



HAL
open science

CRISPR-barcoding pour l'étude fonctionnelle de mutations oncogéniques dans un contexte d'hétérogénéité intra-tumorale.

Alexis Guernet

► **To cite this version:**

Alexis Guernet. CRISPR-barcoding pour l'étude fonctionnelle de mutations oncogéniques dans un contexte d'hétérogénéité intra-tumorale.. Biologie cellulaire. Normandie Université, 2017. Français. NNT : 2017NORMR022 . tel-01635638

HAL Id: tel-01635638

<https://theses.hal.science/tel-01635638v1>

Submitted on 15 Nov 2017

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.



Normandie Université

THÈSE

Pour obtenir le diplôme de doctorat

Spécialité Aspects Moléculaires et Cellulaires de la Biologie

Préparée au sein de l'Université de Rouen

CRISPR-barcoding pour l'étude fonctionnelle de mutations oncogéniques dans un contexte d'hétérogénéité intratumorale

**Présentée et soutenue par
Alexis GUERNET**

**Thèse soutenue publiquement le 27 septembre 2017
devant le jury composé de**

Pr Alexis CORTOT	Professeur des Universités – Praticien Hospitalier, Lille	Rapporteur
Dr Cédric COULOUARN	Chargé de Recherche INSERM, Rennes	Rapporteur
Dr Carine GIOVANNANGELI	Directrice de Recherche INSERM, Paris	Examineur
Dr Brigitte SOLA	Directrice de Recherche INSERM, Caen	Examineur
Dr Youssef ANOUAR	Directeur de Recherche INSERM, Rouen	Directeur de thèse
Dr Luca GRUMOLATO	Maître de Conférences, Rouen	Codirecteur de thèse

**Thèse dirigée par les Drs Luca GRUMOLATO et Youssef ANOUAR,
Unité INSERM 1239, Laboratoire de Différenciation et Communication Neuronale et Neuroendocrine**



Je pense que tout est possible, à qui rêve, ose, travaille et n'abandonne jamais.

Xavier Dolan

REMERCIEMENTS

Ces travaux de thèse ont été réalisés au sein du laboratoire de Différenciation et Communication Neuronale et Neuroendocrine (DC2N), Unité INSERM 1239, dirigé par le Docteur Youssef Anouar.

Ces études ont été financées par la région Normandie et ont également bénéficié de l'aide généreuse des organismes suivants :

- L'Institut National de la Santé et de la Recherche Médicale (INSERM)
- L'Université Rouen Normandie
- L'Institut de Recherche et d'Innovation Biomédicale (IRIB)
- L'Ecole Doctorale Normande de Biologie Intégrative, Santé et Environnement
- La ligue contre le cancer, comité de Seine maritime

Je tiens tout d'abord à remercier les membres de mon jury de thèse.

*Le Professeur **Alexis Cortot**, Professeur des universités et praticien hospitalier, Université de Lille, me fait l'honneur de siéger dans ce jury en qualité de rapporteur. Je tiens à lui exprimer ma gratitude pour l'intérêt qu'il témoigne ainsi à nos recherches.*

*Le Docteur **Cédric Coulouarn**, Chargé de Recherche à l'INSERM, Université de Rennes, a volontiers accepté de siéger dans ce jury en qualité de rapporteur. Je l'en remercie très sincèrement.*

*Le Docteur **Carine Giovannangeli**, Directrice de Recherche au CNRS, Sorbonnes Universités, a bien voulu examiner ces travaux. Je lui signifie ici toute ma reconnaissance et la remercie chaleureusement pour les conseils qu'elle m'a prodigués lors de l'Atelier INSERM sur la révolution CRISPR/Cas9 auquel j'ai participé en 2016.*

*Le Docteur **Brigitte Sola**, Directrice de Recherche à l'INSERM, Université de Caen Normandie, me fait également l'honneur d'être examinatrice de ce travail. Je tiens à la remercier sincèrement pour la faveur qu'elle m'accorde.*

*Un grand Merci au Docteur **Youssef Anouar**, Directeur de Recherche à l'INSERM, Université de Rouen Normandie, qui m'a donné tous les moyens nécessaires à la réalisation et à l'aboutissement de ce travail. Je tiens à lui témoigner ici toute ma gratitude pour la confiance qu'il m'a accordée, ses encouragements et ses conseils qui m'ont poussé à continuer dans la voie de la recherche et de l'excellence scientifique.*

*Je tiens tout particulièrement à remercier le Docteur **Luca Grumolato**, Maître de Conférences à l'Université de Rouen Normandie. Je te suis infiniment reconnaissant de m'avoir fait confiance il y a 5 ans lors de mon stage de Master. Au cours de ces années, tu es devenu un véritable mentor pour moi, un modèle de rigueur et d'excellence scientifique, que j'espère un jour égaler. Tu as su me transmettre ton amour pour la recherche et ton incroyable créativité scientifique. Toutes ces années pendant lesquelles nous avons mené ces projets, préparé des présentations, rédigé des documents, m'ont permis d'évoluer et d'acquérir des compétences qui me serviront tant pour ma vie professionnelle que personnelle. Malgré un emploi du temps souvent très chargé, je te remercie chaleureusement du temps que tu as passé à me former et à faire de moi le chercheur que je suis aujourd'hui malgré mon caractère parfois bien trempé. Tu m'as tout appris dans ce domaine. Enfin, je te remercie d'avoir été présent de mon premier jour de stage en master jusqu'à ma recherche de post-doc, et de n'avoir eu cesse de m'aider et de me prodiguer tes conseils pour la suite de ma carrière.*

Je remercie également les personnes qui font ou qui ont fait partie de notre petite Team.

Ulisse Bocchero, une personne formidable que j'ai eu la chance de côtoyer sur le projet Wnt au cours de mon master. Merci pour toutes ces heures de manip' passées ensemble dans la joie et la bonne humeur.

Hassan Ainani, qui a contribué au développement de CRISPR-barcoding. Ces moments de partage à la paillasse et d'apprentissage de l'arabe resteront gravés dans ma mémoire. Merci pour ta sympathie.

Faisal et ses petits plats. On s'est connu au cours de mon master 2 lors de ton post-doc avec Luca. Tu as été comme un grand frère pour moi au sein du laboratoire. Tu m'as donné des conseils précieux pour la thèse et pour les expériences. Tu as toujours été disponible pour répondre à mes questions et m'éclairer quand j'en avais besoin. Je te remercie beaucoup et remercie pour moi ta femme. Vous revoir la veille de mon départ pour Boston a été pour moi un instant inoubliable. Je vous souhaite le meilleur pour l'avenir.

Ohhhh Sylvain !!!! THE stagiaire qui est maintenant thésard ! Tu es LA rencontre du labo, un mec complètement barré en apparence mais avec un petit cœur d'artichaut en dessous. Le « petit biscuit » de ces dames comme Marjo aime t'appeler. Qui de mieux pour comprendre ce que je vis qu'une personne qui vis sensiblement la même chose au quotidien ? On en aura passé du temps ensemble, c'est simple ces trois dernières années j'ai passé plus de temps avec toi qu'avec Anastasia que ce soit au labo ou à la maison devant la PS3, à corriger des copies, au bar, au restaurant. Que de bons moments que je n'oublierai jamais ! Encore merci mec, si j'en suis là c'est en partie grâce à toi et je te souhaite de faire une thèse aussi passionnante que la mienne, mais vu ton talent je ne me fait aucun soucis pour toi !

Je souhaite également dire toute ma reconnaissance au Docteur David Alexandre, Maître de Conférences à l'Université de Rouen Normandie, pour l'ensemble des discussions que nous avons pu avoir et qui m'ont permis d'aiguiser mon esprit scientifique. Je tiens également à te remercier chaleureusement, pour le soutien, les conseils et l'amitié que tu m'as accordés. Merci pour tes blagues et tes « boujous » qui ont bercé mes jours depuis cette dernière année.

Je remercie particulièrement les membres de l'équipe 1.

Anthony, merci pour les conseils et les discussions scientifiques que nous avons pu avoir au cours de ces dernières années. Je te suis également reconnaissant de la sympathie que tu m'as accordé et du soutien que tu as été, notamment dans la dernière ligne droite pendant laquelle tu venais prendre des nouvelles tous les jours pour voir l'avancé avec toujours un petit schokobons pour me booster.

Ma grande copine Ophélie, ralalala qu'auraient été mes journées au labo sans toi ? Je me le demande même pas. Tu as été une pièce importante du puzzle. Toujours souriante, toujours de bonne humeur, serviable pour s'occuper de mes loulous pendant les congrès ou autres déplacements, tu es une copine en or. Tu es quelqu'un de génial (évidemment tu es de l'équipe 1 XD). J'ai de la chance de t'avoir comme amie et je te souhaite bon courage pour la fin de ta thèse, je sais que tu vas gérer la fougère !

Je remercie Aurélien, pour toutes ces heures à discuter de science ou d'autres sujets. Merci pour ton calme légendaire, ta bienveillance et ton amitié. Tu vas cartonner dans l'avenir j'en suis certain.

Je tiens également à remercier Loubna et Ifat, mes voisines de bureau (ou du moins ex-voisine pour Loubna) pour votre bonne humeur quotidienne et votre sympathie.

Matthieu, la personne la plus bavarde que je connaisse après moi. Merci pour tes conseils, merci pour les verres, les soirées et merci pour toutes ces vengeance orchestrées secrètement.

Charlène, Marion, Abdallah et Houssni, merci pour les nombreuses heures passées à la paillasse ensemble. Ces moments de travail, de joie, de rire, resteront longtemps dans ma mémoire.

Dorthe, la cracotte du labo. Merci pour ta sympathie, tes sourires quotidiens, tes conseils et tes potins. Tu as égayé mes journées, ce fût un véritable plaisir de travailler en ta compagnie.

Le pétillant Hugo et sa joie de vivre débordante. Notre acteur studio, capable de passer du rire à la colère en un claquement de doigt. Merci pour ces franches rigolades, merci d'être un mec simple et cool. Il en faudrait plus des comme toi !

Je tiens évidemment à remercier, Christine, Isabelle, Christophe, Maïté, Lydie, Fanny, Taïmé, ainsi que tous les membres de l'équipe 1 pour l'ambiance chaleureuse qui règne dans l'équipe.

Je tiens à remercier les personnes du laboratoire qui ont été présent pour moi durant toutes ces années.

En premier lieu, je tiens à remercier Pierre-Mickaël, mon frerot de thèse. Merci pour ton soutien quotidien. On a commencé cette aventure ensemble en master 1 et nous avons continué ensemble jusqu'à maintenant. Merci pour tout, pour ces soirées boulot au labo ou chez toi, pour ces verres au 3P, pour ces moments de discussion passionnée dans la salle de culture. Ne désespérons pas peut-être qu'on l'aura notre start-up, qui sait on se retrouvera peut-être un jour ! Bon courage pour la suite, je suis fier de toi !

Un grand merci à la meilleure chanteuse du laboratoire toute catégorie confondue, je parle bien sûr de Mlle Jane-Eileen !!!! Ma tête est remplie de ces souvenirs de nous en train de chanter, de danser en salle de culture. Merci pour ta folie !

Je tiens également à remercier : Benjamin, notre geek au grand cœur ; Sarah pour les pauses déjeuners de dingues en sa compagnie ; le grand sportif Alexandre Mutel pour sa bonne humeur et son soutien ; Aurianne qui m'a épaulé ses dernières semaines et qui est presque devenue ma psy XD ; Karima et son agrafeuse magique ; Dr Julien et ses précieux conseils, tu es un mec génial ; le Docteur Alexandre N ; la handballeuse Julie ; Paulo et son sourire ; Alyson et sa folie constante ; Daninele ; Cynthia ; F-X ; Carmine ; Lisa ; Mahmoud ; Mouna ; POG ; Vadim ; Marine ; Julie M ; Gaëlle ; Hadrien ; Rhita ; Roxanne ; Damien qui est arrivé plus récemment mais qui est une personne entière et sincère qui gagne à être connue.

A tous les thésards, post-Doc, stagiaires. Je vous remercie pour tous les moments que nous avons partagés ensemble. Je suis désolé de toutes les fois où je vous ai fait de mauvaises blagues, pour tous ces sursauts, ces confettis, ces petits mots cachés, ces changements de fond d'écran, ces jets d'eau ou de glace. J'ai passé avec vous 5 années merveilleuses. La thèse n'est pas une période facile, mais vous l'avez rendu belle. Continuez d'être fidèle à vous-même et bon courage pour ceux qui doivent encore passer la thèse, tenez bon !

Un grand merci aux secrétaires, Sabrina, Laurence et Catherine pour leur aide précieuse. Une attention particulière pour toi Catherine, sans qui plus d'une fois je ne serais pas arrivé à destination lors des congrès ou formations. Merci pour ton dévouement sans faille et la gentillesse qui te caractérise.

Je tiens également à remercier tous les membres du laboratoire pour leur gentillesse et leur disponibilité, et particulièrement les Drs Delphine Burel et Julien Chuquet pour la confiance qu'ils m'ont attribuée pour les enseignements. Mais aussi les Drs Sylvie Renouf, Celine Duparc, Isabelle Boutelet, Marie Picot, Estelle Louiset, Hervé Lefebvre, Nicolas Chartrel, Michael Thomas, Gaëtan Prevost, David Vaudry, Olivier Wurtz, Magali Dugay, Benjamin Lefranc, Jérôme Leprince, Helene Castel, Pierrick Gandolfo, Marie-Christine Tonon, Fabrice Morin, Martine Dubois, Laurence Desrues, ainsi que Marjorie Gras.

Un grand merci à notre Hugette qui a tend œuvré pour nous faciliter le quotidien pendant toutes ces années. Nous ne t'oublions pas !

Je tiens enfin à remercier ma famille et mes proches.

Dans la vie il faut savoir rendre à César ce qui appartient à César. Je remercie mes parents qui ont mis tout en œuvre depuis ma plus tendre enfance pour que je puisse réussir dans la vie. Même si aujourd'hui cette vie nous a séparés, je tiens à vous témoigner ainsi qu'à mes sœurs ma sincère gratitude.

A ma sœur Astrid, son conjoint Armand, et leurs enfants Tiana, Arthur et la surprise. Je vous remercie d'avoir cru en moi, d'avoir su être présent au cours de cette dernière année qui a été difficile pour tous. Je vous en suis extrêmement reconnaissant. Je vous souhaite tout le bonheur du monde pour votre nouvelle vie à Carcassonne. Prévoyez l'apéro et la piscine, on arrive !

Je tiens également à remercier, ma tante Yvette, mes cousines Céline et Axelle pour leur soutien sans faille. Merci également à ma belle-famille, Nathalie, Freddy, Alex et Anaïs. Merci également à Jenny et Carla qui m'ont donné ce brin de folie dans ma vie au moment où j'en avais besoin. Merci à mon frerot Vaentin, pour sa compréhension, ce n'est jamais simple de se voir mais je te porte dans mon cœur. Merci également pour le soutien de la famille de Rkia et Laura. Merci à toutes les filles des Majorettes de Chartres, club dans lequel j'ai entraîné pendant 10 ans. Chaque samedi vous m'avez permis de me changer les idées et je vous en remercie !

Rkia, ma grande sœur d'adoption, ma meilleure amie, ma psy, la personne sans qui je serais sûrement devenu fou depuis bien longtemps. Je tiens à te remercier du fond de mon cœur, tu as toujours su être là et trouver les mots pour me booster quand j'étais à bout de souffle et me redonner le sourire, par des petits rien qui font tellement tout. Je t'en suis vraiment reconnaissant. Merci de m'avoir fait prendre confiance en moi et d'avoir su m'épauler et me conseiller chaque jour de ma vie. Je pense que je dois avoir Dieu dans ma poche pour qu'il ait mis sur mon chemin une personne comme toi.

Laura, ma petite sœur de cœur, je te remercie pour ta fraîcheur au quotidien. Tu as su m'apporter cette bulle d'oxygène dont j'avais besoin chaque week-end pour repartir de plus belle au labo, par ces chamailleries, ces délires, ces verres d'eau dans la figure et j'en passe. A la base Dieu nous a fait naître dans deux familles séparées, comme quoi tout le monde peut se tromper.

Un merci tout particulier à mes bébés, Venus et Arya, qui ont su me donner de l'affection et veiller sur moi (à leur façon) quand j'étais seul dans l'appart' à écrire et corriger la thèse.

Je terminerai ici par la personne qui partage ma vie, celle sans qui cette aventure n'aurait jamais été la même. Anastasia, il y a maintenant six ans, on se rencontrait et quelques mois après tu me conseillais de choisir Luca pour mon premier stage en laboratoire de recherche. Quel chemin parcouru, la thèse s'achève, nous sommes maintenant fiancés et nous nous apprêtons à partir à l'aventure à l'étranger. Durant ces années de master puis de thèse, tu as toujours été d'un soutien sans faille, d'une compréhension sans limite et d'un amour inconditionnel. Peu de personnes peuvent comprendre ce que vivent les conjoints des thésards, tu as beaucoup accepté et ceux qui me connaissent vraiment savent à quel point je suis compliqué à vivre (XD). Je ne compte plus le nombre de fois où tu m'as demandé à quelle heure je finissais, et que je rentrais une heure voire deux heures après l'horaire que je t'avais indiquais, ou encore le nombre de samedi ou dimanche passé au labo alors qu'il faisait un temps magnifique. Je te remercie du plus profond de mon cœur, tu es la plus belle rencontre de ma vie, tu es l'amour de ma vie et je n'aurais jamais réussi sans toi. Je t'aime.

PRODUCTION SCIENTIFIQUE

Article et revues dans des périodiques à comité de lecture

Article :

A. Guernet, S.K. Mungamuri, D. Cartier, R. Sachidanandam, A. Jayaprakash, S. Adriouch, M. Vezain, F. Charbonnier, G. Rohkin, S. Coutant, S. Yao, H. Ainani, D. Alexandre, I. Tournier, Boyer, S.A. Aaronson, Y. Anouar & L. Grumolato. (2016) CRISPR-Barcoding for Intratumor Genetic Heterogeneity Modeling and Functional Analysis of Oncogenic Driver Mutations. *Molecular Cell* 63, 526-538

Revues :

A. Guernet, S.A. Aaronson, Y. Anouar & L. Grumolato. (2016) Modeling intratumor heterogeneity through CRISPR-barcodes. *Molecular & Cellular Oncology* 3(6):e1227894

A. Guernet & L. Grumolato. (2017) CRISPR/Cas9 editing of the genome for cancer modeling. *Methods* 121-122, 130-137

Communications orales

A. Guernet, S.K. Mungamuri, D. Cartier, R. Sachidanandam, A. Jayaprakash, S. Adriouch, M. Vezain, F. Charbonnier, G. Rohkin, S. Coutant, S. Yao, H. Ainani, D. Alexandre, I. Tournier, Boyer, S.A. Aaronson, Y. Anouar & L. Grumolato. Functional analysis of oncogenic driver mutations in human cancer cells through CRISPR-barcoding. **19^{èmes} Journées de l'École doctorale NBISE**, 10-11 mars 2016 - Caen, France.

A. Guernet, S.K. Mungamuri, D. Cartier, R. Sachidanandam, A. Jayaprakash, S. Adriouch, M. Vezain, F. Charbonnier, G. Rohkin, S. Coutant, S. Yao, H. Ainani, D. Alexandre, I. Tournier, Boyer, S.A. Aaronson, Y. Anouar & L. Grumolato. CRISPR-barcoding for functional analysis of oncogenic driver mutations in human cancer cells. **Atelier INSERM : La révolution CRISPR/Cas9**, 06-08 avril 2016 - Bordeaux, France.

A. Guernet, S.K. Mungamuri, D. Cartier, R. Sachidanandam, A. Jayaprakash, S. Adriouch, M. Vezain, F. Charbonnier, G. Rohkin, S. Coutant, S. Yao, H. Ainani, D. Alexandre, I. Tournier, Boyer, S.A. Aaronson, Y. Anouar & L. Grumolato. CRISPR-barcoding pour l'étude fonctionnelle de mutations oncogéniques dans un contexte d'hétérogénéité intra-tumorale. **9^{èmes} Journées Scientifiques du Cancéropôle Nord-Ouest**, 18-20 mai 2016 - Deauville, France. *1^{er} prix de communication orale*

Communications affichées

S.F. Mahmood, A. Guernet, U. Bocchero, Y. Anouar and L. Grumolato. ROR1 Oncogenic Addiction in Neuroblastoma: a New Potential Target for Therapy. **2^{ème} Journée de l'Institut de Recherche et d'Innovation Biomédicale**, 21 juin 2013 - Mont Saint Aignan, France.

A. Guernet, S.F. Mahmood, Y. Anouar and L. Grumolato. Dual role of the Wnt non-canonical receptor Ror1 in neuroblastoma cell growth and motility: implication of two distinct mechanisms of receptor activation. **17^{ème} Journée de l'Ecole doctorale NBISE**, 10-11 avril 2014 - Le Havre, France.

A. Guernet, S.F. Mahmood, Y. Anouar and L. Grumolato. Dual role of the Wnt non-canonical receptor Ror1 in neuroblastoma cell growth and motility. **EMBO Conference: Cellular Signaling and Cancer Therapy**, 23-27 mai 2014 - Cavtat, Croatie.

A. Guernet, S.F. Mahmood, Y. Anouar and L. Grumolato. Double rôle du récepteur Wnt non-canonique Ror1 dans la croissance et la motilité des cellules de neuroblastome. **3^{ème} Journée de l'Institut de Recherche et d'Innovation Biomédicale**, 20 juin 2014 - Caugé, France.

A. Guernet, S.F. Mahmood, H. Ainani, Y. Anouar and L. Grumolato. The Wnt noncanonical receptor ROR1 regulates neuroblastoma cell growth and motility through two distinct mechanisms. **Fourth AACR International Conference on FRONTIERS IN BASIC CANCER RESEARCH**, 23-26 octobre 2015 – Philadelphie, Etats-Unis.

A. Guernet, S.K. Mungamuri, D. Cartier, R. Sachidanandam, A. Jayaprakash, S. Adriouch, M. Vezain, F. Charbonnier, G. Rohkin, S. Coutant, S. Yao, H. Ainani, D. Alexandre, I. Tournier, O. Boyer, S.A. Aaronson, Y. Anouar & L. Grumolato. Functional analysis of oncogenic driver mutations in human cancer cells through CRISPR-barcoding. **Fourth AACR International Conference on FRONTIERS IN BASIC CANCER RESEARCH**, 23-26 octobre 2015 – Philadelphie, Etats-Unis.

A. Guernet, S.K. Mungamuri, D. Cartier, R. Sachidanandam, A. Jayaprakash, S. Adriouch, M. Vezain, F. Charbonnier, G. Rohkin, S. Coutant, S. Yao, H. Ainani, D. Alexandre, I. Tournier, Boyer, S.A. Aaronson, Y. Anouar & L. Grumolato. CRISPR-barcoding for intratumor genetic heterogeneity modeling and functional analysis of oncogenic driver mutations. **1^{ère} Journée Normande de Recherche Biomédicale**, 16 septembre 2016 – Rouen, France.

RÉSUMÉ

CRISPR-barcoding pour l'étude fonctionnelle de mutations oncogéniques dans un contexte d'hétérogénéité intratumorale

Les tumeurs sont généralement constituées de différentes sous-populations de cellules cancéreuses génétiquement hétérogènes, responsables en grande partie de la capacité de la tumeur à évoluer rapidement et à s'adapter aux conditions environnementales. Cette diversité génétique a des conséquences majeures pour le patient, notamment au cours de la progression tumorale et pour l'acquisition d'une résistance aux traitements.

Nous avons développé une nouvelle stratégie basée sur la technologie CRISPR/Cas9 qui consiste à introduire, en plus d'une altération de séquence voulue d'un gène d'intérêt, une série de mutations silencieuses, constituant une sorte d'étiquette génétique qui peut être détectée par PCR quantitative ou séquençage de nouvelle génération. En parallèle, un code-barres constitué exclusivement de mutations silencieuses est utilisé comme contrôle interne pour les effets non-spécifiques potentiels qui peuvent être engendrés suite au clivage hors-cible par le système CRISPR/Cas9. Cette approche, que nous avons appelée CRISPR-barcoding, permet de générer et de suivre l'émergence d'un petit groupe de cellules cancéreuses contenant une mutation voulue au sein d'une population de cellules non modifiées, représentant ainsi un nouveau modèle expérimental d'hétérogénéité génétique intratumorale. Grâce à une série de preuves de concept, nous avons montré que CRISPR-barcoding est une nouvelle approche qui permet d'étudier de façon simple et rapide les conséquences fonctionnelles de différents types de modifications génétiques apportées directement au niveau de la séquence génomique.

Dans la deuxième partie de ma thèse, nous avons utilisé cette nouvelle approche pour l'étude de la résistance du cancer bronchique non à petites cellules (CBNPC) à la thérapie ciblée. Les patients de CBNPC dont la tumeur présente une mutation activatrice de l'*epidermal growth factor receptor* (EGFR) sont généralement traités avec des inhibiteurs de ce récepteur. Malheureusement, malgré une réponse initiale, presque invariablement ces patients rechutent, suite au développement d'une résistance causée, dans la majorité des cas, par l'apparition de mutations secondaires ou tertiaire de l'EGFR. Grâce à un criblage réalisé en utilisant un modèle de CBNPC basé sur la stratégie CRISPR-barcoding, nous avons pu identifier l'inhibiteur multikinase sorafénib pour sa capacité à prévenir la résistance de ces cellules aux inhibiteurs d'EGFR. Ce composé présente un mécanisme d'action original, impliquant une diminution

précoce du niveau de phosphorylation de STAT3, suivie par une baisse considérable de l'expression de l'EGFR, aboutissant à une inhibition des voies intracellulaires en aval de ce récepteur, telles que RAS-MAPK et PI3K-AKT-mTOR. Ces données ont été confirmées *in vivo* en utilisant un modèle de xéno greffe de cellules de CBNPC modifiées par CRISPR-barcoding.

En conclusion, l'ensemble de nos résultats montre que le sorafenib peut prévenir l'émergence de cellules de CBNPC résistantes aux inhibiteurs d'EGFR, indiquant que ce composé pourrait représenter une nouvelle stratégie thérapeutique pour le traitement de ce type de tumeur.

Mots clés : CRISPR-barcoding - Hétérogénéité intratumorale - Cancer bronchique non à petites cellules - Résistance à la thérapie ciblée – EGFR - Sorafenib.

ABSTRACT

Functional analysis of oncogenic driver mutations through CRISPR-barcoding in a context of intratumoral heterogeneity

Individual tumors are composed of multiple and genetically distinct subpopulations of transformed cells that can adapt and evolve in a different way based on environmental conditions. This genetic diversity has major consequences for the patient, particularly during tumor progression and for cancer treatment.

We devised a new strategy based on CRISPR/Cas9 technology in which a potentially functional modification in the sequence of a gene of interest is coupled with a series of silent point mutations, functioning as a genetic label for cell tracing. In parallel, a second barcode consisting of distinct silent mutations is inserted in the same cell population and used as a control for CRISPR/Cas9 off-target cleavage. This approach, that we named CRISPR-barcoding, enables detection of cells containing the mutation of interest within a mass-population of unmodified cells using real-time quantitative PCR or deep sequencing. Through a series of proof-of-concept studies, we demonstrated that CRISPR-barcoding is a fast and highly flexible strategy to investigate the functional consequences of a specific genetic modification in a broad range of assays.

In the second part of my thesis, we used CRISPR-barcoding to investigate non-small cell lung cancer (NSCLC) resistance to targeted therapy. Some NSCLCs harbor activating mutations of the epidermal growth factor receptor (EGFR) and are addicted to this signaling pathway. These tumors initially show a good response to EGFR inhibitors (EGFRi), but they almost invariably relapse, due to the acquisition of a resistance, as a result of additional genetic alterations, including secondary and tertiary EGFR mutations. Using a CRISPR-barcoding model, we identified the multikinase inhibitor sorafenib for its ability to prevent EGFRi resistance in NSCLC cells. This compound acts through an original mechanism that involves early reduction of STAT3 phosphorylation and late down-regulation of EGFR, resulting in the inhibition of different downstream pathways activated by this receptor, including, RAS-MAPK and PI3K-AKT-mTOR. These results were confirmed *in vivo*, using a CRISPR-barcoding xenograft model for NSCLC.

Altogether, our data indicate that sorafenib can prevent NSCLC resistance to EGFRi through a novel mechanism, thus providing a new potential therapeutic strategy for the treatment of this type of cancer.

Key words: CRISPR-barcoding – Intratumor heterogeneity – Non-small cell lung cancers – Resistance to targeted therapy – EGFR – Sorafenib.

SOMMAIRE

LISTE DES FIGURES ET DES TABLES.....	23
LISTE DES ABREVIATIONS	25
INTRODUCTION.....	29
1. Hétérogénéité intratumorale	31
1.1. Hétérogénéité tumorale non-génétique.....	31
1.2. Hétérogénéité tumorale génétique	32
1.3. Conséquences de l'hétérogénéité génétique intratumorale	34
1.3.1. Progression tumorale	34
1.3.2. Formation de métastases.....	35
1.3.3. Résistance aux thérapies	36
1.4. Détection et modélisation de l'hétérogénéité intratumorale.....	37
2. CRISPR : une nouvelle technologie d'édition de l'ADN	38
2.1. Découverte de CRISPR	38
2.2. Description du système CRISPR/Cas de type II	39
2.2.1. Phase 1 : Acquisition des <i>spacers</i>	41
2.2.2. Phase 2 : Biogenèse des crRNA	42
2.2.3. Phase 3 : Interférence avec la séquence cible	43
2.3. CRISPR comme outil pour l'édition du génome.....	44
2.3.1. Mécanismes endogènes de réparation de l'ADN.....	45
2.3.2. Technologies d'édition de l'ADN précédant la révolution CRISPR/Cas9.....	46
2.3.2.1. <i>Méganucléases</i>	46
2.3.2.2. <i>Zinc-Finger Nucléases</i>	47
2.3.2.3. <i>Transcription Activator-Like Effector Nucleases</i>	47
2.3.3. Développement de CRISPR/Cas9 en tant qu'outil pour l'édition du génome	48
2.3.4. Autres applications du système CRISPR/Cas9.....	49
2.3.4.1. <i>Remaniements chromosomiques</i>	50
2.3.4.2. <i>Régulation transcriptionnelle</i>	50
2.3.4.3. <i>Modulation épigénétique</i>	53
2.3.4.4. <i>Edition de base par la cytidine désaminase</i>	53
2.3.4.5. <i>Criblage génétique fonctionnel</i>	53
2.3.4.6. <i>Imagerie de région précise du génome</i>	54
2.3.5. Stratégies pour l'introduction du système CRISPR/Cas9 dans les cellules.....	56
2.4. Utilisation de CRISPR pour l'étude de pathologies.....	58

2.5. Limitations de la technologie CRISPR/Cas9.....	59
2.5.1. Effets non-spécifiques de CRISPR/Cas9.....	59
2.5.2. Faible efficacité de réparation par l'HDR.....	61
2.6. Enjeux agro-alimentaires, biomédicaux et éthiques de CRISPR/Cas9	62
3. Cancer bronchique.....	64
3.1. Épidémiologie et facteurs de risque.....	64
3.2. Classification des cancers bronchiques	64
3.2.1. Cancers bronchiques à petites cellules.....	64
3.2.2. Cancers bronchiques non à petites cellules.....	65
3.2.2.1. Mutations de <i>KRAS</i>	66
3.2.2.2. Réarrangements chromosomiques de <i>ALK</i>	69
3.2.2.3. Mutations de <i>PTEN</i>	70
3.2.2.4. Amplification de <i>FGFR1</i>	71
3.2.2.5. Altérations de <i>MET</i>	72
3.2.2.6. Mutations d' <i>EGFR</i>	73
3.3. Thérapie ciblée pour les CBNPCs avec mutation d'EGFR.....	74
3.3.1. Inhibiteurs de 1 ^{ère} génération	74
3.3.2. Résistances aux inhibiteurs d'EGFR de 1 ^{ère} génération	75
3.3.2.1. Résistance primaire.....	75
3.3.2.2. Résistance acquise	76
3.3.3. Nouveaux inhibiteurs capables de cibler la mutation T790M	78
3.3.4. Développement d'une résistance aux inhibiteurs d'EGFR de 3 ^{ème} génération.....	79
Objectifs de la thèse.....	81
RESULTATS	83
1. Travaux publiés	85
CRISPR-barcoding for Intratumor Genetic Heterogeneity Modeling and Functional Analysis of Oncogenic Driver Mutations.....	85
Modeling intratumor heterogeneity through CRISPR-barcodes	121
CRISPR/Cas9 editing of the genome for cancer modeling	127
2. Données non publiées : Nouvelle stratégie thérapeutique pour prévenir la résistance aux inhibiteurs d'EGFR dans le cancer bronchique.....	137
2.1. Introduction	138
2.2. Matériels et méthodes.....	139
2.3. Identification d'un nouveau composé capable de prévenir la résistance au gefitinib des cellules de CBNPC présentant la mutation EGFR-T790M.....	144
2.4. Le sorafenib prévient la résistance des cellules de CBNPC aux inhibiteurs d'EGFR de 3 ^{ème} génération	145
2.5. Mécanisme d'action du sorafenib dans les cellules de CBNPC.....	146

2.6. Le sorafenib prévient la résistance aux inhibiteurs d'EGFR <i>in vivo</i>	149
DISCUSSION	151
1. CRISPR-barcoding	153
1.1. Avantages et limitations de l'approche CRISPR-barcoding	153
1.1.1. Avantages de CRISPR-barcoding.....	153
1.1.2. Limitations de CRISPR-barcoding	153
1.2. Méthodes de détection des codes-barres	156
1.2.1. PCR quantitative	156
1.2.2. Autres technologies de PCR	156
1.2.3. Séquençage à haut débit.....	157
1.3. Utilisation de codes-barres complexes pour tracer différents clones au sein d'une population cellulaire.....	158
1.3.1. Approches antérieures.....	158
1.3.2. Introduction de codes-barres complexes via CRISPR-barcoding.....	158
2. Sorafenib et CBNPC	159
2.1. Mécanisme d'action du sorafenib dans les cellules de CBNPC.....	159
2.2. Autres approches pour contrecarrer la résistance des CBNPCs aux inhibiteurs d'EGFR.....	162
2.3. Le sorafenib pour le traitement du CBNPC avec EGFR muté.....	163
CONCLUSION ET PERSPECTIVES	165
REFERENCES BIBLIOGRAPHIQUES.....	169

LISTE DES FIGURES ET DES TABLES

Figure 1. Modèle des cellules souches cancéreuses (CSC).....	31
Figure 2. Arbre phylogénétique de l'évolution tumorale	33
Figure 3. Evolution tumorale	34
Figure 4. Différents types d'hétérogénéité tumorale.....	35
Figure 5. Modèle de résistance aux thérapies anti-cancéreuses	36
Figure 6. Organisation des systèmes CRISPR/Cas de type II.....	39
Figure 7. Fonctionnement du système CRISPR/Cas chez la bactérie.....	40
Figure 8. Acquisition des séquences <i>spacers</i>	41
Figure 9. Processus de biogenèse des crRNA	43
Figure 10. Interférence du système CRISPR/Cas de type II avec l'ADN.....	44
Figure 11. Mécanismes de réparation de l'ADN.....	45
Figure 12. Autres stratégies d'édition du génome.....	46
Figure 13. Remaniements chromosomiques engendrés en utilisant deux sgRNAs	50
Figure 14. Utilisation de formes mutés de Cas9	51
Figure 15. Exemples d'applications de dCas9	52
Figure 16. Approches CRISPR/Cas9 pour visualiser une région spécifique du génome.....	55
Figure 17. Utilisation de deux sgRNAs et une Cas9 nickase pour augmenter la spécificité du clivage de l'ADN	61
Figure 18. Fréquence des mutations dont l'amplification a été démontrée dans la formation et/ou la progression des CBNPCs	66
Figure 19. Voies de signalisation activées par les récepteurs tyrosine kinases (RTKs).....	67
Figure 20. Mécanismes de résistance aux inhibiteurs d'EGFR dans le CBNPC.....	75
Figure 21. Principe de CRISPR-barcoding	141
Figure 22. Effets du sorafenib dans les cellules PC9 de CBNPC résistantes aux inhibiteurs d'EGFR	144
Figure 23. Effets du sorafenib sur la résistance de différentes lignées cellulaires de CBNPC aux inhibiteurs d'EGFR	145
Figure 24. Le sorafenib prévient la résistance des CBNPC à l'osimertinib induite par la mutation C797S.....	146
Figure 25. Effets du sorafenib dans les cellules PC9	146
Figure 26. Effets du sorafenib dans les cellules HCC827 et HCC4006 dérivées de CBNPC...	147

Figure 27. Le sunitinib ne prévient pas la résistance des cellules de CBNPC aux inhibiteurs d'EGFR	148
Figure 28. L'inhibition de CRAF n'a pas d'effet sur la phosphorylation de STAT3 et la croissance des cellules de CBNPC	148
Figure 29. Effets du sorafenib sur la croissance tumorale	149
Figure 30. Réparation de la mutation oncogénique de <i>ALK</i> dans les cellules Kelly de neuroblastome.....	155
Figure 31. Comparaison des effets du sorafenib et du SC-1 sur la résistance aux inhibiteurs d'EGFR dans les cellules de CBNPC	160
Tableau 1. Séquences des sgRNA et ssODN	141
Tableau 2. Séquences des amorces de qPCR	142
Tableau 3. Profils pharmacologiques des inhibiteurs multikinases sorafenib, regorafenib et sunitinib	147

LISTE DES ABREVIATIONS

5hmC	5-hydroxyméthylcytosine
AAV	virus adeno-associé
ADNg	ADN génomique
ALK	<i>anaplastic lymphoma kinase</i>
ARNm	ARN messenger
cas	<i>CRISPR-associated</i>
Cas9n	<i>Cas9 nickase</i>
Cascade	<i>CRISPR-associated complex for antiviral defense</i>
CBNPC	cancers bronchiques non à petites cellules
CBPC	cancers bronchiques à petites cellules
CPP	<i>cell-penetrating peptides</i>
CRISPR	<i>Clustered Regularly Interspaced Short Palindromic Repeat</i>
crRNA	<i>CRISPR RNA</i>
CS	chromoshadow
CSC	cellules souches cancéreuses
Ct	<i>threshold cycle</i>
dCas9	<i>dead Cas9</i> ou Cas9 catalytiquement inactive
EGFP	<i>enhanced green fluorescent protein</i>
EGFR	<i>epidermal growth factor receptor</i>
ERK	<i>extracellular signal-regulated kinase</i>
FGFR1	<i>fibroblast growth factor receptor 1</i>
FISH	<i>fluorescent in situ hybridization</i>
Foxo	<i>Forkhead box</i>
GAB1	<i>Grb2-associated-binding protein 1</i>
GAP	<i>GTPase-activating protein</i>
GUIDE-seq	<i>genome-wide unbiased identification of double strand breaks enabled by sequencing</i>
Grb2	<i>growth factor receptor-bound protein 2</i>

GSK3	<i>glycogen synthase kinase 3</i>
HDR	<i>homology directed repair</i>
HGF	<i>hepatocyte growth factor</i>
IGF-1R	<i>insulin-like growth factor 1</i>
INCa	Institut National du Cancer
Indels	insertion ou délétion de quelques nucléotides
JAK	<i>janus kinase</i>
KRAB	<i>krüppel associated box</i>
KRAS	<i>Kirsten rat sarcoma 2 viral oncogene homolog</i>
MAPK	<i>Mitogen-Activated Protein Kinase</i>
MCP	<i>MS2 coat protein</i>
MDM2	<i>mouse double minute 2</i>
MEK	<i>mitogen activate protein kinase kinase</i>
mTORC1/2	<i>mammalian target of rapamycin complex 1/2</i>
NF-κB	<i>nuclear factor kappa B</i>
NGS	<i>next-generation sequencing</i>
NHEJ	<i>non-homologous end joining</i>
OGM	organisme génétiquement modifié
PAM	<i>protospacer adjacent motif</i>
PARP	poly (ADP-ribose) polymerase
PBS	<i>phosphate buffered saline</i>
PCP	<i>PP7 coat protein</i>
PD-1	<i>programmed cell death 1</i>
PD-L1	<i>programmed death ligand 1</i>
PDGFR	<i>platelet-derived growth factor receptor</i>
PEI	polyéthylèneimine
PI3K	phosphoinositide-3-kinase
PIP ₂	phosphatidylinositol (4,5)-biphosphate
PIP ₃	phosphatidylinositol (3,4,5)-triphosphate
PLC _γ	phospholipase C gamma

PTB	<i>phosphotyrosine binding</i>
PTEN	<i>Phosphatase and TENsin homolog</i>
qPCR	PCR quantitative
RB1	<i>RB transcriptional corepressor 1</i>
REC	<i>recognition</i>
RFLP	<i>restriction fragment length polymorphism</i>
Rheb	<i>RAS homolog enriched in the brain</i>
Rock	<i>rho associated coiled-coil containing protein kinase</i>
RTK	récepteur tyrosine kinase
SDS	sodium dodécyl sulfate
sgRNA	ARN guide ou <i>single-guide RNA</i>
SH2	<i>Src homology 2</i>
SHP1/2	<i>Src homology 2 domain-containing phosphatase 1/2</i>
shRNA	<i>short hairpin RNA</i>
SOS	<i>son of sevenless</i>
SpCas9	Cas9 de <i>Streptococcus pyogenes</i>
ssODN	<i>single stranded oligodeoxyribonucleotide</i>
STAT	<i>Signal Transducers and Activators of Transcription</i>
TALE	<i>TAL effector</i>
TALEN	<i>Transcription Activator-Like Effectors</i>
TGF- α/β	<i>transforming growth factor-α / β</i>
TP53	<i>Tumor Protein 53</i>
tracrRNA	<i>trans-activating CRISPR RNA</i>
TSC1/2	<i>tuberous sclerosis 1</i>
VEGFR	<i>vascular endothelial growth factor receptor</i>
ZFN	<i>Zinc-Finger Nuclease</i>

INTRODUCTION

1. Hétérogénéité intratumorale

La première description d'une hétérogénéité au sein d'une même tumeur remonte au XIX^{ème} siècle par l'anatomo-pathologiste Rudolf Virchow (Virchow et Chance, 1860). Les tumeurs ne sont pas des entités homogènes, mais elles sont composées de cellules pouvant présenter de profondes différences, non seulement morphologiques, mais également fonctionnelles, telles que des variations de leur capacité proliférative et invasive, ou de leur activité métabolique, ce qui peut avoir des conséquences importantes en clinique (Bloom et Richardson, 1957 ; Heppner et Miller, 1983). Cette diversité provient de deux traits caractéristiques des tumeurs : d'une part la capacité des cellules néoplasiques à se différencier ou se dé-différencier de façon spontanée ou en réponse aux conditions environnementales, et d'autre part l'instabilité génétique, qui provoque l'accumulation de clones au profil mutationnel distinct.

