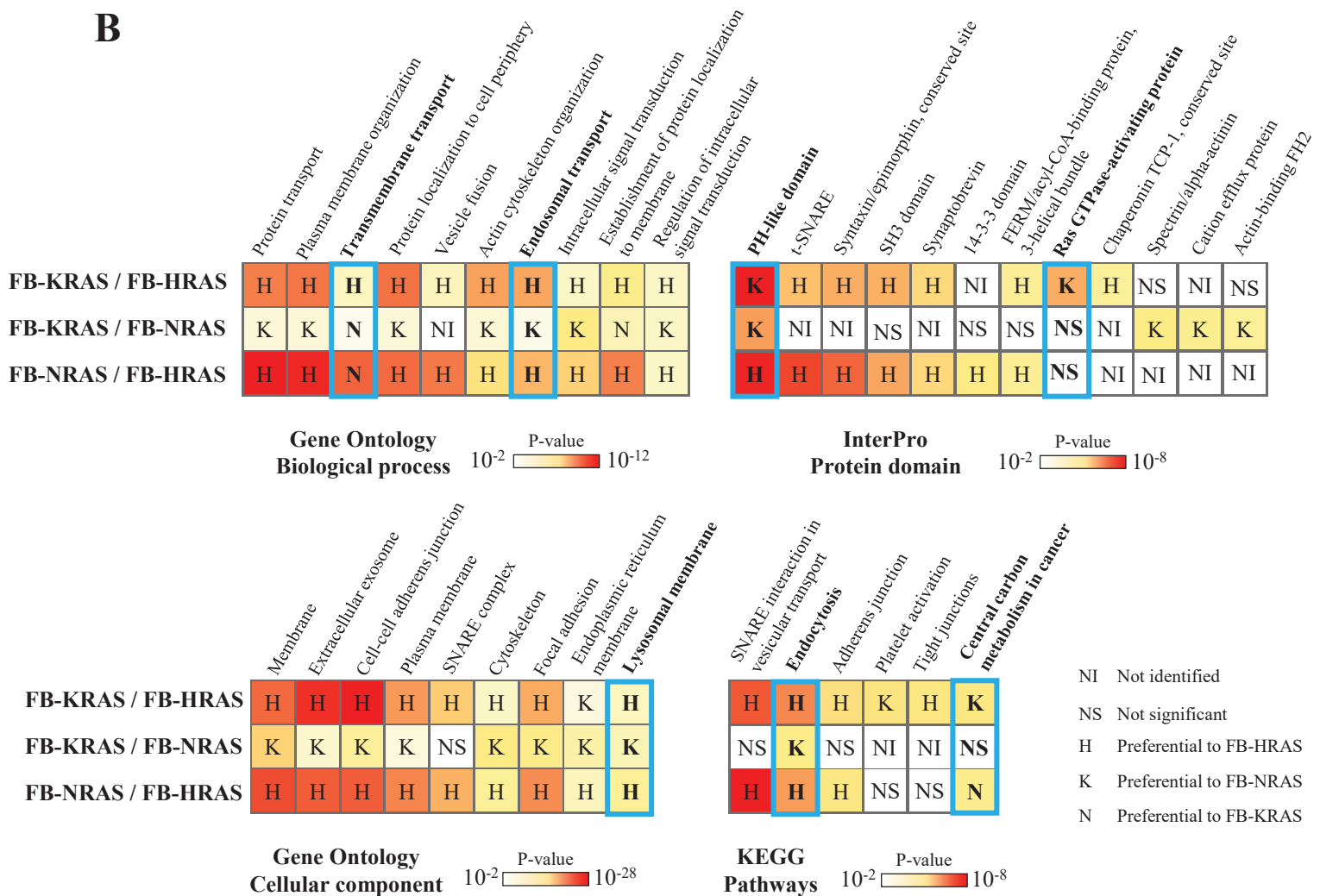


Fig. 8

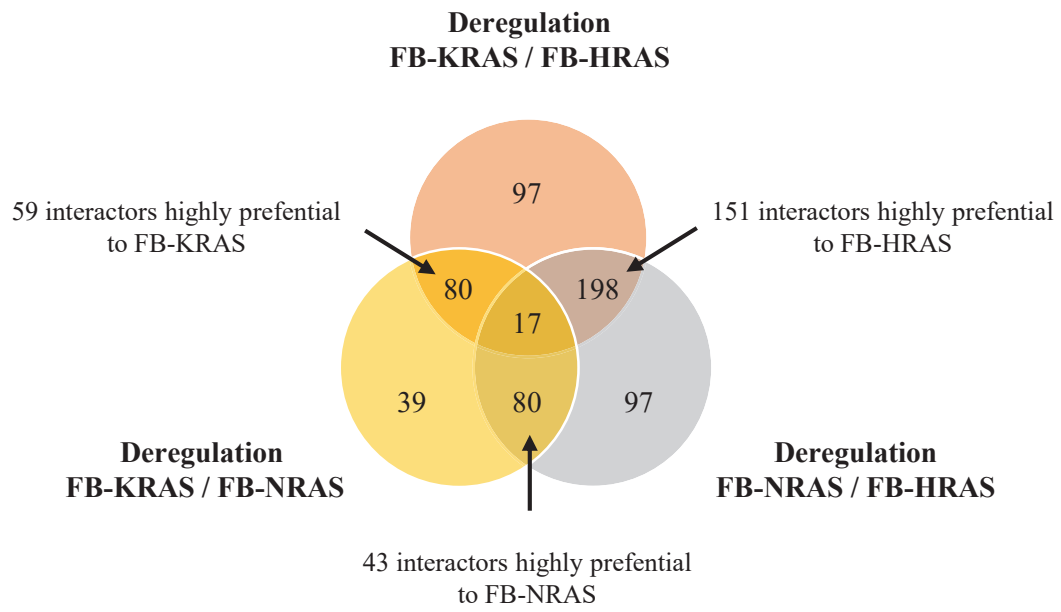


Table 1

	GO BP term	Count	P-value	Proteins identified
Specific to FB-HRAS	Protein transport (GO:0015031)	44	1,47E-06	AHCYL1, ANXA2, BSG, BTN3A1, C14orf133, CHMP5-6, DYNLL1, GDI2, GIPC1, NAPA, NDUFAF2, NOP58, PSEN1, RAB11FIP2, RAB12/21/23, SCAMP1, SEC61A1, SLC9A3R1, SNX3/6/12/17, SRC, SRP9, STX3-4/6-7/12, STXBP1, TBC1D10A, TIMM13/23, TMED10, TOMM6, VAMP2/4/8, VPS33A, YKT6, YWHAH
	Single-organism membrane fusion (GO:0060627)	13	4,91E-06	ANXA2, STX12, STX3-4/6-7, STXBP1, TC2N, VAMP2/4/8, VAT1, YKT6
	Regulation of vesicle-mediated transport (GO:0060627)	19	2,40E-05	ANXA2, B2M, CHMP6, CNN2, NAPA, NCS1, RAB21, RDX, SNX3/6/12/17, SRC, STX4, TBC1D10A, TC2N, VAMP2/8
	Endosomal transport (GO:0016197)	14	6,98E-05	C14orf133, CHMP5-6, RAB12/21, RDX, SNX3/6/12/17, STX6, TBC1D10A, VPS33A, YKT6
	Exocytosis (GO:0006887)	17	1,00E-04	CHMP6, NAPA, NCS1, PSEN1, RAB11FIP2, SCAMP1, SNX6, STX3-4, STXBP1, TC2N, TMED10, VAMP2/4/8, VPS33A, YKT6
Specific to FB-NRAS	Transmembrane transport (GO:0055085)	15	1,41E-03	AKT2, ANK3, CNKSR3, EBP, SLC2A1, SLC5A6, SLC16A10, SLC25A1, SLC30A5-6, SLC35A2/E1, UBB, ZDHHC13-17

Table 2

	InterPro term	Count	P-value	Proteins identified
	Syntaxin/epimorphin, conserved site (IPR006012)	5	2,07E-03	STX3-4/6-7/12
Specific to FB-HRAS	t-SNARE (IPR010989)	5	2,16E-03	STX3-4/6-7/12
	Synaptobrevin (IPR001388)	4	8,66E-03	VAMP2/4/8, YKT6
Specific to FB-KRAS	CRAL-TRIO domain (IPR001251)	4	2,10E-02	NF1, MOSPD2, KALRN, TRIO

Table 3

Gene ID	Protein ID	Gene Name	Top 2 controls		FB-HRAS			FB-NRAS			FB-KRAS		
					Total Spectral Count	Total Unique Peptides	SAINT	Total Spectral Count	Total Unique Peptides	SAINT	Total Spectral Count	Total Unique Peptides	SAINT
673	33188459	BRAF	1	1	58	24	1.00	108	32	1.00	73	23	1.00
3927	5453710	LASP1	3	3	68	17	1.00	4	4		3	3	
6282	5032057	S100A11	4	3	97	8	1.00	0	0		5	2	
5911	10518344	RAP2A	0	0	59	10	1.00	1	1		9	5	
10682	5729810	EBP	1	0	1	1		48	12	1.00	10	5	
92799	223005950	SHKBP1	0	0	2	2		17	9	0.99	0	0	
5226	40068518	PGD	5	4	25	10		0	0		60	23	0.89
10206	16445412	TRIM13	2	2	0	0		16	8		60	17	1.00
55737	17999541	VPS35	2	1	2	2		1	1		36	21	0.98

Fig. 9

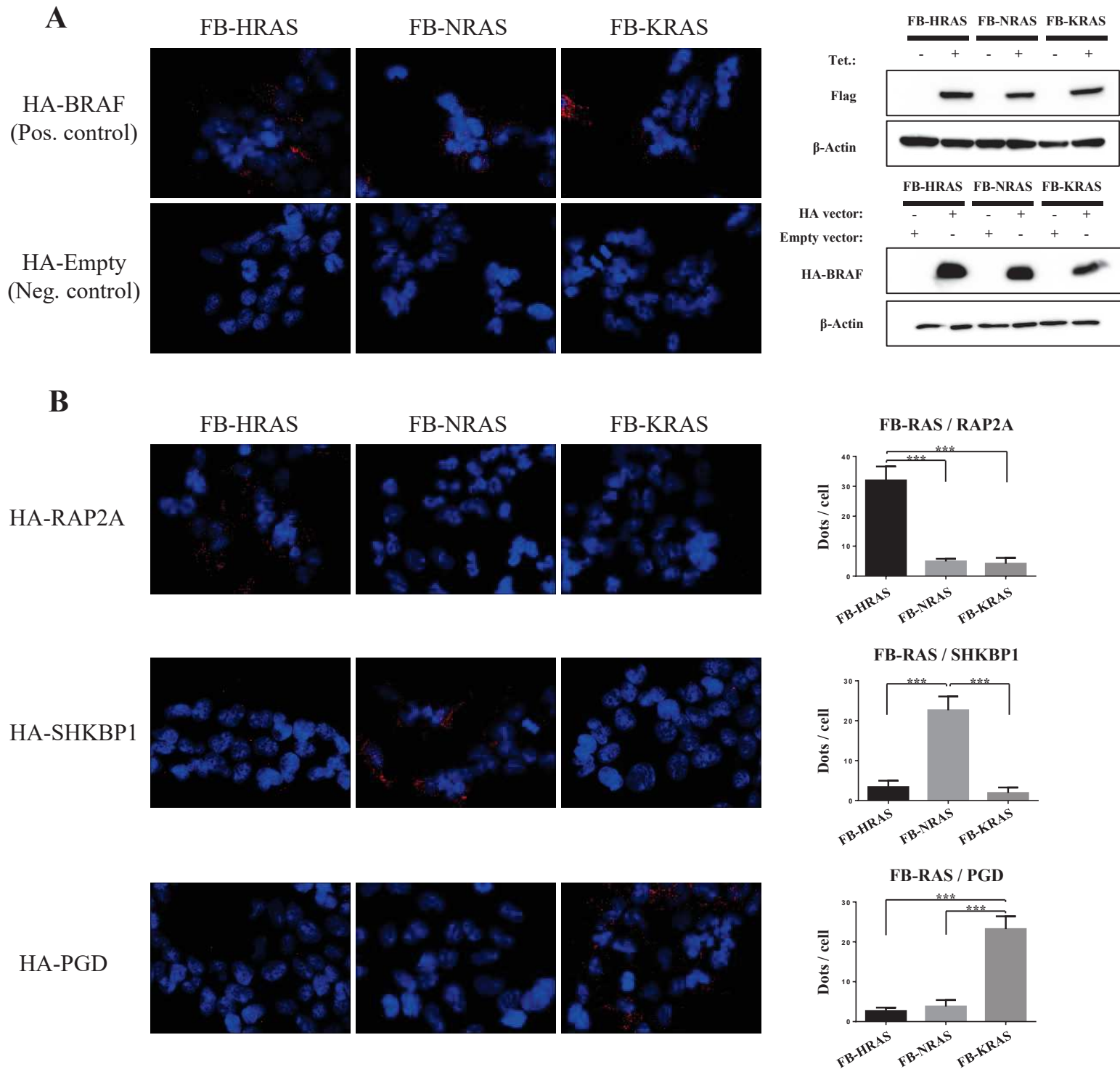
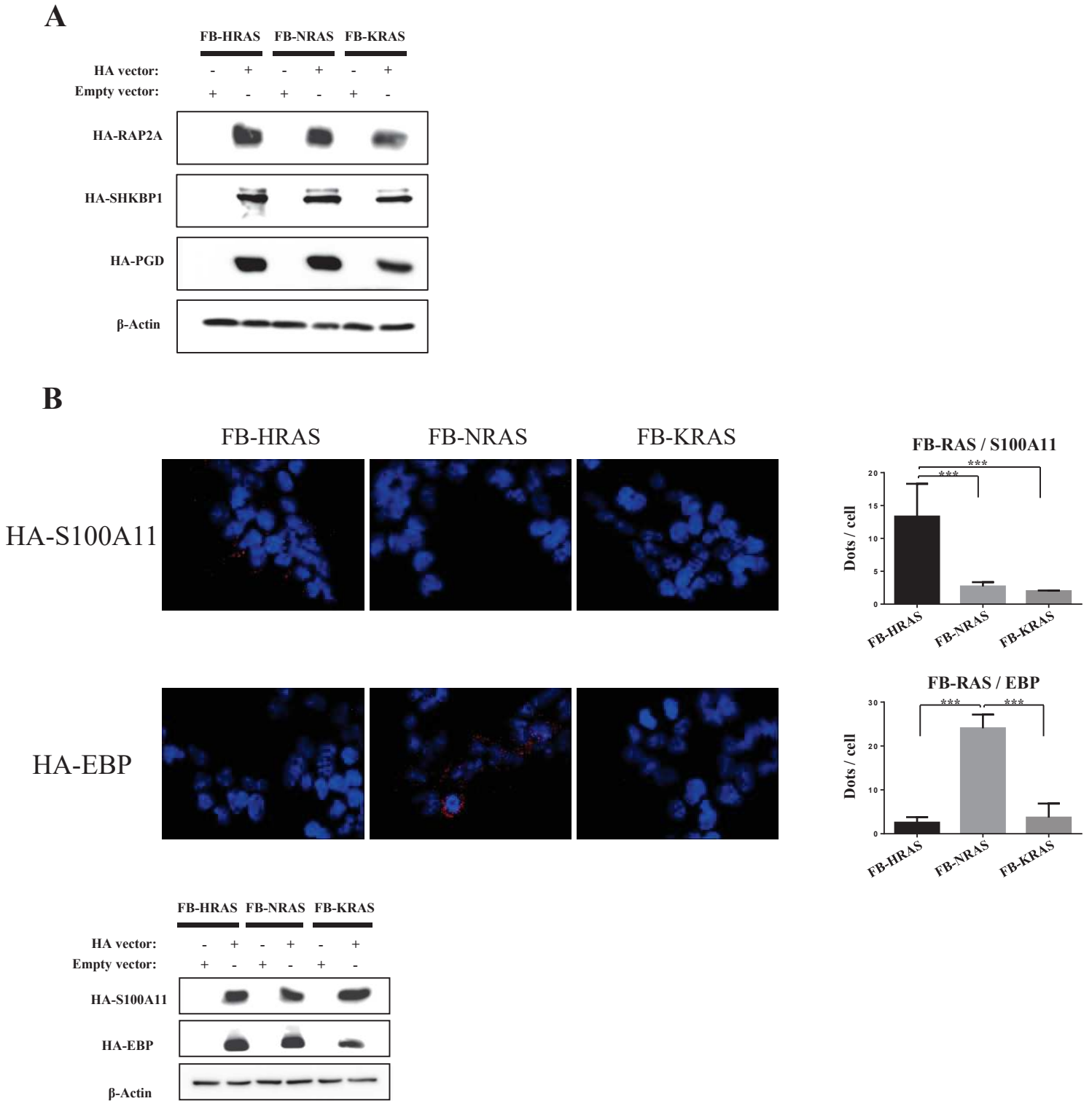


Fig. S1



III. Commentaire

En étudiant les protéines RAS avec la méthode du BioID, ce travail met simultanément en évidence deux points : (i) la fiabilité du BioID par rapport à la technique conventionnelle d'IP-MS à caractériser le réseau d'interactions protéiques de protéines ancrées à la membrane plasmique, (ii) la forte densité du réseau d'interaction protéique de RAS, avec des interacteurs impliqués dans de nombreux processus cellulaires, notamment signalisation, métabolisme, transport vésiculaire et architecture membranaire.

En effet, si les IP-MS ont plutôt relié les protéines RAS au transport vésiculaire de l'appareil de Golgi et au contrôle du cycle cellulaire, c'est-à-dire des procédés biologiques localisés dans le cytoplasme, le BioID à l'inverse montre une très forte interaction des protéines RAS avec des interacteurs membranaires. De plus, les domaines protéiques des interacteurs de RAS identifiés avec le BioID font sens avec ses mécanismes d'interactions connus. Plus précisément, la littérature décrit l'activation des RAS par une cascade de signalisation impliquant l'activation de récepteurs à activité tyrosine kinase, l'interaction de ces récepteurs avec des protéines adaptatrices portant des domaines SH2 et SH3, puis l'interaction de RAS activé avec des effecteurs portant des domaines PH ou de fixation aux petites GTPases. L'ensemble de ces domaines sont significativement identifiés en BioID et permettent de redessiner le schéma d'activation des RAS, alors que ceci n'est pas le cas pour les interacteurs identifiés par IP-MS. Logiquement, plusieurs membres de chacune des voies de signalisations dirigées par RAS ont été identifiés, ainsi que des régulateurs connus en amont de RAS et des récepteurs tyrosine kinase. De manière intéressante, plusieurs interacteurs attribués à chacune de ces voies de signalisation ou familles protéiques n'ont pour la majorité jamais été décrits comme interacteurs physiques de RAS.

Plus étonnement, cette étude a montré une très forte proximité entre RAS et la chaîne respiratoire mitochondriale, soulignant une nouvelle fois la dynamique des réseaux d'interactions. La translocation de KRAS à la mitochondrie et son implication dans la régulation du métabolisme a déjà été décrite (Bivona et al., 2006; Neuzil et al., 2012; Padanad et al., 2016; Weinberg et al., 2010), mais pas la proximité avec l'ATP synthase et la NADH-ubiquinone oxidoreductase. Des expériences complémentaires pour valider ces interactions sont cependant nécessaires (par exemple PLA, BioID des membres de ces complexes) mais cette identification supporte des axes d'étude de RAS à l'interface de la signalisation cellulaire et du métabolisme aujourd'hui moins développés.

Malgré leurs modifications post-traductionnels et leurs adressages à la membrane plasmique distincts, les différentes protéines RAS partagent une grande partie de leur réseau d'interaction. Basé seulement sur le score SAINT, 376 protéines sont communes aux différents isoformes, soit près de 40% des interacteurs identifiés pour chacun des isoformes. Ces interacteurs communs comprennent d'une part les membres des voies de signalisations majeures activées par les protéines RAS (notamment les voies MAPK, PIK3, régulation du cytosquelette d'actine), et valident ainsi l'approche. Les particularités

des isoformes (spécificités exclusives à un isoforme ou fonctions attribuées principalement à un isoforme) se situent alors dans des voies de signalisations moins décrites comme associées aux RAS. Peu d'interacteurs sont identifiés comme fortement préférentiels ou spécifiques à un isoforme en particulier. Peu de catégories significatives ont alors été identifiées, mais celles-ci semblent tout de même mettre en évidence que HRAS pourrait être impliqué dans les mécanismes de transport vésiculaire (endocytose ou exocytose) alors que NRAS serait plus en lien avec le transport transmembranaire, qu'il s'agisse de molécules organiques ou inorganiques et localisées à la membrane plasmique ou aux endomembranes. Enfin, seul le domaine protéique CRAL-TRIO a été significativement identifié comme spécifique aux interacteurs de KRAS. Il s'agit d'un domaine d'interaction avec le groupement prenyl des RAS et seulement 4 protéines ont été annotés à ce domaine, avec parmi elles NF1 qui est un régulateur majeur des RAS et fréquemment muté dans les cancers.

Devant le nombre important d'interacteurs identifiés pour les différents isoformes, une validation à large échelle n'a pas été envisageable. C'est pourquoi j'ai sélectionné 8 protéines en interaction avec un isoforme en particulier (3 interactions préférentiels de HRAS, 2 interactions préférentiels de NRAS et 3 interactions préférentiels de KRAS). Parmi l'ensemble de ces interacteurs, seulement 2 ont été identifiés avec la technique d'IP-MS (LASP1 et PGD). Sur les 8 interacteurs sélectionnés, 5 ont été validés, comprenant PGD et 4 interacteurs non identifiés en IP-MS. La validation de ces 4 interacteurs identifiés seulement en BioID confirme la spécificité de la technique employée, et la fiabilité des analyses présentées.

IV. Discussion

Le BioID apparaît comme une méthode de choix pour l'étude des interactions protéiques, en rupture avec les techniques plus conventionnelles comme l'IP-MS. Le BioID nécessite tout de même de faire des protéines chimères, et précisons ici que BirA* est de taille supérieure aux protéines RAS (32 kDa contre 24 kDa). Tous les interacteurs que j'ai pu identifier ici ne sont alors pas nécessairement en interaction directe avec RAS mais dans un environnement très proche (<10 nm), ce qui est suffisant pour rendre compte de l'environnement protéique dans lequel la protéine cible agit. En plus des faux-positifs dû à la biotinylation des protéines environnantes à RAS, une autre source de faux-positifs concerne les protéines interagissant avec Flag-BirA*. Bien que le contrôle Flag-BirA* seul ait pour but d'éliminer ces interactions, le fait que celui-ci soit cytosolique et non membranaire comme les protéines RAS conduit probablement à une élimination partielle des interactions non spécifiques à RAS. L'approche semi-quantitative employée minimise cependant la prévalence de ces interactions non spécifiques aux protéines RAS, tout en mettant en évidence les spécificités de chacun des isoformes.

Ce travail a mis en évidence la richesse et la densité du réseau d'interaction des RAS tant au niveau fonctionnel qu'au niveau des localisations cellulaires, avec des interacteurs dans tous les compartiments

cellulaires sauf le noyau (membrane plasmique, lysosomes, mitochondries, appareil de Golgi). La répartition spatiale des protéines amène également à considérer une possibilité de modulation du signal de RAS selon leur localisation. Ces localisations multiples ont déjà été décrites pour certains isoformes. Ces isoformes y seraient actifs et induiraient des signalisations similaires à celle des protéines RAS localisées à la membrane plasmique, ou activeraient d'autres voies plus spécifiques au compartiment cellulaire. Par exemple, RAS peut être localisé au niveau des endosomes et l'interaction à cette localisation avec RAF1 et certaines protéines régulatrices GAP est préservée (Grunwald et al., 2013; Kofler-Geles et al., 2009). Cependant, en étant établi dans un contexte protéique différent, l'activation de la voie RAS ne produirait pas nécessairement le même signal mais ce point n'a pas été clairement élucidé par les auteurs (Grunwald et al., 2013; Kofler-Geles et al., 2009). Une signalisation spécifique à un compartiment cellulaire est un point non négligeable car des voies majeures peuvent y être contrôlées. La littérature mentionne par exemple que l'interaction de KRAS avec la mitochondrie pourrait contrôler la mort cellulaire (Neuzil et al., 2012).

Dans notre étude, il semble que les interacteurs identifiés décrivent des sous-ensembles protéiques répartis dans différents compartiments cellulaires, et que les protéines RAS rencontrent de manière séquentielle ou concomitante. Il est intéressant de voir que des groupes d'interacteurs localisent KRAS à proximité des mitochondries, notamment avec la mise en évidence de l'interaction avec l'ATP synthase (complexe V) et la NADH-ubiquinone oxidoreductase (complexe I). Ces deux complexes de la chaîne respiratoire mitochondriale relient KRAS avec la mitochondrie et le métabolisme.

Un groupe Chinois a déjà démontré en 2012 cette proximité entre KRAS et le métabolisme mitochondrial, notamment par l'intermédiaire du complexe I. Notamment, les chercheurs ont montré que la mutation KRAS G12V dérégulait l'activité mitochondriale en faisant basculer le métabolisme d'une phosphorylation oxydative à la glycolyse. Plus particulièrement, ils ont observé la suppression de l'activité du complexe I, et la réduction du potentiel membranaire par la dérégulation de la protéine MTP (Permeability Transition Pore) (Hu et al., 2012). Le potentiel de membrane est ce qui permet le flux de proton par le complexe V et *in fine* la génération d'ATP. Nos expériences n'ont pas permis d'identifier la protéine MTP. Toutefois, cette étude et la nôtre soulignent l'implication de RAS dans le métabolisme mitochondrial, ainsi que sa vulnérabilité lors de mutations activatrices - au moins concernant KRAS. Parallèlement, le métabolisme carboné relatif à la glycolyse et à la chaîne respiratoire mitochondriale a également été mis en évidence par l'identification de l'ontologie KEGG « Central carbon metabolism in cancer » plus spécifique de KRAS et NRAS. Identifier ce terme appuie cette double connexion métabolisme cytoplasmique et mitochondrial et la possibilité du basculement de l'un à l'autre dans un contexte oncogénique.

À l'heure actuelle, aucune étude n'a décrit les réseaux d'interactions protéiques de l'ensemble des isoformes de RAS dans un contexte physiologique, et de façon aussi exhaustive. Notre étude est la

première à le faire, et il reste toutefois qu'elle a été réalisée dans un modèle cellulaire bien défini et que ces réseaux hautement dynamiques pourraient être différents dans un autre contexte cellulaire. Il serait alors intéressant de caractériser le cœur invariable de ces interactions des isoformes de RAS en analysant plusieurs types cellulaires et de définir ainsi les interactions essentielles aux protéines RAS et les interactions modulables et propres à un contexte cellulaire donné.

Un autre point complémentaire à cette étude serait la recherche de complexes protéiques. En effet, ceci a été peu abordé mais à des complexes protéiques sont associés des fonctions cellulaires qu'il aurait été intéressant de valider fonctionnellement. Nous avons par exemple identifié les complexes I et V, et nous pouvons évaluer l'effet de l'extinction des protéines RAS sur le flux de proton de la chaîne respiratoire mitochondriale, ou sur la génération d'ATP par le complexe V.

C. Réseau d'interaction de KRAS : application au NSCLC

I. Introduction de l'étude

Étant donné la prévalence des mutations *KRAS* dans les adénocarcinomes pulmonaires (environ 30% des cas – mutation la plus fréquente) et l'absence de thérapies ciblées, nous nous sommes intéressés à la caractérisation du réseau d'interactions protéiques de la protéine sauvage ainsi que de la protéine mutée G12C. Par l'approche BioID, deux types cellulaires ont été utilisés et en confrontant des dérégulations d'interactions respectivement identifiées, une liste d'interacteurs fiables a été générée et a fait l'objet de validation fonctionnelle. Les données de MS sont disponibles en Annexe 3 et 4.

II. Résultats

II.1 Modèles cellulaires

Comme indiqué dans la première partie des résultats, j'ai choisi d'établir les réseaux d'interactions protéiques de KRAS dans la lignée humaine HEK293, modèle privilégié pour l'étude des interactions protéiques, et dans un modèle pulmonaire d'origine non cancéreuse, la lignée BEAS2B. Une fois les lignées établies, j'ai vérifié l'expression des protéines recombinantes par western bot (Figure 27) :

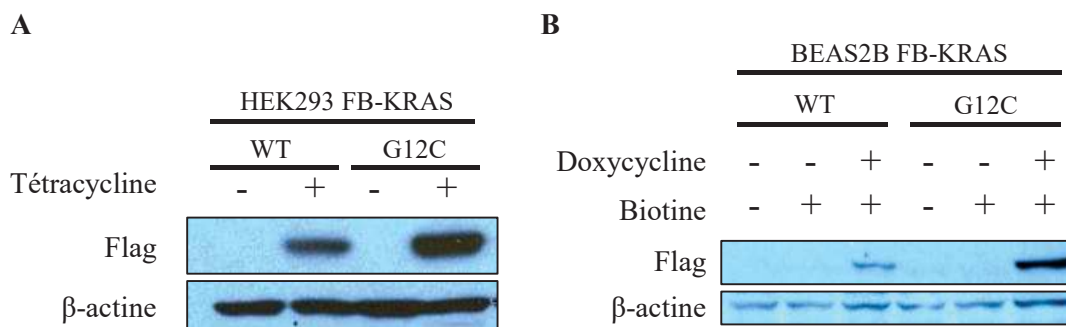


Figure 27. Validation de l'expression du transgène KRAS WT et G12C. (A) Induction de la protéine FB-KRAS WT et G12C dans la lignée HEK293. (B) Induction de la protéine FB-KRAS WT et G12C dans la lignée BEAS2B. FB, FlagBirA*.

Nous observons que l'expression de ces protéines est très bien contrôlée par la tétracycline ou la doxycycline, et que la fuite d'expression du promoteur est négligeable. Qu'il s'agisse de la lignée HEK293 ou BEAS2B, nous remarquons également que la mutation G12C augmente la quantité de protéines dans les cellules par rapport à la protéine sauvage. Il peut s'agir d'une augmentation de stabilité de la protéine mais des expériences complémentaires seraient nécessaires afin de mieux comprendre les mécanismes cellulaires impliqués.

II.2 Description des réseaux d'interactions protéiques

Sur la base des protéines identifiées dans chaque lignée, et en tenant compte de la spécificité des interactions à partir du score SAINT, j'ai comparé les réseaux d'interactions des protéines KRAS WT et KRAS G12C (Figure 28).

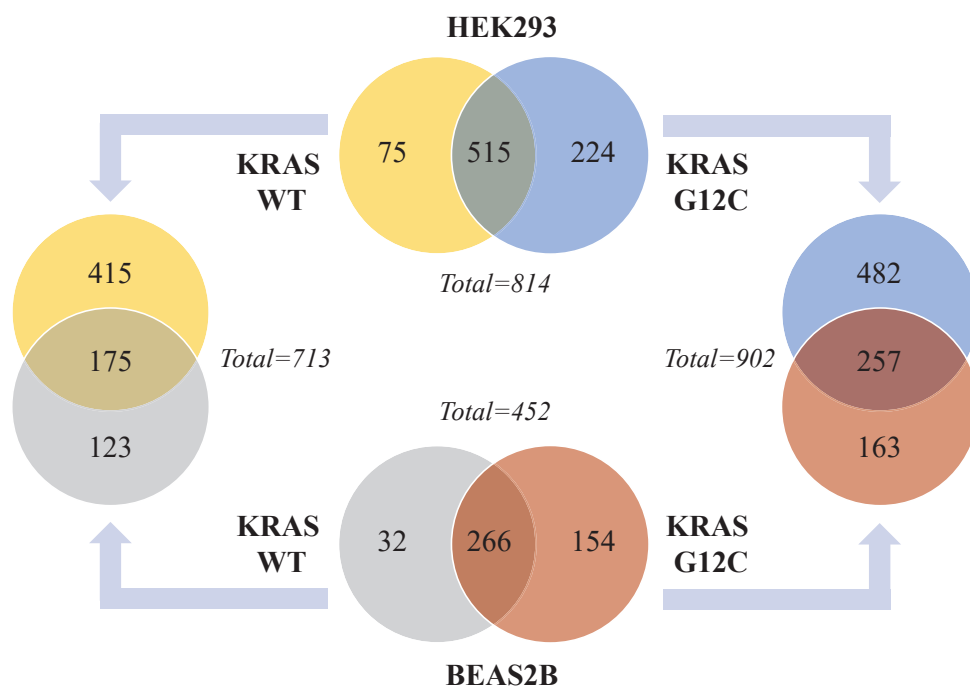


Figure 28. Comparaison des interacteurs de KRAS identifiés à partir des modèles cellulaires HEK293 et BEAS2B. Les diagrammes de Venn alignés au centre comparent les interacteurs de la protéine KRAS WT (gauche) à la protéine mutée G12C (droite) identifiés dans les HEK293 (haut) et BEAS2B (bas). Le diagramme à gauche et droite compare les interacteurs de KRAS WT et G12C, respectivement, identifiés dans les deux lignées.

Les diagrammes de Venn montrent que moins d'interacteurs sont retrouvés à proximité de KRAS (muté ou non) dans la lignée BEAS2B comparé à la lignée HEK293 (452 contre 814). Cependant, quelle que soit la lignée cellulaire utilisée, sur l'ensemble des interactions spécifiques identifiées avec KRAS sauvage ou muté, environ 60% des interacteurs sont communs entre les deux protéines. Comparé à l'ensemble des interactions identifiées, nous observons aussi que la protéine mutée montre un nombre d'interacteurs spécifiques plus important que la protéine sauvage : 28% contre 9% dans les HEK293 et 34% contre 7% dans les BEAS2B. En comparant les interacteurs identifiés avec les HEK293 à ceux identifiés en utilisant les BEAS2B, pour la protéine sauvage ou mutée, on observe que 60% des interacteurs des BEAS2B sont identifiés à partir des HEK293 (175/298 pour KRAS WT et 257/420 pour KRAS G12C). Les cellules HEK293 apparaissent donc comme un bon modèle d'étude des interactions protéiques. Les interacteurs identifiés en utilisant les BEAS2B mais non identifiés à partir des HEK293, et réciproquement, suggèrent des spécificités propres à ces modèles cellulaires.

II.3 Comparaison des interacteurs de KRAS WT

Afin de vérifier la similarité des fonctions biologiques et des localisations identifiées entre les lignées cellulaires HEK293 et BEAS2B exprimant la protéine sauvage, une analyse des termes GO associés aux interacteurs identifiés a été réalisée. Un intérêt a aussi été porté sur les complexes protéiques mis en évidence.

II.3.a *Étude des localisations et fonctions biologiques*

La figure 29 suivante présente le TOP5 des termes GO enrichis dans chaque lignée et associés à la localisation cellulaire et aux fonctions biologiques des interacteurs de KRAS WT. Nous remarquons que les localisations enrichies sont identiques entre les HEK293 et BEAS2B, et avec une très forte significativité ($P\text{-value} < 10^{-10}$). Qu'il s'agisse de la lignée HEK293 ou de la lignée de poumon, les interacteurs identifiés sont principalement localisés à la membrane plasmique (termes « cytoplasmic part », « plasma membrane region » et « plasma membrane part »), aux jonctions cellulaires (« focal adhesion ») ou au niveau des exosomes (« extracellular exosomes »). Le classement des trois ontologies les plus significatives varie toutefois d'une lignée à l'autre avec notamment le terme « adhésions focales » qui est la troisième ontologie significative de la lignée HEK293 et la première de la lignée BEAS2B.

Concernant les procédés biologiques enrichis dans les deux lignées, ils sont également très similaires. De manière cohérente avec des localisations membranaires, ils impliquent les protéines identifiées dans le transport vésiculaire (« protein localization to cell periphery ») et dans l'architecture de la membrane (« cell morphogenesis », « plasma membrane organization »). Des procédés faisant sens avec l'activité de RAS sont également identifiés et ceci plus particulièrement dans la lignée BEAS2B. Il s'agit de la catégorie « enzyme linked receptor protein signaling pathway » enrichies avec le réseau protéique de la lignée HEK293 et « small GTPase mediated signal transduction » enrichi dans la lignée BEAS2B. Deux ontologies sont cependant différentes dans le TOP5 des deux lignées (« cell migration » et « actin cytoskeleton organization »), mais ces termes sont retrouvés lorsqu'on élargi la recherche au TOP10 des fonctions représentées.

En résumé, les réseaux protéiques de KRAS WT identifiés dans les lignées HEK293 et BEAS2B sont globalement très similaires en termes de localisation des interactions et des fonctions biologiques impliquées.

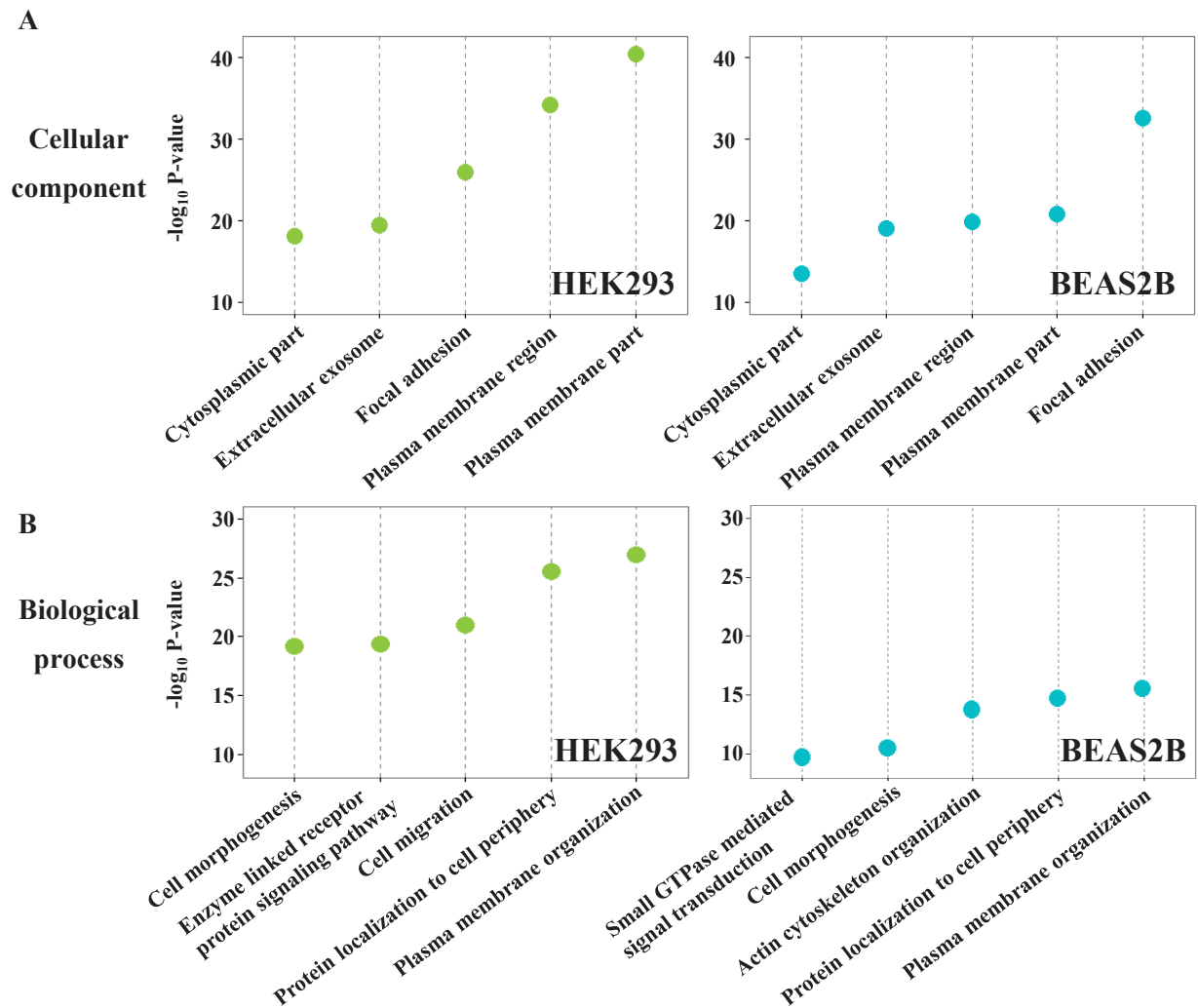


Figure 29. Analyse des localisations et fonctions des interacteurs de KRAS WT. **(A)** La comparaison des TOP5 des composants cellulaires (cellular components) identifiés à partir des interacteurs des HEK293 et BEAS2B montre des localisations identiques impliquant principalement la membrane plasmique. **(B)** Les procédés biologiques (biological process) montrent des fonctions cellulaires associés à la membrane ainsi que liés à la régulation de KRAS. Analyse réalisée avec Gene Ontology (niveau 5) et les ontologies significatives ont été filtrées par Revigo.

II.3.b Mise en évidence des complexes protéiques

Une analyse plus détaillée (niveau 5) des termes GO composant cellulaires permet de mettre en évidence la présence de complexes protéiques en lien avec la protéine KRAS WT (Figure 30).

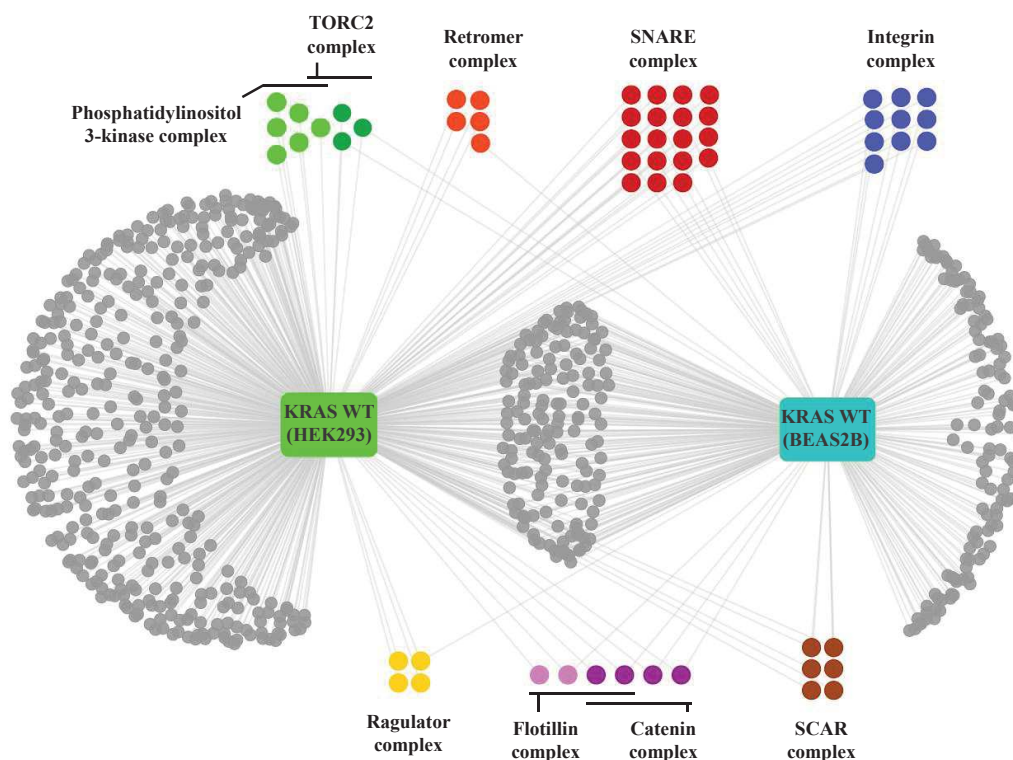


Figure 30. Complexes protéiques enrichis parmi les interacteurs de KRAS WT. L'analyse des interacteurs de KRAS WT mis en évidence avec les lignées cellulaires HEK293 et BEAS2B a permis d'identifier 9 complexes protéiques à proximité de KRAS. Analyse réalisée avec Gene Ontology (niveau 5).

Lorsque la totalité des interacteurs de KRAS WT identifiés dans les deux lignées est analysée, 9 complexes protéiques sont enrichis significativement et regroupent 4 fonctions plus générales : le trafic membranaire (complexes SNARE, retromer, flotillin), les adhésions cellulaires (complexe integrin, catenin), la régulation du cytosquelette d'actine (complexe SCAR) et la signalisation cellulaire/métabolisme (complexes PI3K et TORC2). Ces fonctions rejoignent les procédés biologiques et localisation précédemment identifiés. Précisons toutefois que tous ces complexes ne sont pas forcément enrichis de la même façon dans les lignées HEK293 et BEAS2B. Si les complexes SNARE, integrin et catenin sont partagés entre les deux lignées, alors que les complexes PI3K, retromer, SCAR, flotillin et regulator sont plus spécifiques de la lignée HEK293. Le complexe TORC2 est à l'inverse plus spécifique de la lignée BEAS2B.

Plus généralement, cette analyse montre que l'identification des complexes protéiques à proximité de KRAS WT est plus large dans la lignée HEK293 que dans la lignée BEAS2B. Cependant, dans les deux lignées, un lien très fort avec le trafic membranaire, les adhésions cellulaires et l'interface signalisation/métabolisme est clairement établi. Ces redondances d'interactions entre différents modèles cellulaires soulignent la pertinence de ces interacteurs et la pertinence des fonctions enrichies. De plus, si nous prenons l'exemple du complexe regulator (au total composé de 6 protéines), nous remarquons

que les 4 interacteurs sont identifiés avec les HEK293 et un seul avec les BEAS2B. L'analyse du réseau seulement obtenue avec les BEAS2B n'aurait pas identifié significativement ce complexe, montrant ainsi la complémentarité des modèles cellulaires afin de construire un réseau protéique pertinent de KRAS.

II.4 Perturbation du réseau d'interaction protéique de KRAS WT par la mutation KRAS G12C

II.4.a Mise en évidence des complexes protéiques

Afin d'évaluer l'impact de la mutation G12C sur les réseaux d'interactions de KRAS, j'ai comparé les réseaux d'interactions KRAS WT et KRAS G12C générés à partir des deux lignées cellulaires HEK293 et BEAS2B. De nouveaux complexes protéiques ont été identifiés (Figure 31).

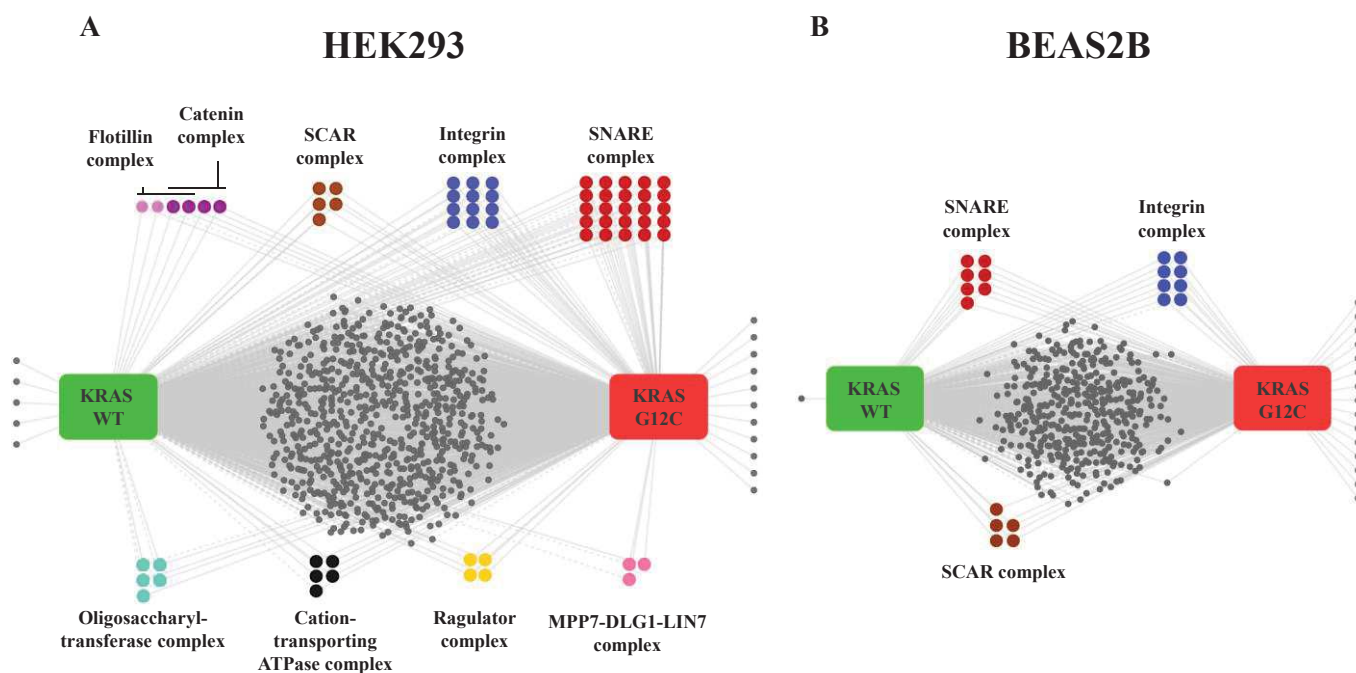


Figure 31. Complexes protéiques significatifs parmi les interacteurs de KRAS WT et G12C. (A) Dans les HEK293, l'analyse des complexes protéiques à proximité de KRAS identifie 9 complexes significativement présents, indépendamment du statut mutationnel. (B) Dans les BEAS2B, l'analyse des complexes protéiques à proximité de KRAS identifie 3 complexes significativement présents, indépendamment du statut mutationnel. Les liens en trait plein correspondent à une interaction avec un score SAINT significatif, et ceux en pointillés à un score SAINT non-significatif mais avec des peptides identifiés pour l'interacteur considéré. Analyse réalisée avec Gene Ontology (niveau 5).

L'analyse des complexes protéiques présents parmi les interacteurs de KRAS WT ou G12C identifiés dans les HEK293 met en évidence 9 complexes significatifs, parmi lesquels 6 sont identiques à ceux déjà identifiés avec l'analyse comparative HEK293 et BEAS2B de la protéine sauvage. Les 3 autres complexes enrichis sont le complexe « Oligosaccharyl-transferase », « MPP7-DLG1-LIN7 » et le complexe « Cation-transporting ATPase ». Ceux-ci sont respectivement impliqués dans la N-glycosylation à la surface du réticulum endoplasmique, dans la stabilisation des jonctions cellulaires et dans la chaîne respiratoire mitochondriale (complexe V). L'identifier ici après ajout du réseau de la protéine mutée suggère un impact potentiel de cette mutation sur le métabolisme mitochondrial. Ce lien avec le métabolisme est conforté par l'identification d'interacteurs des complexes TORC2 et NADH dehydrogenase, dont les fonctions sont à l'interface de la signalisation cellulaire et du métabolisme.

L'analyse du réseau d'interaction des BEAS2B met en évidence des membres du complexe TORC2, mais aucun complexe protéique impliqué dans le métabolisme n'est enrichi. Seuls les complexes protéiques SNARE, integrin et SCAR sont de nouveau caractérisés significativement, ce qui suggère un lien étroit entre KRAS et ces complexes.

II.4.b Dérégulation d'interaction

À partir de la liste des interacteurs identifiés dans la lignée HEK293 (n=814, Figure 28) et BEAS2B (n=452, Figure 28), j'ai sélectionné les interacteurs dérégulés entre KRAS WT et KRAS G12C. Pour cela j'ai fait, pour chacun des interacteurs mis en évidence, le ratio de peptides obtenus entre la forme mutante et sauvage de KRAS. Si le ratio est supérieur à 2 (au moins deux fois plus de peptides identifiés pour la forme mutée, la mutation KRAS G12C enrichit cette interaction) ou inférieur à 0.5 (au moins deux fois plus de peptides identifiés pour la forme sauvage, la mutation KRAS G12C appauvrit cette interaction), alors l'interacteur est défini comme dérégulé.

Le calcul des ratios permet d'identifier 140 interacteurs dérégulés dans les cellules HEK293 et 165 interacteurs dérégulés dans les cellules BEAS2B. Dans les deux cas, ces dérégulations comprennent principalement des enrichissements ou gains d'interactions (Figure 32).

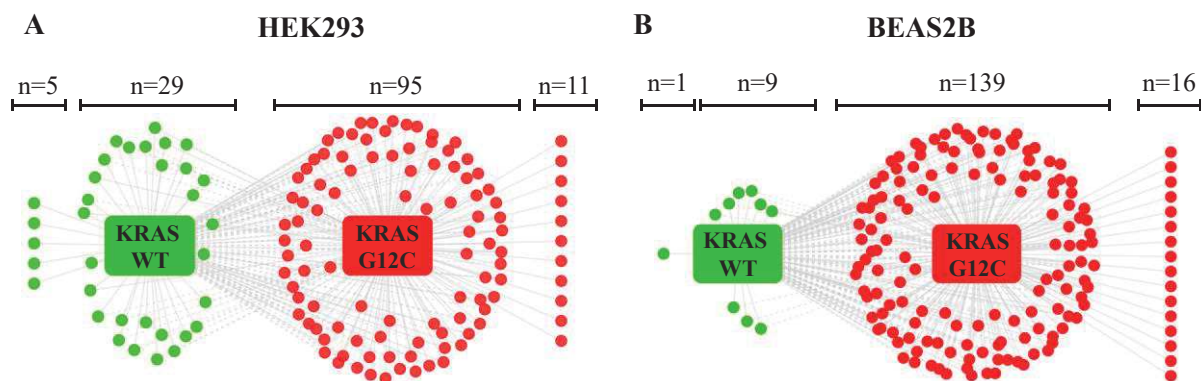


Figure 32. Dérégulations d'interactions par la mutation KRAS G12C. (A) Interactions dérégulées dans la lignée cellulaire HEK293. 5 interacteurs sont spécifiques à la protéine sauvage, 29 interacteurs lui sont préférentiels. La mutation KRAS G12C renforce les interactions avec 95 protéines et fait gagner 11 interacteurs. (B) Interactions dérégulées dans la lignée cellulaire BEAS2B. Un seul interacteur est spécifique de KRAS WT et 9 interacteurs lui sont préférentiels. La protéine mutée interagit plus fortement avec 139 interacteurs, et décrit 16 nouvelles interactions.

Une interaction enrichie comme appauvrie pouvant conduire au final au même effet biologique, l'analyse GO sur les termes procédés biologiques et localisations a été réalisée en prenant en compte l'ensemble des interacteurs identifiés comme dérégulés (Figure 33).

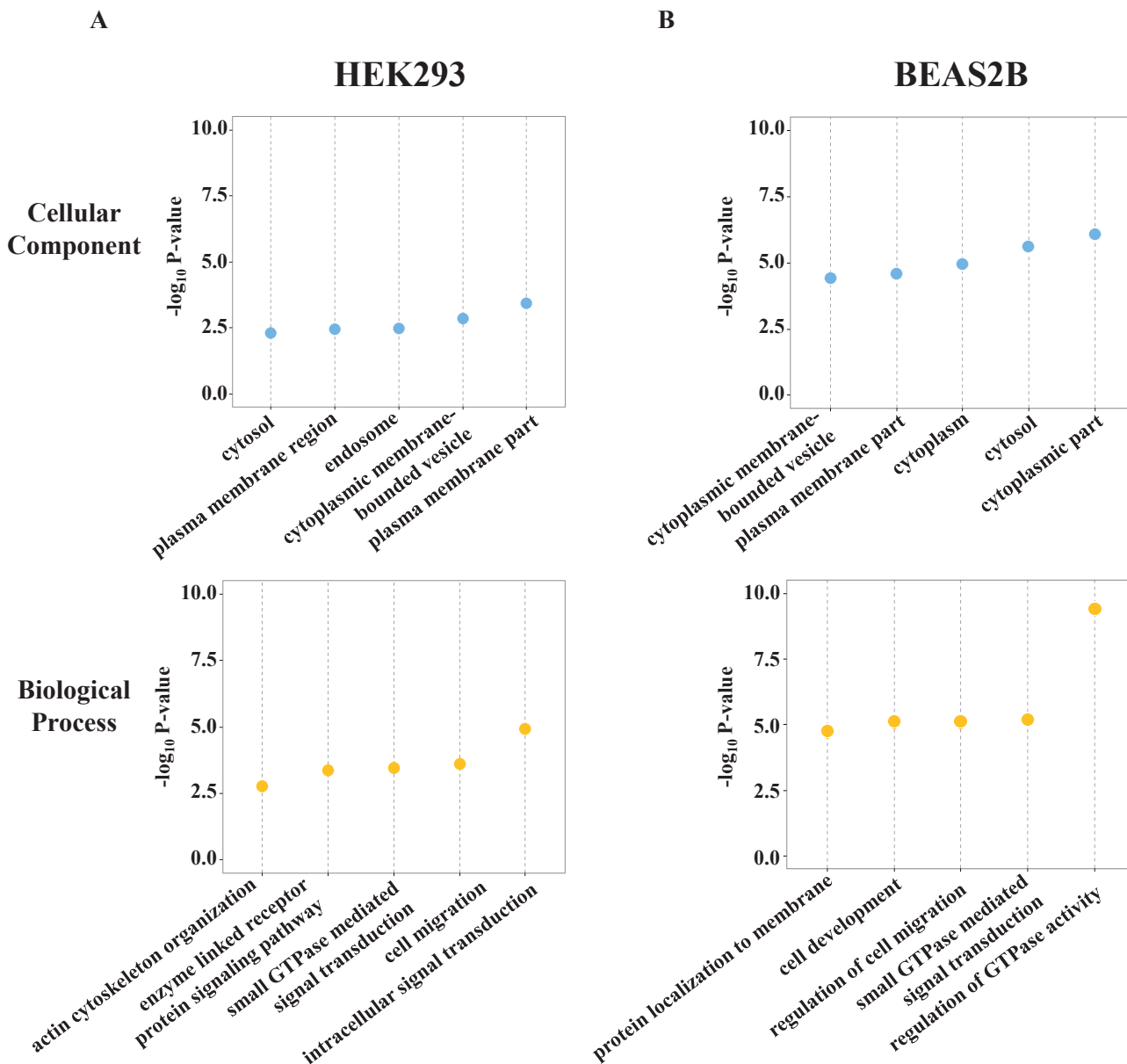


Figure 33. Analyse des localisations et fonctions des interacteurs dérégulés de KRAS. (A) TOP5 des composants cellulaires (cellular components) et procédés biologiques (biological process) identifiés à partir des interacteurs dérégulés dans le modèle cellulaire HEK293. (B) TOP5 des composants cellulaires (cellular components) et procédés biologiques (biological process) identifiés à partir des interacteurs dérégulés dans le modèle cellulaire BEAS2B. Analyse réalisée avec Gene Ontology (niveau 5), termes significatifs filtrés avec Revigo.

Malgré les termes généraux « cytosol » ou « cytoplasme », qu'il s'agisse des HEK293 ou BEAS2B, les interacteurs dérégulés apparaissent en lien étroit avec la membrane plasmique et le transport vésiculaire, comme le montre les termes « cytoplasmic membrane-bounded vesicle », « endosome » et « plasma membrane part ». L'analyse des procédés biologiques met en évidence des fonctions impliquées principalement autour de la régulation des GTPase et les voies contrôlées par ces protéines, ce qui fait sens avec le fait que KRAS est une GTPase. Le terme « cell migration » est également particulièrement intéressant étant donné qu'il s'agit d'une particularité des cancers du poumon non à petites cellules mutés KRAS (Sunaga et al., 2012). Plus globalement, cette analyse indique que la dérégulation du réseau d'interaction KRAS par la mutation G12C implique principalement ses régulateurs et effecteurs.

Une analyse plus fine des termes GO centrée sur les complexes protéiques met en évidence uniquement dans la lignée BEAS2B, deux complexes enrichis spécifiquement pour la forme mutée de KRAS. Il s'agit d'une part du complexe TORC2, qui est composé de 11 protéines dont 3 sont retrouvées préférentielles (RICTOR et MAPKAP1) ou exclusives (MTOR) de la forme KRAS G12C. D'autre part le complexe catenin, composé 9 protéines, est enrichi et les protéines CDH2, CTNNA1 et CTNNB1 sont préférentielles de la forme mutée de KRAS. Le complexe catenin est impliqué dans la régulation des adhésions cellulaires et TORC est impliqué dans la régulation du métabolisme à l'interface avec la signalisation cellulaire. Bien que ces fonctions cellulaires soient altérées lors de processus oncogéniques, ces complexes ne sont pas identifiés parmi les interacteurs dérégulés dans les cellules HEK293, réduisant la pertinence de leur exploration fonctionnelle associée à la mutation KRAS G12C.

Les interacteurs dérégulés les plus pertinents à étudier sont ceux exclusifs, et plus particulièrement ceux exclusifs à la protéine mutée. Effectivement, une perte d'interaction peut être tout autant pathogène qu'un gain d'interaction mais dans la perspective d'un traitement il est plus difficile de forcer une interaction que de l'altérer par une drogue. Le tableau 7 suivant présente les protéines qui interagissent exclusivement avec la protéine sauvage ou mutée dans les cellules HEK293 et les BEAS2B.

Tableau 7. Interacteurs exclusifs de KRAS WT ou G12C. Interacteurs exclusifs de la forme sauvage ou mutée de KRAS identifiés dans les cellules HEK293 ou les cellules BEAS2B. TRAF7 (en rouge) est spécifique de la forme sauvage dans ces deux lignées cellulaires.

Modèle cellulaire	Spécificité	Nombre de protéines identifiées	Protéines
HEK293	KRAS WT	5	ERBB4, PTPRF, SEMA6A, SEMA6D, TRAF7
	KRAS G12C	11	ARC, CYTH2, FAM131B, GRB14, GRK6, HOMER1, MPZL2, RASSF5, RELT, SLC2A3, SPRY2
BEAS2B	KRAS WT	1	TRAF7
	KRAS G12C	16	ARHGEF7, CDC42EP2, EPN1, GPR126, KIAA0368, KRIT1, MTOR, NF1, NOLC1, PPP1R37, RAPGEF2, SCARB1, SNRPD1, SPRED2, STIM1, TIAM1

Une seule protéine (TRAF7, TNF Receptor Associated Factor 7) est retrouvée dans les HEK293 et BEAS2B comme un interacteur exclusif de KRAS WT. Pour les autres interacteurs, comme attendu, ces protéines comprennent des régulateurs de GTPase comme NF1, SPRED2, RAPGEF2 et des protéines de signalisation comme ERBB4, PTPRF et MTOR. Le lien avec le métabolisme, bien que non significatif, est encore mis en évidence ici avec le transporteur de glucose 3 (SLC2A3) identifié comme spécifique à KRAS G12C dans les HEK293 et MTOR, spécifique à KRAS G12C dans les BEAS2B.

Une comparaison incluant l'ensemble des interacteurs dérégulés identifiés dans le modèle cellulaire HEK293 et le modèle cellulaire BEAS2B permet d'identifier à plus grande échelle les interacteurs dérégulés et en commun entre ces deux lignées, et permet par la même de disposer d'une liste réduite d'interacteurs qui pourront être ensuite étudiés spécifiquement (Figure 34).

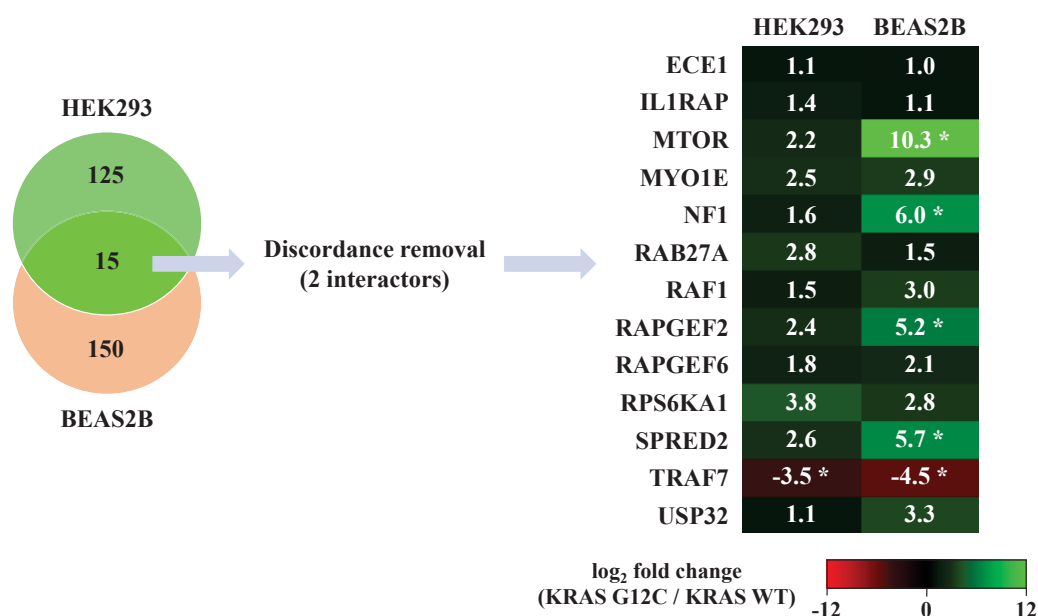


Figure 34. Sélection des interacteurs dérégulés et partagés entre les modèles cellulaires HEK293 et BEAS2B. Le diagramme de Venn permet de sélectionner 15 interacteurs dérégulés et en commun entre les HEK293 et BEAS2B. Après retrait de deux interacteurs dérégulés de façon opposée entre les deux lignées cellulaires, 13 interacteurs sont identifiés. La 'heatmap' à droite représente les variations d'interaction entre la forme sauvage et mutée de KRAS. Les valeurs indiquées sont le logarithme en base 2 du ratio du nombre total de peptide entre la forme mutée et sauvage. Les valeurs notées d'une étoile (*) correspondent à un interacteur exclusif de KRAS sauvage (rouge) ou muté (vert) pour la lignée cellulaire considérée.

Quelques brèves informations sur les interacteurs identifiés :

- ECE1 (Endothelin Converting Enzyme 1): Active par protéolyse l'endothelin 1. Son expression est reliée à un mauvais pronostic dans le cancer du poumon non à petites cellules (Boldrini et al., 2005)
- IL1RAP (Interleukin 1 Receptor Accessory Protein): S'associe au récepteur interleukin 1 pour relayer son activation. Exprimé dans des lignées de cancer du sein triple négatif et cancer du poumon non à petites cellules. Des anticorps à visée thérapeutique contre cette cible sont en

cours de développement par la société Cantargia (communiqué de presse daté du 26 Juin 2017, <http://cantargia.com/>, consulté en Juillet 2017).

- MTOR (Mechanistic Target of Rapamycin): Sérine/threonine kinase clef dans la voie PIK3/Akt. Cette voie de signalisation est déjà connue comme altérée dans le cancer du poumon non à petite cellule et fait l'objet d'essais thérapeutiques (Fumarola et al., 2014).
- MYO1E (Myosin IE): Interagit avec les filaments d'actine et serait impliqué dans le transport vésiculaire. Cette protéine active la prolifération tumorale dans le cancer du sein (Ouderkirk-Pecone et al., 2016).
- NF1 (Neurofibromin 1): GAP qui active l'hydrolyse du GTP des protéines RAS et régule négativement leur activité. Elle est retrouvée mutée dans le cancer du poumon non à petites cellules (Redig et al., 2016).
- RAB27A (RAB27A, Member RAS Oncogene Family): Protéine de la famille des petites GTPases qui intervient dans le trafic membranaire. RAB27A régulerait la production d'exosomes et favoriserait la progression tumorale par altération du micro-environnement cellulaire (Bobrie et al., 2012; Li et al., 2014).
- RAF1 (Raf-1 Proto-Oncogene, Serine/Threonine Kinase): Première serine/threonine kinase de la voie des MAPK activée par les protéines RAS. Protéine essentielle dans le développement tumorale des cancers du poumon non à petites cellules mutés KRAS (Blasco et al., 2011)
- RAPGEF2 (Rap Guanine Nucleotide Exchange Factor 2): GEF qui active les protéines RAS. Cette protéine a été impliquée dans l'invasion tumorale dans le cancer du sein et le développement de métastases (Magliozzi et al., 2013).
- RAPGEF6 (Rap Guanine Nucleotide Exchange Factor 6): GEF qui active les protéines RAS. Cette protéine est peu étudiée, et aucun lien avec l'oncogenèse n'a été établi.
- RPS6KA1 (ribosomal protein S6 kinase, 90kDa, polypeptide 1): Cible de ERK1 et ERK2 dans la voie des MAPK. Cette protéine contrôle notamment la signalisation de MTOR et est impliqué dans la prolifération, la survie et la différenciation cellulaire. Elle a été déjà identifiée comme cible thérapeutique potentielle dans le cancer du poumon non à petites cellules (Poomakkoth et al., 2016).
- SPRED2 (Sprouty Related EVH1 Domain Containing 2): Protéine qui régule l'activation de la voie des MAPK en recrutant NF1 à la membrane plasmique. Une diminution d'expression de cette protéine a été mise en évidence dans le cancer de la prostate (Kachroo et al., 2013).
- TRAF7 (TNF Receptor Associated Factor 7): E3 ubiquitine ligase qui régule l'activation de la voie des MAPK et NF- κ B. Identifiée très récemment comme suppresseur de tumeurs. Cette protéine a été impliqué dans l'initiation et la progression tumorale de différents cancers dont les méningiomes (Zotti et al., 2017) .

- USP32 (Ubiquitin Specific Peptidase 32): Protéine peu étudiée mais identifiée récemment comme ayant un rôle dans la progression tumorale du cancer du poumon à petite cellule, et a été identifiée comme cible thérapeutique potentielle dans ce même cancer (Hu et al., 2017).

Dans cette liste, il apparaît que les couples MTOR-RPS6KA1, NF1-SRPED2, RAPGEF2-RAPGEF6 sont liés d'un point de vue fonctionnel et/ou structurel dans des voies de signalisation ou des complexes fonctionnels.

On remarque également que pour la grande majorité de ces protéines, un lien avec le développement de cancers a été établi que ce soit par des mutations, des sous- ou sur-expressions, et parfois très récemment comme c'est le cas de MYO1E. Ceci conforte une nouvelle fois la méthodologie employée et la pertinence de l'étude des interactions protéiques pour identifier de nouvelles cibles thérapeutiques.

II.5 Effet des protéines cibles sur la viabilité cellulaire

Afin d'évaluer le rôle fonctionnel de certains interacteurs de KRAS, notamment sur phénotype oncogénique, je me suis tout d'abord focalisé sur la liste des 13 interacteurs communs aux lignées HEK293 et BEAS2B, et dérégulés par la mutation KRAS G12C. Pour ce faire, l'expression de ces 13 protéines a été éteinte par ARN interférants (siRNA) dans différentes lignées de cancer du poumon et la viabilité cellulaire a été évaluée. Les 4 lignées cellulaires sont des lignées de NSCLC, et expriment la protéine mutée KRAS G12C (lignées cellulaires H23 et H358) ou la protéine KRAS sauvage mais EGFR muté (lignées cellulaires H1650 et H1975) (Tableau 6).

Dans un premier temps, l'efficacité de l'extinction d'expression des gènes cibles a été validée par western blot. Sur les 13 cibles, seules 6 ont pu être analysées par cette approche (Figure 35), les autres n'ayant pu être validés à ce stade faute d'anticorps spécifiques disponibles. Une fois l'effet des siRNA validé, la viabilité cellulaire a été évaluée dans les 4 lignées de NSCLC (Figure 36), par rapport aux cellules contrôles transfectées par un siRNA non-spécifique.

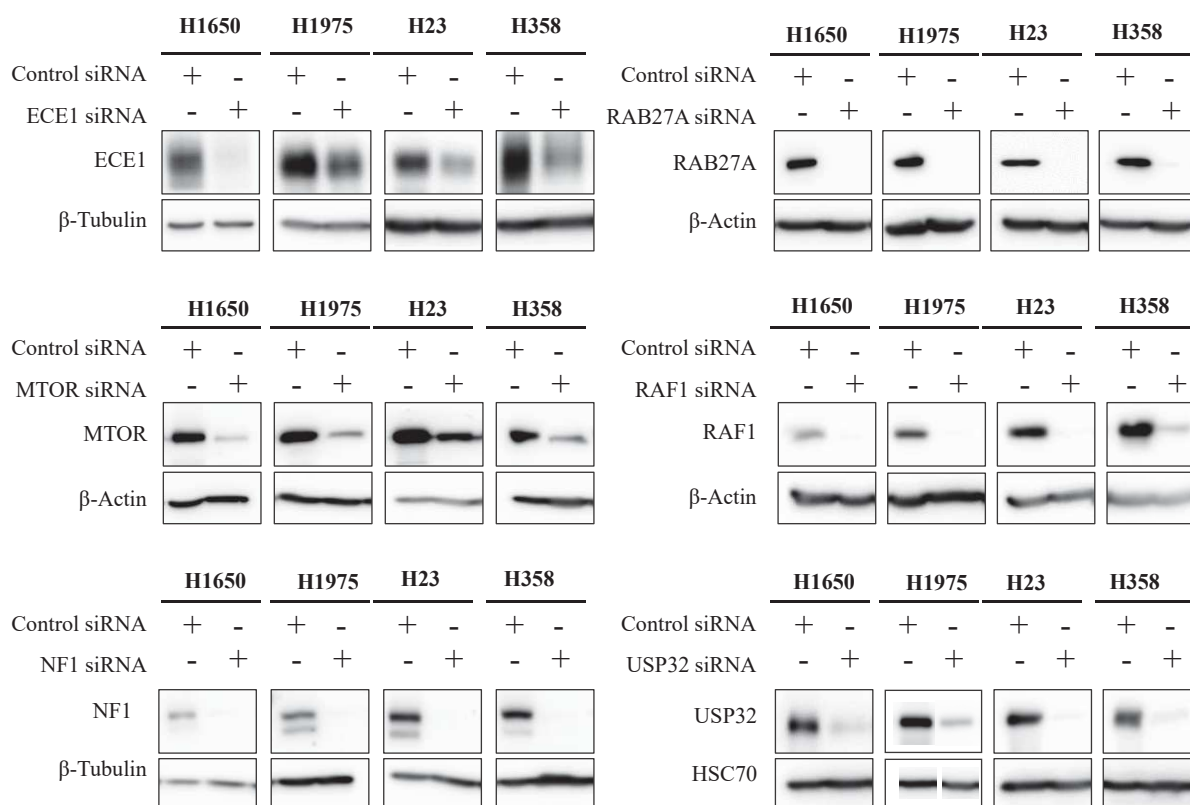


Figure 35. Validation de l'extinction des cibles sélectionnées. La validation de l'extinction des cibles ECE1, MTOR, NF1, RAB27A, RAF1 et USP32 par siRNA sur les lignées cellulaires H1650, H1975, H23 et H358 a été réalisée par western-blot. En témoin de charge, la β -Tubuline, la β -Actine et HSC70 ont été révélées selon les cas.

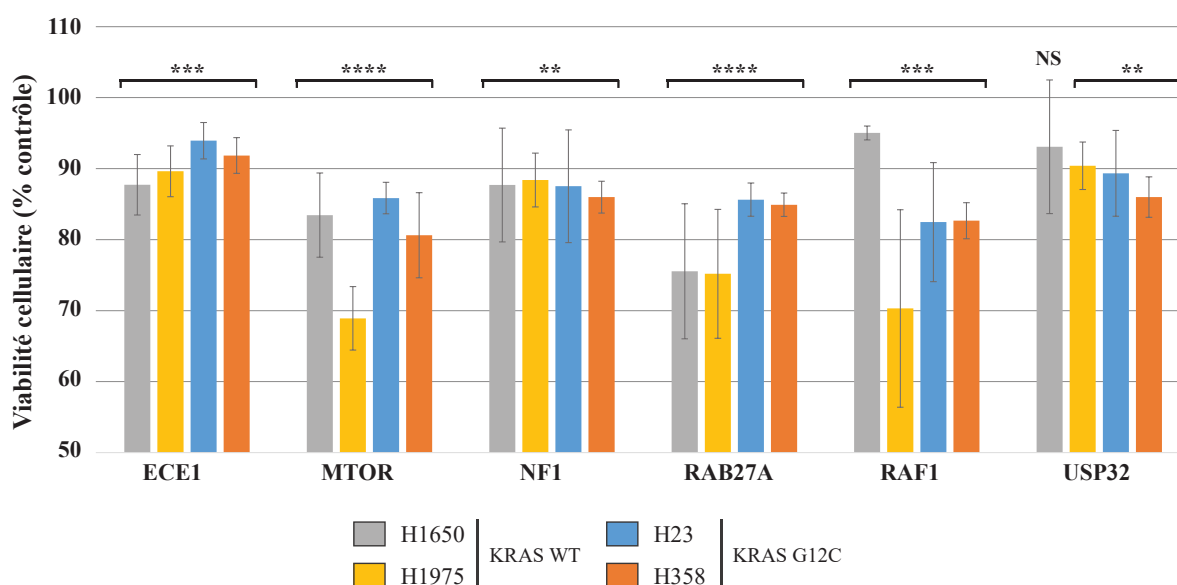


Figure 36. Effet sur la viabilité cellulaire de l'extinction d'expression des cibles sélectionnées. L'activité des protéines ECE1, MTOR, NF1, RAB27A, RAF1 et USP32 a été éteinte par siRNA. Pour chacune des lignées cellulaires, les valeurs sont mesurées en pourcentage de viabilité cellulaire par rapport au contrôle siRNA (non représenté). NS, P-value > 0,05; **, P-value \leq 0,01; ***, P-value \leq 0,001; ****, P-value \leq 0,0001.

Pris globalement, nous observons un effet significatif de l'extinction des cibles sur la viabilité cellulaire des 4 lignées sélectionnées (inhibition de viabilité de 5 à 30%). Cependant, nous n'observons pas de variations significatives dépendantes du statut mutationnel de KRAS.

C'est également le cas avec les 7 autres interacteurs testés (IL1RAP, MYO1E, RAPGEF2, RAPGEF6, RPS6KA1, SPRED2, TRAF7) et pour lesquels nous n'avons pu tester les niveaux d'expression par western blot, faute d'anticorps disponibles (Figure 37). On retrouve un effet significatif sur la réduction de viabilité cellulaire lors de l'extinction d'expression des cibles et ceci quelles que soient les lignées cellulaires étudiées (inhibition de viabilité de 5 à 35%).

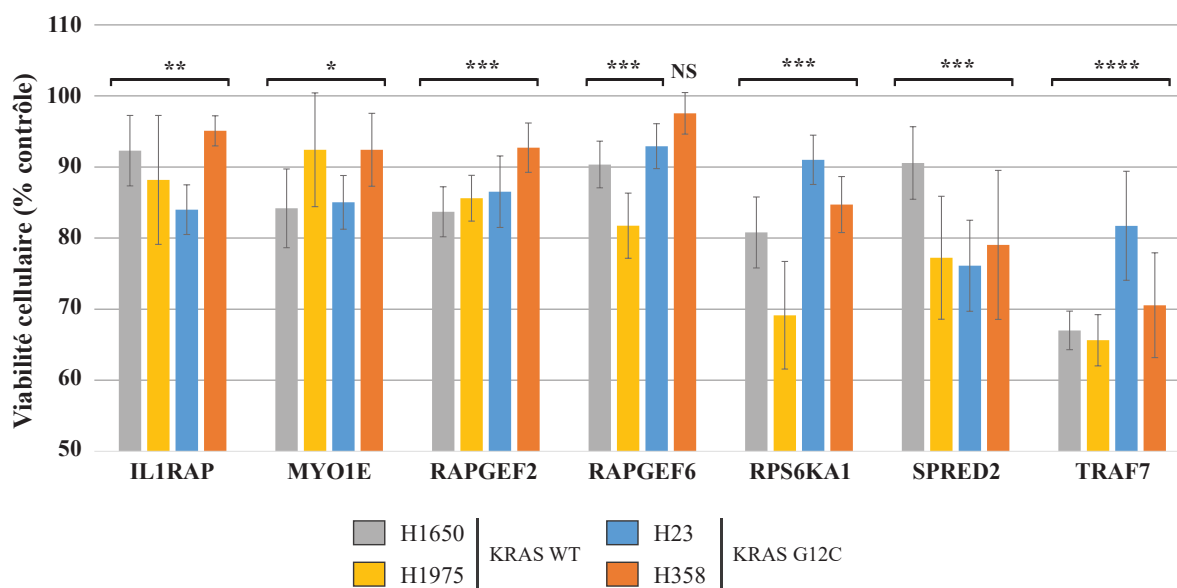


Figure 37. Effet sur la viabilité cellulaire de l'extinction d'expression des 7 dernières cibles sélectionnées. L'activité des protéines IL1RAP, MYO1E, RAPGEF2, RAPGEF6, RPS6KA1, SPRED2 et TRAF7 a été éteinte par siRNA. Pour chacune des lignées cellulaires, les valeurs sont mesurées en pourcentage de viabilité cellulaire par rapport au contrôle siRNA (non représenté). NS, P-value > 0,05; *, P-value ≤ 0,05; **, P-value ≤ 0,01; ***, P-value ≤ 0,001; ****, P-value ≤ 0,0001.

Bien qu'aucun anticorps spécifique de ces 7 cibles n'ait pu être identifié pour véritablement valider l'extinction des protéines ciblées, nous observons qu'*a priori* l'inhibition de TRAF7 conduit à la plus forte réduction de viabilité cellulaire, environ 35% pour les cellules H1650 et H1975. Pour ces 7 cibles, les amorces de PCR quantitative ont été mises au point, et la validation de l'extinction d'expression est en cours.

II.6 Dépendance à GLUT3 de lignées mutées KRAS

En parallèle à l'étude fonctionnelle systématique précédente, je me suis également intéressé au transporteur de glucose 3 (GLUT3 ou SLC2A3) qui interagit spécifiquement avec la forme mutée de KRAS G12C. En effet, une étude récente de criblage à large échelle par l'approche CRISPR a montré

que la délétion du gène *GLUT3* était létale dans la lignée de colon HCT116 portant une mutation KRAS G13D, alors que cette même délétion est sans effet dans un contexte KRAS sauvage (lignée HeLa et A375) (Hart et al., 2015). Bien que la mutation de KRAS de la lignée HCT116 soit différente de la mutation KRAS G12C, ce travail tend à confirmer le lien fonctionnel de KRAS et la protéine GLUT3 que nous avons pu observer par BioID. Ce groupe de recherche m'a transmis une lignée cellulaire d'adénocarcinome pancréatique (HPAFII) présentant une mutation de KRAS proche de la mutation G12C, à savoir la mutation KRAS G12D, afin de valider cette hypothèse.

J'ai donc évalué l'effet de la délétion de *GLUT3* sur la viabilité cellulaire des cellules HPAFII (Figure 38). J'ai utilisé pour cela technologie CRISPR qui permet de cibler spécifiquement la délétion de *GLUT3*. La délétion est réalisée par l'endonucléase Cas9 dont l'activité catalytique est guidée par un gRNA (ARN guide). Les cellules HPAFII utilisées expriment stablement la protéine Cas9, et la séquence du gRNA qui ciblant *GLUT3* a été défini avec l'outil informatique <http://crispr.mit.edu:8079/>. Le contrôle positif de délétion utilise un gRNA qui cible le gène essentiel *PSMD1*. Le contrôle négatif de délétion utilise un gRNA qui cible le gène procaryote *LacZ*.

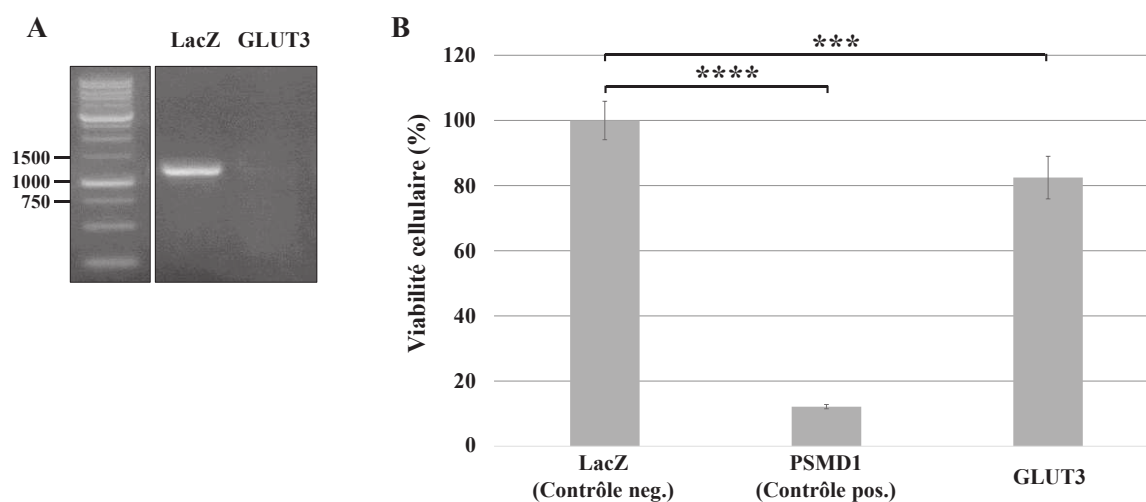


Figure 38. Effet de la délétion de *GLUT3* sur la viabilité des cellules HPAFII. (A) L'amplification de *GLUT3* dans la condition LacZ et l'absence d'amplification de *GLUT3* dans la condition où ce gène a été ciblé valide la délétion de *GLUT3*. (B) La délétion de *GLUT3* réduit d'environ 20% la viabilité des cellules HPAFII mutées KRAS G12D. La délétion du gène *PSMD1* constitue le contrôle positif de délétion, et la condition LacZ, ciblant la délétion du gène procaryote *LacZ*, constitue le contrôle négatif de délétion. Les valeurs présentées sont normalisées par rapport à la condition LacZ. Contrôle neg., contrôle négatif de délétion; Contrôle pos., contrôle positif de délétion; ***, P-value $\leq 0,001$; *****, P-value $\leq 0,0001$.

Dans la lignée HPAFII, la délétion de *GLUT3* réduit d'environ 20% la viabilité cellulaire et la délétion du gène essentiel *PSMD1* conduit quant à elle à une réduction de 90% de la viabilité. Basée sur cette expérience, l'inhibition de GLUT3 réduit significativement la viabilité des cellules HPAFII, mais cette cible n'apparaît pas aussi essentielle que *PSMD1*. Ce lien entre le statut mutationnel de KRAS et

la dépendance de ces cellules à GLUT3 n'est donc pas évident, et des expériences complémentaires doivent être menées pour valider cette hypothèse.

III. Discussion

Le BioID a permis de mettre en évidence des réseaux d'interactions protéiques extrêmement riches autour des protéines KRAS WT et KRAS G12C. Les réseaux d'interactions dépendent des lignées cellulaires utilisées. Pour KRAS G12C ou KRAS WT, plus de 60% des interacteurs identifiés dans les BEAS2B sont retrouvés dans les HEK293. Cependant, le nombre d'interacteur identifié dans les HEK est bien supérieur (814 interacteurs contre 452 dans les BEAS2B), ce qui pose la question de la complémentarité des interacteurs identifiés dans différentes lignées cellulaires.

La lecture de ce réseau peut se faire alors à plusieurs niveaux. À un niveau global, en considérant les interactions dans leur ensemble, on dispose d'un véritable catalogue des interacteurs directs ou indirects à proximité de KRAS WT ou G12C. Avec 902 et 713 protéines identifiées interagissant avec KRAS G12C et KRAS WT, respectivement, il s'agit d'une des listes la plus complète à ce jour et à ma connaissance. À un niveau plus fin, et en comparant les réseaux d'interactions entre les lignées, on peut isoler les interacteurs majeurs pour les fonctions de KRAS et communs à plusieurs lignées, ou les interacteurs les plus pertinents pour une pathologie donnée. C'est le cas de la liste minimale et fiable de 13 interacteurs que j'ai pu mettre en évidence, et pour lesquels j'ai pu montrer que leur extinction dans des lignées de poumon conduit à une réduction significative de la viabilité cellulaire. Cependant, il apparait difficile de relier cet effet sur la viabilité au statut mutationnel de KRAS.

En effet, les modèles cellulaires de cancer du poumon disponibles sont généralement mutés sur le gène *KRAS* ou *EGFR* (mutations mutuellement exclusives). Les lignées exprimant KRAS WT que j'ai utilisées pour les expériences fonctionnelles sont mutées sur le gène de l'*EGFR*, ce qui peut fortement introduire un biais dans l'analyse finale compte tenu de l'intrication des voies de signalisation dirigées par l'EGFR et KRAS. Par ailleurs, la base de données COSMIC référence les mutations de lignées cellulaires couramment utilisées et précise par exemple que la lignée H23 porte 2014 points de mutations et la lignée H1650 présente 277 points de mutations (Forbes et al., 2017). Ainsi, bien qu'EGFR ou KRAS soient considérées comme les protéines mutées majeures, c'est-à-dire celles qui « mènent » le processus tumoral, leurs activités s'établissent dans des environnements protéiques différents, avec des réseaux d'interactions protéiques et régulations différents. En conséquence, de la même façon que KRAS, lorsque l'EGFR est muté celui-ci présente un réseau d'interaction dérégulé pour ses interacteurs directs, mais également pour les interacteurs de ses interacteurs et ainsi de suite. Finalement, avec les 4 lignées de NSCLC utilisés, ce ne sont pas simplement les statuts mutationnels de KRAS et EGFR qui sont comparés mais des contextes cellulaires plus complexes, ne pouvant être reliés strictement aux mutations respectives de ces gènes.

Il faut également noter que je me suis intéressé ici à un phénotype particulier (viabilité cellulaire) et que certaines protéines peuvent être impliquées dans d'autres processus cellulaires tumoraux comme l'invasion ou la migration. Par exemple, bien que l'extinction de USP32 induit une réduction de viabilité d'environ 10% pour les 4 lignées cellulaires, cette protéine semble plutôt impliquée dans la migration du cancer du sein et du poumon (Akhavantabasi et al., 2010; Hu et al., 2017; Pal and Donato, 2014). Il serait alors intéressant de réaliser des expériences de migration et d'invasion car un effet plus significatif que celui affectant la viabilité cellulaire pourrait être observé, et pourrait éventuellement être corrélé au statut mutationnel de KRAS. Étudier l'influence de ces cibles sur le phénotype oncogénique nécessite alors une étude plus complète incluant l'effet de ces protéines sur la prolifération, migration, invasion cellulaire. Étudier la viabilité cellulaire reste cependant une expérience de choix car elle évalue en une seule mesure l'atteinte de la prolifération, l'apoptose ou le contrôle du cycle cellulaire, qui sont des mécanismes fréquemment dérégulés en cancérologie.

Concernant l'étude fonctionnelle de *GLUT3*, la délétion de ce gène dans un contexte KRAS G12D n'a confirmé que partiellement sa dépendance au statut mutationnel de KRAS. Effectivement, la délétion de *GLUT3* induit une réduction de viabilité de 20%, alors que la délétion du gène essentiel *PSMD1* induit une réduction de viabilité de 90%. Cette différence par rapport à l'étude d'origine (Hart et al., 2015) sur la lignée HCT116 KRAS G13D peut s'expliquer de différentes façons. Premièrement d'un point de vue expérimental, car la létalité de la délétion de *GLUT3* dans le contexte KRAS G13D de la lignée HCT116 a été observée après plus de 20 divisions cellulaires, comparé à environ 7 divisions dans l'expérience que j'ai faite. Deuxièmement, les transporteurs de glucose sont une large famille avec plus de 10 membres et tous ne sont pas exprimés de la même façon selon le type cellulaire. Ces différences d'expression créent donc des dépendances à certain transporteur de glucose plutôt que d'autres selon la lignée utilisée.

Malgré ces résultats, *GLUT3* reste une cible intéressante car sa surexpression dans les cellules de poumons H23, H358, H460 et SW1573 (toutes mutées KRAS G12C) a été corrélée à l'augmentation de la consommation de glucose et au nombre de copie de *KRAS* (Kerr et al., 2016). De plus, d'autres articles mettent également en évidence que la surexpression de *GLUT3* est corrélée à la prolifération tumorale des lignées d'adénocarcinome SW1573 et H727, et à l'activité du facteur de transcription HIF1 α , protéine à l'interface de la signalisation cellulaire et du métabolisme et contrôlant notamment la glycolyse (Masin et al., 2014; Pavlova and Thompson, 2016; Pylayeva-Gupta et al., 2011; Vander Heiden, 2011).

D. Autres réseaux d'interactions protéiques

I. Réseau d'interaction de NRAS

I.1 Introduction de l'étude

Une étude supplémentaire de l'interactome des protéines RAS et de ses dérégulations à la suite de mutations oncogénique a été menée sur NRAS, le second isoforme le plus fréquemment muté. Les mutations ont lieu principalement sur le codon 61 et conduisent majoritairement aux mutations NRAS Q61K ou Q61R. Ces mutations sont identifiées dans 20 à 30% des mélanomes cutanés et dans environ 1% des cancers du poumon non à petites cellules. Comme pour KRAS, aucune thérapie ciblée n'a été identifiée. En parallèle de l'étude de KRAS, nous avons donc étudié les réseaux d'interactions protéiques de NRAS WT et NRAS Q61K. Les données de MS sont disponibles en Annexe 5.

I.2 Résultats

I.2.a Expression de NRAS dans la lignée HEK293

Les réseaux d'interactions de NRAS ont été caractérisés exclusivement dans le modèle cellulaire HEK293. L'induction des protéines a été validée par western-blot (Figure 39).

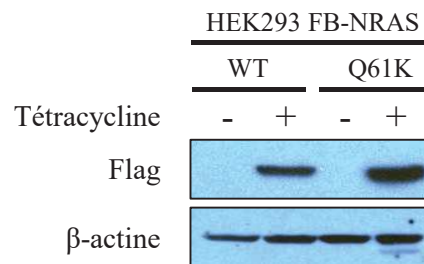


Figure 39. Validation de l'expression du transgène NRAS WT et Q61K dans les cellules HEK293. FB, FlagBirA*.

Comme dans le cas des cellules exprimant KRAS, l'induction de NRAS est fortement contrôlée par la tétracycline, et nous observons plus de protéine mutée que sauvage. Cette augmentation de quantité de protéines suggère un mécanisme commun entre les protéines RAS qu'il serait intéressant d'approfondir.

I.2.b Dynamique du réseau d'interaction de NRAS

L'analyse par BioID des interactions protéiques de NRAS WT et NRAS Q61K a permis d'identifier au total 731 protéines. La figure 40 suivante présente les spécificités d'interactions entre la forme mutante et sauvage de NRAS.

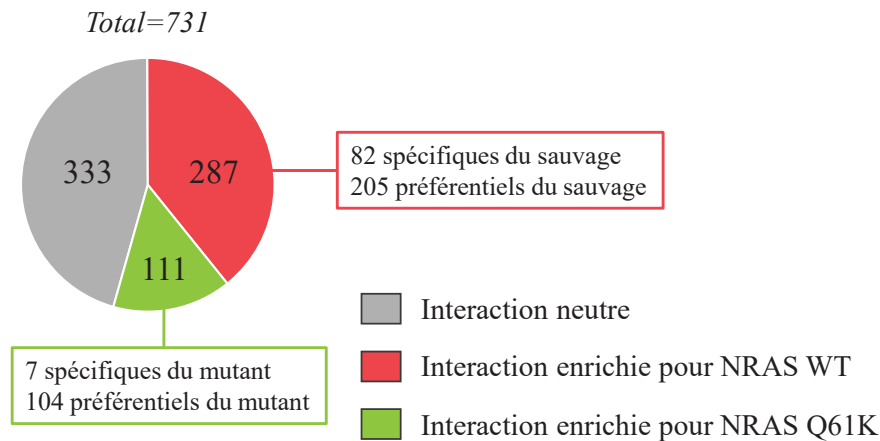


Figure 40. Comparaison des interacteurs de NRAS WT et NRAS Q61K identifiés dans les cellules HEK293. Basé sur le nombre total de peptides identifiés, nous observons 111 interacteurs proches de la forme mutée, parmi lesquels 7 protéines lui sont spécifiques. Plus proche de la forme sauvage, 287 interacteurs sont identifiés et parmi lesquels 82 protéines ne sont pas identifiées à proximité de NRAS Q61K.

Les interacteurs dérégulés de NRAS correspondent dans 70% des cas à une diminution ou une perte d'interaction pour la protéine mutée. Une analyse des termes GO enrichis montre que les ontologies spécifiques aux membranes cellulaires sont retrouvées telles que « plasma membrane region », « cytoplasmic, membrane-bounded vesicle » et « endosome ». Concernant les processus biologiques, les termes « intracellular signal transduction », « enzyme linked receptor protein signaling pathway » et « cell migration » sont une nouvelle fois enrichis. On peut noter par ailleurs que les termes de régulation négative de la croissance (« GO:0030308, negative regulation of cell growth ») et régulation négative de la transduction intracellulaire du signal (« GO:1902532, negative regulation of intracellular signal transduction ») sont enrichis et comprennent plus de 65% d'interacteurs perdus ou interagissant plus faiblement avec le mutant. De manière intéressante, il est important de noter que l'analyse GO de l'ensemble des interacteurs dérégulés montre que les localisations et processus biologiques enrichis correspondent à ceux identifiés comme dérégulés par la mutation KRAS G12C dans le modèle cellulaire HEK293 (Figure 33).

I.2.c Interacteurs dérégulés entre NRAS et KRAS mutés

Ayant identifié les interacteurs dérégulés par la mutation KRAS G12C, nous pouvons nous interroger sur les interacteurs dérégulés et en commun entre NRAS et KRAS à la suite de leur mutations Q61K et G12C respectives (Figure 41).

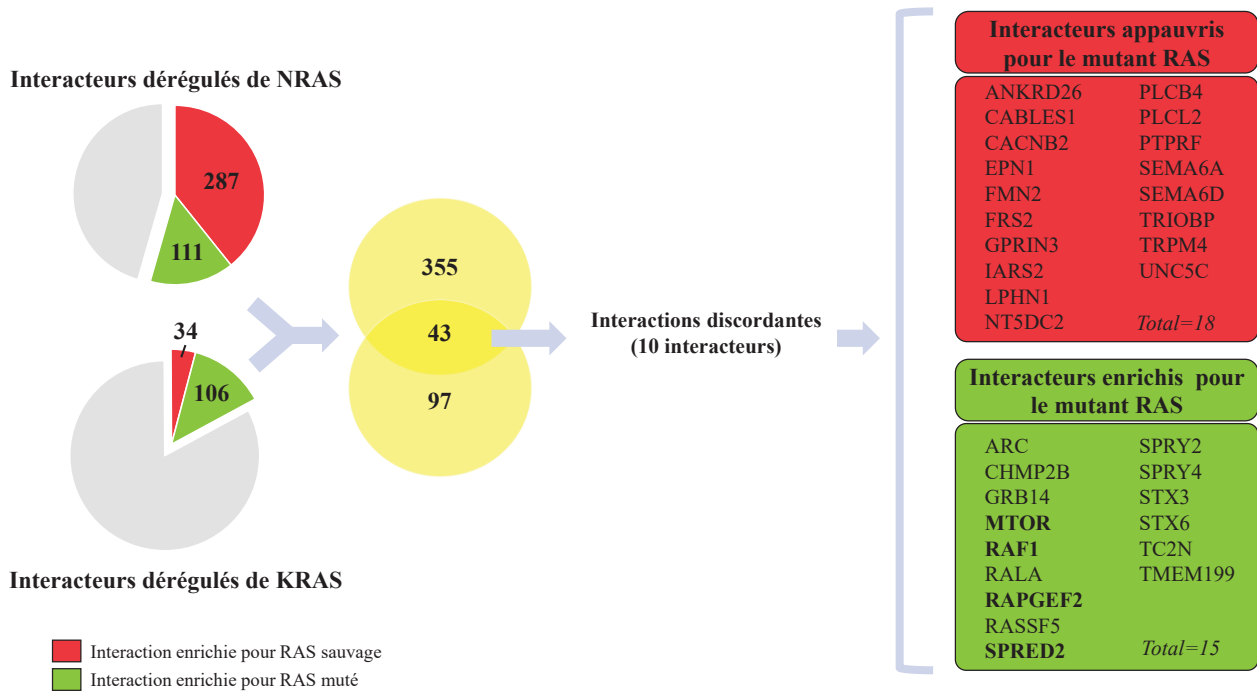


Figure 41. Identification des interacteurs communs dérégulés par les mutations NRAS Q61K et KRAS G12C. La comparaison des interacteurs identifiés comme dérégulés dans les cellules HEK293 par la mutation NRAS Q61K et KRAS G12C permet de mettre en évidence 43 interacteurs communs. Parmi ceux-ci, 18 interagissent plus faiblement avec le mutant RAS (cadre rouge) et 15 interagissent plus fortement avec les mutants RAS (cadre vert). Les interacteurs en gras sont les interacteurs identifiés également comme dérégulés dans la lignée BEAS2B (étude KRAS).

Cette comparaison met en évidence que parmi les 495 interacteurs uniques dérégulés, 43 d'entre eux sont communs entre les dérégulations de NRAS et KRAS. Parmi ceux-ci, 18 interacteurs ont une affinité plus forte pour les formes sauvages de NRAS et KRAS, et 15 interacteurs ont une affinité plus forte pour les formes mutées. De manière intéressante, nous remarquons que parmi les 13 interacteurs sélectionnés lors de l'étude du réseau de KRAS, nous retrouvons une nouvelle fois MTOR, RAF1, RAPGEF2 et SPRED2. Ceci conforte leur pertinence dans l'étude des mutations de KRAS et NRAS.

Une analyse GO de l'ensemble de ces 33 interacteurs montre que seul le terme « intracellular signal transduction » est enrichi, regroupant 19 interacteurs identifiés ($P\text{-value}=2,8.10^{-3}$). Parmi ces interacteurs, et notamment ceux plus proches des formes mutées de NRAS et KRAS on trouve RALA, RAF1, RAPGEF2, RASSF5, MTOR, SPRY2, SPRY4 et SPRED2. Quand on s'intéresse aux domaines

protéiques enrichis, un seul domaine protéique est enrichi. Il s'agit du domaine sprouty identifié sur les protéines SPRY2, SPRY4 et SPRED2. Il s'agit d'un domaine impliqué dans la translocation du cytosol à la membrane et dans la régulation négative de la voie MAPK. Ces protéines interagissant plus fortement avec les formes mutées de KRAS et NRAS, ceci suggère qu'il s'agit de protéines de régulations recrutées par les formes mutées de KRAS et NRAS pour réduire la suractivation de la voie MAPK par les mutants RAS, pouvant constituer ainsi la boucle de régulation négative majeur.

Enfin, afin d'identifier les liens fonctionnels éventuels parmi ces 33 interacteurs dérégulés, une analyse STRING a été conduite (Figure 42).

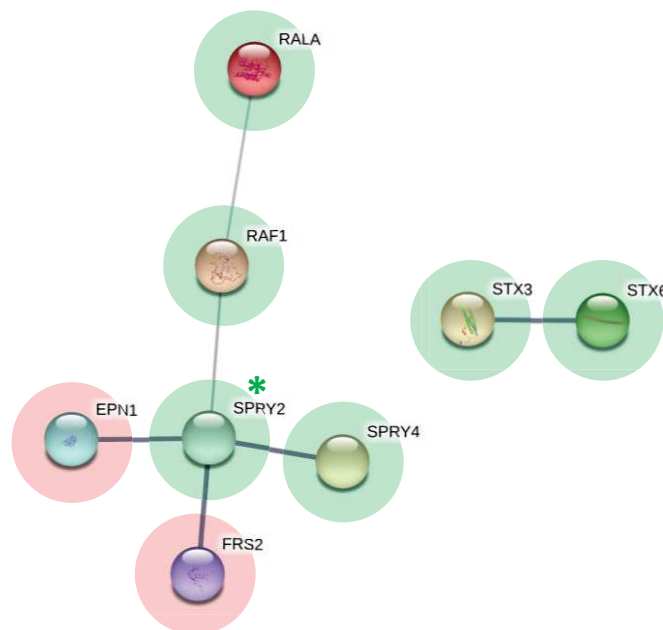


Figure 42. Réseau d'interactions fonctionnelles des protéines dérégulées par les mutants NRAS Q61K et KRAS G12C. Deux réseaux indépendants sont mis en évidence. Les interacteurs entourés de verts interagissent plus fortement avec les mutants et ceux entourés de rouge plus faiblement. L'étoile (*) sur SPRY2 indique qu'il s'agit d'un interacteur spécifique de KRAS G12C et NRAS Q61K. L'épaisseur du lien est proportionnelle à la confiance d'interaction identifiée. Analyse réalisée avec STRING.

Cette analyse met en évidence deux réseaux fonctionnels distincts à proximité des mutants NRAS et KRAS. Le premier de ces réseaux inclut 6 protéines parmi lesquels 4 interagissent préférentiellement (RALA, RAF1, SPRY4) ou spécifiquement (SPRY2) avec les mutants KRAS G12C et NRAS Q61K. Parmi ces 6 protéines, RALA et RAF1 apparaissent les plus intéressantes car elles sont étroitement impliquées dans la signalisation cellulaire. RALA est une petite GTPase en lien avec la signalisation cellulaire *via* des récepteurs tels que RET. RAF1 (ou CRAF) est la première kinase activée de la voie des MAPK. L'identification de RAF1 et non de BRAF ou de ARAF montre la prévalence de cet isoforme lors de la dérégulation des protéines KRAS et NRAS. FRS2 est une protéine adaptatrice des

récepteurs aux domaines tyrosine kinase FGR et NGF, et interagit plus faiblement avec les mutants KRAS et NRAS. Comme nous l'avons vu, les protéines SPRY2 et SPRY4 impliquées dans la régulation négative de la voie des MAPK et apparaissent ainsi essentielles à la régulation des RAS. Ceci d'autant plus que SPRY2 est un interacteur exclusif des formes mutées de NRAS et KRAS. La protéine EPN1 interagit quant à elle plus faiblement avec les mutants RAS. Il s'agit d'une protéine impliquée dans le recyclage membranaire. Le second réseau mis en évidence pour cette analyse concerne également le trafic vésiculaire avec 2 protéines de la famille des syntaxines (STX3 et STX6).

Finalement, ces 33 protéines dérégulées ou plus particulièrement ces deux réseaux constituent un noyau d'interaction important dans les fonctions cellulaires des formes mutées de NRAS et KRAS. Ces réseaux suggèrent également une nouvelle l'interconnexion des interacteurs.

I.3 Discussion

Comme observé lors de l'étude des réseaux d'interactions de KRAS, le BioID permet de mettre en évidence des réseaux d'interactions riches autour de NRAS WT et NRAS Q61K, avec au total plus de 700 interacteurs. Cependant, à l'inverse des dérégulations induites par la mutation KRAS G12C, la mutation NRAS Q61K conduit principalement à des appauvrissements ou des pertes d'interactions par rapport à la protéine sauvage (Figure 40).

La comparaison des réseaux dérégulés par les mutations de NRAS et KRAS a permis de mettre en évidence 33 interacteurs fiables, dérégulés par la mutation NRAS Q61K et KRAS G12C. Ces partenaires sont intéressants dans la mesure où les mutations NRAS et KRAS sont principalement impliquées dans deux pathologies distincts (mélanomes cutanés pour NRAS et adénocarcinomes pulmonaires pour KRAS). Quinze d'entre eux interagissent plus fortement avec les formes mutées de NRAS et KRAS et parmi eux les protéines MTOR, RAF1, RAPGEF2 et SPRED2 sont exprimés dans la lignée de poumon BEAS2B exprimant KRAS G12C. Cette analyse souligne une nouvelle fois l'importance de la voie de signalisation MAPK et de la protéine MTOR dans la dérégulation des réseaux d'interactions des protéines RAS. Des combinaisons de drogues ciblant les protéines MTOR et MEK (en aval de la voie MAPK) ont d'ailleurs été identifiées comme option thérapeutique synergique dans les mélanomes et les cancers du poumon non à petites cellules mutés RAS (Lasithiotakis et al., 2008; Legrier et al., 2007; Posch et al., 2013).

Parmi les protéines interagissant plus spécifiquement avec les formes mutées de RAS, on retrouve des protéines régulatrices en amont de RAS, et jouant un rôle dans son activation, comme GRB14 ou RAPGEF2, mais également des effecteurs en aval de RAS, et participant à différentes voies de signalisation, comme RASSF5, RALA, RAF1 et MTOR. Ces multiples partenaires de RAS mutées suggèrent une plasticité dans la surexpression des différentes voies de signalisation médiées par les mutations des isoformes de RAS. Toutefois, l'analyse de ces réseaux d'interactions doit être faite avec

précaution d'une part car des boucles de régulation négatives fines peuvent être déclenchées en fonction du contexte cellulaire, et d'autre part car ces réseaux ont été constitués dans des modèles cellulaires non-cancéreux (HEK293 et BEAS2B). En effet, parmi les protéines dérégulées et interagissant plus spécifiquement avec les protéines RAS mutées, on retrouve à la fois des suppresseurs de tumeurs comme SPRED2 (Yoshida et al., 2006) ou RASSF5 (Park et al., 2010), et des protéines pro-oncogéniques comme RALA (Lim et al., 2005), RAF1 (Blasco et al., 2011) ou MTOR (Fumarola et al., 2014). De plus certaines protéines apparaissent avoir un rôle oncogénique dans différents types de cancers, c'est de la cas de RALA dont l'implication dans l'invasion et la migration tumorale a été prouvée dans les cancers du poumon, les cancers colorectaux, les cancers pancréatiques et les mélanomes, tous présentant une mutation RAS (Guin et al., 2013; Györfy et al., 2015; Padavano et al., 2015; Zipfel et al., 2010). Il apparaît ainsi difficile de déterminer les mécanismes de régulation mis en jeu sur la seule interprétation des données *in silico*. Par exemple, la protéine SPRED2 est considérée comme un suppresseur de tumeurs et est impliquée dans la régulation négative de la voie des MAPK (Nonami et al., 2004; Wakioka et al., 2001). SPRED2 n'empêche pas l'interaction entre RAS et RAF mais inhibe l'activation de la voie MAPK en supprimant la phosphorylation et l'activation de RAF (Wakioka et al., 2001). Il apparaît donc étonnant de voir que l'interaction de SPRED2 est augmentée avec les protéines RAS mutées. De plus, lors de l'étude de KRAS G12C, nous avons pu observer que l'inhibition de SPRED2 réduisait la viabilité cellulaire de 10 à 25% dans les 4 lignées cellulaires de cancer du poumon, alors que, finalement, un effet inverse était attendu (Figure 37). Ce simple exemple suggère la présence de réseaux de régulation plus complexes, faisant intervenir d'autres partenaires avec d'autres boucles de régulation négatives qui finalement iraient en faveur d'une pro-activation des voies de signalisation de RAS.

Dans le cadre de NRAS, nous avons observé majoritairement des pertes d'interactions, avec la forme mutée de la protéine. Bien que moins intéressant d'un point de vue thérapeutique, ces pertes d'interactions restent néanmoins informatives sur les dérégulations biologiques observées, ou les mécanismes biologiques sous-jacents aux mutations.

EPN1 interagit plus faiblement avec les mutants RAS, et il a été démontré que cette protéine favorise notamment l'endocytose de l'EGFR, lorsque celui-ci est ubiquitinylé (Kazazic et al., 2009). Cette perte d'interaction pourrait alors se traduire par une réduction du recyclage membranaire, conduisant à une plus grande stabilité des récepteurs à la surface de la membrane plasmique. De plus, la dérégulation de cette protéine a été impliquée dans l'altération de l'endocytose et *in fine* l'augmentation de la migration et l'invasion cellulaire des fibrosarcomes (Coon et al., 2010). Un autre réseau soulignant la dérégulation du trafic vésiculaire est celui composé de STX3 et STX6, où à l'inverse d'EPN1 ces deux protéines interagissent plus fortement avec les formes mutées de RAS. L'expression de *STX6* est régulée par TP53, et intervient dans la régulation de l'adhésion cellulaire (Zhang et al., 2008). De plus, une analyse « -omic » a mis en évidence que parmi l'ensemble des protéines SNARE, *STX6* est la plus fréquemment surexprimée dans les cancers. Elle est notamment mutée dans les cancers du côlon, du pancréas, de la

peau, du poumon et sa dérégulation favoriserait la migration cellulaire et la survie de ces cellules (Riggs et al., 2012). Elle possède par ailleurs une activité pro-angiogénique dans les cellules endothéliales HUVEC, et a ainsi été identifiée comme cible thérapeutique (Manickam et al., 2011).

Enfin, cette analyse comparée des dérégulations des interactions protéiques par les mutations NRAS Q61K et KRAS G12C supporte l'étude des voies cellulaires PIK3/Akt, MAPK, RALA et l'endocytose avec en cibles principales respectives MTOR, RAF1, RALA et STX6.

II. Réseau d'interaction de l'EGFR : application au NSCLC

II.1 Introduction à l'étude

Les mutations EGFR comptent pour environ 10% des mutations des adénocarcinomes pulmonaires. Il s'agit principalement des mutations EGFR L858R et de la délétion partielle de l'exon 19 (Del. 19). Ces deux mutations sont sensibles aux ITKs mais malgré la survie sans progression de la maladie d'environ 1 an apportée par ces thérapies ciblées, des mutations de résistance apparaissent avec notamment la mutation EGFR T790M. Toutefois, la sélection spécifique de cette mutation et son effet en termes de partenaires recrutés et voies de signalisation activées est encore peu compris. C'est pourquoi nous nous sommes intéressés à caractériser les réseaux d'interactions de l'EGFR en présence ou non de cette mutation secondaire de résistance. Cette approche permettra d'une part d'identifier les voies suractivées ou altérées par la première mutation, et d'autre part de voir l'impact de la mutation secondaire sur ces mêmes voies ou d'autres voies. L'identification de ces réseaux d'interactions constitue une première étape dans le développement rationnel de drogues ciblant les effecteurs sélectionnés par ces mutations. Dans de rares cas, la mutation EGFR T790M est observée en l'absence de toute autre mutation, et constitue une mutation de mauvais pronostic. C'est pourquoi, le réseau d'interaction de cette mutation unique a aussi été inclus dans cette analyse.

Les réseaux d'interactions de l'ensemble des mutants de l'EGFR (L858R, Del. 19, T790M, L858R-T790M et Del. 19-T790M) ainsi que de la protéine sauvage ont été caractérisés dans les cellules HEK293. Je ne présenterai ici que les interacteurs les plus dérégulés par ces mutations, et m'intéresserai particulièrement aux effets de la mutation secondaire EGFR T790M. Les données de MS sont disponibles en Annexe 6.

II.2 Résultats

II.2.a Protéine sauvage

La caractérisation du réseau d'interaction de l'EGFR sauvage par le BioID a mis en évidence 322 interacteurs. L'analyse des termes GO enrichis associés à ces interacteurs montre que l'ensemble de ces partenaires sont localisés principalement au niveau de la membrane plasmique, et qu'ils sont majoritairement impliqués dans des processus fonctionnels tels que la transduction du signal via des protéines liées à des récepteurs, et notamment les petites GTPases. Les interacteurs identifiés sont en accord avec les fonctions d'EGFR connues, et abondent dans la pertinence de ces résultats.

II.2.b Dynamique du réseau d'interaction de l'EGFR

Bien que l'on puisse supposer que les différentes mutations EGFR (L858R, T790M ou Del.19) montrent des spécificités en termes de partenaires recrutés et de voies de signalisation activées, je ne m'intéresserai qu'aux interacteurs dérégulés communs à l'ensemble des EGFR mutés. En effet, les patients présentant ces mutations sont sensibles aux mêmes ITKs, et finiront par développer les mêmes résistances à ces traitements. De plus, identifier des cibles incluant ces deux types de patient est plus pertinent que des cibles plus spécifiques. Le même raisonnement sera appliqué aux doubles mutations EGFR L858R-T790M et Del.19-T790M.

Comme précédemment, l'identification des dérégulations d'interactions a été réalisée en faisant le ratio des peptides identifiés entre la condition EGFR muté et EGFR sauvage. Si le ratio est supérieur à 2 ou inférieur à 0.5, l'interacteur est alors considéré comme dérégulé.

La figure 43 présente les interacteurs identifiés comme dérégulés pour chacun des mutants EGFR, et les processus biologiques dans lesquels ceux-ci sont impliqués.

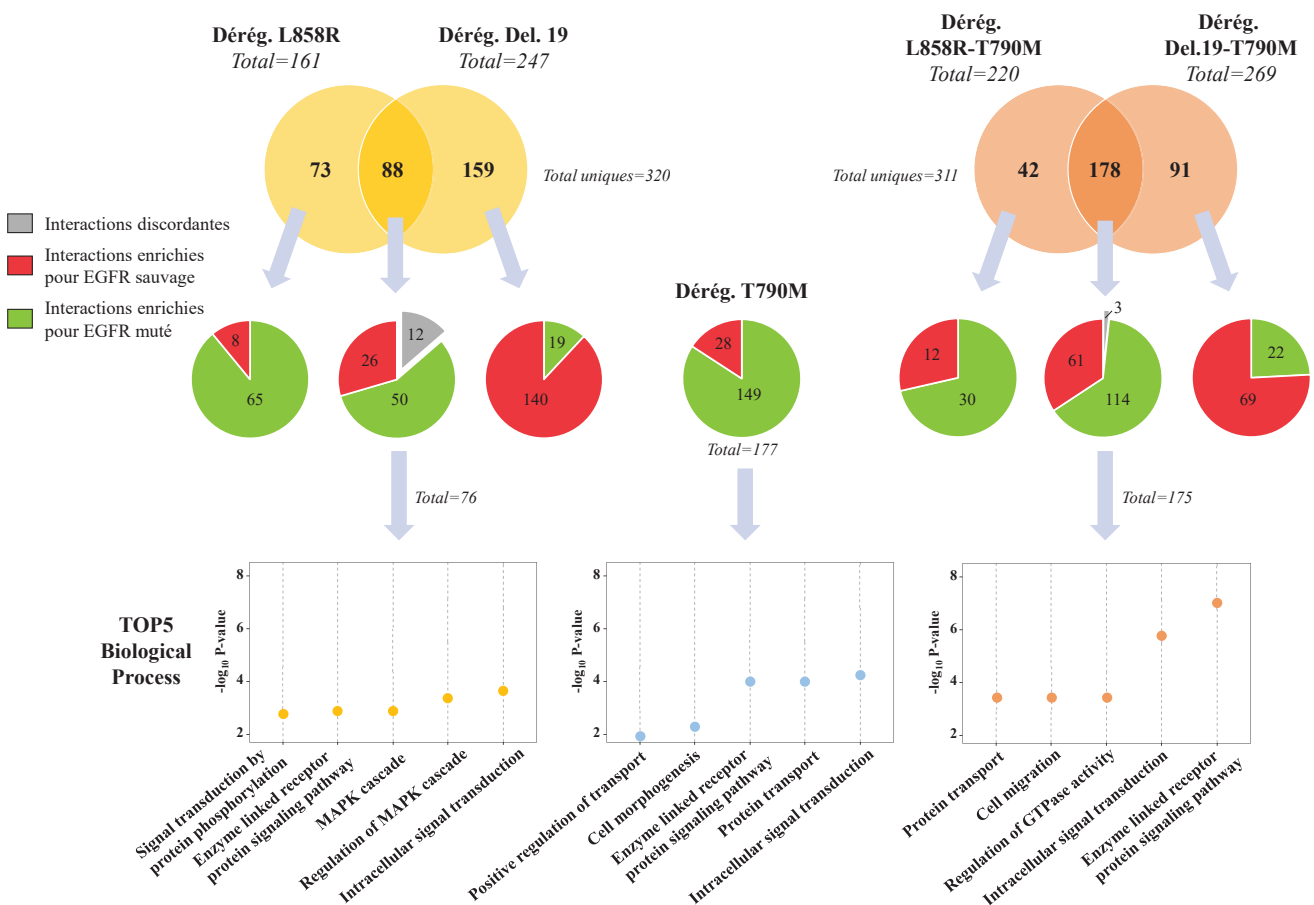


Figure 43. Identification des interacteurs dérégulés de l'EGFR par les mutations primo-activatrices (L858R ou Del.19) ou de résistance (T790M). Les diagrammes de Venn comparent les interactions dérégulées des mutants EGFR L858R et Del.19, EGFR L858R-T790M et Del.19-T790M et permettent d'identifier les interacteurs dérégulés et en commun entre ces formes. Les processus biologiques dans lesquels chacun de ces groupes d'interacteurs sont impliqués sont identifiés par une analyse Gene Ontology (niveau 5). Dereg., Dérégulation.

Cette analyse montre qu'environ 30% de tous les interacteurs dérégulés par les mutations primo-activatrices EGFR L858R et Del.19 sont partagés et ce recouvrement augmente jusqu'à environ 60% lorsque la mutation secondaire de résistance T790M s'ajoute à ces premières mutations. Cette analyse permet finalement de sélectionner 3 groupes : 76 interacteurs dérégulés par les mutations primo-activatrices (L858R et Del. 19), 177 interacteurs dérégulés par la seule mutation de résistance aux ITKs (T790M), et 175 interacteurs dérégulés par les doubles mutations de l'EGFR (L858R-T790M et Del.19-T790M). Lorsqu'on s'intéresse aux TOP5 des termes GO « Biological Process » enrichis pour chacun de ces groupes, on remarque que la voie des MAPK et ses régulateurs sont particulièrement représentés parmi les interacteurs dérégulés par les mutations primo-activatrices, ce qui n'est pas le cas pour les deux autres groupes. En effet, ceux-ci sont plutôt enrichis pour des processus biologiques impliquant le transport cellulaires (« Protein transport »), l'architecture membranaire (« Cell morphogenesis », « Cell migration ») et la transduction intracellulaire de la signalisation (« Enzyme linked receptor protein signaling pathway », « Intracellular signal transduction »).

La comparaison de ces trois groupes d'interacteurs permet de mettre en évidence les partenaires dont les interactions avec l'EGFR sont préférentiellement altérées avec la mutation de résistance T790M, ce que soit dans le sens d'un gain ou une perte d'interaction (Figure 44).

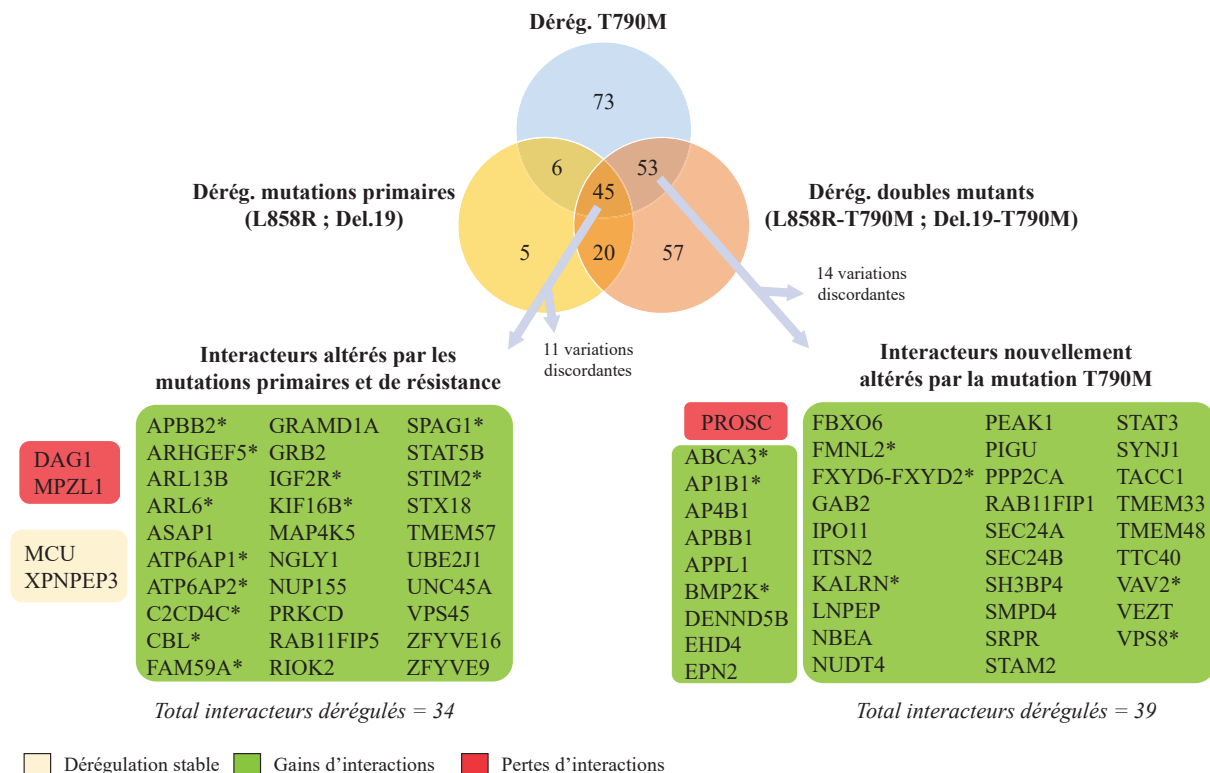


Figure 44. Identification des interacteurs affectés par la mutation EGFR T790M. Le diagramme de Venn permet de sélectionner deux groupes d'interacteurs dérégulés. Le premier est constitué des interacteurs altérés par la mutation EGFR T790M seule, altérés par les mutations primaires EGFR L858R et EGFR Del.19 et dont l'altération est similaire ou intensifiée par la mutation de résistance. Le second groupe est constitué des interacteurs nouvellement dérégulés par le gain de la mutation EGFR T790M, et aussi par cette mutation seule. Les interacteurs marqués d'une étoile (*) ne sont pas identifiés pour la protéine sauvage. Analyse basée exclusivement sur les peptides totaux identifiés. Dérég., Dérégulation.

Finalement, cette comparaison permet de sélectionner deux groupes d'interacteurs dérégulés. Un premier groupe composé de 34 interacteurs altérés par la mutation EGFR T790M seule, altérés par les mutations primaires EGFR L858R et EGFR Del.19 et dont l'altération est similaire ou intensifiée par la mutation de résistance. Il s'agit du groupe le plus intéressant d'un point de vue clinique car ces interacteurs sont représentatifs des patients avant et après traitement aux ITKs, ainsi que de la minorité de patients présentant la mutation de résistance EGFR T790M seule, sans avoir reçus de thérapies ciblées. Une analyse GO (non présentée) montre que la voie des MAPK n'est plus enrichie, et que le seul terme de signalisation cellulaire significativement identifié est « enzyme linked receptor protein signaling pathway » ($P\text{-value}=8.10^{-3}$). Celui-ci inclue les protéines ATP6AP1, CBL, FAM59A, GRB2, IGF2R, KIF16B, MPZL1, PRKCD, STAT5B, ZFYVE16, ZFYVE9. La voie de l'endocytose est également enrichie ($P\text{-value}=3.10^{-3}$) avec les protéines ASAP1, CBL, IGF2R, RAB11FIP5, VPS45, ZFYVE16, ZFYVE9.

Le second groupe intéressant est composé de 39 protéines (53 moins 14 discordantes), dérégulées seulement lors de la mutation secondaire EGFR T790M ou lorsque cette mutation survient seule. Il s'agit des interacteurs nouvellement altérés et induit par la mutation secondaire de résistance. L'analyse des termes GO ne montre aucun enrichissement au niveau des processus biologiques. Le seul terme enrichi significativement concerne la présence du domaine d'interaction avec la membrane plasmique « Pleckstrin homology-like » et regroupe les protéines APBB1, APPL1, GAB2, ITSN2, KALRN, NBEA, VAV2.

Globalement, cette analyse montre que la majorité des dérégulations observées concerne des gains d'interactions, et que 20 interacteurs sont spécifiques des protéines EGFR mutées, c'est à dire non retrouvés à proximité de la protéine sauvage. Afin d'affiner cette sélection, je me suis intéressé d'une part aux interacteurs les plus fortement dérégulés, et d'autre part à la présence d'un réseau fonctionnel qui pourrait être établi entre ces interacteurs (Figure 45).

Concernant les interacteurs altérés par les mutations primaires et de résistance, les interacteurs les plus dérégulés ont été sélectionnés en choisissant arbitrairement un ratio de peptides entre la forme mutée et sauvage supérieur ou égal à 16 ($\log_2 \geq 4$). Concernant les interacteurs nouvellement altérés par la mutation EGFR T790M, les interacteurs les plus dérégulés ont été sélectionnés en choisissant arbitrairement un ratio de peptides entre la forme mutée et sauvage supérieur ou égal à 8 ($\log_2 \geq 3$).

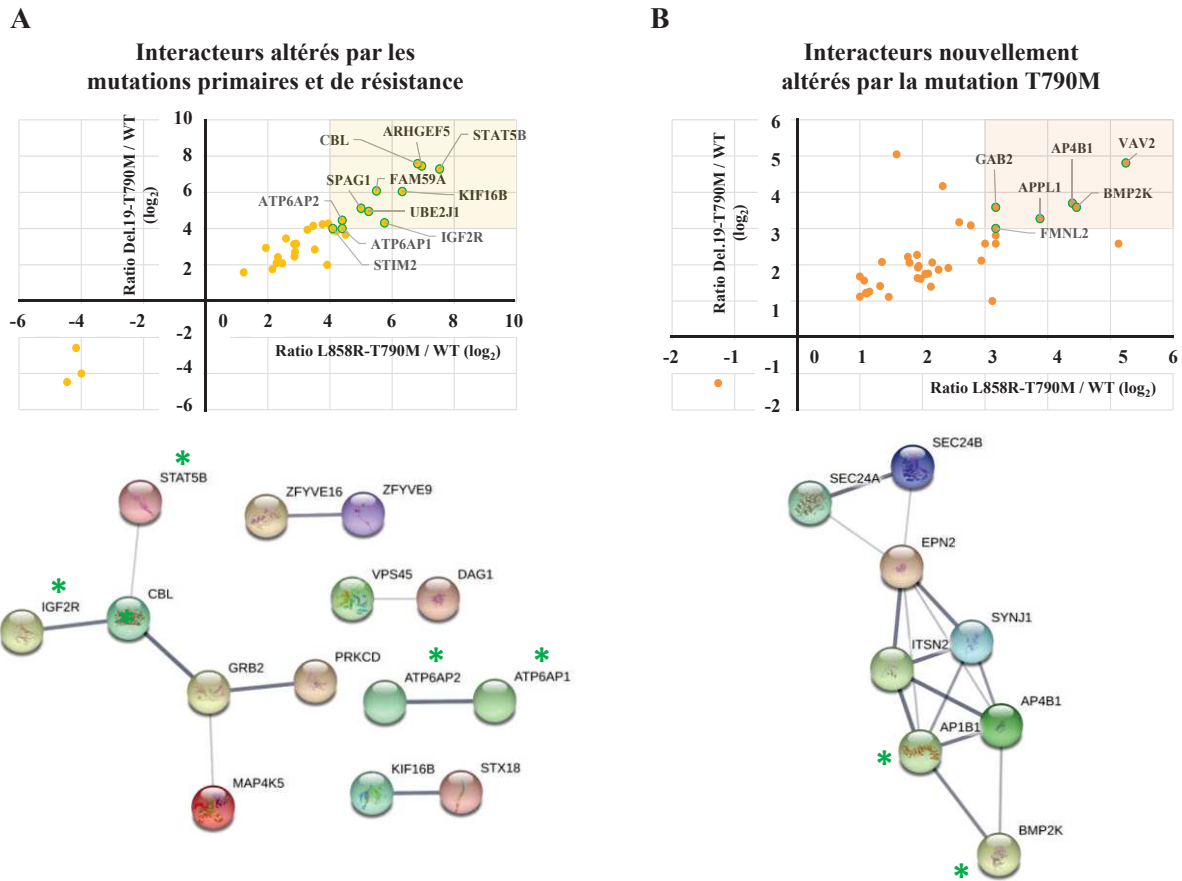


Figure 45. Identification des réseaux fonctionnels et interacteurs les plus spécifiques aux mutations primaires et de résistance, ou nouvellement altérés par la mutation de résistance. (A) Analyse appliquée aux interacteurs altérés par les mutations primaires et de résistance ou par la mutation de résistance seule. Le graphique représente les ratios des peptides totaux identifiés entre les formes EGFR Del.19-T790M et sauvage en fonction des ratios de peptides totaux identifiés entre les formes EGFR L858R-T790M et sauvage. Pour sélectionner les interacteurs les plus pertinents (dans la zone colorée, ceclés en vert), un ratio de 4 (log₂) a été défini arbitrairement. L'analyse STRING montre un réseau fonctionnel incluant 6 protéines et 4 réseaux binaires. Les interacteurs marqués d'une étoile (*) ont été identifiés parmi les plus pertinents. (B) Analyse appliquée aux interacteurs altérés seulement lors de l'acquisition de la mutation de résistance ou par lorsque celle-ci survient seule. Le graphique représente les ratios des peptides totaux identifiés entre les formes EGFR Del.19-T790M et sauvage en fonction des ratios de peptides totaux identifiés entre les formes EGFR L858R-T790M et sauvage. Pour sélectionner les interacteurs les plus pertinents (dans la zone colorée, ceclés en vert), un ratio de 3 (log₂) a été défini arbitrairement. L'analyse STRING montre un seul réseau fonctionnel incluant 8 protéines. Les interacteurs marqués d'une étoile (*) ont été identifiés parmi les plus pertinents.

Parmi les interacteurs dérégulés par les mutations primaires L858R et Del.19, et dont l'altération est accrue par la mutation secondaire, les protéines les plus dérégulées sont ARHGEF5, ATP6AP1, ATP6AP2, CBL, FAM59A, IGF2R, KIF16B, SPAG1, STAT5B, STIM2, UBE2J1. L'analyse du réseau fonctionnel reliant ces interacteurs montre un réseau principal composé de 6 protéines et 4 réseaux binaires. Le premier réseau inclut les protéines CBL, IGF2R et STAT5B qui sont parmi les plus dérégulées, et est complété par les protéines GRB2, MAP4K5 et PRKCD. Il s'agit d'un réseau de signalisation comprenant un récepteur à l'insuline (IGF2R), des protéines adaptatrices (CBL, GRB2), deux kinases (MAP4K5 et PRKCD) et un facteur de transcription (STAT5B). Parmi les réseaux binaires, le réseau le plus pertinent comprend les protéines parmi les plus dérégulées ATP6AP1 et ATP6AP2. Il

s'agit de pompes à protons nécessaires à l'acidification des compartiments intracellulaires telles que les endosomes et l'appareil de Golgi (Figure 45A).

Parmi les interacteurs dérégulés seulement lors de l'acquisition de la mutation de résistance, et aussi lorsque cette mutation survient seule, les protéines les plus dérégulées sont AP4B1, APPL1, BMP2K, FMNL2, GAB2 et VAV2. L'analyse fonctionnelle des interacteurs de ce groupe fait apparaître un réseau dense d'interactions incluant 8 protéines, AP4B1 en constitue le cœur avec des liens notamment vers ITSN2, SYNJ1 et AP1B1. Connectée à ce sous réseau, la protéine kinase BMP2K est identifiée. Plus globalement, ce réseau peut être divisé en deux sous-groupes fonctionnels : un premier sous-réseau, incluant toutes les protéines hormis BMP2K, impliqué dans le trafic vésiculaire et l'endocytose ; et un deuxième sous réseau impliquant BMP2K, une sérine/thréonine kinase peu étudiée, possédant un signal d'adressage nucléaire et jouant un rôle potentiel dans la différenciation cellulaire des ostéoblastes (Figure 45B).

De manière intéressante, quand on analyse le réseau fonctionnel constitué avec l'ensemble des interacteurs des deux groupes identifiés plus haut, on met en évidence la complémentation des réseaux fonctionnels (Figure 46).

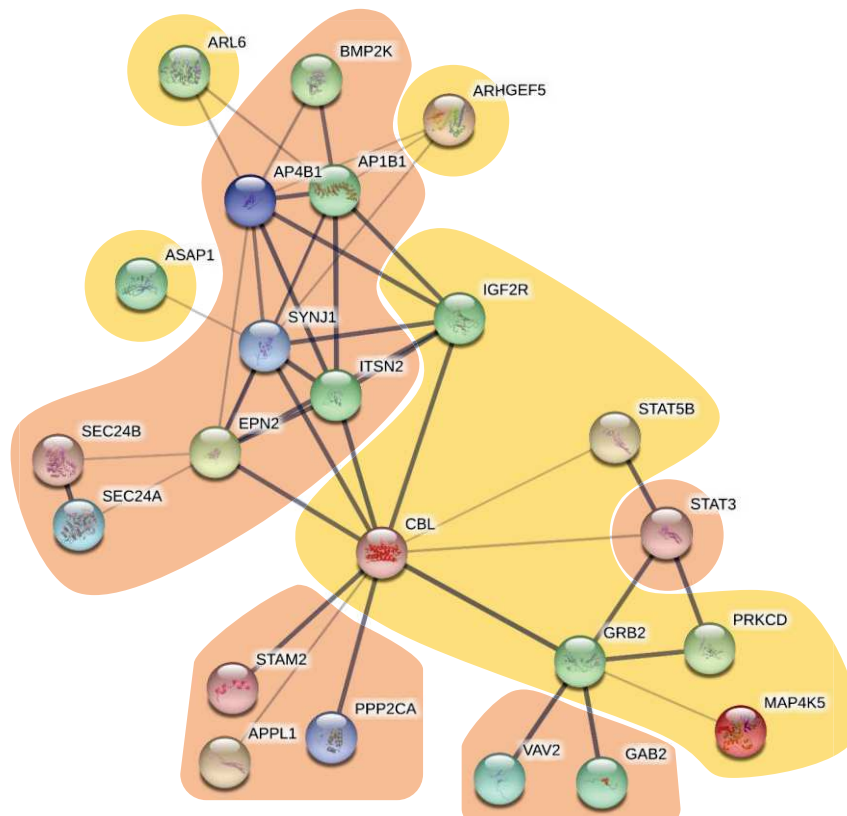


Figure 46. Réseau fonctionnel d'interaction des interacteurs dérégulés par les mutations primaires et lors de l'acquisition de la mutation de résistance. Les interacteurs en jaune forment un premier réseau fonctionnel dérégulé par les mutations primaires EGFR L858R et Del.19. Lors de l'acquisition de la mutation de résistance EGFR T790M, les interacteurs en orange sont dérégulés et se connectent au premier réseau principalement via IGF2R et CBL.

Ce dernier réseau montre comment les dérégulations qui apparaissent seulement lors de l'acquisition de la mutation de résistance se connectent au réseau décrit par les interacteurs altérés par les mutations primaires EGFR L858R et EGFR Del.19, et dont l'altération est similaire ou intensifiée par la mutation de résistance. On observe que CBL et IGF2R sont les deux interacteurs dérégulés par les mutations primo activatrices qui assurent la connexion entre ces deux réseaux.

II.3 Discussion

Sur les 683 interacteurs uniques identifiés par BioID pour les différents mutants EGFR, les analyses successives ont permis de sélectionner 34 interacteurs dérégulés par les mutations primo-activatrices par rapport à l'EGFR sauvage, et dont les dérégulations sont intensifiées par la présence de la mutation de résistance T790M, et 39 autres interacteurs dérégulés seulement par l'acquisition de la mutation T790M. L'ensemble de ces protéines dérégulées est également dérégulé par la mutation de résistance seule.

Pour le premier groupe, une connexion directe avec un réseau fonctionnel impliqué dans la signalisation cellulaire a été mise en évidence. Ce réseau comprend plusieurs partenaires nécessaires à l'activation d'un récepteur tels que l'EGFR, comme le récepteur IGF2R, les protéines adaptatrices CBL et GRB2, les kinases MAP4K5 et PRKCD et le facteur de transcription STAT5B. De plus, CBL, GRB2, PRKCD et STAT5B sont tous référencés comme des interacteurs directs de l'EGFR (NCBI). Concernant le récepteur IGF2R, il est impossible en l'état de conclure sur une dimérisation des deux récepteurs et des expériences complémentaires seront nécessaires pour identifier une éventuelle colocalisation avec le récepteur EGFR activé. Concernant la protéine RB1, cible de PRKCD, nous pouvons également évaluer sa phosphorylation en contexte d'inhibition d'expression de l'IGF2R et/ou de l'EGFR. Plusieurs de ces protéines sont encore peu étudiées mais il est intéressant de remarquer qu'elles ont toutes été reliées à des processus cancéreux. Plus particulièrement, PRKCD est surexprimée dans de nombreux cancers dont le cancer du pancréas, et favorise la progression tumorale (Mauro et al., 2010). Dans les adénocarcinomes pulmonaires, il s'agit aussi d'une cible de KRAS lorsque celui-ci est muté (Symonds et al., 2011). Le gène *IGF2R* a quant à lui été identifié comme suppresseur de tumeur, il est muté dans les cancers du poumon non à petites cellules et associé à l'initiation tumorale et la progression tumorale (Ghosh et al., 2003). Enfin, le proto-oncogène CBL est une protéine adaptatrice à domaine SH3, déjà référencé comme interacteur de l'EGFR et GRB2. Il s'agit d'une E3 ubiquitine ligase multifonctionnelle qui selon le contexte cellulaire régule la signalisation en recrutant des effecteurs à la membrane plasmique, ou la localisation des récepteurs membranaires en favorisant l'endocytose (Liyasova et al., 2015). Cette protéine est fréquemment mutée dans le cancer du poumon non à petites cellules, et sa surexpression accroît la viabilité et la motilité cellulaire (Tan et al., 2010). Finalement, l'ensemble de ce réseau présente des protéines identifiées comme surexprimées ou mutées lors de cancers, et notamment lors du cancer du poumon non à petites cellules.

L'analyse des interacteurs nouvellement dérégulés par la mutation de résistance a permis de mettre en évidence un réseau d'interaction fonctionnel incluant principalement des protéines régulant le trafic membranaire. Au-delà de cette fonction, le plus étonnant est de voir que le cœur de ce réseau se connecte directement au premier réseau caractéristique des mutations primo-activatrices, et ceci *via* les protéines CBL et IGF2R. Pris dans son ensemble, ce réseau représente un modèle fonctionnel de résistance de l'EGFR aux ITKs. Effectivement, lorsque l'EGFR présente une mutation L858R ou Del.19, on identifie à proximité les protéines CBL et IGF2R avec GRB2, STAT5B et PRKCD. En présence de la mutation secondaire de résistance EGFR T790M, nous observons que l'ensemble de ces protéines interagissent encore plus fortement avec l'EGFR, et qu'un autre groupe de protéines, jusqu'alors non altéré par les mutations L858R et Del.19, est identifié à proximité, et en lien avec CBL et IGF2R. Ceci suggère d'une part que ces deux protéines auraient un rôle important dans la signalisation de l'EGFR portant une mutation primo-activatrice, et d'autre part qu'elles permettraient de coordonner des partenaires nouvellement dérégulés par la mutation de résistance T790M. Très récemment, il a d'ailleurs été montré que l'inhibition de CBL créait une vulnérabilité aux traitements ciblant MEK dans la lignée H1975, résistante aux ITKs (Tan et al., 2017).

Afin de valider la pertinence de CBL comme cible thérapeutique des tumeurs présentant une simple ou double mutation de l'EGFR, nous pouvons tout d'abord valider son interaction avec l'EGFR en utilisant la technique de PLA. Ensuite, nous pouvons sélectionner un panel de cellules d'adénocarcinomes pulmonaires et évaluer l'effet de l'inhibition de CBL par ARN interférant sur la viabilité cellulaire en présence ou non d'une exposition à un ITK comme l'erlotinib. Pour les cellules présentant une simple mutation de l'EGFR, nous pouvons ainsi évaluer l'effet synergique de l'inhibition d'expression de CBL et d'un traitement à l'erlotinib. Au-delà de l'évaluation à un instant donné de l'inhibition d'une ou plusieurs cibles thérapeutiques sur la viabilité cellulaire, nous pouvons également évaluer le délai dans l'acquisition de la mutation de résistance. Effectivement, une culture continue de cellules portant une mutation de l'EGFR L858R (par exemple la lignée cellulaire H3255) ou Del.19 (par exemple la lignée cellulaire H1650) et exposée à un ITK comme l'erlotinib permet de générer la mutation de résistance EGFR T790M (Chmielecki et al., 2011; Ghosh et al., 2012). Modifier ces cellules pour exprimer constitutivement un ARN interférant ciblant CBL, et les cultiver avec l'erlotinib permettra d'évaluer le délai dans la survenue de la mutation de résistance comparé au traitement à l'erlotinib seul. Cette dernière expérience permettra d'évaluer l'intérêt d'une combinaison de drogues ciblant simultanément l'EGFR et CBL pour traiter les patients ayant une mutation EGFR L858R ou Del.19, et ceci afin de retarder la survenue de la résistance aux ITKs.

III. Réseau d'interaction de BRAF : application au NSCLC

III.1 Introduction de l'étude

Les mutations affectant BRAF compte pour environ 7% des adénocarcinomes pulmonaires et 40% des mélanomes. La mutation BRAF V600E est la principale altération et les patients porteurs de cette mutation sont éligibles aux thérapies ciblées utilisant le vemurafenib et le dabrafenib. Comme dans le cas de l'EGFR un échappement thérapeutique est observé environ 1 an après le début du traitement et aucune thérapie n'est à ce jour disponible pour traiter ces résistances. Afin d'identifier les partenaires de BRAF dérégulés par la mutation V600E, j'ai caractérisé les réseaux d'interactions de la protéine sauvage et mutée. Les données de MS sont disponibles en Annexe 7.

III.2 Résultats

III.2.a Réseau d'interactions de la protéine sauvage

La protéine BRAF sauvage compte 347 interacteurs identifiés significativement (avec au moins deux peptides uniques). Ces interacteurs sont principalement localisés dans le cytoplasme et dans la mitochondrie. Les interacteurs majeurs de BRAF comme ARAF, RAF1, KRAS, NRAS, MAP2K1, MAP2K2 et MAP2K3 sont identifiées par BioID, ce qui abonde dans la pertinence de notre réseau d'interaction identifié. De manière particulièrement intéressante, nous remarquons une très forte proximité des interacteurs avec le métabolisme carboné et la phosphorylation oxydative (Figure 47).

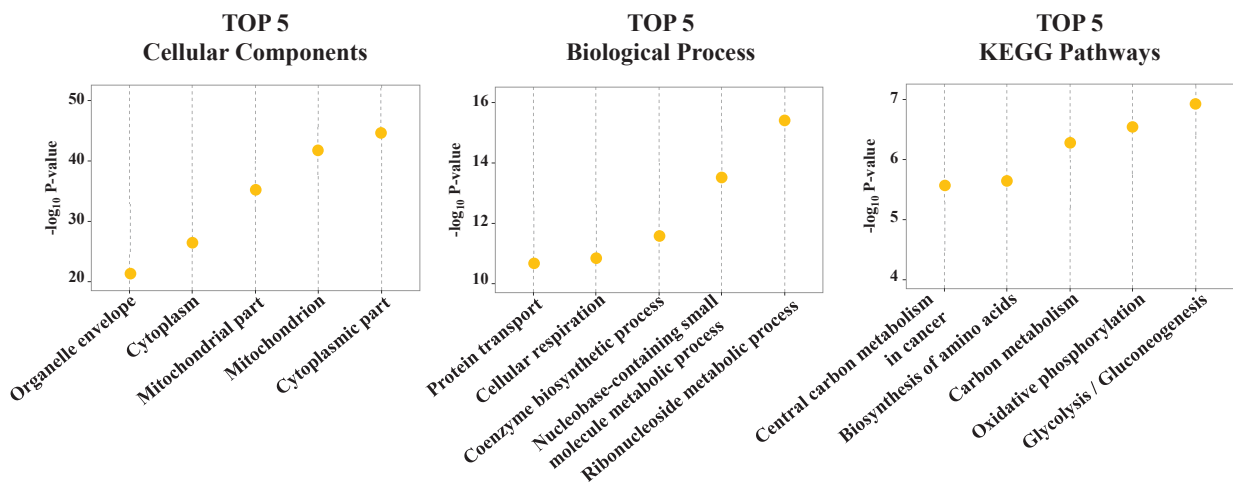


Figure 47. Analyse des composants cellulaires, processus biologiques et voies cellulaires des interacteurs de BRAF WT. Les composants cellulaires et processus biologiques ont été identifiés par une analyse Gene Ontology (niveau 5) et les voies cellulaires par une analyse KEGG. Seuls les TOP5 des termes enrichis sont représentés.

Plus particulièrement, j'identifie des enzymes de la glycolyse comme les hexokinases HK1 et HK2, les phosphofructokinases PFKL, PFKM, PFKP, PKM, le glyceraldehyde-3-phosphate déshydrogénase (GAPDH), et enfin la pyruvate kinase (PKM) qui catalyse la conversion du phosphoénolpyruvate en pyruvate. D'autres enzymes en aval de la glycolyse sont également identifiées comme la lactate déshydrogénase (LDHB) et la pyruvate déshydrogénase (PDHB), cette dernière enzyme permettant la conversion du pyruvate en acétyl-CoA qui rejoindra le cycle de Krebs. La phosphorylation oxydative est quant à elle enrichie dans le réseau d'interaction avec l'identification plusieurs protéines impliquées dans la chaîne respiratoire mitochondriale, et plus particulièrement le complexe I (NDUFA5, NDUFA10, NDUFA13, NDUFS1, NDUFS2, NDUFS3, NDUFS7, NDUFV1, NDUFV2) et le complexe V (ATP5A1, ATP5B, ATP5C1, ATP5L, ATP5O) (Tableau 8).

Tableau 8. Tableau résumant les interacteurs attribués aux catégories KEGG Glycolysis/Gluconeogenesis et Oxidative phosphorylation.

KEGG Catégories	P-value	Nombre de protéines identifiées	Protéines
Glycolysis / Gluconeogenesis (hsa00010)	1,2E-07	16	ALDH1B1, ALDH2, ALDH7A1, DLAT, DLD, ENO1, GAPDH, HK1, HK2, LDHA, LDHB, PDHB, PFKL, PFKM, PFKP, PKM2
Oxidative phosphorylation (hsa00190)	3,3E-07	21	ATP5A1, ATP5B, ATP5C1, ATP5L, ATP5O, ATP6V1H, COX2, NDUFA10, NDUFA13, NDUFA5, NDUFA9, NDUFS1, NDUFS2, NDUFS3, NDUFS7, NDUFV1, NDUFV2, PPA1, UQCR10, UQCRC2, UQCRFS1

III.2.b Dynamique du réseau d'interaction de BRAF

En se basant sur le nombre total de peptides dénombrés entre l'analyse de BRAF sauvage et de la protéine mutée, au final ce sont 404 interacteurs uniques qui sont identifiés. En calculant pour chacun d'eux le ratio de peptides obtenus pour BRAF V600E et WT on observe que seulement 15% de ces interacteurs sont dérégulés, et il s'agit principalement de pertes ou d'appauvrissements d'interactions pour la protéine mutée (Figure 48).

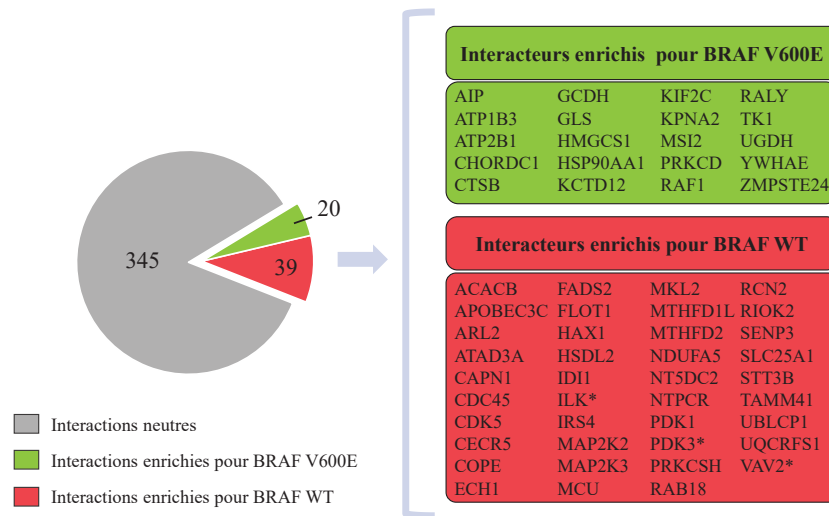


Figure 48. Identification des interacteurs dérégulés entre les formes BRAF WT et BRAF V600E. Les interacteurs dérégulés sont les interacteurs pour lesquels au moins deux fois plus de peptides sont mesurés pour la forme sauvage ou mutée. Sur les 404 interacteurs uniques, 59 protéines sont dérégulées. Les interacteurs marqués d'une étoile (*) sont exclusifs à la protéine sauvage.

Alors que la protéine sauvage montrait une claire connexion avec les voies canoniques du métabolisme carboné (glycolyse) et oxydatif (respiration mitochondriale), très peu de protéines de ces voies sont retrouvées dérégulées par la mutation BRAF V600E. L'altération du métabolisme est ici moins évidente car seulement deux termes associés sont enrichis, il s'agit de « Coenzyme biosynthetic process » et « Acyl-CoA biosynthetic process » qui comprennent tous deux des interacteurs qui interagissent plus faiblement avec la protéine mutée. Concernant les localisations, les termes relatifs à la mitochondrie et au cytosol sont de nouveaux retrouvés, et comprennent aussi principalement des protéines qui interagissent plus faiblement avec la protéine mutée. Plus globalement très peu de termes sont enrichis, ce qui amène à identifier protéine par protéine les plus pertinentes.

Parmi les protéines qui interagissent plus fortement avec BRAF V600E, deux protéines kinases sont présentes, il s'agit de RAF1 (ou CRAF) et PRKCD. En prenant en compte les mécanismes activation de BRAF, une troisième protéine particulièrement intéressante est YWHAE (ou 14-3-3 Epsilon). L'interaction directe entre ces trois protéines RAF1, PRKCD et YWHAE a été démontré (NCBI) et une analyse STRING des interacteurs préférentiels de la forme BRAF mutée permet d'inclure d'autres protéines dans un réseau fonctionnel (Figure 49).

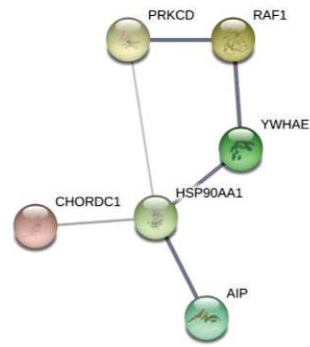


Figure 49. Réseau d'interaction fonctionnel des interacteurs préférentiels de BRAF V600E. Réseau analysé par STRING.

Finalement, ce sont la chaperonne HSP90AA1, la co-chaperonne CHORDC1 et le facteur de transcription AIP qui complètent ce premier réseau. Ces deux dernières protéines sont très peu documentées.

III.3 Discussion

Bien que BRAF soit décrite principalement pour son implication dans la voie des MAPK en tant qu'effecteur des protéines RAS et activateur des protéines MAP2K, l'analyse du réseau d'interaction de la protéine sauvage a montré une très forte proximité de BRAF avec la glycolyse et la chaîne respiratoire mitochondriale. Ce lien doit être validé par d'autres techniques (AP-MS, PLA) mais à ma connaissance, ce lien direct n'a jamais été démontré. Ce qui est particulièrement surprenant est d'identifier des protéines aussi bien en amont de ces voies ainsi qu'en aval. Effectivement, les hexokinases sont les premières enzymes de la glycolyse nécessaire à la conversion du glucose en glucose-6-phosphate et j'identifie également la dernière enzyme de cette voie, la pyruvate kinase, qui converti le phosphoénolpyruvate en pyruvate et génère de l'ATP. Des enzymes intermédiaires de cette voie comme les isoformes de la phosphofructokinase, le glyceraldehyde-3-phosphate déshydrogénase, et les enzymes qui dirigent la conversion du pyruvate selon le statut énergétique des cellules (lactate déshydrogénase, pyruvate déshydrogénase) sont également identifiés, ce qui placerait véritablement BRAF au cœur du métabolisme carboné. De plus, un lien avec le métabolisme mitochondrial est aussi observé par la proximité de BRAF avec les complexes I et V de la chaîne respiratoire, lieu principale de production d'énergie dans les cellules eucaryotes. Plus globalement, il est intéressant de noter que chacune de ces enzymes, qu'elles soient cytoplasmiques ou mitochondriales, produisent ou consomment de l'ATP, ou du pouvoir réducteur NADH, H⁺.

Malgré les outils statistiques utilisés pour identifier les interacteurs fiables de BRAF, considérer des protéines du métabolisme par BioID doit être fait cependant avec précaution. Effectivement, la biotine

est le coenzyme de toutes les carboxylases et celles-ci sont donc naturellement biotinylées. Cependant, par la quantité de protéines en lien avec le métabolisme ici identifiées, cette proximité avec BRAF ne peut être négligée.

La comparaison du réseau de la protéine BRAF sauvage à BRAF V600E a finalement montré que peu de protéines sont dérégulées. Étonnement, malgré le lien étroit de BRAF WT avec des protéines carrefour du métabolisme cellulaire comme la phosphofructokinase ou la lactate déshydrogénase, aucune de ces protéines n'est dérégulée par la mutation activatrice. Seulement 20 protéines interagissent préférentiellement avec la forme mutée mais aucune fonction n'est particulièrement enrichie, ce qui rend difficile l'identification d'une fonction cellulaire particulièrement attribuable à la BRAF V600E.

Au final, seul le réseau comprenant RAF1 et YWHAE semble le plus intéressant, notamment car ces deux protéines participent à l'activation de BRAF (hétérodimérisation avec RAF1, stabilisation du complexe par les protéines 14-3-3) et semblent interagir plus fortement avec la forme mutée. L'implication de RAF1 dans la croissance tumorale, et plus particulièrement dans le cancer du poumon et le mélanome, est richement documentée (Blasco et al., 2011; Karreth et al., 2011; Zanucco et al., 2014), comme celle de la protéine 14-3-3 (Hermeking, 2003; Wilker and Yaffe, 2004). Les protéines PRKCD et HSP90AA1 qui complètent ce réseau sont aussi intéressantes car un lien avec le développement cancéreux a aussi été mis en évidence. Les protéines HSP90 occupent quant à elle une place particulière car il s'agit de protéines ubiquitaires essentielles au repliement et à l'adressage correct d'un large nombre de protéines dont de nombreuses kinases (Taipale et al., 2012; Vaughan et al., 2008). De plus, cette classe particulière de chaperonne est une cible thérapeutique dans de nombreux cancers dont le cancer du poumon et des essais cliniques sont en cours (Acquaviva et al., 2012; Barrott and Haystead, 2013; Whitesell and Lindquist, 2005).

La taille réduite de ce réseau donne l'avantage d'envisager des expériences de validation à faible échelle. L'utilisation d'ARN interférant ciblant ces interacteurs dans la lignée de mélanome A375 portant la mutation BRAF V600E permettrait de vérifier l'effet de ces cibles sur la viabilité et l'invasion cellulaire, ce dernier point étant caractéristique des mélanomes. Si ces expériences sont concluantes, une analyse comparée de ces cibles dans une lignée de poumon permettra de valider la dépendance de ces interacteurs à BRAF plutôt qu'au contexte cellulaire.

IV. Réseau d'interaction de EML4-ALK : application au NSCLC

IV.1 Introduction de l'étude

Les translocations du gène *ALK* avec celui d'*EML4* sont observées dans environ 5% des adénocarcinomes pulmonaires. Les variants 1, 3a ou 3b comptent pour plus de la moitié des cas, mais d'autres variants existent et incluent d'autres exons d'*EML4* ou d'autres gènes fusions. Quel que soit le variant ou le partenaire de fusion, ces translocations incluent toute le domaine tyrosine kinase du récepteur ALK. Finalement, c'est l'identification de la translocation de ce domaine tyrosine kinase qui rend éligible les patients au crizotinib, traitement qui permet une survie sans progression de la maladie quel que soit le type de fusion. Caractériser les réseaux d'interactions des variants principaux (EML4-ALK v1 et EML4-ALK v3a) renseigne alors sur les connexions de ces variants avec l'environnement cellulaire et les voies de signalisation ou les fonctions suractivées par ces chimères. L'identification de ces interacteurs pourrait permettre d'identifier de nouvelles cibles dans le traitement des adénocarcinomes, qu'il s'agisse de monothérapies ou de combinaison de drogues avec le crizotinib. Les données de MS sont disponibles en Annexe 8.

IV.2 Résultats

Pour cette étude, j'ai réalisé les constructions pour le BioID des variants de translocation 1 et 3a, ainsi que des protéines ALK et EML4 sauvages et des portions de ces 2 protéines incluent dans les translocations. À l'heure actuelle, seul l'interactome d'EML4-ALK v1, EML4-ALK v3a et EML4 WT a été analysé par MS. Étonnement, l'interactome du variant 1 apparait moins riche que le variant 3a. En effet, 85 protéines ont été identifiées significativement pour le variant 1 contre 190 pour le variant 3a. Bien que des peptides appartenant à des protéines de signalisation comme PIK3CA ou GAB1 soient identifiées, la forte présence de protéines chaperonnes comme les Hsp40, Hsp70 ou les protéines prefoldin amène à supposer une instabilité du variant 1 fusionné à BirA*. Finalement, seul le variant EML4-ALK v3a sera inclus dans cette analyse et, à l'inverse des analyses précédentes, comme la seule expression de cette protéine de fusion est en elle-même propre à un processus cancéreux, l'ensemble des interacteurs identifiés peuvent être considérés comme dérégulés. Toutefois, afin de disposer d'un filtre supplémentaire, tous les interacteurs significatifs d'EML4 seront retirés de la liste des interacteurs d'EML4-ALK.

Après soustraction des interacteurs d'EML4, 166 interacteurs spécifiques à EML4-ALK sont significativement identifiés. Ceux-ci sont principalement localisés dans le cytosol et impliqués dans l'organisation du cytosquelette, la signalisation et la régulation du cycle cellulaire. Pour cette dernière fonction, le centrosome est significativement enrichi. Lorsqu'on analyse plus en détail les voies de signalisation impliquées, la voie des récepteurs aux facteurs de croissance (« ErbB signaling pathway»),

la voie de l'insuline et la voie Rap1 sont particulièrement enrichies (Figure 50). Toutefois, ces différentes voies incluent des protéines en commun comme les protéines adaptatrices de récepteurs à domaine tyrosine kinase CRK ou CRKL, et les protéines PK3CA, PIK3R1, PIK3R2 du complexe PIK3. Au-delà de ces voies de signalisation spécifiques, il est intéressant de remarquer que les domaines protéiques SH2 et SH3 relatif à la transduction de signaux de signalisation sont fortement enrichis et relie une nouvelle fois EML4-ALK v3a à l'activation de voies de signalisations essentielles aux cellules (Tableau 9).

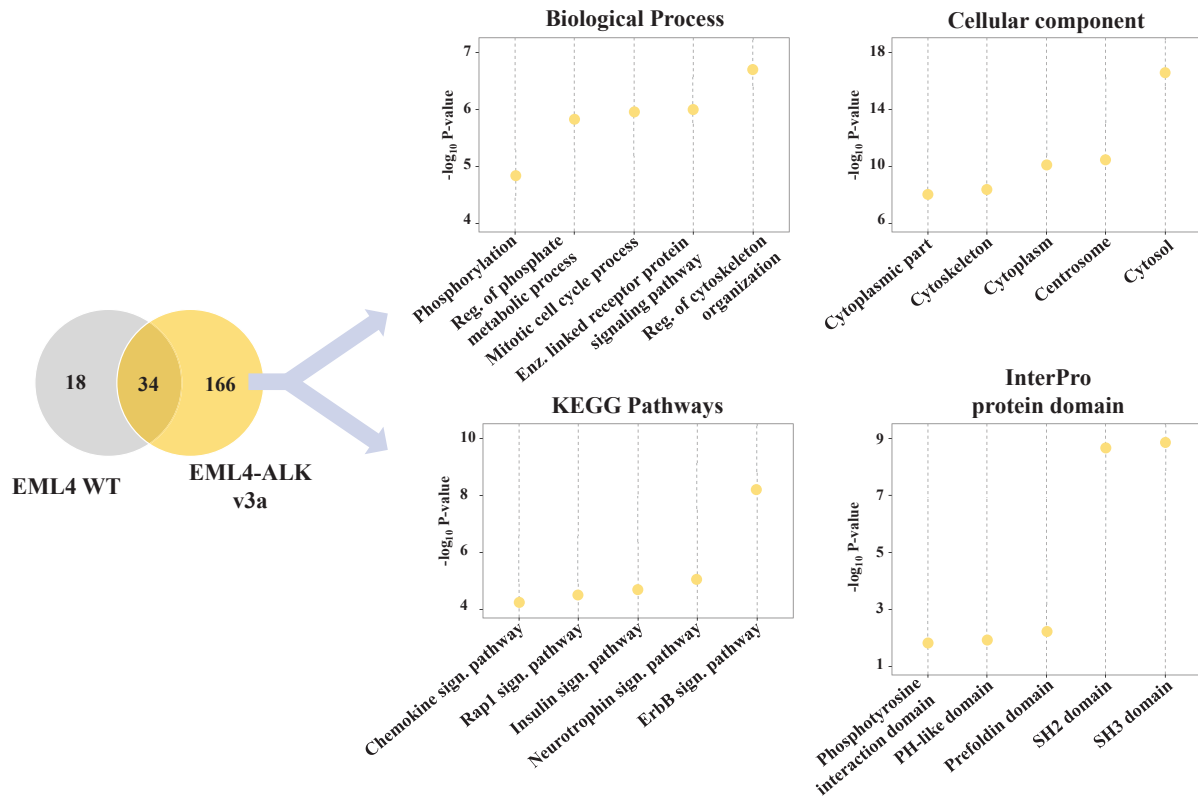


Figure 50. Identification des procédés biologiques, composants cellulaires, voies de signalisation et domaines protéiques enrichis par EML4-ALK v3a. Les procédés biologiques et les composants cellulaires enrichis ont été identifiés par Gene Ontology (niveau 5). Les voies cellulaires enrichies ont été identifiées par KEGG et les domaines protéiques par InterPro. Les termes enrichis présentés sont le TOP5 des catégories respectives. Enz., Enzyme ; Reg., Regulation ; Sign., signaling.

Tableau 9. Sélection des catégories KEGG et InterPro enrichies.

Catégories KEGG	Nombre de protéines identifiées	P-value	Protéines
ErbB signaling pathway (hsa04012)	14	5,30E-09	AKT2, CBL, CRK, CRKL, GAB1, GRB2, NCK1, NCK2, PIK3CA, PIK3R1, PIK3R2, PLCG1, SHC1, STAT5B
Insulin signaling pathway (hsa04910)	12	2,97E-05	AKT2, CBL, CRK, CRKL, GRB2, IRS2, IRS4, PIK3CA, PIK3R1, PIK3R2, SHC1, SORBS1
Rap1 signaling pathway (hsa04015)	14	4,29E-05	AKT2, BCAR1, CRK, CRKL, EPHA2, MLLT4, PFN1, PIK3CA, PIK3R1, PIK3R2, PLCG1, RHOA, SIPA1L2, SIPA1L3
Domaines protéiques InterPro	Nombre de protéines identifiées	P-value	Protéines
SH3 domain (IPR001452)	18	1,49E-09	ARHGEF5, BCAR1, CRK, CRKL, CTTN, EPS8, FBNP1L, GRB2, LASP1, NCK1, NCK2, PIK3R1, PIK3R2, PLCG1, SORBS1, TJP1, TJP2, TP53BP2
SH2 domain (IPR000980)	14	2,09E-09	BCAR3, CBL, CRK, CRKL, GRB2, NCK1, NCK2, PIK3R1, PIK3R2, PLCG1, SHC1, STAT1, STAT2, STAT5B

Le détail des domaines protéiques enrichis montre que certaines protéines de signalisation sont indépendantes des voies de signalisation précédemment citées. C'est le cas de TP53BP2, ligand de TP53, et des facteurs de transcription STAT1 et STAT2 qui peuvent être des protéines essentielles à la signalisation de la fusion EML4-ALK (Tableau 9).

Enfin, l'analyse du réseau fonctionnel de ces 166 protéines met en évidence deux réseaux denses et interconnectés (Figure 51). Le premier réseau est un réseau de signalisation cellulaire comprenant notamment les protéines des voies ErbB et Rap1 comme GRB2, PIK3CA, PIK3C2B, PIK3R1, PIK3R2, RHOA. Le second réseau est celui impliqué dans la régulation de la division cellulaire, et comprend notamment les protéines du centrosome CEP97, CEP152, CEP192 et la protéine de contrôle du cycle cellulaire comme CDK1. Connectant ces deux premiers réseaux, un troisième réseau moins dense et plus étendu comprend des protéines aux fonctions plus diverses comme des protéines des fuseaux mitotiques (KIF2A, KIF14), des chaperonnes (HSP90AB1), la protéine adaptatrice IRS2 et des protéines du trafic vésiculaire comme ARF1 et ARFGAP1 (Figure 51).