1.1. Hétérogénéité tumorale non-génétique

L'hétérogénéité tumorale non-génétique s'explique par l'implication de mécanismes épigénétiques régulés en réponse aux interactions que les cellules tumorales entretiennent entre elles et avec leur microenvironnement, se traduisant par des modifications des niveaux d'expression de différents gènes. Un exemple typique est celui des cellules souches cancéreuses (CSC), caractérisées par leur capacité d'auto-renouvellement et de différenciation. Selon certains auteurs, les CSC seraient les seules capables de générer toutes les autres populations

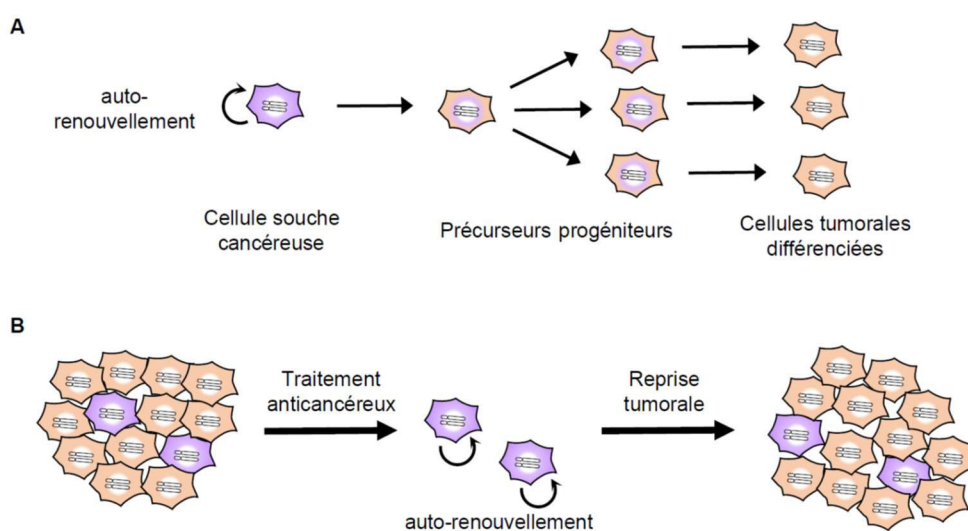


Figure 1. Modèle des cellules souches cancéreuses (CSC)

A. Le développement hiérarchique des CSC. Les CSC (violette) se différencient en cellules progénitrices qui peuvent ensuite se transformer en cellules tumorales différenciées. Ces cellules cancéreuses différenciées ont perdu leur capacité à s'autorenouveler et à reconstituer d'elles-même une tumeur. **B.** Les CSC sont plus résistantes que les cellules tumorales différenciées (orange) aux thérapies anticancéreuses conventionnelles. Après traitement, les CSC produisent de nouvelles cellules tumorales différenciées et relancent ainsi rapidement la croissance de la tumeur.

cellulaires au sein de la tumeur, selon un modèle hiérarchique qui rappelle le développement normal de la plupart des organes et tissus (**Fig. 1A**) (Meacham et Morrison, 2013 ; Michor et Polyak, 2010). Comme les cellules souches normales, les CSC présenteraient également une résistance accrue aux thérapies, leur permettant ainsi de régénérer la tumeur à la suite d'un traitement anticancéreux (**Fig. 1B**) (Meacham et Morrison, 2013).

Le concept de CSC a émergé dans les années 90 grâce à des expériences de xénogreffes de cellules issues d'hémopathies malignes chez des souris immunodéficientes. Ces travaux révélèrent qu'une sous-population phénotypiquement distincte de cellules humaines de leucémies myéloïdes aiguës pouvait reconstituer la pathologie après transplantation, alors que les autres cellules leucémiques ne possédaient pas cette capacité (Lapidot *et al.*, 1994). A la suite de ces travaux, différentes études ont démontré l'existence, au moins chez la souris, de sous-populations de CSC dans différents types de cancers solides, tels que les cancers du sein (Al-Hajj *et al.*, 2003), du colon (Dalerba *et al.*, 2007 ; O'Brien *et al.*, 2007), du pancréas (Li *et al.*, 2007) ou encore les glioblastomes (Singh *et al.*, 2004). Malgré le débat toujours ouvert sur les implications cliniques réelles des CSCs, ce modèle a permis une meilleure compréhension de la plasticité tumorale, ainsi que le développement de toute une série de marqueurs moléculaires pour l'identification et la purification de différentes sous-populations de cellules cancéreuses.

1.2. Hétérogénéité tumorale génétique

Dans un article paru dans *Science* en 1976, Peter Nowell énonça le principe d'évolution clonale des cellules tumorales, selon lequel le cancer est le résultat d'une succession de mutations au sein d'une même cellule, ce qui permet progressivement l'acquisition des caractéristiques biologiques propres du phénotype néoplasique (Nowell, 1976). L'apparition de mutations est un phénomène stochastique (Tomasetti *et al.*, 2017), et le plus souvent ces aberrations génétiques ne confèrent aucun avantage à la cellule cancéreuse (*passenger mutations*). En revanche, dans certains cas on peut assister à des altérations qui constituent un bénéfice sélectif, à l'instar de l'évolution darwinienne des espèces, ce qui permet l'émergence de nouvelles sous-populations clonales qui contiennent la mutation (*driver mutations*) (Kandoth *et al.*, 2013 ; Lawrence *et al.*, 2014 ; Yates et Campbell, 2012). Puisque les cellules cancéreuses présentent souvent un niveau d'instabilité génétique élevé dû à des altérations des processus de maintenance du génome et des systèmes de surveillance (Hanahan et Weinberg, 2011), certains types de tumeurs peuvent accumuler des centaines, voire des milliers, de mutations. En fonction de l'histoire évolutive de la tumeur, ces mutations sont présentes dans la plupart des cellules ou

alors distribuées de façon plus complexe, se traduisant par un degré d'hétérogénéité intratumorale plus ou moins important.

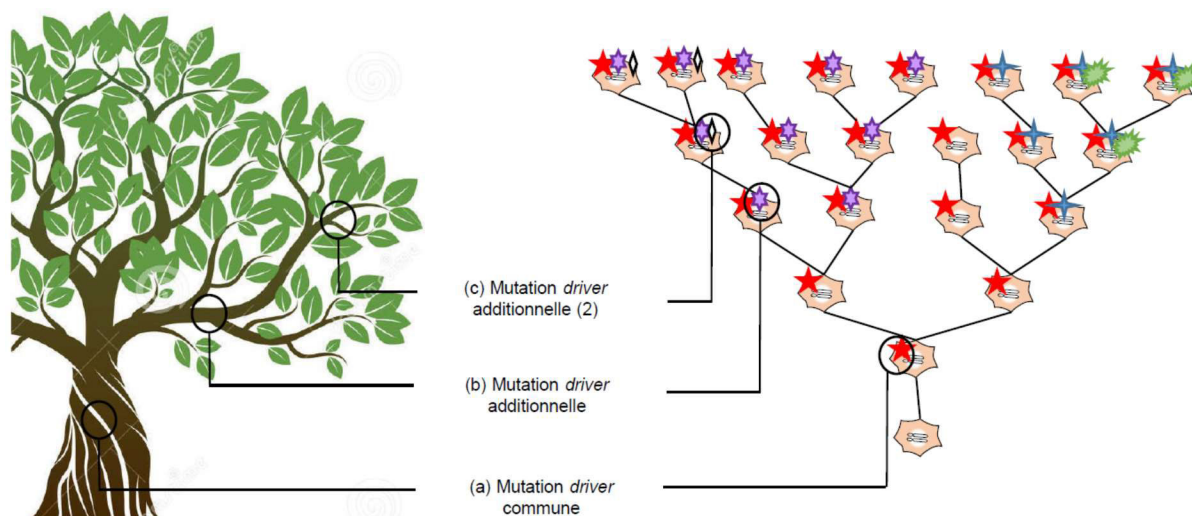


Figure 2. Arbre phylogénétique de l'évolution tumorale

Les tumeurs primaires sont composées de différents sous-clones contenant des mutations communes et individuelles. Les altérations génétiques *driver* partagées par toutes les cellules cancéreuses (a) apparaissent précocement dans le développement tumoral et représentent le tronc de l'arbre phylogénétique. Différentes mutations *driver* additionnelles, retrouvées dans les cellules de certaines régions de la tumeur (b), se développent plus tardivement et correspondent ainsi aux différentes branches proximales de l'arbre. Enfin, des mutations *driver* plus individuelles sont contenues dans seulement une zone très restreinte de la tumeur (c) et apparaissent encore plus tardivement. Ces dernières mutations sont comparées aux ramifications des branches.

Cette diversité génétique est fréquemment représentée sous la forme d'un arbre (**Fig. 2**) (Yap *et al.*, 2012) : les mutations *driver* communes à toutes les cellules tumorales constituent le tronc, à partir duquel se forment différentes branches et ramifications correspondant aux autres mutations, qu'on retrouve seulement dans certaines sous-populations clonales. Les différents sous-clones coexistent au sein d'une tumeur, et leur taille varie en fonction de la progression et de la pression sélective imposée par l'environnement. Selon le type de tumeur, trois formes de clonalité ont été décrites. D'une part, les cancers clonaux simples présentent une acquisition séquentielle de mutations qui entraîne une dominance successive de chaque nouvelle population, remplaçant ainsi la précédente (Hanahan et Weinberg, 2011 ; Vogelstein et Kinzler, 1993). Ces cancers se caractérisent par une faible instabilité génétique, se traduisant par une hétérogénéité tumorale limitée (**Fig. 3A**) (McGranahan et Swanton, 2015). D'autre part, les cancers clonaux complexes sont constitués de sous-populations de cellules cancéreuses contenant non seulement les mêmes mutations *driver* fondatrices, mais également des mutations *passenger* différentes, qui ne confèrent pas de pression de sélection, mais qui sont potentiellement capables d'instaurer une résistance aux thérapies anticancéreuses (Heinrich *et al.*, 2008 ; Negri *et al.*, 2009). Enfin, les cancers clonaux mosaïques sont composés de plusieurs sous-populations caractérisées par des mutations *driver* distinctes, même si toutes les cellules partagent un certain nombre d'altérations communes présentes dès les phases précoces de la

transformation tumorale (**Fig. 3B**). Une hétérogénéité génétique intratumorale a été décrite dans la plupart des cancers solides (Bashashati *et al.*, 2013 ; de Bruin *et al.*, 2014 ; Campbell *et al.*, 2010 ; Carreira *et al.*, 2014 ; Gerlinger *et al.*, 2014 ; Haffner *et al.*, 2013 ; Navin *et al.*, 2011 ; Shah *et al.*, 2012 ; Snuderl *et al.*, 2011 ; Sottoriva *et al.*, 2013 ; Thirlwell *et al.*, 2010 ; Yachida *et al.*, 2010 ; Zhang *et al.*, 2014), ainsi que dans des hémopathies malignes (Landau *et al.*, 2013).

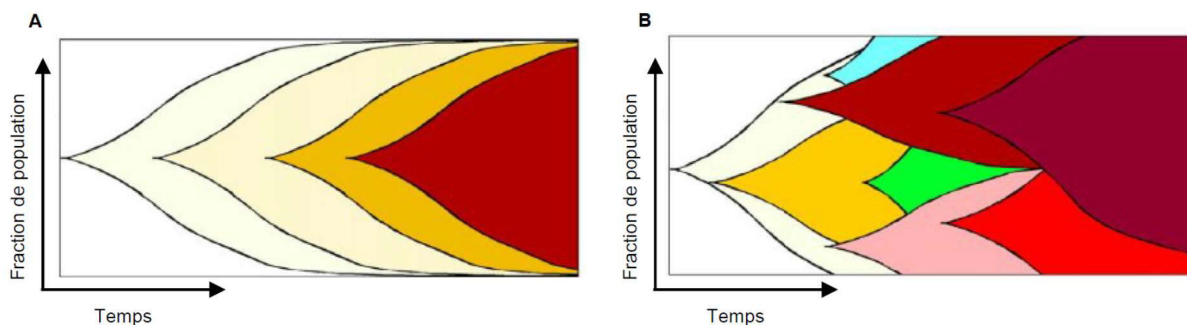


Figure 3. Evolution tumorale

A. Développement au cours du temps d'une tumeur selon un modèle linéaire. Les cellules accumulent successivement des mutations qui sont préférentiellement sélectionnées par l'environnement tumoral. Chaque génération de cellules va ainsi remplacer la précédente. **B.** Les tumeurs sont composées de différents sous-clones contenant des mutations *passenger* ou *driver* distinctes menant à une dominance de certains sous-clones au cours du temps.

Figure issue de Marusyk et Polyak, 2010.

1.3. Conséquences de l'hétérogénéité génétique intratumorale

Depuis maintenant quelques années, différentes équipes à travers le monde s'intéressent au lien entre l'hétérogénéité d'une tumeur, son devenir et la signification clinique des mutations *driver* retrouvées dans différentes sous-populations clonales.

1.3.1. Progression tumorale

Il a été montré dans les leucémies lymphocytaires chroniques que la présence de différents sous-clones est associée à une rechute de la tumeur beaucoup plus rapide (Landau *et al.*, 2013). Ces données, ainsi que celles d'une autre étude basée sur le séquençage de différentes régions de cancers bronchiques non à petites cellules (CBNPC) de stade précoce (Zhang *et al.*, 2014), indiquent que l'hétérogénéité génétique intratumorale favorise la progression tumorale et qu'elle est associée à un pronostic plus sombre. En 2013, le Royaume-Uni a lancé un essai clinique réunissant près de 850 patients de CBNPC au stade précoce. Cet essai, appelé TRACERx, pour *Tracking Non-Small Cell Lung Cancer Evolution Through Therapy*, a pour but de mieux comprendre la relation entre hétérogénéité intratumorale et devenir clinique du patient (<http://clinicaltrials.gov/show/NTC01888601>).

1.3.2. Formation de métastases

En plus de favoriser la progression tumorale, l'hétérogénéité génétique peut faciliter la survenue de métastases. Selon le modèle classique d'évolution du cancer, le clone le plus avancé et le plus agressif de la tumeur primaire devient capable de migrer et coloniser une autre région du corps (**Fig. 4**) (Fearon et Vogelstein, 1990). Dans d'autres cas, la dissémination métastatique peut avoir lieu lors des phases précoces de la tumorigenèse, comme il l'a été récemment démontré pour le cancer du sein (Harper *et al.*, 2016 ; Hosseini *et al.*, 2016). En fonction du clone qui s'est échappé, les différentes métastases seront donc génétiquement différentes, ce qui constituera une hétérogénéité inter-métastatique (**Fig. 4B**). De plus, de nouvelles mutations pourront se manifester au sein des métastases, engendrant une hétérogénéité intra-métastatique (**Fig. 4C**) (Vogelstein *et al.*, 2013). Les cellules cancéreuses constituant la tumeur primaire et celles constituant les métastases ne posséderont donc plus le même profil génétique, on parle donc d'hétérogénéité temporelle (**Fig. 4**). Puisque généralement les thérapies sont définies à partir des mutations retrouvées dans les biopsies de tumeurs primaires, ces types d'hétérogénéité peuvent avoir des conséquences majeures pour la prise en charge des patients (Vogelstein *et al.*, 2013).

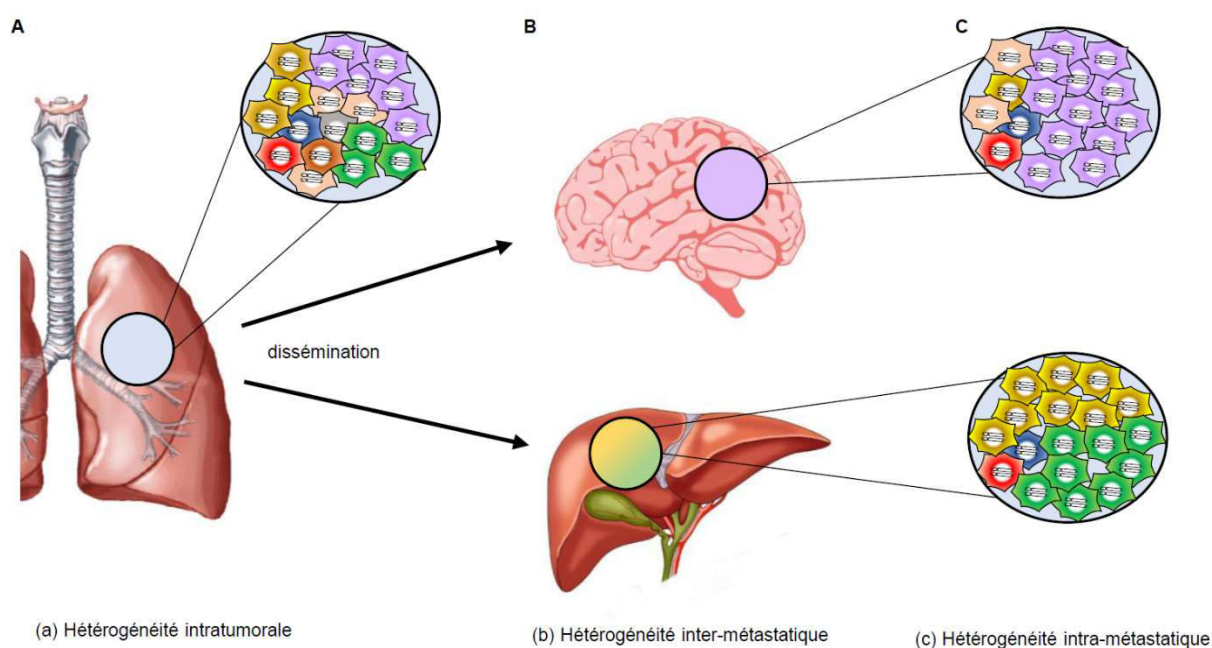


Figure 4. Différents types d'hétérogénéité tumorale

A. L'hétérogénéité intratumorale correspond à la présence de différents sous-clones au sein d'une même tumeur. **B.** Certains clones peuvent s'échapper de la tumeur et aller former des métastases dans d'autres organes du corps. Lorsqu'elles sont formées à partir de sous-clones génétiquement distincts on parle alors d'hétérogénéité inter-métastatique. Certains sous-clones retrouvés dans les métastases peuvent provenir de la tumeur primaire mais d'autres peuvent également avoir émergé suite à l'acquisition de nouvelles altérations génétiques. **C.** L'hétérogénéité intra-métastatique consiste en la présence de multiples sous-clones au sein d'un même foyer métastatique.

1.3.3. Résistance aux thérapies

L'hétérogénéité génétique constitue un réservoir qui alimente la résistance aux traitements anticancéreux, provoquant la rechute de la tumeur. En effet, les différentes sous-populations clonales peuvent présenter des niveaux de sensibilité différents vis-à-vis des agents cytotoxiques utilisés en clinique. Bien que dans un premier temps les patients répondent favorablement aux traitements, la pression sélective qui est exercée par la thérapie peut mener à l'expansion des cellules résistantes déjà présentes dans la tumeur (**Fig. 5A**). Dans d'autres cas, l'acquisition de nouvelles mutations au cours du traitement permet l'émergence de clones qui ne répondent plus au traitement (**Fig. 5B**) (Ding *et al.*, 2012 ; McGranahan et Swanton, 2015).

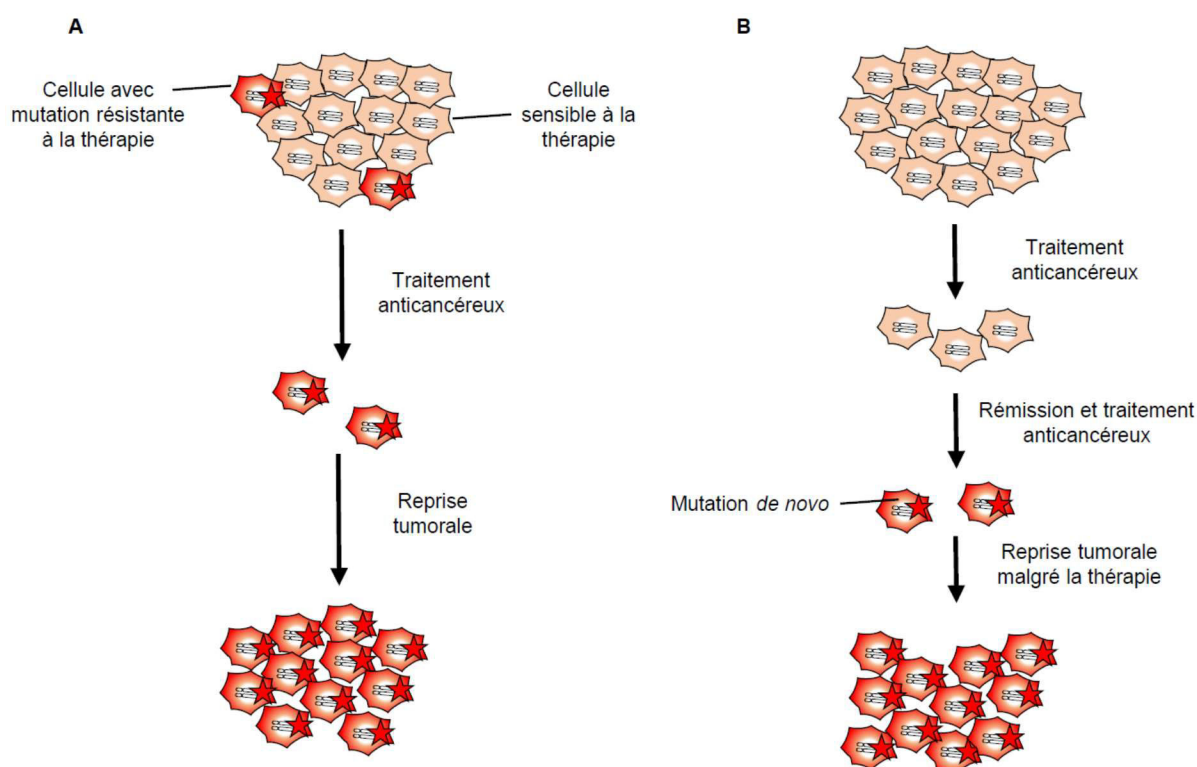


Figure 5. Modèle de résistance aux thérapies anti-cancéreuses

A. Certaines cellules au sein de la tumeur présentent des altérations génétiques qui leur confèrent une résistance envers les thérapies. Ces cellules sont ensuite capables de reformer une tumeur. **B.** Dans un premier temps la stratégie thérapeutique permet de résorber la tumeur. Cependant, après quelques semaines de traitement, des mutations *de novo* peuvent apparaître. Ces mutations sont susceptibles de conférer une résistance à la thérapie et permettre ainsi une résurgence de la tumeur.

Figure adaptée de Meacham et Morrison, 2013.

Ainsi, il a été montré que la résistance aux inhibiteurs de l'*epidermal growth factor receptor* (EGFR) dans le cancer bronchique peut dépendre de la présence d'une petite sous-population de cellules contenant des aberrations génétiques qui les rendent insensibles au traitement, telles que la mutation EGFR-T790M ou l'amplification de l'oncogène MET (Turke *et al.*, 2010 ; Hata *et al.*, 2016). Ces cellules représentent en général une fraction très faible de la tumeur, ce qui explique la réponse initiale, mais au fur et à mesure que leur proportion augmente on assiste à une rechute. Ces études ont également révélé que les tumeurs peuvent également inclure des

cellules moins sensibles au traitement dites « tolérantes », qui constituent une sorte de réservoir pour l'acquisition de nouvelles mutations de résistance (Hata *et al.*, 2016).

L'hétérogénéité des mécanismes de résistance peut varier de façon temporelle, en fonction du protocole thérapeutique, mais également de façon spatiale. Différents facteurs peuvent expliquer la présence de mécanismes de résistance distincts d'un site tumoral à un autre, comme la présence de différents sous-clones génétiquement distincts ou encore une exposition différente aux thérapies. Ce dernier cas est notamment illustré par les métastases dans le cerveau, une zone difficile d'accès pour un bon nombre de composés. Le microenvironnement peut également jouer un rôle important au cours des processus de résistance. Bien que le microenvironnement puisse parfois agir en tant qu'oncosuppresseur, lors de l'établissement d'une tumeur les tissus environnants peuvent être perturbés et altérés, constituant ainsi un stroma tumoral. Contrairement au microenvironnement « normal », les cellules qui forment le stroma tumoral sont détournées pour servir les besoins de la tumeur et contribuer ainsi à sa progression et à sa dissémination, comme par exemple la promotion d'une neoangiogenèse pour alimenter la tumeur, ou encore la sécrétion de facteurs facilitant la dégradation de la matrice extra-cellulaire. En fonction du site d'implantation, le microenvironnement peut varier et influencer ainsi l'évolution tumorale et la réponse aux thérapies (Hanahan et Weinberg, 2011).

1.4. Détection et modélisation de l'hétérogénéité intratumorale

Dans le but d'adapter au mieux la stratégie thérapeutique aux types de mutations présentes dans une tumeur, il est nécessaire de déterminer leur distribution parmi les différentes sous-populations cancéreuses. Traditionnellement, une seule biopsie était pratiquée afin d'établir le profil génétique d'une néoplasie. Bien qu'un prélèvement de plusieurs échantillons à partir d'une tumeur pourrait offrir une meilleure appréciation de la complexité clonale, ce type de stratégie n'est souvent pas envisageable en pratique chez le patient. A l'avenir, les nouvelles approches de biopsie liquide et la modélisation informatique basée sur l'analyse des données de séquençage à haut débit devraient permettre d'établir l'architecture clonale de chaque tumeur et adapter le traitement en conséquence (Altrock *et al.*, 2015 ; Ding *et al.*, 2014 ; Gawad *et al.*, 2016 ; McGranahan et Swanton, 2015). Grâce à ce type d'analyse, il est maintenant possible d'étudier le cancer sous une autre perspective, qui tient compte des aspects évolutifs et écologiques de la progression tumorale. En revanche, ces approches restent descriptives, et dans la plupart des cas la recherche et la validation de nouvelles cibles thérapeutiques sont basées sur des modèles expérimentaux qui ne récapitulent pas l'hétérogénéité intratumorale.

2. CRISPR : une nouvelle technologie d'édition de l'ADN

A l'origine un mécanisme adaptatif de défense immunitaire chez les procaryotes, le système CRISPR/Cas9 a été récemment transformé en un nouvel outil d'édition de l'ADN qui permet de modifier à façon une séquence génomique d'intérêt. Très rapidement, de nombreuses applications basées sur cette technologie ont vu le jour, telles que l'inactivation de gène, l'introduction de mutations ponctuelles, le remaniement chromosomique, ou bien la modulation de l'expression génique.

2.1. Découverte de CRISPR

L'histoire du système CRISPR débute il y a trente ans avec l'identification chez la bactérie *Escherichia coli* de courtes séquences d'ADN palindromiques répétées espacées par des portions d'ADN non-répétées de 20 à 40 nucléotides appelées *spacers* (Ishino *et al.*, 1987 ; Jansen *et al.*, 2002 ; Kunin *et al.*, 2007). Ce n'est pas la première fois que l'on découvre des séquences répétées chez les procaryotes, mais celles-ci, présentes chez plus de 50% des bactéries et près de 90% des archées, ne correspondent en rien à celles retrouvées classiquement et sont regroupées en 2002 sous le terme de CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeat*) (Barrangou et Marraffini, 2014 ; Jansen *et al.*, 2002 ; Makarova *et al.*, 2015 ; Mojica *et al.*, 2000 ; van der Oost *et al.*, 2014). Le locus CRISPR contient également différents gènes associés très conservés entre les procaryotes, appelés *cas* (*CRISPR-associated*), qui codent différentes protéines présentant des domaines nucléase, hélicase ou polymérase (Jansen *et al.*, 2002).

Il faudra attendre presque vingt ans après la publication du papier d'Ishino *et al.* pour que l'on s'aperçoive que les *spacers* dérivent d'éléments génétiques étrangers (*protospacers*), notamment des phages ou des plasmides (Bolotin *et al.*, 2005 ; Mojica *et al.*, 2005 ; Pourcel *et al.*, 2005). Dans leur étude, Mojica et ses collaborateurs remarquent également que chez les archées la présence de *spacers* empêche toute infection par un virus possédant une séquence génétique similaire (Mojica *et al.*, 2005). Quelques années plus tard, l'équipe de Philippe Horvath confirme ces résultats chez les bactéries en démontrant que, lors d'une infection, des fragments d'ADN phagique peuvent s'intégrer au sein du locus CRISPR, entraînant une résistance vis-à-vis d'autres phages présentant ces séquences (Barrangou *et al.*, 2007). Cette étude est rapidement suivie par d'autres qui montrent que le locus CRISPR est transcrit, puis converti en plusieurs courtes portions d'ARN (crRNA : *CRISPR RNA*) contenant chacun un *spacer* unique, qui va guider l'activité de la nucléase Cas et ainsi reconnaître et cliver l'ADN

exogène complémentaire (Brouns *et al.*, 2008 ; Garneau *et al.*, 2010 ; Marraffini et Sontheimer, 2008). L'ensemble de ces observations permet un véritable bond en avant dans la compréhension du système CRISPR et de son rôle en tant que mécanisme adaptatif de défense immunitaire contre les bactériophages.

2.2. Description du système CRISPR/Cas de type II

De nombreuses études ont permis l'identification de différents systèmes CRISPR/Cas, qui ont été répartis en classes, types et sous-types, en fonction de l'organisation du locus et de la machinerie protéique nécessaire pour le clivage de l'ADN exogène (Makarova *et al.*, 2015 ; Shmakov *et al.*, 2015). Eu égard de sa simplicité, c'est le système CRISPR/Cas de type II qui a été choisi pour les applications biotechnologiques d'édition de l'ADN. Dans ce chapitre, nous allons décrire en détail son fonctionnement chez la bactérie.

En plus des gènes *cas1* et *cas2*, communs à tous les types, le locus CRISPR de type II contient le gène signature *cas9* codant l'endonucléase du même nom, qui est impliquée dans l'acquisition des *spacers*, la biogenèse des crRNA et le clivage de l'ADN cible (**Fig. 6**) (van der Oost *et al.*, 2014). L'étude de la structure cristallographique de la protéine Cas9, véritable pièce maîtresse du système de type II, a révélé que cette protéine est constituée de deux lobes distincts. Le premier contient un domaine REC (*recognition*), qui correspond à une région riche en hélice alpha avec une hélice pontée riche en arginine, permettant d'interagir avec le crRNA (Chylinski *et al.*, 2014 ; Jinek *et al.*, 2014 ; Nishimasu *et al.*, 2014). Le second lobe présente un domaine d'interaction avec la séquence PAM (*protospacer adjacent motif* ; motif particulier contenu dans la séquence cible, qui sera décrit ci-dessous), ainsi que deux domaines nucléases appelés HNH et RuvC, responsables du clivage des deux brins de l'ADN cible (Anders *et al.*,

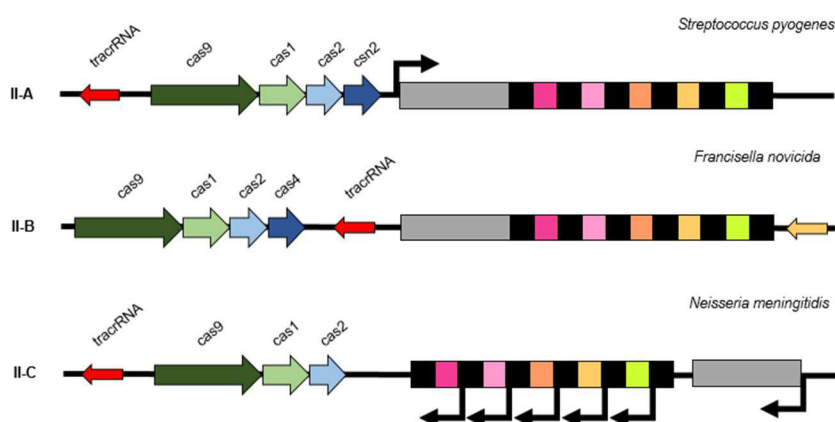


Figure 6. Organisation des systèmes CRISPR/Cas de type II

Les locus CRISPR de type II sont classés en IIA, IIB et IIC sur la base des gènes *cas* présents. Le type IIC contient uniquement les gènes *cas9*, *cas1* et *cas2* ; les types IIA et IIB possèdent un gène *cas* additionnel appelé respectivement *casn2* et *cas4*.

2014 ; Jinek *et al.*, 2014 ; Nishimasu *et al.*, 2014). Chez les différentes espèces, la taille de la protéine Cas9 peut varier, notamment selon la séquence du domaine REC (Jinek *et al.*, 2014). Les sous-types II-A et II-B sont caractérisés par la présence d'autres gènes *cas*, respectivement *csn2* et *cas4* (Fig. 6) (Chylinski *et al.*, 2014 ; Makarova *et al.*, 2015).

Le locus CRISPR contient également une séquence codant un court ARN non traduit, appelé *tracrRNA* (*trans-activating CRISPR RNA*), qui se lie au *crRNA* et agit comme trans-activateur lors du couplage avec l'ADN exogène cible (Deltcheva *et al.*, 2011).