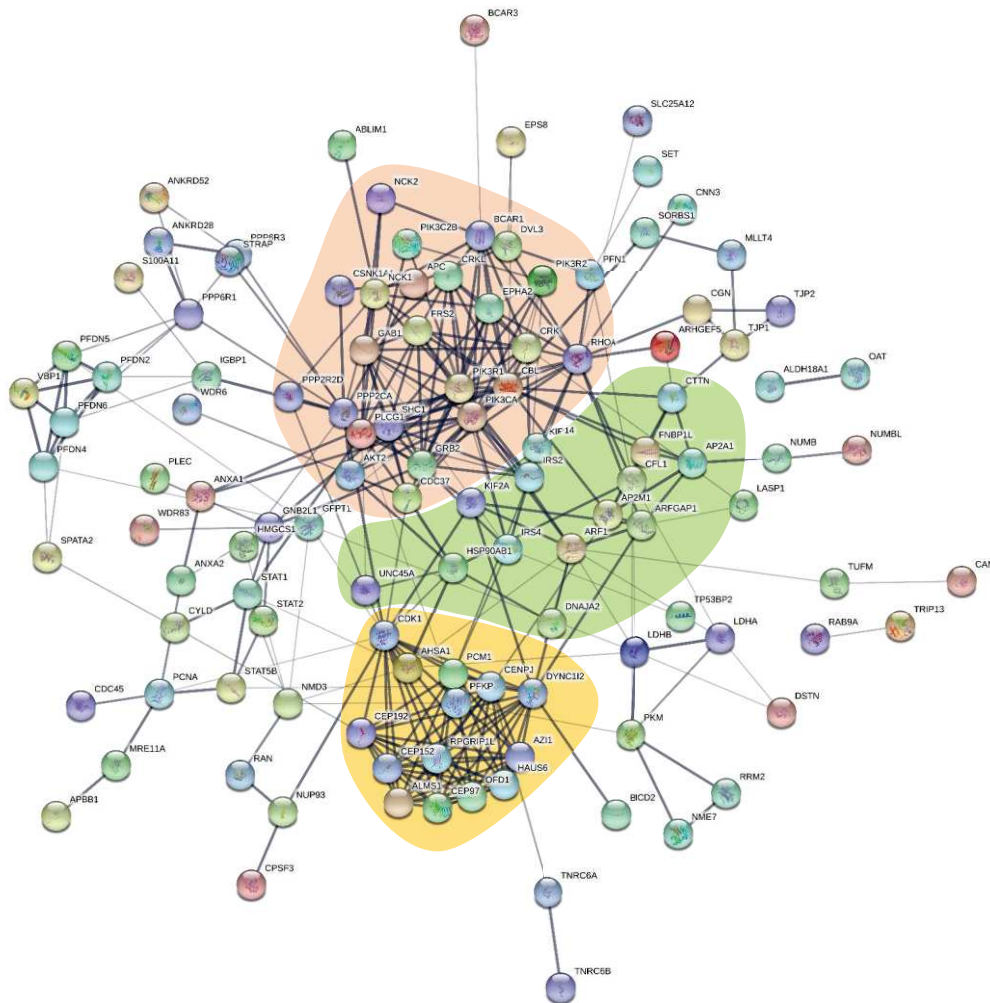


Figure 51. Réseau fonctionnel des interacteurs d'EML4-ALK v3a. Les interacteurs du réseau rouge sont impliqués dans la signalisation cellulaire via les voies ErbB, Rap1 et de récepteur à l'insuline. Les interacteurs du réseau jaune sont impliqués dans la régulation de la mitose. Cette zone comprend notamment des protéines du centrosome. Enfin, les interacteurs du réseau vert connectent les deux premiers réseaux et ont des fonctions moins partagées comme le trafic vésiculaire, la transduction de signaux cellulaires ou la formation des fuseaux mitotiques.

IV.3 Discussion

Analyser le réseau d'interaction d'EML4-ALK est plus complexe qu'analyser celui d'une protéine portant seulement un point de mutation car ici toutes les interactions peuvent être considérées comme pathologiques. Cependant, afin de hiérarchiser ces interactions et en prenant en compte les marqueurs généraux des cancers (Hanahan and Weinberg, 2011), les interacteurs attribués à la voie des récepteurs aux facteurs de croissance ou à l'insuline, et les interacteurs impliqués dans la régulation de la mitose apparaissent les plus pertinents.

Ce qui est intéressant de remarquer est que ces deux fonctions, signalisation et cycle cellulaire, correspondent aux différents domaines fonctionnels présents dans les fusions EML4-ALK. Effectivement, ALK par son domaine kinase interagit avec des protéines de signalisation à domaine SH2 ou SH3, et EML4 intervient lors de la mitose pour stabiliser les microtubules des centrosomes (Pollmann et al., 2006). Comme pour toutes expériences de BioID, nous ne pouvons conclure sur une interaction simultanée ou séquentielle avec ces deux réseaux protéiques, mais ceci suggère toute de même une diversité d'activité cellulaire de la protéine de fusion apportée par les domaines transloqués. Effectivement, identifier des partenaires en lien avec l'organisation des microtubules est tout à fait surprenant car EML4 WT a justement été utilisé comme contrôle pour retirer ces interactions *a priori* spécifiques à EML4 WT. Si ces interactions sont bien spécifiques au domaine d'EML4 transloqué, identifier ces interacteurs seulement avec la protéine fusionnée suggère alors un mécanisme d'interaction du domaine d'EML4 transloqué dépendant de la présence du domaine kinase d'ALK. Lorsque l'interactome de ALK WT sera analysé par MS et afin d'établir une liste encore plus réduite d'interacteur de la protéine de fusion, il sera aussi intéressant de retirer ces partenaires aux 166 identifiés ici comme spécifique à EML4-ALK v3a. De la même façon qu'EML4, cette nouvelle liste permettra d'identifier les interactions avec le domaine kinase seulement observée lorsque celui-ci est fusionné à EML4.

Au cours de ma thèse j'ai participé à l'identification des nouveaux variants de fusion de la protéine ALK dans le cancer du poumon comme CLIP1-ALK et GCC2-ALK (Vendrell et al., 2017b - voir Publication 2). Avec DCTN1-ALK, KIF5B-ALK, TFG-ALK ils constituent des variants de fusion rares dans le cancer du poumon. Au-delà même du partenaire engagé avec ALK, ce qui est intéressant de remarquer est que tous ont une capacité à interagir avec les microtubules, et certains avec le centrosome (DCTN1, KIF5B) (Splinter et al., 2010; Zhapparova et al., 2013). Cette connexion avec le centrosome *via* la protéine fusionnée avec ALK pourrait donc constituer un second versant de la pathogénicité d'EML4-ALK, cette fois-ci plus en lien avec le contrôle du cycle cellulaire. Ceci pourrait d'ailleurs être étendu à d'autres pathologies où l'altération génétique de ALK est plus fréquente comme les lymphomes anaplasiques à grandes cellules où la translocation ALK est observée dans environ 60% des cas, et principalement exprimée en fusion avec NPM1. Cette dernière protéine interagit elle aussi avec les

microtubules et est impliquée dans la duplication du centrosome (Merkel et al., 2011; Naoe et al., 2006). Pour ce type de cancer, l'altération du centrosome a été particulièrement mise en évidence, sans que les causes aient été identifiées (Ventura et al., 2004).

L'efficacité du crizotinib par rapport aux chimiothérapies à base de docetaxel, molécule ciblant les microtubules, a été démontrée et c'est ce qui a permis la mise sur le marché de cette molécule (Shaw et al., 2013). La combinaison du crizotinib avec le docetaxel ou une autre molécule ciblant les microtubules n'a cependant pas été étudiée et pourrait conduire à un effet synergique. Pour un traitement plus ciblé, la protéine CDK1 est la seule cycline qui connecte l'ensemble des protéines du centrosome identifiées ici. Le 8-Hydroxypiperidinemethyl-baicalein (BA-j) est un inhibiteur spécifique de CDK1 récemment mis en évidence (Zhang et al., 2015), et combiné au crizotinib pourrait ouvrir de nouvelles opportunités thérapeutiques dans la prise en charge des patients portants une translocation du gène *ALK*.

Discussion

Les adénocarcinomes pulmonaires sont la première cause de décès en France et dans le monde. Bien que ce type de cancer fasse l'objet d'une recherche intense, aucun traitement ne permet aujourd'hui de véritablement contrôler la maladie pour offrir aux patients une survie globale nettement prolongée. En ciblant les protéines principalement altérées, les thérapies ciblées apportent toutefois un réel bénéfice aux patients en termes de confort de vie et de survie sans progression de la maladie. En dépit des résistances développées, la médecine personnalisée constitue aujourd'hui un paradigme dans la prise en charge thérapeutique des patients de stade avancé, et la recherche continue de nouvelles cibles est un axe majeur de la lutte contre le cancer.

Au cours de ma thèse, j'ai cherché à comprendre quelles pouvaient être les perturbations cellulaires en lien avec l'apparition de mutations activatrices ou de résistances au niveau de différents oncodrivers retrouvés dans le cancer du poumon. J'ai caractérisé par BioID les réseaux d'interactions des protéines les plus fréquemment altérées dans cette maladie et identifié des protéines ou des groupes fonctionnels encore non décrits. Qu'il s'agisse des réseaux d'interactions protéiques des formes sauvages et mutées des protéines EGFR, KRAS, NRAS, HRAS, BRAF ou de la fusion EML4-ALK, tous décrivent des interactions protéine-protéine extrêmement riches, impliquant souvent plusieurs centaines d'interacteurs. Ces réseaux d'interactions sont bien plus importants que ceux obtenus jusqu'à présent par la technique plus conventionnelle d'AP-MS, proposant de nouvelles perspectives quant à la détermination des mécanismes de résistance spécifiques aux ITKs et à l'identification de nouvelles cibles thérapeutiques chez les sujets *KRAS* mutés, par exemple.

La majorité des réseaux que j'ai pu caractériser l'ont été en utilisant le modèle cellulaire HEK293. Bien qu'il s'agisse *a priori* d'un modèle cellulaire non tumoral et non pulmonaire, la comparaison du réseau d'interactions identifié entre HEK293-KRAS (WT et G12C) et BEAS2B-KRAS (WT et G12C) a montré un fort recouvrement des partenaires identifiés suggérant que cette stratégie reste pertinente pour identifier des interacteurs spécifiques d'une mutation considérée. Ces résultats encourageants sont toutefois à pondérer. En effet, alors que nous avons pu mettre en évidence un effet sur la viabilité cellulaire lors de l'extinction de 13 interacteurs spécifiques ou préférentiels à KRAS G12C dans les lignées cellulaires KRAS G12C, celui-ci a également été observé dans des lignées KRAS sauvage. Finalement, le modèle cellulaire BEAS2B pourrait être utilisé pour étudier le rôle des interacteurs associés à la mutation KRAS G12C sur le phénotype cancéreux, ce qui permettrait de relier le phénotype observé à la mutation introduite.

Parmi les explorations possibles dans l'identification de nouvelles cibles thérapeutiques, une attention particulière devrait être portée sur l'étude d'une molécule ciblant une protéine ou une fonction cellulaire dérégulée et commune aux altérations *KRAS*, *EGFR*, *BRAF* ou *ALK*. À titre d'exemple, l'analyse des interacteurs dérégulés pour KRAS, NRAS ou EGFR a montré que plusieurs d'entre eux avait un lien étroit avec le transport vésiculaire et l'endocytose. En lien avec cette observation, il a été

récemment montré que la macropinocytose, une forme d'endocytose particulière, favoriserait la croissance tumorale dans les cellules de pancréas mutées *KRAS* (Bryant et al., 2014; Commisso et al., 2013). Des expériences utilisant des métabolites radioactifs pourraient permettre d'évaluer la capacité d'endocytose, de dégradation et de réutilisation des métabolites de l'environnement en fonction de la présence ou de l'absence des protéines EGFR, KRAS ou NRAS mutées. Un autre exemple peut être donné au sujet de la protéine RAF1. En effet, nous avons montré que la dérégulation de RAF1 est retrouvée dans les lignées présentant des mutants *KRAS*, *NRAS* ou *BRAF*.

L'ensemble de ces observations suggère qu'il pourrait exister des effets dérégulés communs à plusieurs mutations, pourtant généralement exclusives. Ceci ouvrirait alors la porte à des thérapies moins liées à un oncodriver mais englobant un effet plus généralisé ou plus étendu.

Un autre point mis en évidence par nos analyses est l'interaction fonctionnelle des interacteurs dérégulés. En effet, l'analyse des réseaux fonctionnels a montré que plusieurs interacteurs dérégulés étaient reliés entre eux, constituant alors un second réseau fonctionnel, plus fin, d'interacteurs. C'est, par exemple, ce que nous avons observé avec le réseau d'interactions de l'EGFR dans lequel certains interacteurs dérégulés par les mutations primaires de l'EGFR, sont à l'origine du recrutement d'autres interacteurs, expliquant alors les résistances aux thérapies ciblées. Ainsi, CBL et IGF2R, interacteurs de l'EGFR pour les simples mutations (Del.19, L858R et T790M) sont également dérégulés en présence de doubles mutations (Del.19-T790M et L858R-T790M), les faisant apparaître comme des protéines motrices dans l'étude de l'acquisition des résistances aux ITKs.

Enfin, nos résultats permettent d'envisager sereinement la réalisation d'étude BioID chez l'animal (Dingar et al., 2015). Les xénogreffes se développant en trois dimensions, dans un contexte tissulaire et sous l'influence du microenvironnement, elles constituent certainement à ce jour un modèle de caractérisation des réseaux d'interactions protéiques le plus avancée. Réaliser de tels modèles murins pour l'étude des adénocarcinomes pulmonaires, et confronter les résultats à ceux générés dans différentes lignées cellulaires, seraient un pas supplémentaire dans l'identification de réseaux protéiques à même de mettre en lumière les vulnérabilités de ces cancers.

En conclusion, ce travail de thèse a permis d'identifier une série de plusieurs interacteurs spécifiques des isoformes de RAS et des formes mutantes de KRAS, NRAS, EGFR, BRAF et de la fusion EML4-ALK. Parmi la série des interacteurs validés, certains emmargent. Ils doivent maintenant être étudiés par une série d'expériences fonctionnelles afin de déterminer leur rôle potentiel dans les mécanismes de résistance tumorale. Enfin, à court termes, nous envisageons de tester des inhibiteurs et analyser l'effet de ces molécules comme molécules thérapeutiques.

Publications

Publication 1 : Vendrell, J.A., Mau-Them, F.T., Béganton, B., Godreuil, S., Coopman, P., and Solassol, J. (2017). Circulating Cell Free Tumor DNA Detection as a Routine Tool for Lung Cancer Patient Management. *Int. J. Mol. Sci.* 18.



Review

Circulating Cell Free Tumor DNA Detection as a Routine Tool for Lung Cancer Patient Management

Julie A. Vendrell ¹, Frédéric Tran Mau-Them ¹, Benoît Béganton ^{2,3,4,5}, Sylvain Godreuil ⁶, Peter Coopman ^{2,3,4,5} and Jérôme Solassol ^{1,2,3,5,*}

¹ CHU Montpellier, Arnaud de Villeneuve Hospital, Department of Pathology, 34295 Montpellier, France; J-vendrell@chu-montpellier.fr (J.A.V.); f-tran_mau_them@chu-montpellier.fr (F.T.M.-T.)

² IRCM, Institut de Recherche en Cancérologie de Montpellier, 34298 Montpellier, France; benoit.beganton@inserm.fr (B.B.); peter.coopman@inserm.fr (P.C.)

³ INSERM, U1194, 34298 Montpellier, France

⁴ ICM, Institut Régional du Cancer de Montpellier, 34298 Montpellier, France

⁵ Université de Montpellier, 34000 Montpellier, France

⁶ CHU Montpellier, Arnaud de Villeneuve Hospital, Department of Microbiology, 34295 Montpellier, France; S-godreuil@chu-montpellier.fr

* Correspondence: j-solassol@chu-montpellier.fr; Tel.: +33-4-6733-5871; Fax: +33-4-6733-9590

Academic Editor: William Chi-shing Cho

Received: 29 November 2016; Accepted: 18 January 2017; Published: 29 January 2017

Abstract: Circulating tumoral DNA (ctDNA), commonly named “liquid biopsy”, has emerged as a new promising noninvasive tool to detect biomarker in several cancers including lung cancer. Applications involving molecular analysis of ctDNA in lung cancer have increased and encompass diagnosis, response to treatment, acquired resistance and prognosis prediction, while bypassing the problem of tumor heterogeneity. ctDNA may then help perform dynamic genetic surveillance in the era of precision medicine through indirect tumoral genomic information determination. The aims of this review were to examine the recent technical developments that allowed the detection of genetic alterations of ctDNA in lung cancer. Furthermore, we explored clinical applications in patients with lung cancer including treatment efficiency monitoring, acquired therapy resistance mechanisms and prognosis value.

Keywords: circulating DNA; molecular diagnosis; targeted therapies; routine practice; lung cancer

1. Introduction

Lung cancer is the most common cause of cancer death around the world. About 80%–85% of lung cancer cases are non-small-cell lung cancer (NSCLC) patients, the remaining 15%–20% are small-cell lung cancer (SCLC) [1]. NSCLC is divided into three categories called: adenocarcinoma, squamous-cell adenocarcinoma and large cell adenocarcinoma. Among them, adenocarcinoma cases account for around 40% of NSCLC patients. The prognosis for NSCLC is low with a five-year survival rate of less than 20%, and is even worse for SCLC with a five-year survival rate of less than 5% [1].

For a long time, the first-line treatments have been surgery, chemotherapy or radiotherapy. However, the discovery of several oncogenic driver mutations in patients with NSCLC, adenocarcinoma cases in particular, has allowed the development of personalized treatments based on these specific molecular alterations. Therefore, *EGFR* (epidermal growth factor receptor) mutations account for up to 15% of adenocarcinoma and primarily occurred in the tyrosine kinase (TK) domain of the gene. More than 80% of these mutations consist of in-frame deletions in exon 19 and the L858R point mutation in exon 21. Such mutations induced a constitutive activation of EGFR, making it a potential therapeutic target. Thus, *EGFR*-mutated patients can benefit from a specific first-line treatment specifically the TK inhibitors (TKI) that competitively inhibits fixation of

adenosine triphosphate (ATP) in the catalytic binding site of TK domain. Other driver biomarkers in lung cancer (point mutations, rearrangements or amplifications in specific genes including *KRAS*, *NRAS*, *HER2*, *BRAF*, *ALK*, *RET*, and *ROS1*) have also been proposed and some of them might provide additional information for clinical decision-making.

Unfortunately, side effects of personalized treatments have emerged. Among them, the appearance of the T790M mutation located in exon 20 of *EGFR* systematically results in cancer relapse, generally within 1–2 years. The T790M mutation is present in about half of the lung cancer patients with acquired resistance, and is reported to increase the affinity of the receptor to ATP, relative to its affinity to TKIs [2]. Identification of such mutations is required to propose second-line treatment. Recently, third-generation *EGFR* inhibitors, such as osimertinib, mereletinib or rociletinib, have been proposed as relevant therapeutics that could specifically disrupt the growth of *EGFR* T790M-positive tumors and thus increase patient survival [3–5].

2. Tumor Tissue Biopsy Limitations

Molecular characterization of tumors became mandatory, not only for patients to receive the right treatment, but also to follow the evolution of the molecular characteristics and, accordingly, to adapt treatments [6]. Tissue biopsies remain the gold standard to assess molecular alterations. However, this strategy presents several limitations that can impair patient treatment. Indeed, access to tumor tissues is not always optimal. Many patients with NSCLC are diagnosed at an advanced stage of the disease that makes the surgery or the biopsy difficult and even sometimes dangerous. Thus, complications from intrathoracic biopsies have been reported in 17.1% cases in a series of 211 biopsies [7]. In addition, the quality/quantity of the available tumoral material and *EGFR* genotyping failed in approximately 5% of the cases [8]. Finally, the intratumoral heterogeneity of *EGFR* mutation status has been described in several studies (ranges from 13.9% to 27%; [9]) demonstrating that tumor biopsy do not systematically reveal the complete genomic landscape of the whole patient tumoral cell population. Altogether, these issues related to tissue biopsy analysis failure resulted in an unknown *EGFR* status and excluded some patients that could have been eligible to TKI treatment.

Given these limitations, exploring alternative practical, economical and less invasive techniques to monitor the *EGFR* TKI therapy in NSCLC is absolutely needed. Noninvasive approaches, based on samples of plasma or serum, have shown great potential in monitoring the *EGFR* TKI therapy in recent years. Among the different materials derived from liquid biopsies, ctDNA has been successfully applied to detect *EGFR* mutations in NSCLC patients and can give similar molecular information as those given by invasive tumor biopsies [10] (Figure 1). In addition, the dynamic changes in ctDNA *EGFR* mutation status may predict clinical outcome of *EGFR* TKI therapy [11]. In patient drug resistance instances, one alternative to improve early detection rate and overcome the limitation of repeated tissue sampling is to perform genomic analysis using other liquid biopsy markers such as circulating tumor cells (CTCs), circulating RNA, circulating miRNA, platelet markers, etc. Since the use of these different markers for lung cancer management has previously been reported, it will not be discussed here [11–17].

Hereby, we summarized different technical approaches available that have been proposed for the detection of molecular events from ctDNA and considered their possible applications in hospitals and routine laboratories for the management and monitoring of patients with lung cancer.

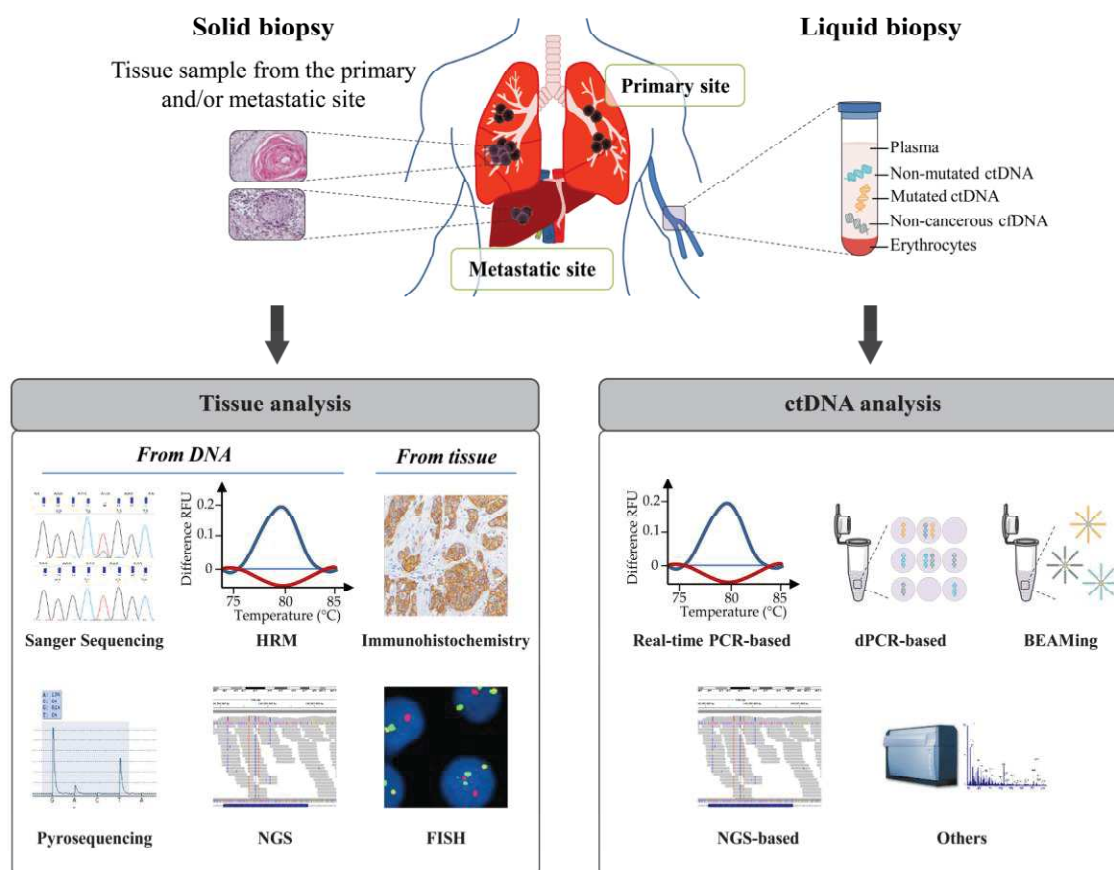


Figure 1. Overview of the available techniques to detect alterations from solid or liquid biopsies. The left side describes the conventional techniques that use tissue sample as starting material, specifically Sanger sequencing, pyrosequencing, High Resolution Melting (HRM), Next-Generation Sequencing (NGS) and Immunohistochemistry. The right side highlights the different methods available for aberration detections from liquid biopsy. They include, in particular, real-time polymerase chain reaction (PCR)-based methods, digital PCR (dPCR), Beads, Emulsion, Amplification, and Magnetics (BEAMing) and NGS-based methods. DNA strand in blue corresponds to non-mutated circulating tumoral DNA (ctDNA), in orange to mutated ctDNA and in grey to non-cancerous cell-free DNA (cfDNA). For each technique, a representation of the principle or the result is given as illustration.

3. The Biology of cfDNA and Circulating Tumoral DNA (ctDNA)

New opportunities arose with the discovery of circulating cell-free DNA (cfDNA) in unaffected individuals [18]. Application includes different fields specifically the non-invasive prenatal diagnosis with the use of cell-free fetal DNA (cffDNA; [19]) and cancer with the use of circulating tumor DNA (ctDNA; [20]).

Origin and mechanisms of cfDNA release in bloodstream are still not completely documented. It is however widely accepted that several conditions such as inflammation, heavy smoking, or pregnancy can induce cfDNA release from cells into the systemic circulation [21–23]. As for patients suffering from heart injury, cfDNA increase over the first 48 h in emergency intensive care unit predicts fatal outcome [24]. The source of ctDNA is also likely multiple and mainly included cell lysis induced by apoptosis and/or necrosis of primary tumors and metastases [25,26] (Figure 1).

cfDNA and ctDNA are highly fragmented with a median size of 170 base pairs or less, which corresponds to the DNA wrapped around a nucleosome plus a linker fragment [27,28]. Several studies have tried to clarify the alleged mechanism of ctDNA (necrosis versus apoptosis) depending on the size of the ctDNA, however, results remain controverted [26,29]. Indeed, Wang et al. [30] and Gao et al. [31] reported that ctDNA is longer than normal cfDNA [30,31]. Paradoxically, Diel et al. [27] and Moulière et al. [29] observed a lower size of ctDNA. Most importantly, ctDNA is

probably composed of short and long fragments with genetic aberrations specifically carried by the shorter ones. This hypothesis has been recently validated in hepatocellular carcinoma patients [32] and in lung cancer patients [22].

4. Technical Approaches for ctDNA Detection and Analysis

Preanalytical conditions may certainly play a crucial role in ctDNA detection. Due to different aspects of ctDNA (high fragmentation, contamination by non-tumoral cfDNA, low amounts and clearance), detection of molecular events from ctDNA materials remains a challenge and requires adapted and ultrasensitive analytical assays. Therefore, specific formaldehyde-free cfDNA collection tubes have recently been commercialized. Such processes not only stabilize but also prevent the release of genomic DNA from nucleated blood cells and reduce the need of immediate plasma preparation. In addition, these tubes allow transport and storage at room temperature and are highly adapted to hospital shipment procedures.

Comparative analysis of ctDNA in plasma and serum have shown that plasma represents the best tool to monitor NSCLC patients in clinical practice [33]. However, ctDNA dilution in patient's cfDNA highly limits liquid biopsy's detection of genetic alteration. Only a few thousands of copy number of cfDNA per milliliter of plasma could be extracted, among which only a small fraction is clinically relevant. Therefore, since genetic alterations that need to be detected from ctDNA are diluted by both the non-tumoral cfDNA and by the non-mutated ctDNA, highly sensitive and specific detection methods are required to provide a relevant ctDNA-based diagnosis. This concern has led to the improvement and the development of several methods of detection such as real-time polymerase chain reaction (PCR), digital PCR (dPCR), Next-Generation Sequencing (NGS), Beads, Emulsion, Amplification, and Magnetics (BEAMing) (Table 1). These methods can be classified into two groups: (i) the targeted approaches that allow detection of specific alterations; and (ii) the untargeted approaches that allow identification of events without a priori, in particular whole-exome sequencing or whole-genome sequencing.

Table 1. Features of techniques used to detect alterations from circulating tumoral DNA (ctDNA).

Techniques	Limit of Detection	Number of Targets	Type of Alteration Detection	Reference
PCR-based approaches				
COLD-PCR	0.10%	1	SNV, indels	[34]
PNA-LNA	0.10%	1	SNV, indels	[35]
Probes improvement	0.01%–0.10%	1	SNV, indels	[36,37]
Digital PCR	0.01%–0.10%	1 to 4	SNV, indels, CNV	[38–40]
BEAMing	0.01%	1 to 20	SNV, indels	[41,42]
NGS-based approaches				
Deep sequencing	0.02%	Panel	SNV, indels	[43]
Base position-error rate correction	0.003%	Panel	SNV, indels	[44]
TAm-Seq	2.00%	Panel	SNV, indels	[45]
CAPP-Seq	0.02%	Panel	SNV, indels, CNV, rearrangements	[46]
cSMART	0.01%	Panel	SNV, indels, rearrangements	[47,48]
Digital sequencing	0.10%	Panel	SNV, indels, CNV, rearrangements	[49]
Bias-Corrected Targeted NGS	0.10%	Panel	SNV, indels, CNV, rearrangements	[50]
SERS-nanotags	0.10%	1 to 3	SNV	[51]
UltraSEEK	0.10%	1 to 7	SNV, indels	[52]

PCR, polymerase chain reaction; COLD-PCR, coamplification at lower denaturation temperature PCR; PNA-LNA, peptide nuclei acid-locked nucleic acid; BEAMing, beads, emulsion, amplification, and magnetics; NGS, next-generation sequencing; TAm-Seq, tagged-amplicon deep sequencing; CAPP-Seq, cancer personalized profiling by deep Sequencing; SERS, surface-enhanced raman spectroscopy; UltraSEEK, high-throughput, multiplexed, ultrasensitive mutation detection; SNV, single nucleotide variation; CNV, copy number variation.

4.1. Real-Time PCR-Based Methods

Allele-specific amplification combined with real-time PCR are commonly used in clinical setting to detect mutations from formalin-fixed paraffin-embedded (FFPE) tumor tissues. Even commercial kits based on the same principle have been developed and are widely used to detect single nucleotide variation (SNV) or small insertion/deletion (indels) (therascreen kit from Qiagen, Hilden, Germany and cobas® from Roche Diagnostics, Meylan, France). However, as they were not fully adapted to the detection of rare genetic events, specific and more sensitive PCR-based methods have been engineered. Notably, custom-designed coamplification at lower denaturation temperature (COLD-PCR) [53,54] or Peptide Nuclei Acid-Locked Nucleic Acid (PNA-LNA) PCR clamp method [35,55,56] have been successfully applied to lung cancer samples. Briefly, COLD-PCR allows the enrichment of low-abundance mutations from a mixture of wild-type, regardless of whether they are known or unknown mutations. Therefore, lower denaturation temperature used during the PCR helps the amplification of heteroduplex mutant/wild-type sequence [34,57]. This PCR method has been further coupled with HRM, pyrosequencing, or Sequencing analysis of the harbored mutations identification [34].

PNA-LNA PCR clamp protocol takes advantage of the increased stability of PNA and LNA probes to highly bind DNA sequences compared to DNA duplex. In this approach, PNA probes firmly bind to DNA to specifically inhibit the amplification of the wild-type allele and thus, increase the specific detection of the mutant allele in real-time PCR cycling. An improved PNA-LNA PCR clamp method has been used to detect *EGFR* mutations in plasma samples [56].

Efforts were also focused on the improvement of allele-specific amplification technique. Indeed, probe-blocking methods have been engineered to block amplification of wild-type templates and thus, to increase detection sensitivity of mutant alleles. Therefore, minor groove binder (MGB) blocker oligonucleotide [37] and modified non-extendable primer blocker (NEPB) [36] have been developed and demonstrated the detection of mutation present at 0.1% in a background of wild-type DNA. Scorpion probes, for which higher sensitivity compared to Taqman probes has been demonstrated, also enable the detection of rare mutations [58–62].

Finally, as there is a tremendous and increased market for the detection of mutation from plasma specimens, new versions of commercial kits have been refined. In particular, the cobas® *EGFR* Mutation Test v2 has been the first liquid biopsy test to be approved by the Food and Drug Administration (FDA) for the detection of *EGFR* mutations.

4.2. Digital PCR (dPCR)

dPCR relies on a real-time PCR, except that DNA templates are partitioned to obtain individual DNA molecule per entities (well, droplet or chamber) that are subsequently amplified by PCR and independently analyzed. Based on the Poisson distribution, it is assumed that small volume reaction compartments must contain 0 or 1 DNA molecules. After end-point PCR quantification of positive compartments, absolute concentration of the target is determined. Several digital PCR platforms are available and based on different process: microfluidic-chamber-based, micro-well chip-based and droplet-based [63]. The most common platforms in clinical laboratories are digital droplet PCR (ddPCR) in which samples are dispersed into thousands of droplets. Droplets containing mutated or non-mutated DNA strand can be discriminated by flow cytometry using fluorescent TaqMan-based probes [63], which allows sensitive detection of mutated ctDNA in a vast background of cfDNA.

Besides high sensitivity estimated at 0.01% to 0.1% [38], dPCR also has a relatively easy workflow, which can be implemented in a clinical setting [64]. Moreover, it has also been applied to detection of copy number variations (CNVs) in the blood sample of lung cancer patients [65]. One disadvantage is that dPCR only screens for known mutations, even if recent works demonstrated the feasibility of multiplex dPCR to detect *EGFR* and *KRAS* mutation in blood samples of cancer patients [40,66].

4.3. Beads, Emulsion, Amplification and Magnetics (BEAMing)

BEAMing is also a targeted approach based on the same principle as the emulsion PCR. Briefly, a first conventional PCR step is performed using primers specific of the targeted sequence that contain known tag sequences. Emulsion PCR of the amplicons is done in presence of tag-coupled magnetic beads that is easily purified. After single-base primer extension or hybridization with fluorescent mutant-specific probes, flow cytometric analysis allows the detection and quantification of mutant versus wild-type alleles [42]. In lung cancer samples, this technique already demonstrated its potency in the detection of *EGFR* activating mutations and the T790M resistance mutation from plasma DNA samples [41,67,68]. Like dPCR methods, BEAMing only allows the screening of known mutations, furthermore it also has a complex workflow and a high cost per sample, making implementation in routine clinical settings less feasible.

4.4. Next-Generation Sequencing (NGS)-Based Approaches

NGS is based on the analysis of millions of short sequences from DNA molecules and their comparison to a reference sequence. Multiple applications have been developed and currently used in oncology, such as targeted sequencing and whole-exome or whole-genome sequencing. Currently, NGS demonstrates a high sensitivity and specificity; nevertheless, random error rate of sequencing platforms is between 0.1% and 1% depending on the platform used [69], making impossible the detection of rare mutations. According to this observation, protocols have been specifically improved and expanded to detect rare mutations in plasma samples. Despite its great advantage to detect multiple somatic alterations simultaneously, NGS remains an expensive and time-consuming technique. Furthermore, extensive data analysis requires highly experienced bioinformaticians to identify with high confidence relevant mutations. Nevertheless, global approaches provide more accurate mutational spectrum of the tumor than targeted analyses and may also allow detection of copy number alterations and large rearrangements [46,49,50,70].

4.4.1. Deep-Sequencing Using Classical NGS Protocols

Since classic NGS experimental protocols are not fully adapted to detect rare mutations, first intents to avoid this problem have been to sequence targeted regions with deep-coverage (>10,000×) [43,71,72]. Another approach was to improve alteration detections using adapted statistical methods. Thus, determination of the base position-error rate (BPER) from control samples allowed detection of true mutations as low as 0.003% and 0.001% for indels after statistical computational [44].

4.4.2. TAM-Seq

Tagged-amplicon deep sequencing (TAM-Seq) has been the first sequencing method adapted to detect rare diagnosis mutations in cfDNA [45]. It is a two-step amplification process that uses the Access Array microfluidic system from Fluidigm. A first preamplification step where all primer sets are used to capture the starting molecules present in the template is processed and is then followed by a second amplification step with limited couple of primers in the microchambers of the Access Array. This process, that is only adapted to point mutation and indels, allows the identification of cancer mutations at allele frequencies as low as 2%, with more than 97% sensitivity and specificity [45].

4.4.3. Cancer Personalized Profiling by Deep Sequencing (CAPP-Seq)

More recently, a capture-based NGS ctDNA detection method, the Cancer Personalized Profiling by deep Sequencing (CAPP-Seq), has been developed [46,73]. The crucial step of this protocol is the design of biotinylated “selectors” that are complementary of previously defined recurrent mutated regions. After hybridization of the “selectors” on the regions of interest and purification, amplification is carried on the reduced library [46]. Diverse classes of mutations present in somatic samples, including single nucleotide variants, indels, rearrangements, and copy number alterations, may thus be detected depending on the designed “selectors”.

In lung cancer, this method could identify mutations in 95% of NSCLC patients with 96% specificity for mutant allele fractions down to approximately 0.02% of tumors [46]. It also has been used to detect resistant mechanism in NSCLC-roceletinib-treated patients such as *EGFR* L798I and *EGFR* C797S mutations [74]. However, CAPP-Seq is still expensive for routine laboratories, with an estimated cost of 200–300 USD [73].

4.4.4. Safe-SeqS

The Safe-Sequencing System (Safe-SeqS) has been proposed as a new tool to increase the sensitivity of massively parallel sequencing system instruments for rare variants identification. A unique identifier (UID) is assigned to each template DNA molecule. Tagged template molecules are then amplified to create UID families and sequences. Variants are considered real if $\geq 95\%$ of the PCR fragments with the same UID contain an identical mutation [75]. The advantage of this approach is to limit base misincorporation errors during sequencing steps or basecalling errors, and to allow rare mutation detection on commercially available sequencers. To our knowledge, Safe-SeqS has been applied to plasma samples of metastatic colorectal cancer [76] and to GIST patients [77], but not to lung cancer patients.

4.4.5. Circulating Single Molecule Amplification and Re-Sequencing Technology (cSMART)

cSMART is another strategy based on a similar approach that can also reduce errors occurring during library preparation or the sequencing phase. Briefly, unique barcodes are added to the end of DNA molecules, they are then circularized by ligation with an oligonucleotide containing a 4 bp degenerate sequence, and are finally amplified using two pairs of reverse PCR targeting primers strategically designed on each side of the hotspot mutation. For detection and quantitation of the targeted mutations, unique single allelic molecules are counted and mutation levels are defined [47]. This method has been used to detect clinically *EGFR* mutations in plasma samples from patients with advanced NSCLC [48]. Despite the critical step of inverse PCR primers design, one advantage of this approach is that gene fusion with unknown partner fusion can be identified.

4.4.6. Digital Sequencing

In digital sequencing experiments, each strand of a double-stranded cfDNA molecule is individually tagged, allowing custom software to compare the two complementary strands and minimize errors occurring during library preparation or the sequencing phase. The digital sequence libraries are amplified and enriched for target genes using capture probes [49]. This process allows detection of SNV, indels, copy number variations (CNVs) and fusion from ctDNA samples [49,70]. As previously described, the authors considered the per-base noise in their bioinformatical process to improve detection of true variant cells.

4.4.7. Bias-Corrected Targeted Next-Generation Sequencing

Recently, Paweletz et al. [50] developed a new method for library preparations that allows the minimizing of the off-target and artifacts. Briefly, multifunctional adaptors that include sequences for single-primer amplification, barcodes for sample identification and tags for sequence identification are used during the tagging step. Small targeting probes (~40 bp) designed to be adjacent to the region of interest are used to capture the targeted regions. Each probe possesses an additional tail sequence that is complementary to a biotinylated pull-down oligonucleotide. After primer extension, captured fragments are amplified with tailed PCR primers and further sequenced. This process has been successfully used for the detection of SNV, CNV and de novo rearrangement detection in 48 ctDNA samples.

4.4.8. Untargeted Sequencing

Whole exome sequencing (WES) and whole genome sequencing (WGS) allow not only the screening of mutations, but also of rearrangements and of copy number variations, providing a more

global genomic profiling of ctDNA aberrations. WES method on cancer patients' plasma has been demonstrated in several studies using hybridization-based exome enrichment technologies [78–81]. Regarding WGS sequencing, Leary and colleagues was the first group to establish genome-wide profiling from plasma samples in completion of the personalized analysis of rearranged ends (PARE) method to allow a better identification of rearranged breakpoints from ctDNA [82]. Another group ran WGS in combination with bisulfite DNA sequencing to simultaneously detect genome-wide hypomethylation and copy number aberrations from hepatocellular carcinoma plasma samples [83,84]. These large-scale methods are clinically relevant as they allowed detection of alterations in an unbiased manner; however, they are still very expensive to be carried out for clinical diagnosis.

4.5. Other Technologies

Recently, alternative technologies using in particular Raman spectrometer or mass-spectrometry have also been development to detect low mutations from ctDNA. Surface-Enhanced Raman Spectroscopy (SERS) nanotags is based on the generation of amplicons by conventional multiplex-PCR with a barcode at the 5'-end that enables the mutation-dependent specific hybridization of SERS-nanotags and a biotin molecule at the 3'-end that allows the specific enrichment of mutated-amplicons. Following a laser excitation, each SERS-nanotag emits a specific signal that enables an easy and direct detection of multiple mutations at the same time using a Raman spectrometer [51].

UltraSEEK (high-throughput, multiplexed, ultrasensitive mutation detection) is a Mass-spectrometry-based technology that has been designed to address the limited multiplexing ability of conventional PCR approaches [52]. Briefly, a multiplex-PCR is first performed to amplify several genes, biotinylated probes, specific of the targeted mutations, are then used to generate biotinylated-mutated-strand. Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) is finally used to detect presence of mutations [52].

Of note, three other approaches have previously been described to detect *EGFR* mutations in plasma DNA from lung cancer patients, including an enzymatic-based technique [85], mass spectrometry genotyping assay [86] and denaturing high performance liquid chromatography [87]. However, these techniques have not been extensively used in routine laboratories.

5. ctDNA and Lung Clinical Applications

5.1. ctDNA at Diagnosis

Since lung cancer is often diagnosed at an advanced stage of the disease, quantification of cfDNA as an early diagnostic tool for lung cancer aroused great interest. Different studies demonstrated that concentration of cfDNA is higher in advanced grade adenocarcinoma patients than in healthy volunteers [88–93]. However, as ctDNA are not detectable in all patients with NSCLC [46,94], the use of cfDNA quantification method is currently limited for lung cancer diagnosis.

One of the most important potential applications for ctDNA in lung cancer diagnosis is the detection of genetic alterations when tumor tissue is not accessible or tissue biopsy DNA extracted is not amplifiable. Many studies have demonstrated whether genetic variations within ctDNA reflects the tumor tissue mutational landscape (Table 2). Interestingly, although specificity is near 100% regardless of the technique used, the sensitivity is usually weaker (Table 2) and may depend on the alterations' type. Indeed, *EGFR* T790M mutation showed a significantly lower detection in plasma compared to other *EGFR* alterations [67,68,95]. However, *EGFR* mutational detection in ctDNA remains a relevant alternative when the diagnostic tissue biopsy is not available [96].

Table 2. Concordance of alteration detections in ctDNA and tissue specimen in lung cancer.

Targeted Genes	Technical Approach		Number of Plasma Samples	Performance			Reference
	Principle	Method		Sensitivity (%)	Specificity (%)	Concordance (%)	
KRAS	PCR-based	COLD-PCR	82	95.7	94.9	95.1	[53]
EGFR	PCR-based	PNA-LNA	30	79.2	100	~80	[56]
EGFR	PCR-based	Therascreen	652	65.7	99.8	94.3	[96]
EGFR	PCR-based	PNA-adapted method	97	78.3	100	ND	[97]
EGFR	PCR-based	Cobas	32	50	69.2	60	[98]
EGFR	PCR-based	Cobas	238	75	96	88	[99]
EGFR del19/L858R T790M		Cobas	110	73.3 63.6	100 98.4	79.8 82.8	[95]
EGFR del19 L858R T790M	PCR-based	Cobas	38	86 90 41	100 100 100	89 97 57	[68]
EGFR del19 L858R T790M	PCR-based	Therascreen	38	82 78 29	100 100 100	87 95 48	[68]
EGFR	PCR-based	PCR-restriction fragment length polymorphism	111	35.6	95.5	71	[100]
KRAS	PCR-based	PCR-restriction fragment length polymorphism	120	77	95	93	[101]
EML4-ALK rearrangement	PCR-based	Taqman probes	32	21	100	66	[102]
KRAS	dPCR	Droplet-based	64	78	100	-	[14]
EGFR	dPCR	Droplet-based	73	-	-	74	[103]
EGFR	dPCR	Droplet-based	46	66.7	100	84.8	[104]
EGFR L858R T790M	dPCR	Droplet-based	38	90 71	100 83	97 74	[68]
EGFR	dPCR	Microfluidic-chamber-based	35	92	100	-	[105]
EGFR	BEAMing	BEAMing	44	72.7	-	73	[41]
EGFR del19 L858R T790M	BEAMing	BEAMing	216	82.3 86.3 70.3	97.5 96.5 69	- - -	[67]

EGFR	BEAMing	BEAMing					
del19				93	100	95	[68]
L858R			38	100	93	95	
T790M				71	67	70	
EGFR, KRAS, BRAF	NGS-based	Deep sequencing	21	100	100	100	[43]
EGFR, KRAS, BRAF, ERBB2, PIK3CA	NGS-based	Deep sequencing	68	58	87	68	[71]
EGFR	NGS-based	Deep sequencing					
del19			288	50.9	98	-	[72]
L858R				51.9	94.1	-	
EGFR	NGS-based	Digital sequencing	50	-	-	97.5	[70]
Panel	NGS-based	Digital sequencing	165	85	99.6	99.3	[49]
EGFR	NGS-based	CAPP-Seq	43	95	100	91	[74]
EGFR, fusion	NGS-based	CAPP-Seq	13	85	96	-	[46]
EGFR	NGS-based	cSMART	61	71.8	70	90.5	[48]
KRAS, EGFR	NGS-based	Capture	31	-	-	71	[94]
EGFR, KRAS, PIK3CA, fusion	NGS-based	Capture	39	68.5	100	78.2	[106]
EGFR, fusion, CNV	NGS-based	Bias-corrected	48	77	100	86	[50]
EGFR	Mass spectrometry	MALDI-TOF	31	80	52.4	61	[86]
EGFR	DHPLC		230	81.8	89.5	87	[87]
EGFR	Meta-analysis		3110	63	95.9	-	[107]

PCR, polymerase chain reaction; COLD-PCR, coamplification at lower denaturation temperature PCR; PNA-LNA, peptide nucleic acid-locked nucleic acid; ND, not done; dPCR, digital PCR; BEAMing, beads, emulsion, amplification, and magnetics; NGS, next-generation sequencing; CAPP-Seq, cancer personalized profiling by deep Sequencing; cSMART, circulating single molecule amplification and re-sequencing technology; MALDI-TOF, matrix-assisted laser desorption/ionization-time of flight; DHPLC, Denaturing high performance liquid chromatography; CNV, copy number variation.

5.2. ctDNA as a Prognostic Marker

ctDNA in lung cancer patients as the new prognostic and predictive tool has been extensively studied and challenged. Indeed, several studies report that high levels of cfDNA result in shorter overall survival (OS) [108,109], whereas other reports show that increased levels of cfDNA are not associated with OS or progression-free survival (PFS) [88,110]. These contrasting results indicate that cfDNA quantification has a limited prognostic value that can result, to some extent, from differences of plasma processing protocols used in the different studies. In contrast, quantification of *EGFR* mutations in cfDNA seems to be more relevant. Patients with high circulating *EGFR* copy number levels have a lower OS and PFS than patients with low *EGFR* copy number levels in plasma [111]. Furthermore, patients with high levels of *EGFR* activating mutations in TKI-naïve plasma sample have longer OS and PFS [103,111]. Regarding the prognostic value of *KRAS* mutation levels in the plasma of lung cancer patients, discordances are reported and were recently reviewed by Garzón and colleagues [112]. Whereas some studies show that patients with detectable *KRAS* mutation have a significantly shorter OS and/or PFS compared to wild-type patients [101,113], no differences between the two groups are reported in a recent study and a meta-analysis [14,114]. Altogether, even if some evidence suggests that *EGFR* status seems to be a more informative prognostic tool than *KRAS* in plasma samples, a reliable cut-off still needs to be determined.

5.3. ctDNA and Lung Cancer Tumor Burden

Another clinical application of cfDNA levels is that they may reflect the total body disease burden and surpass medical imaging for cancer detection. High cfDNA levels are significantly associated with the number of metastatic sites and tumor volume at diagnosis [46,115]. Despite that, Nygaard and colleagues found no correlation between cfDNA levels and tumoral volumetric parameters assessed by positron emission tomography (PET) scans [108], suggesting that cfDNA do not mirror a simple measurement of tumor burden. These discrepancies may primarily be attributed to differences in the methods employed for extraction and quantification.

Currently, PET scans allow the routine radiologic evaluation of treatment response to early detected signs of local recurrences or metastases. However, medical imaging is not always easily accessible, patients are exposed to ionizing radiations, and the metastases need to have reached a significant volume to be detectable. Sozzi et al. [88] first reported a link between an increase in cfDNA levels and further development of metastases or recurrence in the patients. More recently, *EGFR* mutation from plasma samples has successfully been assessed for early evaluation of the TKI treatment efficiency corresponding to the early radiologic response evaluated by chest X-rays [116]. Newman et al. [46] reported similar correlation between ctDNA levels and treatment-related imaging changes. Altogether, these studies emphasize the powerful potential of ctDNA in the follow-up of lung cancer patients in order to evaluate earlier relapse or to identify patients with residual disease. Recently, Thompson et al. [70] demonstrated the feasibility of multiple ctDNA mutation detections for lung cancer patient management using NGS.

5.4. ctDNA in Treatment Efficiency Monitoring

ctDNA also offers the possibility to detect acquired resistance mechanisms, including the second T790M mutation of *EGFR*, amplification of *MET* or *HER2*, and mutations of *PIK3CA* or *BRAF*, for early stage lung cancer patients under first-generation TKI medication [70]. Taniguchi et al. [41] identified the T790M mutation in ctDNA in 43.5% (10/23) of patients who had progressive disease after *EGFR*-TKI treatment. Another study also proved that sequencing of plasma DNA could complement current invasive approaches to identify mutations associated with acquired drug resistance in advanced cancers [79]. In this study, *EGFR* T790M mutation could be detected in plasma during the progression, but not at the initiation of treatment for NSCLC gefitinib-treated patients. More recently, the monitoring of T790M apparition in the ctDNA of the first-generation *EGFR*-TKI treated patients showed an average period of 2.2 months before clinical disease progression [39].

Recently, Oxnard et al. [67] proposed that the T790M ctDNA genotyping should warrant the relevant monitoring of patients treated by osimertinib (AZD9291) prior to undergoing a tumor biopsy.

The specific *EGFR* C797S mutation was also successfully detected in ctDNA in patients who developed resistance to osimertinib [117]. Interestingly, patients who relapsed under rociletinib, a third-generation *EGFR*-TKI, harbored other mechanisms of resistance including increase of *MET* copy number and *EGFR* L798I mutation [74].

6. Conclusions and Perspectives

Although analysis and detection of ctDNA have been asserted many years ago, liquid biopsy has recently emerged as a new potential attractive blood-based biomarker with multiple clinical applications for lung cancer patients including primary molecular diagnosis of tumors, resistant mechanisms monitoring to adapt treatments, and cancer prediction outcomes.

The relatively low sensitivity observed in the different studies reported to date can probably be explained by the lack of consensus in the choice of technical approaches, preferred sample type (serum vs. plasma), storage conditions, detected candidate mutation or suitable detection techniques (Table 2). Therefore, in order to complete the analytic and clinical validations of the sensitivity, specificity and accuracy of such liquid biopsy tests and provide the standardization of all experimental steps, ctDNA-based large-scale studies including internal validation (training and test sets) and external validation should be proposed. Targeted approaches could be taken as references since they certainly have a higher analytic sensitivity than untargeted approaches [118,119].

The use of ctDNA within the scope of clinical trials shows significant benefits and will certainly be more considered in the next years. Ultimately and more specifically the patient will highly benefits from the incorporation of this technology into the standard of care. Whether ctDNA provides a complementary or even an adequate alternative to the gold standard tumor biopsies in the near future remains the subject of many speculations. Current limitations that have been reported in many studies such as reduced sensitivity to detect some mutations in ctDNA compared to tissue biopsies should no longer remain an issue, especially because of the constant improvement in genomic approaches.

Author Contributions: Julie Vendrell, Frédéric Tran Mau-Them, Benoît Béganton and Jérôme Solassol wrote the original paper. Sylvain Godreuil and Peter Coopman reviewed the manuscript and participated in discussions about content. All authors have read and approved the final submitted manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zappa, C.; Mousa, S.A. Non-small cell lung cancer: Current treatment and future advances. *Transl. Lung Cancer Res.* **2016**, *5*, 288–300.
2. Yun, C.H.; Mengwasser, K.E.; Toms, A.V.; Woo, M.S.; Greulich, H.; Wong, K.K.; Meyerson, M.; Eck, M.J. The T790M mutation in *EGFR* kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 2070–2075.
3. Cross, D.A.; Ashton, S.E.; Ghiorghiu, S.; Eberlein, C.; Nebhan, C.A.; Spitzler, P.J.; Orme, J.P.; Finlay, M.R.; Ward, R.A.; Mellor, M.J.; et al. AZD9291, an irreversible *EGFR* TKI, overcomes T790M-mediated resistance to *EGFR* inhibitors in lung cancer. *Cancer Discov.* **2014**, *4*, 1046–1061.
4. Walter, A.O.; Sjin, R.T.; Haringsma, H.J.; Ohashi, K.; Sun, J.; Lee, K.; Dubrovskiy, A.; Labenski, M.; Zhu, Z.; Wang, Z.; et al. Discovery of a mutant-selective covalent inhibitor of *EGFR* that overcomes T790M-mediated resistance in NSCLC. *Cancer Discov.* **2013**, *3*, 1404–1415.
5. Park, K.; Han, J.Y.; Kim, D.W.; Bazhenova, L.A.; Ou, S.H.; Pang, Y.K.; Hin, H.S.; Juan, O.; Son, J.; Janne, P. 190TiP: ELUXA 1: Phase II study of BI 1482694 (HM61713) in patients (pts) with T790M-positive non-small cell lung cancer (NSCLC) after treatment with an epidermal growth factor receptor tyrosine kinase inhibitor (*EGFR* TKI). *J. Thorac. Oncol.* **2016**, *11*, S139, doi:10.1016/S1556-0864(16)30299-4.

6. Tan, D.S.; Yom, S.S.; Tsao, M.S.; Pass, H.I.; Kelly, K.; Peled, N.; Yung, R.C.; Wistuba, I.I.; Yatabe, Y.; Unger, M.; et al. The International Association for the Study of Lung Cancer Consensus Statement on Optimizing Management of EGFR Mutation-Positive Non-Small Cell Lung Cancer: Status in 2016. *J. Thorac. Oncol.* **2016**, *11*, 946–963.
7. Overman, M.J.; Modak, J.; Kopetz, S.; Murthy, R.; Yao, J.C.; Hicks, M.E.; Abbruzzese, J.L.; Tam, A.L. Use of research biopsies in clinical trials: Are risks and benefits adequately discussed? *J. Clin. Oncol.* **2013**, *31*, 17–22.
8. Barlesi, F.; Mazieres, J.; Merlio, J.P.; Debieuvre, D.; Mosser, J.; Lena, H.; Ouafik, L.; Besse, B.; Rouquette, I.; Westeel, V.; et al. Routine molecular profiling of patients with advanced non-small-cell lung cancer: Results of a 1-year nationwide programme of the French Cooperative Thoracic Intergroup (IFCT). *Lancet* **2016**, *387*, 1415–1426.
9. Remon, J.; Majem, M. EGFR mutation heterogeneity and mixed response to EGFR tyrosine kinase inhibitors of non small cell lung cancer: A clue to overcoming resistance. *Transl. Lung Cancer Res.* **2013**, *2*, 445–448.
10. Diaz, L.A., Jr.; Bardelli, A. Liquid biopsies: Genotyping circulating tumor DNA. *J. Clin. Oncol.* **2014**, *32*, 579–586.
11. Sorber, L.; Zwaenepoel, K.; Deschoolmeester, V.; van Schil, P.E.; van Meerbeeck, J.; Lardon, F.; Rolfo, C.; Pauwels, P. Circulating cell-free nucleic acids and platelets as a liquid biopsy in the provision of personalized therapy for lung cancer patients. *Lung Cancer* **2016**, doi:10.1016/j.lungcan.2016.04.026.
12. Gorges, T.M.; Penkalla, N.; Schalk, T.; Joosse, S.A.; Riethdorf, S.; Tucholski, J.; Lucke, K.; Wikman, H.; Jackson, S.; Brychta, N.; et al. Enumeration and Molecular Characterization of Tumor Cells in Lung Cancer Patients Using a Novel In Vivo Device for Capturing Circulating Tumor Cells. *Clin. Cancer Res.* **2016**, *22*, 2197–2206.
13. Guibert, N.; Pradines, A.; Casanova, A.; Farella, M.; Keller, L.; Soria, J.C.; Favre, G.; Mazieres, J. Detection and Monitoring of the *BRAF* Mutation in Circulating Tumor Cells and Circulating Tumor DNA in *BRAF*-Mutated Lung Adenocarcinoma. *J. Thorac. Oncol.* **2016**, *11*, e109–e112.
14. Guibert, N.; Pradines, A.; Farella, M.; Casanova, A.; Gouin, S.; Keller, L.; Favre, G.; Mazieres, J. Monitoring *KRAS* mutations in circulating DNA and tumor cells using digital droplet PCR during treatment of *KRAS*-mutated lung adenocarcinoma. *Lung Cancer* **2016**, *100*, 1–4.
15. Ma, Y.; Xu, P.; Mi, Y.; Wang, W.; Pan, X.; Wu, X.; He, Q.; Liu, H.; Tang, W.; An, H. Plasma MiRNA alterations between NSCLC patients harboring Del19 and L858R EGFR mutations. *Oncotarget* **2016**, *7*, 54965–54972.
16. Park, S.M.; Wong, D.J.; Ooi, C.C.; Kurtz, D.M.; Vermesh, O.; Aalipour, A.; Suh, S.; Pian, K.L.; Chabon, J.J.; Lee, S.H.; et al. Molecular profiling of single circulating tumor cells from lung cancer patients. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E8379–E8386.
17. Zhang, H.; Su, Y.; Xu, F.; Kong, J.; Yu, H.; Qian, B. Circulating microRNAs in relation to EGFR status and survival of lung adenocarcinoma in female non-smokers. *PLoS ONE* **2013**, *8*, e81408.
18. Mandel, P.; Metais, P. Les acides nucléiques du plasma sanguin chez l'homme. *1948*, *142*, 241–243.
19. Lo, Y. M.; Corbetta, N.; Chamberlain, P.F.; Rai, V.; Sargent, I.L.; Redman, C.W.; Wainscoat, J.S. Presence of fetal DNA in maternal plasma and serum. *Lancet* **1997**, *350*, 485–487.
20. Leon, S. Free DNA in the Serum of Cancer Patients and the Effect of Therapy. *Cancer Res.* **1977**, *37*, 646–650.
21. Anker, P.; Lyautey, J.; Lederrey, C.; Stroun, M. Circulating nucleic acids in plasma or serum. *Clin. Chim. Acta* **2001**, *313*, 143–146.
22. Underhill, H.R.; Kitzman, J.O.; Hellwig, S.; Welker, N.C.; Daza, R.; Baker, D.N.; Gligorich, K.M.; Rostomily, R.C.; Bronner, M.P.; Shendure, J. Fragment Length of Circulating Tumor DNA. *PLoS Genet.* **2016**, *12*, e1006162.
23. Wei, Z.; Shah, N.; Deng, C.; Xiao, X.; Zhong, T.; Li, X. Circulating DNA addresses cancer monitoring in non small cell lung cancer patients for detection and capturing the dynamic changes of the disease. *Springerplus* **2016**, *5*, 531, doi:10.1186/s40064-016-2141-5.
24. Xia, D.L.; Zhang, H.; Luo, Q.L.; Zhang, A.F.; Zhu, L.X. Cell-free DNA increase over first 48 hours in emergency intensive care unit predicts fatal outcome in patients with shock. *J. Int. Med. Res.* **2016**, *44*, 1002–1012.
25. Bettegowda, C.; Sausen, M.; Leary, R.J.; Kinde, I.; Wang, Y.; Agrawal, N.; Bartlett, B.R.; Wang, H.; Lubner, B.; Alani, R.M.; et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci. Transl. Med.* **2014**, *6*, 224ra24, doi:10.1126/scitranslmed.3007094.

26. Cheng, F.; Su, L.; Qian, C. Circulating tumor DNA: A promising biomarker in the liquid biopsy of cancer. *Oncotarget* **2016**, *7*, 48832–48841.
27. Diehl, F.; Li, M.; Dressman, D.; He, Y.; Shen, D.; Szabo, S.; Diaz, L.A., Jr.; Goodman, S.N.; David, K.A.; Juhl, H.; et al. Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16368–16373.
28. Lo, Y.M.; Chan, K.C.; Sun, H.; Chen, E.Z.; Jiang, P.; Lun, F.M.; Zheng, Y.W.; Leung, T.Y.; Lau, T.K.; Cantor, C.R.; et al. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Sci. Transl. Med.* **2010**, *2*, 61ra91, doi:10.1126/scitranslmed.3001720.
29. Mouliere, F.; Robert, B.; Arnau Peyrotte, E.; del Rio, M.; Ychou, M.; Molina, F.; Gongora, C.; Thierry, A.R. High fragmentation characterizes tumour-derived circulating DNA. *PLoS ONE* **2011**, *6*, e23418.
30. Wang, B.G.; Huang, H.Y.; Chen, Y.C.; Bristow, R.E.; Kassaei, K.; Cheng, C.C.; Roden, R.; Sokoll, L.J.; Chan, D.W.; Shih Ie, M. Increased plasma DNA integrity in cancer patients. *Cancer Res.* **2003**, *63*, 3966–3968.
31. Gao, Y.J.; He, Y.J.; Yang, Z.L.; Shao, H.Y.; Zuo, Y.; Bai, Y.; Chen, H.; Chen, X.C.; Qin, F.X.; Tan, S.; et al. Increased integrity of circulating cell-free DNA in plasma of patients with acute leukemia. *Clin. Chem. Lab. Med.* **2010**, *48*, 1651–1656.
32. Jiang, P.; Chan, C.W.; Chan, K.C.; Cheng, S.H.; Wong, J.; Wong, V.W.; Wong, G.L.; Chan, S.L.; Mok, T.S.; Chan, H.L.; et al. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E1317–E1325.
33. Molina-Vila, M.A.; de-Las-Casas, C.M.; Bertran-Alamillo, J.; Jordana-Ariza, N.; Gonzalez-Cao, M.; Rosell, R. cfDNA analysis from blood in melanoma. *Ann. Transl. Med.* **2015**, *3*, 309.
34. Li, J.; Wang, L.; Mamon, H.; Kulke, M.H.; Berbeco, R.; Makrigiorgos, G.M. Replacing PCR with COLD-PCR enriches variant DNA sequences and redefines the sensitivity of genetic testing. *Nat. Med.* **2008**, *14*, 579–584.
35. Nagai, Y.; Miyazawa, H.; Tanaka, T.; Udagawa, K.; Kato, M.; Fukuyama, S.; Yokote, A.; Kobayashi, K.; Kanazawa, M.; Hagiwara, K. Genetic heterogeneity of the epidermal growth factor receptor in non-small cell lung cancer cell lines revealed by a rapid and sensitive detection system, the peptide nucleic acid-locked nucleic acid PCR clamp. *Cancer Res.* **2005**, *65*, 7276–7282.
36. Wang, H.; Jiang, J.; Mostert, B.; Sieuwerts, A.; Martens, J.W.; Sleijfer, S.; Foekens, J.A.; Wang, Y. Allele-specific, non-extendable primer blocker PCR (AS-NEPB-PCR) for DNA mutation detection in cancer. *J. Mol. Diagn.* **2013**, *15*, 62–69.
37. Xie, F.; Huang, J.; Qu, S.; Wu, W.; Jiang, J.; Wang, H.; Wang, S.; Liu, Q.; Zhang, S.; Xu, L.; et al. Sensitive detection of trace amounts of KRAS codon 12 mutations by a fast and novel one-step technique. *Clin. Biochem.* **2014**, *47*, 237–242.
38. Zhang, B.O.; Xu, C.W.; Shao, Y.; Wang, H.T.; Wu, Y.F.; Song, Y.Y.; Li, X.B.; Zhang, Z.; Wang, W.J.; Li, L.Q.; et al. Comparison of droplet digital PCR and conventional quantitative PCR for measuring EGFR gene mutation. *Exp. Ther. Med.* **2015**, *9*, 1383–1388.
39. Zheng, D.; Ye, X.; Zhang, M.Z.; Sun, Y.; Wang, J.Y.; Ni, J.; Zhang, H.P.; Zhang, L.; Luo, J.; Zhang, J.; et al. Plasma EGFR T790M ctDNA status is associated with clinical outcome in advanced NSCLC patients with acquired EGFR-TKI resistance. *Sci. Rep.* **2016**, *6*, 20913, doi:10.1038/srep20913.
40. Zonta, E.; Garlan, F.; Pecuchet, N.; Perez-Toralla, K.; Caen, O.; Milbury, C.; Didelot, A.; Fabre, E.; Blons, H.; Laurent-Puig, P.; et al. Multiplex Detection of Rare Mutations by Picoliter Droplet Based Digital PCR: Sensitivity and Specificity Considerations. *PLoS ONE* **2016**, *11*, e0159094.
41. Taniguchi, K.; Uchida, J.; Nishino, K.; Kumagai, T.; Okuyama, T.; Okami, J.; Higashiyama, M.; Kodama, K.; Imamura, F.; Kato, K. Quantitative detection of EGFR mutations in circulating tumor DNA derived from lung adenocarcinomas. *Clin. Cancer Res.* **2011**, *17*, 7808–7815.
42. Diehl, F.; Schmidt, K.; Choti, M.A.; Romans, K.; Goodman, S.; Li, M.; Thornton, K.; Agrawal, N.; Sokoll, L.; Szabo, S.A.; et al. Circulating mutant DNA to assess tumor dynamics. *Nat. Med.* **2008**, *14*, 985–990.
43. Narayan, A.; Carriero, N.J.; Gettinger, S.N.; Kluytenaar, J.; Kozak, K.R.; Yock, T.I.; Muscato, N.E.; Ugarelli, P.; Decker, R.H.; Patel, A.A. Ultrasensitive measurement of hotspot mutations in tumor DNA in blood using error-suppressed multiplexed deep sequencing. *Cancer Res.* **2012**, *72*, 3492–3498.
44. Pecuchet, N.; Rozenholc, Y.; Zonta, E.; Pietraz, D.; Didelot, A.; Combe, P.; Gibault, L.; Bachet, J.B.; Taly, V.; Fabre, E.; et al. Analysis of Base-Position Error Rate of Next-Generation Sequencing to Detect Tumor Mutations in Circulating DNA. *Clin. Chem.* **2016**, *62*, 1492–1503.

45. Forshew, T.; Murtaza, M.; Parkinson, C.; Gale, D.; Tsui, D.W.; Kaper, F.; Dawson, S.J.; Piskorz, A.M.; Jimenez-Linan, M.; Bentley, D.; et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. *Sci. Transl. Med.* **2012**, *4*, 136ra68, doi:10.1126/scitranslmed.3003726.
46. Newman, A.M.; Bratman, S.V.; To, J.; Wynne, J.F.; Eclow, N.C.; Modlin, L.A.; Liu, C.L.; Neal, J.W.; Wakelee, H.A.; Merritt, R.E.; et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat. Med.* **2014**, *20*, 548–554.
47. Lv, W.; Wei, X.; Guo, R.; Liu, Q.; Zheng, Y.; Chang, J.; Bai, T.; Li, H.; Zhang, J.; Song, Z.; et al. Noninvasive prenatal testing for Wilson disease by use of circulating single-molecule amplification and resequencing technology (cSMART). *Clin. Chem.* **2015**, *61*, 172–181.
48. Chai, X.; Ren, P.; Wei, B.; Ma, J.; Mai, L.; Cram, D.S.; Song, Y.; Guo, Y. A comparative study of EGFR oncogenic mutations in matching tissue and plasma samples from patients with advanced non-small cell lung carcinoma. *Clin. Chim. Acta* **2016**, *457*, 106–111.
49. Lanman, R.B.; Mortimer, S.A.; Zill, O.A.; Sebisano, D.; Lopez, R.; Blau, S.; Collisson, E.A.; Divers, S.G.; Hoon, D.S.; Kopetz, E.S.; et al. Analytical and Clinical Validation of a Digital Sequencing Panel for Quantitative, Highly Accurate Evaluation of Cell-Free Circulating Tumor DNA. *PLoS ONE* **2015**, *10*, e0140712.
50. Paweletz, C.P.; Sacher, A.G.; Raymond, C.K.; Alden, R.S.; O'Connell, A.; Mach, S.L.; Kuang, Y.; Gandhi, L.; Kirschmeier, P.; English, J.M.; et al. Bias-Corrected Targeted Next-Generation Sequencing for Rapid, Multiplexed Detection of Actionable Alterations in Cell-Free DNA from Advanced Lung Cancer Patients. *Clin. Cancer Res.* **2016**, *22*, 915–922.
51. Wee, E.J.; Wang, Y.; Tsao, S.C.; Trau, M. Simple, Sensitive and Accurate Multiplex Detection of Clinically Important Melanoma DNA Mutations in Circulating Tumour DNA with SERS Nanotags. *Theranostics* **2016**, *6*, 1506–1513.
52. Mosko, M.J.; Nakorchevsky, A.A.; Flores, E.; Metzler, H.; Ehrich, M.; van den Boom, D.J.; Sherwood, J.L.; Nygren, A.O. Ultrasensitive Detection of Multiplexed Somatic Mutations Using MALDI-TOF Mass Spectrometry. *J. Mol. Diagn.* **2016**, *18*, 23–31.
53. Freidin, M.B.; Freydina, D.V.; Leung, M.; Montero Fernandez, A.; Nicholson, A.G.; Lim, E. Circulating tumor DNA outperforms circulating tumor cells for KRAS mutation detection in thoracic malignancies. *Clin. Chem.* **2015**, *61*, 1299–1304.
54. Castellanos-Rizaldos, E.; Liu, P.; Milbury, C.A.; Guha, M.; Brisci, A.; Cremonesi, L.; Ferrari, M.; Mamon, H.; Makrigiorgos, G.M. Temperature-tolerant COLD-PCR reduces temperature stringency and enables robust mutation enrichment. *Clin. Chem.* **2012**, *58*, 1130–1138.
55. Miyazawa, H.; Tanaka, T.; Nagai, Y.; Matsuo, M.; Sutani, A.; Udagawa, K.; Zhang, J.; Hiram, T.; Murayama, Y.; Koyama, N.; et al. Peptide nucleic acid-locked nucleic acid polymerase chain reaction clamp-based detection test for gefitinib-refractory T790M epidermal growth factor receptor mutation. *Cancer Sci.* **2008**, *99*, 595–600.
56. Watanabe, K.; Fukuhara, T.; Tsukita, Y.; Morita, M.; Suzuki, A.; Tanaka, N.; Terasaki, H.; Nukiwa, T.; Maemondo, M. EGFR Mutation Analysis of Circulating Tumor DNA Using an Improved PNA-LNA PCR Clamp Method. *Can. Respir. J.* **2016**, *2016*, 5297329, doi:10.1155/2016/5297329.
57. Li, J.; Makrigiorgos, G.M. COLD-PCR: A new platform for highly improved mutation detection in cancer and genetic testing. *Biochem. Soc. Trans.* **2009**, *37*, 427–432.
58. Board, R.E.; Wardley, A.M.; Dixon, J.M.; Armstrong, A.C.; Howell, S.; Renshaw, L.; Donald, E.; Greystoke, A.; Ranson, M.; Hughes, A.; et al. Detection of PIK3CA mutations in circulating free DNA in patients with breast cancer. *Breast Cancer Res. Treat.* **2010**, *120*, 461–467.
59. Wang, S.; Han, X.; Hu, X.; Wang, X.; Zhao, L.; Tang, L.; Feng, Y.; Wu, D.; Sun, Y.; Shi, Y. Clinical significance of pretreatment plasma biomarkers in advanced non-small cell lung cancer patients. *Clin. Chim. Acta* **2014**, *430*, 63–70.
60. Goto, K.; Ichinose, Y.; Ohe, Y.; Yamamoto, N.; Negoro, S.; Nishio, K.; Itoh, Y.; Jiang, H.; Duffield, E.; McCormack, R.; et al. Epidermal growth factor receptor mutation status in circulating free DNA in serum: From IPASS, a phase III study of gefitinib or carboplatin/paclitaxel in non-small cell lung cancer. *J. Thorac. Oncol.* **2012**, *7*, 115–121.

61. Kimura, H.; Suminoe, M.; Kasahara, K.; Sone, T.; Araya, T.; Tamori, S.; Koizumi, F.; Nishio, K.; Miyamoto, K.; Fujimura, M.; et al. Evaluation of epidermal growth factor receptor mutation status in serum DNA as a predictor of response to gefitinib (IRESSA). *Br. J. Cancer* **2007**, *97*, 778–784.
62. Kimura, H.; Kasahara, K.; Kawaishi, M.; Kunitoh, H.; Tamura, T.; Holloway, B.; Nishio, K. Detection of epidermal growth factor receptor mutations in serum as a predictor of the response to gefitinib in patients with non-small-cell lung cancer. *Clin. Cancer Res.* **2006**, *12*, 3915–3921.
63. Dong, L.; Meng, Y.; Sui, Z.; Wang, J.; Wu, L.; Fu, B. Comparison of four digital PCR platforms for accurate quantification of DNA copy number of a certified plasmid DNA reference material. *Sci. Rep.* **2015**, *5*, 13174.
64. Malapelle, U.; de Luca, C.; Vigliar, E.; Ambrosio, F.; Rocco, D.; Pisapia, P.; Bellevicine, C.; Troncone, G. EGFR mutation detection on routine cytological smears of non-small cell lung cancer by digital PCR: A validation study. *J. Clin. Pathol.* **2016**, *69*, 454–457.
65. Shoda, K.; Ichikawa, D.; Fujita, Y.; Masuda, K.; Hiramoto, H.; Hamada, J.; Arita, T.; Konishi, H.; Komatsu, S.; Shiozaki, A.; et al. Monitoring the *HER2* copy number status in circulating tumor DNA by droplet digital PCR in patients with gastric cancer. *Gastric Cancer* **2017**, *20*, 126–135.
66. Taly, V.; Pekin, D.; Benhaim, L.; Kotsopoulos, S.K.; le Corre, D.; Li, X.; Atochin, I.; Link, D.R.; Griffiths, A.D.; Pallier, K.; et al. Multiplex picodroplet digital PCR to detect *KRAS* mutations in circulating DNA from the plasma of colorectal cancer patients. *Clin. Chem.* **2013**, *59*, 1722–1731.
67. Oxnard, G.R.; Thress, K.S.; Alden, R.S.; Lawrance, R.; Paweletz, C.P.; Cantarini, M.; Yang, J.C.; Barrett, J.C.; Janne, P.A. Association Between Plasma Genotyping and Outcomes of Treatment with Osimertinib (AZD9291) in Advanced Non-Small-Cell Lung Cancer. *J. Clin. Oncol.* **2016**, *34*, 3375–3382.
68. Thress, K.S.; Brant, R.; Carr, T.H.; Dearden, S.; Jenkins, S.; Brown, H.; Hammett, T.; Cantarini, M.; Barrett, J.C. EGFR mutation detection in ctDNA from NSCLC patient plasma: A cross-platform comparison of leading technologies to support the clinical development of AZD9291. *Lung Cancer* **2015**, *90*, 509–515.
69. Quail, M.A.; Smith, M.; Coupland, P.; Otto, T.D.; Harris, S.R.; Connor, T.R.; Bertoni, A.; Swerdlow, H.P.; Gu, Y. A tale of three next generation sequencing platforms: Comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genom.* **2012**, *13*, 341, doi:10.1186/1471-2164-13-341.
70. Thompson, J.C.; Yee, S.S.; Troxel, A.B.; Savitch, S.L.; Fan, R.; Balli, D.; Lieberman, D.B.; Morrisette, J.D.; Evans, T.L.; Bauml, J.; et al. Detection of Therapeutically Targetable Driver and Resistance Mutations in Lung Cancer Patients by Next-Generation Sequencing of Cell-Free Circulating Tumor DNA. *Clin. Cancer Res.* **2016**, *22*, 5772–5782.
71. Couraud, S.; Vaca-Paniagua, F.; Villar, S.; Oliver, J.; Schuster, T.; Blanche, H.; Girard, N.; Tredaniel, J.; Guilleminault, L.; Gervais, R.; et al. Noninvasive diagnosis of actionable mutations by deep sequencing of circulating free DNA in lung cancer from never-smokers: A proof-of-concept study from BioCAST/IFCT-1002. *Clin. Cancer Res.* **2014**, *20*, 4613–4624.
72. Uchida, J.; Kato, K.; Kukita, Y.; Kumagai, T.; Nishino, K.; Daga, H.; Nagatomo, I.; Inoue, T.; Kimura, M.; Oba, S.; et al. Diagnostic Accuracy of Noninvasive Genotyping of EGFR in Lung Cancer Patients by Deep Sequencing of Plasma Cell-Free DNA. *Clin. Chem.* **2015**, *61*, 1191–1196.
73. Bratman, S.V.; Newman, A.M.; Alizadeh, A.A.; Diehn, M. Potential clinical utility of ultrasensitive circulating tumor DNA detection with CAPP-Seq. *Expert Rev. Mol. Diagn.* **2015**, *15*, 715–719.
74. Chabon, J.J.; Simmons, A.D.; Lovejoy, A.F.; Esfahani, M.S.; Newman, A.M.; Haringsma, H.J.; Kurtz, D.M.; Stehr, H.; Scherer, F.; Karlovich, C.A.; et al. Circulating tumour DNA profiling reveals heterogeneity of EGFR inhibitor resistance mechanisms in lung cancer patients. *Nat. Commun.* **2016**, *7*, 11815, doi:10.1038/ncomms11815.
75. Kinde, I.; Wu, J.; Papadopoulos, N.; Kinzler, K.W.; Vogelstein, B. Detection and quantification of rare mutations with massively parallel sequencing. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9530–9535.
76. Tie, J.; Kinde, I.; Wang, Y.; Wong, H.L.; Roebert, J.; Christie, M.; Tacey, M.; Wong, R.; Singh, M.; Karapetis, C.S.; et al. Circulating tumor DNA as an early marker of therapeutic response in patients with metastatic colorectal cancer. *Ann. Oncol.* **2015**, *26*, 1715–1722.
77. Fredebohm, J.; Mehnert, D.H.; Lober, A.K.; Holtrup, F.; van Rahden, V.; Angenendt, P.; Diehl, F. Detection and Quantification of KIT Mutations in ctDNA by Plasma Safe-SeqS. *Adv. Exp. Med. Biol.* **2016**, *924*, 187–189.
78. Dietz, S.; Schirmer, U.; Merce, C.; von Bubnoff, N.; Dahl, E.; Meister, M.; Muley, T.; Thomas, M.; Sultmann, H. Low Input Whole-Exome Sequencing to Determine the Representation of the Tumor Exome in Circulating DNA of Non-Small Cell Lung Cancer Patients. *PLoS ONE* **2016**, *11*, e0161012.

79. Murtaza, M.; Dawson, S.J.; Tsui, D.W.; Gale, D.; Forshew, T.; Piskorz, A.M.; Parkinson, C.; Chin, S.F.; Kingsbury, Z.; Wong, A.S.; et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* **2013**, *497*, 108–112.
80. Butler, T.M.; Johnson-Camacho, K.; Peto, M.; Wang, N.J.; Macey, T.A.; Korkola, J.E.; Koppie, T.M.; Corless, C.L.; Gray, J.W.; Spellman, P.T. Exome Sequencing of Cell-Free DNA from Metastatic Cancer Patients Identifies Clinically Actionable Mutations Distinct from Primary Disease. *PLoS ONE* **2015**, *10*, e0136407.
81. Klevebring, D.; Neiman, M.; Sundling, S.; Eriksson, L.; Darai Ramqvist, E.; Celebioglu, F.; Czene, K.; Hall, P.; Egevad, L.; Gronberg, H.; et al. Evaluation of exome sequencing to estimate tumor burden in plasma. *PLoS ONE* **2014**, *9*, e104417.
82. Leary, R.J.; Sausen, M.; Kinde, I.; Papadopoulos, N.; Carpten, J.D.; Craig, D.; O'Shaughnessy, J.; Kinzler, K.W.; Parmigiani, G.; Vogelstein, B.; et al. Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing. *Sci. Transl. Med.* **2012**, *4*, 162ra154, doi:10.1126/scitranslmed.3004742.
83. Chan, K.C.; Jiang, P.; Chan, C.W.; Sun, K.; Wong, J.; Hui, E.P.; Chan, S.L.; Chan, W.C.; Hui, D.S.; Ng, S.S.; et al. Noninvasive detection of cancer-associated genome-wide hypomethylation and copy number aberrations by plasma DNA bisulfite sequencing. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 18761–18768.
84. Chan, K.C.; Jiang, P.; Zheng, Y.W.; Liao, G.J.; Sun, H.; Wong, J.; Siu, S.S.; Chan, W.C.; Chan, S.L.; Chan, A.T.; et al. Cancer genome scanning in plasma: Detection of tumor-associated copy number aberrations, single-nucleotide variants, and tumoral heterogeneity by massively parallel sequencing. *Clin. Chem.* **2013**, *59*, 211–224.
85. Kuang, Y.; Rogers, A.; Yeap, B.Y.; Wang, L.; Makrigiorgos, M.; Vetrand, K.; Thiede, S.; Distel, R.J.; Janne, P.A. Noninvasive detection of EGFR T790M in gefitinib or erlotinib resistant non-small cell lung cancer. *Clin. Cancer Res.* **2009**, *15*, 2630–2636.
86. Brevet, M.; Johnson, M.L.; Azzoli, C.G.; Ladanyi, M. Detection of EGFR mutations in plasma DNA from lung cancer patients by mass spectrometry genotyping is predictive of tumor EGFR status and response to EGFR inhibitors. *Lung Cancer* **2011**, *73*, 96–102.
87. Bai, H.; Mao, L.; Wang, H.S.; Zhao, J.; Yang, L.; An, T.T.; Wang, X.; Duan, C.J.; Wu, N.M.; Guo, Z.Q.; et al. Epidermal growth factor receptor mutations in plasma DNA samples predict tumor response in Chinese patients with stages IIIB to IV non-small-cell lung cancer. *J. Clin. Oncol.* **2009**, *27*, 2653–2659.
88. Sozzi, G.; Conte, D.; Mariani, L.; Lo Vullo, S.; Roz, L.; Lombardo, C.; Pierotti, M.A.; Tavecchio, L. Analysis of circulating tumor DNA in plasma at diagnosis and during follow-up of lung cancer patients. *Cancer Res.* **2001**, *61*, 4675–4678.
89. Sozzi, G.; Conte, D.; Leon, M.; Ciricione, R.; Roz, L.; Ratcliffe, C.; Roz, E.; Cirenei, N.; Bellomi, M.; Pelosi, G.; et al. Quantification of free circulating DNA as a diagnostic marker in lung cancer. *J. Clin. Oncol.* **2003**, *21*, 3902–3908.
90. Paci, M.; Maramotti, S.; Bellesia, E.; Formisano, D.; Albertazzi, L.; Ricchetti, T.; Ferrari, G.; Annessi, V.; Lasagni, D.; Carbonelli, C.; et al. Circulating plasma DNA as diagnostic biomarker in non-small cell lung cancer. *Lung Cancer* **2009**, *64*, 92–97.
91. Yoon, K.A.; Park, S.; Lee, S.H.; Kim, J.H.; Lee, J.S. Comparison of circulating plasma DNA levels between lung cancer patients and healthy controls. *J. Mol. Diagn.* **2009**, *11*, 182–185.
92. Szpechcinski, A.; Dancewicz, M.; Kopinski, P.; Kowalewski, J.; Chorostowska-Wynimko, J. Real-time PCR quantification of plasma DNA in non-small cell lung cancer patients and healthy controls. *Eur. J. Med. Res.* **2009**, *14*, 237–240.
93. Ulivi, P.; Silvestrini, R. Role of quantitative and qualitative characteristics of free circulating DNA in the management of patients with non-small cell lung cancer. *Cell Oncol.* **2013**, *36*, 439–448.
94. Villafior, V.; Won, B.; Nagy, R.; Banks, K.; Lanman, R.B.; Talasaz, A.; Salgia, R. Biopsy-free circulating tumor DNA assay identifies actionable mutations in lung cancer. *Oncotarget* **2016**, *7*, 66880–66891.
95. Karlovich, C.; Goldman, J.W.; Sun, J.M.; Mann, E.; Sequist, L.V.; Konopa, K.; Wen, W.; Angenendt, P.; Horn, L.; Spigel, D.; et al. Assessment of EGFR Mutation Status in Matched Plasma and Tumor Tissue of NSCLC Patients from a Phase I Study of Rociletinib (CO-1686). *Clin. Cancer Res.* **2016**, *22*, 2386–2395.
96. Douillard, J.Y.; Ostoros, G.; Cobo, M.; Ciuleanu, T.; McCormack, R.; Webster, A.; Milenkova, T. First-line gefitinib in Caucasian EGFR mutation-positive NSCLC patients: A phase-IV, open-label, single-arm study. *Br. J. Cancer* **2014**, *110*, 55–62.

97. Karachaliou, N.; de las Casas, C.M.; Queralt, C.; de Aguirre, I.; Melloni, B.; Cardenal, F.; Garcia-Gomez, R.; Massuti, B.; Sanchez, J.M.; Porta, R.; et al. Association of EGFR L858R Mutation in Circulating Free DNA With Survival in the EURTAC Trial. *JAMA Oncol.* **2015**, *1*, 149–157.
98. Sundaresan, T.K.; Sequist, L.V.; Heymach, J.V.; Riely, G.J.; Janne, P.A.; Koch, W.H.; Sullivan, J.P.; Fox, D.B.; Maher, R.; Muzikansky, A.; et al. Detection of T790M, the Acquired Resistance EGFR Mutation, by Tumor Biopsy versus Noninvasive Blood-Based Analyses. *Clin. Cancer Res.* **2016**, *22*, 1103–1110.
99. Mok, T.; Wu, Y.L.; Lee, J.S.; Yu, C.J.; Sriuranpong, V.; Sandoval-Tan, J.; Ladrera, G.; Thongprasert, S.; Srimuninnimit, V.; Liao, M.; et al. Detection and Dynamic Changes of EGFR Mutations from Circulating Tumor DNA as a Predictor of Survival Outcomes in NSCLC Patients Treated with First-line Intercalated Erlotinib and Chemotherapy. *Clin. Cancer Res.* **2015**, *21*, 3196–3203.
100. Zhao, X.; Han, R.B.; Zhao, J.; Wang, J.; Yang, F.; Zhong, W.; Zhang, L.; Li, L.Y.; Wang, M.Z. Comparison of epidermal growth factor receptor mutation statuses in tissue and plasma in stage I–IV non-small cell lung cancer patients. *Respiration* **2013**, *85*, 119–125.
101. Wang, S.; An, T.; Wang, J.; Zhao, J.; Wang, Z.; Zhuo, M.; Bai, H.; Yang, L.; Zhang, Y.; Wang, X.; et al. Potential clinical significance of a plasma-based KRAS mutation analysis in patients with advanced non-small cell lung cancer. *Clin. Cancer Res.* **2010**, *16*, 1324–1330.
102. Nilsson, R.J.; Karachaliou, N.; Berenguer, J.; Gimenez-Capitan, A.; Schellen, P.; Teixido, C.; Tannous, J.; Kuiper, J.L.; Drees, E.; Grabowska, M.; et al. Rearranged EML4-ALK fusion transcripts sequester in circulating blood platelets and enable blood-based crizotinib response monitoring in non-small-cell lung cancer. *Oncotarget* **2016**, *7*, 1066–1075.
103. Yang, X.; Zhuo, M.; Ye, X.; Bai, H.; Wang, Z.; Sun, Y.; Zhao, J.; An, T.; Duan, J.; Wu, M.; et al. Quantification of mutant alleles in circulating tumor DNA can predict survival in lung cancer. *Oncotarget* **2016**, *7*, 20810–20824.
104. Oxnard, G.R.; Paweletz, C.P.; Kuang, Y.; Mach, S.L.; O’Connell, A.; Messineo, M.M.; Luke, J.J.; Butaney, M.; Kirschmeier, P.; Jackman, D.M.; et al. Noninvasive detection of response and resistance in EGFR-mutant lung cancer using quantitative next-generation genotyping of cell-free plasma DNA. *Clin. Cancer Res.* **2014**, *20*, 1698–1705.
105. Yung, T.K.; Chan, K.C.; Mok, T.S.; Tong, J.; To, K.F.; Lo, Y.M. Single-molecule detection of epidermal growth factor receptor mutations in plasma by microfluidics digital PCR in non-small cell lung cancer patients. *Clin. Cancer Res.* **2009**, *15*, 2076–2084.
106. Yao, Y.; Liu, J.; Li, L.; Yuan, Y.; Nan, K.; Wu, X.; Zhang, Z.; Wu, Y.; Li, X.; Zhu, J.; et al. Detection of circulating tumor DNA in patients with advanced non-small cell lung cancer. *Oncotarget* **2016**, doi:10.18632/oncotarget.12883.
107. Qiu, M.; Wang, J.; Xu, Y.; Ding, X.; Li, M.; Jiang, F.; Xu, L.; Yin, R. Circulating tumor DNA is effective for the detection of EGFR mutation in non-small cell lung cancer: A meta-analysis. *Cancer Epidemiol. Biomark. Prev.* **2015**, *24*, 206–212.
108. Nygaard, A.D.; Holdgaard, P.C.; Spindler, K.L.; Pallisgaard, N.; Jakobsen, A. The correlation between cell-free DNA and tumour burden was estimated by PET/CT in patients with advanced NSCLC. *Br. J. Cancer* **2014**, *110*, 363–368.
109. Catarino, R.; Coelho, A.; Araujo, A.; Gomes, M.; Nogueira, A.; Lopes, C.; Medeiros, R. Circulating DNA: Diagnostic tool and predictive marker for overall survival of NSCLC patients. *PLoS ONE* **2012**, *7*, e38559.
110. Herrera, L.J.; Raja, S.; Gooding, W.E.; El-Hefnawy, T.; Kelly, L.; Luketich, J.D.; Godfrey, T.E. Quantitative analysis of circulating plasma DNA as a tumor marker in thoracic malignancies. *Clin. Chem.* **2005**, *51*, 113–118.
111. Alegre, E.; Fusco, J.P.; Restituto, P.; Salas-Benito, D.; Rodriguez-Ruiz, M.E.; Andueza, M.P.; Pajares, M.J.; Patino-Garcia, A.; Pio, R.; Lozano, M.D.; et al. Total and mutated EGFR quantification in cell-free DNA from non-small cell lung cancer patients detects tumor heterogeneity and presents prognostic value. *Tumour Biol.* **2016**, *37*, 13687–13694.
112. Garzon, M.; Villatoro, S.; Teixido, C.; Mayo, C.; Martinez, A.; de Los Llanos Gil, M.; Viteri, S.; Morales-Espinosa, D.; Rosell, R. KRAS mutations in the circulating free DNA (cfDNA) of non-small cell lung cancer (NSCLC) patients. *Transl. Lung Cancer Res.* **2016**, *5*, 511–516.
113. Nygaard, A.D.; Garm Spindler, K.L.; Pallisgaard, N.; Andersen, R.F.; Jakobsen, A. The prognostic value of KRAS mutated plasma DNA in advanced non-small cell lung cancer. *Lung Cancer* **2013**, *79*, 312–317.

114. Ai, B.; Liu, H.; Huang, Y.; Peng, P. Circulating cell-free DNA as a prognostic and predictive biomarker in non-small cell lung cancer. *Oncotarget* **2016**, *7*, 44583–44595.
115. Yanagita, M.; Redig, A.J.; Paweletz, C.P.; Dahlberg, S.E.; O'Connell, A.; Feeney, N.; Taibi, M.; Boucher, D.; Oxnard, G.R.; Johnson, B.E.; et al. A Prospective Evaluation of Circulating Tumor Cells and Cell-Free DNA in EGFR-Mutant Non-Small Cell Lung Cancer Patients Treated with Erlotinib on a Phase II Trial. *Clin. Cancer Res.* **2016**, *22*, 6010–6020.
116. Imamura, F.; Uchida, J.; Kukita, Y.; Kumagai, T.; Nishino, K.; Inoue, T.; Kimura, M.; Oba, S.; Kato, K. Monitoring of treatment responses and clonal evolution of tumor cells by circulating tumor DNA of heterogeneous mutant EGFR genes in lung cancer. *Lung Cancer* **2016**, *94*, 68–73.
117. Thress, K.S.; Paweletz, C.P.; Felip, E.; Cho, B.C.; Stetson, D.; Dougherty, B.; Lai, Z.; Markovets, A.; Vivancos, A.; Kuang, Y.; et al. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. *Nat. Med.* **2015**, *21*, 560–562.
118. Heitzer, E.; Ulz, P.; Geigl, J.B.; Speicher, M.R. Non-invasive detection of genome-wide somatic copy number alterations by liquid biopsies. *Mol. Oncol.* **2016**, *10*, 494–502.
119. Sholl, L.M.; Aisner, D.L.; Allen, T.C.; Beasley, M.B.; Cagle, P.T.; Capelozzi, V.L.; Dacic, S.; Hariri, L.P.; Kerr, K.M.; Lantuejoul, S.; et al. Liquid Biopsy in Lung Cancer: A Perspective From Members of the Pulmonary Pathology Society. *Arch. Pathol. Lab. Med.* **2016**, *140*, 825–829.



© 2017 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Publication 2 : Vendrell, J.A., Taviaux, S., Béganton, B., Godreuil, S., Audran, P., Grand, D., Clermont, E., Serre, I., Szablewski, V., Coopman, P., et al. (2017). Detection of known and novel ALK fusion transcripts in lung cancer patients using next-generation sequencing approaches. *Sci. Rep.* 7.

SCIENTIFIC REPORTS



OPEN

Detection of known and novel *ALK* fusion transcripts in lung cancer patients using next-generation sequencing approaches

Julie A. Vendrell¹, Sylvie Taviaux¹, Benoît Béganton², Sylvain Godreuil³, Patricia Audran⁴, David Grand⁵, Estelle Clermont⁵, Isabelle Serre¹, Vanessa Szablewski¹, Peter Coopman², Julien Mazières⁶, Valérie Costes¹, Jean-Louis Pujol⁷, Pierre Brousset^{5,8}, Isabelle Rouquette⁵ & Jérôme Solassol^{1,2}

Rearrangements of the anaplastic lymphoma kinase (*ALK*) gene in non-small cell lung cancer (NSCLC) represent a novel molecular target in a small subset of tumors. Although *ALK* rearrangements are usually assessed by immunohistochemistry (IHC) and fluorescence *in situ* hybridization (FISH), molecular approaches have recently emerged as relevant alternatives in routine laboratories. Here, we evaluated the use of two different amplicon-based next-generation sequencing (NGS) methods (AmpliSeq and Archer[®] FusionPlex[®]) to detect *ALK* rearrangements, and compared these with IHC and FISH. A total of 1128 NSCLC specimens were screened using conventional analyses, and a subset of 37 (15 *ALK*-positive, and 22 *ALK*-negative) samples were selected for NGS assays. Although AmpliSeq correctly detected 25/37 (67.6%) samples, 1/37 (2.7%) and 11/37 (29.7%) specimens were discordant and uncertain, respectively, requiring further validation. In contrast, Archer[®] FusionPlex[®] accurately classified all samples and allowed the correct identification of one rare *DCTN1-ALK* fusion, one novel *CLIP1-ALK* fusion, and one novel *GCC2-ALK* transcript. Of particular interest, two out of three patients harboring these singular rearrangements were treated with and sensitive to crizotinib. These data show that Archer[®] FusionPlex[®] may provide an effective and accurate alternative to FISH testing for the detection of known and novel *ALK* rearrangements in clinical diagnostic settings.

In the past decade, the outcomes of selected subgroups of patients with non-small cell lung cancer (NSCLC) have improved considerably with the emergence of targeted therapies for management of the disease¹. Comprehensive molecular profiling of lung adenocarcinoma has revealed a number of actionable driver alterations that are potential targets for inhibition in approximately 60% of this subtype of lung cancer^{2–4}. Among the alterations, rearrangement of the anaplastic lymphoma kinase (*ALK*) gene and echinoderm microtubule-associated protein-like 4 (*EML4*) occurs in approximately 5% of lung adenocarcinomas, representing the most frequent rearrangements. Other *ALK* fusion partners have been reported, such as *KIF5B*, and *TFG*⁵. All identified *ALK* rearrangements harbor the 5' end of the partner (including the promoter and an oligomerization domain, which is mainly a coiled-coil domain) fused to the entire *ALK* kinase domain, and lead to constitutive ligand-independent kinase activation. Since *ALK* tyrosine kinase activity is necessary for its transforming activity and oncogenicity, several *ALK* kinase

¹CHU Montpellier, Arnaud de Villeneuve Hospital, Department of Pathology, Montpellier, Université de Montpellier, Montpellier, France. ²Institut de Recherche en Cancérologie de Montpellier (IRCM), INSERM U1194, Université de Montpellier, Institut du Cancer de Montpellier (ICM), Montpellier, France. ³CHU Montpellier, Arnaud de Villeneuve Hospital, Department of Bacteriology, Université de Montpellier, Montpellier, France. ⁴Institut du Cancer de Montpellier (ICM), Department of Biopathology, Montpellier, France. ⁵Department of Pathology, Institut Universitaire du Cancer Toulouse Oncopole, CHU de Toulouse, Toulouse, France. ⁶Thoracic Oncology Department, Larrey Hospital, University Hospital of Toulouse, Toulouse, France. ⁷CHU Montpellier, Arnaud de Villeneuve Hospital, Department of Thoracic Oncology, Université de Montpellier, Montpellier, France. ⁸Laboratoire d'excellence Labex TOUCAN, Toulouse, France. Correspondence and requests for materials should be addressed to J.S. (email: j-solassol@chu-montpellier.fr)

inhibitors have been identified and successfully validated, first in preclinical models *in vitro* and *in vivo*, and then in clinical studies. The US Food and Drugs Administration (FDA) has therefore approved the use of some small molecules in advanced *ALK*-rearranged NSCLC patients. Crizotinib, a well-tolerated first generation *ALK* inhibitor^{3,6}, has been shown to be superior to standard chemotherapy both as a first- and second-line treatment^{4,7}, while second generation *ALK* inhibitors, such as alectinib and ceritinib, are effective not only in crizotinib-naïve patients, but also in patients with acquired resistance to crizotinib^{1,8–10}.

Fluorescence *in situ* hybridization (FISH) is currently acknowledged as the “gold standard” for detection of *ALK* rearrangements. The Vysis LSI *ALK* Break Apart FISH Probe Kit has been approved by the FDA as a companion diagnostic test for administration of *ALK* inhibitors in lung cancer patients. The immunohistochemical (IHC) method, which can detect *ALK* protein expression independently of the underlying mechanism mediating its overexpression, is used as a pre-screening test, alongside FISH, to determine *ALK* status in formalin-fixed paraffin embedded (FFPE) tissue specimens. However, even though IHC is widely implemented in pathology laboratories, easy-to-use, and automatically performed, its interpretation remains difficult to standardize and time-consuming. In addition, FISH is expensive, labor intensive, requires expert pathology assessment, and is not amenable to multiplexing.

It has been recognized that the development of molecular approaches strengthens the accuracy of *ALK* fusion diagnosis, by resolving discordant or borderline cases^{11–13}. Several RNA-based methods, including the nCounter assay (NanoString Technologies), reverse transcription-polymerase chain reaction (RT-PCR), multiplex RT-PCR followed by capillary electrophoresis, and RT-quantitative PCR (RT-qPCR) have demonstrated their ability^{14–21}. However, some limitations prevent their full implementation in the clinical setting. They easily highlight already known fusions, but may misdiagnose new variants and fusion partners due to the low precision of the 3'/5' imbalance value. In addition, the multiplex capabilities of some of the techniques are limited. In this context, next-generation sequencing (NGS) amplicon-based approaches have been assessed for the detection of *ALK* fusions in NSCLC patients^{22–26}. Two main molecular amplicon-based NGS approaches emerged, but have not been compared to date.

Here, we evaluated two different amplicon-based NGS methods (Ampliseq and Archer® FusionPlex®) for the detection of *ALK* fusions in order to determine the most relevant approach available for routine clinical practice in pathology laboratories. Among a set of 1128 well-characterized FFPE NSCLC specimens, 10 and 13 samples with or without *ALK* fusion, respectively, were selected for NGS testing and results were compared to IHC and FISH. Interestingly, both amplicon-based assays gave relevant results; however, only one allowed us to detect and to correctly identify the presence of two new and one rare *ALK* rearrangements.

Results

Specimen characteristics. A total of 1128 NSCLC specimens submitted to the University Hospitals of Montpellier or Toulouse (France) for detection of *ALK* translocations were firstly screened using IHC. The *ALK* IHC-positive samples (69, 6.1%) were further explored using FISH. Among them, we randomly selected 15 samples positive for *ALK* rearrangement determined by both IHC and FISH. Twenty-two *ALK*-negative samples were also selected as negative controls. We then performed two amplicon-based NGS assays: the Ion AmpliSeq RNA Lung Cancer Research Fusion Panel and the Archer® FusionPlex® *ALK*, *RET*, *ROS1* v2 kit.

Fusion gene detection using the AmpliSeq kit. The Ion AmpliSeq RNA Lung Cancer Research Fusion Panel is based on an amplicon target enrichment approach that allows amplification and detection of 70 known fusion transcripts for the *ALK*, *RET*, *ROS1*, and *NTRK1* genes using a couple of primers specific of each fusion (Fig. 1). If no common fusion transcripts are detected, a 3'/5' ratio is calculated for the four genes included in the panel and, according to the value obtained, samples are classified into three categories: no evidence, uncertain evidence, or strong evidence of the presence of a fusion. Thus, an imbalanced ratio may reflect the presence of a novel or uncommon fusion transcript in the sample that requires further exploration using a complementary molecular technique.

To assess the reliability of the panel and the inter-run reproducibility, we analyzed RNA extracted from two well-characterized control samples included in FFPE blocks in three independent experiments. These control samples corresponded to an engineered standard sample which harbored two well-characterized fusion transcripts, *EML4-ALK* (E6:A20) and *CCDC6-RET* (C1:R12), and the human lung cancer cell line H2228 (*EML4-ALK*; E6:A20). As expected, using the AmpliSeq panel for library preparations in combination with the corresponding bioinformatics analysis, the variant fusion transcripts harbored by the two control samples were correctly identified in the three independent experiments.

We next retrospectively analyzed the 37 selected tumor samples for which the *ALK* fusion status had been previously determined using conventional techniques (IHC and FISH). An *ALK* rearrangement was clearly detected in 12 cases: five *EML4-ALK* (E6:A20), six *EML4-ALK* (E13:A20), and one *EML4-ALK* (E13:A19) (Table 1). For the 25 remaining samples, as no known fusion transcript was highlighted by the analysis, the 3'/5' imbalance values were interpreted to determine the presence or not of a potential fusion transcript (Supplementary Table S1). Among these, 13 samples did not display evidence of a fusion and were considered negative. However, for 11 other samples, uncertain evidence of an *ALK* and *RET* fusion was reported in 10 samples and one sample, respectively (Tables 1 and S1).

When comparing the results obtained with the AmpliSeq RNA Fusion kit to those from the conventional techniques, concordant diagnoses were reported for 25 samples (67%) (Tables 1 and 2). Since a clear conclusion could not be made for 11 samples (30%), further experiments are required to deliver a molecular diagnosis (Tables 1 and 2). Finally, a discordant result occurred using this NGS kit for one sample (3%); an *ALK* fusion gene was detected using IHC/FISH but not using the AmpliSeq method (S37, Tables 1 and 2).

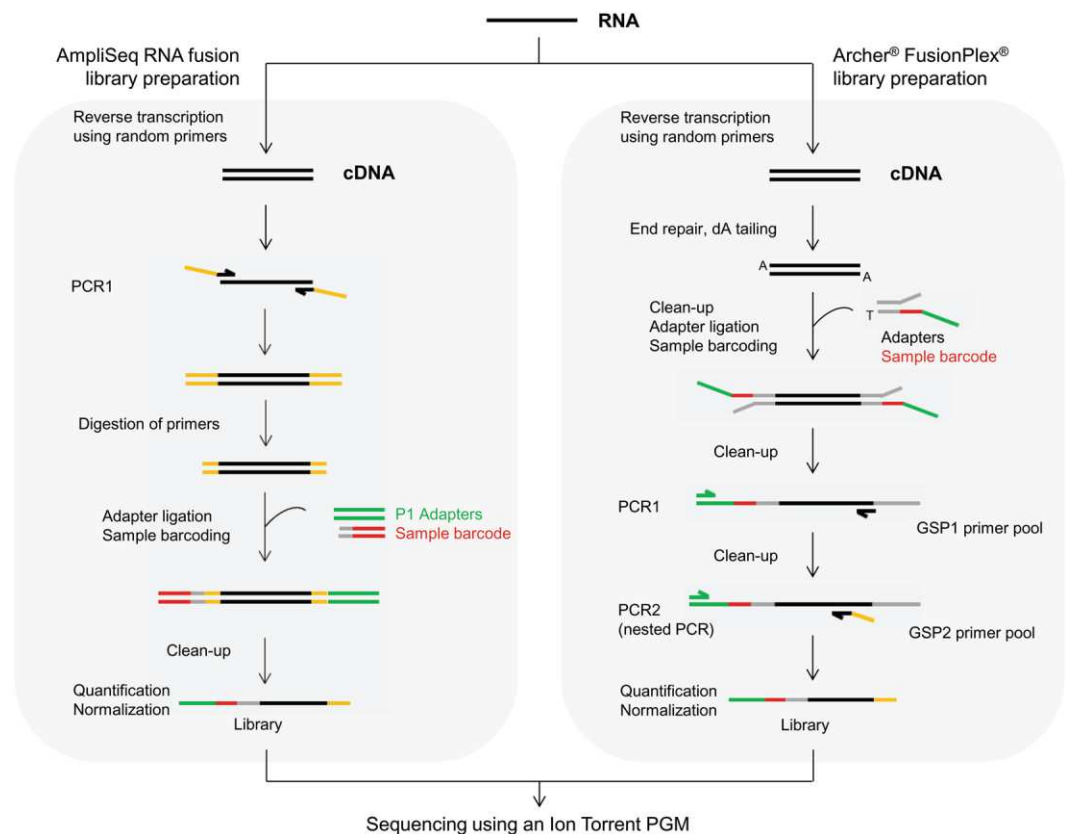


Figure 1. Schematic description of the library preparation workflow using the AmpliSeq RNA fusion kit or the Archer® FusionPlex® kit.

Fusion Gene Detection in Tumor Specimens using the Archer® FusionPlex®. The Archer® FusionPlex® kits are based on a different targeted enrichment method known as AMP (Fig. 1). This open-ended technique has the advantage of being able to sequence fused partners of the targeted genes without *a priori*. This permits the detection of well-described fusion events as well as previously unknown partners, with identification of the detected novel fusion transcript.

To investigate the specificity of the kit, we analyzed the RNA extracted from the two well-characterized control samples in three independent experiments. The Archer® FusionPlex® *ALK*, *RET*, *ROS1* v2 allowed the expected detection of the fusion breakpoints present in the samples: *EML4-ALK* (E6:A20) and *CCDC6-RET* (C1:R12) for the commercialized sample, and *EML4-ALK* (E6:A20) for the human lung cancer cell line H2228.

Using the AMP target enrichment technique on the same 37 tumor samples, we detected the presence of *ALK* fusion transcripts in 15 specimens (Tables 1 and S2). Among them, 12 harbored a common *EML4-ALK* rearrangement (E6:A20 and E13:A20). Interestingly, one novel and two rare *ALK* fusion transcripts were also identified: *GCC2-ALK* (S35), *DCTN1-ALK* (S36), and *CLIP1-ALK* (S37) (Tables 1 and S2). For 22 cases, no rearrangements were highlighted for the *ALK*, *RET*, and *ROS1* genes.

Results obtained using the Archer® FusionPlex® kit correlated perfectly with those from the ‘gold standard’ conventional methods. Indeed, all the specimens reported by IHC and FISH as negative cases ($n = 22$) or positive ($n = 15$) for *ALK* translocations were correctly classified using this molecular approach (Table 2). More importantly, unlike the IHC and the FISH methods, this technique allowed identification of the fusion partners without *a priori*, revealing the presence of uncommon fusion transcripts in three tumor samples in our study.

Validation of the uncommon *ALK* fusion partners and patient clinical outcome. Among the 10 *ALK*-positive IHC/FISH samples, three exhibited a singular gene fusion detected by the Archer® FusionPlex® kit: *CLIP1-ALK* (S35), *DCTN1-ALK* (S36), and *GCC2-ALK* (S37) rearrangements. Primers flanking the specific fusion regions were designed and used for validation. After reverse transcription and PCR, the presence of gene fusions was further analyzed using Sanger sequencing. Of particular note, our results validated the presence of three uncommon *DCTN1-ALK* (D26:A20), *CLIP1-ALK* (C22:A20), and *GCC2-ALK* (G19:A20) fusion genes (Fig. 2). All three uncommon fusion transcripts detected included the first 1065, 1294, and 1482 amino acids (aa) of *DCTN1*, *CLIP1*, and *GCC2*, respectively, fused to the last 562 aa of the *ALK* protein and retained the intact kinase domain of *ALK*, which is located from 1116 to 1392 aa (Fig. 2a,d and f).

As IHC and FISH revealed an *ALK*-positive status (Fig. 2b), patient S36 received crizotinib orally at a dose of 250 mg twice daily in September 2015, which resulted in a significant symptomatic improvement and computed tomography (CT) response after three months of therapy (Fig. 2c). The patient remains on treatment with

Sample ID	IHC results	FISH results	NGS results		Results summary
			AmpliSeq RNA Fusion kit	Archer® FusionPlex® kit	
S01	Negative	Negative	Negative	Negative	Concordance
S02	Negative	Negative	Negative	Negative	Concordance
S03	Negative	Negative	Negative	Negative	Concordance
S04	Negative	Negative	Negative	Negative	Concordance
S05	Negative	Negative	Negative	Negative	Concordance
S06	Negative	Negative	Negative	Negative	Concordance
S07	Negative	Negative	Negative	Negative	Concordance
S08	Negative	Negative	Negative	Negative	Concordance
S09	Negative	Negative	Negative	Negative	Concordance
S10	Negative	Negative	Negative	Negative	Concordance
S11	Negative	Negative	Negative	Negative	Concordance
S12	Negative	Negative	Negative	Negative	Concordance
S13	Negative	Negative	Negative	Negative	Concordance
S14	Positive	Positive	<i>EML4-ALK</i> (E6:A20)	<i>EML4-ALK</i> (E6:A20)	Concordance
S15	Positive	Positive	<i>EML4-ALK</i> (E6:A20)	<i>EML4-ALK</i> (E6:A20)	Concordance
S16	Positive	Positive	<i>EML4-ALK</i> (E6:A20)	<i>EML4-ALK</i> (E6:A20)	Concordance
S17	Positive	Positive	<i>EML4-ALK</i> (E6:A20)	<i>EML4-ALK</i> (E6:A20)	Concordance
S18	Positive	Positive	<i>EML4-ALK</i> (E6:A20)	<i>EML4-ALK</i> (E6:A20)	Concordance
S19	Positive	Positive	<i>EML4-ALK</i> (E13:A20)	<i>EML4-ALK</i> (E13:A20)	Concordance
S20	Positive	Positive	<i>EML4-ALK</i> (E13:A20)	<i>EML4-ALK</i> (E13:A20)	Concordance
S21	Positive	Positive	<i>EML4-ALK</i> (E13:A20)	<i>EML4-ALK</i> (E13:A20)	Concordance
S22	Positive	Positive	<i>EML4-ALK</i> (E13:A20)	<i>EML4-ALK</i> (E13:A20)	Concordance
S23	Positive	Positive	<i>EML4-ALK</i> (E13:A20)	<i>EML4-ALK</i> (E13:A20)	Concordance
S24	Positive	Positive	<i>EML4-ALK</i> (E13:A19)	<i>EML4-ALK</i> (E13:A20)	Discordance in the variant detected by NGS approaches
S25	Positive	Positive	<i>EML4-ALK</i> (E13:A20)	<i>EML4-ALK</i> (E13:A20)	Concordance
S26	Negative	Negative	Uncertain <i>ALK</i>	Negative	Discordance between NGS results
S27	Negative	Negative	Uncertain <i>ALK</i>	Negative	Discordance between NGS results
S28	Negative	Negative	Uncertain <i>ALK</i>	Negative	Discordance between NGS results
S29	Negative	Negative	Uncertain <i>ALK</i>	Negative	Discordance between NGS results
S30	Negative	Negative	Uncertain <i>ALK</i>	Negative	Discordance between NGS results
S31	Negative	Negative	Uncertain <i>ALK</i>	Negative	Discordance between NGS results
S32	Negative	Negative	Uncertain <i>ALK</i>	Negative	Discordance between NGS results
S33	Negative	Negative	Uncertain <i>ALK</i>	Negative	Discordance between NGS results
S34	Negative	Negative	Uncertain <i>RET</i>	Negative	Discordance between NGS results
S35	Positive	Positive	Uncertain <i>ALK</i>	<i>GCC2-ALK</i> (G19:A20)	Discordance between NGS results
S36	Positive	Positive	Uncertain <i>ALK</i>	<i>DCTN1-ALK</i> (D26:A20)	Discordance between NGS results
S37	Positive	Positive	Negative	<i>CLIP1-ALK</i> (C22:A20)	Discordance between NGS results

Table 1. Detection of *ALK* rearrangements in clinical specimens using IHC, FISH, AmpliSeq RNA Fusion kit, and Archer® FusionPlex® Kit.

crizotinib and a recent CT scan demonstrated a significant shrinkage of all tumor sites outside the central nervous system. According to the *ALK*-positive status of patient S37 (Fig. 2e), crizotinib treatment was started in February 2016. Successive images showed a continuation of response to therapy, with stabilization of the skeletal metastases but evidence of local extension of brain metastases. Finally, despite positive IHC/FISH results (Fig. 2g), *ALK*-inhibitor efficiency could not be assessed in patient S35, as this patient is still in remission after surgery.

Discussion

ALK gene rearrangements are usually detectable using IHC or FISH, and guide patient selection for therapy. Currently, expert consensus proposes the use of *ALK* IHC assays as a screening tool in two-step testing, with FISH evaluation used to validate positive or equivocal IHC samples^{10,27–29}. However, several studies have reported *ALK* fusions in samples that had tested negative using IHC, demonstrating that protein expression is not automatically linked to gene rearrangements. This highlights the risk of denying therapy with an *ALK* inhibitor delivery based only on IHC results^{13,30–32}. In addition, the interpretation of *ALK* rearrangement by FISH strongly relies on expert experience that requires long periods of training and can be compromised by technical pitfalls³³. Moreover, neither FISH nor IHC allow for identification of fusion partners and exact breakpoints.

Molecular diagnosis could overcome the limits of both these conventional analyses. However, to date, no technical consensus has emerged. In this study, we used two commercial amplicon-based NGS assays to determine

	Number of samples	NGS diagnosis	
		AmpliSeq Kit	ArcherDx Kit
Routine diagnosis (IHC and/or FISH)	25	Concordant	Concordant
	11	Uncertain ^a	Concordant
	1	Discordant	Concordant

Table 2. Concordance between diagnoses delivered using conventional techniques (IHC and/or FISH) and NGS-based molecular approaches. ^aThe 3'/5' imbalance value obtained could not allow to clearly determined the presence or not of an fusion transcript in the samples. Another technique must be performed to deliver a diagnosis.

the presence of clinically actionable *ALK* fusion transcripts. To maximize efficiency, we used 37 true NSCLC patient-derived oncology specimens previously tested for *ALK* by IHC and FISH. Twelve common *EML4-ALK* E13:A20 (S14–S18) and E6:A20 (S19–S25) variants were detected by both NGS approaches. The two assays displayed different results for three samples (S35, S36, and S37). The AmpliSeq amplicon-based method delivered an “uncertain” result, whereas the Archer[®] AMP-based approach detected an *ALK*-positive fusion for all of them. Through combined RT-PCR and Sanger sequencing analysis, we validated the presence of a rearrangement in each sample. The AmpliSeq and AMP Archer[®] FusionPlex[®] methods identified a *EML4(13)-ALK(19)* fusion and a *EML4(13)-ALK(20)* fusion, respectively, in one FISH-positive sample (S24). This fusion was validated as *EML4(13)-ALK(20)* using RT-PCR and Sanger sequencing (data not shown). Importantly, among the 22 *ALK* fusion-negative samples initially detected by conventional approaches, the AMP Archer[®] FusionPlex[®] assay correctly established a negative result for all specimens, whereas the AmpliSeq assay was “uncertain” for nine of them, rendering further investigations necessary before *ALK* fusion status could be concluded²². Altogether, these results demonstrate the clear advantage of AMP Archer[®] FusionPlex[®] over AmpliSeq amplicon-based methodology in terms of giving clinically relevant, highly accurate results in a timely manner.

The Archer[®] FusionPlex[®] results suggest that it could be routinely used for the molecular diagnosis of NSCLC rearrangements. It is an easy-to-use laboratory test with kits developed for both PGM sequencer (Thermo Fisher Scientific) and MiSeq sequencer (Illumina) technologies. The workflow design provides a result in five working days. Furthermore, the accuracy of the test observed in our cohort demonstrates that confirmation of the result using another molecular approach is not required, as has previously been suggested³⁴. There was no screening failure in our study, even though, in some cases, the RNA analyzed was extracted from samples that contained less than 20% tumor cells. In this respect, as cytological samples are the only source of material for a significant number of patients, we are planning to examine this approach in this setting. Finally, there have been concerns that the bioinformatics aspect of NGS may be challenging for regional/county hospitals. However, using Archer analysis software, we could detect and validate all known and novel rearrangements despite the absence of strong bioinformatical infrastructure in our unit, and without specific pipeline development.

Different techniques based on high-throughput molecular approaches have been improved recently, and used to detect the presence of fusion transcripts in NSCLC samples. Thus, NanoString Technologies developed a technique based on the dual hybridization of a capture probe and a molecularly barcoded reporter probe complementary to a contiguous target sequence, allowing an accurate count of molecules, even where the RNA is poor quality. The nCounter Vantage[™] Lung Fusion Panel included junction probes specific to the fusion breakpoint, and probes upstream and downstream of a potential fusion junction for detection of gene-expression imbalance. Comparisons of NanoString performance with IHC and FISH have clearly shown showed a high degree of concordance with these gold standard techniques^{17,20,35,36}. Amplicon-based NGS fusion panels have also been developed by different suppliers, and two main methods are available: the target enrichment-based (e.g. Thermo Fisher) and the AMP-based approaches (e.g. ArcherDx, Qiagen or MolecularMD). As demonstrated in this study, both molecular methods are highly sensitive, easy to perform, and give comparable results to conventional techniques. The main difference is that, unlike target enrichment-based methods, AMP-based approaches allow identification and correct naming of rare and new fusion transcripts. Indeed, although the imbalance detection used in the target enrichment-based approach is a good method to detect samples harboring new fusion genes, further experiments must then be performed to confirm the identity of the fusion partner^{22,37}. Very recently, Rogers *et al.* also evaluated a new technology developed by Agena Bioscience, based on cDNA synthesis, amplification, labeling, and detection using mass spectrometry, in combination with the Agena LungFusion panel³⁵. In this study, the authors compared the three transcriptome-based approaches (nCounter Vantage[™] Lung Fusion Panel from NanoString Technologies, AmpliSeq RNA Lung fusion panel from Thermo Fisher, and Agena LungFusion panel) to FISH, and showed an overall agreement ranging from 86–96%, depending on the technique. Interestingly, both the Agena panel and AmpliSeq fusion panel reported fusions that were not detectable by FISH.

In the present study, we identified one rare *DCTN1-ALK* fusion transcript (S36), one new *CLIP1-ALK* fusion (S37), and one new *GCC2-ALK* rearrangement (S35). *DCTN1*, for dynactin subunit 1, encodes the largest subunit of dynactin, a macromolecular complex that binds to both microtubules and cytoplasmic dynein. *DCTN1-ALK* fusions have been rarely reported: four inflammatory myofibroblastic tumors (IMT)^{38–40}, six Spitz tumors^{41,42}, and one pancreatic tumor⁴³. It has also been observed in two specimens with NSCLC^{44,45}. Fusion of *ALK* with *DCTN1* induces the constitutive activation of *ALK*, which can be inhibited *in vitro* by treatment of the cells with crizotinib⁴². However, patients' responses to crizotinib was poorly described in these studies and, at present, only one patient with an IMT that responded to *ALK* inhibitor has been reported³⁹. Interestingly, the patient harboring

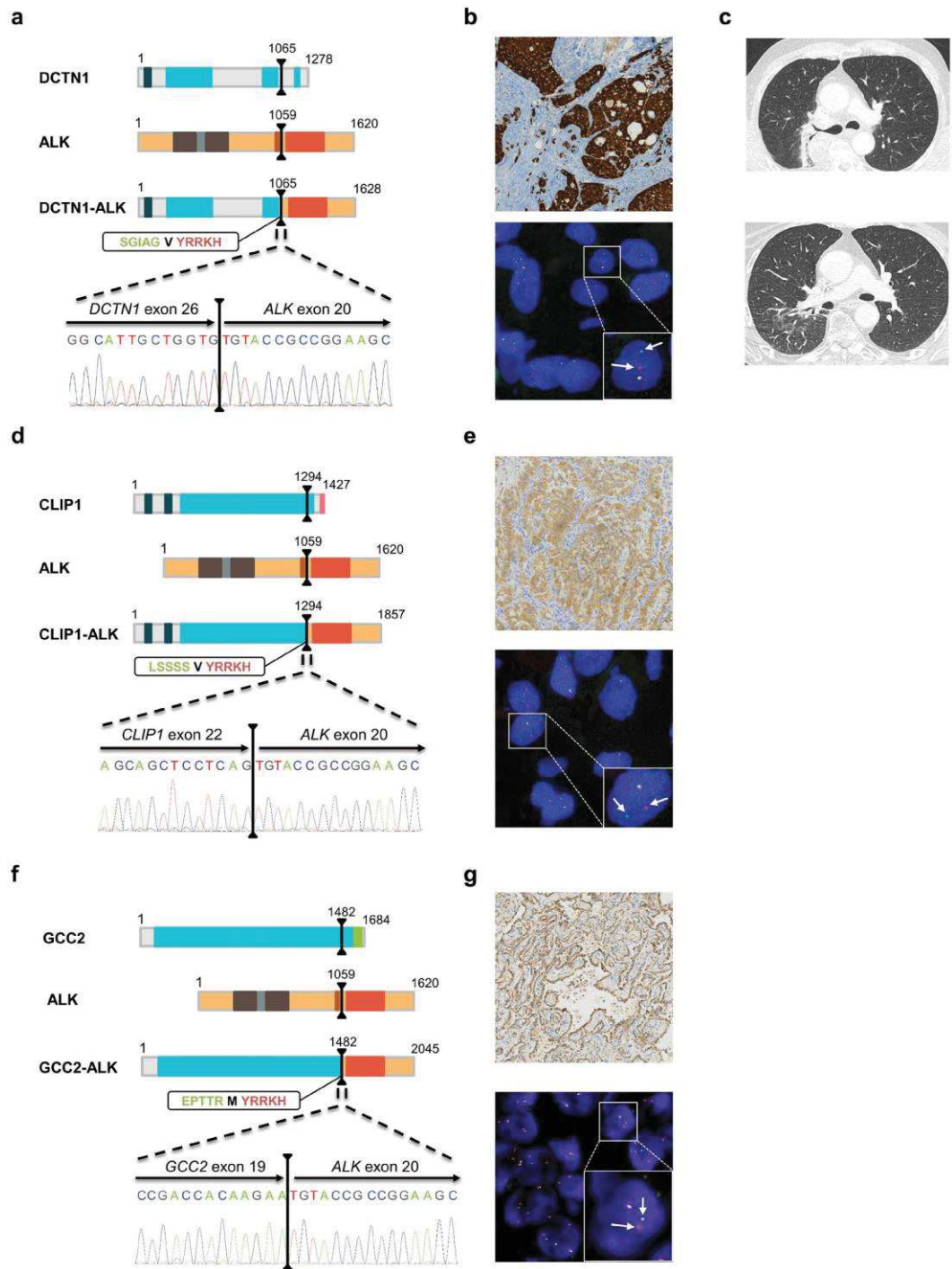


Figure 2. Detection of one rare *DCTN1-ALK* fusion transcript, one new *CLIP1-ALK* fusion, and one new *GCC2-ALK* rearrangement in patient samples. (a,d and f) Schematic representation of the main functional domains of the proteins. The black lines represent the breakpoints and the dashed lines zoom in on the transcript fusion points and the electropherogram of the validation test. The amino acid (aa) sequences at the fusion points are highlighted in a rectangle: the green bolded sequences correspond to the fusion partner, the red bolded sequence to the *ALK* sequence, and the black bolded aa to the aa generated by the fusion of the codon from the fusion partner and *ALK*. The protein functional domains are represented as colored boxes: deep blue box, Cytoskeleton-associated protein glycine-rich (CAP-Gly) domain; light blue box, coiled-coil domain; brown box, Meprin/A5-protein/PTPmu (MAM) domain; grey box, LDL-receptor class A domain; orange box, transmembrane domain; red box, kinase domain; pink box, zinc finger domain; green box, GRIP domain. (b,e and g) IHC and FISH images showing the presence of *ALK* rearrangements in patient samples. Top panel, IHC imaging showing an intense cytoplasmic staining. Bottom panel, representative image of a slide hybridized with a break-apart *ALK* FISH assay. In this given example, the box highlights one nucleus harboring a split (arrows) and a fused signal. (c) Thoracic CT scan of patient S36 before (top panel) and after (bottom panel) three months of crizotinib therapy. IHC, immunohistochemistry; FISH, fluorescence *in situ* hybridization; CT, computed tomography.

this fusion rearrangement in our study showed sensitivity to crizotinib (S36), consistent with the results of the IMT crizotinib-treated patient.

Moreover, we highlighted, for the first time, the presence of a *CLIP1-ALK* fusion in an NSCLC sample. Although one case of a Spitz tumor harboring a *CLIP1-ALK* fusion has previously been reported, the breakpoint described differs⁴⁶. Yeh and colleagues identified a breakpoint located between *CLIP1* exon 13 and *ALK* exon 20, whereas in our study it is between the exon 22 of *CLIP1* and 20 of *ALK*. *CLIP1* protein is a member of the cytoskeleton-associated protein family with a conserved glycine-rich domain. It binds to microtubules and thereby plays an important role in intracellular vesicle trafficking. As observed for patient S36, patient S37 responded to crizotinib therapy and continues crizotinib monotherapy with no evidence of major disease progression.

Finally, we also identified, for the first time to our knowledge, a new *ALK* fusion partner: *GCC2*. The breakpoint is located between *GCC2* exon 19 and *ALK* exon 20. *GCC2*, for GRIP and coiled-coil domain containing 2, encodes for Golgi proteins involved in the tethering of transport vesicles to the *trans*-Golgi network. As previously reported for other *ALK* fusion partners, the large coiled-coil domain harbored by *GCC2*, when fused with *ALK*, may facilitate dimerization and induce the constitutive activation of *ALK*. The patient harboring this fusion is still in remission after surgery, rendering it impossible to determine the *ALK*-inhibitor efficiency in this case (S35).

The response of cell lines or patients to *ALK* inhibitors depending on the *ALK* fusion variant expressed has been insufficiently explored to date^{47–49}. Of greatest interest is the observation that patients with an *EML4-ALK* variant 1 (E13:A20) exhibit better outcomes with crizotinib treatment than patients without this variant⁴⁹, suggesting that *ALK* variants might influence the response duration of crizotinib in *ALK*-positive NSCLC. Moreover, Heuckmann and colleagues demonstrated that the cellular localization of the *EML4-ALK* fusion protein depends on the variant expressed, which may affect the oncogenic activity of the fusion protein⁴⁷. At present, the specific *ALK* variant status and the fusion partner involved is not routinely considered when determining a prognosis or a therapeutic stratification for patients. Further comprehensive studies are now required to monitor patient outcomes according to the specific *ALK* variant status. AMP-based assays, which allow the precise determination of the fusion partner and breakpoint, are a simple tool for acquiring this information. This paves the way for the development of large cohort studies to determine the impact of this information on the healthcare of lung cancer patients.

With new targetable driver genes identified, and with therapeutic options evolving, a new composite decisional algorithm must be defined. As *ALK*-positive lung cancer patients benefit from tyrosine kinase inhibitor therapy in the first-line setting, *ALK* must be tested at the time of diagnosis. Our results suggest that an amplicon-based NGS assay could be performed initially. However, for laboratories that would prefer to continue using IHC as a screening test, Archer[®] FusionPlex[®] could be performed as a second step, to replace FISH. Moreover, since current guidelines recommend routine *ALK* testing as well as *EGFR* testing, it is important to point out that all these actionable driver genes should be tested as part of a one-test multiplex NGS panel, extracting DNA-RNA from the same FFPE sample. *ROS* and *RET* fusions, as well as a broader spectrum of genes (i.e. *KRAS*, *BRAF*, or *ERBB2*), could also be included in such routine tests. Finally, since the optimal amount of RNA recommended for Archer[®] FusionPlex[®] analysis is 200 ng (range from 20 to 250 ng), this parameter could represent a limitation for very small biopsies, even if in our study all specimens were successfully analyzed. This point has been addressed recently by Evangelista *et al.* that implemented NanoString panel for *ALK* fusion detection and demonstrated its applicability in series of 43 lung cancer biopsies using up to 100 ng of RNA with only 7% of sample failure³⁶.

In summary, our study investigated *ALK* fusion detection based on two different commercially NGS-based approaches in FFPE-derived cancer specimens. In contrast to the AmpliSeq amplicon-based approach that was unable to detect several variants, the Archer[®] AMP-based technique successfully identified all *ALK* fusion-positive samples, rendering this method highly applicable for routine *ALK* fusion detection and variant identification. In addition, in contrast to the conventional IHC and FISH techniques, this amplicon-based NGS approach has the distinct advantage of requiring knowledge only of one partner in the fusion. This allows the identification of novel gene rearrangements with previously unknown partners, which could clinically impact patient management.

Materials and Methods

Tumor samples. This study was performed with approval from the Institutional Review Board of both hospitals (Toulouse and Montpellier) and in concordance with regulatory guidelines regarding clinical assay validation. For this non-interventional study, an approved informed consent statement was acquired for all patients. FFPE tissue samples from NSCLC patients that had been submitted in 2014 to the University Hospital of Montpellier or Toulouse (France) for detection of *ALK* translocations were included in this study (n = 1128). Among them, 37 samples with previously determined *ALK* rearrangement status were randomly selected. Table 3 lists the characteristics of the patients and the corresponding specimens enrolled in the NGS assay. All lesions were submitted for pathological examination using standard techniques. The percentage of tumor cells in the specimens ranged from 10–90%. For each sample, *ALK* fusions were explored using IHC, a dual-color break-apart FISH, and NGS approaches using two different assays: Ion AmpliSeq RNA Lung Cancer Research Fusion Panel (Thermo Fisher Scientific, Waltham, MA), and Archer[®] FusionPlex[®] *ALK*, *RET*, *ROS1* v2 Kit (ArcherDX, Boulder, CO). Results were interpreted blindly, without knowledge of the results obtained by the other methods.

Control samples. The lung adenocarcinoma cell line NCI-H2228 (*EML4-ALK* fusion) was purchased from the American Type Culture Collection (Manassas, USA) and cultured as recommended. An FFPE Horizon Diagnostics control sample (Cambridge, UK) that harbored two well-characterized fusion transcripts was acquired (*EML4-ALK* and *CCDC6-RET* fusions).

Characteristics	n	%
Sex		
Male	20	54.0
Female	17	46.0
Age		
<60	6	16.2
>60	31	83.8
Smoking status		
Have smoked	8	21.6
Smoker	9	24.3
Non-smoker	14	37.8
Unknown	6	16.2
Stage		
I	6	16.2
II	7	18.9
III	8	21.6
IV	14	37.8
Unknown	2	5.4
Type of specimen		
Biopsy	19	51.4
Surgical specimen	17	45.9
Unknown	1	2.7
Tumor cell content		
<50%	8	21.6
≥50%	29	78.4

Table 3. Patient and specimen characteristics.

ALK IHC. ALK IHC was carried out using the mouse monoclonal antibody 5A4 (Abcam, Cambridge, UK) before 2015, and subsequently using the rabbit monoclonal antibody D5F3 (Roche, Basel, Switzerland). Briefly, 3–4 μm FFPE tumor tissue sections were deparaffinized and incubated in a PT link (Dako, Glostrup, Denmark) with a high pH buffer according to the manufacturer's recommendations for antigen retrieval. Anti-ALK antibody was then applied for 30 min at 1:50. Slides were incubated at room temperature with EnVision FLEX+ Mouse Linker (Dako) for 15 min. The immune complexes were then detected with the dextran polymer reagent. The percentage of labeled tumor cells and intensity of staining were independently assessed by two pathologists.

ALK FISH. Where IHC analysis was positive, FISH was performed on 3 μm FFPE tissue sections using the ALK FISH DNA break-apart Probe, Split Signal (Dako) according to the manufacturer's recommendations. Slides were pretreated at 98 °C in solution for 10 min and digested with pepsin for 3 min at 37 °C using the histology FISH accessory Kit (Dako). Slides were incubated for 18 h at 45 °C with ALK probes diluted at 1:10, and had been previously denatured for 5 min at 85 °C. Slides were then washed and dehydrated before counterstaining and application of mounting medium. Slides were analyzed with a Zeiss AxioImager Z1 fluorescence microscope (Labexchange, Burladingen, Germany). Slides were analyzed independently by two pathologists. A minimum of 100 nuclei were scored and cases were considered positive when more than 15% of cells displayed split signals.

Total RNA extraction. RNA extraction was performed on the same FFPE blocks as the IHC and/or FISH exploration. RNA was extracted from 10 μm-thick paraffin sections using the RecoverAll™ Total Nucleic Acid Isolation Kit (Thermo Fisher Scientific, Wilmington, USA) according to the manufacturer's recommendations. RNA from control samples was extracted using the same kit. Extracted RNA was quantified using the Qubit® RNA HS Assay kit in combination with a Qubit® 2.0 fluorometer (Thermo Fisher Scientific) and qualified using the RNA 6000 Nano kit in combination with the BioAnalyzer 2100™ (Agilent Technologies, Palo Alto, CA, USA). Molecular testing by the two NGS-based approaches was performed on the same RNA samples.

Ion AmpliSeq RNA Lung Cancer Research Fusion Panel experiment. The AmpliSeq RNA Lung Cancer Research Fusion Panel is based on an amplicon sequencing approach (Table 4). The panel is composed of 83 pairs of unique primers in a single pool that includes: (i) primers that allow the amplification and detection of 70 known *ALK*, *RET*, *ROS1*, and *NTRK1* fusion transcripts; (ii) primers located in the 5' and 3' regions of *ALK*, *RET*, *ROS1*, and *NTRK1* mRNA genes; (iii) primers that target five housekeeping genes to serve as internal controls of the experiment.

For library preparation, 10 ng of total RNA was used according to the manufacturer's recommendations. Briefly, RNA was reverse transcribed using the SuperScript® VILO™ cDNA Synthesis Kit (Thermo Fisher Scientific). Target cDNA was amplified using AmpliSeq primer pool (Fig. 1). Primer sequences were then partially digested using FuPa reagent, and adapters and barcodes were ligated using DNA ligase. Libraries were

Kit	AmpliSeq RNA Fusion	Archer® FusionPlex®
Supplier	Thermo Fisher Scientific	ArcherDx
Molecular approach	Target enrichment	AMP-based method
RNA input	10 ng	200 ng
Genes in the panel	<i>ALK, RET, ROS, NTRK1</i>	<i>ALK, RET, ROS</i>
Panel customizable	Yes, but necessity to design a new panel as PCR primers are multiplexed together	Yes. Easy because only one primer is specific of the target region. Thus, no problem of multiplexing
Detectable alterations	Fusion transcripts	Fusion transcripts, point mutations, small insertion/deletion
Detection of new fusion variants	Yes, but only detection of the presence, not allowed the identification	Yes, with a correct naming of the breakpoint and the fusion partner
Analysis software	Provided by the supplier	Provided by the supplier
Sequencer	PGM only	PGM and Illumina sequencers

Table 4. Characteristics of AmpliSeq RNA fusion and Archer® FusionPlex® approaches.

purified with Agencourt® AMPure® XP (Beckman Coulter, Nyon, Switzerland), amplified by PCR as described in the user guide, purified again, and quantified using a Qubit® 2.0 fluorometer using the Qubit dsDNA HS assay kit (Thermo Fisher Scientific). Based on the calculated library concentration, eight libraries were pooled to equimolar concentration. The emulsion PCR and chip loading were then performed using an Ion Chef in combination with the Ion PGM™ Hi-Q™ Chef kit and the Ion 318™ Chip kit v2 according to the manufacturer's recommendations (Thermo Fisher Scientific). Finally, sequencing was performed on the Ion PGM sequencer using the Ion PGM™ Hi-Q™ Sequencing kit and analyzed by the Ion Reporter™ 4.4 Software (Thermo Fisher Scientific).

For samples where the software did not detect a known fusion transcript, the 3'/5' imbalance value given by the software was used to determine the presence or not of novel or uncommon fusion transcripts²⁵. For each gene present in the panel, a specific threshold has been determined by the supplier to classify samples into three categories: no evidence, uncertain evidence, or strong evidence of the presence of a fusion involving the corresponding genes.

Archer® FusionPlex® *ALK, RET, ROS1* v2 Kit experiment. The Archer® FusionPlex® *ALK, RET, ROS1* v2 kit is based on a targeted enrichment method called anchored multiplex PCR (AMP), derived from the rapid amplification of cDNA ends (RACE) method (Table 4)⁵⁰. After reverse transcription, double-stranded cDNA undergoes end repair, adenylation, and ligation with a half-functional universal adapter (Fig. 1). Obtained cDNA are then amplified by two rounds of nested low-cycle PCR using nested gene-specific primers (GSP1 and GSP2) in combination with the first half-functional universal adapter. GSP2 primers are also 5' tagged with a common sequencing adapter to allow the clonal amplification necessary for the sequencing step.

The panel used is composed of: (i) 29 GSP that allow the detection of gene fusion events involving *ALK, RET, ROS*, and also *ALK* and *RET* specific point mutations, at the same time; (ii) GSP specific for five housekeeping genes.

For this kit, 200 ng of total RNA was used as input for library generation using the Archer Universal RNA reagent Kit v2, Archer Molecular Barcode (MBC) Adapters for Ion Torrent, and the Archer FusionPlex *ALK, RET, ROS* v2 Panel GSPs v2 (ArcherDX, Boulder, CO, USA) according to the manufacturer's instructions. Briefly, RNA was reverse transcribed using random primers, first strand cDNA was synthesized, and RNA quality was assessed using the Archer PreSeq RNA QC assay (ArcherDX). After second strand cDNA synthesis, end repair and A-tailing steps were performed, cDNA was purified using Agencourt® AMPure® XP beads (Beckman Coulter), and MBC adapters were ligated. Purified cDNA was firstly amplified using the GSP1 pool, then purified using Agencourt® AMPure® XP beads, and amplified again using the GSP2 pool. After another purification step, libraries were quantified using D1000 ScreenTapes in combination with a 4200 TapeStation instrument (Agilent Technologies, Santa Clara, CA, USA) and pooled to equimolar concentration. Emulsion PCR, chip loading, and sequencing was performed as described above and results were analyzed using the Archer Analysis v3.3 software. A sample was considered as positive when the fusion breakpoint was supported by at least two unique reads.

RT-PCR and Sanger sequencing. 200 ng of RNA was reverse transcribed with random hexamers using the SuperScript®III First-Strand Synthesis System (Thermo Fisher). Primers specific for the detected fusion events were designed (Supplementary Table 3) and direct Sanger sequencing was performed as previously described⁵¹.

References

- Shaw, A. T. *et al.* Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. *N Engl J Med.* **368**, 2385–94 (2013).
- Network, C. G. A. R. Comprehensive molecular profiling of lung adenocarcinoma. *Nature.* **511**, 543–50 (2014).
- Kwak, E. L. *et al.* Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med.* **363**, 1693–703 (2010).
- Kris, M. G. *et al.* Using multiplexed assays of oncogenic drivers in lung cancers to select targeted drugs. *JAMA.* **311**, 1998–2006 (2014).
- Hallberg, B. & Palmer, R. H. The role of the ALK receptor in cancer biology. *Ann Oncol.* **27**(Suppl 3), iii4–iii15 (2016).
- Koivunen, J. P. *et al.* EML4-ALK fusion gene and efficacy of an ALK kinase inhibitor in lung cancer. *Clin Cancer Res.* **14**, 4275–83 (2008).
- Solomon, B. J. *et al.* First-line crizotinib versus chemotherapy in ALK-positive lung cancer. *N Engl J Med.* **371**, 2167–77 (2014).
- Seto, T. *et al.* CH5424802 (RO5424802) for patients with ALK-rearranged advanced non-small-cell lung cancer (AF-001JP study): a single-arm, open-label, phase 1-2 study. *Lancet Oncol.* **14**, 590–8 (2013).

9. Iwama, E., Okamoto, I., Harada, T., Takayama, K. & Nakanishi, Y. Development of anaplastic lymphoma kinase (ALK) inhibitors and molecular diagnosis in ALK rearrangement-positive lung cancer. *Onco Targets Ther.* **7**, 375–85 (2014).
10. Katayama, R. *et al.* Mechanisms of acquired crizotinib resistance in ALK-rearranged lung Cancers. *Sci Transl Med.* **4**, 120ra17 (2012).
11. Li, W. *et al.* Combinational Analysis of FISH and Immunohistochemistry Reveals Rare Genomic Events in ALK Fusion Patterns in NSCLC that Responds to Crizotinib Treatment. *J Thorac Oncol.* **12**, 94–101 (2017).
12. Pekar-Zlotin, M. *et al.* Fluorescence *in situ* hybridization, immunohistochemistry, and next-generation sequencing for detection of EML4-ALK rearrangement in lung cancer. *Oncologist.* **20**, 316–22 (2015).
13. Demidova, I. *et al.* Detection of ALK rearrangements in 4002 Russian patients: The utility of different diagnostic approaches. *Lung Cancer.* **103**, 17–23 (2017).
14. Soda, M. *et al.* A prospective PCR-based screening for the EML4-ALK oncogene in non-small cell lung cancer. *Clin Cancer Res.* **18**, 5682–9 (2012).
15. Wang, R. *et al.* The use of quantitative real-time reverse transcriptase PCR for 5' and 3' portions of ALK transcripts to detect ALK rearrangements in lung cancers. *Clin Cancer Res.* **18**, 4725–32 (2012).
16. Liu, C., Pepper, K., Hendrickson, H., Cagle, P. T. & Portier, B. P. Clinical Validation of a Novel Commercial Reverse Transcription-Quantitative Polymerase Chain Reaction Screening Assay for Detection of ALK Translocations and Amplifications in Non-Small Cell Lung Carcinomas. *Arch Pathol Lab Med.* **140**, 690–3 (2016).
17. Lira, M. E. *et al.* A single-tube multiplexed assay for detecting ALK, ROS1, and RET fusions in lung cancer. *J Mol Diagn.* **16**, 229–43 (2014).
18. Takeuchi, K. *et al.* Multiplex reverse transcription-PCR screening for EML4-ALK fusion transcripts. *Clin Cancer Res.* **14**, 6618–24 (2008).
19. Sunami, K. *et al.* Multiplex Diagnosis of Oncogenic Fusion and MET Exon Skipping by Molecular Counting Using Formalin-Fixed Paraffin Embedded Lung Adenocarcinoma Tissues. *J Thorac Oncol.* **11**, 203–12 (2016).
20. Reguart, N. *et al.* Identification of ALK, ROS1, and RET Fusions by a Multiplexed mRNA-Based Assay in Formalin-Fixed, Paraffin-Embedded Samples from Advanced Non-Small-Cell Lung Cancer Patients. *Clin Chem.* **63**, 751–760 (2017).
21. Laforet, M. P. *et al.* Design and feasibility of a novel, rapid, and simple fluorescence 26-plex rt-PCR assay for simultaneous detection of 24 fusion transcripts in adult acute myeloid leukemia. *J Mol Diagn.* **15**, 186–95 (2013).
22. Beadling, C. *et al.* A Multiplexed Amplicon Approach for Detecting Gene Fusions by Next-Generation Sequencing. *J Mol Diagn.* **18**, 165–75 (2016).
23. Choi, Y. L. *et al.* A novel fusion of TPR and ALK in lung adenocarcinoma. *J Thorac Oncol.* **9**, 563–6 (2014).
24. Dacic, S. *et al.* ALK FISH patterns and the detection of ALK fusions by next generation sequencing in lung adenocarcinoma. *Oncotarget.* **7**, 82943–82952 (2016).
25. Moskalev, E. A. *et al.* Sensitive and specific detection of EML4-ALK rearrangements in non-small cell lung cancer (NSCLC) specimens by multiplex amplicon RNA massive parallel sequencing. *Lung Cancer.* **84**, 215–21 (2014).
26. Pfarr, N. *et al.* High-throughput diagnostic profiling of clinically actionable gene fusions in lung cancer. *Genes Chromosomes Cancer.* **55**, 30–44 (2016).
27. Yatabe, Y. ALK FISH and IHC: you cannot have one without the other. *J Thorac Oncol.* **10**, 548–50 (2015).
28. Savic, S. *et al.* Screening for ALK in non-small cell lung carcinomas: 5A4 and D5F3 antibodies perform equally well, but combined use with FISH is recommended. *Lung Cancer.* **89**, 104–9 (2015).
29. Leighl, N. B. *et al.* Molecular testing for selection of patients with lung cancer for epidermal growth factor receptor and anaplastic lymphoma kinase tyrosine kinase inhibitors: American Society of Clinical Oncology endorsement of the College of American Pathologists/International Association for the study of lung cancer/association for molecular pathology guideline. *J Clin Oncol.* **32**, 3673–9 (2014).
30. Antonescu, C. R. *et al.* Molecular characterization of inflammatory myofibroblastic tumors with frequent ALK and ROS1 gene fusions and rare novel RET rearrangement. *Am J Surg Pathol.* **39**, 957–67 (2015).
31. Cabillic, F. *et al.* Parallel FISH and immunohistochemical studies of ALK status in 3244 non-small-cell lung cancers reveal major discordances. *J Thorac Oncol.* **9**, 295–306 (2014).
32. Ilie, M. I. *et al.* Discrepancies between FISH and immunohistochemistry for assessment of the ALK status are associated with ALK 'borderline'-positive rearrangements or a high copy number: a potential major issue for anti-ALK therapeutic strategies. *Ann Oncol.* **26**, 238–44 (2015).
33. Von Laffert, M. *et al.* Anaplastic lymphoma kinase (ALK) gene rearrangement in non-small cell lung cancer (NSCLC): results of a multi-centre ALK-testing. *Lung Cancer.* **81**, 200–6 (2013).
34. Guseva, N. V. *et al.* Anchored multiplex PCR for targeted next-generation sequencing reveals recurrent and novel USP6 fusions and upregulation of USP6 expression in aneurysmal bone cyst. *Genes Chromosomes Cancer.* **56**, 266–277 (2017).
35. Rogers, T. M. *et al.* Multiplexed transcriptome analysis to detect ALK, ROS1 and RET rearrangements in lung cancer. *Sci Rep.* **7**, 42259 (2017).
36. Evangelista, A. F. *et al.* Detection of ALK fusion transcripts in FFPE lung cancer samples by NanoString technology. *BMC Pulm Med.* **17**, 86 (2017).
37. Fang, D. D. *et al.* HIP1-ALK, a novel ALK fusion variant that responds to crizotinib. *J Thorac Oncol.* **9**, 285–94 (2014).
38. Wang, X. *et al.* Fusion of dynactin 1 to anaplastic lymphoma kinase in inflammatory myofibroblastic tumor. *Hum Pathol.* **43**, 2047–52 (2012).
39. Subbiah, V. *et al.* STUMP un“stumped”: anti-tumor response to anaplastic lymphoma kinase (ALK) inhibitor based targeted therapy in uterine inflammatory myofibroblastic tumor with myxoid features harboring DCTN1-ALK fusion. *J Hematol Oncol.* **8**, 66 (2015).
40. Lee, J. C. *et al.* ALK oncoproteins in atypical inflammatory myofibroblastic tumours: novel RRBP1-ALK fusions in epithelioid inflammatory myofibroblastic sarcoma. *J Pathol.* **241**, 316–323 (2017).
41. Busam, K. J., Kutzner, H., Cerroni, L. & Wiesner, T. Clinical and pathologic findings of Spitz nevi and atypical Spitz tumors with ALK fusions. *Am J Surg Pathol.* **38**, 925–33 (2014).
42. Wiesner, T. *et al.* Kinase fusions are frequent in Spitz tumours and spitzoid melanomas. *Nat Commun.* **5**, 3116 (2014).
43. Shimada, Y. *et al.* An Oncogenic ALK Fusion and an RRAS Mutation in KRAS Mutation-Negative Pancreatic Ductal Adenocarcinoma. *Oncologist.* **22**, 158–164 (2017).
44. Iyevleva, A. G. *et al.* Novel ALK fusion partners in lung cancer. *Cancer Lett.* **362**, 116–21 (2015).
45. Gao, X. *et al.* Clinical Implications of Variant ALK FISH Rearrangement Patterns. *J Thorac Oncol.* **10**, 1648–52 (2015).
46. Yeh, I. *et al.* Clinical, histopathologic, and genomic features of Spitz tumors with ALK fusions. *Am J Surg Pathol.* **39**, 581–91 (2015).
47. Heuckmann, J. M. *et al.* Differential protein stability and ALK inhibitor sensitivity of EML4-ALK fusion variants. *Clin Cancer Res.* **18**, 4682–90 (2012).
48. Hrustanovic, G. *et al.* RAS-MAPK dependence underlies a rational polytherapy strategy in EML4-ALK-positive lung cancer. *Nat Med.* **21**, 1038–47 (2015).
49. Yoshida, T. *et al.* Differential Crizotinib Response Duration Among ALK Fusion Variants in ALK-Positive Non-Small-Cell Lung Cancer. *J Clin Oncol.* **34**, 3383–9 (2016).
50. Zheng, Z. *et al.* Anchored multiplex PCR for targeted next-generation sequencing. *Nat Med.* **20**, 1479–84 (2014).
51. Marchant, J., Mange, A., Larrieux, M., Costes, V. & Solassol, J. Comparative evaluation of the new FDA approved THxID-BRAF test with High Resolution Melting and Sanger sequencing. *BMC Cancer.* **14**, 519 (2014).

Author Contributions

J.A.V. and J.S. conceived and designed the experiments; J.A.V., B.B., P.A. and V.S. performed experiments; S.G., D.G., P.C., J.L.P., V.C., P.B. and I.R. reviewed the manuscript and participated to fruitful discussions; I.S. and I.R. provided us the biological materials and E.C., J.L.P. and J.M. the corresponding clinical information; J.A.V., S.T. and J.S. wrote the manuscript. All authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-12679-8>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017

Annexes

Annexe 1 : Tableau des interacteurs de RAS identifiés par la technique du BioID

Gene ID	Gene Name	Top 2 controls	HEK293_FB-HRAS WT			HEK293_FB-NRAS WT			HEK293_FB-KRAS WT		
			Total	SAINT	Total Unique	Total	SAINT	Total Unique	Total	SAINT	Total Unique
948469	htrA	1 1	21337	-	104	10470	-	111	6954	-	100
3265	HRAS	1 0	5424	-	55	1	-	1	0	-	0
4893	NRAS	0 0	0	-	0	2230	-	42	0	-	0
3845	KRAS	1 0	0	-	0	0	-	0	988	-	48
10157	AASS	1 1	9	0,98	6	5	-	5	3	-	3
4363	ABCC1	0 0	83	1,00	36	123	1,00	61	149	1,00	76
10057	ABCC5	1 0	101	1,00	35	70	1,00	33	96	1,00	41
5825	ABCD3	14 14	142	-	36	397	1,00	62	221	1,00	61
6059	ABCE1	9 9	213	1,00	38	78	-	33	67	-	34
10006	ABII	10 8	200	1,00	36	190	1,00	29	123	0,99	32
10152	ABII2	2 2	36	0,84	19	75	1,00	24	25	-	13
27	ABL2	2 1	30	0,95	15	20	-	14	12	-	9
2327	ACAP2	2 1	34	0,97	16	3	-	3	15	-	11
64746	ACBD3	10 9	56	-	17	152	1,00	45	52	-	27
2181	ACSL3	7 4	69	0,85	20	43	-	21	65	0,82	24
2182	ACSL4	2 0	99	1,00	32	38	0,99	21	54	1,00	23
58	ACTA1	5 4	79	1,00	4	62	0,86	5	24	-	3
60	ACTB	0 0	300	1,00	29	0	-	0	0	-	0
71	ACTG1	30 29	317	0,99	29	271	-	34	179	-	33
8751	ADAM15	0 0	25	1,00	9	1	-	1	0	-	0
6868	ADAM17	0 0	23	1,00	8	8	-	4	1	-	1
53616	ADAM22	0 0	16	0,99	6	11	-	6	16	1,00	8
56997	ADCK3	3 1	17	-	10	24	0,81	16	7	-	5
109	ADCY3	0 0	64	1,00	16	24	1,00	19	34	1,00	15
112	ADCY6	0 0	9	-	6	8	-	8	22	1,00	12
115	ADCY9	0 0	54	1,00	22	126	1,00	42	112	1,00	43
118	ADD1	18 15	733	1,00	81	725	1,00	69	565	1,00	70
119	ADD2	5 3	221	1,00	56	137	1,00	53	223	1,00	65
120	ADD3	14 13	1050	1,00	117	1096	1,00	102	913	1,00	124
55256	AD1	1 1	0	-	0	0	-	0	11	0,99	9
51094	ADIPOR1	0 0	118	1,00	17	48	1,00	12	26	0,99	10
10939	AFG3L2	4 4	61	0,96	25	45	-	21	48	-	27
116987	AGAP1	0 0	32	1,00	16	15	0,99	9	31	1,00	20
137964	AGPAT6	1 1	8	-	4	8	-	7	30	1,00	13
84803	AGPAT9	0 0	0	-	0	10	-	8	26	1,00	15
8540	AGPS	1 1	13	0,99	6	6	-	5	5	-	3
10768	AHCYL1	7 6	255	1,00	33	74	-	24	112	1,00	27
23382	AHCYL2	3 3	126	1,00	23	37	-	17	67	1,00	20
64853	AIDA	0 0	18	1,00	6	3	-	3	5	-	4
9131	AIFM1	2 0	23	0,85	14	24	0,86	14	11	-	9
9590	AKAP12	12 11	219	1,00	77	281	1,00	108	435	1,00	125
9495	AKAP5	0 0	43	1,00	18	27	0,99	11	70	1,00	26
207	AKT1	0 0	3	-	2	12	0,99	8	17	1,00	10
208	AKT2	1 1	35	1,00	12	95	1,00	25	41	1,00	18
214	ALCAM	0 0	37	1,00	18	30	1,00	16	34	1,00	14
5832	ALDH18A1	19 17	273	1,00	55	363	1,00	86	190	-	65
219	ALDH1B1	1 1	27	1,00	12	18	1,00	9	9	-	7
347902	AMIGO2	0 0	36	1,00	10	22	1,00	15	19	1,00	8
288	ANK3	23 21	166	-	59	518	1,00	148	96	-	51
22852	ANKRD26	1 0	57	1,00	17	66	1,00	49	27	1,00	21
57182	ANKRD50	0 0	33	1,00	16	66	1,00	39	43	1,00	31
203286	ANKS6	0 0	17	0,99	8	24	1,00	16	23	1,00	16
196527	ANOG	0 0	119	1,00	38	47	1,00	23	80	1,00	30
10541	ANP32B	6 4	48	-	11	70	0,96	23	51	-	18
84168	ANTXR1	0 0	16	1,00	10	6	-	4	2	-	2
302	ANXA2	21 19	415	1,00	78	176	-	45	207	-	57
8907	AP1M1	1 0	6	-	5	12	-	9	13	0,99	9
160	AP2A1	8 7	81	23	81	36	90	0,78	37	-	37
26985	AP3M1	7 6	54	-	21	110	1,00	34	61	-	29
322	APBB1	8 6	108	1,00	25	83	-	34	70	-	26
323	APBB2	0 0	19	0,99	8	25	1,00	15	16	1,00	11
27350	APOBEC3C	0 0	11	0,98	5	6	-	4	1	-	1
55198	APPL2	3 1	44	0,91	25	20	-	13	31	0,85	16
369	ARAF	2 2	109	1,00	24	74	1,00	25	58	1,00	26
116984	ARAP2	0 0	0	-	0	7	-	7	22	1,00	16
375	ARF1	15 13	216	1,00	16	174	0,98	14	133	-	15
378	ARF4	3 3	52	0,95	13	62	1,00	16	60	1,00	15
381	ARF5	5 4	113	1,00	12	38	-	4	58	-	9
382	ARF6	4 4	75	1,00	17	49	-	16	36	-	17
27236	ARFIP1	7 6	122	1,00	33	69	-	36	70	-	33
392	ARHGAP1	7 5	187	1,00	31	86	0,78	32	134	1,00	34
94134	ARHGAP12	2 1	8	-	5	28	0,88	20	14	-	13
57584	ARHGAP21	9 6	92	0,89	42	209	1,00	77	120	0,97	66
80728	ARHGAP39	0 0	22	1,00	14	9	-	5	21	1,00	13
394	ARHGAP5	4 4	57	0,91	26	40	-	25	30	-	33
396	ARHGAP10	4 4	106	1,00	18	43	-	15	64	0,99	18
9639	ARHGEF10	1 1	0	-	0	20	0,99	14	6	-	5
200894	ARL13B	2 2	197	1,00	43	140	1,00	33	130	1,00	36
402	ARL2	5 5	61	0,81	20	35	-	17	46	-	22
10550	ARL6IP5	2 1	46	1,00	11	12	-	5	71	1,00	12
421	ARVCF	0 0	64	1,00	28	52	1,00	28	67	1,00	36
41	ASIC1	0 0	0	-	5	23	1,00	10	29	1,00	13
84896	ATAD1	0 0	13	-	10	18	1,00	9	13	1,00	8
55210	ATAD3A	4 4	118	1,00	28	94	1,00	48	49	-	30
79065	ATG9A	0 0	0	-	0	7	-	6	10	0,98	7
286410	ATP11C	0 0	18	-	10	17	1,00	13	25	1,00	16
79572	ATP13A3	0 0	18	-	9	21	1,00	12	45	1,00	26
476	ATP1A1	8 7	981	1,00	105	406	1,00	94	379	1,00	94
481	ATP1B1	0 0	86	1,00	25	47	1,00	18	34	1,00	10
483	ATP1B3	1 1	144	1,00	20	91	1,00	19	84	1,00	19
488	ATP2A2	24 23	355	1,00	61	100	-	43	243	-	67
490	ATP2B1	4 3	322	1,00	69	214	1,00	52	214	1,00	65
493	ATP2B4	0 0	107	1,00	30	35	1,00	21	33	1,00	19
498	ATP5A1	53 51	932	1,00	77	874	1,00	80	504	-	81
506	ATP5B	31 29	755	1,00	87	503	1,00	79	291	-	75
509	ATP5C1	9 8	160	1,00	30	154	1,00	30	132	1,00	25
100526740	ATP5J2-PTCI	2 2	72	1,00	8	66	1,00	9	41	0,99	8
10632	ATP5L	3 2	42	0,96	7	18	-	8	8	-	4
537	ATP6AP1	1 0	0	-	0	0	-	0	33	1,00	15
10159	ATP6AP2	0 0	0	-	0	2	-	1	43	1,00	24
23545	ATP6V0A2	0 0	0	-	0	0	-	0	0	-	0
538	ATP7A	0 0	28	1,00	11	30	1,00	23	43	1,00	28
10079	ATP9A	1 0	11	0,99	7	7	-	6	4	-	3
550	AUP1	0 0	4	-	3	10	-	5	20	1,00	8
567	B2M	0 0	32	1,00	8	2	-	2	0	-	0
374383	BTH6	0 0	13	0,99	6	9	-	8	9	-	8
10458	BAIAP2	6 5	242	1,00	60	232	1,00	64	227	1,00	68
55971	BAIAP2L1	6 6	284	1,00	57	133	1,00	47	106	1,00	40
10409	BASP1	2 2	152	1,00	26	43	0,91	22	41	-	14
10134	BCAP31	2 1	0	-	0	21	0,86	7	1	-	1
10282	BET1	2 1	29	0,95	5	32	0,97	8	11	-	3
51272	BET1L	1 1	44	1,00	8	24	1,00	9	17	1,00	8
637	BID	1 1	8	0,98	4	0	-	0	17	1,00	7
673	BRAF	1 1	58	1,00	24	108	1,00	32	73	1,00	23
55845	BRK1	1 1	39	1,00	14	13	0,99	6	14	1,00	6
682	BSG	2 2	74	1,00	8	20	-	4	28	-	8
11119	BTN3A1	0 0	25	1,00	8	0	-	0	0	-	0
11118	BTN3A2	0 0	0	-	0	21	1,00	11	16	1,00	10
10384	BTN3A3	0 0	0	-	0	3	-	3	10	0,99	9
738	Cl1orf2	4 3	26	-	11	27	-	21	71	1,00	36
28971	Cl1orf67	0 0	13	0,99	7	16	-	11	13	1,00	7
387882	Cl2orf75	0 0	20	1,00	8	0	-	0	0	-	0
63894	Cl4orf133	1 1	25	1,00	14	4	-	3	12	-	7
348110	Cl5orf38	25 0	257	0,97	24	0	-	0	0	-	0
255057	Cl9orf26	0 0	91	1,00	19	227	1,00	33	122	1,00	25
125988	Cl9orf70	3 3	50	0,98	8	0	-	0	8	-	7
126695	Clorf172	0 0	10	-	4	11	-	7	22	1,00	11
84886	Clorf198	14 13	224								

Annexes | Annexe 1 : Tableau des interacteurs de RAS identifiés par la technique du BioID

1267	CNP	4	4	220	1,00	60	99	1,00	35	84	1,00	34
57175	CORO1B	67	64	836	1,00	72	432		53	416		54
23603	CORO1C	18	16	327	1,00	48	150		29	103		24
1327	COX4H1	2	1	28	0,94	7	0		0	4		3
10328	COX4NB	1	1	2		2	5		4	20	0,99	7
1347	COX7A2	1	1	17	0,99	5	1		1	7		4
1362	CPD	4	2	182	1,00	60	171	1,00	75	324	1,00	97
8895	CPNE3	4	4	60	0,91	22	19		15	33		19
144402	CPNE8	9	9	1100	1,00	85	590	1,00	71	341	1,00	58
54504	CPVL	9	9	53		25	105	0,92	31	92		29
1399	CRKL	47	42	532	1,00	49	319		57	269		50
1434	CSEIL	17	16	126		45	104		47	195	1,00	74
1452	CSNK1A1	6	6	68		16	140	1,00	36	83	0,97	25
1456	CSNK1G3	0	0	50	1,00	19	40	1,00	15	30	1,00	19
1464	CSPG4	0	0	23	1,00	11	36	1,00	29	24	1,00	18
1466	CSR2	4	3	73	1,00	21	57	0,93	13	22		11
58190	CTDSP1	0	0	53	1,00	14	28		11	23	1,00	11
1495	CTNNA1	6	6	221	1,00	67	105	1,00	50	143	1,00	62
1499	CTNNB1	0	0	109	1,00	26	96	1,00	25	70	1,00	20
1500	CTNND1	13	11	1289	1,00	156	887		106	913	1,00	132
1508	CTSB	1	1	0		0	10		6	18	1,00	8
2017	CTTN	33	31	343	0,97	57	203		49	156		37
8454	CUL1	6	5	89	1,00	29	29		22	27		20
1525	CXADR	2	1	560	1,00	58	230	1,00	36	244	1,00	40
23191	CYFIP1	10	9	372	1,00	81	428	1,00	98	260	1,00	85
26999	CYFIP2	11	11	340	1,00	68	424	1,00	97	341	1,00	102
23002	DAAM1	0	0	10		7	20	0,99	15	30	1,00	23
1601	DAB2	0	0	8		4	3		3	15	0,99	11
153090	DAB2IP	0	0	11		6	39	1,00	27	37	1,00	27
1605	DAG1	0	0	62	1,00	15	30	1,00	17	36	1,00	13
55157	DARS2	4	3	46	0,90	17	41		21	36		17
26528	DAZAP1	6	3	19		6	1		1	54	0,84	13
28988	DBNL	13	12	186	1,00	38	43		20	77		22
285761	DCBLD1	0	0	34	1,00	10	72	1,00	26	51	1,00	21
1650	DDOST	6	6	155	1,00	36	80	0,88	31	72		20
10212	DDX39A	19	17	209	0,87	52	128		43	176		45
1654	DDX3X	36	34	519	1,00	73	356		57	250		60
22898	DENND3	0	0	0		0	3		3	19	1,00	14
160518	DENND5B	1	0	6		3	10		7	19	1,00	11
55789	DEPDC1B	13	11	408	1,00	67	411	1,00	75	337	1,00	75
8527	DKD	1	1	0		0	6		5	14	0,99	13
1717	DHCR7	4	1	23	0,88	7	24	0,90	7	14		7
1719	DHFR	2	3	39		11	66	1,00	23	86	1,00	21
51635	DHR57	2	1	0		0	29	0,95	15	58	1,00	23
22907	DHX30	5	4	102	1,00	26	46		28	44		27
23181	DIP2A	11	11	155	1,00	37	88		48	55		30
57609	DIP2B	4	3	62	0,99	33	40		29	41		29
1738	DLG	4	3	85	1,00	21	26		15	17		10
1739	DLG1	17	16	326		60	850	1,00	113	483		86
1741	DLG3	0	0	120	1,00	48	144	1,00	49	109	1,00	40
9231	DLG5	0	0	0		0	52	1,00	37	24	1,00	17
22839	DLGAP4	0	0	2		2	19	1,00	14	11		10
10294	DNAJA2	10	8	145	1,00	26	53		25	111		20
51726	DNAJB11	0	0	0		0	4		4	13	0,99	7
10049	DNAJB6	1	0	22	1,00	12	9		8	6		4
1785	DNM2	4	4	50	0,90	24	86	1,00	42	89	1,00	56
1793	DOCK1	0	0	16	0,99	7	23	1,00	18	29	1,00	22
139818	DOCK11	1	0	27		14	69	1,00	45	70	1,00	49
85440	DOCK7	12	11	103		55	171	1,00	93	99		56
23348	DOCK9	0	0	40	1,00	16	52	1,00	41	90	1,00	54
57171	DOLPP1	2	1	30	0,96	5	4		4	12		4
8813	DPM1	5	5	59		18	103	1,00	32	58	0,80	22
4733	DRG1	13	13	185	1,00	33	108		29	81		31
1824	DSC2	2	1	421	1,00	51	346	1,00	53	314	1,00	51
1829	DSG2	23	16	1102	1,00	128	1133	1,00	119	1008	1,00	124
667	DST	23	21	340	1,00	123	895	1,00	246	785	1,00	227
1837	DTNA	3	3	61	0,99	26	27		15	54	0,77	28
8655	DYNLL1	3	2	37	0,83	7	0		0	1		1
83657	DYNLRB2	1	1	15	0,99	6	3		1	6		2
22873	DZIP1	0	0	2		1	14	0,99	13	5		5
9166	EBAG9	0	0	11	0,99	6	24	1,00	7	11		5
10682	EBP	1	0	1		1	48	1,00	12	10		5
1889	ECE1	0	0	7		5	6		5	18	0,99	9
10455	ECE2	1	0	24	1,00	11	13	1,00	6	10		7
80820	EEDP1	0	0	49	1,00	19	46	1,00	22	42	1,00	22
79180	EFHD2	13	10	133	0,89	33	72		22	54		21
1947	EFNB1	1	0	340	1,00	40	258	1,00	37	180	1,00	28
1948	EFNB2	0	0	465	1,00	45	248	1,00	33	184	1,00	23
1949	EFNB3	0	0	105	1,00	21	42	1,00	18	36	1,00	10
23167	EFR3A	1	1	83	1,00	29	83	1,00	41	78	1,00	32
22979	EFR3B	0	0	150	1,00	51	158	1,00	66	129	1,00	57
1956	EGR	0	0	89	1,00	29	85	1,00	32	107	1,00	33
23301	EHBP1	11	8	260	1,00	71	170	1,00	77	259	1,00	97
10938	EHD1	1	0	79	1,00	26	41	1,00	26	74	1,00	35
30844	EHD4	2	1	94	1,00	37	54	0,99	26	85	1,00	41
63916	ELMO2	1	1	17	1,00	8	12		9	21	1,00	13
84173	ELMOD3	1	0	36	1,00	15	19	0,99	8	24	1,00	10
60481	ELOVL5	2	2	0		0	37	0,91	6	29	0,86	7
2010	EMD	11	10	80		23	162	1,00	31	143		36
2023	ENO1	75	66	1560	1,00	71	1268	1,00	64	702		63
2037	EPB41L2	54	47	981	1,00	99	504		48	579		61
23136	EPB41L3	84	78	1965	1,00	155	1918	1,00	156	1405	1,00	139
64097	EPB41L4A	0	0	3		2	5		3	11	0,98	9
54566	EPB41L4B	0	0	164	1,00	27	90	1,00	24	37	1,00	15
57669	EPB41L5	5	4	170	1,00	33	181	1,00	52	223	1,00	56
1969	EPHA2	10	9	471	1,00	87	588	1,00	100	443	1,00	92
2042	EPHA3	0	0	94	1,00	26	87	1,00	37	80	1,00	33
2043	EPHA4	1	1	32	1,00	12	33	1,00	15	45	1,00	22
2045	EPHA7	0	0	168	1,00	53	88	1,00	36	90	1,00	40
2048	EPHB2	0	0	35	1,00	14	43	1,00	20	42	1,00	21
2050	EPHB4	0	0	59	1,00	18	37	1,00	23	42	1,00	22
29924	EPN1	3	2	13		10	51	1,00	21	35	0,91	21
22905	EPN2	1	1	23	1,00	15	22	0,99	12	14	0,99	11
58513	EPS15L1	50	47	562	1,00	122	330		105	241		72
2059	EP88	1	1	70	1,00	32	35	1,00	24	40	1,00	28
2064	ERBB2	0	0	47	1,00	17	28	1,00	17	60	1,00	24
55914	ERBB2IP	21	20	535		102	975	1,00	173	592		119
2066	ERBB4	0	0	9		6	6		5	16	0,99	8
23344	ESYT1	33	31	902	1,00	120	746	1,00	112	639	1,00	130
57488	ESYT2	12	12	398	1,00	71	324	1,00	88	262	1,00	73
55218	EXD2	1	1	0		0	16	1,00	12	15	0,99	10
11336	EXOC3	11	10	279	1,00	78	147	1,00	62	232	1,00	70
10640	EXOC5	0	0	9		8	7		5	28	0,99	16
23265	EXOC7	0	0	17	0,99	9	16	0,99	13	17	1,00	16
50848	F11R	0	0	88	1,00	17	49	1,00	10	17	0,99	9
9415	FADS2	1	1	13		7	14	0,99	10	4		4
93343	FAM125A	1	1	50	1,00	13	6		5	27	1,00	17
84668	FAM126A	1	0	21	0,99	7	22	1,00	8	35	1,00	11
64855	FAM129B	11	10	524	1,00	70	287	1,00	66	313	1,00	69
54463	FAM134B	1	0	0		0	8		5	13	0,99	9
57579	FAM135A	5	5	117		34	129	1,00	49	68		32
221061	FAM171A1	3	1	87	0,99	30	108	1,00	52	75	1,00	37
284069	FAM171A2	0	0	57	1,00	21	90	1,00	37	82	1,00	39
165215	FAM1											