Le fonctionnement du système CRISPR/Cas9 se décompose en trois étapes clés (Fig. 7). La première étape, appelée phase d'adaptation, consiste en la reconnaissance et en l'intégration au sein du locus CRISPR d'une petite portion d'ADN viral en tant que nouveau *spacer*. La seconde phase correspond à la transcription du locus, aboutissant à l'expression des différents composants du système CRISPR/Cas et à la biogenèse de *crRNAs*, contenant chacun un *spacer*

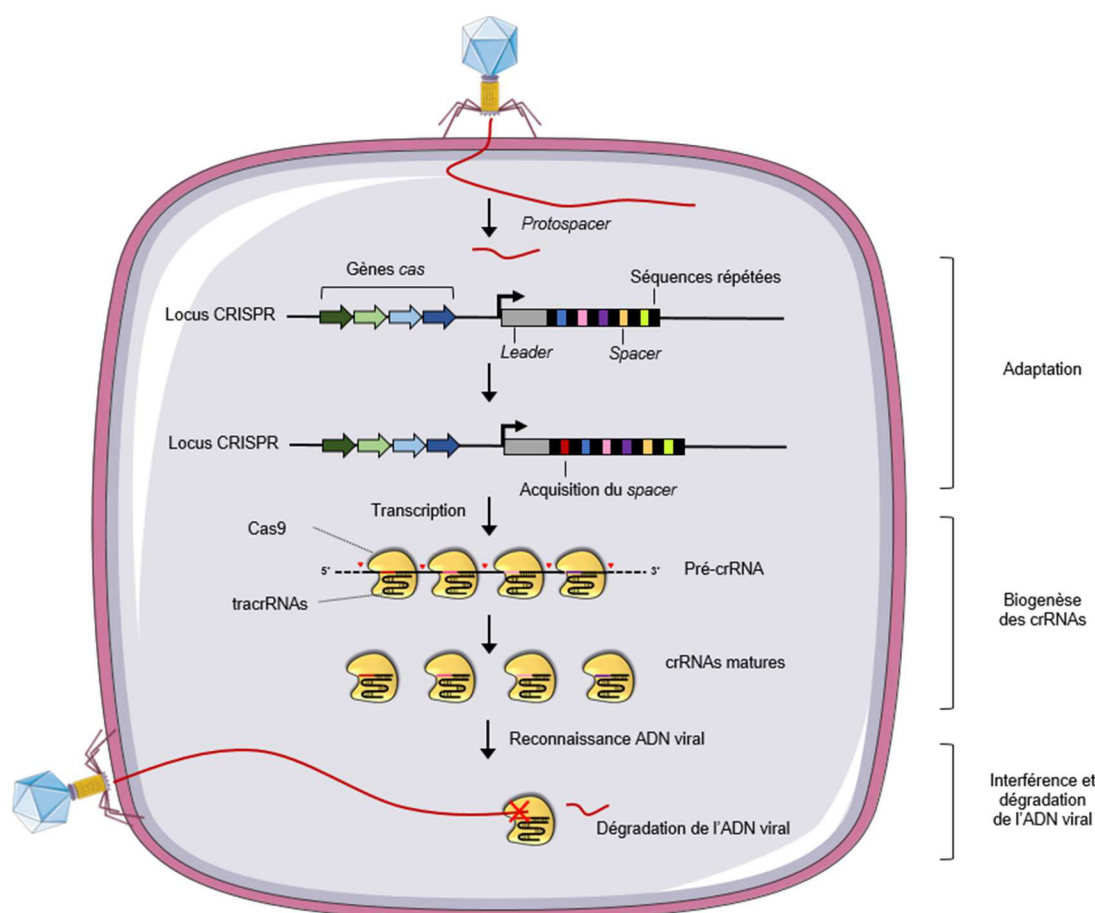


Figure 7. Fonctionnement du système CRISPR/Cas chez la bactérie

Suite à l'introduction dans la cellule d'un ADN étranger issu d'un bactériophage ou d'un plasmide, différentes enzymes Cas vont sélectionner une portion de la séquence étrangère, appelée *protospacer*, et l'introduire en tant que nouvelle séquence *spacer* dans le locus CRISPR de la bactérie juste après la séquence *leader*. Cette phase d'adaptation aboutit à une alternance de *spacers* et de courtes séquences répétées au sein du locus CRISPR. La seconde phase consiste en la biogenèse des *crRNA*, c'est la transcription du locus CRISPR en un long ARN précurseur, le *pre-crRNA*, mûré par la RNase III afin de former des *crRNA* fonctionnels contenant chacun un *spacer* unique. Cette courte séquence d'ARN peut ensuite former un complexe effecteur avec les protéines Cas et reconnaître les éléments génétiques étrangers consécutifs à une infection pour les dégrader.

unique. Enfin, durant la 3^{ème} phase, dite d'interférence, le crRNA conduit la protéine Cas9 vers la séquence du *protospacer* complémentaire au crRNA, pour assurer son clivage grâce à l'activité nucléasique de la Cas9 (van der Oost *et al.*, 2014).

2.2.1. Phase 1 : Acquisition des *spacers*

L'immunité acquise par CRISPR est initiée par la détection, puis l'intégration d'un fragment d'ADN étranger d'environ 30 nucléotides dans le génome de la cellule hôte. La sélection de cette séquence ne se fait pas de façon aléatoire, mais elle dépend de la présence à son extrémité 5' d'un motif de deux à cinq nucléotides appelé PAM (*protospacer adjacent motif*), spécifique de chaque orthologue de Cas9 (Mojica *et al.*, 2009). Deux études menées en parallèle dans des bactéries contenant le système de type II-A (*S. pyogenes* et *S. thermophilus*) montrent que le *protospacer* est reconnu par la protéine Cas9 *via* son domaine d'interaction avec le PAM (Heler *et al.*, 2015 ; Wei *et al.*, 2015b). Cette portion d'ADN étranger est ensuite reconnue par un complexe formé de dimères des deux intégrases Cas1 et Cas2, conservées entre tous les types de CRISPR/Cas (**Fig. 8**) (Heler *et al.*, 2015 ; Nuñez *et al.*, 2014). Une fois lié au *protospacer*, ce complexe va ensuite permettre l'insertion de cette séquence dans le locus CRISPR, encadrée par des portions d'ADN répété et en aval de la séquence promotrice *leader* (**Fig. 8**) (Nuñez *et al.*, 2014, 2015 ; Wang *et al.*, 2015a). Alors que la Cas1 semble avoir un rôle catalytique majeur

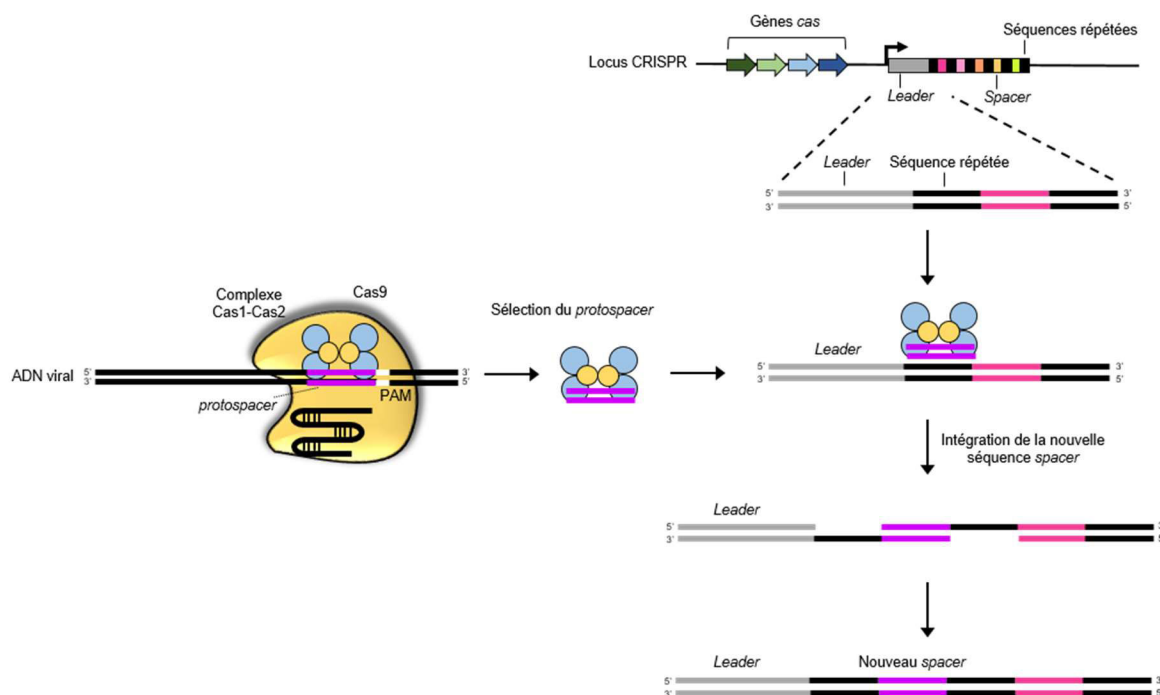


Figure 8. Acquisition des séquences *spacers*

Dans le système CRISPR de type II, la Cas9 reconnaît un motif PAM sur l'ADN étranger et recrute un hétéro-duplex Cas1-Cas2 pour extraire une portion de cet ADN exogène, appelée *protospacer*. Ce *protospacer* est ensuite intégré en tant que nouveau *spacer* à la suite de la 1^{ère} séquence répétée en aval de la séquence *leader* par une réaction de transestérification réalisée par les protéines Cas1-Cas2. Au cours de ce processus, la séquence répétée sera dupliquée, afin de conserver la structure alternant *spacers* et séquences répétées.

dans le mécanisme d'intégration, la Cas2 serait essentielle pour la formation du complexe Cas1-Cas2 (Nuñez *et al.*, 2014, 2015). Enfin, d'autres acteurs peuvent participer au processus d'acquisition, comme la protéine Csn2, qui stabiliserait la coupure double brin de l'ADN au cours de l'intégration, ou encore le tracrRNA, qui formerait un complexe avec la protéine Cas9, mais dont le rôle n'a pas encore été clairement établi (Heler *et al.*, 2015 ; Wei *et al.*, 2015b).

Afin d'éviter toute reconnaissance auto-immune du système CRISPR/Cas, le motif PAM, qui permet la sélection des *protospacers* lors des premières étapes de la phase d'adaptation, n'est pas intégré au locus CRISPR. En effet, puisque le PAM est aussi nécessaire pour la reconnaissance de l'ADN cible par le crRNA, son absence dans les *spacers* évite le clivage du génome bactérien par le système CRISPR/Cas. Ainsi, par l'insertion successive des séquences nucléiques étrangères rencontrées, le locus CRISPR constitue une véritable mémoire immunologique de la bactérie.

2.2.2. Phase 2 : Biogenèse des crRNA

Une fois les *spacers* intégrés dans le locus CRISPR, ils doivent être disponibles pour le système CRISPR/Cas afin de contrecarrer d'éventuelles infections virales. Cette fonction est assurée par l'expression des crRNAs, de courtes séquences d'ARN qui contiennent, entre autre, un *spacer* unique permettant la spécificité d'hybridation avec l'ADN étranger complémentaire. La biogenèse des crRNA varie en fonction des types de CRISPR/Cas. Dans le système de type II-A et II-B, le locus CRISPR est transcrit en un seul et long ARN précurseur, le pre-crRNA. L'hybridation sur cet ARN précurseur de plusieurs tracrRNAs complémentaires aux séquences CRISPR répétées permet le recrutement de protéines Cas9 qui stabilisent chaque interaction pre-crRNA:tracrRNAs (**Fig. 9**) (Deltcheva *et al.*, 2011). Le pre-crRNA est ensuite clivé par la RNase III au niveau des portions répétées, libérant ainsi plusieurs complexes crRNAs:tracrRNAs:Cas9 immatures. Enfin, la séquence répétée et une portion du *spacer* à l'extrémité 5' des crRNAs sont excisées par un mécanisme non encore caractérisé (**Fig. 9**). Cette étape aboutit à la formation de crRNAs matures contenant des *spacers* de 20 nucléotides, qui seront ensuite capables de s'apparier avec les séquences complémentaires d'ADN exogène (Deltcheva *et al.*, 2011 ; Jinek *et al.*, 2012).

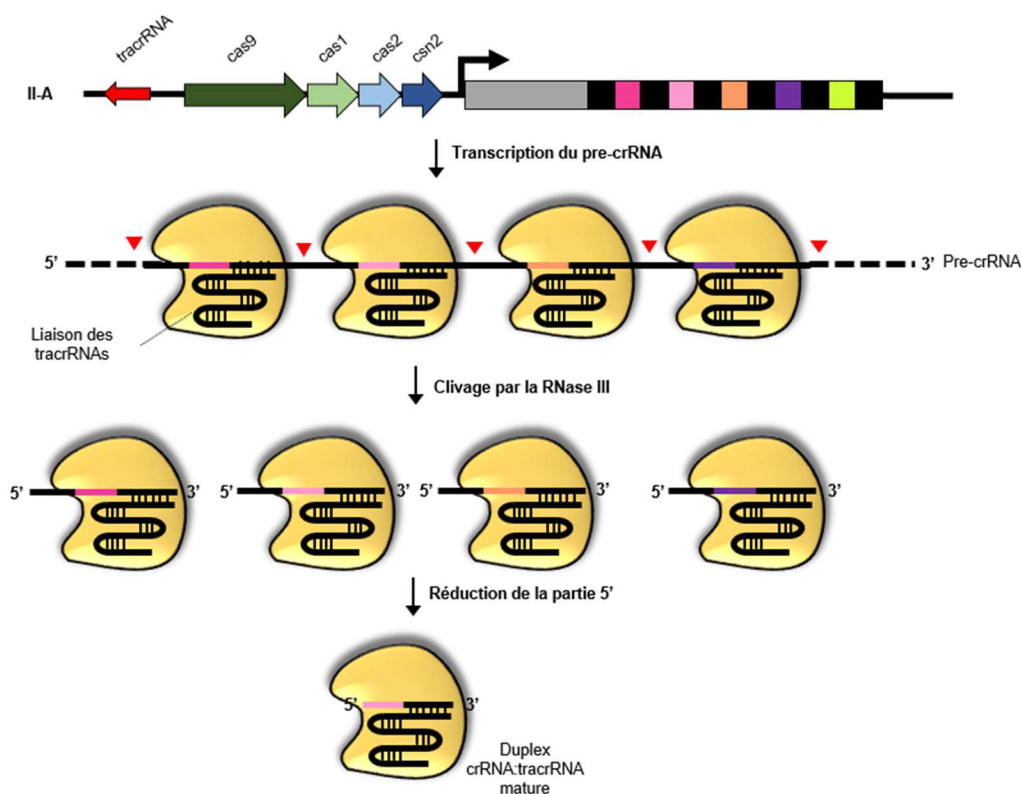


Figure 9. Processus de biogenèse des crRNA

L'expression du locus CRISPR conduit à la synthèse des différentes protéines Cas et à la formation d'un pre-crRNA, un long ARN correspondant à l'ensemble des *spacers* du locus. Un second ARN non-codant, appelé tracrRNA, est également transcrit en parallèle et va s'hybrider avec les séquences répétées présentes sur le pre-crRNA. Le pre-crRNA-tracrRNA est stabilisé par les protéines Cas9, puis clivé par la RNase III pour libérer les complexes crRNA-tracrRNA-Cas9. Enfin ces crRNAs immatures vont être réduits au niveau de leur extrémité 5' par des enzymes non-identifiées à ce jour, résultant en un duplex crRNA-tracrRNA mature, prêt à interférer avec une séquence d'ADN exogène complémentaire à la séquence *spacer*.

Pour le type II-C, au lieu d'un seul long pre-crRNA on en retrouve plusieurs de taille différente, en fonction du site du début de la transcription (Dugar *et al.*, 2013 ; Zhang *et al.*, 2013). Bien qu'une maturation en crRNAs peut avoir lieu grâce à la RNase III, à l'instar des type II-A et II-B, il est possible que ces pre-crRNAs puissent reconnaître leur ADN cible en combinaison avec la Cas9 sans maturation préalable (Zhang *et al.*, 2013).

2.2.3. Phase 3 : Interférence avec la séquence cible

Dans le système CRISPR/Cas de type II, la première étape dans l'interférence avec le *protospacer* consiste en la formation d'un complexe entre le duplex crRNA:tracrRNA et la protéine Cas9. Ce complexe va ensuite balayer l'ADN à la recherche de séquences étrangères complémentaires présentant un motif PAM adjacent (Fonfara *et al.*, 2014 ; Gasiunas *et al.*, 2012 ; Jinek *et al.*, 2012). En l'absence du PAM, la séquence d'ADN ne sera pas reconnue en tant que cible, même si elle est parfaitement complémentaire au crRNA (Sternberg *et al.*, 2014). L'interaction entre Cas9 et PAM provoque une déstabilisation de l'ADN double brin à proximité, facilitant l'appariement du crRNA à la séquence d'ADN étrangère complémentaire

(Jinek *et al.*, 2014 ; Nishimasu *et al.*, 2014), avec la formation d'un hétéroduplex ARN:ADN (Anders *et al.*, 2014, 2014 ; Nishimasu *et al.*, 2014). Cet évènement provoque un changement de conformation de la Cas9 et le clivage des deux brins d'ADN (celui hybridé au crRNA et son complémentaire) par les domaines nucléases HNH et RuvC (**Fig. 10**). Le résultat final est une coupure franche de la séquence cible trois nucléotides en amont du PAM, permettant ainsi la dégradation de l'ADN exogène (**Fig. 10**) (Sternberg *et al.*, 2015). Il a été montré qu'une complémentarité parfaite entre les douze paires de bases jouxtant le PAM et la séquence crRNA est essentielle pour le clivage de l'ADN par la Cas9, alors que des mésappariements peuvent être tolérés sur les huit paires de bases les plus distantes du PAM (Jinek *et al.*, 2012 ; Sternberg *et al.*, 2014).

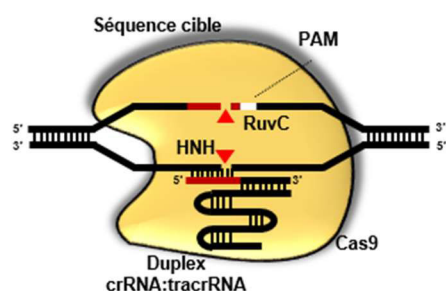


Figure 10. Interférence du système CRISPR/Cas de type II avec l'ADN

En premier lieu, le complexe crRNA-tracrRNA-Cas9 scanne l'ADN à la recherche d'un motif PAM. Une fois trouvé, ce complexe déstabilise l'ADN double brin pour faciliter l'hybridation du crRNA sur la séquence cible complémentaire. Ce processus aboutit au clivage de l'ADN cible trois nucléotides en amont du motif PAM, par les domaines nucléases HNH et RuvC, responsables, respectivement, de la coupure des brins cible et non-cible.

Les autres systèmes CRISPR/Cas présentent des mécanismes plus complexes, impliquant non pas une seule protéine comme dans le type II, mais au contraire un complexe effecteur appelé Cascade (*CRISPR-associated complex for antiviral defense*), généralement composé de plusieurs sous-unités (Makarova *et al.*, 2015). Par souci de concision, ces systèmes ne seront pas décrits en détail dans ce manuscrit.

2.3. CRISPR comme outil pour l'édition du génome

Grâce au développement d'approches basées sur la recombinaison homologe, il a été possible pour la première fois de modifier à façon l'ADN génomique. Malgré le potentiel extraordinaire, l'utilisation de ces stratégies a été limitée par leur faible taux d'efficacité, variable entre $1/10^6$ et $1/10^9$ modifications par cellule (Capecchi, 1989). La mise en place de technologies capables d'induire un clivage double brin spécifique d'une séquence d'intérêt a permis d'augmenter fortement l'efficacité d'édition du génome, en tirant profit des mécanismes cellulaires de réparation de l'ADN.

2.3.1. Mécanismes endogènes de réparation de l'ADN

A la suite d'un clivage double brin du génome, deux mécanismes cellulaires de réparation de l'ADN sont activés, le système de réparation par jonction d'extrémités non-homologues (*non-homologous end joining* ; NHEJ) et la recombinaison homologue (*homology directed repair* ; HDR) (Fig. 11).

Le mécanisme NHEJ permet de réparer l'ADN grâce à des exonucléases et des polymérases spécifiques capables de générer des extrémités cohésives, qui sont ensuite fusionnées par une ligase. Avec une efficacité variable en fonction du type cellulaire, le NHEJ conduit à la restauration de la séquence originale, mais peut fréquemment mener à une altération de l'information génétique par l'insertion ou la délétion de quelques nucléotides (*indels* ; Fig. 11) (Moore et Haber, 1996 ; Weaver, 1995). Ce mécanisme de réparation peu fidèle est utilisé en ingénierie génétique pour induire l'inactivation d'un gène *via* l'altération de son cadre de lecture (Kim, 2016).

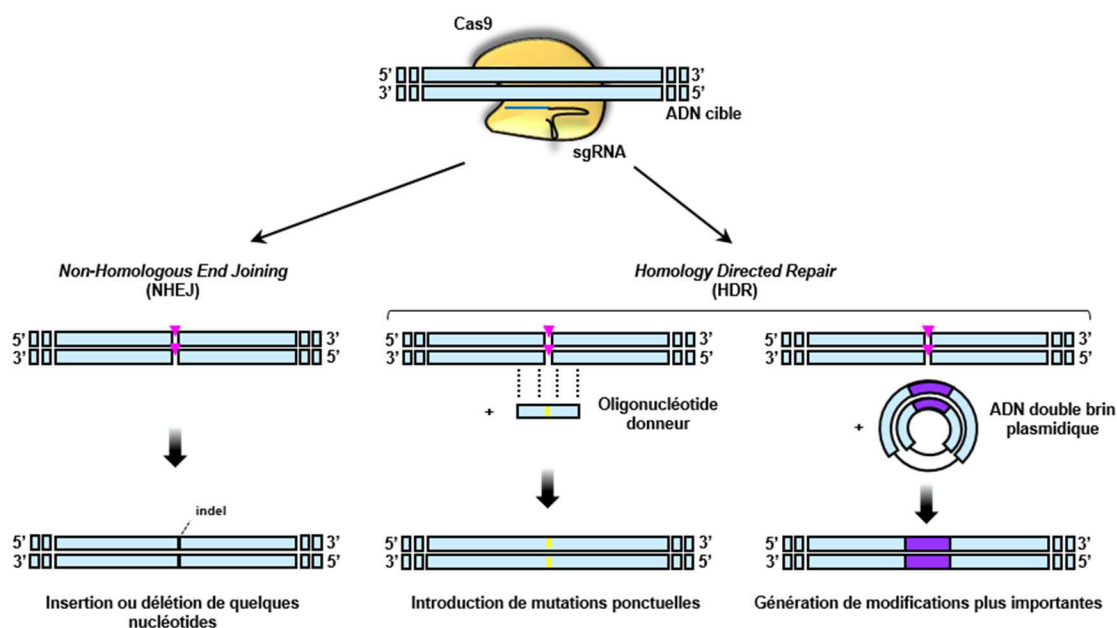


Figure 11. Mécanismes de réparation de l'ADN

La nucléase Cas9 induit une coupure double brin de l'ADN qui mène à l'activation de deux mécanismes cellulaires de réparation de l'ADN. Un mécanisme non-fidèle appelé *Non-homologous end joining* (NHEJ) engendre fréquemment des insertions ou des délétions de quelques nucléotides (*indels*). Le second mécanisme de réparation est l'*homology directed repair* ou HDR, qui utilise un ADN matrice contenant des séquences homologues avec le locus cible. En utilisant comme ADN donneur un oligonucléotide simple brin ou un vecteur plasmidique, l'HDR peut être utilisé pour introduire des mutations ponctuelles ou engendrer des modifications plus importantes du génome.

Le second mécanisme activé à la suite du clivage de l'ADN est l'HDR, qui permet de réparer l'ADN de façon fidèle par recombinaison homologue, en utilisant la chromatide sœur en guise d'ADN matrice au cours des phases S et G2 du cycle cellulaire (Bolderson *et al.*, 2009 ; Pardo *et al.*, 2009). En introduisant dans la cellule un fragment d'ADN exogène, contenant la modification de séquence voulue, flanquée par des extrémités homologues à la région du

génomique clivée, l'HDR permet de remanier à façon un locus d'intérêt (**Fig. 11**). Selon l'étendue de la modification, l'ADN donneur peut correspondre à une longue séquence d'ADN double brin, par exemple lorsque l'on veut insérer un gène entier, ou bien à un oligonucléotide simple brin (ssODN : *single stranded oligodeoxyribonucleotide*), pour introduire des mutations ponctuelles (Kim, 2016).

2.3.2. Technologies d'édition de l'ADN précédant la révolution CRISPR/Cas9

Avant l'avènement de la technologie CRISPR/Cas9, différents outils moléculaires avaient été mis au point pour le clivage spécifique d'une séquence d'intérêt, tels que les méganucléases, les *Zinc-Finger Nuclease* (ZFN) et les *Transcription Activator-Like Effectors* (TALEN) (**Fig. 12**). Ces approches reposent sur la reconnaissance de l'ADN par un domaine protéique.

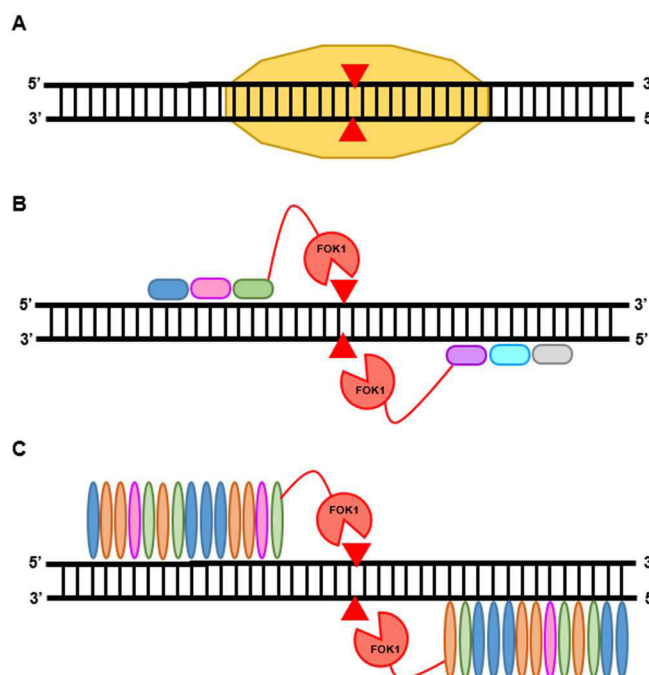


Figure 12. Autres stratégies d'édition du génome

A. Les méganucléases sont des endodéoxyribonucléases présentant un large site de reconnaissance de l'ADN double brin. Ces systèmes ont été créés pour cibler et cliver spécifiquement une région d'ADN donnée. **B.** Les ZFN sont composés d'un domaine contenant plusieurs motifs de liaison à l'ADN en doigt de zinc et d'un domaine de clivage de l'ADN dérivé de l'enzyme de restriction FokI qui fonctionne en dimère. **C.** Les TALEN sont un assemblage de différents domaines dérivés des protéines TALE, chacun capable de reconnaître un nucléotide, agencés à façon et fusionnés au domaine nucléase de FokI. A l'instar des ZFN, les TALEN fonctionnent par dimérisation.

2.3.2.1. Méganucléases

Les méganucléases sont des endodéoxyribonucléases présentes naturellement chez les phages, les bactéries, les archéobactéries, ainsi que divers eucaryotes, tels que la levure *Saccharomyces cerevisiae* ou l'algue verte *Chlamydomonas reinhardtii* (Dürrenberger et Rochaix, 1991 ; Sargueil *et al.*, 1991). Considérées comme des enzymes de restriction

extrêmement spécifiques, ces protéines reconnaissent des séquences d'ADN double brin de douze à quarante paires de bases (**Fig. 12A**) (Rouet *et al.*, 1994). La génération de méganucléases artificielles capables de cibler spécifiquement une séquence d'intérêt (Chevalier *et al.*, 2002 ; Smith *et al.*, 2006 ; Sussman *et al.*, 2004) a permis l'utilisation de ce système pour l'édition du génome (Gao *et al.*, 2010 ; Gürlevik *et al.*, 2013 ; Popplewell *et al.*, 2013 ; Rivière *et al.*, 2014). Cependant, la relation étroite entre domaine catalytique et domaine de liaison à l'ADN rend particulièrement complexe la création de ces outils, ce qui limite considérablement les applications de cette technologie.

2.3.2.2. Zinc-Finger Nucléases

Les ZFN sont des protéines synthétiques basées sur la fusion du domaine de clivage de l'endonucléase FokI avec un domaine de liaison à l'ADN en doigt de zinc dérivé de facteurs de transcription (Kim et Chandrasegaran, 1994 ; Kim *et al.*, 1996). Les premiers domaines à doigt de zinc ont été identifiés dans le facteur de transcription IIIA de *Xenopus laevis* et depuis d'autres domaines de liaison à l'ADN du même type ont été découverts, tels que le facteur Zif268 chez la souris ou encore le facteur GAGA chez la drosophile (Miller *et al.*, 1985 ; Pavletich et Pabo, 1991 ; Pedone *et al.*, 1996). Chaque doigt de zinc peut être modifié pour reconnaître un triplet de nucléotides, et de multiples doigts de zinc (typiquement trois à six) peuvent être assemblés en tandem pour cibler spécifiquement une séquence d'ADN d'intérêt (**Fig. 12B**) (Desjarlais et Berg, 1992 ; Rebar et Pabo, 1994 ; Shimizu *et al.*, 2009). L'enzyme FokI fonctionnant sous forme de dimère, deux ZFN sont requis pour induire une coupure double brin de l'ADN (Bitinaite *et al.*, 1998 ; Smith *et al.*, 2000). Utilisée pour la première fois chez la drosophile il y a plus de 15 ans (Bibikova *et al.*, 2002), cette technologie a été employée pour modifier le génome de nombreux organismes, comme par exemple le poisson (Doyon *et al.*, 2008 ; Meng *et al.*, 2008), la grenouille (Young *et al.*, 2011), la souris (Meyer *et al.*, 2010) ou encore le rat (Geurts *et al.*, 2009 ; Mashimo *et al.*, 2010).

2.3.2.3. Transcription Activator-Like Effector Nucleases

Décrites pour la première fois en 2009, les TALENs sont des protéines chimériques artificielles composées d'un domaine de liaison à l'ADN programmable dérivé des protéines TAL *effectors* (TALE), qui est fusionné au domaine nucléase de FokI (**Fig. 12C**) (Boch *et al.*, 2009 ; Moscou et Bogdanove, 2009). Les TALE sont des protéines naturelles sécrétées par la bactérie *Xanthomonas* et elles sont capables de se fixer à l'ADN grâce à des répétitions de 33 ou 34 acides aminés, chacune permettant de reconnaître spécifiquement l'un des quatre

nucléotides. L'assemblage de ces éléments selon un ordre précis permet de générer un nouveau domaine protéique qui reconnaît la séquence d'ADN d'intérêt (Boch *et al.*, 2009 ; Christian *et al.*, 2010 ; Mak *et al.*, 2012 ; Miller *et al.*, 2011 ; Moscou et Bogdanove, 2009). A l'instar des ZFN, les TALEN doivent être utilisées en dimère afin de cliver les deux brins d'ADN. Ce système a été utilisé pour l'édition du génome chez plusieurs espèces, telles que les levures, les plantes, les nématodes, les poissons ou encore les mammifères (rat, souris et cochon) (Christian *et al.*, 2010 ; Li *et al.*, 2011 ; Sung *et al.*, 2013 ; Tesson *et al.*, 2011).

2.3.3. Développement de CRISPR/Cas9 en tant qu'outil pour l'édition du génome

En 2012, deux groupes montrent pour la première fois que, en utilisant une combinaison de Cas9, crRNA et tracrRNA, il est possible de cliver spécifiquement *in vitro* une séquence double brin d'ADN cible, sous forme d'oligoduplex ou de séquence plasmidique (Gasiunas *et al.*, 2012 ; Jinek *et al.*, 2012). Les auteurs d'une de ces études révèlent également que le même type de coupure par la Cas9 peut être réalisé grâce à une séquence d'ARN chimérique, dérivée de la fusion du crRNA et du tracrRNA, et appelée ARN guide ou *single-guide RNA* (sgRNA) (Jinek *et al.*, 2012).

Seulement quelques mois après la publication de ces deux papiers, des équipes américaines et coréennes décrivent l'utilisation de CRISPR/Cas9 pour induire des modifications spécifiques dans le génome de cellules humaines et de souris (Cho *et al.*, 2013a ; Cong *et al.*, 2013 ; Mali *et al.*, 2013a). Pour cela, ces auteurs développent une Cas9 humanisée, dans laquelle un signal de translocation nucléaire est ajouté pour l'adressage dans le noyau des cellules eucaryotes (Cong *et al.*, 2013, 2013). Rapidement, d'autres études démontrent que ce type d'approche peut être appliqué chez différentes espèces, y compris les poissons (Chang *et al.*, 2013 ; Hwang *et al.*, 2013), les insectes (Bassett *et al.*, 2013 ; Gratz *et al.*, 2013), les nématodes (Dickinson *et al.*, 2013 ; Friedland *et al.*, 2013) et les plantes (Jiang *et al.*, 2013 ; Li *et al.*, 2013b).

L'une des premières applications de la technologie CRISPR/Cas9 est l'inactivation de gène *via* l'introduction d'*indels* par le mécanisme de réparation NHEJ, décalant ainsi le cadre de lecture (**Fig. 11**). En guise de preuve de concept, Le Cong *et al.* ont utilisé cette approche pour invalider le gène *EMX1* après transfection des cellules HEK-293FT avec des plasmides contenant séparément la Cas9 et un sgRNA dirigé contre une portion de ce locus (Cong *et al.*, 2013). La comparaison des résultats obtenus dans différentes études révèle que l'efficacité du clivage par le système CRISPR/Cas9 est très variable et peut atteindre jusqu'à 80% en fonction de la séquence ciblée et du modèle cellulaire (Kim *et al.*, 2014 ; Zuris *et al.*, 2015).

En plus de sa capacité à induire l'inactivation de gènes, le système CRISPR/Cas9 permet également d'introduire des modifications précises de l'ADN par l'intermédiaire du mécanisme de réparation HDR (**Fig. 11**). Cette stratégie a par exemple été utilisée pour introduire les sites de restriction NheI et HindIII dans le gène *EMX1* à l'aide d'un ADN donneur. La présence de ces sites de restriction a ensuite permis d'évaluer l'efficacité de recombinaison par une approche de RFLP (*restriction fragment length polymorphism*) (Cong *et al.*, 2013). Contrairement au mécanisme NHEJ, la réparation par recombinaison homologue présente un taux d'efficacité relativement faible, généralement inférieur à 20%.

Facile d'utilisation et d'une très bonne efficacité, la technologie CRISPR/Cas9 permet de cibler plusieurs gènes simultanément (Cong *et al.*, 2013 ; Mali *et al.*, 2013a ; Wang *et al.*, 2013). Ce type d'application, appelé multiplex, consiste en l'introduction dans la cellule de la Cas9 et de différents sgRNA dirigés contre des gènes distincts. Ainsi, Wang *et al.*, ont inactivé simultanément les trois membres de la famille *Ten-eleven translocation* dans des cellules souches embryonnaires de souris en co-transfectant trois sgRNAs en combinaison avec un plasmide contenant la Cas9 (Wang *et al.*, 2013). Les auteurs ont ciblé ces gènes responsables de la conversion des 5-méthylcytosine en 5-hydroxyméthylcytosine (5hmC) pour la simplicité de visualisation du phénotype des cellules contenant la mutation, qui se traduit par une diminution du niveau de 5hmC.

Le système CRISPR/Cas9 permet également de générer des modèles transgéniques avec une plus grande simplicité, comparé aux méthodes précédentes. En effet, de par une injection directement dans le zygote au stade une cellule, cette technique permet de s'affranchir des longues étapes de sélection de cellules souches embryonnaires contenant la mutation désirée et de croisement de souris chimères, nécessaires avec les méthodes précédentes (Cheng *et al.*, 2013 ; Wang *et al.*, 2013 ; Yang *et al.*, 2013). En particulier, ce type d'approche a ouvert la voie à la génération de modèles transgéniques dans des espèces non-murines, telles que le rat (Hu *et al.*, 2013 ; Ma *et al.*, 2014), le lapin (Honda *et al.*, 2015 ; Yuan *et al.*, 2016), le cochon (Hai *et al.*, 2014 ; Whitworth *et al.*, 2014), la vache (Pinzon *et al.*, 2016 ; Wei *et al.*, 2015a) et le singe (Niu *et al.*, 2014).

2.3.4. Autres applications du système CRISPR/Cas9

En exploitant la spécificité et la flexibilité du système CRISPR/Cas9, de nouvelles approches ont été développées pour des applications très différentes, telles que le remaniement des chromosomes, la modulation de l'expression génique ou la visualisation de séquences particulières dans le génome.

2.3.4.1. Remaniements chromosomiques

Selon la localisation des séquences ciblées, l'utilisation simultanée de deux sgRNAs peut induire des altérations génétiques plus ou moins complexes (**Fig. 13**). Ainsi, cette approche a été utilisée pour supprimer de larges portions de chromosomes (**Fig. 13A**) (Xiao *et al.*, 2013). Dans ce type d'expériences, la réparation de l'ADN peut également générer des inversions chromosomiques, si la région comprise entre les deux séquences ciblées est clivée et ensuite recollée à l'envers. Cette stratégie a notamment été utilisée pour récapituler une aberration génétique fréquente dans le cancer bronchique, l'inversion EML4-ALK sur le chromosome 2 (**Fig. 13B**) (Choi et Meyerson, 2014 ; Maddalo *et al.*, 2014 ; Torres *et al.*, 2014). Lorsque les sgRNAs ciblent des séquences sur deux chromosomes différents, on peut assister à des translocations (**Fig. 13C**), comme par exemple celles à l'origine des oncogènes de fusion *CD74-ROSI* (Choi et Meyerson, 2014) et *RUNXI-ETO* (Torres *et al.*, 2014).

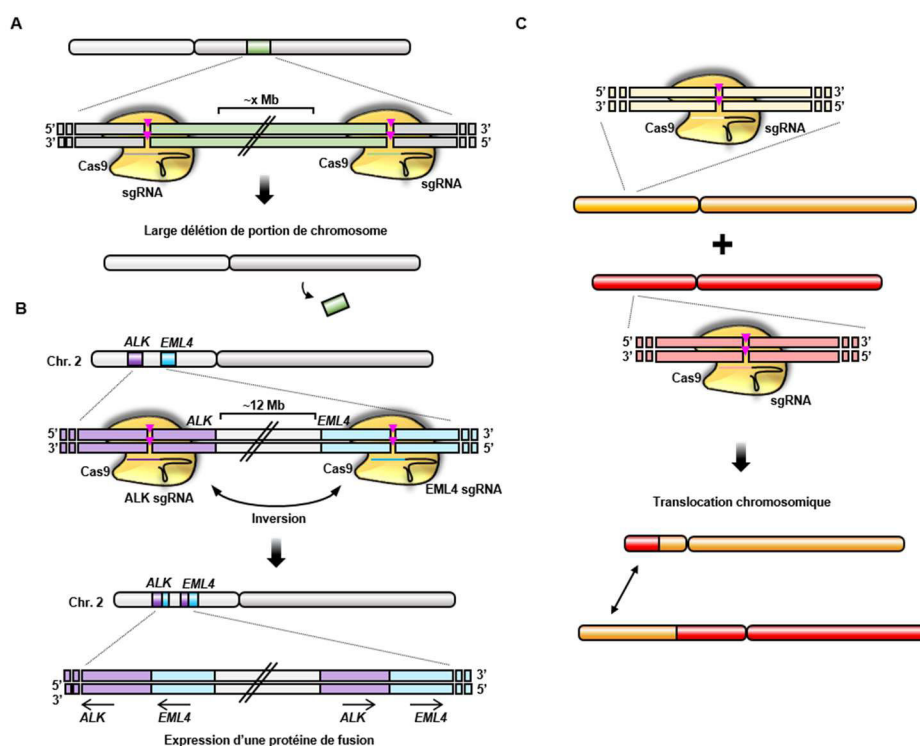


Figure 13. Remaniements chromosomiques engendrés en utilisant deux sgRNAs

A. Une co-expression de Cas9 et de deux sgRNAs peut être utilisée pour supprimer spécifiquement une région génomique. **B.** Représentation d'une inversion chromosomique impliquant les gènes *ALK* et *EML4* en utilisant la protéine Cas9 et deux sgRNAs distincts ciblant spécifiquement les gènes *ALK* et *EML4*. L'inversion chromosomique aboutit à l'expression d'une protéine de fusion EML4-ALK. **C.** Des sgRNAs ciblant deux chromosomes différents peuvent générer des translocations chromosomiques.