Annexes | Annexe 1 : Tableau des interacteurs de RAS identifiés par la technique du BioID

10788	IQGAP2	2	1	27	0,94	15	18	15	57	1,00	40	
8471	IRS4	32	29	2489	1,00	174	2494	1,00	176	1,964	1,00	161
83737	ITCH	0	0	61	1,00	25	43	1,00	27	89	1,00	35
3672	ITGA1	0	0	15	1,00	10	82	1,00	36	12	1,00	11
3673	ITGA2	0	0	323	1,00	76	338	1,00	71	184	1,00	63
3675	ITGA3	0	0	38	1,00	13	25	1,00	11	15	1,00	9
3676	ITGA4	1	0	58	1,00	27	59	1,00	27	49	1,00	27
3678	ITGA5	0	0	302	1,00	48	152	1,00	41	174	1,00	45
3655	ITGA6	0	0	424	1,00	92	466	1,00	96	381	1,00	91
3679	ITGA7	0	0	48	1,00	20	94	1,00	43	55	1,00	28
8516	ITGA8	0	0	8	1,00	4	39	1,00	21	37	1,00	21
3685	ITGA9	0	0	321	1,00	74	243	1,00	61	254	1,00	80
3688	ITGB1	1	1	680	1,00	66	566	1,00	56	416	1,00	57
3693	ITGB5	1	0	0	0	0	14	0,99	9	5	1,00	5
3704	ITPA	1	0	34	1,00	8	2	1,00	2	24	1,00	6
3714	JAG2	0	0	29	1,00	10	10	0,98	5	9	1,00	6
58494	JAM2	0	0	19	1,00	8	3	1,00	3	7	1,00	6
83700	JAM3	0	0	133	1,00	24	42	1,00	15	31	1,00	11
56704	JPH1	2	2	78	1,00	29	74	1,00	29	73	1,00	28
3728	JUP	9	9	645	1,00	113	774	1,00	95	557	1,00	94
8997	KALRN	0	0	6	1,00	4	0	1,00	0	13	0,99	6
9312	KCNB2	0	0	0	0	0	44	1,00	23	38	1,00	23
56479	KCNQ5	0	0	4	1,00	3	14	0,99	8	9	1,00	6
115207	KCTD12	1	1	2	1,00	1	6	1,00	5	34	1,00	16
9772	KIAA0195	0	0	23	1,00	13	24	1,00	13	58	1,00	28
79932	KIAA0319L	0	0	0	0	0	4	1,00	4	17	1,00	13
9847	KIAA0528	3	3	119	1,00	37	118	1,00	41	155	1,00	52
643314	KIAA0754	0	0	15	0,99	10	5	1,00	5	15	0,99	12
27143	KIAA1274	24	22	374	1,00	79	418	1,00	94	290	1,00	85
57648	KIAA1522	0	0	23	1,00	12	43	1,00	28	42	1,00	24
57670	KIAA1549	0	0	58	1,00	19	44	1,00	22	42	1,00	19
57707	KIAA1609	0	0	39	1,00	11	17	0,99	10	21	1,00	13
84451	KIAA1804	1	1	30	1,00	14	40	1,00	27	25	1,00	13
57498	KIDINS220	4	0	56	0,94	30	152	1,00	72	81	0,99	46
55243	KIRREL	0	0	347	1,00	76	329	1,00	71	288	1,00	72
889	KRIT1	0	0	0	0	0	10	0,99	10	2	1,00	2
55004	LAMTOR1	2	2	80	1,00	26	38	0,92	19	52	1,00	23
28956	LAMTOR2	1	0	163	1,00	25	11	1,00	6	89	1,00	19
8649	LAMTOR3	3	2	42	0,92	10	23	1,00	11	53	0,99	12
51520	LARS	34	33	496	1,00	88	215	1,00	77	241	1,00	86
3927	LASP1	3	3	68	1,00	17	4	1,00	4	3	1,00	3
3930	LBR	25	24	247	1,00	35	290	1,00	38	346	1,00	42
253558	LCLAT1	1	1	15	1,00	6	11	0,99	8	6	1,00	6
3939	LDHA	4	4	108	1,00	23	40	1,00	22	35	1,00	20
26119	LDLRAP1	1	1	14	1,00	5	7	1,00	4	13	0,99	8
55327	LINTC	3	2	50	0,91	22	47	0,99	17	38	1,00	18
3996	LLGL1	7	6	667	1,00	93	561	1,00	79	306	1,00	62
3993	LLGL2	0	0	48	1,00	22	29	1,00	18	25	1,00	17
3998	LMAN1	3	3	15	1,00	9	72	1,00	31	31	1,00	17
4001	LMNB1	13	11	143	0,89	48	90	1,00	44	79	1,00	36
4012	LNPEP	1	1	71	1,00	32	86	1,00	43	128	1,00	64
101060252	LOC101060252	3	2	58	1,00	8	18	1,00	8	23	1,00	7
9926	LPGAT1	0	0	9	1,00	3	2	1,00	1	15	0,99	10
22859	LPHN1	0	0	13	0,99	8	75	1,00	42	48	1,00	34
23266	LPHN2	1	0	509	1,00	93	416	1,00	98	357	1,00	93
987	LRBA	36	34	374	1,00	115	526	1,00	187	423	0,97	154
4040	LRP6	0	0	0	0	0	4	1,00	4	15	0,99	13
7804	LRP8	0	0	65	1,00	20	52	1,00	23	57	1,00	26
10128	LRPPRC	7	6	32	1,00	13	27	1,00	16	103	1,00	55
55227	LRRC1	0	0	127	1,00	34	80	1,00	19	66	1,00	21
94030	LRRC4B	0	0	25	1,00	13	20	1,00	11	16	0,99	13
255252	LRRCS7	0	0	124	1,00	30	86	1,00	23	70	1,00	21
56262	LRRCSA	0	0	6	1,00	5	24	1,00	14	31	1,00	20
84230	LRRCSB	0	0	0	0	0	10	1,00	8	15	1,00	10
51599	LSR	2	1	312	1,00	53	219	1,00	34	160	1,00	46
4067	LYN	0	0	515	1,00	60	281	1,00	55	234	1,00	54
11313	LYPLA2	13	9	168	1,00	22	82	1,00	5	21	1,00	5
8216	LZTR1	0	0	357	1,00	76	455	1,00	103	376	1,00	87
84445	LZTS2	0	0	10	1,00	4	23	1,00	19	20	1,00	10
23499	MACF1	2	2	16	1,00	11	16	1,00	15	83	1,00	44
4085	MAD2L1	1	1	36	1,00	8	8	1,00	9	2	1,00	4
84061	MAGT1	0	0	0	0	0	15	1,00	9	19	1,00	9
5604	MAP2K1	4	3	66	1,00	17	63	0,99	34	25	1,00	16
5605	MAP2K2	1	0	23	1,00	7	15	1,00	10	14	1,00	8
9448	MAP4K4	4	3	76	1,00	32	67	1,00	34	67	0,98	31
11183	MAP4K5	10	10	304	1,00	68	124	1,00	55	117	1,00	41
79109	MAPKAP1	1	1	32	1,00	13	59	1,00	22	60	1,00	21
4082	MARCKS	3	2	191	1,00	35	131	1,00	30	101	1,00	26
65108	MARCKSL1	0	0	99	1,00	17	17	1,00	10	64	1,00	11
4139	MARK1	1	1	20	0,99	8	40	1,00	27	22	1,00	18
2011	MARK2	3	1	88	0,90	17	163	0,88	41	143	1,00	47
4140	MARK3	5	4	324	1,00	69	306	1,00	68	173	1,00	57
153562	MARVELD2	0	0	155	1,00	42	82	1,00	34	102	1,00	39
151963	MB21D2	5	5	159	1,00	30	86	1,00	29	49	1,00	20
4155	MBP	0	0	33	1,00	7	1	1,00	1	12	0,99	5
4162	MCAM	0	0	564	1,00	64	365	1,00	47	303	1,00	48
4171	MCM2	4	3	44	1,00	19	4	1,00	4	46	0,77	26
4175	MCM6	14	14	174	1,00	50	99	1,00	50	77	1,00	40
4176	MCM7	25	23	342	1,00	71	414	1,00	84	217	1,00	77
4233	MET	0	0	163	1,00	33	40	1,00	22	55	1,00	25
4238	MFAP3	0	0	33	1,00	11	9	1,00	5	13	0,99	8
256471	MFSD8	0	0	13	0,99	7	0	1,00	0	5	1,00	5
4259	MGST3	2	1	0	0	0	0	1,00	0	18	0,80	5
57553	MICAL3	3	3	33	1,00	18	56	1,00	36	62	1,00	37
85377	MICALL1	9	8	135	1,00	45	124	1,00	53	71	1,00	38
4282	MIF	2	2	6	1,00	4	4	1,00	2	30	0,88	8
50488	MINK1	0	0	12	0,99	8	14	1,00	12	22	1,00	10
4301	MLLT4	78	70	2223	1,00	275	2006	1,00	251	1297	1,00	225
158747	MOSPD2	1	0	2	1,00	1	5	1,00	3	15	1,00	7
9526	MPP1	2	1	12	1,00	6	21	0,86	7	16	1,00	8
4354	MPP1	10	7	293	1,00	54	167	1,00	43	195	1,00	44

64398	MPP5	2	2	34	0,92	11	6	1,00	5	17	1,00	8
51678	MPP6	1	0	115	1,00	34	49	1,00	28	59	1,00	34
143098	MPP7	2	2	310	1,00	68	334	1,00	81	287	1,00	87
23164	MPRIP	3	2	7	1,00	5	55	0,77	29	62	1,00	38
9019	MPZL1	0	0	118	1,00	20	116	1,00	18	43	1,00	16
84245	MRI1	2	1	0	0	0	22	0,89	16	15	1,00	5
4478	MSN	2	2	35	0,98	15	17	1,00	13	45	1,00	14
51537	MTFP1	1	1	18	1,00	6	0	1,00	0	4	1,00	4
4522	MTHFD1	24	18	389	1,00	90	230	0,83	71	182	1,00	67
25902	MTHFD1L	4	2	44	0,86	21	14	1,00	12	37	1,00	23
4534	MTM1	0	0	30	1,00	16	25	1,00	14	14	0,99	11
54893	MTMR10	0	0	52	1,00	18	50	1,00	31	53	1,00	27
54545	MTMR12	1	1	48	1,00	20	9	1,00	8	21	1,00	14
64419	MTMR14	2	2	43	1,00	14	15	1,00	11	29	1,00	13
2475	MTOR	0	0	47	1,00	26	100	1,00	55	154	1,00	82
55149	MTPAP	1	1	0	0	0	13	0,99	9	7	1,00	5
10135	NAMPT	13	12	172	1,00	51	195	1,00	60	90	1,00	41
8775	NAPA	0	0	89								

Annexes | Annexe 1 : Tableau des interacteurs de RAS identifiés par la technique du BioID

84295	PHF6	1	1	21	1,00	8	7	3	18	1,00	8	
26227	PHGDH	16	16	323	1,00	51	128	36	109	1,00	39	
23187	PHLDB1	0	0	15	0,99	9	20	0,99	15	1,00	9	
90102	PHLDB2	1	0	14	1,00	9	43	1,00	30	24	1,00	16
5297	P14KA	4	4	108	1,00	42	163	1,00	92	142	1,00	75
8301	PICALM	6	5	76	0,86	23	54	16	37	20		
128869	PIGU	2	1	19	0,76	6	8	7	28	0,92	11	
5290	PIK3CA	0	0	6	1,00	4	38	1,00	19	34	1,00	20
5291	PIK3CB	0	0	12	1,00	7	24	1,00	16	35	1,00	19
5295	PIK3R1	1	1	118	1,00	44	149	1,00	47	146	1,00	53
5296	PIK3R2	0	0	156	1,00	41	185	1,00	66	235	1,00	69
8503	PIK3R3	0	0	29	1,00	16	48	1,00	25	78	1,00	26
8394	PIP5K1A	2	2	66	1,00	24	77	1,00	31	42	0,99	18
23396	PIP5K1C	0	0	12	0,99	5	8	7	15	0,99	10	
23760	PITPNB	0	0	13	0,99	8	0	0	0	0		
5315	PKM2	59	58	822	1,00	76	320	71	344	73		
5585	PKN1	1	1	2	1,00	1	1	1	13	0,99	9	
5586	PKN2	20	19	247	1,00	67	175	79	166	68		
5318	PKP2	3	2	93	1,00	42	70	1,00	39	100	1,00	41
8502	PKP4	0	0	82	1,00	27	145	1,00	64	119	1,00	55
23236	PLCB1	0	0	434	1,00	102	194	1,00	78	227	1,00	77
5332	PLCB4	0	0	3	1,00	2	10	0,99	10	18	0,99	13
23007	PLCH1	0	0	26	1,00	14	102	1,00	56	89	1,00	43
23228	PLCL2	0	0	27	1,00	15	26	1,00	20	29	1,00	20
5337	PLD1	0	0	164	1,00	45	86	1,00	43	105	1,00	49
59338	PLEKHA1	8	7	569	1,00	73	447	1,00	53	281	1,00	48
59339	PLEKHA2	1	0	21	1,00	6	8	5	6	4		
54477	PLEKHA5	9	9	388	1,00	104	367	1,00	109	304	1,00	107
144100	PLEKHA7	1	0	47	1,00	22	70	1,00	44	51	1,00	32
80301	PLEKHO2	0	0	8	1,00	5	0	0	22	1,00	10	
57048	PLSCR3	0	0	84	1,00	16	2	2	28	1,00	4	
5361	PLXNA1	0	0	117	1,00	54	106	1,00	52	115	1,00	64
23654	PLXNB2	0	0	9	1,00	5	12	12	23	1,00	16	
5420	PODXL	0	0	17	1,00	8	5	2	6	3		
5464	PPA1	17	13	233	1,00	45	170	0,96	38	150	37	
8613	PPAP2B	0	0	39	1,00	7	17	0,99	7	25	1,00	8
8500	PPF1A1	1	0	57	1,00	26	103	1,00	42	66	1,00	38
8496	PPF1B1	2	1	331	1,00	80	411	1,00	114	129	0,97	41
26051	PPP1R16B	0	0	8	1,00	4	16	0,99	12	9	8	
55607	PPP1R9A	0	0	17	0,99	8	30	1,00	19	21	1,00	20
5515	PPP2CA	5	4	56	0,86	17	22	9	31	14		
25824	PRDX5	2	2	0	1,00	0	5	3	32	0,95	16	
5576	PRKAR2A	12	12	150	0,98	39	100	35	138	37		
5577	PRKAR2B	6	5	64	0,80	20	72	30	128	1,00	43	
5584	PRKC1	4	2	223	1,00	33	117	1,00	28	88	1,00	27
10419	PRMT5	8	7	160	1,00	43	335	1,00	75	64	25	
55615	PRR5	1	1	49	1,00	15	25	1,00	14	59	1,00	21
23362	PSD3	2	1	250	1,00	59	126	1,00	61	179	1,00	71
5663	PSEN1	1	1	34	1,00	6	2	1	8	8	4	
5709	PSMD3	14	13	188	1,00	58	128	50	121	51		
9050	PSTPIP2	6	4	94	1,00	26	30	23	65	0,87	21	
9791	PTDSS1	1	1	8	1,00	4	18	1,00	11	7	5	
10728	PTGES3	2	1	4	1,00	3	19	4	34	0,97	7	
5754	PTK7	0	0	147	1,00	45	106	1,00	44	116	1,00	46
5770	PTPN1	9	8	35	1,00	14	75	28	104	0,98	35	
5783	PTPN13	13	13	437	1,00	152	478	1,00	194	482	1,00	195
5784	PTPN14	3	2	29	1,00	16	54	0,99	32	31	23	
5786	PTPRA	0	0	19	1,00	11	5	5	24	1,00	14	
5792	PTPRF	0	0	119	1,00	40	23	1,00	16	32	1,00	20
5793	PTPRG	0	0	73	1,00	31	177	1,00	68	152	1,00	61
5795	PTPRJ	0	0	21	1,00	14	30	1,00	20	17	13	
284119	PTRF	0	0	36	1,00	17	19	1,00	12	45	1,00	19
5818	PVRL1	0	0	10	0,98	5	13	1,00	9	9	6	
25945	PVRL2	3	0	670	1,00	70	306	1,00	45	264	1,00	47
25945	PVRL3	0	0	13	0,99	7	2	2	1	1		
10890	RAB10	2	1	33	0,98	5	24	0,90	5	14	4	
9230	RAB11B	4	3	109	1,00	29	32	15	96	1,00	35	
80223	RAB11FIP1	10	9	153	1,00	37	112	0,80	48	118	0,97	53
22841	RAB11FIP2	2	1	98	1,00	22	13	5	6	6		
26056	RAB11FIP5	5	4	76	1,00	27	47	21	64	0,78	25	
201475	RAB12	0	0	30	1,00	14	12	0,99	9	10	6	
5872	RAB13	2	2	49	1,00	14	33	0,94	14	36	0,98	16
22931	RAB18	3	2	32	1,00	12	37	0,81	19	50	0,99	23
23011	RAB21	5	4	99	1,00	22	29	14	33	14		
51715	RAB23	1	1	224	1,00	40	72	1,00	24	86	1,00	26
5874	RAB27B	0	0	26	0,99	8	15	1,00	8	9	5	
83871	RAB34	2	1	50	1,00	17	22	0,79	15	41	0,99	18
5865	RAB3B	1	1	10	1,00	4	12	0,99	7	15	0,99	9
5868	RAB5A	3	2	62	1,00	20	33	13	36	0,91	13	
5878	RAB5C	4	3	53	0,99	17	51	0,77	12	36	14	
5870	RAB6A	5	5	61	0,91	18	64	20	49	24		
7879	RAB7A	4	4	212	1,00	34	118	1,00	36	136	1,00	37
4218	RAB8A	4	3	59	0,94	13	50	0,93	16	35	14	
9367	RAB9A	3	3	66	1,00	17	25	12	37	18		
285282	RABL3	2	1	16	1,00	6	15	9	28	0,90	10	
5894	RAF1	4	3	175	1,00	37	243	1,00	47	157	1,00	46
26064	RAI14	15	15	421	1,00	112	679	1,00	139	449	1,00	128
5898	RALA	0	0	31	1,00	11	17	1,00	12	7	5	
253959	RALGAPA1	1	0	12	1,00	6	16	12	31	1,00	19	
55103	RALGPS2	1	0	38	1,00	16	7	5	16	1,00	9	
5901	RAN	24	23	229	1,00	36	451	1,00	33	202	35	
5911	RAP2A	0	0	59	1,00	10	1	9	5			
9693	RAPGEF2	0	0	12	0,99	6	7	7	9	6		
5922	RASA2	0	0	16	0,99	10	7	6	21	0,99	17	
22821	RASA3	0	0	53	1,00	27	22	0,99	12	19	1,00	13
9462	RASAL2	0	0	129	1,00	52	186	1,00	77	187	1,00	80
5955	RCN2	0	0	19	0,99	8	0	0	10	8		
5962	RDX	10	10	273	1,00	39	41	22	63	24		

768211	RELL1	0	0	272	1,00	32	122	1,00	24	127	1,00	24
11079	RER1	3	2	24	1,00	12	30	13	39	0,89	12	
23180	RFTN1	1	0	49	1,00	19	81	1,00	29	42	1,00	27
6000	RGS7	0	0	25	1,00	9	4	4	9	7		
79651	RHBDP2	0	0	44	1,00	16	20	12	33	1,00	17	
391	RHOQ	1	1	20	1,00	6	8	6	14	8		
89941	RHOT2	2	1	20	0,81	10	23	0,90	14	21	0,85	15
253260	RICTOR	3	3	187	1,00	56	238	1,00	102	183	1,00	92
55781	RIOK2	0	0	2	1,00	1	6	4	17	0,99	13	
6091	ROBO1	0	0	74	1,00	34	88	1,00	46	54	1,00	32
6092	ROBO2	0	0	75	1,00	33	84	1,00	44	79	1,00	40
6093	ROCK1	1	0	142	1,00	59	128	1,00	69	95	1,00	53
9475	ROCK2	1	0	47	1,00	23	45	1,00	30	69	1,00	38
4919	ROR1	1	0	19	1,00	11	16	1,00	10	5	4	
4920	ROR2	4	1	181	1,00	55	246	1,00	47	176	1,00	50
6102	RP2	0	0	31	1,00	13	37	1,00	16	38	1,00	14
6185	RPN2	12	11	210	1,00	44	162	1,00	53	139	46	
10670	RRAGA	2	0	70	1,00	14	0	0	0	0		
64121	RRAGC	5	4	72	1,00	20	19	11	57	0,91	25	
89970	RSPRY1	0	0	20	0,99	8	3	3	16	1,00	10	
8634	RTCD1	1	1	10	1,00	5	14	1,00	7	22	1,00	14
6282	SI00A11	4	3	97	1,00	8	0	0	5	2		
56681	SAR1A	4	3	52	0,99	10	26	10	43	10		
51128	SAR1B	0	0	11	1,00	4	22	1,00	8	16	7	
6305	SBF1	2	0	159	1,00	58	143	1,00	76	155	1,00	72
9522	SCAMP1	5	4	322	1,00	28						

Annexes | Annexe 1 : Tableau des interacteurs de RAS identifiés par la technique du BioID

6533	SLC6A6	0	0	28	1,00	8	36	1,00	12	19	1,00	8
6535	SLC6A8	1	1	385	1,00	52	349	1,00	39	203	1,00	43
6536	SLC6A9	0	0	127	1,00	30	72	1,00	17	59	1,00	17
6541	SLC7A1	1	0	70	1,00	18	28	1,00	4	29	1,00	7
23657	SLC7A11	1	0	20	1,00	9	36	1,00	8	16	1,00	6
6542	SLC7A2	1	0	188	1,00	35	180	1,00	35	116	1,00	29
8140	SLC7A5	2	1	112	1,00	20	60	1,00	17	63	1,00	12
9057	SLC7A6	0	0	0	0	0	1	0	1	15	1,00	8
6548	SLC9A1	0	0	21	1,00	8	41	1,00	21	32	1,00	14
9368	SLC9A3R1	5	2	57	0,93	22	6	0	4	23	0	9
9351	SLC9A3R2	0	0	77	1,00	20	27	0,99	14	44	1,00	19
133308	SLC9B2	0	0	7	1,00	4	15	0,99	9	9	1,00	7
353189	SLC04C1	0	0	21	0,99	8	0	0	0	0	0	0
8773	SNAP23	4	3	682	1,00	57	492	1,00	40	495	1,00	52
6616	SNAP25	0	0	112	1,00	31	32	1,00	16	56	1,00	24
9342	SNAP29	8	8	323	1,00	39	230	1,00	40	179	1,00	38
116841	SNAP47	3	2	64	1,00	30	56	1,00	34	76	1,00	32
6641	SNTB1	1	1	46	1,00	22	67	1,00	31	33	1,00	20
6645	SNTB2	3	3	40	1,00	18	54	0,95	21	58	1,00	26
6642	SNX1	17	16	215	1,00	56	200	1,00	56	173	1,00	57
29934	SNX12	3	0	48	0,97	12	2	0	2	2	0	2
9784	SNX17	2	2	58	1,00	20	3	0	3	9	0	8
6643	SNX2	8	7	131	1,00	41	76	1,00	36	60	1,00	31
81609	SNX27	2	1	26	0,93	14	6	0	5	62	1,00	29
8724	SNX3	2	1	99	1,00	21	39	0,99	19	31	0,97	14
58533	SNX6	5	3	73	0,84	27	12	0	6	25	1,00	16
51429	SNX9	18	14	279	1,00	53	202	0,81	42	193	1,00	48
6646	SOAT1	1	0	0	0	0	46	1,00	14	6	0	4
9306	SOC56	0	0	16	0,99	8	16	0	8	13	0,99	10
6674	SPAG1	1	1	21	0,99	7	9	0	9	11	0,99	8
9789	SPCS2	1	0	20	1,00	6	2	0	2	26	1,00	9
60559	SPCS3	2	2	18	0	7	10	0	4	27	0,78	9
200734	SPRED2	0	0	27	1,00	9	8	0	5	23	1,00	14
81848	SPRY4	0	0	2	0	1	16	1,00	6	5	0	4
6709	SPTAN1	12	8	327	1,00	128	340	1,00	157	90	1,00	46
6711	SPTBN1	5	5	306	1,00	104	214	1,00	124	293	1,00	140
6712	SPTBN2	0	0	90	1,00	49	67	1,00	45	147	1,00	84
6714	SRC	0	0	302	1,00	49	142	1,00	46	116	1,00	39
23380	SRGAP2	3	3	44	0,76	19	50	0,81	31	62	1,00	34
6723	SRM	4	2	23	1,00	13	34	1,00	16	51	0,98	15
6726	SRP9	1	1	39	1,00	12	14	1,00	5	14	0,99	5
6427	SRSF2	3	3	36	0,78	8	7	0	3	9	0	6
6745	SSR1	0	0	2	1	0	0	0	0	21	1,00	5
10948	STARDB3	0	0	7	1	1	1	1	1	11	0,99	5
261729	STEAP2	0	0	72	1,00	24	23	0,99	13	54	1,00	22
55240	STEAP3	3	2	242	1,00	41	372	1,00	35	225	1,00	36
6786	STIM1	4	4	15	0	7	66	0,92	23	83	1,00	35
57620	STIM2	0	0	26	1,00	10	31	1,00	19	37	1,00	19
6793	STK10	0	0	62	1,00	26	31	1,00	16	50	1,00	21
6794	STK11	0	0	54	1,00	11	22	1,00	11	16	1,00	8
114790	STK11IP	0	0	0	0	0	0	0	0	16	1,00	13
2040	STOM	0	0	45	1,00	16	43	1,00	18	41	1,00	19
11037	STON1	0	0	51	1,00	18	16	1,00	10	28	1,00	14
92335	STRADA	0	0	14	0,99	8	11	0,99	8	12	1,00	4
3703	STT3A	4	4	71	1,00	23	59	0,99	20	37	1,00	17
201595	STT3B	1	1	2	1	1	25	1,00	13	26	1,00	12
8677	STX10	1	0	104	1,00	24	48	1,00	18	60	1,00	21
23673	STX12	1	1	190	1,00	31	68	1,00	17	85	1,00	26
2054	STX2	0	0	6	0	4	12	0	9	10	0,99	8
6809	STX3	0	0	61	1,00	19	5	0	5	14	0	7
6810	STX4	0	0	44	1,00	18	7	0	5	11	0	8
6811	STX5	4	2	4	0	3	49	0,99	25	41	0,94	21
10228	STX6	0	0	109	1,00	19	26	1,00	14	38	1,00	20
8417	STX7	4	3	213	1,00	32	59	0,99	19	99	1,00	26
9482	STX8	1	1	53	1,00	17	41	1,00	19	56	1,00	16
6812	STXBP1	1	0	86	1,00	34	26	1,00	19	26	1,00	16
6813	STXBP2	1	1	18	0	9	15	1,00	11	17	1,00	12
6814	STXBP3	1	0	123	1,00	42	79	1,00	30	87	1,00	38
134957	STXBP5	1	0	17	1,00	11	21	1,00	15	8	0	7
29091	STXBP6	0	0	13	0,99	7	16	0	8	12	0	6
6817	SULT1A1	4	3	9	0	8	64	0,98	26	63	1,00	22
94121	SYTL4	0	0	79	1,00	29	47	1,00	31	41	1,00	25
8407	TAGLN2	44	44	853	1,00	47	267	1,00	31	218	1,00	25
9344	TAOK2	0	0	0	0	0	10	0,99	8	3	0	3
51347	TAOK3	2	1	41	0,99	22	33	0,84	21	38	0,99	21
83874	TBC1D10A	0	0	83	1,00	22	20	1,00	13	38	1,00	17
26000	TBC1D10B	3	1	74	1,00	26	88	1,00	33	53	0,98	24
55633	TBC1D22B	4	1	20	1	11	37	0,81	24	21	1,00	14
55773	TBC1D23	1	0	0	0	0	0	0	0	11	0,99	9
123036	TC2N	0	0	79	1,00	24	28	1,00	20	28	1,00	18
6950	TCP1	59	58	713	1,00	109	475	1,00	84	326	1,00	80
11022	TDKPH	1	0	0	0	0	0	0	0	23	1,00	14
9524	TECR	4	4	68	1,00	22	51	1,00	17	30	1,00	15
7037	TERC	2	0	116	1,00	47	241	1,00	75	175	1,00	74
55901	THSD1	0	0	1	0	1	15	0,99	10	17	1,00	11
7074	TIAM1	0	0	100	1,00	42	45	1,00	30	73	1,00	32
26517	TIMM13	0	0	31	1,00	12	9	0	6	4	0	2
100287932	TIMM23	6	3	77	0,99	17	20	1,00	11	33	1,00	15
10469	TIMM44	2	2	32	0,89	9	3	0	3	16	1,00	11
92609	TIMM50	11	9	172	1,00	26	138	1,00	24	112	0,78	25
10972	TMED10	2	1	24	0,91	5	10	0	4	11	0	8
23670	TMEM2	0	0	439	1,00	95	264	1,00	80	170	1,00	69
65062	TMEM237	0	0	16	0,99	4	26	1,00	11	24	1,00	11
55161	TMEM33	3	3	31	1,00	12	51	0,99	17	44	0,93	15
55706	TMEM48	4	3	10	0	6	67	1,00	34	62	0,99	28
55092	TMEM51	0	0	36	1,00	12	25	1,00	7	37	1,00	11
25963	TMEM87A	1	1	38	1,00	16	60	1,00	20	20	0,99	12
27346	TMEM97	2	1	21	0,84	6	23	0,91	4	17	0,76	5
7110	TMF1	1	1	0	0	0	26	1,00	18	0	0	0
81542	TMX1	1	1	6	0	3	5	0	4	21	1,00	10

25816	TNFAIP8	2	1	11	0	7	22	0,88	8	0	0	0
8795	TNFRSF10B	0	0	32	1,00	12	10	0	5	40	1,00	14
23043	TNIK	6	4	71	0,98	26	72	0	25	89	0,95	39
85456	TNKS1BP1	4	4	21	1	11	33	0	23	69	0,91	46
114034	TOE1	1	1	7	0	4	22	1,00	13	16	0,99	8
56993	TOMM22	8	8	117	1,00	21	148	1,00	20	127	1,00	21
100188893	TOMM6	0	0	16	0,99	7	0	0	0	1	0	1
26092	TOR1AIP1	4	3	10	0	6	53	0,93	27	38	1,00	23
7157	TP53	3	3	0	0	0	43	0,86	11	0	0	0
84231	TRAF7	0	0	0	0	0	15	1,00	11	10	0,98	8
10131	TRAP1	24	23	265	0	72	253	0	79	299	1,00	81
10206	TRIM13	2	2	0	0	0	16	0	8	60	1,00	17
7204	TRIO	3	3	4	0	2	40	0	25	84	1,00	48
11078	TRIOBP	1	0	39	1,00	21	35	1,00	19	53	1,00	22
9319	TRIP13	8	5	91	0,95	31	84	0	38	60	1,00	34
54795	TRPM4	0	0	1	0	1	21	1,00	13	19	1,00	12
125488	TTC39C	0	0	11	0,99	8	64	1,00	19	0	0	0
145567	TTCTB	1	1	66	1,00	21	91	1,00	42	60	1,00	27
7284	TUFM	26	23	347	1,00	67	315	1,00	61	217	1,00	55
7289	TULP3	0	0	120	1,00	28	69	1,00	23	59	1,00	24
7298	TYMS	2	2	36	0,98	11	2	0	3	14	1,00	

Annexe 2 : Tableau des interacteurs de RAS identifiés par la technique d'AP-MS

Gene ID	Gene Name	Top 2 controls	HEK293 FB-HRAS WT			HEK293 FB-NRAS WT			HEK293 FB-KRAS WT		
			Total	SAINT	Total Unique	Total	SAINT	Total Unique	Total	SAINT	Total Unique
948469	brA	0 0	26073	-	110	20183	-	108	17085	-	98
3265	HRAS	0 0	693	-	13	31	-	3	35	-	4
4893	NRAS	0 0	362	-	10	3198	-	34	9	-	4
3845	KRAS	0 0	1	-	1	0	-	0	1150	-	24
65985	AACS	0 0	32	1,00	19	2	-	2	7	-	5
14	AAMP	0 0	39	1,00	13	40	1,00	17	29	0,75	11
16	AARS	0 0	76	1,00	32	67	1,00	32	66	1,00	30
57505	AARS2	11 8	155	1,00	52	116	1,00	51	147	0,97	51
4363	ABCC1	1 0	22	1,00	13	7	-	7	8	-	6
10257	ABCC4	1 1	47	1,00	28	7	-	5	1	-	1
55324	ABCF3	0 0	14	0,75	10	14	-	7	16	-	8
26090	ABHD12	1 0	24	1,00	9	8	-	5	14	-	6
30	ACAA1	2 0	36	0,98	12	40	0,98	12	38	0,91	10
39	ACAT2	12 8	232	1,00	27	186	1,00	24	204	1,00	26
64746	ACBD3	0 0	25	1,00	10	29	1,00	10	43	1,00	19
84320	ACBD6	0 0	6	-	2	13	-	7	17	1,00	10
47	ACLY	3 3	856	1,00	123	783	1,00	131	1066	1,00	132
10965	ACOT2	2 0	72	0,99	27	50	0,97	16	53	0,99	22
2180	ACSL1	2 2	66	1,00	34	45	1,00	23	51	1,00	30
55860	ACTR10	1 1	27	1,00	10	23	1,00	12	33	1,00	12
10120	ACTR1B	0 0	1	-	1	2	-	1	23	0,99	8
10097	ACTR2	0 0	33	1,00	11	47	1,00	14	39	1,00	10
10096	ACTR3	0 0	44	1,00	18	43	1,00	18	44	1,00	17
102	ADAM10	1 0	11	0,99	7	0	-	0	1	-	1
79934	ADCK4	2 0	15	-	7	21	-	8	38	0,99	12
83440	ADPGK	3 2	71	0,99	17	30	-	14	51	0,78	16
54936	ADPRHL2	0 0	11	-	8	13	0,99	8	16	1,00	11
11047	ADRM1	2 2	26	0,76	8	25	0,77	8	18	-	7
158	ADSL	7 7	344	1,00	51	334	1,00	50	334	1,00	49
178	AGL	0 0	24	1,00	15	23	1,00	13	12	0,99	9
10554	AGPAT1	0 0	16	0,99	10	3	-	2	1	-	1
8540	AGPS	2 0	142	1,00	35	140	1,00	36	145	1,00	37
191	AHCY	1 1	271	1,00	55	252	1,00	49	326	1,00	54
23382	AHCYL2	0 0	23	1,00	13	13	-	7	14	0,75	6
9590	AKAP12	1 0	9	0,99	5	12	-	7	6	-	5
208	AKT2	1 0	7	-	6	15	1,00	8	12	0,99	8
214	ALCAM	1 0	37	1,00	16	0	-	0	22	-	8
126133	ALDH16A1	2 1	132	1,00	28	107	1,00	28	135	1,00	28
5832	ALDH18A1	8 5	152	0,98	50	134	0,97	39	161	1,00	45
2919	ALDH1B1	1 0	26	1,00	13	13	0,99	8	14	-	10
715	ALDH5A1	0 0	16	1,00	6	16	0,99	8	13	0,99	5
4329	ALDH6A1	0 0	8	-	5	15	0,99	10	13	0,99	8
501	ALDH7A1	0 0	67	1,00	25	42	1,00	20	36	1,00	17
223	ALDH9A1	0 0	66	1,00	24	69	1,00	24	69	1,00	23
85365	ALG2	0 0	7	-	5	17	1,00	9	16	1,00	8
29880	ALG5	0 0	76	1,00	17	52	1,00	15	49	1,00	9
29929	ALG6	0 0	16	1,00	6	5	-	3	5	-	2
64682	ANAPC1	1 1	49	1,00	21	61	1,00	27	66	1,00	29
51433	ANAPC5	0 0	44	1,00	17	35	1,00	21	51	1,00	20
51434	ANAPC7	1 1	31	1,00	15	42	1,00	20	41	1,00	20
288	ANK3	0 0	20	1,00	13	5	-	4	23	1,00	12
23141	ANKLE2	0 0	13	-	8	14	0,75	11	12	-	11
283373	ANKRD52	1 1	9	0,99	7	4	-	4	19	1,00	11
81611	ANP32E	8 8	153	0,99	18	161	1,00	16	211	0,76	17
311	ANXA11	0 0	46	1,00	21	51	1,00	24	68	1,00	29
307	ANXA4	0 0	32	1,00	20	40	1,00	21	53	1,00	22
162	AP1B1	9 1	94	0,95	36	51	-	28	138	0,91	47
164	AP1G1	2 0	65	1,00	32	61	1,00	30	82	1,00	34
8907	AP1M1	8 8	98	-	37	90	-	37	132	0,99	45
160	AP2A1	7 5	79	-	34	66	-	24	93	0,93	36
161	AP2A2	5 1	70	-	25	58	0,83	21	85	0,98	34
163	AP2B1	14 8	176	0,95	48	141	-	51	226	1,00	59
8546	AP3B1	0 0	55	1,00	33	75	1,00	43	224	1,00	76
26985	AP3M1	0 0	47	1,00	20	48	1,00	20	74	1,00	25
317	APAF1	0 0	2	-	1	17	-	10	25	0,99	11
8539	APIS	0 0	107	1,00	37	62	1,00	23	100	1,00	31
79135	APOO	0 0	14	1,00	8	7	-	4	3	-	2
9716	AQR	0 0	9	-	4	15	0,99	10	23	-	14
369	ARAF	3 2	57	1,00	19	63	1,00	19	67	0,97	21
9138	ARHGEF1	0 0	42	1,00	23	28	1,00	20	72	1,00	32
8874	ARHGEF7	0 0	10	-	5	15	1,00	10	11	0,99	7
10425	ARH2	0 0	21	1,00	11	36	1,00	17	16	0,75	6
402	ARL2	0 0	30	1,00	11	35	1,00	13	49	1,00	16
10550	ARL6IP5	0 0	134	1,00	10	52	1,00	10	77	1,00	10
55207	ARL8B	2 2	76	1,00	18	54	1,00	20	48	1,00	16
83787	ARMC10	0 0	25	1,00	11	18	1,00	8	12	-	5
10094	ARPC3	0 0	65	1,00	10	62	1,00	12	59	1,00	10
51130	ASB3	0 0	6	-	3	7	-	4	12	0,99	5
84164	ASCC2	0 0	15	1,00	7	5	-	3	15	1,00	9
439	ASNA1	0 0	74	1,00	14	25	1,00	11	63	1,00	15
444	ASPH	1 1	13	-	5	10	-	4	18	1,00	6
445	ASS1	0 0	26	1,00	8	31	1,00	12	29	1,00	11
23192	ATG4B	0 0	20	1,00	12	21	1,00	11	22	0,75	9
9474	ATG5	0 0	9	-	5	16	-	8	20	1,00	12
10533	ATG7	0 0	58	1,00	20	182	1,00	21	34	1,00	7
79065	ATG9A	0 0	23	1,00	12	11	0,99	7	9	-	5
51062	ATL1	0 0	19	1,00	10	8	-	6	8	-	6
64225	ATL2	1 0	36	1,00	16	27	1,00	13	44	1,00	14
472	ATM	0 0	20	1,00	10	24	1,00	14	26	1,00	14
286410	ATP11C	0 0	39	1,00	18	5	-	4	4	-	3
57130	ATP13A1	0 0	22	1,00	11	4	-	4	8	-	5
537	ATP6AP1	3 3	64	0,99	17	37	-	18	41	-	14
10159	ATP6AP2	2 0	34	0,95	17	5	-	3	15	-	9
523	ATP6V1A	2 2	130	1,00	32	74	1,00	27	134	1,00	34
545	ATR	2 0	20	0,87	12	14	-	9	9	-	6
25814	ATXN10	8 7	161	1,00	35	135	1,00	29	165	1,00	35
11041	B3GNT1	1 0	43	-	10	19	-	8	16	0,75	10
374383	BTH6	0 0	15	0,99	7	2	-	2	5	-	4
7917	BAG6	5 4	115	1,00	33	30	-	16	63	0,96	25
8815	BANF1	0 0	17	1,00	6	28	1,00	7	17	0,99	5
581	BAX	1 0	49	1,00	12	17	1,00	10	26	1,00	10
10134	BCAP31	2 2	44	1,00	14	19	-	11	28	-	10
56647	BCCIP	5 3	55	0,78	17	59	0,85	18	88	1,00	22
617	BCS1L	1 1	94	1,00	21	36	1,00	10	51	1,00	13
57448	BIRC6	0 0	47	1,00	32	28	1,00	20	38	1,00	27
644	BLVRA	0 0	158	1,00	33	185	1,00	35	143	1,00	27
662	BNIP1	1 0	22	1,00	9	6	-	4	11	0,99	6
673	BRAF	2 0	141	1,00	42	162	1,00	40	157	1,00	40
221927	BRAT1	1 1	81	1,00	23	43	1,00	20	63	1,00	21
9577	BRE	0 0	34	1,00	6	34	1,00	5	24	0,75	3
9044	BTAF1	2 2	69	1,00	32	77	1,00	40	108	1,00	40
701	BUB1B	0 0	17	0,83	13	22	0,88	10	25	0,87	11
9689	BZW1	1 1	31	1,00	17	22	1,00	14	45	1,00	20
28969	BZW2	0 0	55	1,00	19	26	1,00	10	52	1,00	15
26098	C10orf137	0 0	36	0,75	12	18	-	6	39	1,00	12
738	C11orf2	1 1	35	1,00	20	39	1,00	20	34	1,00	19
84520	C14orf142	0 0	22	1,00	6	11	-	3	18	-	4
161424	C14orf21	0 0	20	1,00	8	1	-	1	7	-	4
57020	C16orf62	0 0	35	1,00	16	25	1,00	12	47	1,00	17
64149	C17orf75	4 2	51	0,99	15	58	0,99	18	39	-	12
29919	C18orf8	0 0	13	-	7	21	1,00	11	23	1,00	12
148223	C19orf25	1 1	31	1,00	7	21	1,00	5	28	0,75	4
55732	C1orf12	0 0	24	1,00	11	12	-	5	11	-	6
57136	C20orf3	18 16	215	0,99	48	133	-	44	167	-	41
8209	C21orf33	1 0	25	-	7	21	0,99	8	22	1,00	7
60526	C2orf43	0 0	10	-	3	13	0,99	4	14	0,75	5
54859	C3orf75	0									

Annexes | Annexe 2 : Tableau des interacteurs de RAS identifiés par la technique d'AP-MS

1445	CSK	1	0	98	1,00	27	105	1,00	33	73	1,00	30	57700	FAM160B1	0	0	12	0,99	8	2	2	13	0,99	8	
1457	CSNK2A1	1	1	7		4	22	1,00	9	26	1,00	9	728071	FAM203B	0	0	20	1,00	11	26	1,00	15	35	1,00	19
1477	CSTF1	0	0	19	1,00	11	24	1,00	12	25		7	84293	FAM213A	0	0	70	1,00	15	30	1,00	11	47	1,00	11
1478	CSTF2	0	0	81	1,00	28	22	1,00	11	41	1,00	14	10447	FAM3C	3	1	47	0,98	13	23	0,81	7	37	0,91	8
1479	CSTF3	0	0	16		9	18		8	27	1,00	14	51571	FAM49B	1	1	38	1,00	18	21	1,00	15	33	1,00	16
1487	CTBP1	0	0	106	1,00	21	90	1,00	21	130	1,00	21	51115	FAM82B	0	0	29	1,00	12	40	1,00	17	32	1,00	12
1488	CTBP2	1	1	62	1,00	23	95	1,00	28	95	1,00	26	2177	FANCD2	0	0	115	1,00	37	46	0,98	28	96	1,00	35
1495	CTNNA1	10	6	150	0,85	59	52		25	89		36	55215	FANCI	16	9	425	1,00	106	251	1,00	94	323	1,00	99
8727	CTNNA1	0	0	23	1,00	8	2		1	13	0,99	8	10160	FARP1	0	0	9		4	7		5	10	0,99	5
1499	CTNNB1	1	1	54	1,00	27	18	0,99	9	5		3	2193	FARSA	7	6	106	0,84	28	108	1,00	31	139	0,97	35
56259	CTNNBL1	1	0	35	1,00	13	46	1,00	16	66	1,00	19	10056	FARSB	0	0	85	1,00	28	94	1,00	34	103	1,00	34
56474	CTPS2	1	0	78	1,00	23	86	1,00	27	114	1,00	26	22868	FASTKD2	3	2	71	1,00	20	41		18	52		19
8453	CUL2	3	3	110	1,00	40	57	1,00	27	84	1,00	32	60493	FASTKD5	3	1	45	0,84	12	37	0,83	13	40		11
8452	CUL3	1	0	12		7	18	1,00	13	42	1,00	18	115290	FBXO17	0	0	15	1,00	7	7		4	11	0,99	5
8451	CUL4A	0	0	45	1,00	14	118	1,00	21	112	1,00	22	26263	FBXO22	2	2	53	1,00	18	59	1,00	20	36		15
8450	CUL4B	0	0	88	1,00	26	146	1,00	33	175	1,00	42	26272	FBXO4	0	0	7		5	11		6	17	0,99	6
8065	CUL5	2	2	68	1,00	41	74	1,00	39	81	1,00	42	25793	FBXO7	0	0	18	1,00	8	20		7	20	1,00	6
9820	CUL7	0	0	2		1	2		2	14	0,99	8	2222	FDFT1	6	5	170	1,00	31	124	1,00	26	199	1,00	34
1528	CYBSA	2	1	46	0,99	15	18	0,84	10	26	0,93	13	2224	FDP5	5	5	90	1,00	21	96	1,00	21	147	1,00	20
1727	CYBSR3	13	13	199	0,75	37	61		23	107		28	2235	FECH	5	4	71		20	54		15	65	0,89	16
23191	CYFIP1	1	0	154	1,00	58	154	1,00	55	139	1,00	55	9158	FIBP	0	0	25	1,00	11	25	1,00	12	13		8
26999	CYFIP2	0	0	19		10	69	1,00	27	50	0,75	18	51024	FIS1	1	0	62	1,00	9	25	1,00	9	18	1,00	8
1595	CYP51A1	4	4	172	1,00	35	73	1,00	23	98	1,00	28	51303	FKBP11	1	1	21	1,00	8	9	0,99	4	6		3
1605	DAG1	0	0	14	0,99	6	0		0	4		3	2288	FKBP4	54	48	858	0,75	81	934	1,00	70	1020		72
55157	DARS2	2	0	100	1,00	34	92	1,00	34	90	1,00	32	80308	FLAD1	0	0	61	1,00	18	42	1,00	19	74	1,00	20
26528	DAZAP1	2	2	52	1,00	11	57	1,00	12	63	0,87	12	10211	FLOT1	0	0	24	1,00	14	0		0	2		1
28988	DBNL	0	0	23	0,75	10	22	1,00	9	24	1,00	13	2319	FLOT2	2	1	55	1,00	25	19	0,79	10	16		9
79877	DCAKD	3	2	54	0,99	14	33	0,82	8	40	0,90	12	6624	FSCN1	0	0	93	1,00	27	106	1,00	27	68	1,00	25
1635	DCTD	1	0	18	1,00	7	25	1,00	11	14	1,00	7	2539	G6PD	0	0	5		4	11	0,99	6	9		5
84259	DCUN1D5	1	1	47	1,00	19	41	1,00	15	67	1,00	17	2580	GAK	0	0	21		8	23	1,00	11	31	1,00	13
23576	DDAH1	0	0	15	1,00	12	9		8	13		7	2581	GALC	0	0	13	0,99	8	0		0	6		4
84301	DDI2	0	0	69	1,00	22	40	1,00	16	56	1,00	20	2582	GALE	0	0	38	1,00	16	19		11	34	1,00	16
1656	DDX6	0	0	33	1,00	17	39	1,00	17	35	1,00	11	2589	GALNT1	0	0	15	1,00	10	1		1	3		2
8560	DEGS1	1	0	33	1,00	9	4		3	9		4	2590	GALNT2	1	0	33	1,00	20	0		0	14	0,75	7
51071	DERA	2	2	91	1,00	26	68	1,00	23	99	1,00	25	23193	GANAB	3	2	61	1,00	26	46	0,94	27	45		21
51635	DHR57	0	0	61	1,00	18	11		6	11		9	26130	GAPVD1	0	0	56	1,00	29	54	1,00	35	81	1,00	35
79665	DHX40	2	1	18	0,84	8	27	0,95	8	33	0,97	7	2617	GARS	0	0	11		6	10	0,99	8	22	1,00	10
1729	DIAPH1	3	3	82	1,00	38	73	1,00	31	93	1,00	35	2618	GART	15	12	667	1,00	90	613	1,00	93	588	1,00	89
27292	DIMT1	2	0	5		2	14		10	18	0,81	12	8729	GBF1	1	1	266	1,00	81	127	1,00	51	192	1,00	59
1738	DLD	2	1	58	1,00	9	23	0,92	11	54	0,99	9	25801	GCA	0	0	24	1,00	12	18	1,00	9	27		11
3337	DNAJB1	0	0	0		0	17	1,00	7	23	1,00	7	23464	GCAT	0	0	11	0,99	8	8		6	4		2
7266	DNAJC7	6	5	133	1,00	32	96	1,00	30	125	1,00	30	6936	GCF2	0	0	6		4	5		4	21	1,00	10
10059	DNM1L	1	0	29	1,00	19	13		10	26	1,00	16	10985	GCN1L1	94	79	2214	1,00	332	1390	1,00	291	1598	1,00	287
1785	DNM2	0	0	112	1,00	49	89	1,00	37	130	1,00	45	284161	GDPD1	0	0	32	1,00	13	4		4	11		8
85440	DOCK7	1	1	29	1,00	14	36	1,00	16	54	1,00	27	8487	GEMIN2	0	0	10	0,99	5	0		0	4		3
9980	DOPEY2	1	1	20	1,00	9	17	1,00	12	36	1,00	20	51608	GET4	1	0	20	1,00	7	1		1	10		3
51611	DPH5	0	0	27	1,00	8	42	1,00	14	26	1,00	9	85476	GFM1	12	9	107		35	121		39	144	0,90	41
1808	DPYSL2	4	1	44		23	29		20	57	0,85	27	84340	GFM2	1	0	24	1,00	12	16	0,75	8	29	1,00	11
1809	DPYSL3	0	0	14		7	13		9	21	1,00	13	2673	GFPT1	1	0	47	1,00	26	64	1,00	32	65	1,00	34
1819	DRG2	0	0	9		6	19	1,00	9	36	1,00	14	2677	GGCX	0	0	24	1,00	9	7		4	8		5
11034	DSTN	3	3	59	1,00	10	62	0,99	12	51		10	28964	GIT1	0	0	17	1,00	9	9		4	20	1,00	9
56931	DUS3L	0	0	10		5	13	0,99	10	18	1,00	10	2710	GK	3	0	16		6	29	0,80	17	13		7
11266	DUSP12	1	0	13		8	15		9	14	1,00	7	89944	GLB1L2	0	0	20	1,00	7	4		4	2		2
1783	DYNC1L12	3	3	72	0,99	24	60	1,00	26	93	0,99	30	11146	GLMN	0	0	99	1,00	37	45	1,00	23	52	1,00	22
79659	DYNC2H1	4	4	35		28	87	1,00	52	75	0,99	40	29926	GMPPA	0	0	17	1,00	10	25	1,00	10	23	1,00	12
124454	EARS2	0	0	24	0,75	9	29	1,00	11	19	1,00	10	8833	GMPS	37	31	528	0,97	79	546	1,00	83	568	1,00	79
1889	ECE1	0	0	23	1,00	11	2		2	2		2	10672	GNAI3	0	0	42	1,00	14	0		0	8		4
1892	ECHS1	0	0	14	0,99	6	7		2	7		4	2771	GNAI2	3	1	28	0,79	11	10		3	7		3
80153	EDC3	0	0	12	0,99	6	12	0,99	8	28	1,00	12	2778	GNAS	5	4	96	1,00	31	73	0,97	29	69	1,00	27
23644	EDC4	1	0	51	1,00	19	60	1,00	20	55	1,00	14	10020	GNE	0	0	18		9	44	1,00	19	59	1,00	26
1947	EFNB1	0	0	16	1,00	8	8		6	8		6	8443	GNPAT	2	1	85	1,00	28	67	1,00	22	65	0,99	26
22979	EFR3B	0	0	27	1,00	11	7		3	2		1	64841	GNPNAT1	0	0	50	1,00	12	47	1,00	12	42	1,00	10
79631	EFTUD1	1	1	15	1,00	9	32	1,00	14	35	1,00	14	51280	GOLM1	1	0	22	1,00	15	0		0	1		1
30844	EHD4	0	0	43	1,00	20	30	1,00	16	35	1,00	16	64083	GOLPH3	4	2	66	0,95	11	76	0,98	9	76	0,87</	

Annexes | Annexe 2 : Tableau des interacteurs de RAS identifiés par la technique d'AP-MS

81502	HM13	2	1	24	0,83	10	4	4	7	6		
3163	HMOX2	4	3	93	1,00	24	43	19	75	0,98	23	
51170	HSD17B11	0	0	24	1,00	12	5	4	1	1	1	
51144	HSD17B12	13	7	178	0,95	33	108	31	107	25		
3295	HSD17B4	0	0	50	1,00	28	64	1,00	26	48	1,00	23
83693	HSDL1	0	0	28	1,00	12	24	1,00	10	19	1,00	10
3320	HSP90AA1	98	96	1369	0,75	97	1751	1,00	106	1423	1,00	92
51182	HSPA14	0	0	39	1,00	13	29	1,00	11	38	1,00	11
51668	HSPB11	3	1	57	0,99	10	64	1,00	12	49	0,96	12
23640	HSPBP1	0	0	84	1,00	21	35	1,00	17	70	1,00	19
27336	HTATSF1	0	0	23	1,00	11	23	1,00	12	30	1,00	13
3064	HTT	0	0	68	1,00	43	48	1,00	29	88	1,00	46
3417	IDH1	0	0	9	5	22	1,00	11	13	0,75	7	
3418	IDH2	0	0	109	1,00	32	138	1,00	42	128	1,00	38
3419	IDH3A	7	6	114	1,00	32	120	1,00	33	138	1,00	31
3420	IDH3B	3	3	48	0,93	22	29	20	44	22		
3421	IDH3G	1	0	25	1,00	12	29	1,00	12	38	1,00	14
3422	IDH1	0	0	80	1,00	18	51	1,00	16	44	1,00	17
51124	IER3P1	2	0	19	0,80	6	5	3	13	6		
11020	IFT27	0	0	11	0,99	6	6	4	15	8		
93185	IGSF8	0	0	24	1,00	14	0	0	1	1		
8518	IKBKAP	5	3	46	25	49	0,85	28	68	28		
3551	IKBKB	0	0	4	2	22	1,00	14	28	1,00	13	
3608	ILF2	29	27	378	1,00	48	414	1,00	49	448	1,00	49
3611	ILK	0	0	22	1,00	11	28	1,00	11	23	1,00	8
80895	ILKAP	0	0	50	1,00	17	63	1,00	19	63	1,00	22
3612	IMP1	0	0	14	1,00	10	14	9	11	0,99	10	
54928	IMPAD1	0	0	36	1,00	15	15	0,99	6	21	1,00	7
57508	INTS2	1	0	25	1,00	11	40	1,00	14	28	1,00	14
65123	INTS3	2	1	47	0,99	14	56	1,00	18	87	1,00	22
92105	INTS4	0	0	21	11	29	1,00	16	24	11		
80789	INTS5	0	0	12	3	26	1,00	9	13	8		
25896	INTS7	0	0	38	1,00	23	43	1,00	23	22	1,00	11
55756	INTS9	0	0	9	5	12	8	29	1,00	7		
51194	IPO11	2	2	327	1,00	57	108	1,00	37	157	1,00	39
9670	IPO13	1	0	40	1,00	17	16	1,00	6	26	1,00	10
79711	IPO4	29	27	535	1,00	96	355	0,77	72	672	1,00	96
3843	IPO5	37	33	707	1,00	89	382	70	593	0,75	71	
10527	IPO7	27	26	572	1,00	69	352	0,92	55	656	1,00	77
10526	IPO8	6	5	171	1,00	49	106	0,99	42	150	1,00	44
55705	IPO9	15	12	353	1,00	63	264	1,00	57	478	1,00	72
26145	IRF2BP1	0	0	13	0,99	5	12	0,99	4	11	0,99	4
3673	ITGA2	0	0	24	1,00	15	0	0	6	4		
3655	ITGA6	0	0	12	0,99	7	0	0	4	3		
3685	ITGAV	0	0	17	1,00	8	1	1	1	1		
3728	JUP	1	1	118	1,00	38	74	1,00	27	82	1,00	30
3735	KARS	6	6	105	34	99	0,93	33	172	1,00	45	
10300	KATNB1	2	1	23	0,92	10	8	4	29	0,89	11	
115207	KCTD12	0	0	19	0,75	6	7	5	22	1,00	9	
51780	KDM3B	0	0	17	1,00	9	15	1,00	9	29	1,00	13
2531	KDSR	0	0	25	1,00	10	9	0,99	4	20	1,00	8
8570	KHSRP	4	4	53	0,77	22	43	17	50	21		
23065	KIAA0090	5	4	159	1,00	54	50	26	97	0,99	42	
23392	KIAA0368	10	6	624	1,00	132	296	1,00	89	312	1,00	87
9847	KIAA0528	0	0	15	0,75	8	22	1,00	14	32	1,00	14
23277	KIAA0664	0	0	152	1,00	50	115	1,00	52	142	1,00	52
26128	KIAA1279	0	0	71	1,00	28	68	1,00	24	71	1,00	24
57614	KIAA1468	0	0	75	1,00	27	41	1,00	12	66	1,00	20
57650	KIAA1524	0	0	38	1,00	17	20	1,00	14	36	1,00	18
80856	KIAA1715	0	0	35	1,00	18	10	0,99	10	17	1,00	12
54914	KIAA1797	0	0	13	0,75	7	12	8	13	7		
9735	KNTC1	3	3	108	1,00	53	82	1,00	50	100	1,00	51
3836	KPNA1	2	1	44	0,93	7	39	0,94	8	60	0,99	9
3838	KPNA2	13	13	281	1,00	54	220	1,00	47	239	1,00	48
3839	KPNA3	1	0	27	1,00	16	64	1,00	22	71	1,00	21
112970	KTI12	0	0	21	1,00	9	18	1,00	9	21	0,75	10
10314	LANCL1	1	0	24	1,00	14	15	1,00	9	18	1,00	9
51056	LAP3	0	0	27	1,00	11	35	1,00	17	37	1,00	15
113251	LARP4	1	0	11	4	6	4	23	1,00	7		
3927	LASP1	1	1	10	0,99	5	5	3	12	6		
253558	LCLAT1	0	0	20	1,00	13	8	6	18	1,00	8	
51451	LCMT1	0	0	34	1,00	10	28	1,00	9	13	0,7	7
3939	LDHA	3	2	86	0,93	21	101	1,00	22	109	0,97	23
3978	LIG1	0	0	33	1,00	18	25	1,00	13	41	1,00	18
3980	LIG3	0	0	11	0,99	7	10	7	34	1,00	13	
3998	LMAN1	4	3	86	1,00	29	13	8	33	18		
91289	LMF2	1	0	113	1,00	21	42	1,00	13	56	1,00	13
4012	LNPEP	1	1	86	1,00	34	33	1,00	12	61	1,00	24
100507855	LOC100507855	0	0	54	1,00	19	37	1,00	18	37	1,00	14
9361	LONP1	9	4	178	1,00	53	144	1,00	49	202	0,99	53
23266	LPNH2	1	0	36	1,00	9	13	5	6	1		
987	LRBA	0	0	120	1,00	58	90	1,00	52	92	1,00	48
55631	LRRC40	1	0	57	1,00	24	72	1,00	34	57	1,00	22
10489	LRRC41	0	0	8	4	12	0,99	6	23	1,00	12	
51599	LSR	0	0	21	1,00	10	9	6	23	4		
4047	LSS	0	0	87	1,00	35	32	1,00	17	49	1,00	25
4048	LT4AH	0	0	120	1,00	43	113	1,00	48	124	1,00	43
26046	LTN1	8	5	112	0,84	38	83	40	131	0,78	53	
10434	LYPLA1	0	0	20	1,00	5	19	1,00	8	14	0,99	4
11313	LYPLA2	2	0	54	0,98	6	31	0,89	6	33	0,77	7
10892	MALT1	2	1	45	1,00	14	35	0,97	12	59	0,99	16
4121	MAN1A1	1	0	12	0,99	8	3	3	2	2		
4124	MAN2A1	0	0	17	1,00	9	5	2	10	5		
55201	MAP1S	1	0	14	0,99	8	17	1,00	13	18	1,00	12
6416	MAP2K4	0	0	30	1,00	13	14	0,99	9	16	1,00	12
92935	MARS2	1	1	20	1,00	8	21	1,00	11	18	1,00	5
57506	MAVS	1	1	14	7	6	5	20	1,00	8		
153364	MBLAC2	0	0	14	0,99	10	5	4	10	6		
4171	MCM2	0	0	28	1,00	12	34	1,00	18	20	1,00	12
4173	MCM4	8	6	272	1,00	56	264	1,00	63	298	1,00	57
4175	MCM6	15	14	233	0,98	55	249	1,00	61	257	0,77	62
28985	MCTS1	0	0	67	1,00	16	83	1,00	17	41	1,00	15
10025	MED16	0	0	10	4	7	5	23	1,00	11		
9477	MED20	0	0	11	0,99	5	9	4	17	1,00	6	
9439	MED23	0	0	93	1,00	41	65	1,00	31	46	1,00	23
9862	MED24	0	0	68	1,00	29	25	1,00	10	17	9	
90390	MED30	0	0	5	3	13	0,99	8	5	3		
51072	MEMO1	1	0	30	1,00	13	26	1,00	14	30	1,00	15
4233	MET	2	0	30	0,91	11	2	2	8	5		
10724	MGEA5	2	1	131	1,00	32	143	1,00	42	128	1,00	36
4255	MGMT	0	0	22	1,00	6	12	5	13	0,99	5	
375056	MIA3	0	0	16	1,00	10	0	0	0	0	0	
253714	MMS22L	0	0	56	1,00	22	7	5	17	1,00	11	
27304	MOC3	0	0	16	9	17	1,00	9	25	1,00	10	
158747	MOSPPD2	0	0	36	1,00	17	9	5	5	3		
4350	MPG	0	0	1	5	4	26	1,00	10			
4351	MPI	0	0	48	1,00	17	29	1,00	14	34	1,00	13
4436	MSH2	9	6	185	1,00	55	155	1,00	52	202	1,00	53
2956	MSH6	10	8	184	1,00	57	190	1,00	59	275	1,00	68
124540	MSI2	1	0	22	1,00	9	6	2	18	0,75	9	
51765	MST4	1	0	23	1,00	9	54	1,00	15	61	1,00	21
55154	MSTO1	0	0	12	7	8	6	15	1,00	8		
8776	MTMR1	0	0	6	4	12	0,99	7	7	5		
9107	MTMR6	2	0	23	0,92	13	22	0,91	12	41	0,95	21
2475	MTOR	2	0	304	1,00	113	122	1,00	52	195	1,00	65
4580	MTX1	2	2	74	0,94	20	67	1,00	21	78	0,99	19
10651	MTX2	0	0	12	0,99</							

Annexes | Annexe 2 : Tableau des interacteurs de RAS identifiés par la technique d'AP-MS

56996	SLC12A9	0	0	16	0,99	10	4	2	2	1	1	
6509	SLC1A4	0	0	26	1,00	8	13	6	1	1	1	
11001	SLC27A2	1	1	51	1,00	23	12	9	41	1,00	19	
10999	SLC27A4	1	0	24	1,00	15	7	5	5	4	4	
57181	SLC39A10	1	0	13	1,00	8	3	3	1	1	1	
9497	SLC4A7	0	0	29	1,00	10	3	2	2	1	1	
8884	SLC5A6	0	0	24	1,00	10	0	0	0	0	0	
8140	SLC7A5	0	0	15	0,99	8	3	2	3	2	2	
4087	SMAD2	0	0	14	1,00	6	13	3	33	0,75	9	
4090	SMAD5	0	0	10	0,99	7	21	1,00	6	18	0,99	7
56916	SMARCA1	0	0	12	1,00	8	4	3	15	0,75	8	
55671	SMEK1	0	0	78	1,00	16	65	1,00	15	84	1,00	18
23049	SMG1	0	0	28	1,00	21	24	1,00	15	20	1,00	16
55181	SMG8	0	0	23	0,75	9	11	9	31	1,00	16	
27044	SND1	0	0	99	1,00	44	105	1,00	48	101	1,00	47
8723	SNX4	0	0	12	1,00	5	19	0,99	7	22	8	
6652	SORD	2	1	93	1,00	18	93	1,00	19	92	1,00	19
6272	SORT1	3	1	43	0,98	17	18	9	24	0,77	8	
6654	SOS1	0	0	51	1,00	25	67	1,00	33	77	1,00	33
60559	SPCS3	1	1	16	1,00	8	4	2	6	3	3	
51324	SPG21	1	0	23	1,00	9	8	5	10	7	7	
56907	SPIRE1	0	0	29	1,00	1	0	0	107	0,75	4	
9517	SPTLC2	0	0	36	1,00	17	8	6	10	5	5	
23380	SRGAP2	0	0	24	1,00	5	22	1,00	6	30	1,00	8
6717	SRI	0	0	48	1,00	17	41	1,00	15	49	1,00	17
6723	SRM	0	0	45	1,00	16	52	1,00	14	54	1,00	15
6730	SRP68	6	5	90	0,95	25	97	0,99	36	144	1,00	40
51593	SRRT	2	2	40	0,97	20	45	1,00	22	83	1,00	38
6742	SSBP1	1	1	35	1,00	10	19	1,00	11	17	9	
6749	SSRP1	5	3	105	1,00	41	98	1,00	40	119	1,00	44
29101	SSU72	0	0	19	1,00	12	26	1,00	10	28	1,00	12
10735	STAG2	0	0	97	1,00	32	54	1,00	18	72	1,00	25
8027	STAM	0	0	31	1,00	18	28	1,00	16	36	1,00	19
6772	STAT1	0	0	19	1,00	13	28	1,00	17	22	1,00	14
6774	STAT3	2	1	43	0,91	23	7	4	27	10	10	
55240	STEAP3	0	0	25	0,75	8	0	0	0	0	0	
114790	STK11P	0	0	32	1,00	12	9	4	5	3	3	
201595	STT3B	1	1	27	1,00	15	11	7	14	0,75	9	
10273	STUB1	3	3	115	1,00	25	139	1,00	32	155	1,00	33
23673	STX12	2	1	57	0,99	16	7	4	29	9	9	
8675	STX16	0	0	18	0,99	7	8	5	10	5	5	
53407	STX18	1	1	26	1,00	5	13	0,99	5	25	1,00	4
8417	STX7	3	2	80	1,00	20	23	10	23	11	11	
9482	STX8	0	0	22	1,00	12	0	0	3	2	2	
6813	STXB2	0	0	8	1,00	7	11	0,99	7	18	1,00	11
614	STXB3	1	0	35	1,00	15	33	1,00	13	35	1,00	15
8803	SUCLA2	0	0	66	1,00	25	68	1,00	25	71	1,00	25
8802	SUCLG1	0	0	37	1,00	9	29	1,00	9	32	1,00	9
8801	SUCLG2	0	0	55	1,00	21	31	1,00	17	27	1,00	16
6832	SUPV3L1	0	0	27	1,00	14	26	1,00	14	38	1,00	17
8189	SYMPK	0	0	87	1,00	36	10	7	47	1,00	23	
8407	TAGLN2	2	1	165	1,00	23	157	1,00	21	158	1,00	22
132001	TAMM41	0	0	35	1,00	15	6	5	11	5	5	
6894	TARBP1	0	0	56	1,00	25	45	1,00	22	47	1,00	26
6897	TARS	12	9	310	1,00	65	300	1,00	74	297	1,00	72
123283	TARSL2	0	0	38	1,00	8	16	3	24	6	6	
64786	TBCID15	6	6	88	1,00	28	57	24	86	0,79	28	
57465	TBCID24	0	0	15	1,00	8	10	5	12	8	8	
9882	TBCID4	0	0	12	1,00	5	6	2	17	0,75	6	
23061	TBC1D9B	2	1	21	0,87	13	25	0,90	15	35	0,97	15
1155	TBCB	0	0	23	1,00	14	22	1,00	13	24	1,00	11
6904	TBCD	1	0	290	1,00	77	151	1,00	47	202	1,00	59
6905	TBCE	0	0	93	1,00	23	85	1,00	22	106	1,00	30
29110	TBK1	0	0	21	1,00	13	17	1,00	14	14	6	
6908	TBP	0	0	4	1,00	3	9	0,99	5	6	3	
9238	TBRG4	0	0	49	1,00	19	20	1,00	12	39	1,00	19
6917	TCEA1	2	0	17	0,82	9	14	9	24	0,94	11	
9894	TELO2	8	8	177	1,00	60	153	1,00	53	172	1,00	47
54881	TEX10	0	0	15	1,00	10	7	6	23	1,00	12	
51497	THIL	0	0	71	1,00	25	30	1,00	15	49	1,00	18
9984	THOC1	0	0	39	1,00	23	29	1,00	17	47	1,00	21
57187	THOC2	4	3	66	0,98	22	71	1,00	27	93	1,00	32
8563	THOC5	0	0	20	1,00	14	16	10	43	1,00	19	
8914	TIMELESS	1	0	52	1,00	19	37	1,00	16	54	1,00	20
10469	TIMM44	15	14	212	0,82	44	158	36	161	38	38	
7094	TLN1	2	1	25	0,87	8	11	7	33	0,98	15	
56889	TM9SF3	3	3	77	1,00	14	48	0,90	14	54	12	
9777	TM9SF4	0	0	55	1,00	15	13	6	28	1,00	11	
79613	TMC07	0	0	12	1,00	7	2	2	9	0,99	5	
54732	TMED9	0	0	26	1,00	9	5	2	6	3	3	
55831	TMEM111	0	0	27	1,00	11	0	0	13	8	8	
54867	TMEM214	1	0	42	1,00	15	1	4	2	2	2	
55706	TMEM48	0	0	30	1,00	12	20	1,00	11	9	5	
160418	TMTC3	0	0	32	0,75	8	2	1	19	1,00	6	
81542	TMX1	6	3	58	0,81	21	42	20	45	18	18	
51075	TMX2	0	0	15	0,75	7	0	0	3	1	1	
3842	TNPO1	7	5	449	1,00	66	104	1,00	27	302	1,00	50
30000	TNPO2	0	0	192	1,00	27	46	1,00	13	155	1,00	26
23534	TNPO3	0	0	326	1,00	61	107	1,00	37	169	1,00	45
10953	TOMM34	4	3	99	1,00	26	95	1,00	21	124	1,00	27
9868	TOMM70A	0	0	34	1,00	13	21	1,00	11	21	1,00	11
1861	TOR1A	1	1	34	1,00	19	11	6	15	9	9	
64222	TOR3A	1	0	10	0,99	6	0	0	0	0	0	
112858	TP53RK	0	0	2	1,00	1	4	3	15	0,75	8	
7162	TPBG	0	0	54	1,00	15	11	0,99	4	13	5	
7174	TPP2	0	0	54	1,00	30	70	1,00	40	58	1,00	33
80305	TRABD	1	1	31	1,00	12	14	0,99	10	7	4	
8805	TRIM24	0	0	11	0,99	7	31	1,00	18	46	1,00	20
51592	TRIM33	1	0	15	1,00	5	32	1,00	16	32	1,00	11
9320	TRIP12	0	0	20	1,00	9	29	1,00	12	52	1,00	20
9319	TRIP13	11	10	174	1,00	48	174	1,00	49	178	1,00	53
51504	TRMT112	1	1	9	1,00	5	12	5	9	0,99	7	
6738	TROVE2	6	5	99	1,00	30	91	1,00	30	102	0,88	31
8295	TRRAP	0	0	32	1,00	23	33	1,00	22	23	1,00	19
7249	TSC2	0	0	12	0,99	8	22	1,00	12	21	9	
79042	TSEN34	0	0	6	1,00	3	13	0,99	6	16	0,75	4
10102	TSFM	6	3	92	0,98	22	68	0,81	15	87	0,76	19
55622	TTC27	2	1	27	0,94	13	32	0,96	13	49	0,99	18
9652	TTC37	5	4	187	1,00	64	173	1,00	64	209	1,00	73
8458	TFP2	0	0	15	1,00	6	14	0,99	9	22	1,00	10
9675	TTI1	0	0	76	1,00	24	20	8	36	1,00	12	
7272	TTK	2	1	8	1,00	4	18	0,84	5	21	0,86	9
23170	TTLL12	0	0	62	1,00	29	50	1,00	25	53	1,00	26
80727	TTYH3	0	0	28	1,00	5	12	5	15	0,99	4	
7277	TUBA4A	78	77	1849	0,98	8	1736	1,00	9	1927	1,00	10
27229	TUBGCP4	2	1	55	0,99	19	36	0,98	13	42	0,99	15
25828	TXN2	0	0	21	1,00	3	16	1,00	6	14	3	
84817	TXNDC17	7	6	66	1,00	14	78	1,00	14	91	0,99	14
9352	TXNL1	0	0	27	1,00	11	16	0,75	8	33	1,00	13
7296	TXNRD1	0	0	107	1,00	29	94	1,00	29	99	1,00	27
11065	UBE2C	0	0	24	1,00	10	10	6	8	4	4	
65264	UBE2Z	0	0	17	1,00	9	8	8	13	5	5	
7337	UBE3A	0	0	64	1,00	29	60	1,00	29	70	1,00	33
9690	UBE3C	2	2	105	1,00	32	80	1,00	31	91	1,00	29
9354	UBE4A	0	0	17	1,00	13	11	8	15	0,99	10	
10277	UBE4B	0	0	31	1,00	13	32	1,00	15	38	1,00	14

Annexe 3 : Tableau des interacteurs de KRAS identifiés dans le modèle cellulaire HEK293 par la technique de BioID

Gene ID	Gene Name	Top 2 controls	HEK293_KRAS WT			HEK293_KRAS G12C			Log2 fold change
			Total	SAINT	Total Unique	Total	SAINT	Total Unique	
948469	birA	1528 1528	6954	-	100	11196	-	106	
3845	KRAS	4 4	988	-	48	1606	-	45	
8471	ABCC1		149	1.00	76	145	1.00	67	=
23136	ABCC5		96	1.00	41	159	1.00	54	=
4301	ABCD3	14 14	221	1.00	61	219	1.00	63	=
1829	ABIL	8 10	123	0.99	32	126	1.00	28	=
120	ABL1	1 2	12		9	34	0.97	25	1.4
1500	ACAP2	1 2	15		11	27	0.89	15	=
667	ACOT7	7 8	91		31	115	1.00	31	=
2023	ACSL3	4 7	65	0.80	24	92	1.00	28	=
9497	ACSL4	2 2	54	1.00	23	77	1.00	29	=
23344	ADAM22		16	1.00	8	26	1.00	12	=
10652	ADAP1					10	0.98	5	1.1
6558	ADCY3		34	1.00	15	32	1.00	12	=
55914	ADCY6		22	1.00	12	16	1.00	9	=
10979	ADCY9		112	1.00	43	100	1.00	48	=
2037	ADD1	15 18	565	1.00	70	414	1.00	65	=
118	ADD2	3 5	223	1.00	65	173	1.00	59	=
3728	ADD3	13 14	913	1.00	124	800	1.00	116	=
23513	ADIPOR1		26	0.99	10	65	1.00	10	1.3
498	AGAP1		31	1.00	20	41	1.00	25	=
8773	AGPAT6	1 1	30	0.95	13	49	1.00	23	=
56829	AGPAT9		26	1.00	15	47	1.00	24	=
1739	AHCYL1	6 7	112	1.00	27	121	1.00	29	=
5783	AHCYL2	3 3	67	1.00	20	68	1.00	21	=
81839	AIFM1	2 2	11		9	26	0.91	13	1.2
26064	AKAP12	11 12	435	1.00	125	348	1.00	122	=
1969	AKAP5		70	1.00	26	53	1.00	19	=
9590	AKT1		17	1.00	10	18	1.00	11	=
11180	AKT2	1 1	41	1.00	18	69	1.00	27	=
1207	ALCAM		34	1.00	14	47	1.00	16	=
987	AMIGO2		19	1.00	8	32	1.00	13	=
3688	ANKRD13C					8	0.97	8	1.6
7402	ANKRD26	1 2	27	1.00	21	13	0.88	8	-1.0
3326	ANKRD50		43	1.00	31	46	1.00	31	=
10787	ANKS6		23	1.00	16	16	0.99	13	=
6510	ANOG		80	1.00	30	87	1.00	34	=
55717	ANTXR2		8		4	29	1.00	15	1.7
3655	ANXA2	19 21	207		57	255	0.98	67	=
476	AP1M1		13	0.83	9	17	0.95	14	=
2319	APBB2		16	1.00	11	22	1.00	15	=
100506658	APPL2	1 3	31		16	44	1.00	30	=
8216	ARAF	2 2	58	1.00	26	85	1.00	33	=
51741	ARAP2		22	1.00	16	81	1.00	55	1.8
4853	ARC		0		61	1.00	23	6.0	
23266	ARF1	13 15	133		15	184	1.00	16	=
3930	ARF4	3 3	60	1.00	15	63	1.00	16	=
144402	ARF5	4 5	58		9	98	1.00	12	=
26999	ARF6	4 4	36		17	62	1.00	18	=
55789	ARFIP1	6 7	70		33	87	0.98	35	=
157769	ARHGAP1	5 7	134	1.00	34	151	1.00	38	=
1362	ARHGAP21	6 9	120	0.97	66	107	1.00	65	=
57216	ARHGAP39		21	1.00	13	39	1.00	21	=
4735	ARHGAP3	4 4	64	0.99	18	68	1.00	18	=
1824	ARHGAP26		10		8	14	0.99	12	=
64855	ARL13B	2 2	130	1.00	36	241	1.00	48	=
3996	ARL5B					10	0.98	5	1.1
54477	ARL6IP5	1 2	71	1.00	12	79	1.00	12	=
4162	ARVCF		67	1.00	36	71	1.00	41	=
10131	ASIC1		29	1.00	13	28	1.00	13	=
3187	ATAD1		13	1.00	8	6		4	-1.0
11331	ATG9A		10	0.98	7	11	0.99	8	=
6711	ATP11C		25	1.00	16	23	1.00	16	=
506	ATP13A3		45	1.00	26	68	1.00	33	=
27143	ATP1A1	7 8	379	1.00	94	536	1.00	108	=
55117	ATP1B1		34	1.00	10	36	1.00	12	=
55243	ATP1B3	1 1	84	1.00	19	124	1.00	20	=
143098	ATP2A2	23 24	243		67	281	0.99	77	=
59338	ATP2B1	3 4	214	1.00	65	221	1.00	69	=
3309	ATP2B4		33	1.00	19	36	1.00	15	=
6520	ATP5A1	51 53	504		81	558	1.00	84	=
10397	ATP5B	29 31	291		75	319	0.87	77	=
5819	ATP5C1	8 9	132	1.00	25	170	1.00	28	=
57488	ATP5J2-PTCD1	2 2	41	0.99	8	23		5	=
9685	ATP6AP1	1 1	35	1.00	17	43	1.00	21	=
23191	ATP6AP2		51	1.00	25	56	1.00	28	=
23301	ATP6V0A2		10	0.98	7	4		3	-1.1
7531	ATP6V1F	1 1	9		4	16	0.94	6	=
3685	ATP7A		43	1.00	28	35	1.00	21	=
1525	ATP8B2		11		7	19	0.99	11	=
488	AUP1		20	1.00	8	20	1.00	8	=
114882	BAIAP2	5 6	227	1.00	68	350	1.00	69	=
5099	BAIAP2L1	6 6	106	1.00	40	163	1.00	49	=
5296	BASP1	2 2	41		14	71	1.00	25	=
55714	BET1L	1 1	17		8	21	0.82	8	=
4067	BLOC1S2	2 8	8		7	19	0.81	11	1.2
11336	BLVRA	5 6	62		23	80	1.00	22	=
989	BRAF	1 1	73	1.00	23	142	1.00	37	=
10458	BRSK2	3 3				47	0.93	26	2.8
23236	BTN3A2		16	1.00	10	17	0.99	6	=
55240	BTN3A3		10	0.98	9				-1.9
119	C11orf87		13	1.00	7	23	1.00	13	=
57669	C19orf26		122	1.00	25	139	1.00	30	=
7525	C1orf172		22	1.00	11	26	1.00	10	=
5825	C1orf21		27	1.00	8	36	1.00	9	=
8573	C1orf43		19	1.00	8	15	1.00	7	=
490	C2orf30					25	1.00	8	1.9
10399	C2DC4C		50	1.00	26	109	1.00	39	1.1

302	C7orf59				1	16	0.94	6	27	1.00	13	=	
10298	C8orf42				2	26		10	39	0.98	11	=	
10801	C8orf47				1	3	50	1.00	27	49	1.00	28	
9765	CAB39						55	1.00	23	58	1.00	22	
10723	CABLES1				1	1	33	0.94	22	11	1.00	10	-1.5
6535	CACHD1						94	1.00	49	107	1.00	55	
65979	CACNB2						17	0.99	13	5		5	-1.6
11135	CADM1						74	1.00	13	81	1.00	14	
1434	CADM4						11		6	17	1.00	7	
4354	CANTI1				1	1	9		6	25	0.97	15	1.4
25921	CANX				7	7	97	1.00	32	99	1.00	32	
51429	CASK				6	7	218	1.00	71	217	1.00	64	
5062	CAV1				1	1	38	1.00	7	40	1.00	8	
11252	CBY1						8		7	11	0.98	7	
5245	CCDC47				4	7	88	0.99	29	115	1.00	33	
3843	CCDC8						19	11	31	1.00	17		
7416	CCDC88A						38	1.00	26	50	1.00	31	
983	CCM2						21	1.00	17	14	1.00	13	
9462	CCNY						13	0.99	7	16	0.99	8	
1948	CD276						21	1.00	11	16	0.99	7	
3673	CD44						26	1.00	10	55	1.00	16	1.1
253260	CD47						6		4	15	1.00	8	1.2
1947	CD99						40	1.00	14	76	1.00	18	
7779	CD99L2						5		5	9	0.98	7	
23362	CDC37				6	7	86	0.99	20	108	1.00	20	
9342	CDC42				2	3	160	1.00	8	202	1.00	8	
4172	CDC42BPA						2	85	1.00	50	51	1.00	28
10212	CDC42BPB						1	49	1.00	32	31	0.99	27
4920	CDC42EP1				1	1	198	1.00	40	214	1.00	42	
7037	CDC42EP4				1								

Annexes | Annexe 3 : Tableau des interacteurs de KRAS identifiés dans le modèle cellulaire HEK293 par la technique du BioID