2.3.4.2. Régulation transcriptionnelle

L'endonucléase Cas9 clive l'ADN double brin par l'intermédiaire de ses domaines nucléases. Les domaines RuvC et HNH peuvent être inactivés *via* respectivement les mutations D10A et H840A sans impacter la liaison du complexe sgRNA-Cas9 à sa séquence cible (Bikard

et al., 2013 ; Qi *et al.*, 2013). Une mutation sur un seul de ces domaines confère à la Cas9 une activité de "nickase", puisqu'elle ne peut cliver qu'un seul brin d'ADN cible (**Fig. 14A et B**). En revanche, lorsque les deux domaines sont mutés, la Cas9 perd complètement son activité catalytique (*dead Cas9*, dCas9) (**Fig. 14C**).

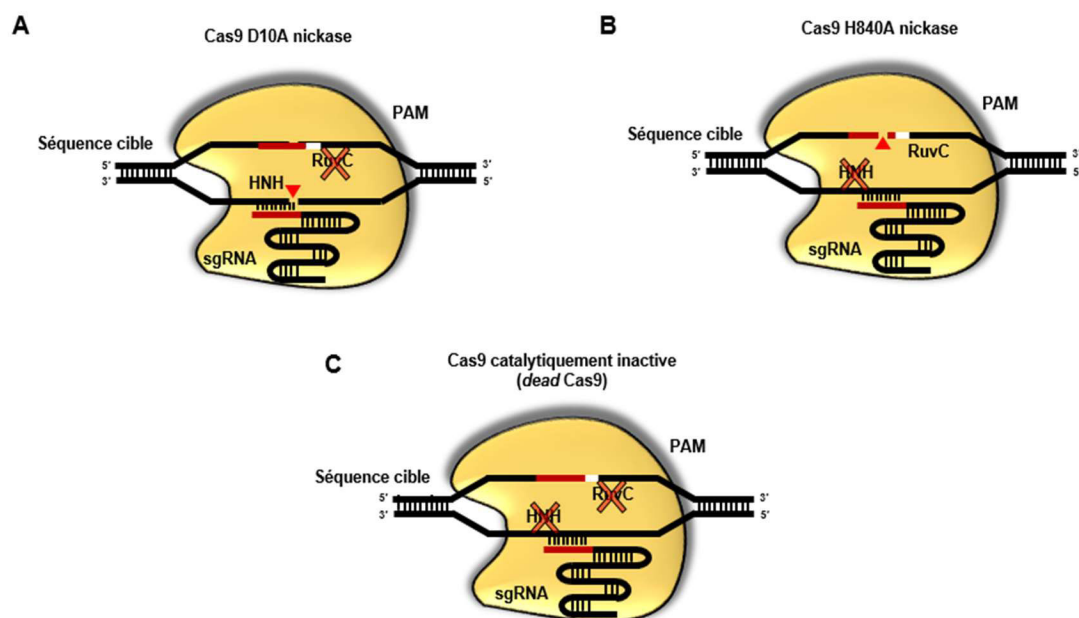


Figure 14. Utilisation de formes mutées de Cas9

A-B. Afin de modifier la nucléase Cas9 en nickase ne coupant plus qu'un seul brin d'ADN, il est possible d'introduire une mutation inactivatrice dans les domaines RuvC (D10A) (**A**) ou HNH (H840A) (**B**), responsables, respectivement, du clivage des brins non-sens et sens de l'ADN. **C.** L'introduction simultanément des mutations inactivatrices D10A et H840A dans les deux domaines nucléases de la Cas9 engendre une Cas9 catalytiquement inactive ou *dead Cas9* (dCas).

Les chercheurs ont exploité la capacité de la dCas9 à reconnaître spécifiquement une séquence d'ADN d'intérêt pour induire son recrutement au niveau du promoteur d'un gène d'intérêt à l'aide d'un sgRNA. Le complexe dCas9-sgRNA va ainsi perturber l'initiation de la transcription, inhibant l'expression du gène (Bikard *et al.*, 2013 ; Qi *et al.*, 2013). La fonction répressive de la dCas9 peut être accentuée dans les cellules eucaryotes *via* sa fusion avec un domaine répresseur, tels que le domaine KRAB (*krüppel associated box*) en doigt de zinc de Kox1, le CS (*chromoshadow*) de HP1 α ou le motif WRPW (Trp-Arg-Pro-Trp) de Hes1 (**Fig. 15A**) (Gilbert *et al.*, 2013 ; Konermann *et al.*, 2013). Cette technologie a été utilisée pour réprimer de près de 60% l'expression du récepteur à la transferrine (CD71) et du récepteur des chimiokines de type 4 (CXCR4) dans différentes cellules eucaryotes (Gilbert *et al.*, 2013).

La protéine dCas9 peut également être convertie en un activateur transcriptionnel synthétique en la fusionnant aux domaines d'activation p65 ou VP64 dérivé de la protéine virale VP16 (**Fig. 15B**) (Gilbert *et al.*, 2013). Pour une meilleure activation transcriptionnelle et une meilleure spécificité de ciblage, plusieurs sgRNA peuvent être utilisés simultanément dans le

but de cibler une région plus large de la séquence promotrice du gène d'intérêt (Maeder *et al.*, 2013 ; Mali *et al.*, 2013b ; Perez-Pinera *et al.*, 2013).

De nombreuses études continuent d'être menées afin d'améliorer l'efficacité des systèmes CRISPR/Cas9 pour la régulation transcriptionnelle. Une équipe américaine a récemment développé un système appelé SunTag, basé sur une protéine dCas9 fusionnée à plusieurs copies du peptide GCN4, qui peut être reconnu par un anticorps monoclonal couplé à un domaine VP64 (Tanenbaum *et al.*, 2014). Ainsi, du fait de la liaison antigène-anticorps, la dCas9 recrute de nombreux domaines VP64 au niveau de la séquence cible, permettant un taux d'expression élevé en utilisant un seul sgRNA (**Fig. 15C**) (Tanenbaum *et al.*, 2014). Un autre exemple vient de la fusion d'un sgRNA avec des ARN aptamères reconnus par des protéines de liaison spécifiques couplées à des activateurs transcriptionnels (**Fig. 15D**) (Konermann *et al.*, 2015 ; Mali *et al.*, 2013b).

L'activation ou la répression transcriptionnelle médiée par CRISPR peut également être régulée de façon inductible. Ainsi, différentes approches ont été développées, notamment en exprimant la dCas9 sous-contrôle de la protéine transactivatrice de la tétracycline (tTA) (Mandegar *et al.*, 2016) ou bien en utilisant des systèmes optogénétiques d'induction par une lumière bleue (Nihongaki *et al.*, 2015a, 2015b ; Polstein et Gersbach, 2015).

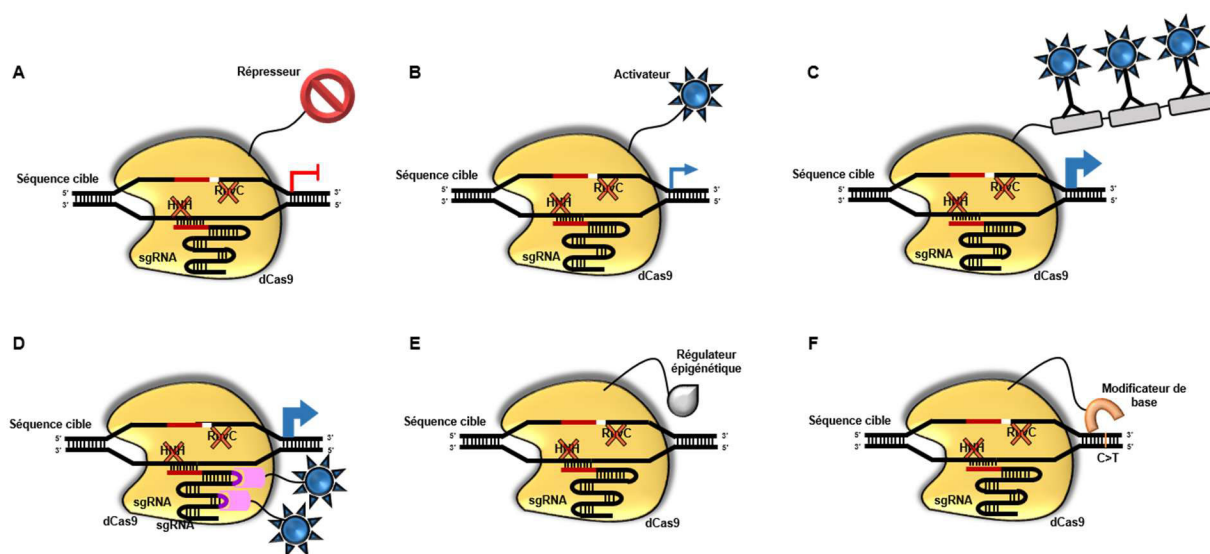


Figure 15. Exemples d'applications de dCas9

A-B. La Cas9 catalytiquement inactive (dCas9) peut être fusionnée avec un répresseur (**A**) (Gilbert *et al.*, 2013 ; Konermann *et al.*, 2013) ou un activateur (**B**) (Gilbert *et al.*, 2013) transcriptionnel dans le but d'inhiber ou d'augmenter l'expression d'un gène d'intérêt. **C.** La protéine dCas9 peut être assemblée avec plusieurs antigènes, qui vont être reconnus par un anticorps couplé à un activateur transcriptionnel. Cette technologie, appelée SunTag, permet de recruter un nombre plus important de régulateurs transcriptionnels au locus d'intérêt pour améliorer l'activation ou la répression du gène (Tanenbaum *et al.*, 2014). **D.** Le niveau d'expression d'un gène peut également être régulé par l'utilisation de sgRNAs chimériques composés de boucles aux motifs permettant de recruter des protéines de liaison spécifiques fusionnées à des activateurs ou des répresseurs transcriptionnels (Konermann *et al.*, 2015 ; Mali *et al.*, 2013b). **E.** La protéine dCas9 peut également être fusionnée avec des régulateurs épigénétiques afin de modifier l'état de la chromatine et moduler l'expression des gènes (Hilton *et al.*, 2015 ; Kearns *et al.*, 2015). **F.** En couplant la protéine dCas9 à une enzyme cytidine désaminase, il est possible de modifier spécifiquement une base à un endroit précis du génome sans coupure de l'ADN (Komor *et al.*, 2016).

2.3.4.3. Modulation épigénétique

Les modifications épigénétiques sur l'ADN ou sur les histones sont cruciales pour la régulation transcriptionnelle et jouent un rôle important dans diverses fonctions biologiques chez les mammifères. A l'aide du système CRISPR/dCas9 il est possible de stimuler l'expression d'un gène d'intérêt en couplant la protéine dCas9 à un activateur tel que l'histone acétyltransférase p300 (Hilton *et al.*, 2015). Cette technique va permettre d'augmenter l'expression des gènes, d'une part en relâchant la chromatine au niveau de la séquence promotrice et d'autre part en recrutant la machinerie transcriptionnelle nécessaire. Dans une autre étude, le couplage de la dCas9 avec l'histone déméthylase LSD1 a été utilisé pour inhiber l'expression de différents gènes, grâce au déplacement de groupes méthyle sur les lysines 4 et 9 de l'histone 3 (**Fig. 15E**) (Kearns *et al.*, 2015).

2.3.4.4. Edition de base par la cytidine désaminase

Une étude parue en 2016 décrit le développement d'une technologie capable de générer une modification ponctuelle de nucléotides sans induire de coupure de l'ADN, grâce à la fusion de l'enzyme cytidine désaminase avec la dCas9 (**Fig. 15F**) (Komor *et al.*, 2016). Le sgRNA permet à la dCas9 de se fixer à la séquence d'intérêt dans laquelle un changement de base est souhaité. Après désamination par la cytidine désaminase, la cytosine est transformée en uracile, qui possède les mêmes propriétés de liaison que la thymine (Conticello, 2008). L'hétéroduplex U:G conduit, après réplication, à deux génotypes (G:C et U:A). Le génotype U:A est ensuite corrigé par les mécanismes de réparation de l'ADN par excision de base, qui remplace l'uracile par une thymine, aboutissant donc à la modification de U:A en T:A. Cette méthode présente une haute efficacité et une faible fréquence d'*indels* ou de ciblage non-spécifique, ce qui confère à la stratégie un potentiel intéressant pour la correction de diverses mutations ponctuelles observées dans certaines pathologies humaines.

2.3.4.5. Criblage génétique fonctionnel

Les approches de criblages génétiques fonctionnels à haut débit peuvent être utilisées pour identifier les gènes impliqués dans différents processus biologiques. La simplicité du système CRISPR/Cas9 a permis le développement de banques de sgRNAs dirigés contre des milliers de gènes (Shalem *et al.*, 2014). A la différence des banques d'ARN interférence (short hairpin RNA, shRNA) qui inhibent l'expression, celles de sgRNAs permettent, en combinaison avec la nucléase Cas9, d'invalider les gènes, et donc d'obtenir des phénotypes généralement plus francs. Ces banques contenant des dizaines de milliers de plasmides lentiviraux, chacun présentant un

sgRNA différent, sont amplifiées, empaquetées puis introduites dans les cellules par infection, en présence de la protéine Cas9. Afin de s'assurer que chaque cellule contient un seul lentivirus, la transduction est effectuée avec une faible multiplicité d'infection. La proportion de chaque sgRNA au sein de la population peut ensuite être mesurée par séquençage à haut débit des provirus intégrés dans le génome des cellules infectées, fonctionnant comme une sorte de code-barres génétique. Ce type d'approche est généralement utilisé pour étudier les effets, négatifs ou positifs, de l'inactivation de gènes sur la croissance cellulaire. Ainsi, les groupes de David Sabatini et Eric Lander ont pu identifier une série de facteurs nécessaires à la prolifération et la survie des cellules humaines en culture (Wang *et al.*, 2015b). Comme exemple de sélection positive, Shalem *et al.* ont introduit une banque lentivirale CRISPR/Cas9 dans les cellules A375 de mélanome humain, pour ensuite les traiter en présence ou en l'absence de l'inhibiteur de BRAF vemurafenib. En analysant le profil de plusieurs milliers de sgRNAs dans les deux conditions, ces auteurs ont identifié différents gènes, dont l'inactivation permet de promouvoir la résistance à cette thérapie ciblée (Shalem *et al.*, 2014). Un an plus tard, le même groupe a utilisé une approche complémentaire, basée sur une banque de sgRNAs en combinaison avec une dCas9 couplée à un transactivateur, afin d'isoler les gènes qui, surexprimés, rendent les cellules de mélanome résistantes à l'inhibition de BRAF (Konermann *et al.*, 2015).

Comme démontré dans un papier paru en 2015, les banques lentivirales CRISPR/Cas9 peuvent être employées pour des études *in vivo*. En inoculant dans des souris immuno-déficientes des cellules cancéreuses murines non-métastatiques préalablement infectées avec une collection de 67 405 sgRNAs différents, Chen *et al.* ont pu identifier plusieurs gènes inactivés dans la tumeur primaire et/ou les métastases, révélant ainsi leur probable fonction d'oncosuppresseurs (Chen *et al.*, 2015).

2.3.4.6. Imagerie de région précise du génome

Afin de comprendre comment l'organisation du génome en chromatine est établie, propagée, maintenue et modifiée au cours du développement et en réponse à des stimuli environnementaux, il est nécessaire d'utiliser une méthode robuste pour visualiser l'ADN dans les cellules vivantes. La FISH (*fluorescent in situ hybridization*) est la technique de référence pour marquer l'ADN, mais elle nécessite pour cela la fixation des cellules (Chaumeil *et al.*, 2013). Au contraire, le système CRISPR/dCas9 peut être adapté pour marquer le génome dans les cellules vivantes et donc étudier des processus dynamiques (**Fig. 16**). L'une des approches les plus simples utilise une fusion dCas9-EGFP afin de marquer et de suivre un locus génomique dans les cellules vivantes (**Fig. 16A**) (Chen *et al.*, 2013). Cependant, la lumière

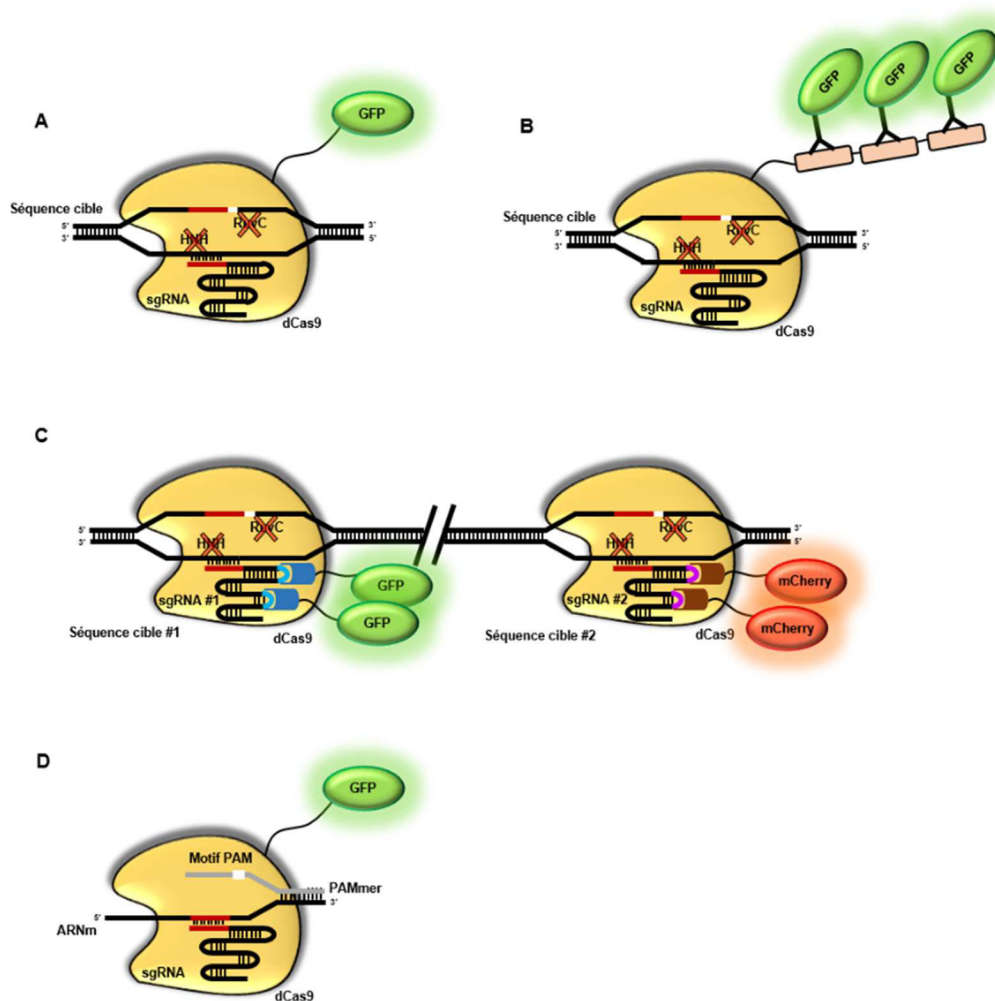


Figure 16. Approches CRISPR/Cas9 pour visualiser une région spécifique du génome

A. La fusion d'une protéine Cas9 catalytiquement inactive (dCas9) avec une protéine fluorescente (GFP) (dCas9-GFP) permet de visualiser un locus d'intérêt (Chen et al., 2013). **B.** Basée sur la technologie SunTag (Fig. 15), une dCas9 fusionnée avec différents antigènes peut être reconnue par des anticorps couplés à la GFP. Cette technologie permet le recrutement de plusieurs molécules fluorescentes à un endroit précis du génome, améliorant ainsi sa détection (Tanenbaum et al., 2014). **C.** Un locus peut également être visualisé par l'utilisation de sgRNAs chimériques composés de boucles aux motifs permettant de recruter des protéines de liaison spécifiques fusionnées à une molécule fluorescente (GFP ou mCherry). En utilisant des motifs spécifiques de différentes protéines de liaison couplées à des protéines fluorescentes distinctes, il est possible de visualiser simultanément deux régions génomiques d'intérêt (Fu et al., 2016 ; Shao et al., 2016). **D.** L'hybridation d'un oligonucléotide contenant un motif PAM, appelé PAMmer, au niveau d'une séquence d'ARNm rends possible le recrutement d'une Cas9. Ainsi, la fusion d'une protéine GFP à la dCas9, permettra de détecter facilement l'ARNm (Nelles et al., 2016).

dégagée par l'unique molécule de GFP fusionnée à la dCas9 est généralement trop faible, nécessitant ainsi l'utilisation de plusieurs sgRNAs pour cibler des régions à proximité du locus d'intérêt. La seconde stratégie est basée sur la méthode dCas9-SunTag, décrite précédemment. Alors que la dCas9 va reconnaître spécifiquement le locus d'intérêt grâce au sgRNA, la GFP fusionnée à un anticorps reconnaît l'antigène porté par la dCas9, amplifiant le signal et permettant ainsi une meilleure révélation (**Fig 16B**) (Tanenbaum *et al.*, 2014). Une autre approche consiste à coupler les sgRNAs avec des motifs d'ARN particuliers, tels que MS2, PP7 ou boxB, qui peuvent être reconnus spécifiquement par des protéines, telles que MCP (*MS2 coat protein*), PCP (*PP7 coat protein*) ou λ N. En couplant ces protéines avec des protéines

fluorescentes, on peut donc marquer simultanément différentes régions d'intérêt dans le génome (**Fig. 16C**) (Fu *et al.*, 2016 ; Shao *et al.*, 2016). Deux à trois boucles ARN composées de motifs distincts peuvent également être combinées avec divers marqueurs fluorescents sur le même sgRNA, qui par conséquent va émettre une lumière distincte, permettant ainsi d'imager jusqu'à six loci différents au sein d'une même cellule (Ma *et al.*, 2016). Cette technologie, appelée CRISPRainbow, représente donc un nouvel outil d'imagerie pour étudier la conformation des chromosomes et des mécanismes de régulations épigénétiques, directement dans des cellules vivantes.

Le système CRISPR/dCas9 peut également être utilisé pour suivre une séquence d'ARN d'intérêt dans les cellules vivantes (Nelles *et al.*, 2016). Cependant, pour fonctionner le complexe Cas9-ARN guide doit reconnaître le motif PAM présent sur le brin non cible, rendant alors impossible le ciblage d'un ADN simple brin ou d'un ARN messenger (ARNm). Afin de parvenir à cibler ce type de séquence, l'équipe du Dr Doudna a mis au point un système de leurre permettant de contourner cette limitation, grâce à l'hybridation sur une séquence simple brin d'un oligonucléotide présentant un motif PAM, appelé PAMmer. Le complexe Cas9-sgRNA peut alors reconnaître le PAM contenu dans l'oligonucléotide PAMmer et induire le clivage de la séquence d'ARNm (O'Connell *et al.*, 2014 ; Sternberg *et al.*, 2014). En utilisant une Cas9 catalytiquement inactive et couplée à une GFP, cette approche permet également d'imager une séquence d'ARNm voulue, sans affecter le niveau de protéines traduites (**Fig. 16D**) (Nelles *et al.*, 2016). Cette technique est particulièrement bien adaptée à l'étude du trafic des ARNm, notamment dans les vésicules de stress dans les cellules vivantes (Nelles *et al.*, 2016).

2.3.5. Stratégies pour l'introduction du système CRISPR/Cas9 dans les cellules

Le système CRISPR/Cas9 peut être utilisé dans une large gamme de types cellulaires et d'organismes en utilisant diverses méthodes d'introduction de matériel génétique. Ce système peut être délivré sous forme de vecteur, d'ARN codant la Cas9 ou de protéine recombinante complexée au sgRNA (Cho *et al.*, 2013b ; Li *et al.*, 2013a ; Lo *et al.*, 2013).

Les vecteurs viraux, tels que les virus adeno-associé (AAV) ou les lentivirus, sont couramment utilisés pour exprimer des gènes d'intérêt dans la plupart des types cellulaires, *in vitro* comme *in vivo*. Le plus souvent, les chercheurs ont recours aux vecteurs AAV pour leur faible potentiel immunogénique et leur faible risque d'intégration chez l'hôte. De nombreuses études ont utilisé les virus AAV pour délivrer le système CRISPR/Cas9 chez la souris dans

différents tissus, tels que le cerveau *via* des injections stéréotaxiques, ou encore les muscles cardiaque et squelettique par injection intramusculaire (Long *et al.*, 2014 ; Nelson *et al.*, 2016 ; Swiech *et al.*, 2015 ; Tabebordbar *et al.*, 2016). Une limitation majeure de l'utilisation des vecteurs AAV est la taille de l'ADN à insérer (~4.7 kb), qui rend difficile l'emploi de la Cas9 issue de *Streptococcus pyogenes* (SpCas9), dont la taille est relativement importante (~4.2 kb). Cette limitation peut cependant être surmontée en utilisant des orthologues plus petits (Cong *et al.*, 2013 ; Esvelt *et al.*, 2013 ; Hou *et al.*, 2013), tels que la Cas9 dérivée de *Staphylococcus aureus*, dont la séquence codante fait environ 3.5 kb (Nelson *et al.*, 2016 ; Nishimasu *et al.*, 2015 ; Ran *et al.*, 2015). Outre la taille réduite, certains orthologues de Cas9 ont la particularité de posséder un motif PAM plus long comparé aux trois bases de la SpCas9 : il s'agit par exemple de la séquence 5'-NNAGAAW-3' pour la Cas9 de *Streptococcus thermophilus* ou 5'-NNNNGATT-3' pour celle de *Neisseria meningitidis* (Garneau *et al.*, 2010 ; Zhang *et al.*, 2013). Eu égard à leur taille, ces motifs sont moins fréquents dans le génome, et, par conséquent, les systèmes basés sur ces autres Cas9 devraient permettre de limiter le nombre de séquences non-spécifiques potentielles.

L'introduction du système CRISPR/Cas9 dans les cellules peut également se faire *via* des vecteurs non-viraux. Parmi les différentes techniques développées, on retrouve l'utilisation de liposomes cationiques, qui peuvent encapsuler la protéine Cas9 et le sgRNA (Zuris *et al.*, 2015), ou bien l'utilisation de nanoparticules polymères cationiques, telles que le polyéthylèneimine (PEI) (Zuckermann *et al.*, 2015). Les composants du système CRISPR/Cas9 peuvent également être fusionnés ou complexés avec des peptides de translocation, aussi appelés *Cell-penetrating peptides*, qui possèdent la capacité de franchir la membrane plasmique (Ramakrishna *et al.*, 2014). Enfin, le système CRISPR/Cas9 peut également être délivré *via* des moyens physiques. Parmi ceux-ci, l'électroporation est la méthode la plus utilisée, principalement parce qu'elle peut être adaptée à une très large variété de types cellulaires en culture et utilisée également *in vivo* et *in utero* (Zuckermann *et al.*, 2015). Comme alternative, l'injection hydrodynamique a aussi été employée pour l'édition du génome dans le foie de souris (Xue *et al.*, 2014 ; Yin *et al.*, 2014).

2.4. Utilisation de CRISPR pour l'étude de pathologies

Grâce à sa capacité à modifier de façon spécifique n'importe quelle séquence d'ADN d'intérêt, la technologie CRISPR/Cas9 a permis de modéliser différentes maladies génétiques, facilitant ainsi la compréhension des mécanismes moléculaires impliqués dans ces pathologies. La dystrophie musculaire de Duchenne est une pathologie qui cause une dégénérescence des muscles squelettiques et cardiaques, aboutissant à une mort prématurée du patient, généralement suite à un arrêt cardiaque. Cette pathologie est causée par des aberrations dans le gène *DMD* codant la dystrophine, telles qu'une mutation non-sens dans l'exon 51, ou bien une délétion des exons 48 à 50. Ces différentes altérations génétiques engendrent un codon stop prématuré menant à une protéine dystrophine tronquée non fonctionnelle, qui est dégradée. Ainsi, des approches de thérapie génique peuvent être envisagées pour rétablir l'expression d'une dystrophine partiellement fonctionnelle (Robinson-Hamm et Gersbach, 2016). L'étude de cette pathologie a été facilitée par le développement de différents modèles murins, dont le plus utilisé présente une mutation non-sens de l'exon 23 du gène *mdx* qui code la dystrophine (Robinson-Hamm et Gersbach, 2016). Dans trois études publiées en parallèle, le système CRISPR/Cas9 a été utilisé pour induire une délétion de l'exon 23 de *mdx* dans ces souris, en injectant en intramusculaire des virus AAV contenant la Cas9 et deux sgRNAs dirigés contre les introns 22 et 23. Ces auteurs ont ainsi montré que la perte de l'exon 23 se traduit par une restauration partielle de l'expression d'une dystrophine tronquée mais fonctionnelle, donnant lieu à une amélioration significative des capacités musculaires de ces souris (Long *et al.*, 2016 ; Nelson *et al.*, 2016 ; Tabebordbar *et al.*, 2016).

CRISPR/Cas9 a également été utilisé pour étudier des maladies infectieuses. Il a été montré par exemple que ce système est capable de cibler et de cliver spécifiquement des régions conservées dans le génome du virus de l'hépatite B, résultant en une forte diminution de l'expression et de la réplication des gènes viraux (Ramanan *et al.*, 2015). Le système CRISPR/Cas9 peut également perturber l'intégration d'un génome viral dans les cellules. Ainsi, des cellules souches pluripotentes ont été modifiées pour exprimer de façon stable la Cas9 et un sgRNA dirigé contre le HIV, permettant donc d'obtenir des cellules immunitaires différenciées résistantes à une infection par ce virus (Liao *et al.*, 2015). Ces différentes études témoignent du potentiel du système CRISPR/Cas9 non seulement comme outil pour la recherche, mais également en tant que nouvelle approche thérapeutique.

L'un des enjeux majeurs de la recherche contre le cancer est de caractériser les mutations oncogéniques et d'évaluer leurs implications dans la formation et la progression des tumeurs.

Grâce à la technologie CRISPR/Cas9, les aberrations génétiques retrouvées dans les tumeurs des patients peuvent être facilement recréées dans des modèles expérimentaux. Nous avons écrit une revue, résumant les diverses stratégies CRISPR/Cas9 utilisées pour récapituler les mutations oncogéniques dans différents systèmes *in vitro* et *in vivo*, ainsi que les nouvelles stratégies développées pour modéliser l'évolution tumorale et l'hétérogénéité génétique (Voir section « Résultats »).

2.5. Limitations de la technologie CRISPR/Cas9

Malgré ses nombreux avantages et potentialités, le système CRISPR/Cas9 présente certaines limitations, notamment liées à la possibilité d'effets hors-cible et à la relativement faible efficacité d'édition de l'ADN par HDR.

2.5.1. Effets non-spécifiques de CRISPR/Cas9

Plusieurs équipes de recherche ont montré que le complexe Cas9/sgRNA peut tolérer des mésappariements de quelques nucléotides entre le sgRNA et la séquence d'intérêt, ce qui peut engendrer des clivages non-spécifiques de l'ADN (Fu *et al.*, 2013 ; Hsu *et al.*, 2013, 2014 ; Mali *et al.*, 2013c ; Pattanayak *et al.*, 2013). Ces effets peuvent être accentués lorsque la protéine Cas9 se trouve en forte concentration dans la cellule (Hsu *et al.*, 2013 ; Pattanayak *et al.*, 2013). Afin de minimiser la survenue de ces coupures non spécifiques, différentes approches ou précautions sont conseillées.

Tout d'abord, il est nécessaire d'identifier en amont l'ensemble des séquences génomiques susceptibles d'être clivées par la Cas9. Pour cela, différentes bases de données ont été mises à disposition, telles que CRISPR Design Tool ou encore E-CRISP, qui permettent de prédire les sites non-spécifiques potentiels sur la base du degré de similarité du sgRNA avec la séquence cible, tout en considérant d'autres paramètres, tels que la localisation des mésappariements au sein de la séquence d'intérêt (Heigwer *et al.*, 2014 ; Hsu *et al.*, 2013). Ces outils bio-informatiques sont le fruit de la comparaison de travaux théoriques et expérimentaux. Les données théoriques ont été obtenues par prédiction *in silico* des sites non-spécifiques basée sur une similarité avec la séquence cible, tandis que les données pratiques ont été générées à la suite d'expériences de clivage par le système CRISPR/Cas9 dans lesquelles les *indels* introduits à la suite de la réparation de l'ADN par NHEJ sont évalués par séquençage à haut débit (Fu *et al.*, 2013 ; Hsu *et al.*, 2013). Bien qu'adapté à la détection de coupures non-spécifiques dans des populations cellulaires homogènes, telles que les clones issus de cellules isolées ou encore les animaux F1 non mosaïques (Iyer *et al.*, 2015 ; Smith *et al.*, 2014 ; Veres *et al.*, 2014), le

séquençage à haut débit manque de sensibilité lorsque l'on analyse des altérations peu fréquentes et variables dans le génome, et il ne peut donc pas être utilisé pour des populations cellulaires hétérogènes.

Différentes équipes de recherche ont développé des techniques permettant de détecter plus efficacement les effets non-spécifiques du système CRISPR/Cas9 (pour revue : Tsai et Joung, 2016). L'une de ces stratégies, appelée GUIDE-seq (*genome-wide unbiased identification of double strand breaks enabled by sequencing*), repose sur l'incorporation d'un oligonucléotide double brin dans les sites de coupure de l'ADN induits par la Cas9. Grâce à l'amplification spécifique de cette séquence avec des portions avoisinantes d'ADN génomique, il est ainsi possible d'identifier par séquençage à haut débit le site de clivage attendu, mais également l'ensemble des coupures hors-cible induites par le complexe Cas9/sgRNA (Tsai *et al.*, 2015).

Différentes stratégies ont été développées afin d'augmenter la spécificité du système CRISPR/Cas9. L'une de ces approches consiste à générer des sgRNAs dans lesquels deux nucléotides G sont ajoutés à l'extrémité 5', limitant ainsi le clivage non-spécifique. Cependant, le mécanisme responsable de cet effet n'a pas encore été complètement caractérisé, et ces sgRNAs peuvent parfois présenter une efficacité réduite (Cho *et al.*, 2014 ; Kim *et al.*, 2015). Il a également été suggéré que l'utilisation de sgRNAs tronqués de 2 à 3 nucléotides à l'extrémité 5' pourrait permettre une meilleure stabilité de liaison, en diminuant ainsi le nombre de coupures hors-cible (Fu *et al.*, 2014 ; Tsai *et al.*, 2015). A partir de ces travaux, une équipe de recherche a développé un site internet pour dessiner des sgRNAs potentiellement plus efficaces et spécifiques, grâce à une plus grande stabilité conférée par la délétion d'un ou deux nucléotides et la présence de quelques mésappariements à l'extrémité 5' (<http://CRISPRscan.org>) (Moreno-Mateos *et al.*, 2015).

La seconde approche utilisée pour augmenter la spécificité du système CRISPR/Cas9 consiste à modifier l'activité de la protéine Cas9. La stratégie la plus utilisée est basée sur le clivage de l'ADN par une Cas9 nickase (Cas9n), en combinaison avec deux sgRNAs qui ciblent les deux brins opposés à proximité de la région d'intérêt (**Fig. 17**). Cette astuce permet de diminuer considérablement le taux de coupure non-spécifique de l'ordre de 50 à 1500 fois, comparé à un système utilisant une Cas9 non modifiée (Cho *et al.*, 2014 ; Mali *et al.*, 2013c ; Ran *et al.*, 2013). En s'appuyant sur le même principe, un autre groupe a mis au point une dCas9 fusionnée au domaine nucléase de Fok1, qui fonctionne uniquement par dimérisation. L'utilisation d'une paire de dCas9-Fok1 couplées à deux sgRNAs différents permet ainsi d'induire un clivage double brin dans un locus d'intérêt, tout en réduisant la probabilité de

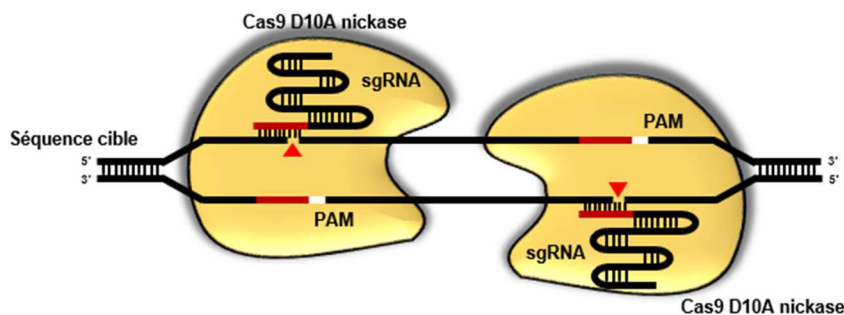


Figure 17. Utilisation de deux sgRNAs et une Cas9 nickase pour augmenter la spécificité du clivage de l'ADN

En combinant simultanément une Cas9 nickase et deux sgRNAs distincts contre deux séquences très proches, il est possible d'induire un clivage double brin de l'ADN avec beaucoup plus de spécificité qu'avec le système CRISPR/Cas9 classique.

provoquer d'autres modifications non voulues du génome (Guilinger *et al.*, 2014 ; Tsai *et al.*, 2014). Enfin, en mutant certains acides aminés de la Cas9, il a été montré qu'il est possible d'augmenter la capacité de cette protéine à discriminer entre séquences cible et hors-cible, et donc d'améliorer la spécificité du système, sans pour autant affecter l'efficacité d'édition de l'ADN (Kleinstiver *et al.*, 2015, 2016 ; Slaymaker *et al.*, 2016).