9158	DOCK1		29	1.00	22	44	1.00	29	=	
24145	DOCK11	1	70	1.00	49	102	1.00	62	=	
115	DOCK5					22	0.99	13	1,9	
10768	DOCK9		90	1.00	54	74	1.00	47	=	
23580	DSC2	1	2	314	1.00	51	319	1.00	55	=
92609	DSG2	16	23	1008	1.00	124	872	1.00	126	=
10294	DST	21	23	785	1.00	227	434	1.00	178	=
1741	EBAG9		11		5	50	1.00	14	2,1	
1956	ECE1		18	0.99	9	40	1.00	25	1,1	
7165	ECI2	1	10		7	15	0.88	10	=	
55971	EEDP1		42	1.00	22	46	1.00	25	=	
114299	EFNB1	1	180	1.00	28	235	1.00	32	=	
6507	EFNB2		184	1.00	23	304	1.00	29	=	
5337	EFNB3		36	1.00	10	52	1.00	18	=	
5770	EFR3A	1	1	78	1.00	32	97	1.00	40	=
10128	EFR3B		129	1.00	57	200	1.00	71	=	
153562	EGRF		107	1.00	33	96	1.00	37	=	
2549	EHBP1	8	11	259	1.00	97	194	1.00	73	=
4082	EHD1	1	1	74	1.00	35	76	1.00	38	=
23048	EHD4	1	2	85	1.00	41	113	1.00	48	=
5318	EJF1		16	1.00	9	20	1.00	10	=	
55255	EJF2AK3		23	1.00	14	22	1.00	7	=	
7529	ELMO2	1	1	21	0.85	13	44	1.00	18	1,0
7534	ELMOD3	1	1	24	0.99	10	21	0.99	9	=
8417	ELOVL5	2	2	29	0.84	7	23		5	=
6734	EMD	10	11	143		36	164	1.00	39	=
821	ENO1	66	75	702		63	756	0.98	67	=
10057	EPB41L2	47	54	579		61	544	0.98	66	=
9230	EPB41L3	78	84	1405	1.00	139	1397	1.00	164	=
949	EPB41L4A		11	0.98	9				-2,0	=
6093	EPB41L4B		37	1.00	15	39	1.00	16	=	
6566	EPB41L5	4	5	223	1.00	56	174	1.00	43	=
57685	EPHA2	9	10	443	1.00	92	859	1.00	116	=
4771	EPHA3		80	1.00	33	76	1.00	28	=	
94081	EPHA4	1	1	45	1.00	22	35	1.00	19	=
11332	EPHA7		90	1.00	40	70	1.00	35	=	
4684	EPHB2		42	1.00	21	68	1.00	29	=	
10163	EPHB4		42	1.00	22	86	1.00	39	1,0	
23348	EPN1	2	3	35	0.89	21	16	1.00	-1,1	=
2045	EPN2	1	1	14		11	36	1.00	19	1,3
6573	EPS8	1	1	40	1.00	28	101	1.00	56	1,3
6709	ERBB2		60	1.00	24	36	1.00	23	=	
51510	ERBB2IP	20	21	592		119	811	1.00	164	=
1785	ERBB4		16	0.99	8				-4,1	=
83737	ESYT1	31	33	639	1.00	130	659	1.00	137	=
28956	ESYT2	12	12	262	1.00	73	269	1.00	85	=
23007	EXD2	1	1	15		10	18	0.83	8	=
23043	EXOC3	10	11	232	1.00	70	265	1.00	72	=
57003	EXOC5		28	0.99	16	21	1.00	16	=	
5584	EXOC7		17	1.00	16	16	1.00	10	=	
9987	F1IR		17	0.99	9	40	1.00	12	1,2	
6814	FAM125A	1	1	27	0.97	17	42	1.00	14	=
11140	FAM126A	1	1	35	1.00	11	35	1.00	12	=
1719	FAM129B	10	11	313	1.00	69	264	1.00	50	=
5214	FAM131B		0		21	1.00	14		4,5	=
51715	FAM134B	1	1	13	0.85	9	21	0.96	10	=
293	FAM171A1	1	3	75	1.00	37	86	1.00	37	=
6844	FAM171A2		82	1.00	39	75	1.00	35	=	
8476	FAM171B		45	1.00	23	74	1.00	35	=	
30844	FAM83B		116	1.00	60	107	1.00	52	=	
10165	FAM91A1	9	10	326	1.00	66	248	1.00	57	=
23673	FAM92A1		28	0.99	11	33	1.00	15	=	
483	FARP1	3	3	146	1.00	48	132	1.00	48	=
1267	FARP2	1	1	34	1.00	16	24	0.99	9	=
7204	FCHO2	9	9	130	1.00	48	162	1.00	52	=
1452	FDF1		11		7	13	0.98	10	=	
23499	FERMT1		8		6	12	0.98	10	=	
6786	FERMT2	16	17	583	1.00	96	818	1.00	111	=
284069	FIBP	3	4	114	1.00	36	104	1.00	35	=
10818	FKBP8	1	3	37	0.98	17	33		15	=
8633	FLOT1	1	1	21	0.92	15	13		12	=
57498	FLOT2	5	7	378	1.00	72	395	1.00	76	=
196527	FLRT3		69	1.00	39	45	1.00	29	=	
2042	FLVCR1		146	1.00	31	180	1.00	32	=	
3321	FMN2		42	1.00	22	9		8	-2,1	=
51248	FMNL2		36	1.00	20	42	1.00	18	=	
6092	FMNL3		33	1.00	21	42	1.00	23	=	
23167	FNBP1	3	3	100	1.00	37	119	1.00	48	=
8503	FNBP1L	1	1	16		12	29	0.99	19	=
2030	FRMPD4		5		4	16	1.00	10	1,5	=
11333	FRS2		82	1.00	26	40	1.00	15	-1,0	=
116841	FTSJ1	1	1	40	1.00	13	38	1.00	12	=
221061	GAB1	3	3	101	1.00	34	107	1.00	40	=
23705	GAB2		7		4	22	1.00	14	1,5	=
10938	GD12	2	4	47	0.92	20	61	1.00	25	=
673	GEN1	1	2	16		10	28	0.94	14	=
56704	GNAI2	1	1	14	0.83	9	15	0.87	11	=
23065	GNB2L1	15	18	210	1.00	32	208	1.00	36	=
7074	GNB5		12	1.00	8	8		5	=	
1650	GNS		20	1.00	10	16	1.00	10	=	
10550	GOLGA7		10	0.98	6	12	0.98	7	=	
85377	GPR89A	3	3	31		13	55	1.00	14	=
738	GPRC5A		15	1.00	6					=
9495	GPRC5C		14		7	24	1.00	13	=	
27236	GPRIN1	6	7	122		48	168	1.00	53	=
1499	GPRIN3		31	1.00	23	10	0.98	9	-1,5	=
139818	GRAMD1A	1	1	9		6	19	0.98	12	1,0
255252	GRB14		0		22	1.00	13		4,5	=
23767	GRID1		37	1.00	22	44	1.00	22	=	
9475	GRRK6		0		15	1.00	10		4,0	=

85456	GULP1		8		5	14	1.00	9	=	
6461	H2AFV	1	2	19		6	23	0.85	7	=
6845	HEATR2		19	0.99	13	12	0.99	9	=	
79084	HLA-A	2	4	132	1.00	36	196	1.00	31	=
23382	HMOX2	4	4	58	0.93	21	77	1.00	21	=
421	HNRNP1	20	23	294	1.00	43	304	1.00	42	=
9448	HNRPDL	6	7	87	0.98	22	69		22	=
51144	HOMER1		0		17	1.00	11		4,2	=
55227	HSD17B12	6	8	66		20	100	0.98	23	=
8500	HSP90AB1	33	33	403		88	463	1.00	97	=
2181	HSPA12A	4	5	174	1.00	60	122	1.00	50	=
9050	HSPA5	22	25	280	1.00	76	286	1.00	75	=
6526	IARS2	1	1	19	0.88	15	9		6	-1,0
396	ID1	3	3	31	0.87	12	39	0.93	15	=
65108	IGF1R		40	1.00	23	38	1.00	25	=	
79834	IGF2R	2	2	171	1.00	73	183	1.00	80	=
23677	IGSF3		80	1.00	44	47	1.00	32	=	
8140	IL1RAP		6		5	18	1.00	11	1,4	=
6817	ILDR2		30	1.00	22	37	1.00	22	=	
11311	ILK		7		4	14	1.00	6	=	
7532	INF2	4	4	53	0.82	29	37		20	=
644	INSR		46	1.00	24	53	1.00	33	=	
57553	IPOS	12	14	189	0.84	64	147	0.83	53	=
23164	IQGAP1	5	5	157	1.00	90	154	1.00	93	=
81539	IQGAP2	1	2	57	1.00	40	50	1.00	37	=
81609	IRS4	29	32	1964	1.00	161	2323	1.00	180	=
23380	ITCH		89	1.00	35	126	1.00	51	=	
55706	ITGA1		12		11	16	1.00	10	=	
1000	ITGA2		184	1.00	63	217	1.00	62	=	
84902	ITGA3		15	1.00	9	28	1.00	15	=	
378	ITGA4	1	1	49	1.00	27	83	1.00	39	=
2064	ITGA5		174	1.00	45	236	1.00	48	=	
79109	ITGA6		381	1.00	91	421	1.00	97	=	
5226	ITGA7		55	1.00	28	64	1.00	34	=	
6643	ITGA8		37	1.00	21	49	1.00	25	=	
8677	ITGAV		254	1.00	80	385	1.00	95	=	
10206	ITGB1	1	1	416	1.00	57	480	1.00	58	=
145567	ITGB5	1	1	5		5	18	0.87	11	1,7
51678	ITM2B		10		6	15	1.00	13	=	
55615	ITPA	1	1	24	0.99	6	28	1.00	8	=
6536	JAG2		9		6	10	0.98	6	=	
7289	JAM2		7		6	18	1.00	10	1,2	=
381	JAM3		31	1.00	11	62	1.00	16	1,0	=
369	JPH1	2	2	73	1.00	28	82	1.00	33	=
51635	JUP	9	9	557	1.00	94	532	1.00	98	=
3163	KALRN		13	0.98	6	23	1.00	10	=	
9772	KCNB2		38	1.00	23	23	0.99	15	=	
6645	KCTD12	1	1	34	0.99	16	40	1.00	18	=
5127	KIAA0090	5	7	73		47	121	1.00	59	=
10788	KIAA0195		58	1.00	28	34	1.00	17	=	

Annexes | Annexe 3 : Tableau des interacteurs de KRAS identifiés dans le modèle cellulaire HEK293 par la technique du BioID

284119	MARVELD2			102	1,00	39	142	1,00	43	=
10493	MB21D2	5	5	49		20	83	1,00	26	=
9351	MBP			12	0,99	5	12		6	=
55161	MCAM			303	1,00	48	298	1,00	54	=
29801	MCM3	15	15	177		70	168	0,87	76	=
57182	MET			55	1,00	25	91	1,00	37	=
538	MFAP3			13	0,98	8	18	1,00	11	=
1020	MICAL3	3	3	62	1,00	37	48	0,98	29	=
9019	MICALL1	8	9	71		38	124	0,99	45	=
56926	MIF	2	2	30	0,86	8				=
4719	MINK1			22	1,00	10	25	1,00	15	=
56681	MLLT4	70	78	1297	1,00	225	1692	1,00	259	=
9372	MOSP2			15	0,92	7	17	0,95	10	=
80820	MPP1	7	10	195	1,00	44	212	1,00	44	=
2048	MPP6			59	1,00	34	118	1,00	47	1,0
2050	MPP7	2	2	287	1,00	87	384	1,00	96	=
56776	MPRIP	2	3	62	0,99	38	24		18	-1,3
57648	MPZL1			43	1,00	16	64	1,00	18	=
57670	MPZL2					0	10	0,98	6	3,5
8394	MSN	2	2	45	1,00	14	68	1,00	22	=
23180	MTHFD1L	2	4	37		23	41	0,91	23	=
6509	MTM1			14	0,99	11	36	1,00	18	1,3
208	MTMR10			53	1,00	27	61	1,00	30	=
100526740	MTMR12	1	1	21	0,90	14	33	0,99	20	=
10409	MTOR			154	1,00	82	693	1,00	193	2,2
57609	MYO1E	5	5	19		15	111	1,00	59	2,5
9223	NAPA			17	1,00	15	19		12	=
83871	NCAM1			91	1,00	34	57	1,00	27	=
5906	NCKAP1	12	17	389	1,00	99	366	1,00	101	=
2040	NCLN	4	4	43		26	50	0,85	27	=
6811	NCS1			21	1,00	11	41	1,00	19	=
94121	NDRG1	23	23	264		37	428	1,00	43	=
4267	NDRG3	2	2	54	1,00	24	51	1,00	24	=
2059	NDUFAP2			34	1,00	12	52	1,00	15	=
24140	NDUFS1	2	3	43	0,98	28	38	0,94	25	=
3480	NDUFS2	1	1	34	1,00	22	39	1,00	22	=
5216	NDUFS3	3	5	57		23	62	0,96	22	=
8795	NDUF57	1	2	16		7	23	0,84	9	=
4942	NDUFV1			10		7	10	0,98	5	=
51735	NDUFV2	1	1	17	0,97	8	13		7	=
11079	NEDD4	1	1	8		5	25	0,96	15	1,5
857	NEDD4L	2	2	25		16	55	1,00	30	1,1
55704	NF1	1	1	19	0,88	13	60	1,00	43	1,6
9312	NF2	1	2	92	1,00	41	114	1,00	43	=
55327	NISCH	1	2	169	1,00	73	144	1,00	70	=
9943	NOP16	2	2	34	0,90	12	29		8	=
5111	NOS1AP			11	0,98	9	10		9	=
6102	NOTCH1			155	1,00	51	228	1,00	63	=
10228	NOTCH2	2	2	374	1,00	79	406	1,00	79	=
51347	NTSDC2	1	1	25	0,99	20	5		4	-2,1
83874	NUDC2	1	1	19	0,98	8	21	0,96	8	=
26092	NUDT1			13	0,88	5				=
153090	NUMB	6	7	125	1,00	40	168	1,00	45	=
54566	OAT	3	3	39		23	56	0,98	30	=
23770	OCLAD1	2	2	28	0,83	10	27		10	=
2894	OCLN	2	5	378	1,00	56	503	1,00	71	=
8516	OSBPL11	3	3	46	0,86	26	44	0,93	22	=
25902	OSBPL6	1	1	9		6	24	0,87	17	1,3
8301	OSBPL8	8	9	237	1,00	66	270	1,00	71	=
9367	OTUD7B			21	1,00	7	36	1,00	17	=
57620	OXSRI1	3	4	38		19	50	0,96	22	=
55092	P4HB			16	0,86	12	11		8	=
382	PACS1	2	2	29	0,84	19	27		19	=
56650	PACSIN2	9	10	192	1,00	43	219	1,00	44	=
55157	PACSIN3			28	1,00	15	32	1,00	12	=
1605	PAG1			56	1,00	20	69	1,00	25	=
1949	PAK2	16	16	192		55	298	1,00	65	=
114793	PAK4	7	8	206	1,00	50	300	1,00	60	=
5872	PALM	2	4	148	1,00	44	205	1,00	46	=
5868	PALM2	3	3	106	1,00	32	92	1,00	23	=
55737	PANX1			113	1,00	23	129	1,00	21	=
537	PARD3B			10		9	14	0,99	11	=
29924	PARD6B			54	1,00	22	85	1,00	29	=
84668	PARVA	1	1	25	0,95	12	20	0,90	6	=
3939	PCDH7			237	1,00	77	243	1,00	81	=
5291	PCDH9			30	1,00	17	8		6	-1,8
23516	PCNA	3	4	38		14	53	0,99	16	=
91833	PDAP1	9	10	76		20	106	0,86	23	=
109	PDE8A			16	1,00	9	12		11	=
214	PDHB	2	3	26		11	50	1,00	24	=
481	PDZD11	2	3	80	1,00	18	92	1,00	18	=
9855	PEAK1	3	4	64	0,99	36	93	1,00	55	=
115207	PEKP	7	8	86	0,87	39	90	0,81	42	=
91942	PFN1	4	5	40		14	75	1,00	16	=
4720	PGD	4	5	60	0,88	23	65	0,99	22	=
51491	PGRMC2	11	12	166	1,00	28	215	1,00	28	=
5290	PHACTR4	8	9	199	1,00	58	229	1,00	71	=
10728	PHB	15	16	191	0,99	40	206	1,00	44	=
7782	PHB2	21	23	294	1,00	50	321	1,00	56	=
65109	PHF6	1	1	18	0,85	8	17	0,82	8	=
493	PHLDB1			16	1,00	9	18	1,00	14	=
91768	PHLDB2	1	1	24	0,95	16	27	1,00	23	=
91010	PHLDB3			5		4	20	1,00	16	1,8
79651	PI4KA	4	4	142	1,00	75	134	1,00	79	=
6641	PICALM	5	6	37		20	72	0,91	23	=
25824	PIGU	1	2	28	0,83	11	31	0,96	11	=
5792	PIK3CA			34	1,00	20	97	1,00	40	1,5
6548	PIK3CB			35	1,00	19	34	1,00	19	=
6747	PIK3R1	1	1	146	1,00	53	152	1,00	58	=
116987	PIK3R2			235	1,00	69	215	1,00	71	=
55198	PIK3R3			78	1,00	26	59	1,00	24	=

653519	PIPSK1A	2	2	42	0,99	18	69	1,00	25	=
285513	PIPSK1C			15	0,99	10	23	1,00	13	=
3422	PKP2	2	3	100	1,00	41	110	1,00	44	=
83700	PKP4			119	1,00	55	141	1,00	65	=
3998	PLCB1			227	1,00	77	290	1,00	102	=
56262	PLCB4			18	0,99	13	7		6	-1,2
5784	PLCH1			89	1,00	43	73	1,00	42	=
253959	PLCL2			29	1,00	20	13		12	-1,1
54910	PLD1			105	1,00	49	123	1,00	47	=
8724	PLEKHA1	7	8	281	1,00	48	412	1,00	62	=
64854	PLEKHA2	1		59	1,00	41	23	0,99	14	1,8
137964	PLEKHA5	9	9	304	1,00	107	401	1,00	121	=
1456	PLEKHA7		1	51	1,00	32	70	1,00	42	=
23002	PLEKHO2			22	1,00	10	23	1,00	15	=
387597	PLSCR3						35	1,00	7	=
4282	PLXNA1			115	1,00	64	114	1,00	58	=
5101	PLXNB2			23	1,00	16	25	1,00	19	=
9990	PODXL			6		3	13	0,99	7	1,0
41	POR			11		8	22	1,00	11	=
5218	PPAP2B			25	1,00	8	26	1,00	8	=
1793	PPF1A1		1	66	1,00	38	72	1,00	42	=
60481	PPFIBP1	1	2	129	0,93	41	124	1,00	44	=
55690	PPP1R16B			9		8	17	0,99	13	=
23228	PPP1R9A			21	1,00	20	19	1,00	15	=
6560	PRDX5	2	2	32	0,94	16	36	0,97	16	=
6541	PREB	4	5	57		22	96	1,00	27	=
81671	PRKAR2A	12	12	138		37	146	0,95	43	=
79971	PRKAR2B	5	6	128	1,00	43	99	1,00	37	=
57048	PRKCI	2	4	88	1,00	27	87	1,00	20	=
10640	PRR5	1	1	59	1,00	21	62	1,00	23	=
137392	PSD3	1	2	179	1,00	71	139	1,00	57	=
54940	PSMD3	13	14	121		51	160	1,00	52	=
29763	PSTPIP2	4	6	65	0,86	21	83	1,00	22	=
128869	PTGES3	1	2	34	0,96	7	35	0,99	8	=
285282	PTK7			116	1,00	46	110	1,00	44	=
11037	PTPN1	8	9	104	0,97	35	118	0,99	34	=
123036	PTPN13	13	13	482	1,00	195	581	1,00	217	=
9213	PTPN14	2	3	31		23	45	0,99	30	=
23480	PTPRA			24	1,00	14	13	1,00	8	=
22852	PTPRF			32	1,00	20			0	-5,0
81563	PTPRG			152	1,00	61	119	1,00	49	=
5128	PTPRJ			17		13	27	1,00	19	=
93343	PTRF			45	1,00	19	34	1,00	20	=
23478	PVRL2	3	3	264	1,00	47	374	1,00	57	=
6584	RAB11B	3	4	96	1,00	35	111		32	=
60559	RAB11FIP1	9	10	118	0,96	53	129	0,96	56	=

Annexes | Annexe 3 : Tableau des interacteurs de KRAS identifiés dans le modèle cellulaire HEK293 par la technique de BioID

57707	SCAMP3			3	4	54	0.90	21	53	0.98	19	=
54545	SCARB1					96	1.00	27	136	1.00	33	=
23413	SCRIB	15	16	552	1.00	138	450	1.00	127	=		
56957	SDK1			23	0.99	17	15	1.00	11	=		
55607	SDK2			17	0.98	10	15	0.99	11	=		
5922	SEC11A	3	3	27	13	43	0.93	14	=			
6745	SEC22B	3	3	50	0.99	13	66	1.00	14	=		
81542	SEC61G	1	1	37	0.98	1	19	0.83	5	=		
550	SEMA4C			31	1.00	19	17	1.00	14	=		
10328	SEMA6A			50	1.00	29			0	-5.7		
2799	SEMA6D			11	0.98	10			0	-3.6		
84445	SENP3	2	2	22	9	30	0.82	13	=			
25963	SEPT10	1	1	14	10	23	0.95	14	=			
79720	SEPT11	10	12	145	47	181	1.00	55	=			
347902	SEPT2	19	22	317	1.00	47	343	1.00	48	=		
25912	SEPT7	14	15	230	1.00	51	276	1.00	59	=		
83987	SEPT9	16	16	206	53	206	1.00	55	=			
23155	SFXN1	8	9	92	32	110	0.94	35	=			
22898	SH2B2			5	1	12	0.98	7	1.1			
160518	SH3BP4			63	1.00	30	88	1.00	37	=		
94239	SH3GL3	1		6	4	16	0.89	10	1.3			
54919	SHB	1		68	1.00	35	76	1.00	36	=		
55699	SHOC2			19	1.00	13	20	1.00	14	=		
84061	SIRPA			9	6	22	1.00	9	1.2			
4643	SLC12A2	4	6	618	1.00	120	685	1.00	143	=		
4763	SLC12A4			29	1.00	22	37	1.00	24	=		
134492	SLC12A6	1	1	30	0.97	21	18	1.00	12	=		
22821	SLC12A7			203	1.00	58	204	1.00	54	=		
8036	SLC16A1	7	9	95	32	108	1.00	34	=			
6575	SLC19A1			90	1.00	23	119	1.00	25	=		
6533	SLC19A2					12	0.99	7	1.7			
54795	SLCIA3	1	1	106	1.00	23	70	1.00	21	=		
8936	SLCIA4			42	1.00	22	58	1.00	22	=		
3177	SLCIA5	7	9	387	1.00	41	445	1.00	52	=		
1508	SLC20A2			19	15	39	1.00	20	1.0			
1889	SLC22A5			27	1.00	11	35	1.00	12	=		
5332	SLC25A13	9	10	85	39	117	0.99	50	=			
84295	SLC25A3	8	12	115	25	151	1.00	26	=			
134510	SLC25A5	15	16	145	27	190	1.00	28	=			
8673	SLC25A6	4	4	86	1.00	14	89	1.00	12	=		
207	SLC26A2			15	0.99	8	11	1.00	5	=		
783	SLC29A1			78	1.00	16	99	1.00	20	=		
51272	SLC29A2			18	1.00	5	15	1.00	4	=		
23265	SLC2A3				0	23	1.00	12	4.6			
79932	SLC30A1			180	1.00	34	222	1.00	37	=		
8775	SLC30A4			34	1.00	9	36	1.00	8	=		
4729	SLC30A5	1	2	26	0.88	14	27	0.95	18	=		
50848	SLC38A1			62	1.00	17	78	1.00	21	=		
5795	SLC38A2	2	3	118	1.00	19	190	1.00	20	=		
55781	SLC39A10			148	1.00	42	182	1.00	47	=		
54549	SLC39A14	1	1	35	0.97	10	43	1.00	9	=		
55901	SLC39A6			56	1.00	24	54	1.00	26	=		
6813	SLC3A2	1	1	267	1.00	60	70	1.00	8	-1.9		
53616	SLC43A1			13	7	22	1.00	11	=			
323	SLC4A7	6	6	693	1.00	108	783	1.00	121	=		
11118	SLC5A3			65	1.00	18	73	1.00	24	=		
389541	SLC6A15	2	3	290	1.00	68	287	1.00	72	=		
2066	SLC6A6			19	1.00	8	24	1.00	11	=		
10209	SLC6A8	1	1	203	1.00	43	219	1.00	45	=		
54874	SLC6A9			59	1.00	17	72	1.00	18	=		
94030	SLC7A1	1		29	1.00	7	50	1.00	14	=		
348654	SLC7A11	1		16	6	42	1.00	10	1.3			
5034	SLC7A2	1		116	1.00	29	144	1.00	35	=		
5151	SLC7A5	1	2	63	1.00	12	75	1.00	17	=		
374291	SLC7A6			15	1.00	8	7	1.00	6	-1.0		
23187	SLC9A1			32	1.00	14	29	1.00	16	=		
55103	SLC9A3R2			44	1.00	19	62	1.00	26	=		
51128	SNAP23	3	4	495	1.00	52	538	1.00	51	=		
89970	SNAP25			56	1.00	24	76	1.00	28	=		
23657	SNAP29	8	8	179	1.00	38	232	1.00	42	=		
114790	SNAP47	2	3	76	1.00	32	97	1.00	39	=		
6794	SNTB1	1	1	33	0.97	20	35	1.00	25	=		
9052	SNTB2	3	3	58	1.00	26	51	0.99	21	=		
23527	SNX1	16	17	173	57	227	1.00	65	=			
1601	SNX2	7	8	60	31	88	0.88	40	=			
55218	SNX27	1	2	62	1.00	29	69	1.00	33	=		
643314	SNX3	1	2	31	0.96	14	51	1.00	14	=		
9926	SNX9	14	18	193	1.00	48	297	1.00	54	=		
4040	SOC56			13	0.99	10	15	1.00	9	=		
84230	SPAG1	1	1	11	8	26	0.94	16	1.2			
3675	SPCS2	1		26	1.00	9	33	1.00	10	=		
158747	SPCS3	2	2	27	9	42	0.99	12	=			
23396	SPG21				13	0.99	5	2.2				
1836	SPRED2			23	1.00	14	145	1.00	36	2.6		
9057	SPRY2				0	24	1.00	11	4.6			
5865	SPRY4			5	4	128	1.00	24	4.4			
84928	SPTAN1	8	12	90	46	244	1.00	146	1.4			
25978	SPTBN1	5	5	293	1.00	140	187	1.00	101	=		
8527	SPTBN2			147	1.00	84	87	1.00	55	=		
22905	SPTLC2			9	6	11	0.98	9	=			
2771	SRC			116	1.00	39	134	1.00	42	=		
55890	SRGAP2	3	3	62	1.00	34	71	1.00	39	=		
5605	SRM	2	4	51	0.98	15	40	0.84	14	=		
4534	SRPR	8	9	98	31	123	1.00	37	=			
151011	SSR1			21	1.00	5	28	1.00	6	=		
6809	SSR3	1	2			49	1.00	5	=			
51726	STAR3			11	0.98	5	13	0.99	6	=		
84896	STEAP2			54	1.00	22	51	1.00	23	=		
8907	STEAP3	2	3	225	1.00	36	224	1.00	36	=		
28971	STIM1	4	4	83	1.00	35	89	1.00	34	=		
219771	STIM2			37	1.00	19	44	1.00	22	=		

8904	STK10			50	1.00	21	72	1.00	30	=		
4521	STK11			16	1.00	8	26	1.00	15	=		
54463	STK11HP			16	1.00	13	10	1.00	8	=		
8997	STK4	8	9					108	0.99	36	5.2	
4238	STOM			41	1.00	19	59	1.00	25	=		
8501	STOML2	6	6	51	1.00	22	71	0.90	28	=		
9396	STON1			28	1.00	14	17	1.00	14	=		
55591	STT3B	1	1	26	0.86	12	25	1.00	12	=		
54805	STX10	1	1	60	1.00	21	54	1.00	20	=		
10681	STX12	1	1	85	1.00	26	124	1.00	28	=		
4155	STX18	6	1	10	3	19	0.95	6	=			
27	STX2			10	0.98	8	7	1.00	7	=		
3672	STX3			14	7	29	1.00	14	1.0			
5910	STX4			11	8	17	1.00	11	=			
9131	STX5	2	4	41	0.93	21	54	0.96	24	=		
57198	STX6			38	1.00	20	89	1.00	34	1.2		
64097	STX7	3	4	99	1.00	26	132	1.00	28	=		
199731	STX8	1	1	56	1.00	16	65	1.00	17	=		
9722	STXBP1	1	1	26	0.99	16	46	1.00	31	=		
79643	STXBP2	1	1	17	0.82	12	25	0.94	16	=		
22839	STXBP3	1	1	87	1.00	38	113	1.00	38	=		
80031	SULT1A1	3	4	63	1.00	22	53	0.98	17	=		
9166	SYTL4			41	1.00	25	50	1.00	26	=		
2222	TANC1	6	6	46	30	83	0.98	47	=			
5447	TAOK3	1	2	38	1.00	21	42	1.00	24	=		
8934	TBC1D10A			38	1.00	17	81	1.00	29	1.1		
6674	TBC1D10B	1	3	53	1.00	24	60	1.00	28	=		
10948	TBC1D12			6	5	11	0.98	7	=			
6810	TBC1D9					16	1.00	13	1.5			
23545	TBRG4	1	8	7	14	0.91	12	=				
10384	TC2N			28	1.00	18	65	1.00	30	1.2		
26084	TDRKH	1	23	0.99	14	26	1.00	16	=			
79065	TENM3											

Annexe 4 : Tableau des interacteurs de KRAS identifiés dans le modèle cellulaire BEAS2B par la technique du BioID

Gene ID	Gene Name	Top 2 Controls		BEAS2B_KRAS WT			BEAS2B_KRAS G12C			Log2 fold change
				Total	SAINT	Total Unique	Total	SAINT	Total Unique	
948469	birA	003	014	4043	-	80	5880	-	86	
3845	KRAS			600	-	40	667	-	34	
4363	ABCC1			97	1,00	38	168	1,00	59	=
5825	ABCD3	8	6	87		21	138	1,00	37	=
6059	ABCE1	12	7	71		32	121	0,95	43	=
10006	ABII	4	3	93	1,00	23	125	1,00	25	=
10152	ABII2			41	1,00	13	41	1,00	16	=
2182	ACSL4	2	2	20		11	35	0,88	16	=
115	ADCY9			12		6	36	1,00	18	1,5
118	ADD1	9	5	377	1,00	55	458	1,00	57	=
120	ADD3	12	11	408	1,00	73	517	1,00	85	=
60312	AFAP1			18	1,00	9	31	1,00	13	=
10768	AHCYL1			33	1,00	18	39	1,00	20	=
9590	AKAP12	20	15	509	1,00	127	815	1,00	145	=
9495	AKAP5			20		6	23	1,00	9	=
214	ALCAM			47	1,00	20	81	1,00	29	=
57182	ANKRD50			18	1,00	12	83	1,00	31	2,1
23294	ANKS1A			12	0,99	7	78	1,00	31	2,6
81611	ANP32E			18	0,99	6	21	1,00	8	=
309	ANXA6	3		38	0,79	15	32	0,88	16	=
161	AP2A2	12	7	55		17	123	0,82	35	1,4
163	AP2B1	14	9	145	1,00	44	211	1,00	61	=
8943	AP3D1	6	4	37		13	82	0,98	20	1,1
322	APBB1			57	1,00	18	88	1,00	29	=
369	ARAF			55	1,00	16	138	1,00	28	1,3
382	ARF6			46	1,00	14	54	1,00	16	=
55738	ARFGAP1			2		2	20	1,00	9	2,8
26286	ARFGAP3	2		10		5	20	0,80	8	=
392	ARHGAP1			42	1,00	17	86	1,00	28	1,0
57584	ARHGAP21			50	1,00	20	84	1,00	33	=
9411	ARHGAP29	10	2	66		28	147	0,99	56	1,1
80728	ARHGAP39			9		3	50	1,00	16	2,4
396	ARHGDLA	2	2	39	0,97	12	49	0,97	14	=
8874	ARHGFE7							1,00	9	4,1
400	ARL1			54	1,00	15	36	1,00	16	=
200894	ARL13B			30	1,00	15	51	1,00	22	=
10550	ARL6IP5			67	1,00	11	102	1,00	12	=
50807	ASAP1	2		9		4	31	0,87	17	1,7
444	ASPH	1						0,99	7	1,2
25923	ATL3	2		23	0,81	11	22	1,00	10	=
476	ATPIA1	13	7	305	1,00	82	394	1,00	85	=
483	ATPIB3			35	1,00	7	59	1,00	15	=
490	ATP2B1			39	1,00	14	38	1,00	14	=
493	ATP2B4	2		53	1,00	21	59	1,00	28	=
10632	ATP5L							0,99	9	1,7
5205	ATP8B1			23	1,00	12	79	1,00	36	1,7
10458	BAIAP2	11	10	457	1,00	67	596	1,00	79	=
10409	BASP1	7	3	58		15	71	0,96	16	=
274	BIN1			6		2	20	1,00	8	1,6
673	BRAF							1,00	41	6,2
55299	BRIX1			35	1,00	14	30	1,00	16	=
682	BSG	8	7	118	1,00	23	141	1,00	25	=
11118	BTN3A2			19	0,99	7	9	1,00	4	-1,0
57103	C12orf5							1,00	5	1,3
387882	C12orf75			13		7	27	1,00	8	1,0
79415	C17orf62			10		5	11	0,98	6	=
255057	C19orf26			84	1,00	21	94	1,00	20	=
125988	C19orf70			29		4	56	1,00	5	=
25966	C2CD2			45	1,00	8	95	1,00	16	1,1
126567	C2CD4C			16	0,99	8	12		6	=
805	CALM2			39	1,00	16	52	1,00	17	=
55832	CAND1	6	4	120	1,00	33	157	1,00	40	=
823	CAPN1	4	2	40	0,80	15	37		20	=
8573	CASK			19		10	37	1,00	17	=
857	CAV1	6	5	58		11	103	1,00	11	=
57003	CCDC47			16	0,99	7	22	0,99	12	=
55704	CCDC88A			48		5	73	1,00	19	=
83605	CCM2			17		9	48	1,00	19	1,4
960	CD44	8	4	122	1,00	23	167	1,00	23	=
996	CDC27	4		15		8	41	0,90	19	1,4
998	CDC42			138	1,00	8	144	1,00	8	=
8476	CDC42BPA			7		4	52	1,00	25	2,7
9578	CDC42BPB			5		3	42	1,00	24	2,8
11135	CDC42EP1			101	1,00	33	159	1,00	40	=
10435	CDC42EP2							0,98	5	3,6
23580	CDC42EP4			111	1,00	32	154	1,00	42	=
83461	CDC43			38	1,00	11	60	1,00	13	=
64866	CDCP1			14		7	25	1,00	14	=
1000	CDH2			17	0,99	9	46	1,00	15	1,4
55165	CEP55	2		34		12	60	1,00	16	=
84902	CEP89			7		4	26	1,00	12	1,8
128866	CHMP4B	2	2					0,87	10	1,8
51510	CHMP5			70	1,00	10	76	1,00	12	=
1153	CIRBP	2		33	0,96	10	26	0,90	9	=
23155	CLCC1	2		17		7	37	0,97	14	1,1
1186	CLCN7			54	1,00	22	93	1,00	32	=
9076	CLDN1			34	1,00	8	55	1,00	8	=
56650	CLDNND1			19	0,99	9	12		7	=
1213	CLTC	22	20	324	1,00	100	338	1,00	98	=
1267	CNP	6	4	74	0,80	25	89	1,00	31	=
8904	CPNE1	4	4	42		10	56	0,80	14	=
144402	CPNE8			28	1,00	10	48	1,00	11	=
1464	CSPG4			1240	1,00	197	1488	1,00	211	=
58190	CTDSP1			9		4	23	0,99	10	1,3
1495	CTNNA1	5	2	51	0,91	21	101	1,00	45	1,0
1499	CTNNA2			11		4	28	1,00	14	1,3
1500	CTNND1	4	2	819	1,00	108	938	1,00	114	=
1525	CXADR			67	1,00	18	90	1,00	15	=

23191	CYFIP1	13	10	399	1,00	99	532	1,00	124	=
1601	DAB2			40	0,98	14	53	1,00	19	=
1605	DAG1			45	1,00	15	51	1,00	19	=
26528	DAZAP1	2		34	0,82	11	23	0,86	8	=
285761	DCBLD1			24	1,00	12	47	1,00	25	=
131566	DCBLD2			8		4	37	1,00	11	2,1
8886	DDX18	4	4	54	0,79	23	58	0,98	26	=
8653	DDX3Y			63	1,00	17	112	1,00	26	=
22898	DENND3			14		7	63	1,00	26	2,1
55667	DENND4C			39	1,00	17	102	1,00	33	1,4
55789	DEPDC1B			94	1,00	30	156	1,00	34	=
79139	DERL1			23	1,00	9	23	1,00	10	=
22907	DHX30	4		24		5	39	0,83	8	=
1729	DIAPH1	3		8		5	34	0,78	15	2,0
81624	DIAPH3	6	4	153	1,00	42	284	1,00	76	=
27292	DIMT1			14	0,99	8	9		5	=
23181	DIP2A			5		3	36	1,00	15	2,6
10395	DLC1			3		2	34	1,00	19	3,1
1739	DLG1	8	3	482	1,00	93	690	1,00	101	=
9231	DLG5			24		14	114	1,00	49	2,2
10294	DNAJA2			27	1,00	9	43	1,00	15	=
10059	DNM1L	6		44		16	70	0,91	25	=
55619	DOCK10			52	1,00	17	147	1,00	35	1,5
85440	DOCK7			53	1,00	25	143	1,00	56	1,4
9980	DOPEY2							1,00	8	3,2
1829	DSG2			104	1,00	40	188	1,00	65	=
667	DST	4		110	1,00	50	187	1,00	84	=
11034	DSTN	5		45	0,83	14	52	0,90	15	=
83658	DYNLRB1			10		4	22	1,00	11	1,1
6993	DYNLT1									

Annexes | Annexe 4 : Tableau des interacteurs de KRAS identifiés dans le modèle cellulaire BEAS2B par la technique du BioID

57498	KIDINS220			8		5		34	1,00	19		2,0
55243	KIRREL			146	1,00	46		244	1,00	61		=
3838	KPNA2		5	4	33		15	73	1,00	26		1,1
889	KRT1							11	1,00	15		5,0
55004	LAMTOR1			23	1,00	9		37	1,00	16		=
143903	LAYN		1		22	1,00	11	20		8		=
3996	LLGL1			150	1,00	43		188	1,00	53		=
3998	LMAN1		1		6		3	23	1,00	13		1,8
91289	LMF2			18	1,00	5		4				-1,9
4012	LNPEP			31	1,00	13		76	1,00	22		1,3
23266	LPIN2			286	1,00	79		349	1,00	76		=
987	LRBA			24		11		36	1,00	20		=
10128	LRPPRC			21		10		68	1,00	32		1,6
255252	LRRCS7			33	1,00	11		34	1,00	15		=
4067	LYN			14	0,99	6		26	1,00	14		=
8216	LZTR1			85	1,00	27		65	1,00	18		=
9448	MAP4K4		3		27	0,82	14	36	0,88	22		=
55700	MAP7D1		1		15	1,00	9	12		5		=
79109	MAPKAP1			24	1,00	11		70	1,00	25		1,5
4082	MARCKS		4	4	67	0,96	14	69	1,00	15		=
2011	MARK2		3	2	73	1,00	24	116	1,00	38		=
4140	MARK3			45	1,00	18		83	1,00	29		=
151963	MB21D2			23	1,00	8		52	1,00	16		1,1
4162	MCAM		4	4	905	1,00	52	1036	1,00	68		=
4233	MET			9		5		56	1,00	13		2,5
85377	MICALL1			16		10		70	1,00	20		2,1
50488	MINK1			3		2		21	0,99	11		2,5
84365	MK167IP		1		7		3	24	1,00	11		1,6
4301	MLLT4		43	29	646	1,00	166	1280	1,00	222		1,0
64210	MMS19		2		7		4	33	0,97	6		2,1
4354	MPP1			21	1,00	8		30		11		=
64398	MPP5			4		2		20	1,00	7		2,1
143098	MPP7			129	1,00	45		167	1,00	52		=
23164	MPRI1		13	10	107		54	227	1,00	85		1,1
9019	MPZL1			11		6		27	1,00	11		1,2
9902	MRC2			39	1,00	11		57	1,00	15		=
51154	MRT04			15	0,99	9		27	1,00	16		=
4478	MSN		4	2	62	0,99	20	114	1,00	27		=
8776	MTMR1			58	1,00	15		71	1,00	19		=
2475	MTOR			0		0		1049	1,00	220		10,3
4430	MYO1B		11	9	137	0,89	52	180	1,00	57		=
4641	MYO1C		23	15	187		66	241	0,91	75		=
4643	MYO1E			5		3		43	1,00	25		2,9
26509	MYOF			54	1,00	20		109	1,00	46		1,0
55226	NAT10		3		38	0,89	13	45	0,89	14		=
10787	NCKAP1		2		227	1,00	70	373	1,00	87		=
10397	NDRG1		2		153	1,00	28	200	1,00	38		=
4704	NDUFA9			23	1,00	9		51	1,00	18		1,1
4723	NDUFV1			16	1,00	5		8				-0,9
140609	NEK7		2		23	0,88	9	14		7		=
4756	NEO1		1		25	1,00	13	7		5		-1,7
4763	NF1			0		0		65	1,00	32		6,0
9221	NOLC1			0		0		0	0,99	6		4,6
4851	NOTCH1			5		4		27	1,00	12		2,2
4853	NOTCH2			226	1,00	55		353	1,00	69		=
9520	NPEPPS		12	9	139	0,98	59	124		49		=
50814	NSDHL		7	7	90	0,93	20	109	0,86	24		=
4916	NTRK3			15		7		34	1,00	12		1,1
8650	NUMB		10	9	150	1,00	36	179	1,00	43		=
9631	NUP155			10		6		31	1,00	15		1,5
1,01E+08	OCLN			142	1,00	34		177	1,00	38		=
114885	OSBPL11			4		3		32	1,00	14		2,7
114882	OSBPL8		3		30		15	69	1,00	22		1,2
114883	OSBPL9							69	1,00	8		2,2
9943	OXSRI1		3		39	0,91	15	42	0,93	24		=
5033	P4HA1			22	1,00	10		32	1,00	13		=
8974	P4HA2		2		16		8	27	0,86	15		=
55690	PAC31			61	1,00	29		145	1,00	57		1,2
11252	PAC32			90	1,00	34		106	1,00	35		=
5062	PAK2		13	11	145	0,98	51	220	1,00	58		=
10298	PAK4			24	0,99	12		35	1,00	17		=
5064	PALM		1		116	1,00	37	102	1,00	31		=
114299	PALM2			6		3		0	0,99	14		2,1
445815	PALM2-AKAP2			41	1,00	19		68	1,00	29		=
24145	PANX1			64	1,00	16		95	1,00	21		=
56288	PARD3			18	1,00	12		56	1,00	27		1,6
10016	PDCD6		1		35	1,00	12	36	1,00	16		=
5159	PDGFRB			25	1,00	12		23	1,00	11		=
5162	PDHB		2	2	34	0,90	12	27	0,79	15		=
51248	PDZD11			42	1,00	12		43	1,00	14		=
8682	PEA15			15		5		15	0,99	6		=
79834	PEAK1		1		245	1,00	81	380	1,00	110		=
5216	PFNI		2	2	28	0,79	10	18		7		=
65979	PHACTR4		2		141	1,00	50	202	1,00	53		=
5245	PHB		17	11	151		33	186	1,00	35		=
23187	PHLDB1			14		7		54	1,00	22		1,9
5297	PI4KA			19	0,99	8		37	1,00	10		=
8394	PIP5K1A		3		11		5	42	0,94	16		1,8
5586	PKN2			80	1,00	38		120	1,00	58		=
8502	PKP4			3		2		25	1,00	14		2,7
5337	PLD1			8		5		21	1,00	11		1,3
54477	PLEKHA5		2	1	94	1,00	40	221	1,00	77		1,2
80301	PLEKHO2		4		39	0,88	12	60	0,94	21		=
5358	PLS3		3		19		9	28	0,78	12		=
57048	PLSCR3			48	1,00	6		65	1,00	6		=
8500	PPF1A1			18	0,99	10		32	1,00	15		=
8496	PPFIB1			220	1,00	66		241	1,00	71		=
284352	PPP1R37			0		0		15	0,99	9		4,0

5520	PPP2R2A			4		87	0,99	23	104	1,00	24	=
5566	PRKACA			4		60	0,92	19	71	0,98	19	=
51422	PRKAG2								19	0,99	8	2,7
5576	PRKAR2A		4	1	130	1,00	38	113	1,00	34		=
5584	PRKCI				7		5	25	1,00	12		1,7
136541	PRSS58		2		38	0,97	8	41	0,97	8		=
23362	PSD3				41	1,00	14	90	1,00	28		1,1
5704	PSMC4		2		20	0,79	11	28	0,90	11		=
5718	PSMD12		4		27		10	43	0,85	15		=
5747	PTK2				14	0,99	7	29	1,00	12		1,0
5793	PTPRG		1		24	0,99	11	65	1,00	29		1,4
5795	PTPRJ				79	1,00	32	131	1,00	40		=
284119	PTRF		14	12	149		33	188	1,00	31		=
5819	PVRL2		2		181	1,00	36	250	1,00	38		=
25945	PVRL3				30	1,00	13	57	1,00	16		=
9230	RAB11B		5	2	73	1,00	24	90	1,00	23		=
80223	RAB11FIP1				19	1,00	8	40	1,00	16		1,0
26056	RAB11FIP5				35	1,00	13	71	1,00	28		1,0
5861	RAB1A				46	1,00	12	55	1,00	11		=
23011	RAB21		2	1	50	0,99	18	41	0,99	12		=
51715	RAB23				33	1,00	13	54	1,00	18		=
5873	RAB27A				9		6	28	1,00	14		1,5
83871	RAB34				28	1,00	12	17		9		=
11021	RAB35				12		4	26	1,00	9		1,1
5869	RAB5B		1		26	1,00	8	15	0,99	6		=
7879	RAB7A		10	8	123		28	141	1,00	36		=
5894	RAF1				99	1,00	22	812	1,00	61		3,0
26064	RAI14		51	50	772	1,00	145	975	1,00	160		=
5898	RALA				15	1,00	9	12		6		=
253959	RALGAPA1				10		6	18	0,99	11		=
5902	RANBP1		1		18	0,99	6	6		4		-1,4
57826	RAP2C				0		0	10	1,00	6		1,8
9693	RAPGEF2				0		0	10	1,00	17		5,2
51735	RAPGEF6		4	3	34		12	146	1,00	44		2,1
65059	RAPH1		35	29	187		49	616	1,00	96		1,7
22821	RASA3				20	0,99	11	41	0,99	20		1,0
9462	RASAL2				111	1,00	40	183	1,00	59		=
5962	RDX		2		6		4	39	0,93	17		2,5
768211</												

Annexes | Annexe 4 : Tableau des interacteurs de KRAS identifiés dans le modèle cellulaire BEAS2B par la technique du BioID

6714	SRC			62	1,00	11	77	1,00	8	=
6731	SRP72	8	3	71		14	78	0,79	18	=
6734	SRPR			20		10	41	1,00	12	1,0
6745	SSR1			4		2	21	0,99	7	2,1
55240	STEAP3			9		6	33	1,00	12	1,8
6786	STIM1			0		0	27	1,00	14	4,8
57620	STIM2	1		6		3	26	1,00	10	1,9
10963	STIP1			24	1,00	10	25		10	=
6793	STK10			99	1,00	36	171	1,00	51	=
6811	STX5			22	1,00	9	36	1,00	16	=
8417	STX7			23	1,00	9	28	1,00	13	=
6812	STXBP1			19	0,99	9	21	1,00	11	=
6814	STXBP3			11		7	34	1,00	15	1,5
85360	SYDE1			5		3	19	0,99	8	1,7
94121	SYTL4			51	1,00	20	76	1,00	29	=
6879	TAF7	4	2	18		11	56	0,94	27	1,6
26115	TANC2			41	1,00	23	68	1,00	33	=
11138	TBC1D8			11		7	21	1,00	13	=
7037	TIRC	3	3	140	1,00	47	234	1,00	71	=
55901	THSD1			30	1,00	15	64	1,00	20	1,1
7074	TIAM1			1		0	1	1,00	23	5,6
23670	TMEM2			349	1,00	97	472	1,00	111	=
65062	TMEM237			15	0,99	6	25	1,00	9	=
23043	TNIK			18	1,00	10	40	1,00	21	1,1
64759	TNS3	7	4	179	1,00	62	239	1,00	72	=
84231	TRAF7			21	1,00	10	0		0	-4,5
7204	TRIO	6	6	140	1,00	56	260	1,00	80	=
11078	TRIOBP	4		117	1,00	41	133	1,00	45	=
9320	TRIP12			1		1	19	0,99	13	2,7
54795	TRPM4			33	1,00	16	60	1,00	27	=
7289	TULP3			53	1,00	20	96	1,00	27	=
84669	USP32							1,00	9	3,3
64854	USP46			17	0,98	7	21	1,00	7	=
9712	USP6NL			12		7	39	1,00	15	1,6
7402	UTRN	40	29	533	1,00	181	804	1,00	227	=
6844	VAMP2			57	1,00	9	45	1,00	8	=
9341	VAMP3	2		71	1,00	7	63	0,99	8	=
81839	VANGL1			386	1,00	69	423	1,00	71	=
57216	VANGL2			118	1,00	43	160	1,00	46	=
114990	VASN			17	0,99	6	5		1	-1,6
10493	VAT1			97	1,00	24	106	1,00	27	=
23230	VPS13A	5		16		9	45	0,79	23	1,4
65082	VPS33A			1	1,00	1	3	1,00	6	=
10163	WASF2	4	3	78	1,00	20	133	1,00	23	=
55717	WDR11	9	4	313	1,00	81	426	1,00	95	=
91833	WDR20			16	0,99	8	35	1,00	13	1,1
54521	WDR44	4		40	0,83	21	67	0,98	38	=
51741	WWOX	2		25		13	34	0,97	16	=
7514	XPO1	7	6	102	1,00	37	119	1,00	46	=
7525	YES1			34	1,00	14	59	1,00	22	=
10652	YKT6			566	1,00	29	808	1,00	30	=
56829	ZC3HAV1	6	6	174	1,00	52	278	1,00	72	=
25921	ZDHHC5			191	1,00	42	209	1,00	50	=
29801	ZDHHC8			10		6	44	1,00	19	2,0

Annexe 5 : Tableau des interacteurs de NRAS identifiés par la technique du BioID

Gene ID	Gene Name	TOP 2 control	HEK293_NRAS WT			HEK293_NRAS Q61K			Log2 fold change		
			Total	SAINT	Total Unique	Total	SAINT	Total Unique			
4893	NRAS	1	2230	-	42	3064	-	37			
948469	braA	5733	10470	-	111	16972	-	93			
4363	ABCC1		123	1,00	61	36	1,00	14	-1,7		
10057	ABCC5	1	70	1,00	33	30		15	-1,2		
5825	ABCD3	14	397	1,00	62	312	1,00	57	=		
6059	ABCE1	9	78		33	116	0,97	29	=		
10006	AB11	10	190	1,00	29	131		33	=		
10152	AB12	2	75	1,00	24	33	0,83	16	-1,2		
64746	ACBD3	10	152	1,00	45	134		36	=		
2181	ACSL3	7	43		21	88	1,00	24	1,0		
2182	ACSL4	2	38	0,99	21	102	1,00	28	1,4		
58	ACTA1	5	62	0,87	5	109	1,00	4	=		
71	ACTG1	30	271		34	610	1,00	36	1,2		
56997	ADCCK3	3	24	0,80	16	3		2	-2,6		
109	ADCY3		24	1,00	19	4		3	-2,3		
115	ADCY9		126	1,00	42	15	0,99	7	-3,0		
118	ADD1	18	725	1,00	69	597	1,00	56	=		
119	ADD2	5	137	1,00	53	45		15	-1,6		
120	ADD3	14	1096	1,00	102	701	1,00	94	=		
51094	ADIPOR1		48	1,00	12	78	1,00	14	=		
116987	AGAP1		15	0,99	9	15		8	=		
84803	AGPAT9		10		8	18	0,99	9	=		
10768	AHCYL1	7	74		24	104	0,89	26	=		
23382	AHCYL2	3	37		17	66	1,00	16	=		
9131	AIFM1	2	24	0,85	14	0		0	-4,6		
9590	AKAP12	12	281	1,00	108	134		55	-1,1		
9495	AKAP5		27	0,99	11	23	0,99	8	=		
207	AKT1		12	0,99	8	0		0	-3,7		
208	AKT2	1	95	1,00	25	33	1,00	11	-1,5		
214	ALCAM		30	1,00	16	14		7	-1,0		
5832	ALDH18A1	19	363	1,00	86	222	0,94	66	=		
219	ALDH1B1	1	18	1,00	9	19		9	=		
224	ALDH3A2	14	102		31	190	0,95	38	=		
347902	AMIGO2		22	1,00	15	37	1,00	13	=		
288	ANK3	23	518	1,00	148	252		87	-1,0		
22852	ANKRD26	1	66	1,00	49	3		3	-4,1		
57182	ANKRD50		66	1,00	39	3		2	-4,1		
203286	ANKS6		24	1,00	16	2		1	-3,1		
196527	AN06		47	1,00	23	45	1,00	18	=		
10541	ANP32B	6	70	0,96	23	48		14	=		
302	ANXA2	21	176		45	384	1,00	62	1,1		
26985	AP3M1	7	110	1,00	34	27	20	-2,0	=		
323	APBB2		25	1,00	15	0		0	-4,7		
369	ARAF	2	74	1,00	25	150	1,00	32	1,0		
23237	ARC		0		0	167	1,00	27	7,4		
375	ARF1	15	174	0,98	14	291	1,00	19	=		
378	ARF4	3	62	1,00	16	72	1,00	13	=		
392	ARHGAP1	7	86	0,78	32	112	1,00	26	=		
94134	ARHGAP12	2	28	0,87	20	4		2	-2,5		
57584	ARHGAP21	9	209	1,00	77	46	21	-2,2	=		
396	ARHGDI1	4	43		15	52	0,83	12	=		
9639	ARHGEF10	1	20	0,99	14	0		0	-4,4		
200894	ARL13B	2	140	1,00	33	179	1,00	36	=		
10550	ARL6IP5	2	12		5	61	1,00	12	2,3		
421	ARVCF		52	1,00	28	12		5	-2,0		
41	ASIC1		23	1,00	10	7		3	-1,6		
84896	ATAD1		18	1,00	9	10	0,98	5	=		
55210	ATAD3A	4	94	1,00	48	95	1,00	29	=		
286410	ATP11C		17	1,00	13	0		0	-4,2		
79572	ATP13A3		21	1,00	12	2		1	-2,9		
476	ATP1A1	8	406	1,00	94	1078	1,00	92	1,4		
481	ATP1B1		47	1,00	18	96	1,00	17	1,0		
483	ATP1B3	1	91	1,00	19	190	1,00	21	1,1		
490	ATP2B1	4	214	1,00	52	248	1,00	51	=		
493	ATP2B4		35	1,00	21	18		5	=		
498	ATP5A1	53	874	1,00	80	1271	1,00	81	=		
506	ATP5B	31	503	1,00	79	798	1,00	77	=		
509	ATP5C1	9	154	1,00	30	143	1,00	21	=		
100526740	ATP5J2-PTCD1	2	66	1,00	9	75	1,00	8	=		
10632	ATP5L	3	18		8	35	0,78	7	=		
537	ATP6AP1	1	33	1,00	15	22		10	=		
10159	ATP6AP2		43	1,00	24	19	1,00	9	-1,1		
538	ATP7A		30	1,00	23	0		0	-5,0		
550	AUP1		10		5	17	1,00	8	=		
10458	BAIAP2	6	232	1,00	64	314	1,00	64	=		
55971	BAIAP2L1	6	133	1,00	47	211	1,00	44	=		
10409	BASP1	2	43	0,91	22	86	1,00	14	1,0		
10134	BCAP31	2	21	0,86	7	8		3	-1,3		
10282	BET1	2	32	0,97	8	36	0,90	7	=		
51272	BET1L	1	24	1,00	9	38	1,00	8	=		
673	BRAF	1	108	1,00	32	139	1,00	37	=		
55845	BRK1	1	13	0,99	6	21	1,00	4	=		
682	BSG	2	20		4	231	1,00	18	3,5		
11118	BTN3A2		21	1,00	11	0		0	-4,5		
255057	C19orf26		227	1,00	33	149	1,00	27	=		
81563	C1orf21		43	1,00	10	45	1,00	8	=		
25912	C1orf43		10	0,98	8	0		0	-3,5		
9854	C2CDL2		11	0,99	8	0		0	-3,6		
126567	C2CD4C		82	1,00	24	125	1,00	32	=		
79568	C2orf47	1	13	0,99	8	0		0	-3,8		
203111	C8orf47	3	29	0,85	14	12		8	-1,2		
51719	CAB39		68	1,00	21	27	1,00	13	-1,3		
91768	CABLES1	1	20	1,00	11	0		0	-4,4		
57685	CACHD1		107	1,00	46	18		9	-2,5		
783	CACNB2		52	1,00	30	0		0	-5,7		
23705	CADM1		53	1,00	13	37	1,00	12	=		
805	CALM2	8	25		13	116	1,00	17	2,2		
813	CALU	1	0		0	31	1,00	11	5,0		
55832	CAND1		21	18	249	0,96	72	279	1,00	65	=
124583	CANT1		1	1	39	1,00	23	42	1,00	18	=
821	CANX	7	7		51	20		183	1,00	37	1,8
8573	CASK	7	6	363	1,00	80		125	1,00	45	-1,5
57513	CASKIN2	2	1	28	0,94	22		2	1	-3,3	
857	CAVI	1	36	1,00	8	51	1,00	8	=		
25776	CBY1		13	0,99	6	4		2	-1,5		
57003	CDC47	7	4	81	0,99	27	44	14	=		
83987	CDC8		6		5	20	1,00	10	1,6		
55704	CDC88A		33	1,00	27	16	0,99	10	-1,0		
83605	CCM2		28	1,00	15	2		1	-3,3		
7203	CCT3	36	27	269		57	458	1,00	64	=	
10575	CCT4	27	26	299	0,82	49	433	1,00	54	=	
80381	CD276		16	1,00	6	0		0	-4,1		
960	CD44		45	1,00	13	79	1,00	12	=		
4267	CD99		26	1,00	9	54	1,00	4	1,0		
998	CDCA2	3	2	79	1,00	8	132	1,00	5	=	
8476	CDC42BPA		97	1,00	51	3		3	-4,6		
9578	CDC42BPB	1	44	1,00	28	0		0	-5,5		
11135	CDC42EP1	1	1	211	1,00	43	125	1,00	32	=	
23580	CDC42EP4	1	1	66	1,00	26	75	1,00	20	=	
83461	CDC43	4	4	125	1,00	25	280	1,00	27	1,2	
1000	CDH2	1	1	21	0,99	14	4	2	-2,1		
10423	CDIPT	3	3	40	0,84	12	22	7	=		
983	CDK1	20	17	245	0,92	43	297	1,00	52	=	
5218	CDK14		12		10	34	1,00	10	1,4		
5127	CDK16	8	5	101	0,95	41	69	27	=		
5128	CDK17	1	27		14	35	1,00	17	=		
1020	CDK5	2	1	82	1,00	20	3	3	-4,4		
1952	CELSR2		21	1,00	11	0		0	-4,5		
84902	CEP89		78	1,00	34	14	7	-2,4	=		
10390	CEPT1	1	1	32	1,00	6	3	1	-3,0		
29956	CERS2	1	30	1,00	12	15	0,99	5	-1,0		
1072	CFL1	11	10	148	1,00	25	278	1,00	27	=	
25978	CHMP2B	1	1	24	1,00	12	55	1,00	9	1,2	

Annexes | Annexe 5 : Tableau des interacteurs de NRAS identifiés par la technique du BioID

1956	EGFR			85	1,00	32	38	1,00	18	-1,1
23301	EHPB1	11	8	170	1,00	77	33	1,00	16	-2,3
10938	EHD1	1		41	1,00	26	11		5	-1,8
30844	EHD4	2	1	54	0,99	26	18	10		-1,5
84173	ELMOD3	1		19	0,99	8	6		3	-1,5
60481	ELOVL5	2	2	37	0,91	6	19		4	=
2010	EMD	11	10	162	1,00	31	164	1,00	30	=
2023	ENO1	75	66	1268	1,00	64	1901	1,00	58	=
2036	EPB41L1	6	5	77		26	185	1,00	43	1,3
23136	EPB41L3	84	78	1918	1,00	156	2190	1,00	121	=
54566	EPB41L4B			90	1,00	24	88	1,00	24	=
57669	EPB41L5	5	4	181	1,00	52	61	0,95	18	-1,6
1969	EPHA2	10	9	588	1,00	100	1147	1,00	100	1,0
2042	EPHA3			87	1,00	37	44	1,00	22	-1,0
2043	EPHA4	1	1	33	1,00	15	0		0	-5,1
2045	EPHA7			88	1,00	36	4		3	-4,2
2048	EPHB2			43	1,00	20	14		7	-1,6
2050	EPHB4			37	1,00	23	43	1,00	13	=
29924	EPN1	3	2	51	1,00	21	2		1	-4,1
22905	EPN2	1	1	22	0,99	12	11		6	=
2059	EPS8	1	1	35	1,00	24	19	1,00	11	=
2064	ERBB2			28	1,00	17	0		0	-4,9
55914	ERBB2IP	21	20	975	1,00	173	552	1,00	124	=
23344	ESYT1	33	31	746	1,00	112	927	1,00	116	=
57488	ESYT2	12	12	324	1,00	88	281	1,00	62	=
2130	EWSR1	10	8	33		18	230	1,00	25	2,8
55218	EXD2	1	1	16	1,00	12	0		0	-4,1
11336	EXOC3	11	10	147	1,00	62	106		37	=
23265	EXOC7			16	0,99	13	0		0	-4,1
50848	F11R			49	1,00	10	55	1,00	13	=
9415	FADS2	1	1	14	0,99	10	0		0	-3,9
93343	FAM125A	1	1	6		5	18	1,00	5	1,4
84668	FAM126A	1		22	1,00	8	13		5	=
64855	FAM129B	11	10	287	1,00	66	230	1,00	43	=
57579	FAM135A	5	5	129	1,00	49	40		21	-1,7
221061	FAM171A1	3	1	108	1,00	52	35	0,81	15	-1,6
284069	FAM171A2			90	1,00	37	41	1,00	20	-1,1
165215	FAM171B			108	1,00	45	52	1,00	19	-1,0
222584	FAM83B			216	1,00	85	137	1,00	47	=
157769	FAM91A1	10	9	229	1,00	57	99		28	-1,2
137392	FAM92A1			37	1,00	17	18	0,99	6	-1,0
84188	FAR1	2	2	33	0,95	18	18		8	=
10160	FARP1	3	3	109	1,00	39	36		12	-1,6
9855	FARP2	1		20	1,00	11	1		1	-3,4
115548	FCHO2	9	9	243	1,00	54	43		20	-2,5
2222	FDFT1			14		8	24	1,00	9	=
10979	FERMT2	17	16	764	1,00	98	1274	1,00	95	=
9158	FIBP	4	3	104	1,00	33	10		7	-3,3
10211	FLOT1	1	1	26	1,00	21	5		3	-2,2
2319	FLOT2	7	5	451	1,00	71	322	1,00	57	=
23767	FLRT3			76	1,00	32	19		8	-1,9
28982	FLVCR1			132	1,00	32	153	1,00	33	=
56776	FMN2			19	1,00	14	2		1	-2,7
23048	FNBP1	3	3	49	0,88	24	9		4	-2,3
10818	FRS2			124	1,00	32	4		2	-4,6
24140	FTSJ1	1		38	1,00	12	24	1,00	8	=
2521	FUS	14	11	85		24	191	1,00	26	1,2
2549	GAB1	3	3	143	1,00	45	128	1,00	33	=
2597	GAPDH	35	29	509	1,00	54	757	1,00	51	=
2618	GART	31	31	303		74	364	1,00	64	=
2673	GFPT1	4	4	56	0,90	29	55		22	=
2697	GIA1			28		15	24	1,00	10	=
2762	GMDS	1		12	0,99	10	0		0	-3,7
10399	GNB2L1	18	15	169		31	280	1,00	40	=
9950	GOLGA5	11	11	398	1,00	94	327	1,00	76	=
57120	GOPC	2	1	49	0,99	31	8		4	-2,5
26003	GORASP2	3	3	81	1,00	23	44	0,86	9	=
9527	GOSR1	2	1	28	0,87	17	13		6	-1,1
55890	GPRC5C			19	1,00	11	22	1,00	10	=
114787	GPRIN1	7	6	188	1,00	52	130	0,98	36	=
285513	GPRIN3			34	1,00	26	2		1	-3,5
2888	GRB14			2		2	42	0,99	16	3,8
2894	GRID1			36	1,00	21	4		1	-2,9
83743	GRWD1	7	6	55		19	94	1,00	16	=
54919	HEATR2			12	0,99	11	0		0	-3,7
3105	HLA-A	4	2	240	1,00	42	318	1,00	41	=
51144	HSD17B12	8	6	146	1,00	29	180	1,00	29	=
3320	HSP90AA1	32	29	218		65	849	1,00	93	2,0
3326	HSP90AB1	33	33	299		72	956	1,00	91	1,7
7184	HSP90B1	15	12	26		15	224	1,00	55	3,1
259217	HSPA12A	5	4	151	1,00	53	39		18	-1,9
3304	HSPA1B	148	145	1240		75	1747	1,00	72	=
3309	HSPA5	25	22	275		74	272	0,98	60	=
3329	HSPD1	55	50	138		54	726	1,00	85	2,4
55699	IARS2	1	1	24	1,00	21	0		0	-4,6
23463	ICMT	1	1	19	1,00	8	0		0	-4,3
3420	IDH3B	2	2	29		20	34	0,81	19	=
3480	IGF1R			29	1,00	22	0		0	-4,9
3482	IGF2R	2	2	83	1,00	44	5		4	-3,8
3321	IGSF3			123	1,00	60	5		3	-4,4
387597	ILDR2			24	1,00	14	6		3	-1,8
3643	INSR			59	1,00	34	5		3	-3,3
8826	IQGAP1	5	5	81	1,00	50	24		14	-1,7
8471	IRS4	32	29	2494	1,00	176	2626	1,00	141	=
83737	ITCH			43	1,00	27	26	1,00	14	=
3672	ITGA1			82	1,00	36	1		1	-5,4
3673	ITGA2			338	1,00	71	216	1,00	40	=
3675	ITGA3			25	1,00	11	10		4	-1,2
3676	ITGA4	1		59	1,00	27	16		9	-1,8
3678	ITGA5			152	1,00	41	277	1,00	34	=

3655	ITGA6			466	1,00	96	340	1,00	72	=
3679	ITGA7			94	1,00	43	83	1,00	24	=
8516	ITGA8			39	1,00	21	18	1,00	8	-1,1
3685	ITGAV			243	1,00	61	243	1,00	64	=
3688	ITGB1	1	1	566	1,00	56	629	1,00	42	=
3693	ITGB5	1		14	0,99	9	0		0	-3,9
3714	JAG2			10	0,98	5	18	1,00	6	=
83700	JAM3			42	1,00	15	69	1,00	15	=
56704	JPH1	2	2	74	1,00	29	35		13	-1,1
3728	JUP	9	9	774	1,00	95	461	1,00	75	=
9312	KCNB2			44	1,00	23	0		0	-5,5
56479	KCNQ5			14	0,99	8	0		0	-3,9
115207	KCTD12	1	1	6		5	19	1,00	8	1,5
10945	KDELR1	3	3	21		5	46	0,87	4	1,1
9674	KIAA0040			5		5	14	0,99	7	1,3
9772	KIAA0195			24	1,00	13	0		0	-4,6
9847	KIAA0528	3	3	118	1,00	41	13		8	-3,1
27143	KIAA1274	24	22	418	1,00	94	398	1,00	83	=
57648	KIAA1522			43	1,00	28	2		1	-3,9
57670	KIAA1549			44	1,00	22	3		2	-3,5
57707	KIAA1609			17	0,99	10	14	0,99	7	=
84451	KIAA1804	1	1	40	1,00	27	23	1,00	14	=
57498	KIDINS220	4		152	1,00	72	26		18	-2,5
55243	KIRREL			329	1,00	71	342	1,00	64	=
3838	KPNA2	10	9	59		24	225	1,00	46	1,9
889	KRIT1			10	0,98	10	0		0	-3,5
55004	LAMTOR1	2	2	38	0,92	19	32	0,95	14	=
28956	LAMTOR2	1		11		6	36	1,00	11	1,6
253558	LCLAT1	1	1	11	0,99	8	7		6	=
55327	LINC7	3	2	47	0,99	17	34		10	=
3996	LLGL1	7	6	561	1,00	79	454	1,00	66	=
3993	LLGL2			29	1,00	18	4		3	-2,6
3998	LMAN1	3	3	72	1,00	31	68	1,00	22	=
4012	LNPEP	1	1	86	1,00	43	15		6	-2,4
101060252	LOC101060252	3	2	18		8	48	0,98	8	1,4
22859	LPHN1			75	1,00	42	4		2	-3,9
23266	LPHN2	1		416	1,00	98	189	1,00	53	-1,1
987	LRBA	36	34	526	1,00	187	147		60	-1,8
7804	LRP8			52	1,00	23	7		4	-2,7
55227	LRRC1			80	1,00	19	108	1,00	23	=
94030	LRRC4B			20	1,00	11	0		0	-4,4
25252	LRRC57			86	1,00	23	112	1,00	24	=
56262	LRRC8A			24	1,00	14	3		2	-2,6
51599	LSR	2	1	219	1,00	34	268	1,00	36	=
4067	LYN			281	1,00	55	273	1,00	47	=
8216	LZTR1			455	1,00	103	0		0	-8,8
84445	LZT52			23	1,00	19	4		3	-2,3

Annexes | Annexe 5 : Tableau des interacteurs de NRAS identifiés par la technique du BioID

4853	NOTCH2	2	2	609	1,00	79	452	1,00	76	=
9520	NPEPPS	9	8	81		44	137	1,00	57	=
4869	NPM1	47	43	421		21	598	1,00	21	=
54888	NSUN2	13	12	91		41	196	1,00	46	1,1
64943	NT5DC2	1		13	0,99	10	0		0	-3,8
8650	NUMB	7	6	159	1,00	47	189	1,00	30	=
9253	NUMBL	8	8	95	0,81	37	64		23	=
9688	NUP93	6	5	72	0,84	40	50		23	=
4942	OAT	3	3	34		23	88	1,00	28	1,3
54940	OCLAD1	2	2	50	0,98	8	25		5	-1,0
100506658	OCLN	5	2	346	1,00	51	839	1,00	58	1,3
114885	OSBPL11	3	3	38	0,81	17	13		4	-1,5
114882	OSBPL8	9	8	309	1,00	60	228	1,00	44	=
55690	PACS1	2	2	67	1,00	37	10		5	-2,6
11252	PACSIN2	10	9	199	1,00	36	178	1,00	31	=
55824	PAG1			44	1,00	26	3		2	-3,5
5062	PAK2	16	16	180	0,91	56	137		36	=
10298	PAK4	8	7	347	1,00	63	382	1,00	46	=
5064	PALM	4	2	103	1,00	29	161	1,00	34	=
114299	PALM2	3		79	1,00	30	63	0,99	20	=
24145	PANX1			46	1,00	17	25		10	=
117583	PARD3B			19	1,00	13	0		0	-4,3
84612	PARD6B			107	1,00	34	38	1,00	18	-1,5
5099	PCDH7			300	1,00	86	174	1,00	56	=
5111	PCNA	4	3	31		14	63	1,00	13	1,0
11333	PDAP1	10	9	75		16	131	1,00	18	=
5162	PDHB	3	2	34		12	41	0,79	12	=
23042	PDXDC1	7	7	186	1,00	65	94		40	-1,0
51248	PDZD11	3	2	118	1,00	19	114	1,00	17	=
79834	PEAK1	4	3	115	1,00	62	66	0,86	32	=
53918	PELO	2	1	9		8	21	0,78	6	1,1
5211	PFKL	16	11	150	0,81	54	53		21	-1,5
5214	PFKP	8	7	131	1,00	42	119	1,00	35	=
10424	PGRMC2	12	11	152	1,00	28	168	0,78	25	=
65979	PHACTR4	9	8	299	1,00	74	208	1,00	49	=
5245	PHB	16	15	214	1,00	38	252	1,00	36	=
11331	PHB2	23	21	349	1,00	49	413	1,00	48	=
26227	PHGDH	16	16	128		36	309	1,00	40	1,3
23187	PHLDB1			20	0,99	15	3		2	-2,4
90102	PHLDB2	1		43	1,00	30	2		1	-3,9
5297	PI4KA	4	4	163	1,00	92	43		16	-1,9
5290	PIK3CA			38	1,00	19	10		5	-1,8
5291	PIK3CB			24	1,00	16	4		2	-2,3
5295	PIK3R1	1	1	149	1,00	47	111	1,00	24	=
5296	PIK3R2			185	1,00	66	94	1,00	28	-1,0
8503	PIK3R3			48	1,00	25	4		2	-3,3
8394	PIP5K1A	2	2	77	1,00	31	88	1,00	31	=
5315	PKM2	59	58	320		71	809	1,00	73	1,3
5318	PKP2	3	2	70	1,00	39	46	0,99	23	=
8502	PKP4			145	1,00	64	18	1,00	9	-2,9
23236	PLCB1			194	1,00	78	172	1,00	52	=
5332	PLCB4			10	0,98	10	0		0	-3,5
23007	PLCH1			102	1,00	56	7		5	-3,7
23228	PLCL2			26	1,00	20	0		0	-4,8
5337	PLD1			86	1,00	43	6		4	-3,6
59338	PLEKHA1	8	7	447	1,00	53	582	1,00	61	=
54477	PLEKHA5	9	9	367	1,00	109	397	1,00	96	=
144100	PLEKHA7	1		70	1,00	44	24	1,00	13	-1,5
57048	PLSCR3			2		2	35	1,00	6	3,6
5361	PLXNA1			106	1,00	52	39	1,00	20	-1,4
5464	PPA1	17	13	170	0,96	38	271	1,00	43	=
8613	PPAP2B			17	0,99	7	25	1,00	9	=
8500	PPF1A1	1		103	1,00	42	21	1,00	11	-2,2
8496	PPF1B1	2	1	411	1,00	114	116	1,00	40	-1,8
26051	PPP1R16B			16	0,99	12	1		1	-3,1
55607	PPP1R9A			30	1,00	19	0		0	-5,0
5515	PPP2CA	5	4	22		9	58	0,79	12	1,4
5052	PRDX1	29	29	184		33	432	1,00	41	1,2
10113	PREB	5	4	39		19	113	1,00	23	1,5
5584	PRKCI	4	2	117	1,00	28	117	1,00	23	=
10419	PRMT5	8	7	335	1,00	75	55		19	-2,6
55615	PRR5	1	1	25	1,00	14	30	1,00	11	=
23362	PSD3	2	1	126	1,00	61	30	0,94	11	-2,0
9791	PTDSS1	1	1	18	1,00	11	4		2	-1,9
5754	PTK7			106	1,00	44	98	1,00	34	=
5783	PTPN13	13	13	478	1,00	194	376	1,00	135	=
5784	PTPN14	3	2	54	0,99	32	23		12	-1,2
5792	PTPRF			23	1,00	16	0		0	-4,6
5793	PTPRG			177	1,00	68	23	1,00	15	-2,9
5795	PTPRJ			30	1,00	20	5		4	-2,4
284119	PTRF			19	1,00	12	8		4	-1,2
5818	PVRL1			13	1,00	9	0		0	-3,8
5819	PVRL2	3		306	1,00	45	601	1,00	50	1,0
10890	RAB10	2	1	24	0,90	5	40	0,99	4	=
9230	RAB11B	4	3	32		15	128	1,00	26	2,0
80223	RAB11FIP1	10	9	112	0,80	48	61		21	=
201475	RAB12			12	0,99	9	4		3	-1,4
5872	RAB13	2	2	33	0,93	14	45	0,92	11	=
22931	RAB18	3	2	37	0,81	19	0		0	-5,2
23011	RAB21	5	4	29		14	63	0,89	19	1,1
51715	RAB23	1	1	72	1,00	24	144	1,00	27	1,0
5874	RAB27B			15	1,00	8	0		0	-4,0
83871	RAB34	2	1	22	0,78	15	19	0,79	8	=
5865	RAB3B	1	1	12	0,99	7	0		0	-3,7
5878	RAB5C	4	3	51		12	62	0,99	16	=
7879	RAB7A	4	4	118	1,00	36	247	1,00	34	1,1
4218	RAB8A	4	3	50	0,92	16	59	0,86	10	=
285282	RABL3	2	1	15		9	28	0,91	9	=

5887	RAD23B	1	1	0		0	21	0,99	9	4,5
5894	RAF1	4	3	243	1,00	47	491	1,00	53	1,0
26064	RAH14	15	15	679	1,00	139	212	1,00	69	-1,7
5898	RALA			17	1,00	12	37	1,00	12	1,1
5901	RAN	24	23	451	1,00	33	373	1,00	33	=
5906	RAP1A	4	2	0		0	79	1,00	16	6,3
9693	RAPGEF2			7		7	34	1,00	17	2,1
22821	RASA3			22	0,99	12	1		1	-3,5
9462	RASAL2			186	1,00	77	34	1,00	20	-2,4
83593	RASSF5	2	2	8		5	76	1,00	16	3,1
27316	RBMX	5	4	27		14	85	1,00	18	1,6
768211	RELL1			122	1,00	24	237	1,00	21	1,0
11079	RER1	3	2	30		13	44	0,98	12	=
23180	RFTN1	1		81	1,00	29	30	1,00	15	-1,4
89941	RHOT2	2	1	23	0,90	14	3		2	-2,6
253260	RICTOR	3	3	238	1,00	102	69	1,00	28	-1,8
6091	ROBO1			88	1,00	46	23		11	-1,9
6092	ROBO2			84	1,00	44	2		1	-4,8
6093	ROCK1	1		128	1,00	69	33	1,00	21	-1,9
9475	ROCK2	1		45	1,00	30	8		6	-2,4
4919	ROR1	1		16	1,00	10	0		0	-4,1
4920	ROR2	4	1	246	1,00	47	185	1,00	43	=
6102	RP2			37	1,00	16	14	0,99	10	-1,3
6185	RPN2	12	11	162	1,00	53	253	1,00	48	=
8634	RTCD1	1	1	14	1,00	7	7		4	=
6282	RT00A11	4	3	0		0	83	1,00	5	6,4
56681	SARIA	4	3	26		10	53	0,84	12	1,0
51128	SARIB			22	1,00	8	28		9	=
6305	SBF1	2		143	1,00	76	47	0,99	25	-1,6
9522	SCAMP1	5	4	153	1,00	26	316	1,00	28	1,0
10067	SCAMP3	4	3	62	1,00	17	67	0,99	19	=
949	SCARB1			129	1,00	32	127	1,00	19	=
23513	SCRIB	16	15	839	1,00	167	655	1,00	123	=
221935	SDK1			16	1,00	14	0		0	-4,1
54549	SDK2			23	1,00	17	4		2	-2,3
9554	SEC22B	3	3	26		11	60	0,98	12	1,2
54910	SEMA4C			39	1,00	22	0		0	-5,3
57556	SEMA6A			51	1,00	25	0		0	-5,7
80031	SEMA6D			13	0,99	10	0		0	-3,8
151011	SEPT10	1	1	25	1,00	17	6		2	-1,9
55752	SEPT11	12	10	196	1,00	50	188	1,00	42	=
4735	SEPT2	22	19	466	1,00	45	412	1,00	42	=
55964	SEPT3	1	1	32	1,00	15	2		1	-3,5
23157	SEPT6	12	11	132	0,89	36	146	0,83	29	=
989	SEPT7	15	14	368	1,00	53	278	1,00	42	=
23176	SEPT8	6	4	77	1,00	25	52		20	=
10801	SEPT9	16	16	210	0,99	57	247	0,99	50	=
6421	SEPPQ	38	37	194		44	536	1,00	61	1,5

Annexes | Annexe 5 : Tableau des interacteurs de NRAS identifiés par la technique du BioID

6548	SLC9A1		41	1,00	21	8	4	-2,2		
9351	SLC9A3R2		27	0,99	14	31	1,00	13	=	
133308	SLC9B2		15	0,99	9	0	0	-4,0		
8773	SNAP23	4	3	492	1,00	40	636	1,00	47	=
6616	SNAP25			32	1,00	16	7	5	-2,0	
9342	SNAP29	8	8	230	1,00	40	230	1,00	29	=
116841	SNAP47	3	2	56	1,00	34	23	13	-1,2	
6641	SNTB1	1	1	67	1,00	31	0	0	-6,1	
6645	SNTB2	3	3	54	0,95	21	11	5	-2,2	
6642	SNX1	17	16	200	1,00	56	178	47	=	
8724	SNX3	2	1	39	0,99	19	99	1,00	20	1,3
51429	SNX9	18	14	202	0,81	42	198	50	=	
6646	SOAT1	1		46	1,00	14	0	0	-5,6	
9789	SPCS2	1		2		2	33	1,00	12	3,5
60559	SPCS3	2	2	10		4	34	0,79	11	1,7
200734	SPRED2			8		5	50	1,00	16	2,5
10253	SPRY2			0		0	29	1,00	11	4,9
81848	SPRY4			16	1,00	6	112	1,00	22	2,7
6709	SPTAN1	12	8	340	1,00	157	101	46	-1,7	
6711	SPTBN1	5	5	214	1,00	124	63	32	-1,7	
6712	SPTBN2			67	1,00	45	0	0	-6,1	
6714	SRC			142	1,00	46	162	1,00	26	=
23380	SRGAP2	3	3	50	0,81	31	0	0	-5,7	
6726	SRP9	1	1	14	1,00	5	25	1,00	7	=
6427	SRSF2	3	3	7		3	39	0,82	8	2,3
6747	SSR3	2	1	18		5	34	0,98	5	=
6767	ST13	4	3	6		4	80	1,00	23	3,5
261729	STEAP2			23	0,99	13	0	0	-4,6	
55240	STEAP3	3	2	372	1,00	35	140	1,00	27	-1,4
6786	STIM1	4	4	66	0,92	23	12	5	-2,4	
57620	STIM2			31	1,00	19	10	3	-1,5	
6793	STK10			31	1,00	16	30	1,00	11	=
6794	STK11			22	1,00	11	3	1	-2,5	
2040	STOM			43	1,00	18	38	1,00	16	=
92335	STRADA			11	0,99	8	0	0	-3,6	
3703	STT3A	4	4	59	0,99	20	52	18	=	
8677	STX10	1		48	1,00	18	33	1,00	12	=
23673	STX12	1	1	68	1,00	17	68	1,00	17	=
6809	STX3			5		5	54	1,00	14	3,2
6810	STX4			7		5	28	1,00	11	1,9
6811	STX5	4	2	49	0,99	25	15	8	-1,6	
10228	STX6			26	1,00	14	52	1,00	16	1,0
8417	STX7	4	3	59	0,99	19	133	1,00	23	1,2
9482	STX8	1	1	41	1,00	19	29	1,00	9	=
6812	STXBP1	1		26	1,00	19	9	5	-1,4	
6813	STXBP2	1	1	15	1,00	11	6	3	-1,2	
6814	STXBP3	1		79	1,00	30	54	1,00	19	=
134957	STXBP5	1		21	1,00	15	0	0	-4,5	
29091	STXBP6			16		8	25	1,00	11	=
6817	SULT1A1	4	3	64	0,98	26	54	0,93	19	=
94121	SYT14			47	1,00	31	23	1,00	14	-1,0
8407	TAGLN2	44	44	267		31	767	1,00	35	1,5
9344	TAOK2			10	0,98	8	0	0	-3,5	
51347	TAOK3	2	1	33	0,83	21	12	8	-1,4	
83874	TBC1D10A			20	1,00	13	20	0,99	9	=
26000	TBC1D10B	3	1	88	1,00	33	23	7	-1,9	
55633	TBC1D22B	4	1	37	0,80	24	4	2	-2,9	
123036	TC2N			28	1,00	20	75	1,00	24	1,4
9524	TECR	4	4	51		17	63	0,99	15	=
55714	TENM3			258	1,00	110	27	1,00	17	-3,2
7019	TFAM			9		9	18	0,99	9	=
7037	TFRC	2		241	1,00	75	159	1,00	52	=
7056	THBD			0		0	17	0,99	5	4,2
55901	THSD1			15	0,99	10	0	0	-4,0	
7074	TIAM1			45	1,00	30	23	13	=	
26517	TIMM13			9		6	25	1,00	9	1,4
92609	TIMM50	11	9	138	1,00	24	160	1,00	24	=
147007	TMEM199			3		3	14	0,99	4	1,9
23670	TMEM2			264	1,00	80	381	1,00	83	=
65062	TMEM237			26	1,00	11	6	2	-1,9	
55161	TMEM33	3	3	51	0,99	17	38	15	=	
55706	TMEM48	4	3	67	1,00	34	23	11	-1,5	
55092	TMEM51			25	1,00	7	12	6	-1,0	
25963	TMEM87A	1	1	60	1,00	20	68	1,00	21	=
7110	TMF1	1	1	26	1,00	18	0	0	-4,8	
7112	TMPO	32	25	341		37	706	1,00	58	1,0
25816	TNFAIP8	2	1	22	0,87	8	103	1,00	11	2,2
114034	TOE1	1	1	22	1,00	13	0	0	-4,5	
56993	TOMM22	8	8	148	1,00	20	124	1,00	19	=
26092	TOR1AIP1	4	3	53	0,92	27	58	0,79	18	=
7157	TP53	3	3	43	0,91	11	0	0	-5,5	
10131	TRAP1	24	23	253		79	329	1,00	75	=
11078	TRIOBP	1		35	1,00	19	0	0	-5,2	
54795	TRPM4			21	1,00	13	0	0	-4,5	
125488	TTC39C			64	1,00	19	13	8	-2,2	
145567	TTC7B	1	1	91	1,00	42	22	1,00	11	-2,0
7284	TUFM	26	23	315	1,00	61	352	1,00	61	=
7289	TULP3			69	1,00	23	67	1,00	16	=
7314	UBB	3	2	82	1,00	20	0	0	-6,4	
7334	UBE2N	2	2	24		9	45	1,00	11	=
23190	UBXN4	2	2	42	0,97	9	24	7	=	
219699	UNC5B	1		232	1,00	47	155	1,00	33	=
8633	UNC5C			56	1,00	23	0	0	-5,8	
65109	UPP3B	2	2	33	0,81	15	5	3	-2,5	
7385	UQCRC2	15	14	209	1,00	45	155	37	=	
57478	USP31			40	1,00	23	0	0	-5,4	
84669	USP32			29	0,99	20	11	5	-1,3	
64854	USP46			47	1,00	18	36	1,00	11	=
9712	USP6NL			244	1,00	75	172	1,00	51	=
8239	USP9X	6	5	66	0,93	26	29	12	-1,2	
6844	VAMP2	5	4	82	0,79	10	327	1,00	15	2,0