2.5.2. Faible efficacité de réparation par l'HDR

Le taux d'édition du génome par HDR est relativement faible et très variable en fonction du modèle cellulaire, de la séquence des sgRNAs, ou bien de la région d'ADN ciblée (Mali *et al.*, 2013a ; Wang *et al.*, 2013). Certains groupes ont montré que l'inhibition de différents composants de la voie de réparation NHEJ (KU70 KU80, ligase IV) par ARN interférence ou avec des petites molécules, tels que la SCR7, un inhibiteur de la ligase IV, permettait d'augmenter l'efficacité de recombinaison par HDR (Chu *et al.*, 2015 ; Maruyama *et al.*, 2015 ; Yu *et al.*, 2015). En parallèle, grâce à un criblage de 4 000 petites molécules, Yu *et al.* ont identifié la brefeldin A, un inhibiteur du transport intracellulaire des protéines dans le réticulum endoplasmique vers l'appareil de Golgi, et le L755507, un agoniste partiel des récepteurs β_3 adrénergiques, pour leur capacité à stimuler l'HDR, bien que les mécanismes impliqués ne soient pas encore connus (Yu *et al.*, 2015). Une étude récente a également suggéré que le choix d'un ssODN donneur complémentaire à la séquence du brin non ciblé par le sgRNA pourrait permettre une fréquence de réparation par l'HDR plus élevée (Richardson *et al.*, 2016). Même en utilisant ces différents artifices, dans une expérience de CRISPR/Cas9 seulement une fraction des cellules au sein d'une population pourra contenir la mutation souhaitée, rendant donc inévitable la dérivation de clones (Ran *et al.*, 2013). Ainsi, ce type d'approche n'est pas compatible avec des modifications génétiques ayant un impact négatif sur la prolifération ou la survie cellulaire.

2.6. Enjeux agro-alimentaires, biomédicaux et éthiques de CRISPR/Cas9

Le système CRISPR/Cas9 peut avoir des retombées importantes tant dans le domaine agricole que dans le secteur biomédical.

En 2016, en isolant grâce à CRISPR le gène responsable de l'oxydation, une équipe américaine a mis au point un champignon de Paris qui ne brunit pas quand on le coupe (Waltz, 2016). Cette étude, qui ouvre la voie vers la modification et l'amélioration des aliments, soulève une nouvelle fois la question de la définition des organismes génétiquement modifiés (OGM). Aux Etats-Unis, pays précurseur dans l'utilisation de ces mutants, les autorités sanitaires ont accordé la mise sur le marché de cet aliment, considérant que la réglementation des OGM ne s'applique pas à ce mutant, étant donné qu'aucun gène n'a été ajouté. En revanche, le débat fait rage en Europe, certains voyant d'un mauvais œil le fait que ces mutants puissent ne pas être considérés comme des OGM. Ces opposants soulignent également la difficulté de traçabilité de ces produits de par la complexité d'identification de ce type de mutation.

Le système CRISPR/Cas9 a également permis un réel bon en avant dans la thérapie génique. En 2016, les premiers essais cliniques utilisant cette technologie ont été mis en place. Ces différentes thérapies visent à extraire les cellules T des patients et inactiver par CRISPR/Cas9 le récepteur PD-1 (*programmed cell death 1*). Les cellules modifiées sont ensuite réintroduites chez le sujet et devraient désormais être capables de renforcer la réponse immunitaire dirigée contre les cellules cancéreuses.

Cependant, l'utilisation d'une technologie aussi puissante pose de nombreuses interrogations éthiques au sein de la communauté scientifique, notamment depuis la publication d'une étude menée par une équipe chinoise dans laquelle les chercheurs ont, pour la première fois, modifié génétiquement un embryon humain non viable (Liang *et al.*, 2015). Néanmoins, grâce à cette approche, des chercheurs sont parvenus à corriger une mutation hétérozygote dans un embryon humain, en utilisant le gène maternel sauvage homologue comme matrice, en l'absence de détection de mutations non-spécifiques ou de phénomène de mosaïque (Ma *et al.*, 2017). Cette étude, publiée en août dernier, suggère le fort potentiel du système CRISPR/Cas9 à corriger les mutations héréditaires dans les embryons humains et ainsi palier à des pathologies prédictibles.

La puissance de CRISPR/Cas9 permet également d'envisager une éradication des vecteurs porteurs de maladies transmissibles, tels que les moustiques grâce à une technique appelée « *gene driver* » qui permet de booster la propagation d'une mutation dans une population (Gantz *et al.*, 2015 ; Hammond *et al.*, 2016). La cassette « *gene driver* » contient un gène qui

code la protéine Cas9, un gène qui code un ARN guide capable de reconnaître une séquence d'ADN cible bien particulière, et des séquences en bordure de la cassette qui permettent d'insérer la totalité de cette cassette au site de coupure cible. En introduisant dans les moustiques une cassette « *gene driver* » contenant la résistance contre le paludisme par exemple et en relâchant une dizaine d'individus, on peut théoriquement obtenir en quelques dizaines de générations une population entièrement résistantes. Cependant, cette capacité de la cassette à se copier elle-même, et la capacité impressionnant de remplacement de la population d'origine par la population mutante suscite des inquiétudes de la part de la communauté scientifique.

Malgré les avantages majeurs du système CRISPR/Cas9 sur l'agro-alimentaire ou encore le biomédicale, il est essentiel de mettre en place un consortium international pour définir une utilisation responsable et éthique de CRISPR/Cas9. En effet, il est nécessaire de préserver notre espèce telle quelle est et d'éviter de modifier le patrimoine génétique germinale, qui est transmis à la descendance.

3. Cancer bronchique

3.1. Épidémiologie et facteurs de risque

Les cancers du poumon se développent à partir des cellules des bronches ou des cellules tapissant les alvéoles pulmonaires. Cette pathologie représente aujourd'hui la première cause de décès par cancer en France et dans le monde, devant le cancer colorectal, avec près de 1,6 millions de décès par an à travers le globe et un taux de survie net à 5 ans inférieur à 15% (Institut National du Cancer, INCa, 2017). En France, alors qu'on note une diminution chez les hommes, l'incidence de ce cancer est en progression constante chez les femmes, avec un nombre de nouveaux cas annuel multiplié par 7 en 30 ans. Le facteur de risque principal associé au développement de la maladie est de loin le tabagisme. D'autres facteurs sont également susceptibles d'accroître le risque de survenue d'un cancer bronchique, tels que la pollution atmosphérique, l'exposition à certains cancérigènes comme l'amiante et le radon, ou encore des variations génétiques familiales ou ethniques (INCa, 2017). Le traitement des cancers bronchiques est un enjeu majeur de santé publique. En 2009, en France, le coût global pour la prise en charge des patients a été évalué à plus de 468 millions d'euros, tandis qu'il représentait 4.2 milliards d'euros pour l'ensemble des pays de l'Union Européenne (Luengo-Fernandez *et al.*, 2013).

3.2. Classification des cancers bronchiques

En fonction de leurs caractéristiques histologiques, tels que la morphologie et l'origine des cellules qui les composent, la plupart des cancers pulmonaires peuvent être classés dans deux larges groupes : les cancers bronchiques à petites cellules (CBPC) et les cancers bronchiques non à petites cellules (CBNPC) (INCa, 2017).

3.2.1. Cancers bronchiques à petites cellules

Les CBPC représentent 15 à 20% des cancers bronchiques et se développent presque exclusivement chez les fumeurs. Ces tumeurs sont dites neuroendocrines puisqu'elles sont issues des cellules neuroendocrines qui tapissent l'épithélium pulmonaire. Les CBPC sont particulièrement agressifs et se caractérisent par une évolution très rapide et une tendance à former des métastases lors des stades précoces de la maladie, principalement au niveau des os, du foie et du cerveau. Généralement diagnostiqués à des stades avancés, les CBPC répondent dans un premier temps assez bien, à la chimiothérapie et à la radiothérapie. Cependant, la grande majorité des patients finissent par rechuter et succombent de leur tumeur. Les altérations

génétiques les plus répandues concernent les gènes suppresseurs de tumeurs *Tumor Protein 53 (TP53)*, *RB transcriptional corepressor 1 (RB1)* et *Phosphatase et TENsin homolog (PTEN)*, l'amplification/surexpression de l'oncogène *MYC* et des mutations inactivatrices des récepteurs NOTCH 1, 2, 3 (George *et al.*, 2015 ; van Meerbeeck *et al.*, 2011).

3.2.2. Cancers bronchiques non à petites cellules

La grande majorité des cancers du poumon (80-85%) sont dits non à petites cellules, de par la taille variable des cellules qui les composent. Ces CBNPC sont répartis en fonction de leur origine cellulaire en trois types principaux : les carcinomes du poumon à grandes cellules, les adénocarcinomes et les carcinomes épidermoïdes (ou à cellules squameuses).

Le carcinome à grandes cellules est le type le moins courant qui ne représente que 20% des CBNPC. Cette forme de tumeur, dont la grande majorité est liée au tabac, peut être située dans n'importe quelle région du poumon et se caractérise par des cellules indifférenciées et une croissance plus rapide que les autres types de CBNPC (Chen *et al.*, 2014). Les deux catégories prédominantes sont les adénocarcinomes et les carcinomes épidermoïdes, qui représentent près de 80% des CBNPC. Les adénocarcinomes se développent généralement dans la partie périphérique du poumon à partir des cellules alvéolaires de type II et ils constituent le type de cancer bronchique majoritaire chez les non-fumeurs et chez les femmes. Ces tumeurs présentent une histologie glandulaire et expriment différents marqueurs en accord avec une origine pulmonaire distale, tels que le surfactant C, la kératine 7 et le facteur de transcription NKX2-1 (Chen *et al.*, 2014). A l'inverse, les carcinomes épidermoïdes, dont la survenue est fortement liée au tabac, se caractérisent par une localisation centrale dans le poumon. Les cellules qui composent ce type de CBNPC présentent des marqueurs caractéristiques des cellules squameuses provenant des voies aériennes supérieures, tels que p63, SOX2 et kératine 5 (Chen *et al.*, 2014).

La suspicion de la présence d'un CBNPC repose sur la survenue de symptômes tels que la toux, une augmentation des expectorations, un essoufflement, un enrouement, des douleurs thoraciques et du sang dans les crachats. Le diagnostic a lieu après examen radiologique et biopsie d'un morceau de tumeur afin d'établir le stade de la maladie. En fonction de cette analyse et de l'état du patient, différentes stratégies thérapeutiques peuvent être envisagées, telles que la chirurgie, la radiothérapie ou la chimiothérapie adjuvante (association de cisplatine/vinorelbine) ou néoadjuvante (sels de platine). Récemment, de nouvelles approches basées sur l'immunothérapie ont été mises au point pour le traitement des CBNPC, avec l'approbation en 2015 du nivolumab (*Opdivo* ; Bristol-Myers Squibb) et en 2016 du

pembrolizumab (*Keytruda* ; Merck) et du atezolizumab (*Tecentriq* ; Roche). Ces anticorps monoclonaux ciblent respectivement le récepteur PD-1 (nivolumab et pembrolizumab), présent à la surface des lymphocytes T, et son ligand *programmed death ligand 1* (PD-L1) (atezolizumab), exprimé par les cellules tumorales (Tsao *et al.*, 2016). En bloquant l'interaction PD1/PD-L1, ces agents thérapeutiques permettent l'activation des lymphocytes T, restaurant ainsi leur capacité à détecter et éliminer les cellules tumorales.

La découverte de différentes aberrations oncogéniques dans les CBNPC, telles que des mutations de *Kirsten rat sarcoma 2 viral oncogene homolog* (*KRAS*), de *PTEN*, de l'*EGFR*, des amplifications du *fibroblast growth factor receptor 1* (*FGFR1*) et de *MET*, ou encore des remaniements chromosomiques impliquant le gène *anaplastic lymphoma kinase* (*ALK*) (**Fig. 18**), a permis non seulement de mieux comprendre les mécanismes impliqués dans la formation et la progression de ces tumeurs, mais également d'identifier des cibles potentielles pour le traitement.

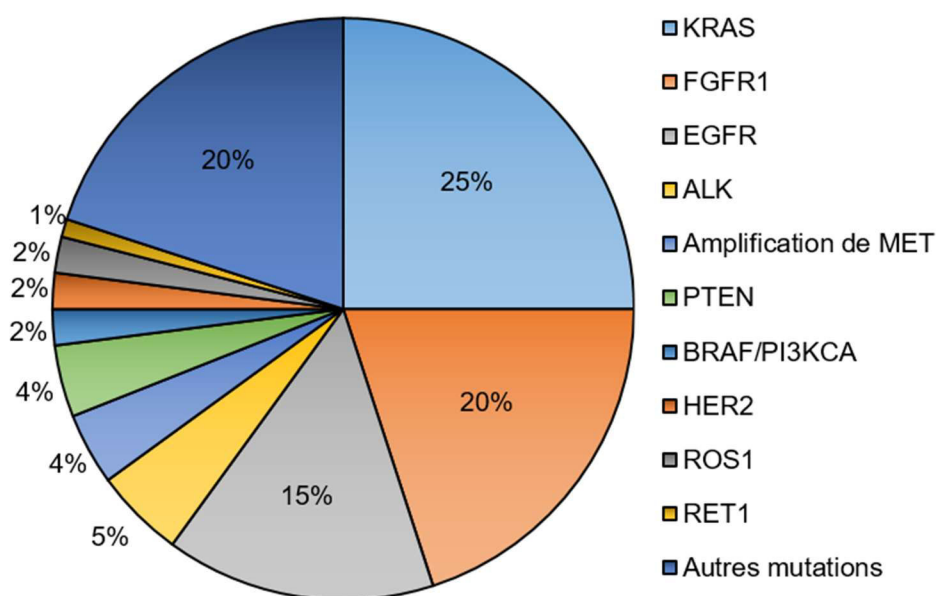


Figure 18. Fréquence des mutations dont l'implication a été démontrée dans la formation et/ou la progression des CBNPCs
Adaptée de Chan et Hughes, 2015 ; Tsao *et al.*, 2016.

3.2.2.1. Mutations de KRAS

L'altération de l'oncogène KRAS constitue l'une des modifications génétiques la plus répandue dans le cancer. Ce membre de la famille des GTPases RAS joue un rôle majeur dans la voie de transduction des récepteurs tyrosines kinases (RTKs). Après fixation de leurs ligands, ces récepteurs se dimérisent et s'autophosphorylent, permettant le recrutement de la protéine adaptatrice *growth factor receptor-bound protein 2* (Grb2) sur des résidus tyrosines

phosphorylés. Grb2 forme un complexe avec *son of sevenless* (SOS), qui stimule l'échange GDP-GTP par RAS. La protéine RAS couplée au GTP recrute et active la Ser/Thr RAF, qui induit la cascade de phosphorylation des *mitogen-activated protein kinases* (MAPK) (**Fig. 19**).

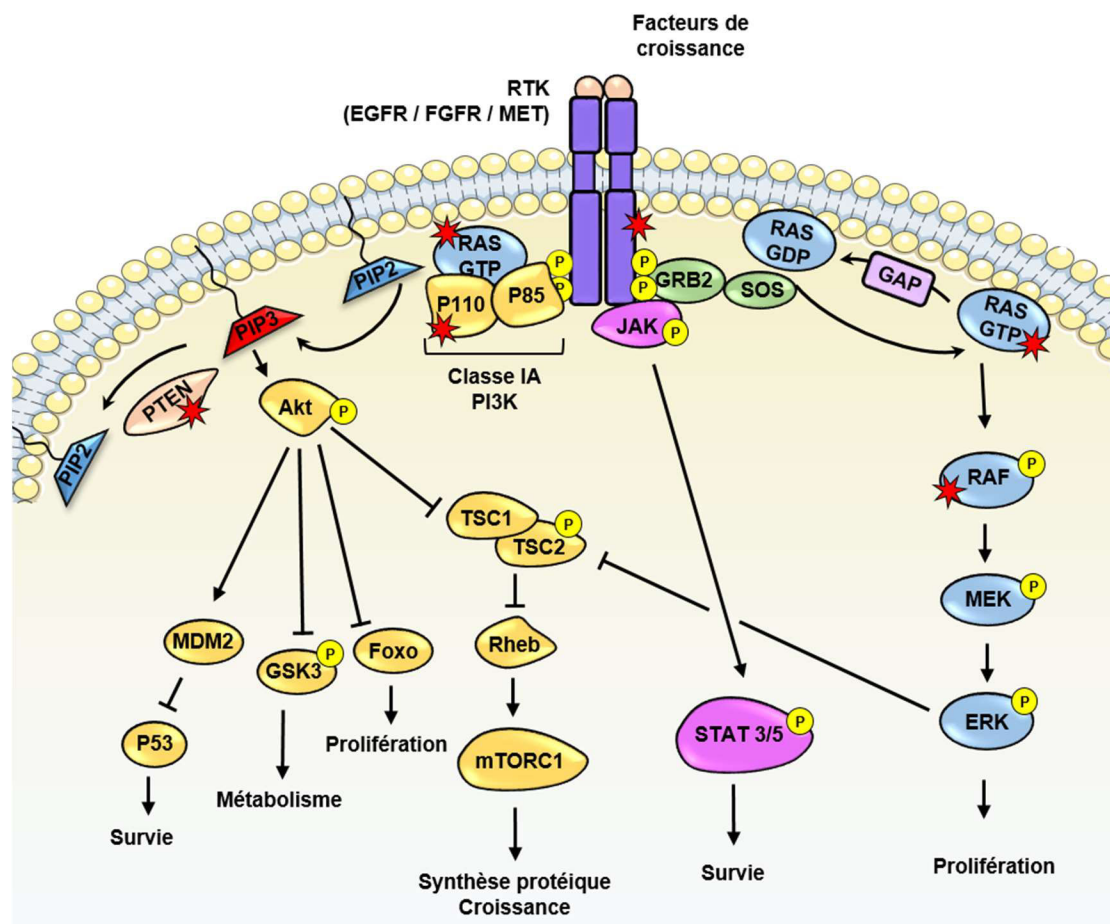


Figure 19. Voies de signalisation activées par les récepteurs tyrosine kinases (RTKs)

L'activation de ces récepteurs par leurs ligands stimule différentes voies de signalisation impliquées dans la régulation de la prolifération et la survie cellulaire, telles que les voies RAS-RAF-MEK-ERK, JAK-STAT ou encore PI3K-Akt. Dans les CBNPC, des mutations (étoiles rouges) dans certains récepteurs tyrosines kinases (EGFR, FGFR, MET) ou composants intracellulaires de ces voies, tels que PTEN, RAS, PI3KCA (p110) ou RAF, ont été identifiées (**Fig. 18**). Ces aberrations génétiques conduisent à une stimulation constitutive contribuant à la tumorigenèse des CBNPCs.

Abréviations : EGF, *epidermal growth factor* ; ERK, *extracellular signal-regulated kinases* ; Foxo, *Forkhead box* ; GAP, *GTPase-activating protein* ; GSK3, *glycogen synthase kinase 3* ; GRB2, *growth factor receptor-bound protein* ; JAK, *Janus kinase* ; MDM2, *mouse double minute 2* ; MEK, *mitogen-activated protein kinase kinase* ; mTORC1, *mammalian target of rapamycin complex 1* ; PI3K, *phosphatidylinositol 3-kinase* ; PIP2, *phosphatidylinositol-4,5 biphosphate* ; PIP3, *phosphatidylinositol(3,4,5) triphosphate* ; PTEN, *phosphatase et tensin homolog* ; Rheb, *RAS homolog enriched in the brain* ; SOS, *son of sevenless* ; STAT, *signal transducer et activator of transcription* ; TSC1/2, *tuberous sclerosis 1*.

RAS peut également former un complexe avec la phosphoinositide 3-kinase (PI3K) et entraîner la stimulation de voies de transduction médiées par AKT, impliquées dans la régulation du métabolisme, la synthèse protéique, la prolifération et la survie cellulaire (**Fig. 19**).

Présents dans 20 à 25% des CBNPC, les mutations de KRAS surviennent fréquemment au niveau des codons 12 et 13 dans l'exon 2, dont la plus courante est la G12C (c.34G > T), suivie par G12V (c.35G > T) et G12D (c.35G > A) (Wood *et al.*, 2016). KRAS peut également présenter des substitutions du codon 61 de l'exon 3 (Alamgeer *et al.*, 2013 ; Bhattacharya *et al.*,

2015). En fonction de la mutation, des voies de signalisation différentes peuvent être activées. Par exemple, les altérations G12V et G12C sont principalement associées à une stimulation de la GTPase Ral, impliquée dans l'invasion cellulaire, tandis que la modification G12D active de préférence PI3K (Ihle *et al.*, 2012). L'ensemble de ces aberrations génétiques sont principalement observées chez les individus Caucasiens et chez les fumeurs, et sont mutuellement exclusives d'autres altérations impliquant les gènes *EGFR* ou *ALK* (Gainor *et al.*, 2013 ; Suda *et al.*, 2010).

Depuis plus de 30 ans, les scientifiques tentent de comprendre comment inhiber efficacement KRAS. Cependant cette protéine reste une cible extrêmement difficile pour de nombreuses raisons, comme sa très forte affinité pour le GTP, l'absence de sites de liaison allostériques connus, les modifications post-traductionnelles complexes que la protéine subit et les multiples voies de signalisation alternatives qui peuvent se mettre en place lors de l'inhibition de KRAS (Gysin *et al.*, 2011 ; Ostrem *et al.*, 2013). L'une des approches envisagées consistait à bloquer l'ajout d'un groupement farnesyl à l'aide d'un inhibiteur de farnesyl transferase, afin d'empêcher le recrutement de KRAS à la membrane cellulaire (Sousa *et al.*, 2008). Malheureusement, à la suite de cette inhibition, un mécanisme d'échappement, impliquant l'incorporation d'un groupement géranyl, est mis en place très rapidement, rendant alors complètement inefficace l'inhibiteur en clinique (Adjei *et al.*, 2003 ; Riely *et al.*, 2011). Depuis, d'autres stratégies visant à cibler les voies de transduction de KRAS grâce à des inhibiteurs spécifiques ont été développées, comme par exemple l'inhibiteur de MEK, le selumetinib (AstraZeneca) (Garon *et al.*, 2010). Bien que prometteuses dans des modèles précliniques, ces approches pour le traitement de CBNPC présentant une mutation de KRAS n'ont pas montré de réel bénéfice pour la survie des patients, comparées à la chimiothérapie (docetaxel) (Jänne *et al.*, 2017).

Récemment, des agents capables de bloquer directement KRAS ont été identifiés (Hunter *et al.*, 2014 ; Lito *et al.*, 2016 ; Ostrem *et al.*, 2013). Par exemple, Ostrem *et al.* ont développé une molécule qui se lie irréversiblement à la protéine contenant la mutation G12C, ce qui permet d'éviter toute interférence avec la protéine sauvage (Ostrem *et al.*, 2013). Bien que ces agents soient encore à des stades précoces de développement clinique, ces études laissent présager qu'un jour de nouveaux inhibiteurs de KRAS pourront être utilisés pour le traitement des CBNPC.

3.2.2.2. Réarrangements chromosomiques de ALK

ALK est un récepteur à activité tyrosine kinase impliqué dans différents cancers. Des mutations de cet oncogène ont été décrites dans les CBNPC pour la première fois en 2007 dans un sous-groupe de patients japonais (7%), présentant une inversion sur le chromosome 2, qui donne lieu à une fusion de *ALK* avec le gène *EML4* (Soda *et al.*, 2007). *ALK* peut également faire l'objet de réarrangements chromosomiques avec d'autres gènes, tels que *KIF5B*, *KLC1*, *TFG*, *TPR*, *HIP1*, *STRN*, *DCTN1*, *SQSTM1* et *BIRC6* (Iyevleva *et al.*, 2015). Ces aberrations génétiques favorisent la dimérisation d'ALK et, par conséquent, son activation constitutive, promouvant ainsi la prolifération et la survie des cellules tumorales. Ces remaniements chromosomiques sont principalement retrouvés chez les populations jeunes qui n'ont pas ou peu fumé de tabac (Wong *et al.*, 2009).

Différentes petites molécules capables de cibler cet oncogène ont été développées. Parmi ces composés, le crizotinib (*Xalkori*, Pfizer), représente la première génération d'inhibiteurs d'ALK à avoir été approuvés en 2011 et il est maintenant devenu le traitement de référence des CBNPC au stade avancé présentant un remaniement de *ALK* (Cui *et al.*, 2011 ; Kwak *et al.*, 2010). A part ALK, cette molécule peut également inhiber c-MET et ROS1 (Cui *et al.*, 2011 ; Shaw *et al.*, 2014 ; Yasuda *et al.*, 2012). Malheureusement, la plupart des patients développe une résistance au crizotinib après un à deux ans de traitement (Shaw et Engelman, 2014). Les mécanismes moléculaires impliqués dans ce phénomène sont variés et peuvent être classés en deux catégories de résistance, respectivement dépendante ou indépendante de ALK (Duchemann *et al.*, 2015). Dans le premier cas, on retrouve des mutations secondaires du récepteur, dont les plus fréquentes sont C1156Y et L1196M, ou bien son amplification. (Toyokawa et Seto, 2015). Concernant les résistances indépendantes de ALK, elles sont issues de mécanismes qui engendrent l'activation de voies de signalisation parallèles, tels que des mutations de KRAS et EGFR, l'amplification de MET ou de KIT, ou encore une activation autocrine et/ou paracrine de l'*insulin-like growth factor 1 receptor* (IGF-1R) (Katayama *et al.*, 2012 ; Lovly *et al.*, 2014 ; Toyokawa et Seto, 2015).

Pour contrecarrer la résistance au crizotinib, une deuxième génération d'inhibiteurs d'ALK a été approuvée. Parmi ces molécules, le ceritinib (*Zykadia*, Novartis) présente une affinité 20 fois plus élevée pour le récepteur, comparé au crizotinib (Chen *et al.*, 2013 ; Friboulet *et al.*, 2014 ; Marsilje *et al.*, 2013), alors que l'alectinib (*Alecensa*, Roche) permet de cibler différentes mutations d'ALK contribuant à la résistance aux inhibiteurs de 1^{ère} génération (Iwama *et al.*, 2017 ; Perez *et al.*, 2014 ; Seto *et al.*, 2013). Ces deux stratégies thérapeutiques ont été

approuvées pour traiter les patients avec un CBNPC métastatique avec altération d'ALK, n'ayant pas répondu au crizotinib ou ayant progressé après traitement par le crizotinib.

Un autre inhibiteur d'ALK de seconde génération, le brigatinib (Alunbrig, Takeda) a été approuvé en avril 2017 pour le même type d'indication que le ceritinib et l'alectinib (Markham, 2017). Ce composé permet une survie sans progression plus longue comparé aux autres molécules, probablement grâce à sa capacité de cibler un plus large panel de mutations du récepteur acquises au cours du traitement par le crizotinib. En particulier, le brigatinib est efficace envers la mutation G1202R, la plus courante chez les patients qui progressent après un traitement avec des inhibiteurs de seconde génération (Zhang *et al.*, 2016). Face à la résistance aux différentes stratégies thérapeutiques induites par l'apparition de nouvelles mutations, une troisième génération d'inhibiteurs d'ALK est en cours d'étude clinique. Parmi ces composés, le lorlatinib est un compétiteur réversible de l'ATP qui permet de contrecarrer l'ensemble des mutations de résistance de ALK connues à ce jour (Infarinato *et al.*, 2016 ; Tucker *et al.*, 2015 ; Zou *et al.*, 2015). Cet inhibiteur est actuellement à l'essai dans une étude clinique de phase III visant à montrer sa supériorité en 1^{ère} ligne de traitement en monothérapie face au crizotinib, chez des patients porteurs de CBNPC avec mutation de ALK (NCT03052608).

3.2.2.3. Mutations de PTEN

Des mutations inactivatrices de l'oncosuppresseur *PTEN* sont retrouvées dans 4 à 8% des CBNPC et sont généralement associées à la consommation de tabac (Jin *et al.*, 2010 ; Kohno *et al.*, 1998 ; Lee *et al.*, 2010). *PTEN* est une phosphatase impliquée dans la déphosphorylation du phosphatidylinositol (3,4,5)-triphosphate (PIP₃) en phosphatidylinositol (4,5)-biphosphate (PIP₂), s'opposant donc directement à l'action de PI3K et empêchant l'activation de Akt. Ainsi, quelles que soient les mutations de *PTEN*, cela va provoquer la perte de fonction et/ou d'expression de cette phosphatase, aboutissant à une accumulation du PIP₃ au niveau de la membrane plasmique et donc à un signal constitutif de prolifération et survie cellulaire (**Fig. 19**) (Jin *et al.*, 2010). A ce jour, aucune approche thérapeutique efficace n'existe pour rétablir la fonction de *PTEN*. Une stratégie consiste à inhiber PI3K, Akt ou mTOR. Une autre piste d'investigation évoquée est le recours à la létalité synthétique, en ciblant un second gène dont la perte additionnée à celle de *PTEN* engendrerait spécifiquement la mort des cellules cancéreuses. Il a été montré que les cellules issues d'embryons de souris *Pten*^{-/-} présentent une augmentation du taux d'instabilité génomique, due à une réduction du niveau de Rad51, une protéine impliquée dans le mécanisme de réparation des coupures double brin de l'ADN par recombinaison homologue (Shen *et al.*, 2007). Mendes-Pereira *et al.* ont démontré dans des

lignées cellulaires issues de cancer colorectal que les cellules déficientes en PTEN sont particulièrement sensibles aux inhibiteurs des poly(ADP-ribose) polymérase (PARP), des enzymes activées lors des coupures simple brin de l'ADN (Mendes-Pereira *et al.*, 2009, 2012). L'inhibition de PARP empêche la réparation de ces lésions, qui se transforment alors en clivage double brin au cours de la réplication du génome. Dans les cellules contenant une délétion de PTEN, ces cassures ne peuvent pas être réparées par recombinaison homologue du fait de la diminution du niveau d'expression de Rad51, ce qui va entraîner une instabilité génétique responsable, à terme, d'une mort cellulaire. Ce type d'approche, capable de cibler spécifiquement les cellules mutantes, permettrait ainsi de combler un désert thérapeutique pour les patients porteurs de CBNPC avec perte de PTEN.

3.2.2.4. Amplification de FGFR1

Le FGFR1 est un récepteur à activité tyrosine kinase, dont les ligands appartiennent à la famille du *fibroblast growth factor*. Ce récepteur stimule différentes voies intracellulaires, dont la voie des MAPK (**Fig. 19**), et il est amplifié dans 20% des CBNPC, plus particulièrement dans les carcinomes épidermoïdes présentant des cellules squameuses (Dutt *et al.*, 2011 ; Turner et Seckl, 2010 ; Weiss *et al.*, 2010). Les cellules présentant une amplification de FGFR1 sont dépendantes de ce récepteur pour leur croissance (Weiss *et al.*, 2010), constituant ainsi une cible thérapeutique idéale. Différents inhibiteurs spécifiques de FGFR ont été développés (par exemple le AZD4547 et le BGJ398), dont l'activité et le bénéfice chez les patients sont en cours d'évaluation dans des essais cliniques (pour revue Desai et Adjei, 2016). En parallèle, des inhibiteurs multikinases capables de cibler non seulement FGFR, mais également d'autres RTKs, tels que le *vascular endothelial growth factor receptor* (VEGFR) ou *platelet-derived growth factor receptor* (PDGFR), ont été générés. Parmi ces inhibiteurs, le brivatinib, le dovitinib, le lenvatinib et le nintedanib ont été étudiés dans différents types de cancer pour leur propriété anti-angiogénique et sont maintenant évalués en clinique pour leur capacité à cibler le FGFR (Desai et Adjei, 2016). Par exemple, une étude parue en 2016 révèle une faible efficacité du dovitinib (inhibiteur de FGFR, PDGFR, VEGFR, FLT3 et c-KIT) dans les carcinomes bronchiques à grandes cellules de stade avancé contenant une amplification de FGFR1. Ces résultats nécessitent des études complémentaires pour identifier les biomarqueurs qui pourraient être corrélés avec une efficacité de ce composé (Lim *et al.*, 2016).

3.2.2.5. Altérations de MET

Met est un proto-oncogène qui code un récepteur à activité tyrosine kinase du même nom. L'interaction de MET avec son ligand, l'*hepatocyte growth factor* (HGF) favorise sa dimérisation et son autophosphorylation sur deux résidus tyrosine situés dans son domaine catalytique (Y1234 et Y1235) (Gherardi *et al.*, 2012 ; Longati *et al.*, 1994). D'autres résidus tyrosines localisés en dehors du domaine catalytique sont par la suite phosphorylés, comme la tyrosine 1003 du domaine juxta-membranaire ou les tyrosines 1349 et 1356 de l'extrémité C-terminale (Gherardi *et al.*, 2012). Les phosphorylations sur les résidus en C-terminal entraînent le recrutement de protéines effectrices, telles que Grb2, *Grb2-associated-binding protein 1* (GAB1), la phospholipase C-gamma (PLC γ), SRC, *SRC homology 2 domain containing phosphatase 2* (SHP2), PI3K et *signal transducer et activator of transcription 3* (STAT3). La diversité des protéines recrutées par MET permet de réguler un large panel de voies de signalisation, incluant MAPK, PI3K-AKT et STAT (**Fig. 19**) (Gherardi *et al.*, 2012). Selon le contexte cellulaire, ces voies peuvent mener à une augmentation de la prolifération, de la survie, de la motilité ou de l'invasion cellulaire, et notamment à une transition épithélio-mésenchymateuse, lors des processus d'embryogenèse, d'organogenèse et de cicatrisation.

Différents mécanismes responsables d'une activation aberrante de MET ont été décrits dans les cancers du poumon. Approximativement 3% à 4% des patients porteurs de CBNPC présentent une amplification de ce récepteur, et certaines molécules ont montré une certaine efficacité, tels que l'inhibiteur de ALK crizotinib, qui cible également MET, et, dans une moindre mesure, le tivantinib (Camidge *et al.*, 2014 ; Kong-Beltran *et al.*, 2006 ; Onozato *et al.*, 2009 ; Scagliotti *et al.*, 2015, 2012). Ainsi, dans un essai clinique de phase I (NCT00585195), il a été montré que le crizotinib induit une réponse chez un patient sur six (16.6%) avec amplification modérée de MET, et chez trois patients sur six (50%) avec amplification élevée (plus de 10 copies) de ce récepteur (Camidge *et al.*, 2014). Environ 3-10% des CBNPC présentent des mutations non-sens de l'exon 14 de *MET*, telles que R970C, P991S et T992I, qui favorisent la croissance cellulaire par un mécanisme encore non élucidé (Cortot *et al.*, 2017). D'autres altérations génétiques localisées dans les sites d'épissage en amont et en aval de l'exon 14 ont été détectés dans 1-8% des tumeurs bronchiques (Schrock *et al.*, 2016). Ces aberrations génétiques peuvent correspondre à des mutations ponctuelles, des délétions, des insertions ou bien des mutations complexes (*indels*) et sont responsables d'un défaut d'épissage conduisant à un saut de l'exon 14, codant le domaine juxta-membranaire de MET. La délétion de ce domaine entraîne donc la perte de la tyrosine 1003, qui sert de site de liaison

à CBL, une ubiquitine ligase E3 impliquée dans le renouvellement de MET (Kong-Beltran *et al.*, 2006). La perte de l'exon 14 résulte ainsi en une diminution de la dégradation de MET et en une stimulation constante des voies de transduction de cette protéine (Cortot *et al.*, 2017 ; Kong-Beltran *et al.*, 2006 ; Krishnaswamy *et al.*, 2009 ; Onozato *et al.*, 2009 ; Tyner *et al.*, 2010). Selon une étude récente, le crizotinib pourrait être efficace pour le traitement de tumeurs présentant une délétion de l'exon 14 de MET (Lu *et al.*, 2017). Parallèlement à cette étude, différents essais cliniques de phase I ou II utilisant des inhibiteurs spécifiques de MET, tels que le capmatinib (NCT02750215), le tepotinib (NCT02864992) ou des inhibiteurs multikinases, tels que le cabozantinib, sont en cours afin d'évaluer le bénéfice de ces traitements pour les patients (<http://clinicaltrials.gov>).

3.2.2.6. Mutations de l'EGFR

La famille des récepteurs au facteur de croissance épidermique (EGFR) a été beaucoup étudiée dans les cancers solides. Cette famille de glycoprotéines transmembranaires est composée de quatre membres : EGFR (ErbB1, HER1), ErbB2 (HER2), ErbB3 (HER3) et ErbB4 (HER4). Parmi ces récepteurs à activité tyrosine kinase, deux ne sont pas autonomes, puisque ErbB2 n'a pas la capacité d'interagir avec un ligand, tandis qu'ErbB3 ne possède pas d'activité kinase. Ces protéines doivent donc former un hétérodimère avec un autre membre de la famille pour fonctionner (Citri et Yarden, 2006). L'activation de l'EGFR se fait *via* la liaison à un ligand, dont les principaux sont l'EGF et le *transforming growth factor- α* (TGF- α). La fixation du ligand entraîne un changement de conformation des récepteurs qui provoque la formation d'homodimères ou hétérodimères, suivie par une transphosphorylation sur des tyrosines. Ces résidus phosphorylés permettent la fixation de protéines cytoplasmiques contenant des domaines *phosphotyrosine binding* (PTB) et *Src homology 2* (SH2) qui vont induire l'activation de différentes voies intracellulaires. Par exemple, les protéines adaptatrices Grb2 et *son of sevenless* (Sos) sont responsables de la stimulation de RAS et de l'activation de la voie MAPK. Certaines tyrosines phosphorylées peuvent également entraîner le recrutement de PI3K, responsable de la phosphorylation du *phosphatidylinositol-4,5 biphosphate* (PIP2) en *phosphatidylinositol (3,4,5) triphosphate* (PIP3). Les PIP3 ainsi générés permettent l'adressage membranaire et l'activation d'AKT. L'EGFR peut également recruter les *Janus kinases* (JAK), qui vont phosphoryler et activer les protéines *signal transducer et activator of transcription 3 et 5* (STAT3/5), des facteurs de transcription qui régulent l'expression de gènes impliqués dans la prolifération, survie et invasion cellulaire (**Fig. 19**) (Chong et Jänne, 2013 ; Krause et Van Etten, 2005).

L'activation constitutive de la voie de l'EGFR, provoquée par des altérations génétiques de ce récepteur, s'est avérée étroitement liée à l'initiation et à la progression du CBNPC. Alors que seulement 10% des patients caucasiens présentent des mutations d'EGFR, la proportion augmente jusqu'à 50% pour les ethnies est-asiatiques. De plus, ce type d'aberration génétique est plus fréquent chez les patients porteurs d'adénocarcinome, les non-fumeurs et les femmes (Fukuoka *et al.*, 2003 ; Kris *et al.*, 2003 ; Miller *et al.*, 2004). Les deux mutations d'EGFR les plus couramment retrouvées dans les cancers du poumon sont des petites délétions dans l'exon 19 (en particulier E746-A750del) et la substitution L858R dans l'exon 21, qui ensemble constituent plus de 90% des cas (**Fig. 20A**) (Chong et Jänne, 2013 ; Lynch *et al.*, 2004 ; Paez *et al.*, 2004 ; Pao *et al.*, 2004). Des mutations dans l'exon 18 (G719X) ou 21 (L861Q), moins fréquentes que les précédentes, sont également retrouvées dans les CBNPC (Pao et Chmielecki, 2010). Ces substitutions dans le domaine catalytique de l'EGFR déstabilisent la conformation inactive du récepteur et entraînent son activation constitutive, indépendamment de la liaison de ligands (Dixit et Verkhivker, 2014 ; Shan *et al.*, 2012).

Les CBNPC présentant des mutations d'EGFR dépendent, pour leur prolifération et leur survie, de l'activation de voies de signalisation en aval de ce récepteur, représentant ainsi un exemple typique d'addiction oncogénique.

3.3. Thérapie ciblée pour les CBNPCs avec mutation d'EGFR

3.3.1. Inhibiteurs de 1^{ère} génération

En 2003 et 2004, les inhibiteurs d'EGFR de 1^{ère} génération gefitinib (Iressa ; AstraZeneca) et erlotinib (Tarceva ; Genentech/OSI Pharmaceuticals) ont été approuvés pour le traitement des CBNPC de stade avancé. Ces inhibiteurs compétitifs d'ATP se sont avérés particulièrement efficaces pour les tumeurs présentant des mutations activatrices d'EGFR (Lynch *et al.*, 2004 ; Paez *et al.*, 2004 ; Pao *et al.*, 2004). Il a été montré que ces mutations réduisent l'affinité du domaine catalytique pour l'ATP (**Fig. 20A**) (Cortot et Jänne, 2014 ; Yun *et al.*, 2008), ce qui permet au gefitinib et à l'erlotinib d'inhiber plus efficacement la forme mutée du récepteur, limitant ainsi les effets secondaires résultant d'une perturbation de la voie EGFR dans les tissus non tumoraux. Par la suite, six différents essais cliniques de phase III démontrèrent que, comparés à la chimiothérapie classique, l'erlotinib et le gefitinib pouvaient induire une augmentation significative de la survie sans progression chez les patients dont la tumeur métastatique présentait des mutations d'EGFR (Han *et al.*, 2012 ; Maemondo *et al.*, 2010 ; Mitsudomi *et al.*, 2010 ; Mok *et al.*, 2009 ; Rosell *et al.*, 2012 ; Zhou *et al.*, 2011). Ces études ont mené à l'approbation en 2013 et 2015 de l'erlotinib et du gefitinib en tant que traitement de

première ligne pour ce type de cancer. Par conséquent, les tumeurs des patients sont systématiquement séquencées afin de rechercher la présence de mutations d'EGFR et prescrire ainsi la stratégie thérapeutique la mieux adaptée.

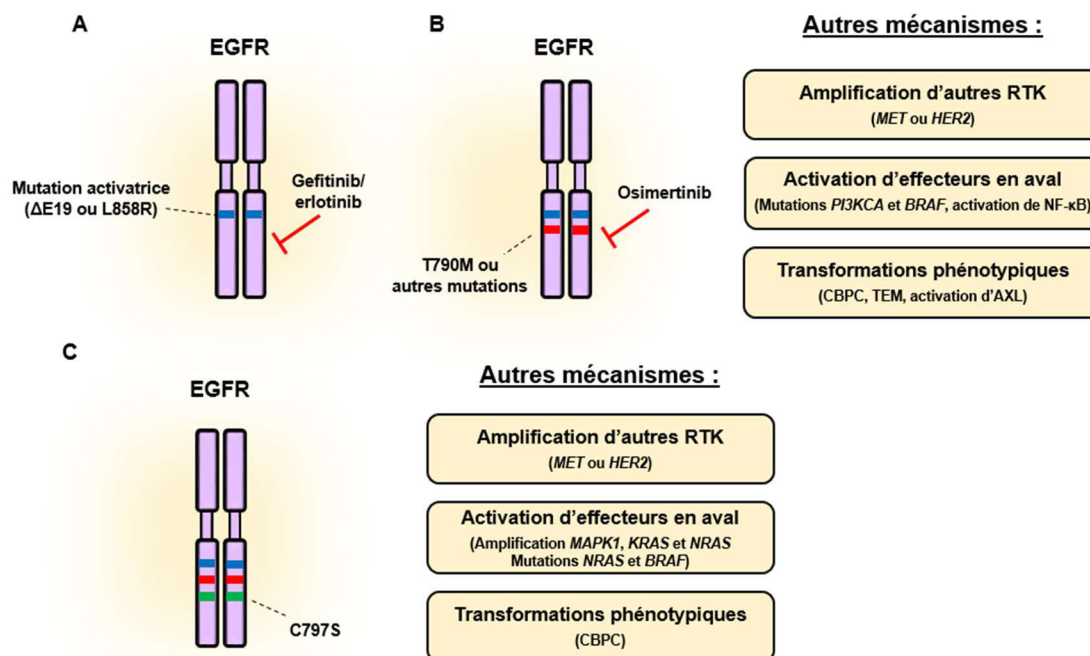


Figure 20. Mécanismes de résistance aux inhibiteurs d'EGFR dans le CBNPC

A. Une délétion de quelques nucléotides dans l'exon 19 ou une mutation dans l'exon 20 dans le domaine catalytique entraîne une activation constitutive de l'EGFR, qui peut être bloquée par les inhibiteurs de 1^{ère} génération, tels que le gefitinib ou l'erlotinib. **B.** A la suite d'un traitement avec un inhibiteur d'EGFR de 1^{ère} génération, différents mécanismes de résistance peuvent apparaître, tels qu'une mutation secondaire de l'EGFR (généralement la T790M dans le domaine catalytique) l'amplification d'un autre RTK, l'activation d'effecteurs en aval ou encore des transformations phénotypiques. La résistance induite par l'apparition de la mutation T790M peut être contrecarrée par l'utilisation d'un inhibiteur de 3^{ème} génération, l'osimertinib. **C.** Après quelques mois de traitement avec l'osimertinib, des mécanismes d'échappement se mettent en place, parmi lesquels le plus fréquent correspond à une mutation tertiaire d'EGFR (C797S). CBPC, cancer bronchique à petites cellules ; TEM, transition épithélio-mésenchymateuse.

3.3.2. Résistances aux inhibiteurs d'EGFR de 1^{ère} génération

Deux types de résistance aux inhibiteurs d'EGFR peuvent être observés. La résistance primaire, ou innée, correspond aux patients qui, malgré la présence de mutations activatrices du récepteur, ne montrent aucun signe de réponse à ces thérapies. La résistance acquise, ou secondaire, fait suite à une réponse initiale, et dépend de l'émergence de nouvelles mutations ou transitions phénotypiques des cellules tumorales.

3.3.2.1. Résistance primaire

La majorité des patients de CBNPC avec EGFR muté répondent aux inhibiteurs de ce récepteur. Toutefois, dans 4 à 10% des cas, ces tumeurs présentent une résistance primaire provoquée par différents types de mécanismes. Il a été suggéré que la consommation de tabac pourrait entraîner une activation du cytochrome CYP 1A1, impliqué dans le métabolisme des

inhibiteurs de l'EGFR (Alfieri *et al.*, 2011), ainsi que l'augmentation du niveau d'espèces réactives de l'oxygène, qui peut engendrer une inhibition de la dégradation de l'EGFR (Filosto *et al.*, 2011, 2012) et une activation de Src (Filosto *et al.*, 2013). D'autres auteurs ont montré que certains polymorphismes décrits chez les populations asiatiques, entraînant un épissage alternatif de la protéine pro-apoptotique BIM, sont responsables d'une plus faible sensibilité aux inhibiteurs d'EGFR (Ng *et al.*, 2012).

Différentes protéines en aval du récepteur peuvent également être mutées dans les CBNPC. Des mutations somatiques activatrices de PIK3CA, qui code pour la sous-unité catalytique de PI3K, ont été détectées dans 1-3% des CBNPC, parfois en association avec des mutations de l'EGFR (Kawano *et al.*, 2006). Une perte d'expression de PTEN (<5%) a aussi été associée à une diminution de la sensibilité aux thérapies ciblant l'EGFR (Sos *et al.*, 2009).

Les tumeurs ne présentant pas de mutations de l'EGFR et dont la survie et la croissance dépendent de mutations d'autres oncogènes sont généralement insensibles aux inhibiteurs de l'EGFR. En particulier, les mutations activatrices des codons 12 et 13 de KRAS sont considérées comme un critère prédictif négatif pour la réponse des CBNPC à ce type de traitement (Linardou *et al.*, 2008 ; Pao *et al.*, 2005a). A l'instar de KRAS, d'autres mutations peuvent également engendrer une résistance primaire envers ces inhibiteurs, telles que des mutations de BRAF (Brose *et al.*, 2002 ; Naoki *et al.*, 2002) ou l'expression d'une protéine de fusion EML4-ALK (Shaw *et al.*, 2009 ; Soda *et al.*, 2007).

3.3.2.2. Résistance acquise

Malgré une réponse initiale, la quasi-totalité des patients porteurs de CBNPC présentant une activation d'EGFR rechutent suite au développement d'une résistance envers les inhibiteurs de ce récepteur dans les 10 à 12 mois suivant le début du traitement. Différentes études à partir de biopsies de patients ont montré que, dans plus de la moitié des cas, cette résistance apparaît suite à l'acquisition d'une seconde mutation, la T790M, dans le domaine catalytique du récepteur (**Fig. 20B**) (Arcila *et al.*, 2011 ; Sequist *et al.*, 2011 ; Yu *et al.*, 2013). Cette substitution augmente l'affinité du récepteur pour l'ATP, rendant alors l'EGFR moins sensible aux inhibiteurs compétitifs, tels que le gefitinib ou l'erlotinib (Kobayashi *et al.*, 2005 ; Pao *et al.*, 2005b). Ainsi, la fenêtre thérapeutique ouverte par la mutation L858R est refermée par la seconde altération T790M, qui restaure un niveau d'affinité pour l'ATP similaire à celui du récepteur sauvage (Cortot et Jänne, 2014 ; Yun *et al.*, 2008).

D'autres altérations d'EGFR ont été décrites comme conférant une résistance aux inhibiteurs de 1^{ère} génération, telles que les mutations L747S (Costa *et al.*, 2008) et D761Y (Balak *et al.*, 2006) dans l'exon 19 et la T854A dans l'exon 21 (Bean *et al.*, 2008). Ces substitutions, beaucoup plus rares comparées à la T790M, sont localisées à proximité du site de liaison à l'ATP, et elles pourraient donc affecter l'affinité de ces inhibiteurs pour le récepteur.

D'autres mécanismes de résistance ont été décrits, notamment l'activation d'autres RTKs ou d'effecteurs plus en aval, ou des transformations phénotypiques (**Fig. 20B**).

Un exemple typique d'activation d'autres récepteurs est représenté par l'amplification du proto-oncogène *MET*. Il a été montré que, en formant des hétérodimères avec ErbB3, *MET* entraîne l'activation de la voie de signalisation PI3K/AKT, permettant ainsi de contourner les effets des inhibiteurs d'EGFR (Engelman *et al.*, 2007 ; Sequist *et al.*, 2011). Selon les études, ce mécanisme de résistance a été identifié dans 5-22% des CBNPC qui ont progressé à la suite d'un traitement avec l'erlotinib ou le gefitinib (Cortot et Jänne, 2014 ; Engelman *et al.*, 2007 ; Sequist *et al.*, 2011 ; Yu *et al.*, 2013). De façon similaire, une amplification de *HER2*, mutuellement exclusive de la mutation T790M, a également été associée à une résistance acquise à ces inhibiteurs (Takezawa *et al.*, 2012).

Des mutations d'effecteurs en aval de la voie EGFR, tels que *PIK3CA* et *BRAF*, peuvent promouvoir la résistance de ces tumeurs, mais elles sont rarement observées chez les patients (Ohashi *et al.*, 2012 ; Sequist *et al.*, 2011). Parmi les autres voies de signalisation qui ont été impliquées dans l'acquisition d'une résistance, l'on peut citer NF- κ B. Une étude menée par Bivona *et al.* a montré que la répression par ARN interférence de certains effecteurs de cette voie augmente la sensibilité aux inhibiteurs d'EGFR. Ces auteurs ont également identifié une corrélation significative entre l'activation de NF- κ B et un faible taux de survie chez les patients de CBNPC avec mutation d'EGFR (Bivona *et al.*, 2011).

En réponse à un traitement avec des inhibiteurs d'EGFR, les CBNPC, généralement des adénocarcinomes, peuvent subir une transformation cellulaire amenant à requalifier le type de cancer en CBPC (Arcila *et al.*, 2011 ; Sequist *et al.*, 2011 ; Yu *et al.*, 2013). Ces tumeurs conservent les mutations d'EGFR initiales et répondent aux chimiothérapies utilisées pour les CBPC. Il est à noter que dans ces études aucune cellule de type CBPC n'avait été détectée dans les tumeurs avant le traitement par un inhibiteur d'EGFR. Dans d'autres cas, une résistance des CBNPC à l'erlotinib/gefitinib a été associée à une transition épithélio-mésenchymateuse (Sequist *et al.*, 2011 ; Uramoto *et al.*, 2010). Les cellules, à l'origine de type épithélial, acquièrent ainsi des caractéristiques mésenchymateuses, telles qu'une augmentation des

capacités invasives, une perte de l'E-cadhérine et l'expression de la vimentine et de la N-cadherine. Les mécanismes menant à cette transformation n'ont pas été complètement caractérisés, mais ils pourraient impliquer une surexpression de NOTCH-1 et du TGB- β et/ou l'activation de MET (La Monica *et al.*, 2013 ; Xie *et al.*, 2012). La surexpression du RTK d'AXL, observée dans environ 20% des CBNPC et parfois accompagnée d'un taux élevé de son ligand GAS6, a également été proposée comme mécanisme responsable d'une résistance aux inhibiteurs d'EGFR, *via* l'induction d'une transition épithélio-mésenchymateuse (Zhang *et al.*, 2012).

3.3.3. Nouveaux inhibiteurs capables de cibler la mutation T790M

Dans le but de contourner la résistance induite par la mutation T790M, des nouveaux inhibiteurs capables de se lier de façon covalente à l'EGFR ont été développés. L'une de ces molécules de 2^{ème} génération, l'afatinib (*Giotrif* ; Boehringer Ingelheim), a été approuvée pour la prise en charge de CBNPC présentant des mutations activatrices d'EGFR, même si dans les essais clinique il s'est révélé peu efficace chez les patients porteurs de la T790M (Miller *et al.*, 2012).

Plus récemment, d'autres molécules qui se fixent de façon irréversible au résidu C797 dans la poche de liaison à l'ATP du domaine catalytique de l'EGFR ont été générés (**Fig. 20B**). Des essais cliniques publiés en 2015 ont démontré l'efficacité de l'un de ces inhibiteurs, le rociletinib (*Xegafri* ; Clovis Oncology), avec un taux de réponse objectif de 59% chez les patients porteurs de CBNPC présentant la mutation T790M (Sequist *et al.*, 2015). Cependant, au vu de certains effets secondaires majeurs (Tran et Klemperer, 2016), les comités d'éthiques ont souhaité des résultats complémentaires avant d'approuver la mise sur le marché, demande à laquelle la compagnie n'a pas donné suite. Dans un autre essai clinique publié en 2015, l'inhibiteur d'EGFR de 3^{ème} génération osimertinib (*Tagrisso* ; AstraZeneca) s'est également avéré efficace chez les patients de CBNPC présentant la mutation T790M (Jänne *et al.*, 2015). La même année, l'osimertinib a été approuvé pour le traitement de tumeurs ayant développé une résistance aux inhibiteurs d'EGFR de 1^{ère} génération, et des essais sont en cours pour évaluer son efficacité en première ligne chez les patients de CBNPC présentant une mutation activatrice de ce récepteur.

3.3.4. Développement d'une résistance aux inhibiteurs d'EGFR de 3^{ème} génération

Malgré de bons résultats pour la survie et des taux de réponse très satisfaisants chez les patients traités avec l'osimertinib, les tumeurs développent inévitablement une résistance après quelques mois de traitement. Le mécanisme le plus fréquent décrit à ce jour correspond à une substitution de la cystéine 797 par une sérine (C797S), qui apparaît après 6 à 17 mois de traitement (Thress *et al.*, 2015). Cette mutation empêche la fixation de l'osimertinib au récepteur (**Fig. 20C**) conduisant à une insensibilité de ces tumeurs envers toutes les thérapies ciblant l'EGFR actuellement disponibles (Ercan *et al.*, 2015 ; Thress *et al.*, 2015).

D'autres mécanismes alternatifs de résistance ont été identifiés (**Fig. 20C**), tels que les amplifications de *MET* et *HER2*, la transformation en CBPC (Ham *et al.*, 2016 ; Planchard *et al.*, 2015), l'amplification de *MAPK1*, la mutation ou l'amplification de *NRAS*, la mutation de *BRAF* ou l'amplification de *KRAS* (Eberlein *et al.*, 2015 ; Ercan *et al.*, 2012 ; Planchard *et al.*, 2015). Une étude récente décrivant les effets de l'osimertinib en première ligne pour le traitement de cancer bronchique avec mutation activatrice de l'EGFR (Ramalingam *et al.*, 2017) a révélé que, malgré une survie sans progression d'environ 22 mois, plus de la moitié des patients développent une résistance associée à des mécanismes différents, tels que l'amplification de *MET*, l'amplification simultanée de *KRAS* et *EGFR*, des mutations de *PIK3CA*, *MEK1*, *JAK2* ou *KRAS*, la mutation C797S de l'EGFR, ou encore une insertion dans l'exon 20 de *HER2*. En accord avec les résultats précliniques, aucun des patients n'a présenté une émergence de la mutation T790M au cours du traitement, témoignant ainsi de la capacité des inhibiteurs de 1^{ère} et de 3^{ème} génération à sélectionner des mécanismes de résistance distincts.

OBJECTIFS DE LA THESE

Les tumeurs sont constituées de différentes populations sous-clonales génétiquement hétérogènes, responsables en grande partie de la capacité des cellules cancéreuses à évoluer rapidement et à s'adapter aux conditions environnementales. Dans l'introduction de cette thèse, nous avons rappelé qu'à l'heure actuelle seules des modélisations informatiques sont proposées pour décrire l'architecture clonale d'une tumeur donnée. En revanche, aucune approche expérimentale capable de récapituler l'émergence de cellules cancéreuses génétiquement distinctes n'est disponible. De ce fait, dans la plupart des cas, la recherche et la validation de nouvelles cibles thérapeutiques sont donc basées sur des modèles expérimentaux qui ne prennent pas en considération l'hétérogénéité intratumorale. Afin de palier à cette limitation, notre travail a consisté, dans un premier temps, à développer une nouvelle approche expérimentale pour l'étude de mutations oncogéniques dans un contexte d'hétérogénéité intratumorale, que nous avons appelée CRISPR-barcoding. Cette étude, qui a fait l'objet d'une publication dans *Molecular Cell*, constitue la première partie du chapitre Résultats.

Dans un second temps, nous avons appliqué cette nouvelle approche à l'étude des résistances des cancers bronchiques aux inhibiteurs d'EGFR, qui apparaissent inévitablement après plusieurs mois de traitement. Dans ce contexte, nous avons effectué un criblage à l'aide de CRISPR-barcoding qui a permis de mettre en évidence que le sorafénib, un inhibiteur multikinase, était capable de prévenir la résistance des cellules de cancer bronchique aux inhibiteurs d'EGFR. Nous avons ensuite entrepris la caractérisation du mécanisme d'action de cette molécule, qui pourrait représenter un nouvel outil thérapeutique pour le traitement de ces tumeurs. Cette deuxième étude, détaillée dans la seconde partie du chapitre Résultats, a donné lieu au dépôt d'une demande de brevet européen.

RESULTATS

1. Travaux publiés

CRISPR-barcoding for Intratumor Genetic Heterogeneity Modeling and Functional Analysis of Oncogenic Driver Mutations

Alexis Guernet, Sathish Kumar Mungamuri, Dorthe Cartier, Ravi Sachidanandam, Anitha Jayaprakash, Sahil Adriouch, Myriam Vezain, Françoise Charbonnier, Guy Rohkin, Sophie Coutant, Shen Yao, Hassan Ainani, David Alexandre, Isabelle Tournier, Olivier Boyer, Stuart A. Aaronson, Youssef Anouar & Luca Grumolato

Molecular Cell (2016), 63, 526-538.

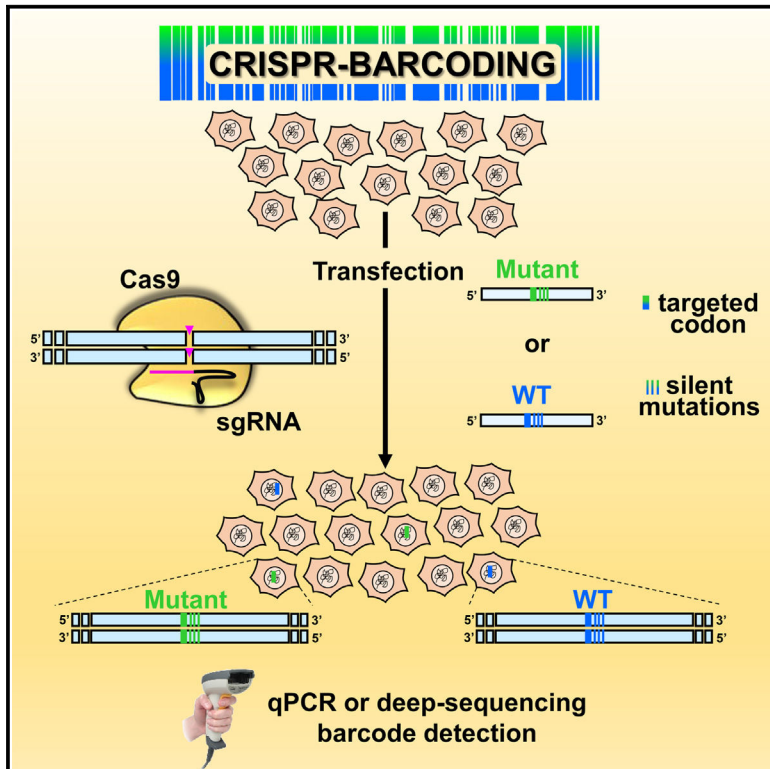
Dans cet article, nous décrivons le développement d'une nouvelle approche, appelée CRISPR-barcoding, permettant de générer et de suivre l'émergence d'un petit groupe de cellules contenant une mutation voulue au sein d'une population de cellules non modifiées. Nous avons démontré que cette stratégie, qui permet de développer de nouveaux modèles expérimentaux d'hétérogénéité intratumorale, peut être utilisée pour diverses applications, telles que :

- L'insertion de nouvelles mutations oncogéniques
- La réparation de mutations d'oncosuppresseurs ou d'oncogènes,
- L'étude des mécanismes de résistance du cancer à la thérapie ciblée,
- L'étiquetage et le suivi de plusieurs milliers de clones au sein d'une même population de cellules tumorales.

Molecular Cell

CRISPR-Barcoding for Intratumor Genetic Heterogeneity Modeling and Functional Analysis of Oncogenic Driver Mutations

Graphical Abstract



Authors

Alexis Guernet,
Sathish Kumar Mungamuri,
Dorthe Cartier, ..., Stuart A. Aaronson,
Youssef Anouar, Luca Grumolato

Correspondence

luca.grumolato@univ-rouen.fr

In Brief

Guernet et al. used CRISPR/Cas9 technology to genetically label and trace tumor cells within a mass population through insertion of a series of point mutations at a specific genomic location. The effects of a particular experimental condition on the fraction of barcoded cells can be assessed by qPCR or deep sequencing.

Highlights

- CRISPR-barcoding is a versatile method to trace clonal dynamics in tumor populations
- Sensitive barcode detection allows DNA editing even when HDR efficiency is low
- A silent barcode is used as a control for potential Cas9 off-target cleavage
- Highly complex barcodes are used to label thousands of cells within a mass population



CRISPR-Barcoding for Intratumor Genetic Heterogeneity Modeling and Functional Analysis of Oncogenic Driver Mutations

Alexis Guernet,^{1,2} Sathish Kumar Mungamuri,³ Dorthe Cartier,^{1,2} Ravi Sachidanandam,³ Anitha Jayaprakash,³ Sahil Adriouch,^{2,4} Myriam Vezain,^{2,5} Françoise Charbonnier,^{2,5} Guy Rohkin,³ Sophie Coutant,^{2,5} Shen Yao,³ Hassan Ainani,^{1,2} David Alexandre,^{1,2} Isabelle Tournier,^{2,5} Olivier Boyer,^{2,4,6} Stuart A. Aaronson,³ Youssef Anouar,^{1,2} and Luca Grumolato^{1,2,*}

¹Normandie Univ, UNIROUEN, INSERM, DC2N, 76000 Rouen, France

²Institute for Research and Innovation in Biomedicine (IRIB), 76000 Rouen, France

³Department of Oncological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

⁴IRIB Flow Cytometry and Cell Sorting Facility, 76000 Rouen, France

⁵IRIB Genomics Facility, 76000 Rouen, France

⁶Department of Immunology, CHU Rouen, 76000 Rouen, France

*Correspondence: luca.grumolato@univ-rouen.fr

<http://dx.doi.org/10.1016/j.molcel.2016.06.017>

SUMMARY

Intratumor genetic heterogeneity underlies the ability of tumors to evolve and adapt to different environmental conditions. Using CRISPR/Cas9 technology and specific DNA barcodes, we devised a strategy to recapitulate and trace the emergence of subpopulations of cancer cells containing a mutation of interest. We used this approach to model different mechanisms of lung cancer cell resistance to EGFR inhibitors and to assess effects of combined drug therapies. By overcoming intrinsic limitations of current approaches, CRISPR-barcoding also enables investigation of most types of genetic modifications, including repair of oncogenic driver mutations. Finally, we used highly complex barcodes inserted at a specific genome location as a means of simultaneously tracing the fates of many thousands of genetically labeled cancer cells. CRISPR-barcoding is a straightforward and highly flexible method that should greatly facilitate the functional investigation of specific mutations, in a context that closely mimics the complexity of cancer.

INTRODUCTION

Staggering advances in sequencing technologies have provided a detailed overview of the multiple genetic aberrations in cancer and also demonstrated that, within an individual tumor, such mutations are often distributed in intricate patterns of multiple and heterogeneous subclonal populations (Gerlinger et al., 2012; McGranahan and Swanton, 2015). This complex genetic reservoir fuels the capacity of tumor cells to adapt to various environmental conditions, with major clinical implications for cancer progression and resistance to therapeutic intervention (Gillies

et al., 2012; Greaves and Maley, 2012; McGranahan and Swanton, 2015). Although the ability of tumors to adapt and evolve through the emergence of clonal subpopulations containing de novo mutations has been known for decades (Nowell, 1976), tumors are generally treated as genetically homogeneous entities (Clevers, 2011). Computational modeling based on analysis of deep-sequencing data sets, including single-cell sequencing, now provides the means to dissect the clonal architecture of a tumor through the identification of its different cell subpopulations (Ding et al., 2014; McGranahan and Swanton, 2015). However, although these new powerful technologies can capture a picture of the genetic complexity and diversity of a given tumor, they cannot experimentally recapitulate cancer evolution through the emergence of new mutations.

CRISPR (clustered regularly interspaced short palindromic repeats) is a new DNA editing technology based on the Cas9 nuclease from *S. pyogenes* and a single-guide RNA (sgRNA), which as a complex can specifically recognize and cleave a genomic sequence of interest (Doudna and Charpentier, 2014; Hsu et al., 2014). The double-strand DNA break induced by CRISPR/Cas9 can trigger two distinct cellular mechanisms for DNA repair: error-prone nonhomologous end-joining (NHEJ) and high-fidelity homology-directed repair (HDR). While NHEJ-induced indels can be exploited to inactivate a gene of interest, HDR enables precise DNA editing (Doudna and Charpentier, 2014; Hsu et al., 2014). Despite its tremendous potential, widespread use of HDR has been curbed by intrinsic limitations, mostly related to its low efficiency and the necessity to derive clones.

By turning the low efficiency of HDR-mediated DNA editing to our advantage, we devised a strategy to recapitulate and trace intratumor heterogeneity and the emergence of genetically distinct cancer subpopulations on the basis of silent DNA barcodes coupled to a desired mutation in the sequence of a gene of interest. These genetic labels can then be “read” by real-time qPCR from genomic DNA (gDNA) to measure the relative proportion of the modified cells within an unmodified mass population. We used this CRISPR-barcoding strategy to model

different mechanisms of non-small-cell lung cancer (NSCLC) resistance to epidermal growth factor receptor (EGFR) inhibitors, and we established a multiplex system to evaluate the efficacy of combined drug therapies aimed at preventing or delaying the emergence of resistant cells. Through a similar approach, we assessed for the first time the effects of repairing oncogenic driver mutations in addicted cancer cells directly at the genome level. Finally, we used a highly complex set of CRISPR-barcodes as a means to simultaneously label several thousand different breast and lung cancer cells, and we compared their relative fitness to grow *in vivo* after inoculation in immunodeficient mice or *in vitro* upon treatment with a targeted therapeutic agent.

DESIGN

Compared with other DNA editing tools, including zinc finger nucleases and transcription activator-like effector nucleases, which are based on protein-DNA recognition, the CRISPR/Cas9 technology is remarkably more flexible and easy to use (Boettcher and McManus, 2015). However, despite the undeniable potential of this new technology, major intrinsic limitations need to be considered when applying CRISPR/Cas9 for HDR genome manipulation of cultured cells. First, in certain contexts this system can tolerate a few mismatches between the sgRNA and its target sequence, which can result in off-target DNA cleavage (Doudna and Charpentier, 2014; Fu et al., 2013; Hsu et al., 2014). Although tools have been designed to identify *in silico* potential off-target sequences, occasional minor shifts in the secondary structure of the RNA-DNA duplex that could make such predictions more problematic have been recently reported (Lin et al., 2014). Another limitation is related to the fact that the efficiency of HDR-mediated DNA editing can be extremely low, depending on the cell model, the sgRNA sequence, and the targeted DNA. Thus, in a given experiment, only a fraction of cells within a population will contain the desired genetic modification. Hence, the derivation and the subsequent analysis of a certain number of individual clones constitute an almost inevitable step (Ran et al., 2013). Besides the potential issues related to clonal variability, this approach requires a reasonably high efficiency of CRISPR-mediated editing, and it is obviously not compatible with modifications that have a negative impact on cell growth.

To overcome these drawbacks, we have devised CRISPR-barcoding, a new strategy in which a potentially functional modification in the sequence of a gene of interest is coupled with a series of silent point mutations, serving as a genetic label for cell tracing. In parallel, a second barcode consisting of distinct silent mutations is inserted in the same cell population and used as a control for possible CRISPR off-target effects. gDNA from the resulting mixture of CRISPR-modified and unmodified cells is then probed using qPCR to assess the relative proportion of each barcode. By exposing the cells to a given selective condition, this approach can be used to functionally characterize the effects of different types of mutations of a particular gene of interest. We applied this strategy to manipulate the endogenous sequence of different oncogenes and tumor suppressors, including *EGFR*, *KRAS*, anaplastic lymphoma kinase (*ALK*),

TP53, and adenomatous polyposis coli (*APC*) in various cancer cell models. We then assessed the effects of such modifications on signaling pathway activation, cell growth and invasion, or resistance to chemotherapy, both *in vitro* and *in vivo*. Through a similar approach, we also labeled breast cancer cells using a degenerate barcode inserted at a specific location of their genome, to investigate the contribution of cell heterogeneity to tumor formation. By overcoming the limitations associated with the low efficiency and potential off-target effects of DNA editing, our work demonstrates that CRISPR-barcoding can be easily implemented in functional studies to investigate the effects of a specific genetic modification.

RESULTS

Modeling Drug Resistance in NSCLC Cells

A major limitation of cancer-targeted therapies is the almost inevitable development of acquired resistance, as a result of the emergence of a subpopulation of cancer cells that become insensitive to the treatment, generally through particular genetic aberrations. Although such mutations could conceivably originate *de novo* during the treatment, recent evidence from different types of tumors indicates that resistant clones are often already present before the onset of the therapy (Bhang et al., 2015; Diaz et al., 2012; Misale et al., 2012; Turke et al., 2010).

As a typical example of drug resistance, gefitinib is a small-molecule ATP-competitive reversible EGFR inhibitor used in the clinic for the treatment of advanced NSCLC harboring *EGFR*-activating mutations. After an initial response, the tumors invariably relapse, in half of the cases because of the appearance of the T790M “gatekeeper” secondary mutation in the catalytic domain of this receptor, resulting in an increased affinity for ATP (Chong and Jänne, 2013). To generate a new model of NSCLC resistance to EGFR inhibitors, we designed a specific sgRNA and a donor single-stranded DNA oligonucleotide (ssODN) containing the T790M mutation, as well as a few additional silent mutations to serve as a genetic barcode. As a control for potential CRISPR/Cas9 off-target cleavage, we generated a ssODN (EGFR-T790T) containing a distinct set of silent mutations (Figure 1A). To prevent the potential incorporation of the two DNA sequences into different alleles within the same cell, EGFR inhibitor-sensitive PC9 cells were transfected with the CRISPR/Cas9 plasmid, encoding Cas9 and the sgRNA, together with one of the two ssODNs. Immediately after transfection, the cells were pooled to form a mixed population of unmodified, EGFR-T790M, and EGFR-T790T cells, and PCR primers were designed to specifically recognize each barcode (Figure 1B).

As shown in Figures 1C and S1A, gefitinib treatment provoked a gradual increase in the EGFR-T790M to EGFR-T790T barcode ratio, consistent with acquired resistance to the inhibitor conferred by the mutation. Similar results were obtained using a different sgRNA (Figure S1B). Because of the EGFR-T790T internal control, the CRISPR-barcoding approach allows specificity over a wide range of inhibitor concentrations. Indeed, an enrichment of EGFR-T790M containing cells was readily detectable after 4 days of gefitinib treatment at a concentration of 10 nM, while a stronger effect was observed at higher doses (Figure 1D).

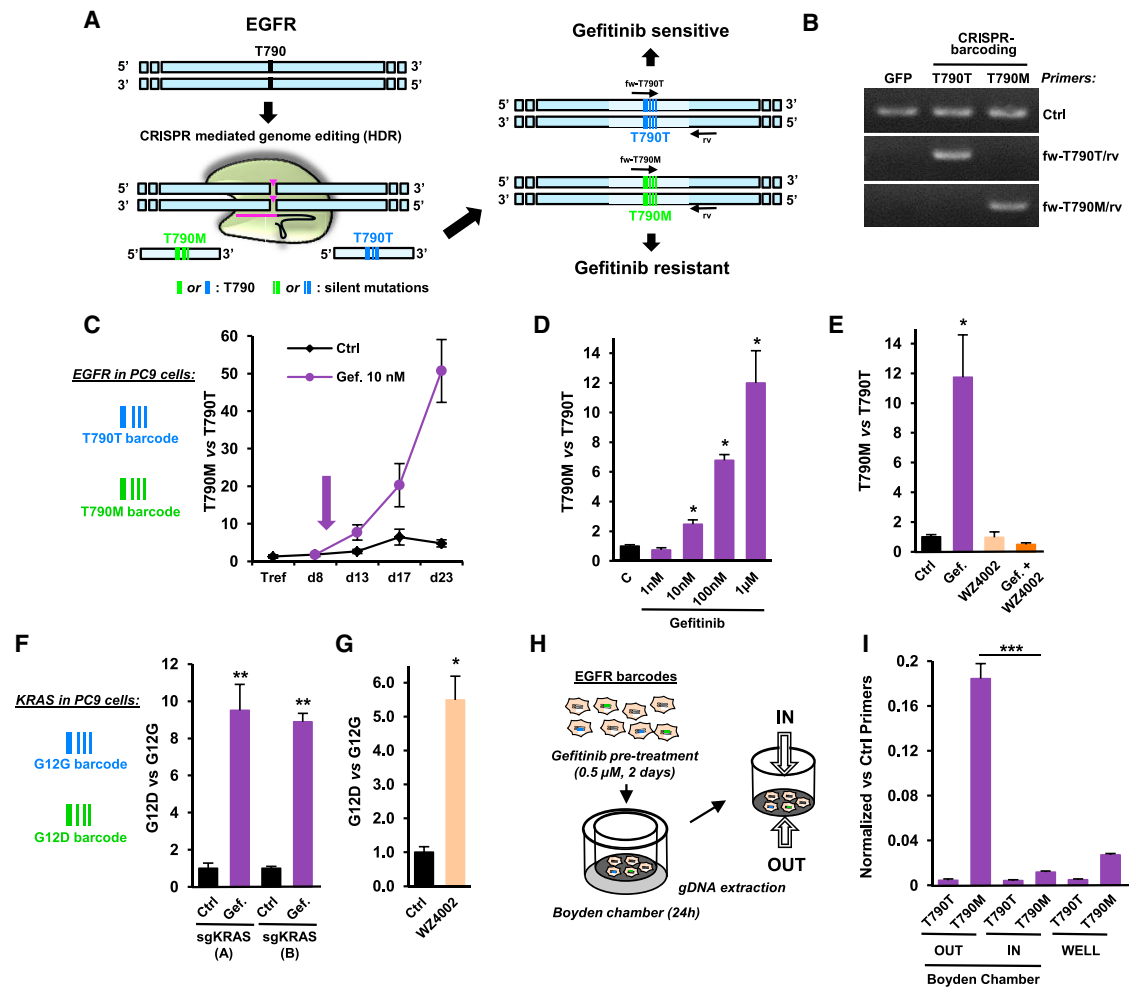


Figure 1. CRISPR-Barcoding to Recapitulate NSCLC Resistance to EGFR Inhibitors

(A) Schematic representation of the incorporation of EGFR-T790M and EGFR-T790T (control) barcodes in NSCLC cells.