9341	VAMP3	5	5	145	1,00	11	422	1,00	15	1,5
6845	VAMP7			47	1,00	15	78	1,00	23	=
8673	VAMP8	1	1	26	1,00	13	48	1,00	17	=
81839	VANGL1	1	1	586	1,00	71	353	1,00	62	=
57216	VANGL2	1		398	1,00	64	142	1,00	39	-1,5
9218	VAPA	12	11	135	0,90	17	70	10	=	
10493	VAT1	2	1	49	1,00	16	84	1,00	20	=
7416	VDAC1	16	15	177	0,99	36	241	1,00	36	=
7417	VDAC2	20	20	202		32	228	0,87	28	=
7419	VDAC3	14	13	192	1,00	29	162	0,83	26	=
11311	VPS45	5	3	63	0,91	35	12	7	-2,3	
10490	VTIBB			12	1,00	7	1	1	-2,7	
10163	WASF2	6	6	38		17	140	1,00	18	1,9
55717	WDR11	17	17	370	1,00	87	237	1,00	67	=
91833	WDR20	1		50	1,00	16	52	1,00	13	=
55255	WDR41	6	6	47		24	91	1,00	25	=
57599	WDR48	2	2	134	1,00	46	28	11	-2,2	
11180	WDR6	4	2	569	1,00	101	311	1,00	71	=
79084	WDR77	6	6	312	1,00	30	61	20	-2,3	
79971	WLS	1		22	1,00	13	10	0,98	4	-1,1
51741	WVVOX	31	30	389	1,00	49	527	1,00	52	=
7525	YES1	3		330	1,00	47	443	1,00	53	=
10652	YKT6	13	13	647	1,00	36	1977	1,00	35	1,6
7529	YWHAB	5	4	36		12	86	1,00	11	1,2
7531	YWHAE	17	15	224	0,89	38	249	1,00	32	=
7534	YWHAZ	8	7	38		14	125	0,94	15	1,7
56829	ZC3HAV1	30	26	385	1,00	87	400	1,00	74	=
54503	ZDHHIC13			8	0,97	6	0	0	-3,2	
79683	ZDHHIC14			20	1,00	11	2	1	-2,8	
23390	ZDHHIC17			19	1,00	13	3	2	-2,3	
84243	ZDHHIC18			30	1,00	15	20	1,00	7	=
253832	ZDHHIC20			27	1,00	11	40	1,00	14	=
25921	ZDHHIC5	2		311	1,00	62	253	1,00	58	=
29801	ZDHHIC8			46	1,00	20	28	1,00	11	=

Annexe 6 : Tableau des interacteurs de l'EGFR identifiés par la technique du BioID

Gene Name	Top 2 controls			HEK293_EGFR WT			HEK293_EGFR L858R			HEK293_EGFR Del19			HEK293_EGFR T790M			HEK293_EGFR T790M-L858R			HEK293_EGFR T790M-Del19		
	Total	SAINT	Total Unique	Total	SAINT	Total Unique	Total	SAINT	Total Unique	Total	SAINT	Total Unique	Total	SAINT	Total Unique	Total	SAINT	Total Unique	Total	SAINT	Total Unique
brA	23981	-	94	25088	-	98	17412	-	76	25603	-	92	26694	-	89	22931	-	84			
EGFR	6355	-	144	6499	-	173	5574	-	136	6908	-	191	6216	-	173	6724	-	187			
ABCA3	1			21		4	0	0,75	0	37	0,75	7									
ABCB7	1	1	18	25	0,95	10	13	0,75	9	23	0,83	14	31	0,99	11	44	1,00	14			
ABCE1	9	9	175	189	1,00	39	191	1,00	40	202	1,00	38	215	1,00	42	210	1,00	38			
ABI1	10	8	115	131	1,00	23	51	0,75	15	159	1,00	25	124	0,86	23	95	2,00	24			
ACBD3	10	9	200	265	1,00	44	123	0,75	30	295	1,00	46	298	1,00	43	264	1,00	40			
ACSL3	7	4	73	107	1,00	33	79	0,75	23	104	1,00	30	95	1,00	25	120	1,00	34			
ACTA1	5	4	87	75	1,00	6	104	1,00	6	78	0,95	3	85	0,98	3	71	1,00	4			
ACTG1	30	29	492	393	0,78	33	431	1,00	30	420	1,00	38	450	1,00	33	396	0,78	34			
ADCY9			13	23	1,00	9	4	0,75	3	37	1,00	15	16	0,99	9	10					
ADD1	18	15	312	251	1,00	38	142	0,75	33	355	1,00	52	197	1,00	33	120					
ADD2	5	3	122	97	1,00	32	3	0,75	2	155	1,00	53	31		13	12					
ADD3	14	13	430	344	1,00	72	155	0,75	40	539	1,00	77	174	0,99	41	112					
AEG3L2	4	4	34	50	0,79	17	37	0,75	15	22		6	29		9	60	0,95	18			
AGFG1	13	10	163	226	1,00	41	197	1,00	39	202	1,00	39	307	1,00	39	243	1,00	39			
AGPAT9			12	17	1,00	5	14	0,99	5	23	0,99	8	28	1,00	13	36	1,00	11			
AGPS	1	1	20	23	0,83	11	8	0,75	11	17		9	12		6	15		8			
AHCYL1	7	6	71	36		13	35	0,75	13	81	0,95	18	21		8	24		11			
AHCYL2	3	3	63	44		14	6	0,75	3	62	1,00	17	29		14	20		11			
AHNAK	247	238	2762	2404		498	1727	0,75	434	3199	1,00	604	2770	1,00	527	2658	0,75	526			
AHSA1	14	13	163	236	1,00	36	186	1,00	33	197	1,00	33	223	1,00	35	249	1,00	39			
AIFM1	2		70	60	1,00	25	51	0,97	24	35	0,96	19	61	1,00	23	70	1,00	24			
AK1	1	1	15	5	0,83	3	4	0,75	2	20	0,75	5	16		3	19		4			
AKAP12	12	11	75	73		30	23	0,75	12	168	1,00	59	61		24	34		13			
AKAP5			39	17	0,75	9	0	0,75	0	42	1,00	12	3		2	1		1			
ALDH18A1	19	17	416	433	1,00	82	411	1,00	81	357	1,00	76	458	1,00	81	473	1,00	80			
ALDH1B1	1	1	32	42	0,97	22	32	0,97	17	30	0,77	15	24	0,90	12	29	0,90	11			
ALDH3A2	14	14	174	203	1,00	37	156	0,90	34	185	1,00	38	236	1,00	38	247	1,00	45			
ALG5	3	1	31	32		6	43	0,95	14	39	0,80	10	20		5	31	0,83	10			
ANK3	23	21	503	429	1,00	103	129	0,75	53	563	1,00	134	294	1,00	86	176		54			
ANKFY1	5	4	13	39		21	28	0,75	14	21		11	136	1,00	45	94	1,00	36			
ANKRD26	1		6	4		0	0	0,75	0	25	1,00	15	0		0	0		0			
AP1B1	7	6	68	176	1,00	38	85	0,75	26	227	1,00	43	293	1,00	54	232	1,00	43			
AP1M1	1		52	106	1,00	28	40	1,00	22	88	1,00	29	171	1,00	40	130	1,00	33			
AP2A1	8	7	352	460	1,00	77	267	1,00	64	544	1,00	85	502	1,00	81	373	1,00	74			
AP2A2	12	11	192	267	0,95	67	137	0,75	40	302	1,00	68	352	1,00	75	216	1,00	58			
AP2B1	22	21	492	682	1,00	89	492	1,00	91	749	1,00	110	853	1,00	105	691	1,00	92			
AP2M1	7	6	388	445	1,00	47	377	1,00	44	525	1,00	56	572	1,00	44	399	1,00	49			
AP2S1	1		53	79	1,00	16	74	1,00	18	104	1,00	21	140	1,00	20	76	1,00	20			
AP3M1	7	6	83	104	0,80	36	76	0,75	26	88	0,79	30	91	0,97	33	98	1,00	33			
AP4B1	1	1	3	3		2	3		2	2		2	20	0,76	12	12		7			
APBB1	8	6	30	49		19	35	0,75	18	66		25	147	1,00	32	112	0,78	28			
APBB2				7		5	0	0,75	0	0		0	14	0,99	7			0			
APPL1	2	2	2	11		6	0	0,75	0	10		6	43	0,99	17	28		14			
APRT				5		3	10	0,98	7	13		7			5			3			
ARCN1	13	12	75	119		34	112	0,75	36	105		29	132		35	156	0,75	38			
ARF1	15	13	163	172	0,78	16	158	0,86	16	203	1,00	16	159		16	158		16			
ARF4	3	3	49	60	1,00	15	63	0,99	18	82	1,00	18	66	1,00	16	65	1,00	15			
ARF5	5	4	70	72	1,00	9	67	0,99	12	95	1,00	11	65	0,94	10	67	0,93	10			
ARF6	4	4	59	67	0,94	16	55	0,75	15	53		16	70	0,94	16	58		18			
ARHGAP1	7	5	20	7		21	27	0,75	11	32		12	98	1,00	30	65		22			
ARHGAP21	9	6	34	19		27	2	0,75	1	92	0,78	43	36		16	17		9			
ARHGAP39				0		0	0	0,75	0	17	0,99	8	0		0	0		0			
ARHGEF5				93	1,00	28	32	0,75	11	59	1,00	23	123	1,00	43	174	1,00	57			
ARL1	7	6	76	87	0,95	27	66	0,75	20	94	1,00	23	111	1,00	26	109	1,00	26			
ARL13B	2	2	15	60	1,00	26	50	1,00	24	48	0,88	22	115	1,00	34	87	1,00	31			
ARL6				0		0	0	0,75	0	0		0	5		3	0	0,88	3			
ARL6IP5	2	1	6	24	0,77	6	13	0,75	5	28	0,87	8	10		3	23		7			
ASAP1	4	3	3	26		15	10	0,75	6	22		16	54	0,92	23	88	1,00	33			
ATAD1			21	18	0,99	10	16	0,75	8	7		5	25	0,99	11	12	0,98	7			
ATAD3A	4	4	126	121	1,00	39	103	0,75	32	109	1,00	30	118	1,00	34	117	1,00	34			
ATF7IP	2	1	3	3		3	3	0,75	2	0		0	48	0,76	8	8		3			
ATP1A1	8	7	186	211	1,00	55	230	1,00	58	169	1,00	39	158	1,00	41	185	1,00	53			
ATP1B3	1	1	22	9		7	28	0,86	10	20		6	25	0,85	7	22	0,90	9			
ATP2A2	24	23	363	376	1,00	65	376	1,00	59	301	1,00	53	362	1,00	55	337	0,99	57			
ATP2B1	4	3	119	145	1,00	45	97	0,91	38	105	1,00	34	126	1,00	39	101	1,00	35			
ATPSA1	53	51																			

Annexes | Annexe 6 : Tableau des interacteurs de l'EGFR identifiés par la technique du BioID

821	CANX	7	7	177	1,00	37	186	1,00	40	191	1,00	38	146	1,00	41	173	1,00	37	181	1,00	41
8573	CASK	7	6	84	0,76	33	88	0,79	35	5	0,75	2	101	0,85	37	60		26	21		11
57513	CASKIN2	2	1	24		15	29	0,88	16		0,75	0	64	1,00	27	18		9	7		4
857	CAV1	1		14		4	20	0,89	6		0,75	3	19	0,87	6	12		4	7		2
867	CBL	9	8			0	40		20	35	0,75	21	23		9	112	0,95	41	191	1,00	53
57003	CCDC47	7	4	53		18	60		22	51	0,75	18	49		16	51		17	69	0,80	20
83987	CCDC8										0,75	0	9	0,97	5						0
55704	CCDC88A			5		4					0,75	0	15	0,99	9						0
7203	CCT3	36	27	368	1,00	55	418	1,00	54	398	1,00	60	376	1,00	53	363	1,00	51	386	0,99	54
10575	CCT4	27	26	350	1,00	46	444	1,00	53	371	0,76	49	383	1,00	51	366	1,00	46	414	1,00	51
908	CCT6A	45	44	495	0,75	59	656	1,00	62	470	0,75	65	495	0,81	65	530	1,00	57	534	0,75	61
960	CD44										0,75	0			9	0,97	5		6		3
11140	CDC37	7	6	39		14	88	0,85	17	158	1,00	20	56		20	131	1,00	18	202	1,00	20
8476	CDC42BPA	2		26	0,78	16	21		12		0,75	2	63	1,00	32	8		4			0
11135	CDC42EP1	1	1	60	1,00	19	39	0,99	15		0,75	2	71	1,00	19	15		7			3
23580	CDC42EP4	1	1	32	0,93	12	20	0,87	10		0,75	1	53	1,00	17	19		9	8		4
83461	CDC43	4	4	133	1,00	22	118	1,00	20	109	1,00	25	130	1,00	25	146	1,00	23	126	1,00	22
1000	CDH2	1	1	24	0,89	11	8		5		0,75	0	37	0,99	17						0
10423	CDIPT	3	3	38		10	49	0,82	12	36	0,75	10	23		8	47		10	32		8
983	CDK1	20	17	328	1,00	46	380	1,00	43	334	1,00	46	359	1,00	46	433	1,00	45	430	1,00	45
54901	CDKAL1	8	7	43		16	76		24	50	0,75	23	67		20	122	1,00	27	120	0,99	31
84902	CEP89						22	0,99	11		0,75	0	26	0,99	14	27	1,00	11	10	0,97	5
29956	CERS2	1					8		4	14	0,75	7			14	0,75	6	8		4	4
1072	CFL1	11	10	179	1,00	26	228	1,00	27	204	1,00	26	228	1,00	29	196	1,00	26	219	1,00	28
23155	CLCC1	3	1	24		9	47	0,99	15	23	0,75	7	55	0,99	17	57	1,00	18	46	0,77	14
9685	CLINT1	22	21	203		48	237	0,84	54	207	0,75	42	297	1,00	64	307	1,00	59	246		52
1207	CLNS1A	13	12	512	1,00	26	482	1,00	22	175	0,92	22	655	1,00	29	300	1,00	24	253	1,00	23
10845	CLPX	1		8		4	13		6	11	0,75	6	15	0,85	8			2	8		3
26505	CNNM3			30	1,00	14	21	1,00	10		0,75	1	38	1,00	17			2	4		3
1267	CNP	4	4	67	1,00	24	52	0,76	19	42	0,75	17	50		15	33		14	58	0,77	23
23603	CORO1C	18	16	274	1,00	31	225	1,00	31	281	1,00	29	215	1,00	34	255	1,00	29	265	1,00	32
4513	COX2	6	4	61		15	59		16	62	0,75	16	55		15	61		15	53		15
144402	CPNE8	9	9	378	1,00	49	317	1,00	50	277	1,00	44	297	1,00	48	266	1,00	49	162	1,00	46
1398	CRK	45	43	309		49	479	0,76	50	203	0,75	41	354		43	477		49	469		51
1399	CRKL	47	42	469		51	514	0,76	53	364	0,75	44	445		51	489	0,75	48	433		48
1445	CSK	6	3	32		13	44		19	33	0,75	16	32		16	49		21	62	0,82	26
1452	CSNK1A1	6	6	57		21	75	0,77	21	77	0,90	22	92	0,99	23	135	1,00	30	73		22
1456	CSNK1G3			5		4	4		3	4	0,75	3	5		3	9		6	16	0,99	11
1466	CSRP2	4	3	55	0,83	13	43		10	55	0,75	12	38		11	31		9	23		10
1476	CSTB	5	5	93	1,00	16	31		14	103	1,00	16	25		11	35		11	56		14
1488	CTBP2	6	5	52		15	75	0,98	21	61	0,75	16	51		15	55		21	58		20
1495	CTNNA1	6	6	130	1,00	59	125	1,00	50	31	0,75	19	141	1,00	55	91	0,77	39	77	0,78	32
1499	CTNNB1			27	0,75	10	23	0,75	9		0,75	0	50	1,00	20	14		6	4		2
1500	CTNND1	13	11	291	0,75	63	367	1,00	79	145	0,86	54	468	1,00	90	276	1,00	67	183	0,99	60
1525	CXADR	2	1	111	1,00	21	84	1,00	20	45	0,99	15	86	1,00	22	62	1,00	17	51	1,00	15
1727	CY5R3	5	5	51		19	36		20	69	0,76	21	17		10	44		13	52		16
23191	CYFIP1	10	9	146	1,00	55	126	1,00	43	74	0,75	28	221	1,00	71	91		37	102		42
26999	CYFIP2	11	11	142	0,95	57	123		40	54	0,75	20	223	1,00	73	66		25	40		13
1601	DAB2						4		3		0,75	0	15	0,99	8	10		6	4		3
153090	DAB2IP						6		3		0,75	0	10	0,97	4			0	2		2
1603	DAD1	2	2	29		11	36	0,96	12	29	0,75	9	22		8	20		8	27		9
1605	DAG1			17	0,75	7	8		5	5	0,75	3	8		3			0	0		1
1615	DARS	13	11	150	0,99	50	141	0,76	51	175	1,00	59	127		49	109		41	145	0,80	45
55157	DARS2	4	3	76	0,91	29	79	1,00	28	61	0,99	25	40		23	54	0,82	18	66	0,98	25
1629	DBT	2	1				7		3			0	22	0,76	5						0
79877	DKAKD			31	1,00	12	30	1,00	11	15	0,75	6	22	1,00	9	19	0,75	11			0
84259	DCUN1D5						15		4		0,75	0							14	0,99	6
1650	DDOST	6	6	106	1,00	29	107	1,00	28	123	1,00	33	88	0,89	25	95	0,77	26	103	1,00	28
65992	DDRGG1	4	4	13		7	31		12	14	0,75	6	18		8	57		20	84	1,00	23
10212	DDX39A	16	17	280	1,00	49	282	1,00	46	269	1,00	48	284	1,00	48	273	1,00	48	252	1,00	49
1654	DDX3X	39	34	314		57	347		57	491	1,00	67	368		64	414	1,00	61	417	1,00	61
51202	DDX47	5	4	60		26	64	0,78	26	64	0,76	28	68	0,87	29	62		23	49		25
11056	DDX52	7	7	54		25	69		28	88	0,81	35	41		20	57		25	54		26
79039	DDX54	9	6	83		28	88		22	123	0,90	36	52		22	57		22	74		26
160518	DENND5B	1					11		4		0,75	0			22	0,97	9	12			6
55789	DEPDC1B	13	11	138		30	131		31	77	0,75	24	177	1,00	39	160	1,00	38	108		25
1719	DHFR	4	3	83	1,00	16	98	1,00	15	119	1,00	22	112	1,00	15	68	1,00	12	62	0,96	12
22907	DHX30	5	4	86	0,99	28	101	1,00	33	88	0,99	31	95	1,00	37	102	1,00	33	102	0,94	32
27292	DIMT1	4	4	36		16	67	0,95	29	64	0,87	28	34		17	77	1,00	25	97	1,00	31
1738	DLD	4	3	74	1,00	20	85	1,00	22	65	0,99	20	66	0,99	22	61	0,83	17	79	0,99	20
1739	DLG1	17	16	469	1,00	91	386	1,00	76	136	0,75	46	395	0,75	76	319	1,00	77	192		63
9231	DLG5			20	1,00	12	14		7		0,75	0	64	1,00	35	24	1,00	16	13		7
3301	DNAA1	12	10	120		31	183	1,00	33	143	0,75	30	157	1,00	33	202	1,00	36	200	1,00	38
10294	DNAA2	10	8	167	1,00	22	200	1,00	26	183	1,00	24	197	1,00	24	195	1,00	25	218	1,00	25
9093	DNAA3	8	7	104	0,99	21	111	1,00	23	109	0,99	19	95	0,77	19	86		19	92		18
51726	DNAB11			15		6	30	1,00	9	13	0,75	6	28	1,00	7	34	1,00	8	33	1,00	10
10049	DNAB6	1		18	0,86	8	32	1,00	12	21	0,98	12	16	0,89	8	29	1,00	15	32	1,00	12
55735	DNAUC1	5	5	57		18	63		15	51	0,75	14</									

Annexes | Annexe 6 : Tableau des interacteurs de l'EGFR identifiés par la technique du BioID

2037	EPB41L2	54	47	245	32	255	34	538	0,75	80	340	46	329	40	772	1,00	91							
23136	EPB41L3	84	78	1874	1,00	114	1818	1,00	120	1052	1,00	99	1810	1,00	129	1395	1,00	113	1367	1,00	110			
54566	EPB41LAB			36	1,00	13	20	0,99	8	0,75	0	40	1,00	13	7	3								
57669	EPB41L5	5	4	170	1,00	35	132	1,00	30	31	0,75	13	234	1,00	48	38	14	20			8			
1969	EPHA2	10	9	313	1,00	81	635	1,00	94	366	1,00	81	501	1,00	85	1241	1,00	98	1143	1,00	104			
2042	EPHA3						9		4	0,75	0	16	0,75	7							0			
22905	EPN2	1	1	24		12	62	1,00	26	28	0,79	13	50	1,00	24	102	1,00	31	83	1,00	28			
83481	EPPK1	9	6	84		36	144	1,00	54	125	1,00	54	93		43	149	1,00	54	221	1,00	78			
58513	EPS15L1	50	47	741	1,00	118	545		79	705	1,00	123	941	1,00	141	871	0,75	121	622		82			
2059	EPS8	1	1	5		3	23	0,78	12			21	11	15	7	9					5			
2064	ERBB2						96	0,99	10	10	0,75	6	100	1,00	18	21	0,99	11	102	0,99	13			
55914	ERBB2IP	21	20	634	1,00	138	695	1,00	151	413	1,00	105	934	1,00	161	727	1,00	139	567	1,00	129			
51290	ERGIC2			9		5	19	0,99	5	0,75	0	13	0,75	4	15	0,99	6	20	0,99	6				
54206	ERRF1						7		4	37	1,00	14			0	58	1,00	16	112	1,00	28			
23344	ESYT1	33	31	657	1,00	104	610	1,00	105	506	1,00	96	660	1,00	110	615	1,00	102	628	1,00	101			
57488	ESYT2	12	12	214	1,00	54	176	1,00	48	120	0,75	37	155	0,99	43	127		40	100		33			
2107	ETF1	24	15	168		46	209		44	200	0,75	47	172		41	236	0,75	52	252	0,92	51			
11336	EXOC3	11	10	154	0,76	53	212	1,00	60	80	0,75	31	147	0,99	53	259	1,00	63	126	0,79	43			
23197	FAF2	4	3	36		11	48	0,90	14	37	0,75	14	33		12	53	0,91	11	61	0,91	16			
64855	FAM129B	11	10	260	1,00	54	262	1,00	51	93	0,75	33	300	1,00	57	174	1,00	50	136	0,92	42			
57579	FAM135A	5	5	62		31	56		17	116	1,00	43	105		34	22		6	155	0,75	39			
221061	FAM171A1	3	1	22		13	44	0,77	17	0,75	0	48	0,98	22	16	9					1			
284069	FAM171A2			28	1,00	12	20	1,00	7	0,75	0	30	1,00	12	17	0,98	10	3			2			
165215	FAM171B	15	0,75	8		55	1,00	25	19	0,75	9	28	1,00	16	40	1,00	21	46	1,00	24				
54540	FAM193B	47	1,00	4		23	1,00	4	59	1,00	4	30	1,00	5	38	1,00	4	30	0,75	3				
64762	FAM59A					0	25	1,00	14	13	0,75	8	14	0,98	7	44	1,00	20	67	1,00	27			
150946	FAM59B					0	2		1	0,75	1			0	15	8		47	1,00	18				
222584	FAM83B			28	1,00	12	25	1,00	13	12	0,75	6	62	1,00	32	55	1,00	27	20	0,99	9			
84188	FAR1	2	2	22		11	35	0,96	13	27	0,75	12	45	1,00	18	36	0,80	14	45	0,99	17			
10160	FARP1	3	3	92	1,00	29	61	1,00	20	23	0,75	11	100	1,00	32	37		15	14		8			
26270	FBXO6																					17	1,00	11
115548	FCHO2	9	9	68		25	118	0,95	39	44	0,75	21	176	1,00	51	142	1,00	43	112		39			
2222	FDFT1						4		3	19	1,00	8	8		4	8		5						
10979	FERM12	17	16	397	1,00	71	311	1,00	69	219	1,00	51	291	1,00	65	222	1,00	56	189	0,75	51			
2273	FHL1	5	5	34		12	28		10	63	0,93	8	20		9	21		7	38		12			
10211	FLOT1	1	1	16		8	35	0,78	13	23	0,75	10	28	0,78	10	28	0,96	11	28	0,86	13			
2319	FLOT2	7	5	294	1,00	48	280	1,00	47	105	0,95	40	307	1,00	55	252	1,00	48	148	1,00	42			
28982	FLVCR1			30	1,00	14	13		7	5	0,75	4	19	0,99	8	22	1,00	8	5		3			
114793	FMNL2					0	5		3	0,75	0	16	0,99	6	8	4					3			
22862	FNDC3A	1	1	10		6	24	0,90	11	13	0,75	8	20	0,75	12	47	1,00	20	60	1,00	20			
10818	FRS2			87	1,00	18	55	1,00	14	17	0,75	4	110	1,00	22	33	1,00	10	24	1,00	7			
100533181	FXYP6-FXYD2	1				0	20		0	0,75	0	33	0,95	5	5	4		8	4		4			
2549	GAB1	3	3	109	1,00	28	196	1,00	37	139	1,00	37	163	1,00	40	185	1,00	40	202	1,00	41			
9846	GAB2						16	0,75	8			6		4	26	1,00	11	35	1,00	14				
2584	GALK1	3	3	33		13	52	0,90	18	39	0,75	15	52	0,99	16	42	1,00	17	42	0,93	15			
2597	GAPDH	35	29	325		36	392	1,00	34	410	0,75	43	347		37	396	1,00	35	383	1,00	35			
2617	GARS	5	4	61	0,89	26	89	1,00	29	100	1,00	37	81	0,98	30	81	1,00	28	121	1,00	33			
2618	GART	31	31	399	0,91	72	470	1,00	71	551	1,00	78	395	1,00	77	412	1,00	65	541	1,00	78			
2639	GCDH			14		9	9		8	10	0,75	4	17	0,75	8			4			3			
2730	GCLM	1		7		3	14		6	12	0,75	7	15	0,91	8	11		5	15	0,79	6			
2673	GPFT1	4	4	52		20	95	1,00	28	75	1,00	29	99	1,00	34	60	0,93	25	108	1,00	33			
2746	GLUD1	7	4	75	0,90	26	70		20	74	0,86	23	93	0,99	33	95	1,00	33	101	1,00	25			
2762	GMDS	1				12	6		6	13	0,75	6	8		4	15	0,75	7	20		11			
2778	GNAS	7	6	83		28	68		30	78	0,75	31	78		30	70		28	66		26			
10399	GNB2L1	18	15	260	1,00	37	288	1,00	47	277	1,00	36	297	1,00	39	273	1,00	39	315	1,00	45			
57120	GOPC	2	1	31		17	18		11	11	0,75	6	34	0,93	17	23		9			3			
55890	GPRC5C									0,99		9			9	5								
114787	GPRIN1	7	6	196	1,00	48	150	1,00	35	125	1,00	35	228	1,00	54	153	1,00	37	141	1,00	36			
285513	GPRIN3			23	0,99	11	8		4	0,75	0	39	1,00	17										
57655	GRAMD1A	1								0,75	5	12		6	14	7		17	0,86	7				
2885	GRB2	4	4	16		8	74	1,00	16	100	1,00	21	65	1,00	18	121	1,00	19	149	1,00	24			
83743	GRWD1	7	6	92	0,91	19	94	1,00	17	68	0,75	17	95	0,99	18	85	0,90	19	87	0,99	19			
2931	GSK3A	2	2	20		6	22		8	23	0,75	9	9		6	30	0,77	10	19		7			
10456	HAX1	1	1	20	0,86	9	20		5	20	0,75	9	15		4			4			2			
3105	HLA-A	4	2	47	0,98	17	24		13	32	0,75	15	35		17	30		15	24		14			
3251	HPRT1	5	3	37		13	25		13	106	0,95	21	14		6	40		13	26		11			
3028	HSD17B10	11	11	126		33	134	0,98	39	148	0,99	34	104		36	113		33	114		31			
51144	HSD17B12	8	6	122	1,00	23	158	1,00	22	140	1,00	19	140	1,00	24	120	1,00	22	137	1,00	23			
84263	HSDL2	2	2	23		10	35	0,91	11	15	0,75	6	37	0,94	14	32		11	17		6			
3320	HSP90AA1	32	29	405	1,00	68	461	1,00	73	717	1,00	89	351	1,00	72	436	0,75	65	783	1,00	86			
3326	HSP90AB1	33	33	579	1,00	81	700	1,00	87	932	1,00	90	553	1,00	89	582	1,00	75	1046	1,00	94			
7184	HSP90B1	15	12	104		38	101		34	178	0,75	56	94		30	115		36	155	0,98	49			
259217	HSPA12A	5	4	93	1,00	33	86	1,00	35	9	0,75	4	128	1,00	38	45		20	30		16			
3304	HSPA12B	148	145	1249		72	1371		74	1513	0,75	67	1483		82	1582		74	1804	1,00	76			
3309	HSPA5	25	22	520	1,00	72	575	1,00	76	535	1,00	77	913	1,00	98	778	1,00	81	663	1,00	75			
10525	HYOU1	4	3	64		23	93	1,00	33	80	1,00	30	92	1,00	30	104	1,00	35	103	1,00	32			
55699	IARS2	1	1	35	1,00	17	27	0,96	14	24</														

Annexes | Annexe 6 : Tableau des interacteurs de l'EGFR identifiés par la technique du BioID

57498	KIDINS220	4		110	1,00	47		81	0,99	39		41	0,75	19		207	1,00	81		102	0,99	41		49	0,89	21
55614	KIF16B							24	0,99	15		6	0,75	4		14	0,99	5		79	1,00	32		65	1,00	26
55243	KIRREL			26	1,00	12		24	0,99	11		4	0,75	2		25	1,00	11		22	1,00	11		9		4
3838	KPNA2	10	9	87		27		101		25		147	1,00	32		79		25		150	1,00	32		194	1,00	38
3845	KRAS	1		8		3						14	0,75	6		13		6						6		3
8270	LAGE3																							12	0,98	5
51520	LARS	34	33	366	0,76	83		412	1,00	85		417	1,00	88		377	1,00	86		297		79		326		82
3930	LBR	25	24	286	0,99	33		267		34		275	0,75	36		339	1,00	39		334	1,00	36		301	1,00	37
3939	LDHA	4	4	109	1,00	26		112	1,00	28		94	1,00	26		129	1,00	31		117	1,00	28		131	1,00	31
3945	LDHB	7	6	88	0,91	22		97	0,94	25		81	0,75	20		69		22		81		21		89	0,94	22
3996	LLGL1	7	6	324	1,00	59		267	1,00	56		118	1,00	34		276	1,00	58		183	1,00	50		135	1,00	40
4012	LNPEP	1	1										0,75	1		11		6		25	0,85	11		5		3
101060252	LOC101060252	3	2	57	0,99	9		58	1,00	9		77	1,00	8		53	0,99	8		50	0,99	8		60	1,00	8
79888	LPCAT1	2	1	16		6		21		5		23	0,75	7						25	0,99	8		25	0,99	8
23266	LPHN2	1		55	1,00	25		51	1,00	25			0,75	2		99	1,00	43		29	1,00	10		12	0,78	6
987	LRBA	36	34	180		79		333		105		173	0,75	73		484	1,00	153		384	0,89	122		232		88
121227	LRIG3												0,75	2												
7804	LRP8							17	0,75	8		8	0,75	5						44	1,00	20		6		4
55227	LRRC1			76	1,00	23		62	1,00	20		18	0,75	7		42	0,75	12		22		9		22		7
55379	LRRC59	11	9	83		26		87		23		104	0,75	27		60		18		119		28		143	1,00	30
51599	LSR	2	1	79	1,00	20		77	1,00	21		43	0,99	15		87	1,00	24		62	1,00	17		47	0,91	13
4067	LYN							10	0,97	5			0,75	1												
4085	MAD2L1	1	1	20		5		22	0,85	6		15	0,75	7		34	0,97	8		23	0,85	8		22	0,93	7
84549	MAK16	1		4		3		13		7		7	0,75	4		15	0,82	10		7		5		15	0,91	7
5604	MAP2K1	4	3	43		18		55	0,83	19		39	0,75	16		50	0,78	15		49		17		52		16
5605	MAP2K2	1		12		7		42	1,00	16		26	1,00	14		22		8		13	0,81	7		14	0,88	8
9448	MAP4K4	4	3	60	0,95	21		27		12		28	0,75	12		36		15		36		16		48		21
11183	MAP4K5	10	10	62		27		210	1,00	58		165	1,00	54		195	1,00	47		314	1,00	70		338	1,00	67
4082	MARCKS	3	2	36	0,76	15		37		13		16	0,75	7		64	0,99	17		32		11		20		8
4139	MARK1	1	1	9		5							0,75	1		21	0,78	11								
2011	MARK2	3	1	148	1,00	40		104	1,00	30		89	1,00	27		190	1,00	48		93	1,00	33		76	1,00	26
4140	MARK3	5	4	142	1,00	34		97	1,00	29		65	0,92	25		173	1,00	43		69		20		50		17
4141	MARS	16	16	122		38		163		44		182	0,85	53		147		41		146		40		171		47
153562	MARVELD2			20	1,00	12		26	1,00	9			0,75	0		24	1,00	12								
151963	MB21D2	5	5	44		14		55		19		27	0,75	13		56		17		83	0,99	26		57		20
4162	MCAM			34	1,00	14		13	0,98	6			0,75	1		27	1,00	18		3		2				
4175	MCM6	14	14	174	1,00	57		106		44		133	0,75	48		162	0,92	55		101		36		114		38
4176	MCM7	25	23	467	1,00	84		487	1,00	79		401	1,00	75		539	1,00	83		446	1,00	78		423	1,00	74
90550	MCU	1										8	0,75	4		17	0,90	8						8		4
23195	MDN1			5		4		8		8		13	0,99	9		4		4		9		7				
4233	MET			26	0,99	6		19	0,99	5		9	0,75	4		17	1,00	8		16	1,00	5		10	0,98	3
4301	MLLT4	78	70	1491	1,00	207		1470	1,00	204		1007	1,00	194		1802	1,00	229		1324	1,00	196		1052	1,00	181
22915	MMRN1							15	0,98	7			0,75	2						12		6		6		3
64210	MMS19	1	1	7		5		27	0,97	12		22	0,89	8		24	0,90	11		30	0,99	13		22	0,90	9
51678	MPP6	1		15	0,89	6		9		5		11	0,75	5		15		5		10		5		9		4
143098	MPP7	2	2	227	1,00	64		278	1,00	64		46	0,98	23		279	1,00	62		253	1,00	59		129	1,00	39
9019	MPZL1			15	0,99	7						5	0,75	3												
64963	MRPS11	2	1	21		6		30		9		9	0,75	4		24	0,77	7		14		6		18		5
51537	MTFP1	1	1	13		6		16		6		15	0,75	8		16		6		21	0,91	8		22	0,84	6
4522	MTHFD1	24	18	276	1,00	84		332	1,00	81		366	1,00	98		313	1,00	84		239		70		257	0,81	77
25902	MTHFD1L	4	2	28		11		38	0,80	19		31	0,75	16		22		13		31		15		29		15
55149	MTPAP	1	1	9		5		9		5		11	0,75	6		12		8		13		8		20	0,83	12
4430	MYO1B	18	15	278	1,00	69		216	1,00	54		200	0,75	56		367	1,00	78		244	1,00	56		269	1,00	66
4641	MYO1C	4	4	33		12		44		18		55	0,90	21		45		19		46		18		67	0,75	21
4642	MYO1D	1						10		5			0,75	1		22	0,88	10								
10135	NAMPT	13	12	220	1,00	64		289	1,00	62		252	1,00	66		249	1,00	62		276	1,00	57		307	1,00	63
26960	NBEA			5		5		9		6		13	0,75	9		26	1,00	13		14	0,99	12		15	1,00	10
4077	NBR1	11	10	156		2		633	0,75	3		599	1,00	5		800	0,75	3		749	0,75	3		367	0,75	3
8440	NCK2	7	5	21		11		43		17		21	0,75	11		23		10		55		21		84	0,98	31
10787	NCKAP1	17	12	233	1,00	71		249	1,00	65		80	0,75	36		350	1,00	90		186	1,00	62		150		49
56926	NCLN	4	4	53	0,76	23		66	1,00	28		57	0,90	28		42		17		47		16		48		25
23413	NCS1			7		4		22	0,75	10		9	0,75	5		14	0,99	6		22		9		20		8
4705	NDUFA10	5	5	63		18		72	0,84	19		58	0,75	19		78	1,00	25		55		17		76	0,76	18
4704	NDUFA9	5	4	78	0,87	27		49		20		55	0,75	25		64	0,84	21		67	0,99	24		59	0,84	27
4720	NDUF52	1	1	6		4		17		8		30	0,99	12		8		4		5		4		7		4
4722	NDUF53	5	3	43		14		37		13		73	0,95	16		35		15		33		14		46		17
374291	NDUF57	2	1	14		7		20		6		18	0,75	6		21		8		32	0,97	8		22		6
4771	NF2	2	1	26		15		45	1,00	17		4	0,75	3		62	1,00	23		25		13		18		11
55768	NGLY1							52	1,00	19		33	0,99	12		32	1,00	14		76	1,00	24		96	1,00	28
11188	NISCH	2	1	13		6		5		3			0,75	0		28	0,92	15		6		4		4		2
28987	NOB1			3		2		10</																		

Annexes | Annexe 6 : Tableau des interacteurs de l'EGFR identifiés par la technique du BioID

10130	PDIA6	8	6	55	17	49	18	106	0,99	26	91	0,88	28	60	22	82	23					
5163	PKD1			5	3				0,75		12	0,98	7									
23042	PDXDC1	7	7	63	26	99	0,84	38	39	0,75	17	115	1,00	40	103	1,00	37	60	27			
51248	PDZD11	3	2	74	1,00	16	57	1,00	12	32	0,75	8	98	1,00	16	63	1,00	12	47	0,77	11	
79834	PEAK1	4	3	30	20	68	0,96	30	28	0,75	15	104	1,00	43	118	1,00	47	120	1,00	52		
53918	PELO	2	1	34	0,91	14	42	1,00	15	26	0,75	14	39	0,99	15	38	0,92	18	52	1,00	17	
5214	PFKP	8	7	153	1,00	37	173	1,00	44	209	1,00	48	172	1,00	42	178	1,00	45	199	1,00	46	
10424	PGRMC2	12	11	137	0,88	29	128		26	96	0,75	24	150	1,00	28	165	1,00	28	162	1,00	30	
65979	PHACTR4	9	8	221	1,00	50	260	1,00	56	101	0,75	33	323	1,00	59	292	1,00	58	228	1,00	55	
5245	PHB	16	15	222	1,00	42	233	1,00	43	231	1,00	37	242	1,00	44	256	1,00	44	248	1,00	44	
11331	PHB2	23	21	468	1,00	51	455	1,00	55	456	1,00	58	453	1,00	58	493	1,00	54	470	1,00	51	
26227	PHGDH	16	16	247	1,00	43	231	1,00	43	226	0,97	39	224	0,92	42	193	0,75	38	202	1,00	39	
90102	PHLDB2	1		15	0,83	9							46	1,00	21	23	0,98	15	9		5	
653583	PHLDB3												18	0,99	9							
5297	PI4KA	4	4	29	10	38	14					60	0,81	21	40	12	12	12	4			
8301	PICALM	6	5	94	0,90	29	131	1,00	31	102	1,00	30	123	1,00	33	161	1,00	39	139	1,00	34	
128869	PIGU	2	1	11	6	16	5					26	0,75	8	32	0,95	9	25	0,79	7		
5295	PIK3R1	1	1	31	0,95	15	31	0,99	13	7	0,75	4	62	1,00	21	24	0,92	9	26	0,87	8	
5296	PIK3R2			45	1,00	21	32	1,00	14		0,75	1	74	1,00	31	24	1,00	12	26	0,99	14	
8394	PIP5K1A	2	2	21	10	29	15					33	0,75	6	33	0,75	16	46	0,82	17	34	14
5315	PKM2	59	58	745	1,00	68	739	1,00	71	721	0,97	67	709	1,00	69	843	1,00	65	787	1,00	69	
5318	PKP2	3	2	72	1,00	27	72	1,00	28	31	0,75	14	104	1,00	39	46	0,81	21	35	16	16	
8502	PKP4			7	4	6	3					11	0,98	8	4			2				
23236	PLCB1			17	1,00	8		0,97			0,75	1	20	1,00	11							
23007	PLCH1			12	8	13	6					45	1,00	23								
59338	PLEKHA1	8	7	174	1,00	36	197	1,00	39	84	0,75	24	198	1,00	42	179	1,00	44	146	1,00	32	
54477	PLEKHA5	9	9	279	1,00	75	251	1,00	76	110	0,75	42	379	1,00	96	227	1,00	70	168	1,00	54	
144100	PLEKHA7	1		5	3	9	5					26	0,97	13								
5351	PLOD1	1		14	0,82	9	17	0,83	12	8	0,75	6	24	0,95	12	19	0,96	11	14		9	
25953	PNKD					11	0,98	5														
56902	PNO1	2	1	24	0,81	7	42	1,00	10	18	0,75	5	14	9	29	0,85	9	25	0,89	9		
5445	PON2					6	3								10	6	16	1,00	10			
5464	PPA1	17	13	375	1,00	41	400	1,00	53	418	1,00	47	334	1,00	53	300	1,00	46	346	1,00	46	
8500	PPF1A1	1				8	6					26	1,00	16								
8496	PPF1BP1	2	1	134	1,00	55	69	1,00	32	8	0,75	4	174	1,00	60	27	0,79	13	12		5	
5501	PPP1CC	8	7	96	16	95	13	28	0,75	10	29	10	156	0,75	16	71		12			12	
23368	PPP13B	2		7	4							24	0,82	11								
5515	PPP2CA	5	4	20	8	45	12	28	0,75	6	53	11	13	43	11	61	0,95	16				
10113	PREB	5	4	69	0,98	21	48	16	76	0,87	20	45	15	69	0,94	17	70	0,94	17			
5580	PRKCD	3				10	6					6	4	26	13	37	0,86	17				
5584	PRKCI	4	2	122	1,00	27	92	1,00	23	61	0,94	22	98	1,00	21	84	1,00	24	64	1,00	19	
5591	PRKDC	82	65	699	216	782	220	907	0,75	258	756	242	609	191	722	216						
10419	PRMT5	8	7	73	28	71	27	22	0,75	11	106	0,98	34	23	12	23	9					
11212	PROSC	1	1	11	5			17	0,77	8												
23362	PSD3	2	1	102	1,00	40	66	1,00	31	5	0,75	4	190	1,00	57	11	7					
10213	PSMD14	4	3	39	12	53	0,82	18	49	0,92	16	53	0,87	19	48	0,84	17	54	0,82	19		
5709	PSMD3	14	13	190	1,00	54	220	1,00	51	228	1,00	55	239	1,00	60	223	1,00	54	216	1,00	51	
5754	PTK7			16	0,75	10						23	0,99	12								
51495	PTPLAD1	19	16	276	1,00	27	239	1,00	27	259	1,00	29	220	1,00	28	232	1,00	27	248	1,00	27	
5770	PTPN1	9	8	72	28	111	1,00	35	78	0,75	28	125	1,00	39	147	1,00	36	143	1,00	38		
5783	PTPN13	13	13	365	1,00	143	367	1,00	131	193	0,78	93	620	1,00	185	305	1,00	114	196	1,00	81	
80324	PUS1	5	5	69	0,84	24	80	1,00	31	55	0,75	20	73	0,99	24	58	24	63	22			
5819	PVRL2	3		98	1,00	27	101	1,00	28	79	1,00	26	69	0,98	23	72	0,99	25	64	0,99	24	
29920	PYCR2	2	1	21	9	30	0,94	9	9	0,75	5	22	9	20	9	25	0,84	7				
10890	RAB10	2	1	29	0,90	5		0,91		25	0,75	6	25	5								
9230	RAB11B	4	3	70	0,78	22	62	0,75	20	102	1,00	31	79	1,00	25	54	0,75	16	70	1,00	21	
80223	RAB11FIP1	10	9	35	12	85	31	53	0,75	18	71	26	157	1,00	45	94		25				
22841	RAB11FIP2	2	1	19	7	45	0,99	12	23	0,75	6	33	0,75	9	39	0,91	17	59	0,78	16		
26056	RAB11FIP5	5	4			21	11	5	0,75	2	5	3	67	0,98	29	37	19					
201475	RAB12			17	1,00	9	13	7	10	0,75	6	16	0,75	9	9	4	14	0,98	7			
5872	RAB13	2	2	24	6	21	9	25	0,75	9	31	0,81	11	20	8	22	8					
22931	RAB18	3	2	72	1,00	22	59	0,98	20	77	1,00	22	42	18	40	0,75	15	34	13			
5861	RAB1A	4	2	41	0,78	12	14	4	32	0,75	13	34	9	24	7	15	5					
81876	RAB1B	6	6	53	14	134	0,75	20	47	0,75	15	95	16	76	16	105	0,75	20				
23011	RAB21	5	4	83	0,99	21	108	0,94	21	81	0,96	18	87	0,99	22	93	1,00	23	94	1,00	24	
51715	RAB23	1	1	31	0,97	12	26	0,82	9	20	0,85	10	21	0,85	8	27	0,94	6	20	0,78	9	
5862	RAB2A	5	4	65	19	54	0,78	16	56	0,82	16	43	14	56	0,80	15	54	18				
83871	RAB34	2	1	26	10	34	0,88	11	18	0,75	8	42	1,00	16	38	0,99	14	35	0,75	11		
11021	RAB35			13	5	12	0,98	5	7	0,75	3	11	4	15	0,99	5	14	5				
5868	RAB5A	3	2	73	1,00	17	54	1,00	14	47	0,94	13	60	1,00	18	50	0,98	17	45	0,90	14	
5869	RAB5B	1	1	19	6	19	0,83	8	16	0,75	5	24	0,95	9	18	6	13	5				
5878	RAB5C	4	3	59	0,99	14	43	10	75	1,00	17	56	0,78	15	33	11	33	11				
5870	RAB6A	5	5	69	0,91	15	72	0,98	15	69	0,95	14	69	0,76	15	72	0,85	13	82	0,92	15	
7879	RAB7A	4	4	57	0,98	22	59	0,98	21	87	1,00	27	44	18	47	16	49	0,78	16			
9367	RAB9A	3	3	91	1,00	18	76	1,00	18	91	1,00	20	71	1,00	16	97	1,00	20	69	0,85	16	
285282	RABL3	2	1	26	12	39	0,96	17	17	0,75	7	30	11	34	0,98	14	41	1,00	14			
26064	RAI14	15	15	378	1,00	108	314	1,00	89	115	0,75	51	531	1,00	119	277	1,00	83	243	1,00	81	
5901	RAN	24	23	328	1,00	32	332	1,00	30	380	1,00	31	348	1,00	36	335	1,00	33</				

Annexes | Annexe 6 : Tableau des interacteurs de l'EGFR identifiés par la technique du BioID

57142	RTN4	11	11	115	20	131	0,82	17	127	0,75	19	104	15	122	16	92	13				
56681	SARIA	4	3	69	1,00	12	71	0,99	14	64	0,99	12	55	0,90	11	74	1,00	14	63	0,75	12
51128	SARIB			18		4	23	1,00	8	39	1,00	8	32	1,00	8	43	1,00	8	36	1,00	8
6305	SBF1	2		7		4	5		3	0,75	0	34	0,90	16	4	3		6			3
9522	SCAMP1	5	4	39		18	82	1,00	25	57	0,75	22	44		18	81	0,78	23	72	0,80	25
10067	SCAMP3	4	3	39		12	66	0,99	17	50	0,75	17	69	1,00	16	44	0,78	16	41		12
23513	SCRIB	16	15	611	1,00	135	550	1,00	128	198	0,75	75	749	1,00	149	424	1,00	105	309	1,00	86
55681	SCYL2	10	10	66		29	102		37	68	0,75	25	95		34	138	1,00	37	102		32
9554	SEC22B	3	3	36		12	52	0,96	14	46	0,96	11	54	0,99	13	58	1,00	14	66	1,00	13
10802	SEC24A	4	4	16		7	55		16	29	0,75	11	44		17	63	0,95	21	81	1,00	25
10427	SEC24B	6	5	19		8	72	0,97	25	15	0,75	5	65	0,77	22	67	0,85	21	92	0,90	26
29927	SEC61A1	3	3	19		7	29		9	43	0,75	16	43	0,76	15	16		7	15		6
11231	SEC63	4	3	40		17	51	0,93	22	44	0,75	24	27		14	44		20	73	1,00	28
26168	SEN3	2	2	21		10	44	0,98	11	18	0,75	7	32	0,89	12	25		9	32		11
55752	SEPT11	12	10	174	1,00	36	147		31	104	0,75	35	214	1,00	38	129	0,82	30	103		27
4735	SEPT2	22	19	302	1,00	37	323	1,00	42	185	0,75	32	362	1,00	47	275	1,00	39	190		37
23157	SEPT6	12	11	123		31	108		23	53	0,75	22	146	0,95	31	77		23	70		23
989	SEPT7	15	14	279	1,00	38	245	1,00	40	192	1,00	33	375	1,00	46	213	1,00	36	170	0,76	33
23176	SEPT8	6	4	58		22	46		16	20	0,75	9	64	0,84	23	41		19	31		16
10801	SEPT9	16	16	260	1,00	48	248	1,00	49	151	0,75	43	351	1,00	64	219	1,00	46	196		41
871	SERPINH1	15	13	128		35	132		41	160	0,76	43	124		37	158		39	130		32
94081	SFXN1	9	8	150	1,00	29	176	1,00	33	140	1,00	28	142	1,00	32	139	1,00	29	161	1,00	33
23677	SH3BP4					14	6		6	0,75	0	18	1,00	8	17	0,99	9	11	0,98	6	5
6461	SHB	1		12		4	13	0,81	5	0,75	2	24	0,99	8	21	0,97	8	9			5
6464	SHC1	6	5	298	1,00	39	556	1,00	49	473	1,00	42	409	1,00	46	511	1,00	43	434	0,88	40
6558	SLC12A2	6	4	201	1,00	51	211	1,00	49	67	0,75	27	273	1,00	68	119	1,00	43	80	0,98	33
6566	SLC16A1	9	7	67		23	79		18	58	0,75	18	97	0,78	25	66		19	85		22
6573	SLC19A1									0,75	0	18	0,99	10							0
6510	SLCIA5	9	7	239	1,00	31	206	1,00	30	181	1,00	26	233	1,00	31	193	1,00	28	158	1,00	27
8604	SLC25A12			28	0,75	12	51	1,00	20	14	0,75	4	7		2	33	0,75	14	43	1,00	18
10165	SLC25A13	10	9	138	1,00	49	157	1,00	50	173	1,00	56	166	1,00	49	138	1,00	43	146	1,00	46
10478	SLC25A17			9		6	8		5	16	1,00	9				11		6	18		8
5250	SLC25A3	12	8	158	1,00	28	184	1,00	33	180	1,00	28	126		25	148	1,00	29	166	1,00	27
291	SLC25A4	4	4	42		15	34		13	36	0,75	16	62	0,77	18	43		16	24		11
292	SLC25A5	16	15	190	1,00	24	206	1,00	27	170	0,75	20	217	1,00	28	225	1,00	24	189	0,81	23
293	SLC25A6	4	4	77	1,00	10	85	1,00	10	61	0,76	8	57	0,81	8	59		9	88	1,00	11
7779	SLC30A1			82	1,00	23	65	1,00	17	21	0,75	9	65	1,00	22	71	1,00	15	43	1,00	11
64924	SLC30A5	2	1	21		11	40	0,96	14	15	0,75	8	42	0,99	15	38	0,99	16	31	0,91	11
55676	SLC30A6	3	2	45	0,80	16	56	1,00	19	37	0,75	14	41		14	64	1,00	18	55	1,00	17
54407	SLC38A2	3	2	118	1,00	12	132	1,00	11	109	1,00	11	79	1,00	9	88	1,00	11	82	1,00	9
57181	SLC39A10			20	1,00	12	10		7	0,75	1	30	1,00	16	21	1,00	14	10	7		7
6520	SLC3A2	1	1	151	1,00	45	79	1,00	31	86	1,00	30	100	1,00	39	104	1,00	35	52	1,00	21
9497	SLC4A7	6	6	193	1,00	41	199	1,00	45	68	0,75	23	281	1,00	64	165	1,00	38	93	0,98	26
55117	SLC6A15	3	2	70	1,00	23	68	0,99	19	44	0,93	16	98	1,00	31	60	1,00	18	42	0,92	15
6535	SLC6A8	1	1	63	1,00	21	43	1,00	14	31	0,99	12	58	1,00	19	34	0,99	15	22	0,88	7
6542	SLC7A2	1		10		6	20	0,83	9	0,75	0	25	0,75	9							0
8140	SLC7A5	2	1	31	0,93	10	28	0,92	8	13	0,75	6	40	1,00	10	23		7	24		8
60682	SMAP1	6	4	70	0,99	10	76	0,92	12	87	1,00	12	70	0,92	13	72	0,85	12	63	0,76	12
64744	SMAP2	6	4	56		16	100	1,00	22	121	1,00	21	72	0,98	22	120	0,99	21	90	0,84	20
55627	SMPD4	3	1	8		4	14		6	0,75	2	17		8	22	8		37	0,89	19	
8773	SNAP23	4	3	151	1,00	28	130	1,00	27	77	1,00	25	159	1,00	34	93	1,00	29	68	1,00	22
6616	SNAP25	3		2		2			1	0,75	0	16	1,00	8							0
9342	SNAP29	8	8	89		25	102	0,88	24	58	0,75	22	135	1,00	27	91		23	90		23
6638	SNRPN	6	6	69		13	72		12	77	0,92	14	78	0,86	14	68		11	61		13
6641	SNTB1	1	1	24	0,78	16	11		8	7	0,75	4	36	0,99	19						0
6645	SNTB2	3	3	41		21	22		11	12	0,75	6	66	0,99	25	14		9	13		7
6642	SNX1	17	16	165		46	226	0,96	49	164	0,75	46	234	1,00	55	254	1,00	58	204	0,75	50
9784	SNX17	2	2	19		9	14		8	10	0,75	4	24		12	62	1,00	20	25		14
6643	SNX2	8	7	83		27	112	0,90	34	76	0,75	26	94		33	121	1,00	35	75		28
8724	SNX3	2	1	9		5	10		7	41	0,99	16	12		4	34	0,96	11	30	0,90	13
58533	SNX6	5	3	23		8	44		17	37	0,75	16	21		11	50	0,84	16	33		14
6674	SPAG1	1	1			17	8		2	0,75	1	11		4	31	1,00	9	34	0,96	10	
9789	SPCS2	1		11		5	21	0,87	7	18	0,87	7			3	16	0,90	7	14		6
6709	SPTAN1	12	8	165	0,76	72	142		60	29	0,75	14	240	1,00	100	87		37	53		22
6711	SPTBN1	5	5	67	0,80	34	47		23	8	0,75	6	112	1,00	57	22		13	9		5
10558	SPTLC1			10		6	9		5			12	0,98	6					12		6
6723	SRM	4	2	51	0,97	19	54	0,85	20	34	0,75	14	65	0,76	26	55	0,99	21	62	0,85	22
6734	SRPR	9	8	57		27	140	1,00	47	89	0,75	36	120	1,00	40	200	1,00	55	240	1,00	53
58477	SRPRB	4	4	52		16	66	1,00	19	58	0,80	20	67	0,95	19	86	1,00	22	79	1,00	22
6427	SRSF2	3	3	30		11	19		9	43	0,93	12	23		9	16		9	26		10
6742	SSBP1	1	1	92	1,00	19	59	0,98	15	81	1,00	19	43	0,99	14	70	1,00	19	62	1,00	15
6748	SSR4	5	3	63	0,95	15	82	1,00	14	57	0,87	12	61	0,88	15	69	0,99	15	90	1,00	16
10254	STAM2	3	2	5		4	33		14	5	0,75	4	23		12	40	0,96	18	50	0,99	16
6774	STAT3	24	23	71		25	176		55	84	0,75	36	162		53	382	1,00	85	270		68
6777	STAT5B	5	2				185	1,00	51	126	1,00	41	46		24	553	1,00	83	470	1,00	84
55240	STEAP3	3	2	85	1,00	24	89	1,00	23	43	0,75	16	102	1,00	25	90	1,00	24	39	0,93	16
57620	STIM2																				

Annexes | Annexe 6 : Tableau des interacteurs de l'EGFR identifiés par la technique du BioID

7083	TK1	2	2			10	6	5	0,75	3	8	4	41	0,98	9	48	0,89	8				
56889	TM9SF3	3	3	25	10	46	0,77	15	23	0,75	9	41	16	35	7	22		8				
10972	TMED10	2	1	21	6	17			34	0,97	8	18	5	29	0,92	8	16	5				
147007	TMEM199					13	0,98	5		0,75	1	10		5	21	1,00	6	14	0,99	6		
23670	TMEM2			3	2	6		4		0,75	0	2	2	23	0,75	14	11	0,98	6			
55161	TMEM33	3	3	22	10	28		9	36	0,75	15	51	0,79	14	45	0,78	14	49	0,96	13		
55706	TMEM48	4	3	31	13	54	0,90	25	43	0,75	18	63	1,00	25	68	1,00	26	73	1,00	27		
55092	TMEM51											13	0,98	6								
55219	TMEM57	1							17	0,93	8	9	4	28	0,98	13	45	1,00	18			
7112	TMPO	32	25	578	0,75	55	556	0,77	46	201	0,75	19	466	1,00	49	483	0,75	34	412	0,75	35	
81542	TMX1	1	1	22	0,85	9	21		7	34	0,99	10	27	0,96	10	11		4	19	0,77	9	
3371	TNC					12	0,75	5		0,75	4	17	0,99	4	14	0,99	4					
23043	TNIK	6	4	27		12	55	25	26	0,75	11	65	0,78	26	63		28	31		16		
23534	TNPO3	3	2			38	14		27	0,75	7	21		9	20		10	59	1,00	16		
56993	TOMM22	8	8	135	1,00	18	152	1,00	18	150	1,00	20	113	0,99	17	130	1,00	16	136	1,00	17	
26092	TOR1AIP1	4	3	50	0,87	13	63	1,00	20	48	0,75	17	50	0,87	14	98	1,00	30	106	1,00	26	
10131	TRAP1	24	23	349	1,00	71	335	1,00	74	357	1,00	77	359	1,00	74	305	1,00	67	336	1,00	70	
10206	TRIM13	2	2	17		7	32	0,92	9	15	0,75	4	30		10	21		10	26		10	
11078	TRIOBP	1		17		10						25	0,94	10								
9319	TRIP13	8	5	112	1,00	38	162	1,00	37	120	1,00	38	116	1,00	33	147	1,00	35	135	1,00	38	
10102	TSFM	3	2	51	0,94	19	52	1,00	15	36	0,84	14	45	0,94	14	53	0,99	19	34		11	
54777	TTC40			4		2	9		3		0,75	0	10		4	18		8	18	0,75	6	
145567	TTC7B	1	1	19		9	31	0,99	10	11	0,75	4	44	1,00	15	26	0,78	12	9		5	
203068	TUBB	22	21	427	1,00	16	466	1,00	16	456	1,00	17	416	1,00	16	444	1,00	16	471	1,00	16	
34733	TUBB2B	7	7	118	1,00	20	128	1,00	18	155	1,00	17	134	1,00	22	86	0,77	19	123	1,00	21	
10382	TUBB4A			62		7	181	0,75	15	50	0,75	6	162	1,00	17	56		6	203	1,00	17	
10383	TUBB4B	12	11	134		11	111		9	197	0,78	14	55		6	159	0,75	12	67		5	
84617	TUBB6	8	6	114	1,00	22	167	1,00	32	139	1,00	30	153	1,00	34	161	1,00	27	198	1,00	31	
7283	TUBG1	6	5	79	0,78	23	95	0,91	30	59	0,75	23	105	1,00	32	90	1,00	29	103	1,00	32	
7284	TUFM	26	23	318	1,00	57	374	1,00	61	335	1,00	60	318	1,00	62	331	1,00	55	339	1,00	59	
7289	TULP3								2	0,75	2	18	0,75	8	11	0,98	5					
51465	UBE2J1	4	2			24	10		10	0,75	5	17	6	75	1,00	21	61	0,99	19			
7334	UBE2N	2	2	37	0,96	10	55	1,00	12	28	0,75	12	28		10	61	1,00	11	53	1,00	14	
27338	UBE2S					5	3		6	0,75	3			8	4	18	1,00	8				
29914	UBIAD1	2	2	11		5	29	0,75	6	0	0,75	3	27		5	20		6	14		6	
51035	UBXN1	4	4	24		10	29		12	23	0,75	10	31	14	56	0,97	15	42	15			
23190	UBXN4	2	2	33	0,80	7	43	0,99	11	22	0,75	8	34	0,76	7	46	0,91	10	29		9	
7358	UGDH	3	2	17		9	32		17	33	0,75	18	9	6	45	0,94	19	57	1,00	23		
7372	UMPS	1		9		5	15	0,82	8	13	0,81	7	5		4	19	0,97	9	16	0,83	9	
10497	UNC13B	1								0,75	0	20	0,79	11								
55898	UNC45A	17	13	23		12	76		31	73	0,75	31	96		39	91		38	183	0,95	63	
219699	UNC5B	1		24	0,88	12	79	1,00	27	28	0,88	12	44	1,00	19	42	1,00	17	20	0,97	12	
8633	UNC5C			24	1,00	7						19	1,00	6								
7385	UQCRC2	15	14	216	1,00	44	247	1,00	45	185	0,75	41	189	1,00	40	209	0,97	43	218	1,00	43	
7386	UQCRCF1	2	1	36	0,98	11	39	0,99	11	20	0,75	8	46	1,00	10	30	0,95	11	29	0,77	11	
9712	USP6NL			241	1,00	56	201	1,00	54	90	1,00	33	267	1,00	71	167	1,00	41	105	1,00	34	
7402	UTRN	30	30	725	1,00	222	600	1,00	189	333	0,75	131	923	1,00	246	454	1,00	165	455	1,00	161	
9341	VAMP3	5	5	80	0,75	7	62		7	69	0,84	8	77	1,00	8	75	0,75	7	73	0,91	8	
81839	VANGL1	1	1	117	1,00	40	119	1,00	37	16	0,75	8	210	1,00	53	67	1,00	26	31		14	
57216	VANGL2	1		41	1,00	19	49	1,00	18		0,75	0	64	1,00	28	15		7				
9218	VAPA	12	11	130	0,79	12	190	1,00	16	138	0,79	9	131		14	216	1,00	13	266	1,00	14	
9217	VAPB	9	8	115	0,85	18	119	0,99	17	122	1,00	19	115	0,99	19	147	1,00	19	138	1,00	17	
7410	VAV2					13	0,98	5						37	1,00	16		27	1,00	11		
10451	VAV3									0,75	0											
7416	VDAC1	16	15	177	0,97	32	151		26	174	0,75	32	159		31	126		28	147		28	
7419	VDAC3	14	13	194	1,00	26	178	1,00	24	182	1,00	30	153	0,80	23	156	0,87	24	150		26	
55591	VEZT	1				10		5		0,75	0	26	0,99	9	23	0,88	8	17	0,86	5		
7431	VIM	69	67	621		80	745	0,75	83	721	0,75	84	735	1,00	87	766	1,00	84	800	1,00	82	
65082	VPS33A	3	2			29	11		11	24	0,75	9	7	4	68	1,00	26	64	0,99	18		
11311	VPS45	5	3			49	20		20	30	0,75	14	28	14	56	0,81	24	35	14			
23355	VPS8	2	2			4			4		0,75	0			34	0,87	18	5	3			
8936	WASF1	1								0,75	0	27	0,99	14	12		7	8		7		
10163	WASF2	6	6	84	0,88	19	69		16	39	0,75	13	123	1,00	17	67		17	43		14	
11180	WDR6	4	2	104	1,00	43	111	1,00	35	84	1,00	32	184	1,00	58	116	1,00	37	82	1,00	29	
79084	WDR77	6	6	60		20	75	0,80	20	25	0,75	9	109	1,00	23	37		14	15		10	
51741	WWOX	31	30	447	1,00	37	411	1,00	37	353	0,99	34	392	1,00	40	348		36	400	1,00	37	
63929	XPNPEP3			21	0,75	6																
7514	XPO1	19	16	176		55	222		62	219	1,00	60	210	0,98	60	162		49	208	0,95	64	
11260	XPOT	9	8	100		38	125	1,00	38	109	0,78	40	116	1,00	38	98	0,78	31	127	1,00	40	
9213	XPR1			13		7	15	0,98	7	3	0,75	3	30	1,00	14	13		5	8		4	
8565	YARS	17	17	157		48	186		48	207	0,93	55	162		44	162		46	195	0,97	51	
51067	YARS2	2	2	52	1,00	18	42	0,95	16	49	0,95	19	45	0,99	15	41	0,89	16	77	1,00	22	
7525	YES1	3		66	0,97	25	37	0,78	15	26	0,75	14	48	0,96	21	23		15	31		11	
90522	YIF1B									2	0,75	2								13	0,99	5
10652	YKT6	13	13	254	1,00	24	265	1,00	21	169	1,00	22	284	1,00	25	250	1,00	24	232	1,00	21	
7529	YWHAB	5	4	68	0,80	8	49		5	0,75	4	87	1,00	10	64	0,98	9	61	6		6	
7531	YWHAE	17	15	158		26	158		23	95	0,75	22	177	0,93	27	167		26	175	0,79	28	
7532	YWHAG	6	5	29		9	34		8			72	0,83	12	36		8	39		8		
7534	YWHAZ	8	7	88		13	86		16	65	0,75	11	100	0,99	16	89	0,85	16	104	0,98		

Annexe 7 : Tableau des interacteurs de BRAF identifiés par la technique du BioID

Gene ID	Gene Name	Top 2 Controls	HEK293_BRAF WT-BF			HEK293_BRAF V600E-BF			Log2 fold change		
			Total	SAINT	Total Unique	Total	SAINT	Total Unique			
948469	brA	250	14626	-	84	13840	-	81			
673	BRAF	2	9246	-	113	8483	-	104			
23456	ABC10	2	39	0,90	22	38	0,82	16			
22	ABC7	1	60	1,00	28	58	1,00	25			
5825	ABCD3	14	195	0,99	49	112	0,99	33			
6059	ABCE1	9	278	1,00	46	290	1,00	43			
31	ACACA	314	5239	1,00	3311						
32	ACACB	58	1366	1,00	380				-1,84		
11332	ACOT7	8	110	0,97	28	110	0,96	30			
23597	ACOT9	6	113	1,00	30	121	1,00	25			
56997	ADCK3	3	55	0,99	19	35	0,87	18			
54936	ADPRHL2	0	21	1,00	12	12	1,00	7			
11047	ADRM1	1	22	1,00	11	19	1,00	11			
8540	AGPS	1	32	1,00	16	26	1,00	15			
10598	AHSA1	14	260	1,00	39	306	1,00	46			
9131	AIFM1	2	47	0,98	22	30	0,91	17			
9049	AIP	2	1			35	0,98	12	4,17		
208	AKT2	1	31	0,99	14	40	1,00	15			
5832	ALDH18A1	19	507	1,00	94	454	1,00	88			
219	ALDH1B1	1	87	1,00	31	60	1,00	26			
217	ALDH2	0	22	1,00	13	12	0,98	9			
501	ALDH7A1	6	76	0,94	36	44	0,94	23			
29880	ALG5	3	46	0,97	11	33	0,82	10			
302	ANXA2	21	129	1,00	40	230	0,86	56			
26985	AP3M1	7	119	1,00	35	68	1,00	31			
27350	APOBEC3C	0	29	1,00	11	7		4	-1,91		
369	ARAF	2	56	1,00	14	62	1,00	22			
372	ARCN1	13	156	0,93	44	147	0,93	38			
375	ARF1	15	238	1,00	18	204	1,00	16			
378	ARF4	3	92	1,00	20	81	1,00	17			
381	ARF5	3	110	1,00	12	91	1,00	11			
382	ARF6	4	64	1,00	18	47	1,00	17			
9138	ARHGEF1	0	19	0,99	11	11		6			
9181	ARHGEF2	8	78	1,00	36	118	1,00	53			
400	ARL1	7	125	1,00	27	73	1,00	21			
402	ARL2	5	66	0,95	20	31		16	-1,07		
55207	ARL8B	3	55	0,98	16	49	0,97	14			
445	ASS1	12	200	1,00	27	170	0,99	27			
55210	ATAD3A	4	206	1,00	49	94	1,00	20	-1,12		
476	ATP1A1	8	191	1,00	51	226	1,00	50			
483	ATP1B3	1	9		4	25	1,00	8	1,38		
488	ATP2A2	24	364	1,00	79	407	1,00	69			
490	ATP2B1	4	25		15	56	0,99	29	1,13		
498	ATP5A1	53	1221	1,00	85	1236	1,00	83			
506	ATP5B	31	1389	1,00	94	1533	1,00	91			
509	ATP5C1	9	229	1,00	32	185	1,00	33			
515	ATP5F1	2	27		12	31	0,82	14			
100526740	ATP5J2-PTCD1	2	64	1,00	8	47	1,00	9			
10632	ATP5L	3	71	1,00	9	59	0,98	9			
539	ATP5O	7	107	0,77	18	86	0,82	20			
51606	ATP6V1H	4	49	0,95	15	38		12			
25814	ATXN10	5	3		21	55	0,80	20			
56647	BCCIP	4	66	0,97	22	98	1,00	23			
644	BLVRA	6	126	1,00	31	84	1,00	22			
682	BSG	2	29		10	31	0,79	12			
79096	C11orf49	16	183	1,00	32	166	1,00	27			
90488	C12orf23	1	22	1,00	10	17		8			
51637	C14orf166	12	213	1,00	32	147	0,95	28			
125988	C19orf70	3	48	1,00	11	27	1,00	10			
25980	C20orf4	3	34	0,77	16	24		9			
8209	C21orf33	1	42	1,00	15	37	1,00	14			
51493	C22orf28	19	298	1,00	64	280	1,00	61			
79624	C6orf211	0	46	1,00	17	41	1,00	18			
84267	C9orf64	4	74	1,00	28	55	0,99	17			
790	CAD	17	324	1,00	112	247	1,00	94			
813	CALU	1	30	1,00	14	22		10			
55832	CAND1	21	485	1,00	86	402	1,00	82			
821	CANX	7	204	1,00	46	198	1,00	48			
823	CAPN1	4	68	0,97	26	25		11	-1,41		
11140	CDC37	7	291	1,00	29	399	1,00	30			
8318	CDC45	0	26	1,00	13	10		5	-1,30		
10423	CDIPT	3	67	1,00	15	68	1,00	15			
983	CDK1	20	474	1,00	53	428	1,00	49			
1020	CDK5	2	26	0,94	12	11		7	-1,17		
27440	CECR5	3	49	0,87	14	20		9	-1,25		
1072	CFL1	11	209	1,00	28	219	1,00	27			
25978	CHMP2B	1	18	1,00	8	21		8			
26973	CHORDC1	7	34		10	83	0,86	14	1,26		
26097	CHTOP	2	26		9	41	0,99	13			
1153	CIRBP	3	45	0,92	15	40	0,81	11			
8192	CLPP	3	37		12	53	0,99	14			
10845	CLPX	1	34	1,00	19	48	1,00	22			
80347	COASY	4	74	0,96	24	40		15			
1314	COPA	11	139	0,98	51	121		47			
1315	COPB1	25	286	1,00	68	265		70			
11316	COPE	4	61	1,00	20	30		15	-1,00		
22820	COPG1	14	165	0,98	45	151	0,85	48			
22818	COPZ1	5	50		16	59	0,78	16			
57175	CORO1B	67	673	1,00	52	977	1,00	56			
23603	CORO1C	18	275	1,00	34	364	1,00	36			
4513	COX2	6	83	0,83	12	50		13			
8904	CPNE1	2	40	0,99	17	22		7			
1434	CSEIL	17	158		64	191	0,93	65			
1466	CSR2	4	52	0,98	12	41		12			
1488	CTBP2	6	84	1,00	20	98	1,00	23			
1503	CTPS	30	405	1,00	60	469	1,00	56			
1508	CTSB	1	14		6	31	1,00	10	1,09		
1727	CYB5R3	5	35		16	69	0,86	22			
1603	DAD1	2	2		29	0,79	11	28	9	=	
1615	DARS	13	11		279	1,00	64	253	1,00	63	=
55157	DARS2	4	3		114	1,00	39	112	1,00	40	=
79877	DAKD	0	0		24	1,00	9	12		5	=
1650	DDOST	6	6		164	1,00	43	166	1,00	41	=
1717	DHCR7	2	1		26	1,00	10	18		6	=
1719	DHFR	4	3		114	1,00	23	67	1,00	15	=
25979	DHRS7B	5	3		58	0,90	16	65	0,83	17	=
22907	DHX30	5	4		197	1,00	61	142	1,00	48	=
1660	DHX9	50	49		626	1,00	125	663	1,00	120	=
27292	DIMT1	4	4		52		23	60	0,83	20	=
1737	DLAT	1	1		36	1,00	17	50	1,00	18	=
1738	DLD	4	3		115	1,00	28	126	1,00	28	=
1785	DNM2	4	4		114	1,00	45	123	1,00	40	=
8813	DPM1	5	5		141	1,00	28	88	1,00	27	=
1778	DYNC1H1	32	31		731	1,00	254	489	1,00	194	=
1891	ECH1	4	3		71	1,00	18	16		10	-2,08
10455	ECI2	1	0		46	1,00	15	28	1,00	12	=
30844	EHD4	2	1		45	0,99	23	31	0,96	14	=
1994	ELAVL1	10	9		148	1,00	41	141	1,00	37	=
2023	ENO1	75	66		3153	1,00	58	3080	1,00	55	=
83481	EPPK1	9	6		287	1,00	86	269	1,00	74	=
2058	EPRS	33	30		367	1,00	112	318		93	=
11160	ERLIN2	6	5		79	0,97	30	62		24	=
9415	FADS2	1	1		20	1,00	9			-2,81	=
25940	FAM98A	4	4		84	1,00	25	58	0,83	22	=
84188	FAR1	2	2		53	1,00	20	35	0,78	13	=
60493	FASTKD5	2	1		21	0,82	10	15		6	=
10211	FLOT1	1	1		59	1,00	23	6		4	-3,10
2584	GALK1	3	3		86	1,00	27	82	1,00	24	=
2597	GAPDH	35	29		481	1,00	45	546	1,00	48	=
2617	GARS	5	4		296	1,00					

Annexes | Annexe 7 : Tableau des interacteurs de BRAF identifiés par la technique du BioID

56922	MCCC1	254	238	3358	1,00	108	2290	97	=	
64087	MCCC2	158	144	2218	1,00	102	1782	1,00	91	=
90550	MCU	1	0	23	1,00	10	8	6	-1,42	
57496	MKL2	61	56	648	1,00	80	272	52	-1,25	
64210	MMS19	1	1	45	1,00	14	35	1,00	15	=
9526	MPDU1	2	1	26	0,90	8	14	6	=	
124540	MSI2	0	0	6		5	17	1,00	10	1,36
4522	MTHFD1	24	18	681	1,00	126	625	1,00	114	=
25902	MTHFD1L	4	2	56	0,94	23	19	13	-1,51	
10797	MTHFD2	1	1	19	1,00	8	6	3	-1,51	
55149	MTPAP	1	1	42	1,00	20	27	1,00	15	=
4641	MYO1C	4	4	52	1,00	15	48	0,82	15	=
10135	NAMPT	13	12	349	1,00	65	388	1,00	68	=
4673	NAP1L1	35	30	418	1,00	41	390	0,84	40	=
4691	NCL	25	24	290	1,00	52	286	1,00	53	=
56926	NCLN	4	4	79	1,00	30	59	22	=	
4705	NDUFA10	5	5	97	0,99	25	66	0,91	20	=
51079	NDUFA13	2	1	29	0,95	12	23	0,80	13	=
4698	NDUFA5	3	2	62	1,00	15	27	6	-1,17	
4704	NDUFA9	5	4	106	1,00	35	97	1,00	36	=
4719	NDUFS1	3	2	100	1,00	45	81	1,00	34	=
4720	NDUFS2	1	1	59	1,00	26	65	1,00	22	=
4722	NDUFS3	5	3	127	1,00	34	158	1,00	29	=
374291	NDUFS7	2	1	35	0,98	8	29	0,96	8	=
4723	NDUFV1	0	0	38	1,00	16	36	1,00	13	=
4729	NDUFV2	1	0	44	1,00	16	22	8	=	
51068	NMD3	2	1	28	0,93	14	53	1,00	23	=
23530	NNT	5	5	88	0,98	40	71	0,78	33	=
64318	NOC3L	6	5	66		24	79	0,78	27	=
4841	NONO	43	38	538	1,00	55	532	1,00	50	=
51602	NOP58	15	13	193	1,00	34	238	1,00	34	=
9520	NPEPPS	9	8	333	1,00	89	283	1,00	84	=
4869	NPPI	47	43	478	20	600	1,00	20	=	
4893	NRAS	0	0	39		8	72	1,00	16	=
4905	NSF	7	5	77	0,86	31	54	26	=	
54888	NSUN2	13	12	346	1,00	74	357	1,00	86	=
64943	NTSDC2	1	0	41	1,00	21	4	3	-3,07	
84284	NTPCR	4	2	57	0,94	13	18	9	-1,61	
10101	NUBP2	3	2	57	1,00	17	33	14	=	
11163	NUDT4	1	1	14	0,99	9	19	8	=	
23279	NUP160	4	4	58	0,93	31	72	1,00	39	=
23165	NUP205	11	9	126	0,88	52	137	52	=	
79902	NUP85	5	4	52		21	61	0,83	24	=
9688	NUP93	6	5	190	1,00	64	157	1,00	58	=
4942	OAT	3	3	281	1,00	48	234	1,00	45	=
5094	PCCBP2	21	19	259	1,00	30	276	1,00	29	=
5095	PCCA	225	196	3044	1,00	130	1840	1,00	119	=
5096	PCCB	180	147	2173	1,00	92	1478	86	=	
5111	PCNA	4	3	111	1,00	23	95	1,00	18	=
11333	PDAP1	10	9	91		19	142	1,00	22	=
10016	PDCC6	7	5	80	0,96	23	58	21	=	
10015	PDCC6IP	11	11	234	1,00	70	212	1,00	63	=
5160	PDHA1	2	1	65		18	54	0,99	24	=
5162	PDHB	3	2	109	1,00	37	106	1,00	33	=
5163	PKD1	0	0	23	1,00	9	10	5	-1,13	
5165	PKD3	0	0	15	1,00	10	0	0	-4,00	
53918	PELO	2	1	45	0,99	18	58	1,00	19	=
5211	PFKL	16	11	339	1,00	62	206	1,00	48	=
5213	PFKM	7	6	107	1,00	31	64	24	=	
5214	PFKP	8	7	223	1,00	47	242	1,00	51	=
5245	PHB	16	15	285	1,00	42	294	1,00	43	=
11331	PHB2	23	21	522	1,00	59	525	1,00	56	=
84295	PHF6	1	1	27	1,00	9	19	7	=	
26227	PHGDH	16	16	315	1,00	45	332	1,00	41	=
128869	PIGU	2	1	28	1,00	11	16	4	=	
5315	PKM2	59	58	930	1,00	77	953	1,00	77	=
5351	PLOD1	1	0	55	1,00	25	45	1,00	24	=
56902	PNO1	2	1	41	0,99	15	29	0,96	14	=
5464	PPA1	17	13	558	1,00	64	542	1,00	66	=
5515	PP2CA	5	4	96	1,00	19	109	1,00	20	=
5520	PP2R2A	20	20	226	0,97	40	221	39	=	
5052	PRDX1	29	29	302		39	355	1,00	39	=
5580	PRKCD	3	0	9		4	31	0,80	18	1,68
5589	PRKCSH	2	1	22	0,87	8	8	5	-1,35	
5591	PRKDC	82	65	1803	1,00	383	1421	1,00	334	=
5634	PRPS2	3	1	66	0,99	19	58	0,98	13	=
51495	PTPLAD1	19	16	232	1,00	29	215	0,81	27	=
80324	PUS1	5	5	58		24	65	0,84	27	=
5831	PYCR1	1	1	24	1,00	12	32	1,00	11	=
29920	PYCR2	2	1	37	0,96	13	27	9	=	
5834	PYGB	6	3	86	1,00	31	97	1,00	30	=
5836	PYGL	13	12	177	1,00	56	125	39	=	
9230	RAB11B	4	3	56		23	77	1,00	26	=
201475	RAB12	0	0	24	1,00	10	30	1,00	13	=
5872	RAB13	2	2	41	0,98	13	29	11	=	
51552	RAB14	8	7	105	0,84	27	106	0,97	29	=
22931	RAB18	3	2	89	1,00	27	36	15	-1,28	
5861	RAB1A	4	2	38	0,87	11	46	0,93	12	=
81876	RAB1B	6	6	114	1,00	15	60	15	=	
23011	RAB21	5	4	70	0,99	16	96	1,00	19	=
5862	RAB2A	5	4	57	0,90	18	58	0,87	17	=
83871	RAB34	2	1	52	1,00	19	56	1,00	18	=
5865	RAB3B	1	1	19		7	30	1,00	12	=
5868	RAB5A	3	2	94	1,00	17	62	1,00	18	=
5869	RAB5B	1	1	32	1,00	10	21	1,00	11	=
5878	RAB5C	4	3	81	1,00	16	72	0,97	17	=
5870	RAB6A	5	5	87	1,00	22	70	0,97	16	=
7879	RAB7A	4	4	66	0,97	25	80	1,00	23	=
9367	RAB9A	3	3	85	1,00	22	85	1,00	20	=
5894	RAF1	4	3	21		10	153	1,00	28	2,81
22913	RALY	2	1	10		4	21	0,83	13	1,00