(B) The specificity of the qPCR primers for EGFR barcodes was assessed using gDNA from 293T cells transfected with a control GFP plasmid or the CRISPR/Cas9 vector with either the EGFR-T790T or the EGFR-T790M ssODN. The control (Ctrl) primers allowed amplification of a different, unmodified region of the EGFR gene.

(C) Nine days after transfection for EGFR CRISPR-barcoding, PC9 cells were treated in the presence or the absence of gefitinib (10 nM, arrow). qPCR was performed from gDNA extracted at the indicated time points. The results correspond to the mean values of the EGFR-T790M to EGFR-T790T barcode ratio (\pm SEM; $n = 3$) of one representative of three independent experiments.

(D) PC9 cells containing the EGFR-T790M and EGFR-T790T barcodes were treated for 4 days with the indicated concentrations of gefitinib. Mean \pm SEM; $n = 4$ of one representative of three independent experiments.

(E) Effects of 4 days of gefitinib (0.1 μ M) and/or WZ4002 (0.5 μ M) on the EGFR-T790M to EGFR-T790T ratio. Mean \pm SEM; $n = 4$ of one representative of three independent experiments.

(F) Effects of gefitinib (1 μ M, 5 days) on the KRAS-G12D to KRAS-G12G ratio in PC9 cells using two distinct sgRNAs (A or B). Mean \pm SEM; $n = 5$ of one representative of four independent experiments. * $p < 0.05$ and ** $p < 0.01$ compared with control (Mann-Whitney test).

(G) Effects of WZ4002 (0.5 μ M, 4 days) on the KRAS-G12D to KRAS-G12G ratio. Mean \pm SEM; $n = 4$ of one representative of three independent experiments.

(H) Strategy for PC9 cell invasion studies using CRISPR-barcoding.

(I) qPCR analysis of EGFR-T790T and EGFR-T790M barcodes from PC9 cells outside (OUT) or inside (IN) the Boyden chamber, or seeded in parallel in a regular well (WELL), as illustrated in (H). The mean values (\pm SEM; $n = 3$) of one representative of three independent experiments were normalized to the total amount of gDNA using EGFR_Ctrl primers (B). *** $p < 0.001$ compared with the cells inside the Boyden chamber (Student's t test).

See also Figure S1.

Recent clinical trials have shown the efficacy of third-generation EGFR irreversible inhibitors against NSCLCs that developed resistance to gefitinib and erlotinib through the T790M mutation (Jänne et al., 2015). We tested the effects of one of these com-

pounds, WZ4002 (Zhou et al., 2009), in our CRISPR-barcoding model for drug resistance. Figure 1E shows that WZ4002 completely abolished the enrichment of the EGFR-T790M barcode in the presence of gefitinib, consistent with the high affinity

of this compound for the receptor containing the gatekeeper mutation. To model a different type of genetic aberration leading to resistance to EGFR inhibitors in NSCLCs, we used a similar approach to generate a subpopulation of PC9 cells containing a mutation of the oncogene *KRAS* (Figure S1C), a well-known negative predictor for primary responsiveness of NSCLCs to EGFR inhibitors (Pao and Chmielecki, 2010). As shown in Figures 1F and S1D, the *KRAS* mutant barcode was significantly enriched in the presence of gefitinib, consistent with a downstream activation of EGFR signaling in these cells. Of note, a similar enrichment was observed in the presence of WZ4002 (Figure 1G), indicating that therapies based on receptor inhibition could result in the selection of clonal subpopulations containing alternative mechanisms of resistance.

EGFR can stimulate cell migration and invasion through different mechanisms, including Src and STAT5 activation, or by promoting epithelial to mesenchymal transition (Avraham and Yarden, 2011; Huveneers and Danen, 2009). However, these effects are often masked by the overwhelming impact of this receptor on cell proliferation and survival. To investigate the specific contribution of EGFR inhibition in restraining NSCLC invasiveness, independently of its effects on cell growth, EGFR-T790M-barcoded PC9 cells were pre-treated for 2 days with gefitinib, then transferred into Boyden chambers containing Matrigel. Twenty-four hours later, the cells on both sides of the polycarbonate membrane were collected, and gDNA was extracted for qPCR analysis (Figure 1H). As shown in Figure 1I, the relative proportion of the EGFR-T790M barcode was dramatically increased in cells that migrated through the Matrigel-coated membrane, indicating that, in addition to the effect on cell growth, EGFR inhibition has also a profound impact on the invasive properties of NSCLC cells. As a control, in the absence of gefitinib the proportion of the two *EGFR* barcodes remained stable in each compartment (Figure S1E). By measuring the fraction of barcoded cells on both sides of the chamber, our barcoding strategy allows specific and quantitative analysis of cell invasion, regardless of any potential effect on cell proliferation and/or survival.

Multiplex CRISPR-Barcoding for In Vitro and In Vivo Studies

Identification of chromosomal rearrangements involving the gene encoding ALK receptor tyrosine kinase in a significant fraction of metastatic NSCLCs (Soda et al., 2007) led to the approval of specific ALK inhibitors for the treatment of this type of cancer (Solomon et al., 2014). The most frequent of such rearrangements corresponds to an inversion on chromosome 2, resulting in the expression of a constitutively active fusion protein, composed of the echinoderm microtubule-associated protein-like 4 (*EML4*) and ALK (Hallberg and Palmer, 2013). As recently described in other models (Choi and Meyerson, 2014; Maddalo et al., 2014), we used CRISPR/Cas9 technology to generate the *EML4-ALK* inversion in PC9 cells. To increase the efficiency of the chromosomal rearrangement, the cells were co-transfected with a ssODN encompassing the *EML4-ALK* fusion sequence, and qPCR primers were designed to specifically recognize the inverted locus (Figures S2A–S2C), providing a genetic barcode for cells expressing the oncogenic *EML4-ALK*

protein. As shown in Figure 2A, gefitinib treatment increased the fraction of PC9 cells containing *EML4-ALK*, indicating that this chromosomal rearrangement could represent a new mechanism of resistance to EGFR inhibition.

The fact that different mechanisms of resistance can coexist within the same tumor or in metastases from the same patient suggests that multiple resistant clones may be present before the onset of therapy (Chong and Jänne, 2013). To recapitulate this type of genetic heterogeneity, we co-introduced the *EGFR*, *KRAS*, and *EML4-ALK* mutations described above in different subpopulations of the same PC9 cell culture, hereafter designed as PC9-EKE (Figure 2B). Consistent with the data obtained by mutating individual genes, treatment of this heterogeneous cell population with gefitinib increased the proportion of cells containing the EGFR-T790M, *KRAS*-G12D, and *EML4-ALK* barcodes (Figure 2C). This type of multiplex model is particularly suited to evaluate the efficacy of combined therapies aimed at preventing or delaying the emergence of resistant cells. As shown in Figure 2D, although WZ4002 and the ALK inhibitor TAE684 could efficiently prevent gefitinib resistance conferred by the T790M mutation or the *EML4-ALK* fusion, respectively, these compounds had no effect on the other resistant populations, indicating that strategies based on targeting single mechanisms of resistance likely have limited effectiveness in the long term.

It has recently been proposed that targeting the MAPK pathway can be an effective strategy to inhibit the growth of NSCLCs addicted to different oncogenes (Tricker et al., 2015). Using our multiplex model of NSCLC resistance to gefitinib, we tested the effects of trametinib, the first MEK inhibitor approved for cancer treatment. Figure 2E shows that a clinically relevant concentration of trametinib blocked the growth of all gefitinib-resistant populations, indicating that combination therapy including downstream inhibitors of EGFR signaling may prevent or delay NSCLC resistance.

To illustrate the potential applications of our strategy for in vivo studies, PC9-EKE cells were subcutaneously and bilaterally injected in immunocompromised mice. In the presence of gefitinib, the growth of tumors was initially arrested and, in certain cases, reversed, but it eventually resumed after a few weeks (Figure 3A), mimicking the pattern typically observed in the clinic for NSCLC patients treated with this class of inhibitors. To assess whether PC9 tumor relapse resulted from amplification of the small fraction of cells made resistant through CRISPR/Cas9, barcode frequency was measured by qPCR in gDNA from each tumor sample and normalized to the levels observed in the same batch of cells prior to mouse injection. Figure 3B shows that the proportion of all three mutations was markedly increased in the tumors from gefitinib treated mice, indicating that all three mechanisms of resistance were selected in vivo. Of note, the relative enrichment of the three barcodes was different in each tumor; for instance, in mouse number 8, the right-flank tumor showed extremely high levels of EGFR-T790M, while the *KRAS*-G12D barcode was predominant in the left-flank tumor (Figures 3B and S2D). This profile is reminiscent of the genetic heterogeneity that can be observed in tumors and metastases from the same patients and suggests that the selection of a particular mechanism of resistance may be guided by competition among different clonal subpopulations, as well as by stochastic events.

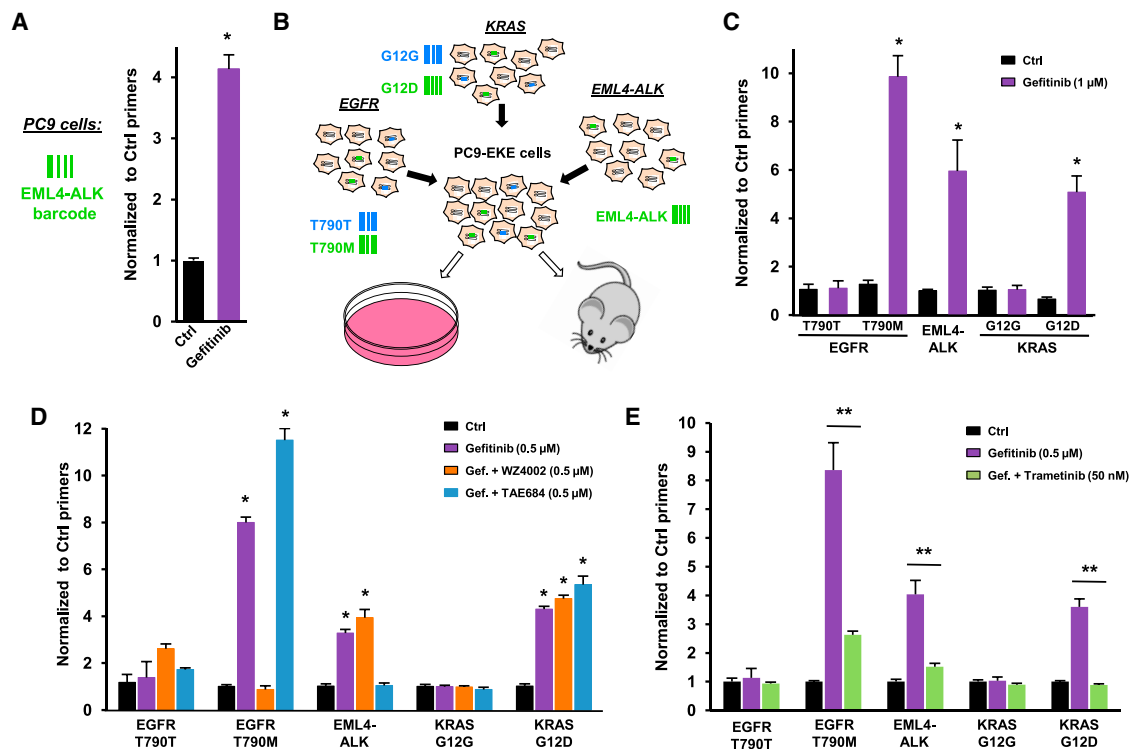


Figure 2. Multiplex Model for NSCLC Resistance to EGFR Inhibition

(A) *EML4-ALK* chromosomal inversion was induced in PC9 cells using CRISPR/Cas9, and the fraction of cells containing the *EML4-ALK* barcode was assessed in the presence or the absence of gefitinib (1 μ M, 5 days). The mean values (\pm SEM; $n = 4$) of one representative of three independent experiments were normalized to the total amount of gDNA using EGFR_Ctrl primers.

(B) Multiplexed CRISPR-barcoding in PC9 cells for in vitro or in vivo studies.

(C) Effects of gefitinib (1 μ M, 5 days) in PC9-EKE cells. The proportion of each barcode was normalized using EGFR_Ctrl primers. Mean (\pm SEM; $n = 4$) of one representative of three independent experiments.

(D) Effects of combinations of gefitinib (0.5 μ M, 5 days) with WZ4002 (0.5 μ M) and TAE684 (0.5 μ M) in PC9-EKE cells. Mean (\pm SEM; $n = 4$) of one representative of three independent experiments.

(E) Effects of gefitinib (0.5 μ M, 5 days) with or without trametinib (50 nM) in EKE cells. Mean (\pm SEM; $n = 5$) of one representative of three independent experiments.

* $p < 0.05$ and ** $p < 0.01$ compared with Ctrl or gefitinib-treated cells (Mann-Whitney test).

See also Figure S2.

CRISPR-Barcoding to Investigate the Effects of Gene Inactivation

Analysis of the distribution of the most common driver mutations within individual tumors has uncovered variable patterns of clonal frequencies (Shah et al., 2012). The coexistence of cancer cells containing or not a particular oncogenic mutation can reflect not only a different stage of tumor evolution but also an improved ability of the tumor to adapt to different environmental conditions. We used our CRISPR-barcoding strategy to model the presence of a subclonal population lacking the tumor suppressor TP53, a major inducer of growth arrest and apoptosis in response to a variety of cellular stressors (Muller and Vousden, 2014). In diploid cells, CRISPR/Cas9 editing can affect one or both alleles, resulting in heterozygous or homozygous mutations, but it is well established that upon CRISPR/Cas9 cleavage, the efficiency of DNA repair through NHEJ is considerably higher compared with HDR (Doudna and Charpentier, 2014; Hsu et al., 2014). If in a particular cell the CRISPR/Cas9 activity is sufficient for the insertion of a barcode on one allele, we reasoned that in

most cases it should also provoke inactivation of the second allele through NHEJ, thus resulting in a dominant effect of the barcode. We designed a ssODN (TP53-STOP) to introduce a STOP codon to replace Phe109, located at the beginning of the TP53 DNA-binding domain (Figures S3A and S3B), as well as a control ssODN (TP53-WT) containing only silent mutations. MCF7 (breast) and HCT-116 (colon) cancer cells, which contain a wild-type (WT) *TP53* gene, were transfected for CRISPR-barcoding and treated or not with Nutlin 3, an inhibitor of MDM2-mediated degradation of TP53 (Vassilev et al., 2004). As shown in Figures 4A, 4B, S3C, and S3D, Nutlin 3 significantly increased the fraction of MCF7 and HCT-116 cells containing the TP53-STOP barcode, consistent with a selective advantage for cells lacking TP53 activity. We used a similar barcoding approach to model TP53 “DNA contact” mutation R273H (Figure S3E), a frequent genetic aberration in human tumors resulting in a dominant-negative form of TP53 (Muller and Vousden, 2014). Nutlin 3 treatment significantly increased the fraction of HCT-116 cells containing the TP53-R273H barcode (Figures S3F and S3G),

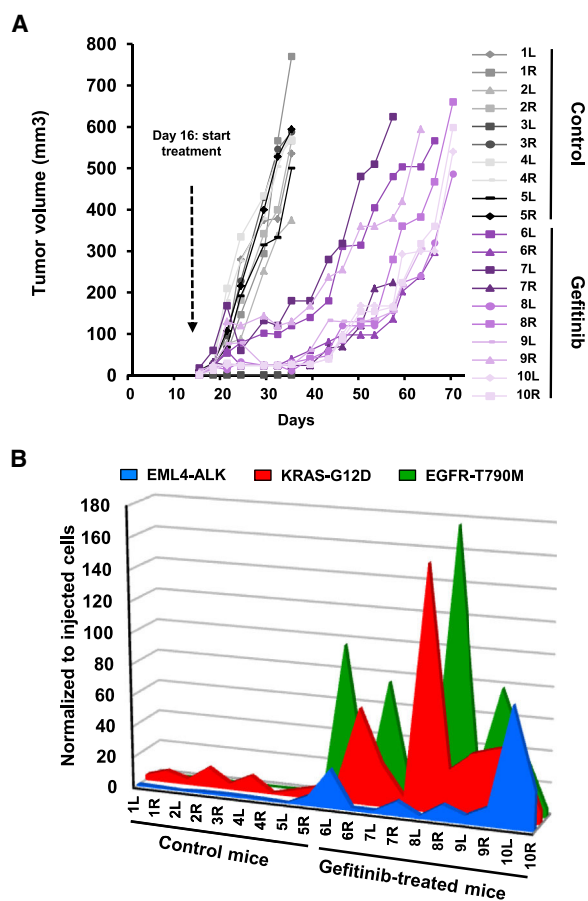


Figure 3. NSCLC In Vivo Resistance to Gefitinib

(A) Multiplex CRISPR-barcoded PC9 cells were subcutaneously injected in the left (L) and right (R) flanks of SCID mice, and the volume of the tumors was measured by caliper. At day 16, mice 6–10 were treated with gefitinib (25 mg/kg/day). The mice were sacrificed when at least one of the tumors reached 550 mm³.

(B) The fraction of the EML4-ALK, KRAS-G12D, and EGFR-T790M barcodes in each tumor was assessed by qPCR and normalized to the total amount of gDNA. The proportions of the three barcodes in the cells prior to mouse injection were analyzed in parallel and arbitrarily set to 1.

See also Figure S2.

indicating that this mutation exerts an effect similar to TP53 loss in our model.

Despite exerting similar or stronger effects than Nutlin 3 on TP53 induction and cell growth inhibition (Figure S3H and data not shown), the genotoxic agent doxorubicin did not affect the relative proportion of the two barcodes (Figures 4B and S3D), supporting the notion that TP53 is not strictly required for the activation of DNA damage checkpoints (Jackson and Bartek, 2009). It has been shown that inhibition of the ataxia telangiectasia mutated (ATM) kinase, a major regulator of the response to DNA damage, can specifically sensitize TP53 null cells to genotoxic stress by promoting mitotic catastrophe and subsequent cell death (Jiang et al., 2009). Consistent with this model, co-treatment with doxorubicin and the ATM inhibitor KU-55933 reduced the fraction of HCT-116 cells containing the TP53-

STOP barcode (Figure 4C). In a seemingly paradoxical manner, a recent genome-wide small hairpin RNA (shRNA) screen revealed that ATM inhibition exerts an opposite effect upon nongenotoxic activation of TP53, in that it promoted apoptosis of TP53 WT cells, possibly through the inhibition of an ATM-dependent prosurvival pathway (Sullivan et al., 2012). Indeed, Figure 4D shows that Nutlin 3 and KU-55933 synergistically increased the fraction of HCT-116 cells containing the TP53-STOP barcode, thus indicating that different combination strategies targeting the same oncogenic lesion can potentially lead to opposite therapeutic outcomes.

Repair of a Mutated Tumor Suppressor Gene through CRISPR-Barcoding

Unlike other CRISPR/Cas9 strategies, our barcoding approach enables tracing of the mutated cells immediately after DNA editing without the need to derive clones, thus providing a unique means to investigate the effects of HDR-mediated modifications, regardless of their potential impact on cell growth. The APC tumor suppressor gene encodes a major component of the destruction complex that promotes the degradation of free β -catenin, and it is mutated in more than 80% of colorectal cancers (Clevers and Nusse, 2012; Polakis, 2012). We used CRISPR-barcoding to investigate the effects of repairing the homozygous frameshift mutation at position c.4248 of this gene in DLD-1 colon cancer cells (Figure S4A). We designed specific sgRNA constructs and two donor DNAs for APC HDR, containing either the repaired (APC-WT) or a STOP codon, as well as a two series of silent mutations to serve as specific barcodes (Figure S4B).

To assess the effects of APC restoration on Wnt signaling activation, we used a lentiviral reporter containing a destabilized form of GFP, driven by a minimal promoter containing 14 Wnt-responsive elements (Figures S4C and S4D). Five days after CRISPR-barcoding transfection, DLD-1 cells stably containing the Wnt reporter were sorted using fluorescence-activated cell sorting (FACS) to isolate the 10% of cells displaying the highest (GFP^{hi}) or lowest (GFP^{lo}) levels of GFP (Figure 5A). As shown in Figures 5B and S4E, qPCR analysis from gDNA revealed that the APC-WT to APC-STOP ratio was strongly increased in GFP^{lo} compared with GFP^{hi} or unsorted cells, consistent with an enrichment of the repaired APC allele in cells displaying low levels of Wnt reporter activity.

Using inducible shRNAs, it has recently been shown that restoration of APC expression can promote regression of small intestine and colon tumors (Dow et al., 2015). To investigate the effects on colon cancer cell growth of directly repairing the mutated APC gene, DLD-1 cells were trypsinized 3 days after CRISPR-barcoding transfection (Tref), replated into four different flasks, and propagated for about 4 weeks. At each cell passage, gDNA was extracted for qPCR analysis (Figure S5A). We found that the fraction of DLD-1 cells containing a repaired APC gene strongly decreased over time, reaching a plateau at 2 weeks after transfection (Figures 5C and S5B–S5D), indicating that aberrant activation of Wnt signaling is strictly required for the growth of these colon cancer cells. We confirmed these data by deep sequencing the samples from Figure 5C and found that the proportion of APC-WT reads dramatically decreased over time

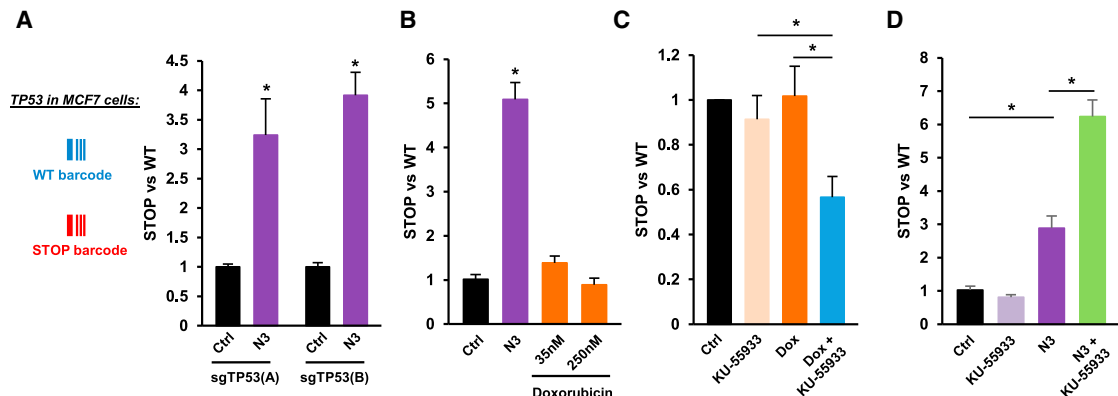


Figure 4. Inactivation of the Tumor Suppressor TP53 through CRISPR-Barcoding

(A) Effects of Nutlin-3 (N3; 10 μ M, 7 days) on the TP53-STOP to TP53-WT ratio in MCF7 cells using two distinct sgRNAs (A or B). Mean \pm SEM; $n = 4$ of one representative of three independent experiments. * $p < 0.05$ (Mann-Whitney test).

(B) Effects of N3 (10 μ M; 7 days) or doxorubicin (Dox) on the TP53-STOP to TP53-WT ratio in HCT-116 cells. Mean \pm SEM; $n = 4$ of one representative of three independent experiments.

(C) The cells in (B) were treated for 7 days with KU-55933 (10 μ M) and/or Dox (50 nM), and the TP53-STOP to TP53-WT ratio was assessed by qPCR. Mean \pm SEM of seven independent experiments.

(D) Effects of KU-55933 (10 μ M) and/or N3 (10 μ M, 7 days) on the TP53-STOP to TP53-WT ratio in HCT-116 cells. Mean \pm SEM; $n = 4$ of one representative of three independent experiments.

See also Figure S3.

compared with the reads containing the APC-STOP barcode (Figure 5D). Of note, a small fraction of APC-WT reads carried additional frameshift mutations, likely originating from errors during the HDR process or the synthesis of the ssODN. Such mutations, which are expected to prevent restoration of a full-length APC protein, showed a 6-fold increase at day 27 compared to the Tref and most likely accounted for the persistence of a small but stable population of cells retaining the APC-WT barcode at longer time points (Figures 5C, 5D, and S5B–S5D).

Repair of a Mutated Oncogene in Neuroblastoma Cells

Although different strategies can conceivably be used to investigate the oncogenic properties of a putative gain-of-function mutation, the definitive demonstration of its role as a driver force in a particular cancer cell would require the restoration of a WT sequence, which is not possible with current DNA editing approaches (see above). We chose the receptor tyrosine kinase ALK, mutated in 6%–10% of neuroblastomas (Cheung and Dyer, 2013; Hallberg and Palmer, 2013), and we used CRISPR-barcoding to correct the ALK-F1174L activating mutation in Kelly neuroblastoma cells (Figures S6A and S6B) by designing three distinct ssODNs: (1) ALK-F1174F, to repair the oncogenic mutation; (2) ALK-F1174L, as an internal control for the mutated receptor; and (3) ALK-STOP, to generate a truncated protein lacking most of the catalytic domain. Because the parental cells contain a WT and a F1174L allele, incorporation of each barcode could result in four possible combinations; for three of them, including the likely most frequent barcode/indel, the phenotypically dominant sequence expressed in the cell is expected to be the one encoded by the barcode (Figure S6C). As shown in Figures 6A and 6B, ALK-F1174F and ALK-STOP barcodes similarly decreased over time, while the fraction of ALK-F1174L cells remained stable. The same results were obtained using a distinct

sgRNA (Figures S6D and S6E), indicating that these neuroblastoma cells are addicted to the F1174L mutation of ALK.

The silent mutations that compose a CRISPR-barcode may conceivably have an impact on the stability of the corresponding transcript. Although irrelevant for a STOP barcode, variations in the expression levels of other modified sequences could potentially result in false-positive or false-negative effects. With the possible exception of TP53-STOP, the relative proportions of the different APC, TP53, and ALK barcodes were similar in gDNA versus cDNA extracted in parallel from the same cells (Figure S6F). As an additional control, we repeated the ALK barcoding experiment by switching the silent mutations between ALK-F1174F and ALK-F1174L. The “swapped” barcodes confirmed a similar growth disadvantage of cells containing the WT or the STOP codon (Figures 6C and S6G), thus excluding an effect of the silent mutations on ALK mRNA stability.

To further demonstrate the specificity of our approach, Figure 6D shows that the fractions of ALK-F1174F and ALK-STOP cells decreased significantly less in the presence of a small-molecule ALK inhibitor, indicating that the selective pressure that favors the expression of a mutant receptor is abolished upon inhibition of ALK catalytic activity.

Highly Complex CRISPR-Barcodes to Probe Intratumor Heterogeneity

Randomly integrated lentiviral libraries have been used to tag and trace the fate of the different cells comprising a tumor mass population (Bhang et al., 2015; Nguyen et al., 2014). As an alternative approach, we adapted CRISPR-barcoding to label distinct subsets of cancer cells through insertion of a degenerate sequence at a specific genomic location. We chose the adeno-associated virus integration site 1 (AAVS1) locus on chromosome 19, a genomic “safe harbor” widely used for transgene

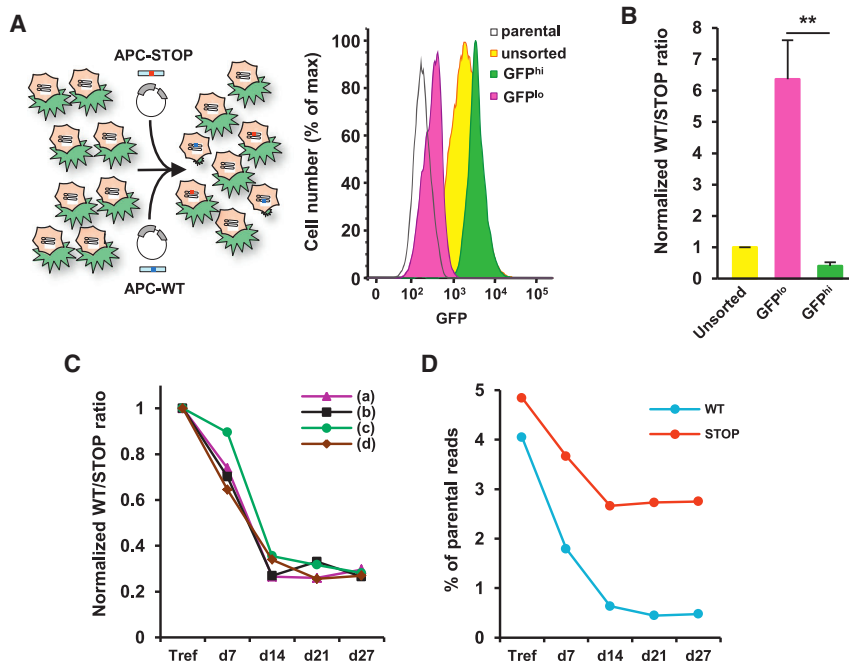


Figure 5. Repair of the Mutated APC Gene in Colon Cancer Cells

(A) FACS sorting of DLD-1 cells containing the Wnt-responsive GFP reporter and transfected for APC CRISPR-barcoding.

(B) APC-WT to APC-STOP ratio in the cells FACS sorted in (A). Mean \pm SEM; $n = 3$ of one representative of three independent experiments. ** $p < 0.01$ (Student's *t* test).

(C) APC-WT to APC-STOP ratio at different time points after transfection of DLD-1 cells. Mean of one representative of three independent experiments.

(D) Deep sequencing analysis of representative gDNA samples used in (C). The percentage of reads containing the APC-WT (WT) and APC-STOP (STOP) barcodes compared with the endogenous sequence is shown.

See also Figures S4 and S5.

insertion (DeKelver et al., 2010), and used BT474 and PC9 cells, derived from breast cancer and NSCLC, respectively (Figure 7A). The cells were sequentially transfected for CRISPR-barcoding with two distinct sgRNAs and a partially degenerate ssODN containing a Sall restriction site for restriction fragment length polymorphism analysis. As shown by surveyor assay, the AAVS1 locus was edited in the vast majority of the transfected cells (Figure S7A and data not shown), while HDR efficiency was about 22% and 8% for BT474 and PC9, respectively (Figure S7B, Table S1, and data not shown).

After barcoding, BT474 cells were either maintained in culture or injected bilaterally in the fat pads of immunocompromised mice (Figures 7A and S7C). gDNA was derived from both tumors and cells in culture at 21, 28, and 35 days after inoculation, and the barcoded AAVS1 locus was deep-sequenced. In each sample, we detected several thousands of distinct sequences, among which the most frequent corresponded to the parental sequence and to small indels at the two CRISPR/Cas9 cleavage sites induced by NHEJ (Figure S7D, corresponding to the dots in the top right section of the plots shown in Figures 7B–7D and S7E). While the distribution of the barcodes was similar in the three tissue culture samples (Figure 7B), comparison between each tumor and its matched cell culture sample resulted in more scattered profiles (Figures 7C and S7E). Of note, high correlation was found among the different tumors, originated either in the same or in different animals (Figures 7D and S7E), indicating that although the great majority of BT474 cells could contribute to tumor formation, certain subpopulations of cells showed an intrinsically different fitness for *in vivo* growth.

We used a similar approach to compare the effects of EGFR inhibition in PC9 and PC9-EKE cells (Figure 7A). Deep-sequencing analysis revealed that in the presence of gefitinib (1 μ M, 14 days) a few hundred barcodes were consistently en-

riched compared with the control in all four replicates of both cell lines (Figure 7E), indicating that certain subpopulations are constitutively less sensitive to the inhibitor. Of note, although the number of barcodes upregulated between 10- and 50-fold was similar in both lines, PC9-EKE cells showed a higher fraction of barcodes upregulated more than 50-fold (Figure 7F). Consistent with a recent study using lentiviral barcoded cells (Hata et al., 2016), our data suggest that before the beginning of the treatment, the mass population already contains not only fully resistant clones, probably resulting from preexisting additional genetic aberrations, such as those we artificially inserted through CRISPR/Cas9 in PC9-EKE cells, but also slower growing, drug-tolerant cells, which were similarly represented in both parental and EKE cells and could constitute a sort of reservoir for new mutations potentially leading to acquired resistance.

DISCUSSION

DNA barcoding is a taxonomic method to identify known or new species through sequencing of standardized genomic regions, which is widely used in ecology to assess and monitor environmental biodiversity (Valentini et al., 2009). On the basis of the idea that cancer can be viewed as an evolutionary and ecological process (Merlo et al., 2006), we devised a new type of genetic barcoding to label and follow the fate of a cluster of tumor cells exposed to genetic drift and selective pressure. By coupling the barcode to a potentially functional mutation, this approach provides a unique means to model intratumor heterogeneity and cancer progression.

The extraordinary capacity of cancer cells to evade therapy largely depends on the complex genetic diversity within individual tumors. Compared with current approaches, our barcoding strategy not only provides strong evidence that a certain mutation confers resistance, but also allows this conclusion to be made very rapidly (Figures 1C and 1D). Of note, a recent study demonstrated that therapeutic inhibition of oncogenic driver mutations can induce a complex array of paracrine signals that

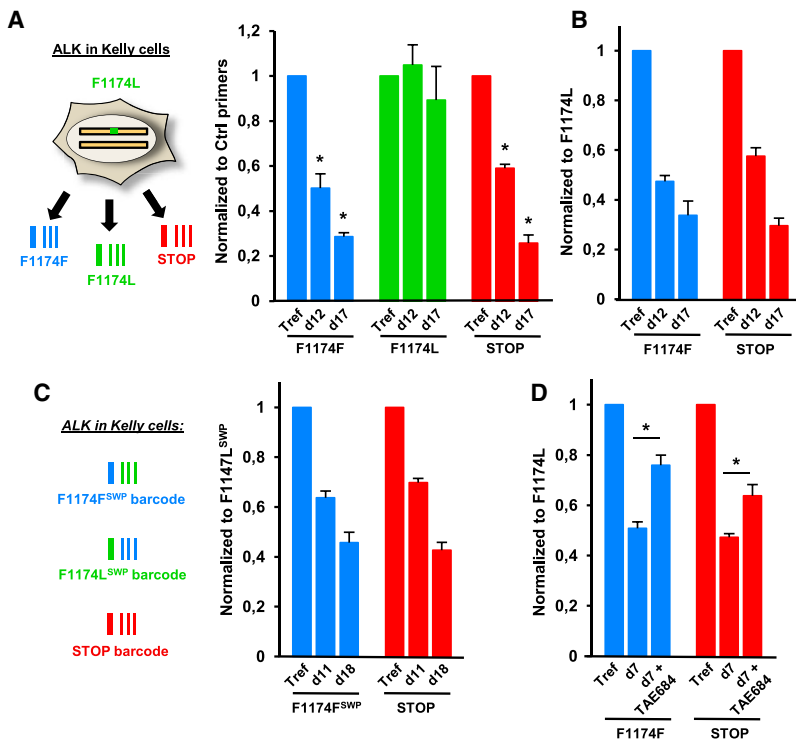


Figure 6. Repair of the Oncogenic ALK-F1174L Mutation in Neuroblastoma Cells

(A) Relative proportions of the three ALK barcodes in Kelly cells at 3 (Tref) 12 or 17 days after transfection, normalized using ALK_Ctrl primers. Mean \pm SEM; n = 4 of one representative of three independent experiments. *p < 0.05 compared with the F1174L control barcode (Mann-Whitney test).

(B) The data in (A) were normalized to the ALK-F1174L barcode.

(C) The silent mutations on the F1174F and F1174L barcodes were swapped (SWP), and an experiment was performed as in (A) and (B). Mean \pm SEM; n = 4 of one representative of three independent experiments.

(D) Effects of TAE684 (0.2 μ M, 7 days) on normalized ALK-F1174F and ALK-STOP barcodes in Kelly cells. gDNA derived from the same cells at time zero was used as the reference (Tref). Mean \pm SEM; n = 4 of one representative of three independent experiments. *p < 0.05 compared with untreated cells (Mann-Whitney test).

See also Figure S6.

promote the emergence, expansion, and metastatic spread of resistant clones (Obenauf et al., 2015). Compared with other approaches, our system reproduces the potential crosstalk between genetically distinct cancer cells both in the presence and in the absence of therapy, thus providing a model that more closely recapitulates the complexity of the response to treatment.

One of the major advantages of CRISPR-barcoding is that it can be easily adapted to multiplexing, which allows comparison of the effects of different mutations within the same heterogeneous cell population. As an example, we used this type of approach to assess the efficacy of different combinations of inhibitors in lung cancer cells containing various resistance mutations. Of note, whereas in our model we could not discriminate between homo- or heterozygous mutations, new strategies have been recently developed to specifically target either one or both alleles (Paquet et al., 2016). By integrating known mechanisms used by a particular type of tumor to escape therapy, multiplex models of resistance can be applied to design and validate combinations of compounds with the best potential of achieving tumor remission or long-term stabilization of the disease.

Besides making it possible to recapitulate and model intratumor genetic heterogeneity, CRISPR-barcoding can also be used to investigate the effects of a particular mutation, regardless of its potential effects on cell growth and/or survival. Indeed, although the CRISPR revolution has greatly simplified DNA editing and genome manipulation, the low efficiency of HDR imposes strong limitations as to the types of modifications and cell models that can be used. From a cancer perspective, with current approaches it is extremely difficult, if not impos-

sible, to repair mutations that drive the growth of cancer cells. By enabling monitoring of the cell phenotype immediately after DNA editing, we demonstrated how our strategy could be used to establish and trace any kind of modification in most cancer types, including reversal of powerful oncogenic events, such as APC or ALK mutations, thus providing ultimate proof of the selective growth advantages such mutations confer. Finally, we describe a strategy to separately label at a specific genomic location thousands of different cancer subpopulations, as an alternative to the time-consuming generation and optimization of complex viral libraries.

Although the proofs of concept reported here are related to cancer models, this technology can be implemented in different fields of biological research. Through the concurrent use of a control barcode, consisting of silent mutations, our approach directly tackles the potential non-specific effects that can derive from off-target DNA cleavage induced by the sgRNA-Cas9 complex (Doudna and Charpentier, 2014; Fu et al., 2013; Hsu et al., 2014).

Despite recent attempts to increase HDR efficiency through small molecules (Yu et al., 2015; Chu et al., 2015; Maruyama et al., 2015) or improved design of the donor DNA (Richardson et al., 2016), the fact that only a fraction of the transfected cells contains the desired mutation constitutes a major limitation for CRISPR/Cas9 technology. By combining the extreme sensitivity of barcode detection through qPCR or deep sequencing, without the need to select and amplify individual clones, CRISPR-barcoding does not require particularly high levels of DNA editing efficiency (Table S1). Hence, our approach can be used in most cell types to generate different kinds of genetic manipulations, including missense point mutations, gene inactivation through insertion of a STOP codon, and chromosomal rearrangements, regardless of their potential effects on cell growth and/or survival.