5901	RAN	24	23	452	1,00	41	474	1,00	40	=
10432	RBM14	10	10	220	1,00	53	193	1,00	50	=
9584	RBM39	13	11	157	1,00	30	180	1,00	35	=
27316	RBMX	5	4	58		15	56	0,90	16	=
5955	RCN2	0	0	37	1,00	16	14	7	-1,34	
54931	RG9MTD1	6	5	68	0,84	25	100	1,00	33	=
6009	RHEB	5	4	69	0,98	15	64	0,99	14	=
89941	RHOT2	2	1	29	0,91	14	21	0,85	12	=
55781	RIOK2	0	0	50	1,00	22	13	7	-1,87	
6184	RPN1	19	17	247	1,00	64	292	1,00	69	=
6185	RPN2	12	11	433	1,00	67	437	1,00	68	=
57521	RPTOR	3	2	54	1,00	27	47	0,82	21	=
22800	RRAS2	1	0	25	1,00	8	22	8	=	
6240	RRM1	4	3	53	0,97	24	33	16	=	
8634	RTCD1	1	1	25	1,00	11	39	1,00	14	=
56681	SARIA	4	3	82	1,00	12	86	1,00	15	=
9554	SEC22B	3	3	30		9	41	0,81	13	=
29927	SEC61A1	3	3	96	1,00	20	83	0,98	19	=
11231	SEC63	4	3	49	0,97	24	35	17	=	
26168	SENP3	2	2	37	0,94	12	7	3	-2,25	
871	SERPINH1	15	13	187	1,00	52	196	1,00	48	=
6418	SET	26	24	298	1,00	30	317	0,77	28	=
6421	SFPQ	38	37	399		63	466	1,00	62	=
94081	SFXN1	9	8	207	1,00	38	192	1,00	35	=
6472	SHMT2	15	14	140		41	172	0,79	39	=
22933	SIRT2	1	0	20	1,00	11	12	8	=	
6499	SKIV2L	5	5	60	0,89	36	34	21	=	
6510	SLC1A5	9	7	162	1,00	25	124	0,99	22	=
6576	SLC25A1	5	5	87	1,00	26	41	14	-1,07	
1468	SLC25A10	4	2	24	0,90	11	16	6	=	
8402	SLC25A11	11	10	179	1,00	40	132	0,79	33	=
8604	SLC25A12	0	0	56	1,00	29	54	1,00	23	=
10165	SLC25A13	10	9	282	1,00	63	268	1,00	59	=
5250	SLC25A3	12	8	170	1,00	27	196	1,00	27	=
291	SLC25A4	4	4	86	1,00	20	81	1,00	19	=
292	SLC25A5	16	15	240	1,00	28	246	1,00	28	=
293	SLC25A6	4	4	57	0,96	8	44	8	=	
8243	SMC1A	15	13	200	1,00	76	191	1,00	66	=
9126	SMC3	14	13	219	1,00	87	181	0,76	76	=
6607	SMN2	3	3	39	0,92	8	64	1,00	19	=
6638	SNRPN	6	6	89	1,00	13	90	1,00	13	=
6643	SNX2	8	7	89	0,93	35	67	27	=	
6723	SRM	4	2	88	1,00	28	74	1,00	24	=
6727	SRP14	3	3	49	0,99	18	31	8	=	
58477	SRPRB	4	4	44		18	62	0,83		

Annexe 8 : Tableau des interacteurs de l'étude des fusions ALK identifiés par la technique du BioID

Gene ID	Gene Name	Top 2 controls	HEK293_FB-EML4 WT			HEK293_FB-EML4-ALK-V3a			HEK293_EML4-ALK-V3a-BF			
			Total	SAINT	Total Unique	Total	SAINT	Total Unique	Total	SAINT	Total Unique	
948469	birA	2528	2528	17147	-	104	16398	-	111	12857	-	81
238	ALK	0	0	0	0	0	7884	-	128	5051	-	112
2009	EML1	0	0	996	-	117	517	-	84	293	-	69
24139	EML2	0	0	42	-	17	31	-	14	0	-	0
6059	ABCE1	9	9	53	-	14	108	-	26	122	0.94	32
3983	ABLIM1	9	8	4	-	2	122	0.98	44	71	-	33
10598	AHSA1	14	13	122	-	24	168	1.00	33	168	1.00	31
9131	AIFM1	2	0	12	-	7	13	-	6	26	0.75	10
208	AKT2	0	0	0	-	0	27	1.00	9	9	-	6
5832	ALDH18A1	19	17	141	-	37	227	0.75	54	177	-	52
7840	ALMS1	6	5	0	-	0	115	1.00	47	45	-	20
23243	ANKRD28	11	7	0	-	0	151	1.00	37	76	-	19
283373	ANKRD52	4	3	0	-	0	60	0.99	28	28	-	14
10541	ANP32B	6	4	125	1.00	17	89	0.95	12	68	0.75	9
81611	ANP32E	8	8	186	1.00	28	84	-	13	103	0.83	19
301	ANXA1	3	0	0	-	0	61	0.97	21	69	0.98	21
302	ANXA2	21	19	131	-	40	387	1.00	66	335	1.00	64
160	AP2A1	8	7	12	-	5	125	1.00	31	75	-	25
1173	AP2M1	7	6	4	-	2	97	0.97	26	39	-	8
322	APBB1	8	6	0	-	0	90	0.87	19	76	-	15
324	APC	2	0	0	-	0	39	0.96	17	30	0.85	18
375	ARF1	15	13	139	-	16	178	0.98	16	176	-	16
378	ARF4	3	3	63	1.00	13	42	-	12	52	0.95	13
381	ARF5	5	4	36	-	5	19	-	2	52	-	10
55738	ARFGAP1	15	14	159	-	28	228	1.00	34	199	0.99	31
7984	ARHGEF5	0	0	0	-	0	0	-	0	32	1.00	15
488	ATP2A2	24	23	124	-	39	183	-	44	305	0.99	60
498	ATP5A1	53	51	646	1.00	69	605	0.75	63	655	1.00	65
506	ATP5B	31	29	456	1.00	68	390	-	70	622	1.00	75
509	ATP5C1	9	8	108	0.97	22	109	0.95	23	127	0.76	21
100526740	ATP5J2-PTCD1	2	2	49	0.99	9	10	-	4	0	-	0
148229	ATP8B3	0	0	14	-	2	9	-	2	0	-	0
22994	AZ11	0	0	0	-	0	18	0.75	7	9	-	6
9532	BAG2	5	4	33	-	9	49	-	13	36	-	11
9529	BAG5	0	0	0	-	0	3	-	2	0	-	0
9564	BCAR1	3	2	0	-	0	69	1.00	24	56	0.99	22
8412	BCAR3	0	0	0	-	0	25	1.00	13	51	1.00	18
23299	BICD2	3	0	0	-	0	53	0.94	21	51	0.95	25
55352	C17orf79	3	2	0	-	0	49	0.96	8	10	-	4
79002	C19orf43	12	11	7	-	4	4	-	2	76	-	12
340277	C7orf46	2	2	2	-	2	38	0.96	10	19	-	7
157922	CAMSAP1	19	14	141	-	50	301	1.00	79	351	1.00	90
23271	CAMSAP2	4	4	26	-	11	51	0.78	23	70	0.79	25
821	CANX	7	7	37	-	19	35	-	15	90	0.94	25
867	CBL	9	8	5	-	3	154	0.78	41	128	0.99	35
124808	CCDC43	10	9	92	-	18	116	0.76	23	116	0.79	21
10694	CCT8	217	211	1283	-	100	1174	-	99	1653	-	107
11140	CDC37	7	6	8	-	3	97	0.98	16	58	-	17
8318	CDC45	0	0	0	-	0	0	-	0	16	1.00	8
983	CDK1	20	17	157	-	33	219	0.75	35	236	1.00	38
55835	CENPJ	0	0	0	-	0	24	1.00	14	3	-	2
22995	CEP152	0	0	0	-	0	23	1.00	13	4	-	2
9859	CEP170	41	37	537	1.00	119	699	0.75	113	292	-	48
55125	CEP192	0	0	0	-	0	170	1.00	64	46	1.00	26
80817	CEP44	4	3	51	0.79	16	52	0.97	13	44	-	11
64793	CEP85	3	2	0	-	0	55	0.96	20	22	-	12
79598	CEP97	3	3	0	-	0	47	0.76	17	19	-	9
1072	CFL1	11	10	123	-	24	128	0.96	24	134	0.98	23
57530	CGN	0	0	0	-	0	28	1.00	13	6	-	4
25978	CHMP2B	0	0	23	1.00	9	20	0.75	6	11	-	4
26586	CKAP2	12	9	156	0.99	39	136	0.87	33	201	1.00	36
1266	CNN3	24	20	208	-	41	432	1.00	42	282	0.98	40
57175	CORO1B	67	64	801	1.00	66	714	-	49	464	-	45
23603	CORO1C	18	16	285	1.00	38	357	1.00	37	248	1.00	31
51692	CPSF3	15	14	82	-	35	249	1.00	66	89	-	33
1398	CRK	45	43	361	-	45	1223	1.00	64	1354	1.00	55
1399	CRKL	47	42	376	-	60	399	-	43	547	1.00	44
7812	CSDE1	56	54	419	-	88	598	1.00	93	469	-	89
1452	CSNK1A1	6	6	24	-	8	118	1.00	25	59	-	17
1466	CSRP2	4	3	44	-	11	0	-	0	3	-	2
1508	CTSB	0	0	0	-	0	0	-	0	5	-	2
2017	CTTN	33	31	340	-	56	697	1.00	64	497	1.00	59
1540	CYLD	6	3	0	-	0	334	1.00	72	247	1.00	56
28988	DBNL	13	12	149	0.86	29	336	1.00	35	209	1.00	36
10238	DCAF7	4	4	14	-	4	63	0.91	16	67	0.99	18
9201	DCLK1	2	0	181	1.00	41	13	-	6	23	-	10
166614	DCLK2	0	0	58	1.00	16	0	-	0	0	-	0
10521	DDX17	30	29	240	-	63	243	-	62	227	-	56
9188	DDX21	33	33	192	-	48	178	-	52	235	-	56
1654	DDX3X	36	34	508	1.00	67	850	1.00	82	635	1.00	82
1737	DLAT	0	0	19	0.75	7	0	-	0	0	-	0
10294	DNAJA2	10	8	50	-	14	120	0.94	24	154	1.00	26
10049	DNAJB6	0	0	12	-	7	20	1.00	11	27	1.00	12
54431	DNAIC10	4	3	0	-	0	103	1.00	36	57	0.98	19
11034	DSTN	0	0	0	-	0	0	-	0	20	0.75	9
1857	DVL3	26	26	192	-	39	275	-	34	303	0.75	34
1781	DYNCH12	0	0	0	-	0	4	-	2	15	1.00	8
11319	ECD	21	20	60	-	20	280	1.00	37	242	0.82	39
2023	ENO1	75	66	1077	1.00	63	792	-	57	962	1.00	52
1969	EPHA2	10	9	0	-	0	23	-	8	367	1.00	71
2059	EPSS8	0	0	0	-	0	22	1.00	11	6	-	4
64762	FAM59A	0	0	0	-	0	8	-	3	22	1.00	13
222584	FAM83B	0	0	0	-	0	22	0.75	10	6	-	4
23219	FBXO28	0	0	94	1.00	25	52	1.00	20	40	1.00	16
2316	FLNA	705	669	3242	-	314	4755	-	317	4600	-	300
54874	FNBP1L	0	0	0	-	0	30	1.00	15	13	-	6
10818	FRS2	0	0	0	-	0	80	1.00	20	62	1.00	20
2549	GAB1	3	3	0	-	0	67	0.99	23	145	1.00	30
2597	GAPDH	35	29	698	1.00	50	308	-	41	366	0.75	39
2673	GFPT1	4	4	4	-	4	16	-	54	0.77	22	43
2762	GMD5	0	0	0	-	0	0	-	23	1.00	10	6
10399	GNB2L1	18	15	168	-	35	311	1.00	37	326	1.00	36
2876	GPX1	3	3	66	1.00	19	0	-	0	0	-	0
2885	GRB2	4	4	0	-	0	158	1.00	31	174	1.00	28
54801	HAUS6	15	13	87	-	34	219	1.00	56	217	1.00	60
3006	HIST1H1C	10	8	0	-	0	0	-	0	28	-	3
3148	HMG2	2	2	50	0.99	13	4	-	2	25	-	8
3157	HMGCS1	2	0	0	-	0	0	-	0	45	0.98	17
51155	HN1	12	11	48	-	11	47	-	14	55	-	13
3320	HSP90AA1	32	29	168	-	45	272	-	59	287	-	60
3326	HSP90AB1	33	33	309	-	60	434	1.00	74	418	1.00	71
3304	HSPA1B	148	145	1258	-	85	1386	-	82	1434	-	78
3305	HSPA1L	73	71	381	-	18	541	-	24	482	-	18
3309	HSPA5	25	22	522	1.00	74	438	1.00	68	580	1.00	70
3310	HSPA6	16	15	111	-	9	141	-	15	87	-	9
3312	HSPA8	123	99	993	-	87	1319	-	73	1021	-	70
3420	IDH3B	2	0	0	-	0	0	-	0	29	0.87	11
3476	IGBP1	6	6	48	-	19	87	0.91	21	46	-	15
8660	IRS2	5	4	3	-	2	221	1.00	45	255	1.00	53
8471	IRS4	32	29	234	-	56	60	-	21	1174	1.00	116
3726	JUNB	0	0	0	-	0	0	-	0	43	1.00	13
115207	KCTD12	0	0	0	-	0	17	0.75	10	27	0.75	7
283638	KIAA0284	0	0	8	-	2	50	1.00	23	20	-	10
57608	KIAA1462	0	0</									

Annexes | Annexe 8 : Tableau des interacteurs de l'étude des fusions ALK identifiés par la technique du BioID

5901	RAN	24	23	220	30	277	0,76	28	247	29
5965	RECQL	14	11	80	29	41		19	50	15
387	RHOA	4	3	11	6	30		9	51	0,93
79657	RPAP3	27	27	42	19	116		40	112	37
23322	RPGRIP1L	0	0	0	0	43	1,00	24	0	0
6241	RRM2	20	18	50	17	267	1,00	50	149	33
8634	RTCD1	0	0	12	4	17	0,99	9	4	2
6282	S100A11	0	0	0	0	17		6	45	1,00
10807	SDCCAG3	5	4	0	0	74	0,90	19	50	17
81929	SEH1L	3	2	0	0	0		0	0	0
26135	SERBP1	50	44	293	44	263		45	342	36
6418	SET	26	24	338	18	369	1,00	26	391	1,00
6421	SFPO	38	37	239	42	172		33	177	41
6464	SHC1	6	5	31	13	499	1,00	43	750	1,00
57568	SIPA1L2	2	0	0	0	100	1,00	36	32	0,82
23094	SIPA1L3	2	2	0	0	47	1,00	23	3	2
22933	SIRT2	0	0	0	0	0		0	45	1,00
220134	SKA1	5	5	101	0,99	21	48	14	79	0,98
348235	SKA2	6	6	94	0,99	15	74	15	99	1,00
221150	SKA3	16	14	316	1,00	54	275	1,00	41	323
122060	SLAIN1	0	0	4	2	40	0,75	11	33	1,00
57606	SLAIN2	0	0	5	2	135	1,00	32	102	1,00
6510	SLC1A5	9	7	5	2	40		10	102	0,83
8604	SLC25A12	0	0	7	2	19		6	48	1,00
10165	SLC25A13	10	9	51	24	113	0,94	40	101	39
292	SLC25A5	16	15	235	1,00	33	291	1,00	27	218
293	SLC25A6	4	4	89	1,00	17	86	1,00	13	80
10580	SORBS1	6	6	0	0	95	0,97	35	86	0,75
9825	SPATA2	0	0	0	0	8		4	27	1,00
152185	SPICE1	0	0	0	0	17		8	30	1,00
6726	SRP9	0	0	23	0,75	8	5	2	16	1,00
6427	SRSF2	3	3	46	0,84	9	29	8	57	1,00
6742	SSBP1	0	0	0	0	12		4	39	1,00
6772	STAT1	24	24	81	29	370	1,00	84	372	1,00
6773	STAT2	3	2	0	0	43	0,92	19	0	0
6777	STAT5B	5	2	0	0	21		8	82	1,00
3925	STMN1	18	16	75	13	72		12	97	13
11171	STRAP	31	28	319	53	376	1,00	51	278	48
3703	STT3A	4	4	26	11	21		10	64	0,84
10910	SUGT1	22	20	90	27	104		32	123	37
6817	SULT1A1	4	3	3	2	41		16	53	0,91
8407	TAGLN2	44	44	610	1,00	41	760	1,00	43	647
84260	TCHP	13	12	9	5	194	1,00	50	125	36
7019	TFAM	0	0	14	6	0		0	0	0
92609	TIMM50	11	9	69	19	83		16	115	0,8
7082	TJP1	32	23	3	2	974	1,00	183	513	1,00
9414	TJP2	10	10	0	0	401	1,00	108	150	0,9
27327	TNRC6A	2	0	0	0	26	0,79	12	4	2
23112	TNRC6B	9	7	0	0	134	0,84	52	57	31
56993	TOMM22	8	8	99	0,87	17	75	16	108	0,96
7159	TP53BP2	14	14	16	8	178	0,82	50	129	45
55521	TRIM36	0	0	39	1,00	12	8	5	5	3
9319	TRIP13	8	5	30	12	60		26	83	0,75
84790	TUBA1C	21	20	0	0	0		0	0	0
7284	TUFM	26	23	232	63	187		49	303	0,99
9352	TXNL1	33	32	106	32	73		25	8	3
9040	UBE2M	3	2	51	0,99	14	27	8	26	8
55898	UNC45A	17	13	52	23	344	1,00	76	544	1,00
7411	VBP1	6	6	47	12	71		13	115	1,00
7431	VIM	69	67	466	88	809	0,96	89	700	86
11180	WDR6	4	2	0	0	0		0	95	1,00
84292	WDR83	5	4	0	0	60	0,82	17	21	9
51741	WWOX	31	30	14	6	36		13	243	37
7520	XRCC5	23	21	567	1,00	75	210	44	155	38
2547	XRCC6	33	30	702	1,00	91	253	59	256	60
54464	XRN1	37	36	134	47	503	1,00	100	393	80

Bibliographie

- Abd Halim, K.B., Koldsø, H., and Sansom, M.S.P. (2015). Interactions of the EGFR juxtamembrane domain with PIP 2 -containing lipid bilayers: Insights from multiscale molecular dynamics simulations. *Biochim. Biophys. Acta BBA - Gen. Subj.* *1850*, 1017–1025.
- Abel, E.V., Basile, K.J., Kugel, C.H., Witkiewicz, A.K., Le, K., Amaravadi, R.K., Karakousis, G.C., Xu, X., Xu, W., Schuchter, L.M., et al. (2013). Melanoma adapts to RAF/MEK inhibitors through FOXD3-mediated upregulation of ERBB3. *J. Clin. Invest.* *123*, 2155–2168.
- Acquaviva, J., Smith, D.L., Sang, J., Friedland, J.C., He, S., Sequeira, M., Zhang, C., Wada, Y., and Proia, D.A. (2012). Targeting KRAS-mutant non-small cell lung cancer with the Hsp90 inhibitor ganetespib. *Mol. Cancer Ther.* *11*, 2633–2643.
- Ahearn, I.M., Haigis, K., Bar-Sagi, D., and Philips, M.R. (2011). Regulating the regulator: post-translational modification of RAS. *Nat. Rev. Mol. Cell Biol.* *13*, 39–51.
- Akhavantabasi, S., Akman, H.B., Sapmaz, A., Keller, J., Petty, E.M., and Erson, A.E. (2010). USP32 is an active, membrane-bound ubiquitin protease overexpressed in breast cancers. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* *21*, 388–397.
- Altan, M., and Chiang, A.C. (2015). Management of Small Cell Lung Cancer: Progress and Updates. *Cancer J. Sudbury Mass* *21*, 425–433.
- ANSM (2014). Rapport d'activité 2014. ANSM.
- Aoki, Y., Niihori, T., Inoue, S., and Matsubara, Y. (2016). Recent advances in RASopathies. *J. Hum. Genet.* *61*, 33–39.
- Apolloni, A., Prior, I.A., Lindsay, M., Parton, R.G., and Hancock, J.F. (2000). H-ras but not K-ras traffics to the plasma membrane through the exocytic pathway. *Mol. Cell. Biol.* *20*, 2475–2487.
- Arcaro, A. (2015). Targeted therapies for small cell lung cancer: Where do we stand? *Crit. Rev. Oncol. Hematol.* *95*, 154–164.
- Arthur, J.S.C., and Ley, S.C. (2013). Mitogen-activated protein kinases in innate immunity. *Nat. Rev. Immunol.* *13*, 679–692.
- Asahina, H., Yamazaki, K., Kinoshita, I., Sukoh, N., Harada, M., Yokouchi, H., Ishida, T., Ogura, S., Kojima, T., Okamoto, Y., et al. (2006). A phase II trial of gefitinib as first-line therapy for advanced non-small cell lung cancer with epidermal growth factor receptor mutations. *Br. J. Cancer* *95*, 998–1004.
- Athuluri-Divakar, S.K., Vasquez-Del Carpio, R., Dutta, K., Baker, S.J., Cosenza, S.C., Basu, I., Gupta, Y.K., Reddy, M.V.R., Ueno, L., Hart, J.R., et al. (2016). A Small Molecule RAS-Mimetic Disrupts RAS Association with Effector Proteins to Block Signaling. *Cell* *165*, 643–655.
- Barrott, J.J., and Haystead, T.A.J. (2013). Hsp90, an unlikely ally in the war on cancer. *FEBS J.* *280*, 1381–1396.
- Bayliss, R., Choi, J., Fennell, D.A., Fry, A.M., and Richards, M.W. (2016). Molecular mechanisms that underpin EML4-ALK driven cancers and their response to targeted drugs. *Cell. Mol. Life Sci.* *73*, 1209–1224.
- Bean, J., Brennan, C., Shih, J.-Y., Riely, G., Viale, A., Wang, L., Chitale, D., Motoi, N., Szoke, J., Broderick, S., et al. (2007). MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proc. Natl. Acad. Sci.* *104*, 20932–20937.

- Berndt, N., Hamilton, A.D., and Sebti, S.M. (2011). Targeting protein prenylation for cancer therapy. *Nat. Rev. Cancer* *11*, 775–791.
- Bivona, T.G., Quatela, S.E., Bodemann, B.O., Ahearn, I.M., Soskis, M.J., Mor, A., Miura, J., Wiener, H.H., Wright, L., Saba, S.G., et al. (2006). PKC Regulates a Farnesyl-Electrostatic Switch on K-Ras that Promotes its Association with Bcl-Xl on Mitochondria and Induces Apoptosis. *Mol. Cell* *21*, 481–493.
- Blasco, R.B., Francoz, S., Santamaría, D., Cañamero, M., Dubus, P., Charron, J., Baccarini, M., and Barbacid, M. (2011). c-Raf, but not B-Raf, is essential for development of K-Ras oncogene-driven non-small cell lung carcinoma. *Cancer Cell* *19*, 652–663.
- Bobrich, M.A., Schwabe, S.A., Brobeil, A., Viard, M., Kamm, M., Mooren, F.C., Krüger, K., Tag, C., and Wimmer, M. (2013). PTP51: a new interaction partner of the insulin receptor and PKA in adipose tissue. *J. Obes.* *2013*, 476240.
- Bobrie, A., Krumeich, S., Reyat, F., Recchi, C., Moita, L.F., Seabra, M.C., Ostrowski, M., and Théry, C. (2012). Rab27a supports exosome-dependent and -independent mechanisms that modify the tumor microenvironment and can promote tumor progression. *Cancer Res.* *72*, 4920–4930.
- Boldrini, L., Gisfredi, S., Ursino, S., Faviana, P., Lucchi, M., Melfi, F., Mussi, A., Basolo, F., and Fontanini, G. (2005). Expression of endothelin-1 is related to poor prognosis in non-small cell lung carcinoma. *Eur. J. Cancer* *41*, 2828–2835.
- Bonaldi, C., Andriantafika, F., Chyderiotis, S., Boussac-Zarebska, M., Cao, B., Benmarhnia, T., and Gremy, I. (2016). Les décès attribuables au tabagisme en France. Dernières estimations et tendance, années 2000 à 2013. *BEH* *30–31*, 528–540.
- Borghaei, H., Paz-Ares, L., Horn, L., Spigel, D.R., Steins, M., Ready, N.E., Chow, L.Q., Vokes, E.E., Felip, E., Holgado, E., et al. (2015). Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. *N. Engl. J. Med.* *373*, 1627–1639.
- Boros, J., Donaldson, I.J., O'Donnell, A., Odrowaz, Z.A., Zeef, L., Lupien, M., Meyer, C.A., Liu, X.S., Brown, M., and Sharrocks, A.D. (2009). Elucidation of the ELK1 target gene network reveals a role in the coordinate regulation of core components of the gene regulation machinery. *Genome Res.* *19*, 1963–1973.
- Bos, J.L., Rehmann, H., and Wittinghofer, A. (2007). GEFs and GAPs: Critical Elements in the Control of Small G Proteins. *Cell* *129*, 865–877.
- Brand, T.M., Iida, M., and Wheeler, D.L. (2011). Molecular mechanisms of resistance to the EGFR monoclonal antibody cetuximab. *Cancer Biol. Ther.* *11*, 777–792.
- Bryant, K.L., Mancias, J.D., Kimmelman, A.C., and Der, C.J. (2014). KRAS: feeding pancreatic cancer proliferation. *Trends Biochem. Sci.* *39*, 91–100.
- Camidge, D.R., Pao, W., and Sequist, L.V. (2014). Acquired resistance to TKIs in solid tumours: learning from lung cancer. *Nat. Rev. Clin. Oncol.* *11*, 473–481.
- Cargnello, M., and Roux, P.P. (2011). Activation and Function of the MAPKs and Their Substrates, the MAPK-Activated Protein Kinases. *Microbiol. Mol. Biol. Rev.* *75*, 50–83.
- Castellano, E., and Downward, J. (2011). RAS Interaction with PI3K: More Than Just Another Effector Pathway. *Genes Cancer* *2*, 261–274.

- Castellano, E., and Santos, E. (2011). Functional Specificity of Ras Isoforms: So Similar but So Different. *Genes Cancer* 2, 216–231.
- Changeux, J.-P., and Christopoulos, A. (2016). Allosteric Modulation as a Unifying Mechanism for Receptor Function and Regulation. *Cell* 166, 1084–1102.
- Chen, Y.A., and Scheller, R.H. (2001). SNARE-mediated membrane fusion. *Nat. Rev. Mol. Cell Biol.* 2, 98–106.
- Chen, C.-L., Hu, Y., Udeshi, N.D., Lau, T.Y., Wirtz-Peitz, F., He, L., Ting, A.Y., Carr, S.A., and Perrimon, N. (2015). Proteomic mapping in live *Drosophila* tissues using an engineered ascorbate peroxidase. *Proc. Natl. Acad. Sci. U. S. A.* 112, 12093–12098.
- Chiapparino, A., Maeda, K., Turei, D., Saez-Rodriguez, J., and Gavin, A.-C. (2016). The orchestra of lipid-transfer proteins at the crossroads between metabolism and signaling. *Prog. Lipid Res.* 61, 30–39.
- Chmielecki, J., Foo, J., Oxnard, G.R., Hutchinson, K., Ohashi, K., Somwar, R., Wang, L., Amato, K.R., Arcila, M., Sos, M.L., et al. (2011). Optimization of Dosing for EGFR-Mutant Non-Small Cell Lung Cancer with Evolutionary Cancer Modeling. *Sci. Transl. Med.* 3, 90ra59-90ra59.
- Choi, H., Larsen, B., Lin, Z.-Y., Breitkreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z.S., Tyers, M., Gingras, A.-C., and Nesvizhskii, A.I. (2011). SAINT: probabilistic scoring of affinity purification–mass spectrometry data. *Nat. Methods* 8, 70–73.
- Chung, H.H., Benson, D.R., Cornish, V.W., and Schultz, P.G. (1993). Probing the role of loop 2 in Ras function with unnatural amino acids. *Proc. Natl. Acad. Sci. U. S. A.* 90, 10145–10149.
- Clapham, D.E. (2007). Calcium signaling. *Cell* 131, 1047–1058.
- Commisso, C., Davidson, S.M., Soydaner-Azeloglu, R.G., Parker, S.J., Kamphorst, J.J., Hackett, S., Grabocka, E., Nofal, M., Drebin, J.A., Thompson, C.B., et al. (2013). Macropinocytosis of protein is an amino acid supply route in Ras-transformed cells. *Nature* 497, 633–637.
- Coon, B.G., Burgner, J., Camonis, J.H., and Aguilar, R.C. (2010). The epsin family of endocytic adaptors promotes fibrosarcoma migration and invasion. *J. Biol. Chem.* 285, 33073–33081.
- Cox, A.D., and Der, C.J. (2010). Ras history: The saga continues. *Small GTPases* 1, 2–27.
- Cox, A.D., Fesik, S.W., Kimmelman, A.C., Luo, J., and Der, C.J. (2014). Drugging the undruggable RAS: Mission Possible? *Nat. Rev. Drug Discov.* 13, 828–851.
- Craig, R., and Beavis, R.C. (2004). TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* 20, 1466–1467.
- Cutler, J.A., Tahir, R., Sreenivasamurthy, S.K., Mitchell, C., Renuse, S., Nirujogi, R.S., Patil, A.H., Heydarian, M., Wong, X., Wu, X., et al. (2017). Differential signaling through p190 and p210 BCR-ABL fusion proteins revealed by interactome and phosphoproteome analysis. *Leukemia* 31, 1513–1524.
- Dankort, D., Filenova, E., Collado, M., Serrano, M., Jones, K., and McMahon, M. (2007). A new mouse model to explore the initiation, progression, and therapy of BRAFV600E-induced lung tumors. *Genes Dev.* 21, 379–384.
- Deutsch, E.W., Mendoza, L., Shteynberg, D., Farrah, T., Lam, H., Tasman, N., Sun, Z., Nilsson, E., Pratt, B., Prazan, B., et al. (2010). A guided tour of the Trans-Proteomic Pipeline. *Proteomics* 10, 1150–1159.

- Dingar, D., Kalkat, M., Chan, P.-K., Srikumar, T., Bailey, S.D., Tu, W.B., Coyaud, E., Ponzielli, R., Kolyar, M., Jurisica, I., et al. (2015). BioID identifies novel c-MYC interacting partners in cultured cells and xenograft tumors. *J. Proteomics* 118, 95–111.
- Doll, R., Peto, R., Boreham, J., and Sutherland, I. (2004). Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ* 328, 1519.
- Downward, J. (2003). Targeting RAS signalling pathways in cancer therapy. *Nat. Rev. Cancer* 3, 11–22.
- Drosten, M., Dhawahir, A., Sum, E.Y.M., Urosevic, J., Lechuga, C.G., Esteban, L.M., Castellano, E., Guerra, C., Santos, E., and Barbacid, M. (2010). Genetic analysis of Ras signalling pathways in cell proliferation, migration and survival. *EMBO J.* 29, 1091–1104.
- Dumaz, N., Hayward, R., Martin, J., Ogilvie, L., Hedley, D., Curtin, J.A., Bastian, B.C., Springer, C., and Marais, R. (2006). In melanoma, RAS mutations are accompanied by switching signaling from BRAF to CRAF and disrupted cyclic AMP signaling. *Cancer Res.* 66, 9483–9491.
- Dunant, A., Pignon, J.-P., and Le Chevalier, T. (2005). Adjuvant chemotherapy for non-small cell lung cancer: contribution of the International Adjuvant Lung Trial. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 11, 5017s–5021s.
- Fehrenbacher, L., Spira, A., Ballinger, M., Kowanetz, M., Vansteenkiste, J., Mazieres, J., Park, K., Smith, D., Artal-Cortes, A., Lewanski, C., et al. (2016). Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet Lond. Engl.* 387, 1837–1846.
- Feng, Z., Hu, W., Chen, J.X., Pao, A., Li, H., Rom, W., Hung, M.-C., and Tang, M. (2002). Preferential DNA damage and poor repair determine ras gene mutational hotspot in human cancer. *J. Natl. Cancer Inst.* 94, 1527–1536.
- Ferguson, K.M. (2008). Structure-based view of epidermal growth factor receptor regulation. *Annu. Rev. Biophys.* 37, 353–373.
- Fernandez-Medarde, A., and Santos, E. (2011). Ras in Cancer and Developmental Diseases. *Genes Cancer* 2, 344–358.
- Forbes, S.A., Beare, D., Boutselakis, H., Bamford, S., Bindal, N., Tate, J., Cole, C.G., Ward, S., Dawson, E., Ponting, L., et al. (2017). COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res.* 45, D777–D783.
- Fumarola, C., Bonelli, M.A., Petronini, P.G., and Alfieri, R.R. (2014). Targeting PI3K/AKT/mTOR pathway in non small cell lung cancer. *Biochem. Pharmacol.* 90, 197–207.
- Furth, M.E., Aldrich, T.H., and Cordon-Cardo, C. (1987). Expression of ras proto-oncogene proteins in normal human tissues. *Oncogene* 1, 47–58.
- Gajiwala, K.S. (2013). EGFR: tale of the C-terminal tail. *Protein Sci. Publ. Protein Soc.* 22, 995–999.
- Garon, E.B., Rizvi, N.A., Hui, R., Leighl, N., Balmanoukian, A.S., Eder, J.P., Patnaik, A., Aggarwal, C., Gubens, M., Horn, L., et al. (2015). Pembrolizumab for the Treatment of Non-Small-Cell Lung Cancer. *N. Engl. J. Med.* 372, 2018–2028.
- Gazdar, A.F. (2009). Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors. *Oncogene* 28, S24–S31.

- Ghosh, G., Lian, X., Kron, S.J., and Palecek, S.P. (2012). Properties of resistant cells generated from lung cancer cell lines treated with EGFR inhibitors. *BMC Cancer* 12, 95.
- Ghosh, P., Dahms, N.M., and Kornfeld, S. (2003). Mannose 6-phosphate receptors: new twists in the tale. *Nat. Rev. Mol. Cell Biol.* 4, 202–212.
- Girotti, M.R., Lopes, F., Preece, N., Niculescu-Duvaz, D., Zambon, A., Davies, L., Whittaker, S., Saturno, G., Viros, A., Pedersen, M., et al. (2015). Paradox-breaking RAF inhibitors that also target SRC are effective in drug-resistant BRAF mutant melanoma. *Cancer Cell* 27, 85–96.
- Goebell, P.J., and Knowles, M.A. (2010). Bladder cancer or bladder cancers? Genetically distinct malignant conditions of the urothelium. *Urol. Oncol.* 28, 409–428.
- Goldstraw, P., Ball, D., Jett, J.R., Le Chevalier, T., Lim, E., Nicholson, A.G., and Shepherd, F.A. (2011). Non-small-cell lung cancer. *Lancet Lond. Engl.* 378, 1727–1740.
- Grunwald, A., Gottfried, I., Cox, A.D., Haklai, R., Kloog, Y., and Ashery, U. (2013). Rasosomes originate from the Golgi to dispense Ras signals. *Cell Death Dis.* 4, e496.
- Guin, S., Ru, Y., Wynes, M.W., Mishra, R., Lu, X., Owens, C., Barn, A.E., Vasu, V.T., Hirsch, F.R., Kern, J.A., et al. (2013). Contributions of KRAS and RAL in non-small-cell lung cancer growth and progression. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 8, 1492–1501.
- Gupta, G.D., Coyaud, É., Gonçalves, J., Mojarad, B.A., Liu, Y., Wu, Q., Gheiratmand, L., Comartin, D., Tkach, J.M., Cheung, S.W.T., et al. (2015). A Dynamic Protein Interaction Landscape of the Human Centrosome-Cilium Interface. *Cell* 163, 1484–1499.
- Győrffy, B., Stelnic-Klotz, I., Sigler, C., Kasack, K., Redmer, T., Qian, Y., and Schäfer, R. (2015). Effects of RAL signal transduction in KRAS- and BRAF-mutated cells and prognostic potential of the RAL signature in colorectal cancer. *Oncotarget* 6, 13334–13346.
- Hallberg, B., and Palmer, R.H. (2016). The role of the ALK receptor in cancer biology. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* 27 Suppl 3, iii4-iii15.
- Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674.
- Hancock, J.F. (2003). Ras proteins: different signals from different locations. *Nat. Rev. Mol. Cell Biol.* 4, 373–384.
- Hart, T., Chandrashekar, M., Aregger, M., Steinhart, Z., Brown, K.R., MacLeod, G., Mis, M., Zimmermann, M., Fradet-Turcotte, A., Sun, S., et al. (2015). High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* 163, 1515–1526.
- He, L., and Hristova, K. (2012). Consequences of replacing EGFR juxtamembrane domain with an unstructured sequence. *Sci. Rep.* 2.
- Hermeking, H. (2003). The 14-3-3 cancer connection. *Nat. Rev. Cancer* 3, 931–943.
- Hiley, C.T., Le Quesne, J., Santis, G., Sharpe, R., de Castro, D.G., Middleton, G., and Swanton, C. (2016). Challenges in molecular testing in non-small-cell lung cancer patients with advanced disease. *Lancet Lond. Engl.* 388, 1002–1011.
- Hirsch, F.R., Suda, K., Wiens, J., and Bunn, P.A. (2016). New and emerging targeted treatments in advanced non-small-cell lung cancer. *Lancet Lond. Engl.* 388, 1012–1024.

- Holderfield, M., Deuker, M.M., McCormick, F., and McMahon, M. (2014). Targeting RAF kinases for cancer therapy: BRAF-mutated melanoma and beyond. *Nat. Rev. Cancer* *14*, 455–467.
- Hrustanovic, G., Olivas, V., Pazarentzos, E., Tulpule, A., Asthana, S., Blakely, C.M., Okimoto, R.A., Lin, L., Neel, D.S., Sabnis, A., et al. (2015). RAS-MAPK dependence underlies a rational polytherapy strategy in EML4-ALK-positive lung cancer. *Nat. Med.* *21*, 1038–1047.
- Hu, J., Stites, E.C., Yu, H., Germino, E.A., Meharena, H.S., Stork, P.J.S., Kornev, A.P., Taylor, S.S., and Shaw, A.S. (2013). Allosteric activation of functionally asymmetric RAF kinase dimers. *Cell* *154*, 1036–1046.
- Hu, W., Wei, H., Li, K., Li, P., Lin, J., and Feng, R. (2017). Downregulation of USP32 inhibits cell proliferation, migration and invasion in human small cell lung cancer. *Cell Prolif.* *50*.
- Hu, Y., Lu, W., Chen, G., Wang, P., Chen, Z., Zhou, Y., Ogasawara, M., Trachootham, D., Feng, L., Pelicano, H., et al. (2012). K-ras(G12V) transformation leads to mitochondrial dysfunction and a metabolic switch from oxidative phosphorylation to glycolysis. *Cell Res.* *22*, 399–412.
- Huang, L., and Tepasamorndech, S. (2013). The SLC30 family of zinc transporters – A review of current understanding of their biological and pathophysiological roles. *Mol. Aspects Med.* *34*, 548–560.
- Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* *4*, 44–57.
- Ihle, N.T., Byers, L.A., Kim, E.S., Saintigny, P., Lee, J.J., Blumenschein, G.R., Tsao, A., Liu, S., Larsen, J.E., Wang, J., et al. (2012). Effect of KRAS oncogene substitutions on protein behavior: implications for signaling and clinical outcome. *J. Natl. Cancer Inst.* *104*, 228–239.
- Imai, K., and Takaoka, A. (2006). Comparing antibody and small-molecule therapies for cancer. *Nat. Rev. Cancer* *6*, 714–727.
- INCa (2014). Plan cancer 2014-2019. INCa.
- INCa (2017). Les cancers en France 2016. INCa.
- INSERM (2008). Cancer et environnement. INSERM.
- Iwahara, T., Fujimoto, J., Wen, D., Cupples, R., Bucay, N., Arakawa, T., Mori, S., Ratzkin, B., and Yamamoto, T. (1997). Molecular characterization of ALK, a receptor tyrosine kinase expressed specifically in the nervous system. *Oncogene* *14*, 439–449.
- Jackman, D.M., and Johnson, B.E. (2005). Small-cell lung cancer. *Lancet Lond. Engl.* *366*, 1385–1396.
- Jang, J.S., Lee, A., Li, J., Liyanage, H., Yang, Y., Guo, L., Asmann, Y.W., Li, P.W., Erickson-Johnson, M., Sakai, Y., et al. (2015). Common Oncogene Mutations and Novel SND1-BRAF Transcript Fusion in Lung Adenocarcinoma from Never Smokers. *Sci. Rep.* *5*.
- Jia, Y., Yun, C.-H., Park, E., Ercan, D., Manuia, M., Juarez, J., Xu, C., Rhee, K., Chen, T., Zhang, H., et al. (2016). Overcoming EGFR(T790M) and EGFR(C797S) resistance with mutant-selective allosteric inhibitors. *Nature* *534*, 129–132.
- Johnson, L., Greenbaum, D., Cichowski, K., Mercer, K., Murphy, E., Schmitt, E., Bronson, R.T., Umanoff, H., Edelman, W., Kucherlapati, R., et al. (1997). K-ras is an essential gene in the mouse with partial functional overlap with N-ras. *Genes Dev.* *11*, 2468–2481.

- Kachroo, N., Valencia, T., Warren, A.Y., and Gnanapragasam, V.J. (2013). Evidence for downregulation of the negative regulator SPRED2 in clinical prostate cancer. *Br. J. Cancer* *108*, 597–601.
- Kapoor, A., Yao, W., Ying, H., Hua, S., Liewen, A., Wang, Q., Zhong, Y., Wu, C.-J., Sadanandam, A., Hu, B., et al. (2014). Yap1 activation enables bypass of oncogenic Kras addiction in pancreatic cancer. *Cell* *158*, 185–197.
- Karnoub, A.E., and Weinberg, R.A. (2008). Ras oncogenes: split personalities. *Nat. Rev. Mol. Cell Biol.* *9*, 517–531.
- Karreth, F.A., Frese, K.K., DeNicola, G.M., Baccharini, M., and Tuveson, D.A. (2011). C-Raf is required for the initiation of lung cancer by K-Ras(G12D). *Cancer Discov.* *1*, 128–136.
- Kaszuba, K., Grzybek, M., Orłowski, A., Danne, R., Róg, T., Simons, K., Coskun, Ü., and Vattulainen, I. (2015). N-Glycosylation as determinant of epidermal growth factor receptor conformation in membranes. *Proc. Natl. Acad. Sci. U. S. A.* *112*, 4334–4339.
- Kazazic, M., Bertelsen, V., Pedersen, K.W., Vuong, T.T., Grandal, M.V., Rødland, M.S., Traub, L.M., Stang, E., and Madhus, I.H. (2009). Epsin 1 is involved in recruitment of ubiquitinated EGF receptors into clathrin-coated pits. *Traffic Cph. Den.* *10*, 235–245.
- Kerr, E.M., Gaude, E., Turrell, F.K., Frezza, C., and Martins, C.P. (2016). Mutant Kras copy number defines metabolic reprogramming and therapeutic susceptibilities. *Nature* *531*, 110–113.
- Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. (2008). ProteoWizard: open source software for rapid proteomics tools development. *Bioinforma. Oxf. Engl.* *24*, 2534–2536.
- Kim, D.I., and Roux, K.J. (2016). Filling the Void: Proximity-Based Labeling of Proteins in Living Cells. *Trends Cell Biol.* *26*, 804–817.
- Kim, H.J., and Bar-Sagi, D. (2004). Modulation of signalling by Sprouty: a developing story. *Nat. Rev. Mol. Cell Biol.* *5*, 441–450.
- Kim, D.I., Kc, B., Zhu, W., Motamedchaboki, K., Doye, V., and Roux, K.J. (2014). Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proc. Natl. Acad. Sci.* *111*, E2453–E2461.
- Knight, J.D.R., Choi, H., Gupta, G.D., Pelletier, L., Raught, B., Nesvizhskii, A.I., and Gingras, A.-C. (2017). ProHits-viz: a suite of web tools for visualizing interaction proteomics data. *Nat. Methods* *14*, 645–646.
- Koera, K., Nakamura, K., Nakao, K., Miyoshi, J., Toyoshima, K., Hatta, T., Otani, H., Aiba, A., and Katsuki, M. (1997). K-ras is essential for the development of the mouse embryo. *Oncogene* *15*, 1151–1159.
- Kofer-Geles, M., Gottfried, I., Haklai, R., Elad-Zefadia, G., Kloog, Y., and Ashery, U. (2009). Rasosomes spread Ras signals from plasma membrane “hotspots.” *Biochim. Biophys. Acta* *1793*, 1691–1702.
- Kumar, A., Petri, E.T., Halmos, B., and Boggon, T.J. (2008). Structure and clinical relevance of the epidermal growth factor receptor in human cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* *26*, 1742–1751.

- Kupzig, S., Bouyoucef-Cherchalli, D., Yarwood, S., Sessions, R., and Cullen, P.J. (2009). The ability of GAP1IP4BP to function as a Rap1 GTPase-activating protein (GAP) requires its Ras GAP-related domain and an arginine finger rather than an asparagine thumb. *Mol. Cell. Biol.* *29*, 3929–3940.
- Lake, D., Corrêa, S.A.L., and Müller, J. (2016). Negative feedback regulation of the ERK1/2 MAPK pathway. *Cell. Mol. Life Sci.* *73*, 4397–4413.
- Larance, M., and Lamond, A.I. (2015). Multidimensional proteomics for cell biology. *Nat. Rev. Mol. Cell Biol.* *16*, 269–280.
- Lasithiotakis, K.G., Sinnberg, T.W., Schitteck, B., Flaherty, K.T., Kulms, D., Maczey, E., Garbe, C., and Meier, F.E. (2008). Combined inhibition of MAPK and mTOR signaling inhibits growth, induces cell death, and abrogates invasive growth of melanoma cells. *J. Invest. Dermatol.* *128*, 2013–2023.
- Lavoie, H., and Therrien, M. (2015). Regulation of RAF protein kinases in ERK signalling. *Nat. Rev. Mol. Cell Biol.* *16*, 281–298.
- Legrier, M.-E., Yang, C.-P.H., Yan, H.-G., Lopez-Barcons, L., Keller, S.M., Pérez-Soler, R., Horwitz, S.B., and McDaid, H.M. (2007). Targeting protein translation in human non small cell lung cancer via combined MEK and mammalian target of rapamycin suppression. *Cancer Res.* *67*, 11300–11308.
- Li, W., Hu, Y., Jiang, T., Han, Y., Han, G., Chen, J., and Li, X. (2014). Rab27A regulates exosome secretion from lung adenocarcinoma cells A549: involvement of EPI64. *APMIS Acta Pathol. Microbiol. Immunol. Scand.* *122*, 1080–1087.
- Light, Y., Paterson, H., and Marais, R. (2002). 14-3-3 antagonizes Ras-mediated Raf-1 recruitment to the plasma membrane to maintain signaling fidelity. *Mol. Cell. Biol.* *22*, 4984–4996.
- Lim, K.-H., Baines, A.T., Fiordalisi, J.J., Shipitsin, M., Feig, L.A., Cox, A.D., Der, C.J., and Counter, C.M. (2005). Activation of RalA is critical for Ras-induced tumorigenesis of human cells. *Cancer Cell* *7*, 533–545.
- Linggi, B., and Carpenter, G. (2006). ErbB receptors: new insights on mechanisms and biology. *Trends Cell Biol.* *16*, 649–656.
- Lito, P., Solomon, M., Li, L.-S., Hansen, R., and Rosen, N. (2016). Allele-specific inhibitors inactivate mutant KRAS G12C by a trapping mechanism. *Science* *351*, 604–608.
- Liu, G., Zhang, J., Larsen, B., Stark, C., Breikreutz, A., Lin, Z.-Y., Breikreutz, B.-J., Ding, Y., Colwill, K., Pasculescu, A., et al. (2010). ProHits: integrated software for mass spectrometry-based interaction proteomics. *Nat. Biotechnol.* *28*, 1015–1017.
- Liyasova, M.S., Ma, K., and Lipkowitz, S. (2015). Molecular pathways: cbl proteins in tumorigenesis and antitumor immunity-opportunities for cancer treatment. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *21*, 1789–1794.
- Lönn, P., and Landegren, U. (2017). Close Encounters - Probing Proximal Proteins in Live or Fixed Cells. *Trends Biochem. Sci.* *42*, 504–515.
- Lu, S., Jang, H., Muratcioglu, S., Gursoy, A., Keskin, O., Nussinov, R., and Zhang, J. (2016). Ras Conformational Ensembles, Allostery, and Signaling. *Chem. Rev.* *116*, 6607–6665.
- Lynch, T.J., Bell, D.W., Sordella, R., Gurubhagavatula, S., Okimoto, R.A., Brannigan, B.W., Harris, P.L., Haserlat, S.M., Supko, J.G., Haluska, F.G., et al. (2004). Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non-Small-Cell Lung Cancer to Gefitinib. *N. Engl. J. Med.* *350*, 2129–2139.

- Magliozzi, R., Low, T.Y., Weijts, B.G.M.W., Cheng, T., Spanjaard, E., Mohammed, S., van Veen, A., Ovaa, H., de Rooij, J., Zwartkruis, F.J.T., et al. (2013). Control of Epithelial Cell Migration and Invasion by the IKK β - and CK1 α -Mediated Degradation of RAPGEF2. *Dev. Cell* 27, 574–585.
- Manickam, V., Tiwari, A., Jung, J.-J., Bhattacharya, R., Goel, A., Mukhopadhyay, D., and Choudhury, A. (2011). Regulation of vascular endothelial growth factor receptor 2 trafficking and angiogenesis by Golgi localized t-SNARE syntaxin 6. *Blood* 117, 1425–1435.
- Masin, M., Vazquez, J., Rossi, S., Groeneveld, S., Samson, N., Schwalie, P.C., Deplancke, B., Frawley, L.E., Gouttenoire, J., Moradpour, D., et al. (2014). GLUT3 is induced during epithelial-mesenchymal transition and promotes tumor cell proliferation in non-small cell lung cancer. *Cancer Metab.* 2, 11.
- Matikas, A., Kentepozidis, N., Georgoulas, V., and Kotsakis, A. (2016). Management of Resistance to Crizotinib in Anaplastic Lymphoma Kinase-Positive Non-Small-cell Lung Cancer. *Clin. Lung Cancer* 17, 474–482.
- Mauro, L.V., Grossoni, V.C., Urtreger, A.J., Yang, C., Colombo, L.L., Morandi, A., Pallotta, M.G., Kazanietz, M.G., Bal de Kier Joffé, E.D., and Puricelli, L.L. (2010). PKC Delta (PKCdelta) promotes tumoral progression of human ductal pancreatic cancer. *Pancreas* 39, e31-41.
- Maus, M.K.H., Stephens, C., Zeger, G., Grimminger, P.P., and Huang, E. (2012). Identification of Novel Variant of EML4-ALK Fusion Gene in NSCLC: Potential Benefits of the RT-PCR Method. *Int. J. Biomed. Sci. IJBS* 8, 1–6.
- Mayer, B.J. (2015). The discovery of modular binding domains: building blocks of cell signalling. *Nat. Rev. Mol. Cell Biol.* 16, 691–698.
- Mehta, V., and Trinkle-Mulcahy, L. (2016). Recent advances in large-scale protein interactome mapping. *F1000Research* 5.
- Mehus, A.A., Anderson, R.H., and Roux, K.J. (2016). BioID Identification of Lamin-Associated Proteins. In *Methods in Enzymology*, (Elsevier), pp. 3–22.
- Merkel, O., Hamacher, F., Sifft, E., Kenner, L., Greil, R., and European Research Initiative on Anaplastic Large Cell Lymphoma (2011). Novel therapeutic options in anaplastic large cell lymphoma: molecular targets and immunological tools. *Mol. Cancer Ther.* 10, 1127–1136.
- Miller, V.A., Hirsh, V., Cadranel, J., Chen, Y.-M., Park, K., Kim, S.-W., Zhou, C., Su, W.-C., Wang, M., Sun, Y., et al. (2012). Afatinib versus placebo for patients with advanced, metastatic non-small-cell lung cancer after failure of erlotinib, gefitinib, or both, and one or two lines of chemotherapy (LUX-Lung 1): a phase 2b/3 randomised trial. *Lancet Oncol.* 13, 528–538.
- Naoe, T., Suzuki, T., Kiyoi, H., and Urano, T. (2006). Nucleophosmin: a versatile molecule associated with hematological malignancies. *Cancer Sci.* 97, 963–969.
- Neuzil, J., Rohlena, J., and Dong, L.-F. (2012). K-Ras and mitochondria: Dangerous liaisons. *Cell Res.* 22, 285–287.
- Noble, C., Mercer, K., Hussain, J., Carragher, L., Giblett, S., Hayward, R., Patterson, C., Marais, R., and Pritchard, C.A. (2008). CRAF autophosphorylation of serine 621 is required to prevent its proteasome-mediated degradation. *Mol. Cell* 31, 862–872.
- Nonami, A., Kato, R., Taniguchi, K., Yoshiga, D., Taketomi, T., Fukuyama, S., Harada, M., Sasaki, A., and Yoshimura, A. (2004). Spred-1 Negatively Regulates Interleukin-3-mediated ERK/Mitogen-activated Protein (MAP) Kinase Activation in Hematopoietic Cells. *J. Biol. Chem.* 279, 52543–52551.

- Ouderkirk-Pecone, J.L., Goreczny, G.J., Chase, S.E., Tatum, A.H., Turner, C.E., and Krendel, M. (2016). Myosin 1e promotes breast cancer malignancy by enhancing tumor cell proliferation and stimulating tumor cell de-differentiation. *Oncotarget* 7, 46419–46432.
- Padanad, M.S., Konstantinidou, G., Venkateswaran, N., Melegari, M., Rindhe, S., Mitsche, M., Yang, C., Batten, K., Huffman, K.E., Liu, J., et al. (2016). Fatty Acid Oxidation Mediated by Acyl-CoA Synthetase Long Chain 3 Is Required for Mutant KRAS Lung Tumorigenesis. *Cell Rep.* 16, 1614–1628.
- Padavano, J., Henkhaus, R.S., Chen, H., Skovan, B.A., Cui, H., and Ignatenko, N.A. (2015). Mutant K-RAS Promotes Invasion and Metastasis in Pancreatic Cancer Through GTPase Signaling Pathways. *Cancer Growth Metastasis* 8, 95–113.
- Paez, J.G., Jänne, P.A., Lee, J.C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F.J., Lindeman, N., Boggon, T.J., et al. (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 304, 1497–1500.
- Paik, P.K., Arcila, M.E., Fara, M., Sima, C.S., Miller, V.A., Kris, M.G., Ladanyi, M., and Riely, G.J. (2011). Clinical Characteristics of Patients With Lung Adenocarcinomas Harboring *BRAF* Mutations. *J. Clin. Oncol.* 29, 2046–2051.
- Pal, A., and Donato, N.J. (2014). Ubiquitin-specific proteases as therapeutic targets for the treatment of breast cancer. *Breast Cancer Res.* 16.
- Palmieri, F. (2013). The mitochondrial transporter family SLC25: Identification, properties and physiopathology. *Mol. Aspects Med.* 34, 465–484.
- Pardoll, D.M. (2012). The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* 12, 252–264.
- Park, J., Kang, S.I., Lee, S.-Y., Zhang, X.F., Kim, M.S., Beers, L.F., Lim, D.-S., Avruch, J., Kim, H.-S., and Lee, S.B. (2010). Tumor suppressor ras association domain family 5 (RASSF5/NORE1) mediates death receptor ligand-induced apoptosis. *J. Biol. Chem.* 285, 35029–35038.
- Parsons, A., Daley, A., Begh, R., and Aveyard, P. (2010). Influence of smoking cessation after diagnosis of early stage lung cancer on prognosis: systematic review of observational studies with meta-analysis. *BMJ* 340, b5569.
- Patel, J.H., Loboda, A.P., Showe, M.K., Showe, L.C., and McMahon, S.B. (2004). Opinion: Analysis of genomic targets reveals complex functions of MYC. *Nat. Rev. Cancer* 4, 562–568.
- Patricelli, M.P., Janes, M.R., Li, L.-S., Hansen, R., Peters, U., Kessler, L.V., Chen, Y., Kucharski, J.M., Feng, J., Ely, T., et al. (2016). Selective Inhibition of Oncogenic KRAS Output with Small Molecules Targeting the Inactive State. *Cancer Discov.* 6, 316–329.
- Pavlova, N.N., and Thompson, C.B. (2016). The Emerging Hallmarks of Cancer Metabolism. *Cell Metab.* 23, 27–47.
- Pedrioli, P.G.A. (2010). Trans-proteomic pipeline: a pipeline for proteomic analysis. *Methods Mol. Biol. Clifton NJ* 604, 213–238.
- Pircher, A., Ploner, F., Popper, H., and Hilbe, W. (2010). Rationale of a relaunch of gefitinib in Caucasian non-small cell lung cancer patients. *Lung Cancer Amst. Neth.* 69, 265–271.
- Planchard, D., Kim, T.M., Mazieres, J., Quoix, E., Riely, G., Barlesi, F., Souquet, P.-J., Smit, E.F., Groen, H.J.M., Kelly, R.J., et al. (2016a). Dabrafenib in patients with BRAF(V600E)-positive advanced

- non-small-cell lung cancer: a single-arm, multicentre, open-label, phase 2 trial. *Lancet Oncol.* *17*, 642–650.
- Planchard, D., Besse, B., Groen, H.J.M., Souquet, P.-J., Quoix, E., Baik, C.S., Barlesi, F., Kim, T.M., Mazieres, J., Novello, S., et al. (2016b). Dabrafenib plus trametinib in patients with previously treated BRAF(V600E)-mutant metastatic non-small cell lung cancer: an open-label, multicentre phase 2 trial. *Lancet Oncol.* *17*, 984–993.
- Platz, A., Egyhazi, S., Ringborg, U., and Hansson, J. (2008). Human cutaneous melanoma; a review of *NRAS* and *BRAF* mutation frequencies in relation to histogenetic subclass and body site. *Mol. Oncol.* *1*, 395–405.
- Pollmann, M., Parwaresch, R., Adam-Klages, S., Kruse, M.-L., Buck, F., and Heidebrecht, H.-J. (2006). Human EML4, a novel member of the EMAP family, is essential for microtubule formation. *Exp. Cell Res.* *312*, 3241–3251.
- Poomakkoth, N., Issa, A., Abdulrahman, N., Abdelaziz, S.G., and Mraiche, F. (2016). p90 ribosomal S6 kinase: a potential therapeutic target in lung cancer. *J. Transl. Med.* *14*.
- Posch, C., Moslehi, H., Feeney, L., Green, G.A., Ebaee, A., Feichtenschlager, V., Chong, K., Peng, L., Dimon, M.T., Phillips, T., et al. (2013). Combined targeting of MEK and PI3K/mTOR effector pathways is necessary to effectively inhibit NRAS mutant melanoma in vitro and in vivo. *Proc. Natl. Acad. Sci.* *110*, 4015–4020.
- Poulikakos, P.I., Persaud, Y., Janakiraman, M., Kong, X., Ng, C., Moriceau, G., Shi, H., Atefi, M., Titz, B., Gabay, M.T., et al. (2011). RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E). *Nature* *480*, 387–390.
- Pulford, K., Morris, S.W., and Turturro, F. (2004). Anaplastic lymphoma kinase proteins in growth control and cancer. *J. Cell. Physiol.* *199*, 330–358.
- Purba, E., Saita, E., and Maruyama, I. (2017). Activation of the EGF Receptor by Ligand Binding and Oncogenic Mutations: The “Rotation Model.” *Cells* *6*, 13.
- Pylayeva-Gupta, Y., Grabocka, E., and Bar-Sagi, D. (2011). RAS oncogenes: weaving a tumorigenic web. *Nat. Rev. Cancer* *11*, 761–774.
- Rauen, K.A. (2007). HRAS and the Costello syndrome. *Clin. Genet.* *71*, 101–108.
- Reddy, V.P., Gay, L.M., Elvin, J.A., Vergilio, J.-A., Suh, J., Ramkissoon, S., Daniel, S., Severson, E.A., Ali, S.M., Schrock, A.B., et al. (2017). *BRAF* fusions in clinically advanced non-small cell lung cancer: An emerging target for anti-*BRAF* therapies. *J. Clin. Oncol.* *35*, 9072–9072.
- Redig, A.J., Capelletti, M., Dahlberg, S.E., Sholl, L.M., Mach, S., Fontes, C., Shi, Y., Chalasani, P., and Jänne, P.A. (2016). Clinical and Molecular Characteristics of NF1-Mutant Lung Cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *22*, 3148–3156.
- Rhee, H.-W., Zou, P., Udeshi, N.D., Martell, J.D., Mootha, V.K., Carr, S.A., and Ting, A.Y. (2013). Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* *339*, 1328–1331.
- Richards, M.W., O’Regan, L., Roth, D., Montgomery, J.M., Straube, A., Fry, A.M., and Bayliss, R. (2015). Microtubule association of EML proteins and the EML4-ALK variant 3 oncoprotein require an N-terminal trimerization domain. *Biochem. J.* *467*, 529–536.
- Ridley, A.J. (2015). Rho GTPase signalling in cell migration. *Curr. Opin. Cell Biol.* *36*, 103–112.

- Riggs, K.A., Hasan, N., Humphrey, D., Raleigh, C., Nevitt, C., Corbin, D., and Hu, C. (2012). Regulation of integrin endocytic recycling and chemotactic cell migration by syntaxin 6 and VAMP3 interaction. *J. Cell Sci.* *125*, 3827–3839.
- Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* *196*, 801–810.
- Roy, S., Lane, A., Yan, J., McPherson, R., and Hancock, J.F. (1997). Activity of plasma membrane-recruited Raf-1 is regulated by Ras via the Raf zinc finger. *J. Biol. Chem.* *272*, 20139–20145.
- Sasaki, T., Rodig, S.J., Chirieac, L.R., and Jänne, P.A. (2010). The biology and treatment of EML4-ALK non-small cell lung cancer. *Eur. J. Cancer* *46*, 1773–1780.
- Schopp, I.M., Amaya Ramirez, C.C., Debeljak, J., Kreibich, E., Skribbe, M., Wild, K., and Béthune, J. (2017). Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. *Nat. Commun.* *8*, 15690.
- Schubbert, S., Zenker, M., Rowe, S.L., Böll, S., Klein, C., Bollag, G., van der Burgt, I., Musante, L., Kalscheuer, V., Wehner, L.-E., et al. (2006). Germline KRAS mutations cause Noonan syndrome. *Nat. Genet.* *38*, 331–336.
- Schubbert, S., Shannon, K., and Bollag, G. (2007). Hyperactive Ras in developmental disorders and cancer. *Nat. Rev. Cancer* *7*, 295–308.
- Sequist, L.V., Waltman, B.A., Dias-Santagata, D., Digumarthy, S., Turke, A.B., Fidias, P., Bergethon, K., Shaw, A.T., Gettinger, S., Cosper, A.K., et al. (2011). Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Sci. Transl. Med.* *3*, 75ra26.
- Shao, D.D., Xue, W., Krall, E.B., Bhutkar, A., Piccioni, F., Wang, X., Schinzel, A.C., Sood, S., Rosenbluh, J., Kim, J.W., et al. (2014). KRAS and YAP1 converge to regulate EMT and tumor survival. *Cell* *158*, 171–184.
- Shaw, A.T., and Solomon, B. (2011). Targeting anaplastic lymphoma kinase in lung cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *17*, 2081–2086.
- Shaw, A.T., Kim, D.-W., Nakagawa, K., Seto, T., Crinó, L., Ahn, M.-J., De Pas, T., Besse, B., Solomon, B.J., Blackhall, F., et al. (2013). Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. *N. Engl. J. Med.* *368*, 2385–2394.
- Shaw, A.T., Gandhi, L., Gadgeel, S., Riely, G.J., Cetnar, J., West, H., Camidge, D.R., Socinski, M.A., Chiappori, A., Mekhail, T., et al. (2016). Alectinib in ALK-positive, crizotinib-resistant, non-small-cell lung cancer: a single-group, multicentre, phase 2 trial. *Lancet Oncol.* *17*, 234–242.
- Shi, H., Moriceau, G., Kong, X., Lee, M.-K., Lee, H., Koya, R.C., Ng, C., Chodon, T., Scolyer, R.A., Dahlman, K.B., et al. (2012). Melanoma whole-exome sequencing identifies (V600E)B-RAF amplification-mediated acquired B-RAF inhibitor resistance. *Nat. Commun.* *3*, 724.
- Shostak, K., and Chariot, A. (2015). EGFR and NF- κ B: partners in cancer. *Trends Mol. Med.* *21*, 385–393.
- Simanshu, D.K., Nissley, D.V., and McCormick, F. (2017). RAS Proteins and Their Regulators in Human Disease. *Cell* *170*, 17–33.
- Skoulidis, F., and Papadimitrakopoulou, V.A. (2017). Targeting the Gatekeeper: Osimertinib in EGFR T790M Mutation-Positive Non-Small Cell Lung Cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *23*, 618–622.

- Soda, M., Choi, Y.L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S., Watanabe, H., Kurashina, K., Hatanaka, H., et al. (2007). Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature* 448, 561–566.
- Söderberg, O., Gullberg, M., Jarvius, M., Ridderstråle, K., Leuchowius, K.-J., Jarvius, J., Wester, K., Hydbring, P., Bahram, F., Larsson, L.-G., et al. (2006). Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat. Methods* 3, 995–1000.
- Solomon, B.J., Mok, T., Kim, D.-W., Wu, Y.-L., Nakagawa, K., Mekhail, T., Felip, E., Cappuzzo, F., Paolini, J., Usari, T., et al. (2014). First-line crizotinib versus chemotherapy in ALK-positive lung cancer. *N. Engl. J. Med.* 371, 2167–2177.
- Song, Z. (2013). Roles of the nucleotide sugar transporters (SLC35 family) in health and disease. *Mol. Aspects Med.* 34, 590–600.
- Splinter, D., Tanenbaum, M.E., Lindqvist, A., Jaarsma, D., Flotho, A., Yu, K.L., Grigoriev, I., Engelsma, D., Haasdijk, E.D., Keijzer, N., et al. (2010). Bicaudal D2, dynein, and kinesin-1 associate with nuclear pore complexes and regulate centrosome and nuclear positioning during mitotic entry. *PLoS Biol.* 8, e1000350.
- Stephen, A.G., Esposito, D., Bagni, R.K., and McCormick, F. (2014). Dragging Ras Back in the Ring. *Cancer Cell* 25, 272–281.
- Sun, S., Schiller, J.H., and Gazdar, A.F. (2007). Lung cancer in never smokers — a different disease. *Nat. Rev. Cancer* 7, 778–790.
- Sunaga, N., Imai, H., Shimizu, K., Shames, D.S., Kakegawa, S., Girard, L., Sato, M., Kaira, K., Ishizuka, T., Gazdar, A.F., et al. (2012). Oncogenic KRAS-induced interleukin-8 overexpression promotes cell growth and migration and contributes to aggressive phenotypes of non-small cell lung cancer. *Int. J. Cancer* 130, 1733–1744.
- Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS ONE* 6, e21800.
- Symonds, J.M., Ohm, A.M., Carter, C.J., Heasley, L.E., Boyle, T.A., Franklin, W.A., and Reyland, M.E. (2011). Protein kinase C δ is a downstream effector of oncogenic K-ras in lung tumors. *Cancer Res.* 71, 2087–2097.
- Taipale, M., Krykbaeva, I., Koeva, M., Kayatekin, C., Westover, K.D., Karras, G.I., and Lindquist, S. (2012). Quantitative Analysis of Hsp90-Client Interactions Reveals Principles of Substrate Recognition. *Cell* 150, 987–1001.
- Tan, Y.-H.C., Krishnaswamy, S., Nandi, S., Kanteti, R., Vora, S., Onel, K., Hasina, R., Lo, F.-Y., El-Hashani, E., Cervantes, G., et al. (2010). CBL is frequently altered in lung cancers: its relationship to mutations in MET and EGFR tyrosine kinases. *PLoS One* 5, e8972.
- Tan, Y.-H.C., Mirzapoiiazova, T., Won, B.M., Zhu, L., Srivastava, M.K., Vokes, E.E., Husain, A.N., Batra, S.K., Sharma, S., and Salgia, R. (2017). Differential responsiveness of MET inhibition in non-small-cell lung cancer with altered CBL. *Sci. Rep.* 7.
- Tchevkina, E., and Komelkov, A. (2012). Protein Phosphorylation as a Key Mechanism of mTORC1/2 Signaling Pathways. In *Protein Phosphorylation in Human Health*, C. Huang, ed. (InTech), p.
- Teo, G., Liu, G., Zhang, J., Nesvizhskii, A.I., Gingras, A.-C., and Choi, H. (2014). SAINTexpress: improvements and additional features in Significance Analysis of INTERactome software. *J. Proteomics* 100, 37–43.

- Thompson, J.C., Yee, S.S., Troxel, A.B., Savitch, S.L., Fan, R., Balli, D., Lieberman, D.B., Morrisette, J.D., Evans, T.L., Bauml, J., et al. (2016). Detection of Therapeutically Targetable Driver and Resistance Mutations in Lung Cancer Patients by Next-Generation Sequencing of Cell-Free Circulating Tumor DNA. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 22, 5772–5782.
- Tian, T., Harding, A., Inder, K., Plowman, S., Parton, R.G., and Hancock, J.F. (2007). Plasma membrane nanoswitches generate high-fidelity Ras signal transduction. *Nat. Cell Biol.* 9, 905–914.
- To, M.D., Wong, C.E., Karnezis, A.N., Del Rosario, R., Di Lauro, R., and Balmain, A. (2008). Kras regulatory elements and exon 4A determine mutation specificity in lung cancer. *Nat. Genet.* 40, 1240–1244.
- Tran, N.H., Wu, X., and Frost, J.A. (2005). B-Raf and Raf-1 are regulated by distinct autoregulatory mechanisms. *J. Biol. Chem.* 280, 16244–16253.
- Tsao, H., Chin, L., Garraway, L.A., and Fisher, D.E. (2012). Melanoma: from mutations to medicine. *Genes Dev.* 26, 1131–1155.
- Tyczynski, J.E., Bray, F., and Parkin, D.M. (2003). Lung cancer in Europe in 2000: epidemiology, prevention, and early detection. *Lancet Oncol.* 4, 45–55.
- Vander Heiden, M.G. (2011). Targeting cancer metabolism: a therapeutic window opens. *Nat. Rev. Drug Discov.* 10, 671–684.
- Vargas, A.J., and Harris, C.C. (2016). Biomarker development in the precision medicine era: lung cancer as a case study. *Nat. Rev. Cancer* 16, 525–537.
- Vaughan, C.K., Mollapour, M., Smith, J.R., Truman, A., Hu, B., Good, V.M., Panaretou, B., Neckers, L., Clarke, P.A., Workman, P., et al. (2008). Hsp90-Dependent Activation of Protein Kinases Is Regulated by Chaperone-Targeted Dephosphorylation of Cdc37. *Mol. Cell* 31, 886–895.
- Vendrell, J.A., Mau-Them, F.T., Béganton, B., Godreuil, S., Coopman, P., and Solassol, J. (2017a). Circulating Cell Free Tumor DNA Detection as a Routine Tool for Lung Cancer Patient Management. *Int. J. Mol. Sci.* 18.
- Vendrell, J.A., Taviaux, S., Béganton, B., Godreuil, S., Audran, P., Grand, D., Clermont, E., Serre, I., Szablewski, V., Coopman, P., et al. (2017b). Detection of known and novel ALK fusion transcripts in lung cancer patients using next-generation sequencing approaches. *Sci. Rep.* 7.
- Ventura, R.A., Martin-Subero, J.I., Knippschild, U., Gascoyne, R.D., Delsol, G., Mason, D.Y., and Siebert, R. (2004). Centrosome abnormalities in ALK-positive anaplastic large-cell lymphoma. *Leukemia* 18, 1910–1911.
- Vigil, D., Cherfils, J., Rossman, K.L., and Der, C.J. (2010). Ras superfamily GEFs and GAPs: validated and tractable targets for cancer therapy? *Nat. Rev. Cancer* 10, 842–857.
- Villanueva, J., Vultur, A., Lee, J.T., Somasundaram, R., Fukunaga-Kalabis, M., Cipolla, A.K., Wubbenhorst, B., Xu, X., Gimotty, P.A., Kee, D., et al. (2010). Acquired resistance to BRAF inhibitors mediated by a RAF kinase switch in melanoma can be overcome by cotargeting MEK and IGF-1R/PI3K. *Cancer Cell* 18, 683–695.
- Wagle, N., Van Allen, E.M., Treacy, D.J., Frederick, D.T., Cooper, Z.A., Taylor-Weiner, A., Rosenberg, M., Goetz, E.M., Sullivan, R.J., Farlow, D.N., et al. (2014). MAP Kinase Pathway Alterations in BRAF-Mutant Melanoma Patients with Acquired Resistance to Combined RAF/MEK Inhibition. *Cancer Discov.* 4, 61–68.

- Wakioka, T., Sasaki, A., Kato, R., Shouda, T., Matsumoto, A., Miyoshi, K., Tsuneoka, M., Komiya, S., Baron, R., and Yoshimura, A. (2001). Spred is a Sprouty-related suppressor of Ras signalling. *Nature* *412*, 647–651.
- Wang, M., and Casey, P.J. (2016). Protein prenylation: unique fats make their mark on biology. *Nat. Rev. Mol. Cell Biol.* *17*, 110–122.
- Wang, S., Song, Y., and Liu, D. (2017). EAI045: The fourth-generation EGFR inhibitor overcoming T790M and C797S resistance. *Cancer Lett.* *385*, 51–54.
- Wang, Y.-N., Yamaguchi, H., Hsu, J.-M., and Hung, M.-C. (2010). Nuclear trafficking of the epidermal growth factor receptor family membrane proteins. *Oncogene* *29*, 3997–4006.
- Ware, K.E., Hinz, T.K., Kleczko, E., Singleton, K.R., Marek, L.A., Helfrich, B.A., Cummings, C.T., Graham, D.K., Astling, D., Tan, A.-C., et al. (2013). A mechanism of resistance to gefitinib mediated by cellular reprogramming and the acquisition of an FGF2-FGFR1 autocrine growth loop. *Oncogenesis* *2*, e39.
- Weber, C.K., Slupsky, J.R., Kalmes, H.A., and Rapp, U.R. (2001). Active Ras induces heterodimerization of cRaf and BRaf. *Cancer Res.* *61*, 3595–3598.
- Weinberg, F., Hamanaka, R., Wheaton, W.W., Weinberg, S., Joseph, J., Lopez, M., Kalyanaraman, B., Mutlu, G.M., Budinger, G.R.S., and Chandel, N.S. (2010). Mitochondrial metabolism and ROS generation are essential for Kras-mediated tumorigenicity. *Proc. Natl. Acad. Sci.* *107*, 8788–8793.
- Whitesell, L., and Lindquist, S.L. (2005). HSP90 and the chaperoning of cancer. *Nat. Rev. Cancer* *5*, 761–772.
- Wilker, E., and Yaffe, M.B. (2004). 14-3-3 Proteins—a focus on cancer and human disease. *J. Mol. Cell. Cardiol.* *37*, 633–642.
- Xiong, Q.-Y., Yu, C., Zhang, Y., Ling, L., Wang, L., and Gao, J.-L. (2017). Key proteins involved in insulin vesicle exocytosis and secretion. *Biomed. Rep.* *6*, 134–139.
- Yadav, K.K., and Bar-Sagi, D. (2010). Allosteric gating of Son of sevenless activity by the histone domain. *Proc. Natl. Acad. Sci. U. S. A.* *107*, 3436–3440.
- Yoshida, T., Hisamoto, T., Akiba, J., Koga, H., Nakamura, K., Tokunaga, Y., Hanada, S., Kumemura, H., Maeyama, M., Harada, M., et al. (2006). Spreds, inhibitors of the Ras/ERK signal transduction, are dysregulated in human hepatocellular carcinoma and linked to the malignant phenotype of tumors. *Oncogene* *25*, 6056–6066.
- Yun, C.-H., Mengwasser, K.E., Toms, A.V., Woo, M.S., Greulich, H., Wong, K.-K., Meyerson, M., and Eck, M.J. (2008). The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 2070–2075.
- Zanucco, E., El-Nikhely, N., Götz, R., Weidmann, K., Pfeiffer, V., Savai, R., Seeger, W., Ullrich, A., and Rapp, U.R. (2014). Elimination of B-RAF in oncogenic C-RAF-expressing alveolar epithelial type II cells reduces MAPK signal intensity and lung tumor growth. *J. Biol. Chem.* *289*, 26804–26816.
- Zhang, S., Bao, Y., Ju, X., Li, K., Shang, H., Ha, L., Qian, Y., Zou, L., Sun, X., Li, J., et al. (2015). BA-j as a novel CDK1 inhibitor selectively induces apoptosis in cancer cells by regulating ROS. *Sci. Rep.* *5*, 13626.

Zhang, Y., Shu, L., and Chen, X. (2008). Syntaxin 6, a Regulator of the Protein Trafficking Machinery and a Target of the p53 Family, Is Required for Cell Adhesion and Survival. *J. Biol. Chem.* 283, 30689–30698.

Zhapparova, O.N., Fokin, A.I., Vorobyeva, N.E., Bryantseva, S.A., and Nadezhdina, E.S. (2013). Ste20-like protein kinase SLK (LOSK) regulates microtubule organization by targeting dynactin to the centrosome. *Mol. Biol. Cell* 24, 3205–3214.

Zipfel, P.A., Brady, D.C., Kashatus, D.F., Ancrile, B.D., Tyler, D.S., and Counter, C.M. (2010). Ral activation promotes melanomagenesis. *Oncogene* 29, 4859–4864.

Zotti, T., Scudiero, I., Vito, P., and Stilo, R. (2017). The Emerging Role of TRAF7 in Tumor Development: TRAF7 IN TUMOR DEVELOPMENT. *J. Cell. Physiol.* 232, 1233–1238.

Abstract

Lung cancer is the leading cause of cancer-related death in France and in the world. It is a cancer of poor prognosis, diagnosed at the late stage III or IV, with a low 5-years survival of 6% and 1%, respectively. This cancer encompasses several histological types, and among them adenocarcinoma are the most frequent and account for 40% of the diagnosis. Genetic analysis of stage IV tumors reveals redundant alterations, and generally exclusives from each other's, of *KRAS*, *EGFR*, *BRAF* and *ALK* genes. The identification of these genetic alterations enables, within companion test, to make eligible patients for targeted therapies when molecules are available.

Even though these targeted therapies represent a true revolution in patient's care, the majority of these patients cannot benefit from such due to the absence of targetable activating mutations (e.g. no *EGFR*, *BRAF* or *ALK* alterations, presence of *KRAS* mutation ...). Additionally, even if they can be treated by an oral molecule, the observed benefit unfortunately remains transient and all these patients will eventually relapse. Some acquired resistance mechanisms are particularly well described, such is the case of the *EGFR* T790M secondary mutation directly affecting the protein targeted by the therapy, and accounting for half of resistance events.

To better understand the molecular mechanisms associated with the resistance events observed in patients, and to propose new therapeutics for patient non-eligible for targeted therapies, I studied the impact of the canonical lung cancer mutations on protein interaction networks. I thus focused my work on the network characterization of *EGFR*, *KRAS*, *BRAF* activating mutations and *ALK* translocations. Considering that mutations of proteins from the RAS family (*HRAS*, *NRAS* and *KRAS*) account for more than 25% of cancers, and given the higher rate of *KRAS* mutations *versus* *HRAS* and *NRAS*, I have also conducted a comparative interactomic study between these three isoforms to uncover *KRAS* interactome specificities.

During my thesis, I showed that the protein networks characterized using BioID (proximity-dependent biotinylation identification) are much richer compared to those identified with the more conventional technique of affinity purification, and that they enable efficient identification of interactors specifically deregulated upon activating mutation. Additionally, the HEK293 cell model (Human Embryonic Kidney) and BEAS2B cell model (non-cancerous lung cell line) showed a high overlapping degree of the interactors identified, suggesting that the strategy used is relevant to identify interactome modifications related to specific lung cancer genetic alterations. This thesis allowed the identification of several interactors specific to the mutants of *KRAS*, *EGFR*, *BRAF*, *NRAS* and *EML4-ALK* fusion. Thirteen interactors specific to the mutated-*KRAS* have been functionally validated in lung-cancer cell lines models. Finally, using BioID data I have been able to propose a preliminary model of *EGFR* resistance to targeted therapies. This model suggests that *CBL* and *IGF2R* could partner to *EGFR* and play a role in the therapeutic escape.

Overall, this thesis opens new perspectives to determine mechanisms of resistance to targeted therapies, and to identify new therapeutic targets for patients harboring an activating mutation or translocation.

Key-words: Lung cancer, targeted therapies, resistance, interactome, cell signaling

Résumé

Le cancer du poumon est la première cause de mortalité liée au cancer en France et dans le monde. Il s'agit d'un cancer de mauvais pronostic, diagnostiqué généralement aux stades tardifs III ou IV et avec une survie à 5 ans faible, respectivement de 6 et 1%. Ce cancer comprend différents types histologiques, parmi lesquels les adénocarcinomes sont les plus fréquents et comptent pour 40% des diagnostics. L'analyse génétique des tumeurs des patients en stade IV révèle des altérations récurrentes, et généralement mutuellement exclusives, au niveau des gènes *KRAS*, *EGFR*, *BRAF* et *ALK*. L'identification de ces altérations génétiques permet dans le cadre de tests compagnons de décider de l'éligibilité des patients aux thérapies ciblées lorsqu'elles sont disponibles.

Bien que ces thérapies ciblées représentent une véritable révolution dans la prise en charge des patients, une majorité de patients ne peuvent bénéficier de ces traitements car ils ne présentent pas au niveau tumoral de mutations activatrices actionnables (absence d'altération de l'*EGFR*, *BRAF* ou *ALK*, présence d'une mutation *KRAS* ...). De plus, dans l'éventualité où ils peuvent être traités par une molécule orale, le bénéfice observé reste malheureusement de courte durée et tous finiront à terme par échapper au traitement. Certains mécanismes de résistances acquises sont par ailleurs particulièrement bien décrits comme la mutation secondaire *EGFR* T790M qui affecte directement la protéine traitée par la thérapie ciblée, et compte pour environ la moitié des cas de résistances.

Ainsi, afin de mieux comprendre quels sont les mécanismes moléculaires associés aux phénomènes de résistances thérapeutiques observés chez les patients, et dans l'objectif de proposer de nouvelles voies thérapeutiques pour les patients non éligibles aux thérapies ciblées, j'ai étudié l'impact des mutations canoniques du cancer du poumon sur les réseaux d'interactions protéiques. Plus particulièrement, j'ai focalisé mon travail sur la caractérisation des réseaux d'interactions des mutations activatrices de l'*EGFR*, *KRAS*, *BRAF* et des translocations *ALK*. Les protéines de la famille RAS (*HRAS*, *NRAS*, *KRAS*) étant mutées dans plus de 25% des cancers, et étant donné le taux plus élevé de mutations de *KRAS* versus *HRAS* et *NRAS*, j'ai également mené une étude comparative d'interactomes entre ces trois isoformes afin de mettre en évidence les interactions protéiques spécifiques à l'isoforme *KRAS*.

Au cours de ma thèse, j'ai montré que les réseaux d'interactions identifiés par la technique du BioID (proximity-dependent biotinylation identification) sont bien plus riches que ceux obtenus par des techniques plus conventionnelles de purification par affinité, et qu'ils permettent de mettre en évidence des interacteurs spécifiquement dérégulés par l'apparition de mutations activatrices. Par ailleurs, les modèles cellulaires HEK293 (cellules humaines embryonnaires de rein) et BEAS2B (cellules pulmonaires d'origine non-cancéreuse) ont montré un fort recouvrement des interacteurs identifiés, ce qui suggère que la stratégie employée est pertinente pour identifier des interacteurs spécifiques de mutations. Ce travail de thèse a permis d'identifier une série de plusieurs interacteurs spécifiques des formes mutantes de *KRAS*, d'*EGFR*, de *BRAF*, de *NRAS* et de la fusion *EML4-ALK*. Treize interacteurs spécifiques de la forme mutée de *KRAS* ont été validés fonctionnellement dans des modèles cellulaires de cancers du poumon. Enfin, en utilisant les données de BioID, j'ai également proposé un modèle préliminaire de résistance de l'*EGFR* aux thérapies ciblées. Ce modèle suggère que *CBL* et *IGF2R* pourraient interagir avec l'*EGFR* et être impliquées dans l'échappement thérapeutique observé.

Ce travail de thèse ouvre ainsi de nouvelles perspectives quant à la détermination des mécanismes de résistance spécifiques aux thérapies ciblées, et à l'identification de nouvelles cibles thérapeutiques chez les patients qui présentent une mutation activatrice ou une translocation.

Mots-clés : Cancer du poumon, thérapies ciblées, résistance, interactome, signalisation cellulaire