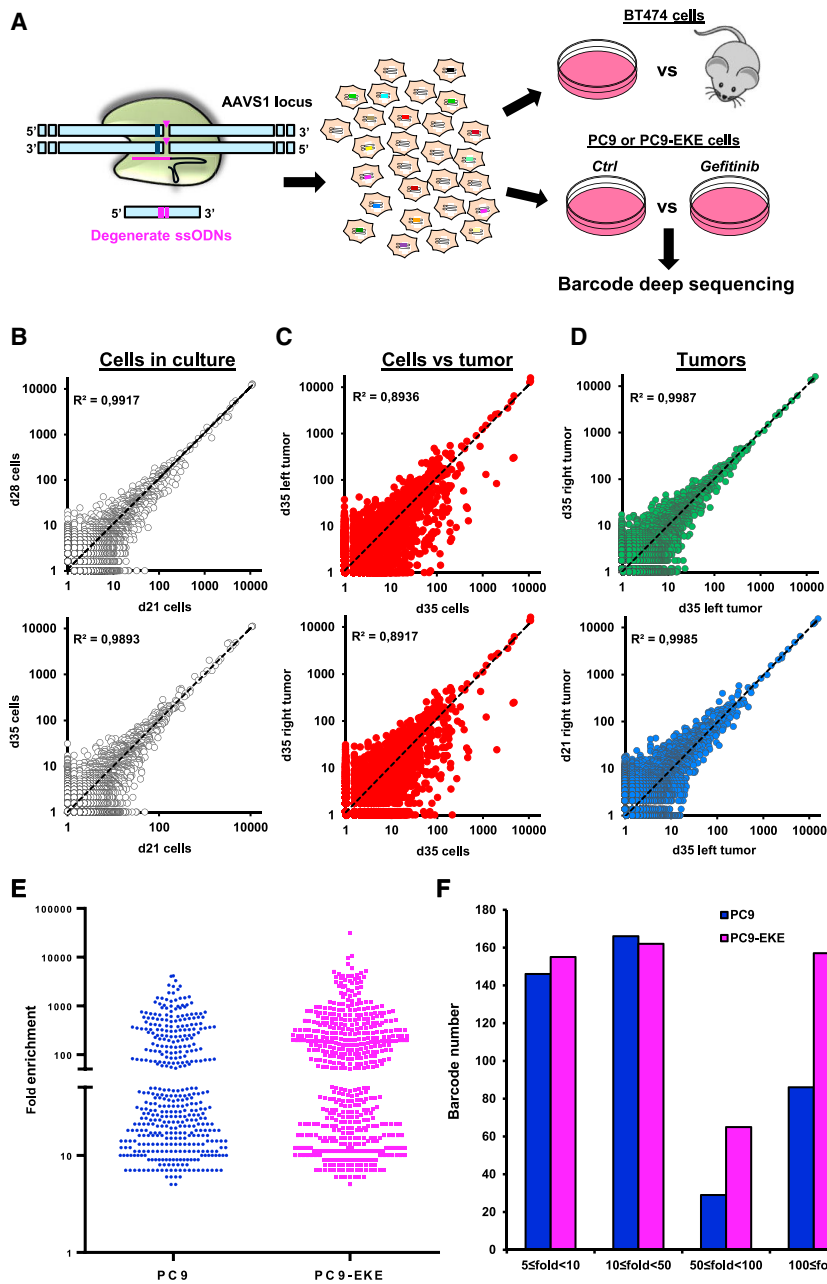


Figure 7. Probing Intratumor Heterogeneity Using Complex CRISPR-Barcodes

(A) Insertion of a degenerate sequence in the AAVS1 locus of BT474 or PC9 cells.

(B) Barcode distribution in BT474 cells maintained in culture for 21, 28, or 35 days. Linear regression was calculated and the coefficient of determination (r^2) is shown.

(C) Barcode profile in tumors versus the corresponding cell cultures (35 days after inoculation).

(D) Barcode distribution in different tumors from the same (up) or different (low) mice.

(E) Distribution of enriched AAVS1 barcodes in PC9 and PC9-EKE cells treated with or without gefitinib (1 μ M, 14 days). Each dot represents a unique barcode upregulated at least 5-fold compared with the control cells in each of the four gefitinib replicates.

(F) The barcodes in (E) are represented according to their mean fold enrichment in the presence of gefitinib.

See also Figure S7.

to “read” the barcodes, including deep sequencing (Figures 5D and 7), TaqMan genotyping probes (data not shown) and Droplet Digital PCR, this approach could also serve as a customizable platform for high-throughput screening of compounds targeting specific genetic alterations, such as those underlying resistance to cancer therapy.

Limitations

Although a limitation of this approach could conceivably derive from the fact that the silent mutations forming the barcode might have an effect on mRNA stability, we demonstrated how this potential issue can be addressed by simply swapping the silent mutations between barcodes. Another potential drawback is the need for some sort of selective pressure in order to establish a positive or negative effect on the relative fraction of cells containing the desired DNA modification.

In conclusion, CRISPR-barcoding is a fast and highly flexible means to investigate the effects of different kinds of genomic modifications in a broad range of functional assays. On the basis of a simple idea, this approach enables a series of novel applications, which would be otherwise extremely difficult, if not impossible, with other methods. Indeed, we demonstrated how this strategy can be shaped to recapitulate and trace intratumor genetic heterogeneity, through generation of novel cancer models for a better understanding of specific oncogenic mechanisms, as well as to develop and validate new therapeutic protocols. Given its adaptability to multiplexing, as well as the possibility to use different methods

Although CRISPR-barcoding cannot be adapted to all kinds of assays, we provided several proofs of concept to illustrate how it can be easily implemented to investigate effects on cell growth and survival, invasion, transcriptional activity, and resistance and sensitivity to a particular agent in culture and in vivo.

Despite being considerably faster and easier to implement compared with viral libraries, the use of highly complex CRISPR-barcodes to trace cellular heterogeneity requires good levels of HDR efficiency, which could limit its potential applications. However, considering the new approaches for improved HDR efficiency (Yu et al., 2015; Chu et al., 2015;

Maruyama et al., 2015; Richardson et al., 2016) and the fact that indels generated by NHEJ can also act as barcodes (Figure S7D), this strategy can be probably used in a relatively broad array of cell models. Finally, depending on the targeted locus, one cell can conceivably contain distinct barcodes on different alleles, which should be taken into account during the analysis of the data, for example through the identification of barcodes displaying identical patterns. To overcome this issue, a haploid region of the genome could be chosen as a target in certain models, such as a locus on the X or Y chromosomes in cells derived from a male individual.

EXPERIMENTAL PROCEDURES

Cell Culture, Transfection, and Inhibitors

293T (human embryonic kidney), DLD-1, HCT-116 (colorectal carcinoma), MCF7, and BT474 (breast cancer) cells were obtained from ATCC; PC9 cells (NSCLC) were obtained from ECACC-Sigma-Aldrich; and Kelly cells (neuroblastoma) were a kind gift from Dr. C. Einvik. HCT-116 and MCF7 cells are WT for *TP53* (<http://p53.iarc.fr/CellLines.aspx>). All cells were grown in DMEM (Life Technologies), except Kelly and PC9 cells, grown in RPMI medium (Life Technologies), both supplemented with 10% fetal bovine serum (Life Technologies) and 0.6% penicillin/streptomycin (Life Technologies). Cells were transfected with a Nucleofector II device (Lonza) using the Amaxa Nucleofector kit (Lonza) and electroporation program recommended by the manufacturer. 293T cells were transfected using polyethylenimine (Polysciences). The efficiency of each transfection was assessed in parallel using a GFP-containing plasmid.

Nutlin 3 was purchased from SelleckChem, and doxorubicin was purchased from Abcam. Gefitinib, WZ4002, TAE684, trametinib, and KU-55933 were purchased from Santa Cruz Biotechnology.

CRISPR-Barcodeing

sgRNA target sequences (Table S2) were designed using the CRISPR Design tool hosted by the Massachusetts Institute of Technology (<http://crispr.mit.edu>) to minimize potential off-target effects. Oligos encoding the targeting sequence were then annealed and ligated into the pSpCas9(BB)-2A-Puro (Ran et al., 2013) vector digested with BbsI (New England Biolabs). The sequence of the ssODNs (Integrated DNA Technologies) used for CRISPR/Cas9-mediated HDR, containing one missense/nonsense mutation coupled to different silent mutations, are provided in Table S2. The set of silent mutations is designed to enable PCR specificity and to avoid recognition by the corresponding sgRNA used to cleave the endogenous sequence. For each targeted locus, cells were co-transfected with 2 μ g of the CRISPR/Cas9 plasmid and 2 μ l of either the control or the sense/nonsense ssODN (50 μ M) to prevent the potential incorporation of the two donor DNA sequences into different alleles within the same cell. Immediately after transfection, the cells were pooled in the same flask. For AAVS1 barcodeing, BT474 cells were subjected to two rounds of transfection with a 2-week interval using CRISPR/Cas9 plasmids encoding two distinct sgRNAs, together with a ssODN containing nine degenerate nucleotides and a Sall restriction site.

qPCR

gDNA was extracted using the NucleoSpin Tissue kit (Macherey-Nagel). The sequence of the different PCR primers, designed using Primer-BLAST (National Center for Biotechnology Information), is provided in Table S3. To avoid potential amplification from ssODN molecules not integrated in the correct genomic locus, one of the two primers was designed to target the endogenous genomic sequence flanking the region sharing homology with the ssODNs. Primer specificity for each particular barcode was assessed. qPCR was performed from 100 ng of gDNA using SYBR Green (Life Technologies) on a 7900 HT Fast-Real-Time, a Q-PCR ABI PRISM 7500, or a QuantStudio Flex PCR System (Life Technologies). qPCR analysis was performed using the standard curve or the Pfaffl methods (Pfaffl, 2001).

Statistical Analysis

Statistical analysis was performed with GraphPad Prism software using the Mann-Whitney test or Student's two-tailed t test. Linear regression of the AAVS1 barcode normalized frequencies between different samples was calculated using Excel (Microsoft).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2016.06.017>.

AUTHOR CONTRIBUTIONS

L.G. conceived the original idea and designed and supervised the study. A.G., S.K.M., D.C., D.A., S.A.A., Y.A., and L.G. designed and/or analyzed experiments. A.G., S.K.M., D.C., S.Y., H.A., and L.G. performed experiments. R.S., A.J., and G.R. performed deep sequencing analysis. M.V., F.C., S.C., and I.T. performed deep sequencing analysis for the *APC* gene. S.A. and O.B. performed cell sorting and FACS analysis. L.G. wrote the manuscript, with input and editing from the other authors.

ACKNOWLEDGMENTS

We thank G. Liu, J.Y. Lee, J.M. Flaman, D. Vaudry, and members of INSERM U982 for helpful discussion. Some of the experiments were performed at the Cell Imaging Platform of Normandie (PRIMACEN). We thank G. Riou (IRIB Flow Cytometry Facility) and M. Di Giovanni (PRIMACEN) for technical assistance. This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM), the Université de Rouen Normandie, the Ligue Contre le Cancer de Haute-Normandie, the National Cancer Institute (P01CA080058), and the Breast Cancer Research Foundation. A.G. is recipient of a doctoral fellowship from the Normandie Region. L.G. was supported by a Chair of Excellence program from INSERM and the Université de Rouen Normandie. A patent application related to this work has been submitted by INSERM and the Université de Rouen Normandie.

Received: February 12, 2016

Revised: May 10, 2016

Accepted: June 10, 2016

Published: July 21, 2016

REFERENCES

- Avraham, R., and Yarden, Y. (2011). Feedback regulation of EGFR signalling: decision making by early and delayed loops. *Nat. Rev. Mol. Cell Biol.* 12, 104–117.
- Bhang, H.E., Ruddy, D.A., Krishnamurthy Radhakrishna, V., Caushi, J.X., Zhao, R., Hims, M.M., Singh, A.P., Kao, I., Rakiec, D., Shaw, P., et al. (2015). Studying clonal dynamics in response to cancer therapy using high-complexity barcodeing. *Nat. Med.* 21, 440–448.
- Boettcher, M., and McManus, M.T. (2015). Choosing the right tool for the job: RNAi, TALEN, or CRISPR. *Mol. Cell* 58, 575–585.
- Cheung, N.K., and Dyer, M.A. (2013). Neuroblastoma: developmental biology, cancer genomics and immunotherapy. *Nat. Rev. Cancer* 13, 397–411.
- Choi, P.S., and Meyerson, M. (2014). Targeted genomic rearrangements using CRISPR/Cas technology. *Nat. Commun.* 5, 3728.
- Chong, C.R., and Jänne, P.A. (2013). The quest to overcome resistance to EGFR-targeted therapies in cancer. *Nat. Med.* 19, 1389–1400.
- Chu, V.T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., and Kühn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* 33, 543–548.
- Clevers, H. (2011). The cancer stem cell: premises, promises and challenges. *Nat. Med.* 17, 313–319.

- Clevers, H., and Nusse, R. (2012). Wnt/ β -catenin signaling and disease. *Cell* 149, 1192–1205.
- DeKaveler, R.C., Choi, V.M., Moehle, E.A., Paschon, D.E., Hockemeyer, D., Meijnsing, S.H., Sancak, Y., Cui, X., Steine, E.J., Miller, J.C., et al. (2010). Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res.* 20, 1133–1142.
- Diaz, L.A., Jr., Williams, R.T., Wu, J., Kinde, I., Hecht, J.R., Berlin, J., Allen, B., Bozic, I., Reiter, J.G., Nowak, M.A., et al. (2012). The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* 486, 537–540.
- Ding, L., Wendl, M.C., McMichael, J.F., and Raphael, B.J. (2014). Expanding the computational toolbox for mining cancer genomes. *Nat. Rev. Genet.* 15, 556–570.
- Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346, 1258096.
- Dow, L.E., O'Rourke, K.P., Simon, J., Tschaharganeh, D.F., van Es, J.H., Clevers, H., and Lowe, S.W. (2015). Apc restoration promotes cellular differentiation and reestablishes crypt homeostasis in colorectal cancer. *Cell* 161, 1539–1552.
- Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826.
- Gerlinger, M., Rowan, A.J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., et al. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* 366, 883–892.
- Gillies, R.J., Verdusco, D., and Gatenby, R.A. (2012). Evolutionary dynamics of carcinogenesis and why targeted therapy does not work. *Nat. Rev. Cancer* 12, 487–493.
- Greaves, M., and Maley, C.C. (2012). Clonal evolution in cancer. *Nature* 481, 306–313.
- Hallberg, B., and Palmer, R.H. (2013). Mechanistic insight into ALK receptor tyrosine kinase in human cancer biology. *Nat. Rev. Cancer* 13, 685–700.
- Hata, A.N., Niederst, M.J., Archibald, H.L., Gomez-Caraballo, M., Siddiqui, F.M., Mulvey, H.E., Maruvka, Y.E., Ji, F., Bhang, H.E., Krishnamurthy Radhakrishna, V., et al. (2016). Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. *Nat. Med.* 22, 262–269.
- Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262–1278.
- Huvnereers, S., and Danen, E.H. (2009). Adhesion signaling - crosstalk between integrins, Src and Rho. *J. Cell Sci.* 122, 1059–1069.
- Jackson, S.P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature* 461, 1071–1078.
- Jänne, P.A., Yang, J.C., Kim, D.W., Planchard, D., Ohe, Y., Ramalingam, S.S., Ahn, M.J., Kim, S.W., Su, W.C., Horn, L., et al. (2015). AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. *N. Engl. J. Med.* 372, 1689–1699.
- Jiang, H., Reinhardt, H.C., Bartkova, J., Tommiska, J., Blomqvist, C., Nevanlinna, H., Bartek, J., Yaffe, M.B., and Hemann, M.T. (2009). The combined status of ATM and p53 link tumor development with therapeutic response. *Genes Dev.* 23, 1895–1909.
- Lin, Y., Cradick, T.J., Brown, M.T., Deshmukh, H., Ranjan, P., Sarode, N., Wile, B.M., Vertino, P.M., Stewart, F.J., and Bao, G. (2014). CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res.* 42, 7473–7485.
- Maddalo, D., Manchado, E., Concepcion, C.P., Bonetti, C., Vidigal, J.A., Han, Y.C., Ogdowski, P., Crippa, A., Rektman, N., de Stanchina, E., de Stanchina, E., et al. (2014). In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature* 516, 423–427.
- Maruyama, T., Dougan, S.K., Truttmann, M.C., Bilate, A.M., Ingram, J.R., and Ploegh, H.L. (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* 33, 538–542.
- McGranahan, N., and Swanton, C. (2015). Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell* 27, 15–26.
- Merlo, L.M., Pepper, J.W., Reid, B.J., and Maley, C.C. (2006). Cancer as an evolutionary and ecological process. *Nat. Rev. Cancer* 6, 924–935.
- Misale, S., Yaeger, R., Hobor, S., Scala, E., Janakiraman, M., Liska, D., Vaitorta, E., Schiavo, R., Buscarino, M., Siravegna, G., et al. (2012). Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature* 486, 532–536.
- Muller, P.A., and Vousden, K.H. (2014). Mutant p53 in cancer: new functions and therapeutic opportunities. *Cancer Cell* 25, 304–317.
- Nguyen, L.V., Cox, C.L., Eirew, P., Knapp, D.J., Pellacani, D., Kannan, N., Carles, A., Moksa, M., Balani, S., Shah, S., et al. (2014). DNA barcoding reveals diverse growth kinetics of human breast tumour subclones in serially passaged xenografts. *Nat. Commun.* 5, 5871.
- Nowell, P.C. (1976). The clonal evolution of tumor cell populations. *Science* 194, 23–28.
- Obenauf, A.C., Zou, Y., Ji, A.L., Vanharanta, S., Shu, W., Shi, H., Kong, X., Bosenberg, M.C., Wiesner, T., Rosen, N., et al. (2015). Therapy-induced tumour secretomes promote resistance and tumour progression. *Nature* 520, 368–372.
- Pao, W., and Chmielecki, J. (2010). Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. *Nat. Rev. Cancer* 10, 760–774.
- Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., Olsen, K.M., Gregg, A., Noggle, S., and Tessier-Lavigne, M. (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 533, 125–129.
- Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29, e45.
- Polakis, P. (2012). Wnt signaling in cancer. *Cold Spring Harb. Perspect. Biol.* 4, a008052.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308.
- Richardson, C.D., Ray, G.J., DeWitt, M.A., Curie, G.L., and Corn, J.E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* 34, 339–344.
- Shah, S.P., Roth, A., Goya, R., Oloumi, A., Ha, G., Zhao, Y., Turashvili, G., Ding, J., Tse, K., Haffari, G., et al. (2012). The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* 486, 395–399.
- 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.
- Solomon, B., Wilner, K.D., and Shaw, A.T. (2014). Current status of targeted therapy for anaplastic lymphoma kinase-rearranged non-small cell lung cancer. *Clin. Pharmacol. Ther.* 95, 15–23.
- Sullivan, K.D., Padilla-Just, N., Henry, R.E., Porter, C.C., Kim, J., Tentler, J.J., Eckhardt, S.G., Tan, A.C., DeGregori, J., and Espinosa, J.M. (2012). ATM and MET kinases are synthetic lethal with nongenotoxic activation of p53. *Nat. Chem. Biol.* 8, 646–654.
- Tricker, E.M., Xu, C., Uddin, S., Capelletti, M., Ercan, D., Ogino, A., Pratilas, C.A., Rosen, N., Gray, N.S., Wong, K.K., and Jänne, P.A. (2015). Combined EGFR/MEK inhibition prevents the emergence of resistance in EGFR-mutant lung cancer. *Cancer Discov.* 5, 960–971.
- Turke, A.B., Zejnullahu, K., Wu, Y.L., Song, Y., Dias-Santagata, D., Lifshits, E., Toschi, L., Rogers, A., Mok, T., Sequist, L., et al. (2010). Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC. *Cancer Cell* 17, 77–88.

- Valentini, A., Pompanon, F., and Taberlet, P. (2009). DNA barcoding for ecologists. *Trends Ecol. Evol.* *24*, 110–117.
- Vassilev, L.T., Vu, B.T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C., et al. (2004). In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* *303*, 844–848.
- Yu, C., Liu, Y., Ma, T., Liu, K., Xu, S., Zhang, Y., Liu, H., La Russa, M., Xie, M., Ding, S., and Qi, L.S. (2015). Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* *16*, 142–147.
- Zhou, W., Ercan, D., Chen, L., Yun, C.H., Li, D., Capelletti, M., Cortot, A.B., Chirieac, L., Iacob, R.E., Padera, R., et al. (2009). Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature* *462*, 1070–1074.

Molecular Cell, Volume 63

Supplemental Information

CRISPR-Barcoding for Intratumor Genetic

Heterogeneity Modeling and Functional Analysis

of Oncogenic Driver Mutations

Alexis Guernet, Sathish Kumar Mungamuri, Dorthe Cartier, Ravi Sachidanandam, Anitha Jayaprakash, Sahil Adriouch, Myriam Vezain, Françoise Charbonnier, Guy Rohkin, Sophie Coutant, Shen Yao, Hassan Ainani, David Alexandre, Isabelle Tournier, Olivier Boyer, Stuart A. Aaronson, Youssef Anouar, and Luca Grumolato

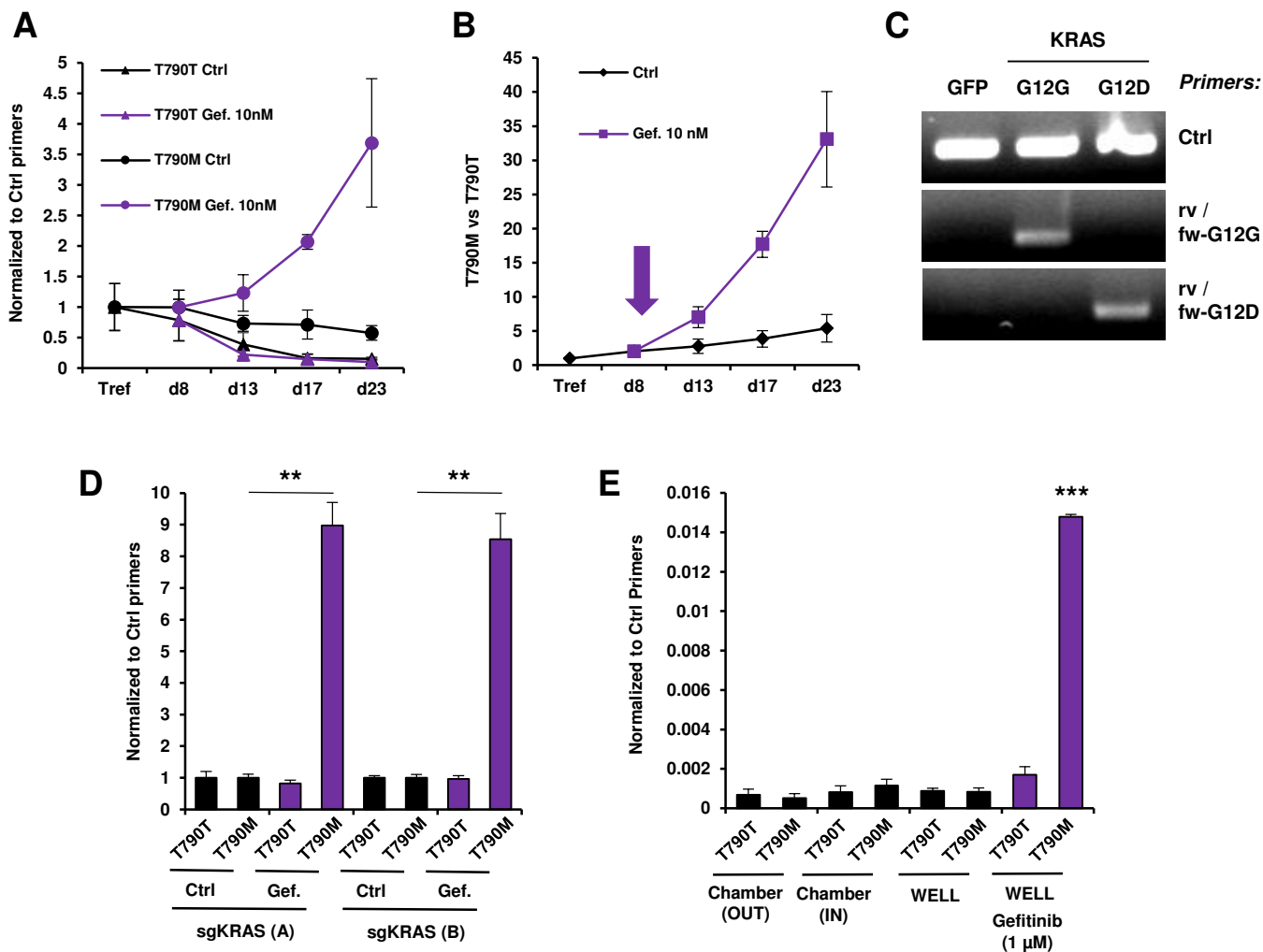


Figure S1, related to Figure 1. CRISPR-Barcoding to Investigate Cancer Resistance to Chemotherapy

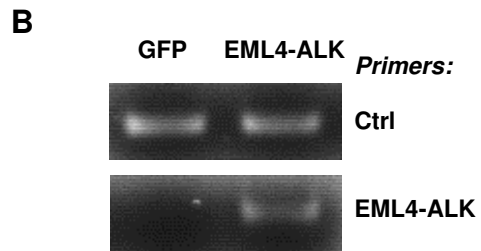
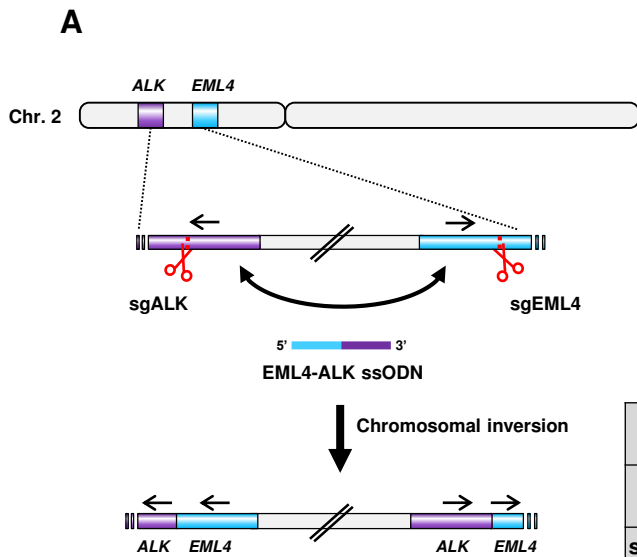
(A) The data from Figure 1C are expressed as relative proportion of each barcode normalized to the amount of total DNA measured using qPCR primers amplifying an unmodified region of the EGFR gene.

(B) Nine days after transfection for EGFR CRISPR-barcoding using sgEGFR(B), PC9 cells were treated in the presence or the absence of gefitinib (10 nM, arrow). qPCR was performed from gDNA extracted at the indicate time points. The results correspond to the mean values of the T790M- to T790T-EGFR barcode ratio (\pm SEM; n=3) of one representative of three independent experiments.

(C) PCR using control or barcode-specific primers was performed from gDNA extracted from 293T cells transfected with a control GFP plasmid or the sgKRAS vector with the KRAS-G12G or the KRAS-G12D ssODN.

(D) The data from Figure 1F are expressed as relative proportion of each barcode normalized to the amount of total DNA measured using qPCR primers amplifying an unmodified region of the KRAS gene.

(E) Untreated PC9 cells were used in a CRISPR-barcoding invasion assay as described in Figure 4F. qPCR analysis for the EGFR-T790T and -T790M barcodes was performed from cells outside (OUT) or inside (IN) the Boyden chamber, or seeded in parallel in a regular well (WELL). PC9 cells plated in parallel and treated for 3 days with gefitinib (0.5 μ M) were used as a control. The mean values (\pm SEM; n=3) of one representative of two independent experiments were normalized to the total amount of gDNA measured using qPCR primers amplifying an unmodified region of the EGFR gene (Ctrl primers). *** p<0.001 compared to untreated cells (Student's t test).



C

	Ctrl Primers (Ct)	EML4-ALK Primers (Ct)
GFP	19.42	undetermined
sgALK + sgEML4	19.64	26.76
sgALK + sgEML4 + EML4-ALK ssODN	19.61	24.43

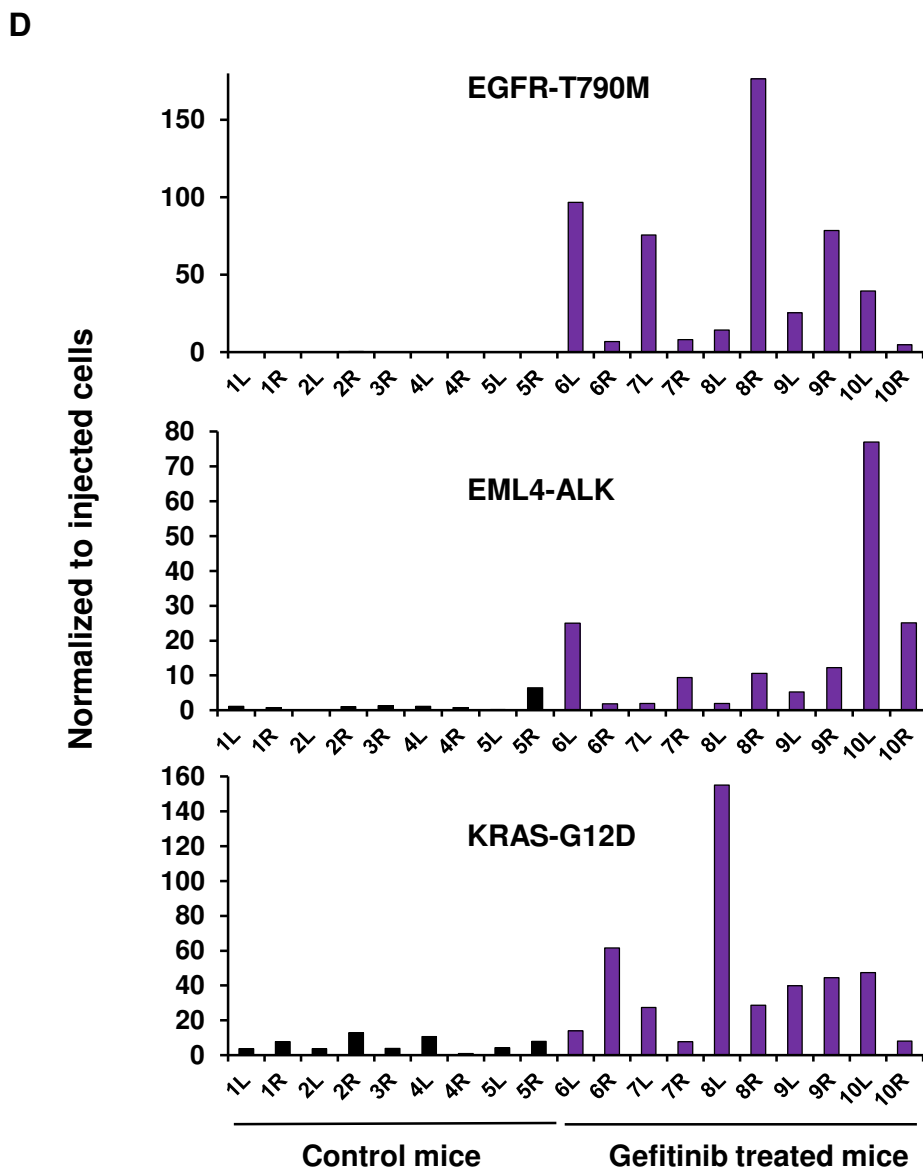


Figure S2, related to Figures 2 and 3. Multiplex Model of NSCLC Resistance to Gefitinib

(A) Schematic representation of the strategy to induce CRISPR-mediated rearrangement of human chromosome 2. The paracentric inversion of ALK and EML4 was induced using two sgRNAs specific for the ALK and EML4 genes and a ssODN corresponding to the EML4-ALK fusion sequence. The direction of ALK and EML4 coding sequences is indicated (arrows).

(B) The specificity of the qPCR primers for the EML4-ALK fusion was assessed using gDNA from 293T cells.

(C) The efficiency of CRISPR-mediated inversion on chromosome 2 in 293T cells was assessed by qPCR in the presence or the absence of EML4-ALK-ssODN. Co-transfection of the donor DNA allowed an increase of efficiency of about 2 Cts.

(D) The data from Figure 3B are represented as three different diagrams corresponding to each mutation.

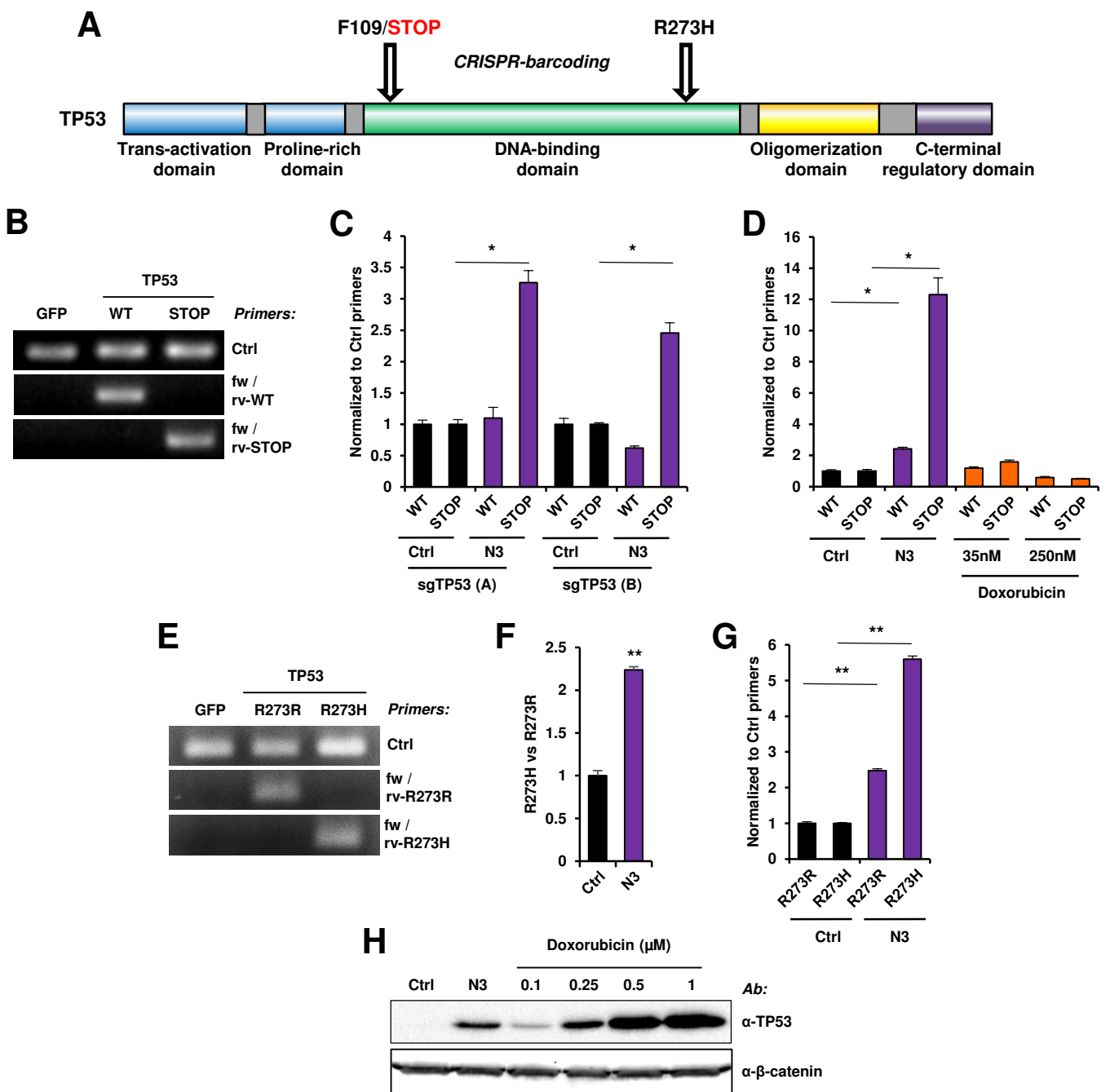


Figure S3, related to Figure 4. Inactivation of the *TP53* Gene

(A) Schematic representation of the TP53 protein. CRISPR-barcoding was used to introduce either a STOP codon at position 109, at the beginning of the DNA-binding domain, or the R273H DNA contact mutation.

(B) The specificity of the qPCR primers for TP53 F109 barcodes was assessed using gDNA from 293T cells transfected with a control GFP plasmid or the CRISPR/Cas9 vector with the TP53-WT or the TP53-STOP ssODN.

(C-D) The data from Figure 4A-B are expressed as relative proportion of each barcode normalized to the amount of total DNA measured using qPCR primers amplifying an unmodified region of the TP53 gene. * $p < 0.05$ (Mann-Whitney's test).

(E) The specificity of the TP53-R273R and TP53-R273H qPCR primers was assessed as in B.

(F) HCT-116 cells containing the TP53-R273 barcodes were treated for seven days in the presence or the absence of Nutlin 3 (N3, 10 μ M) and the TP53-R273H to TP53-R273R ratio was measured by qPCR. Mean \pm SEM; $n=5$ of one representative of three independent experiments. ** $p < 0.01$ (Mann-Whitney's test).

(G) The data from Figure S3F are expressed as relative proportion of each barcode normalized to the amount of total DNA measured using qPCR primers amplifying an unmodified region of the TP53 gene. ** $p < 0.01$ (Mann-Whitney's test).

(H) HCT-116 cells were treated overnight with DMSO (Ctrl), Nutlin-3 (10 μ M; N3) or the indicated concentrations of doxorubicin, followed by western blot using anti-TP53 and anti- β -catenin (loading control) antibodies.

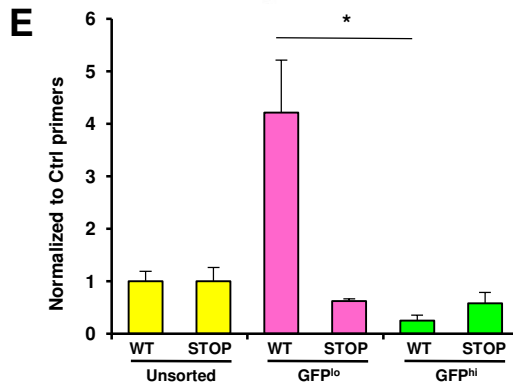
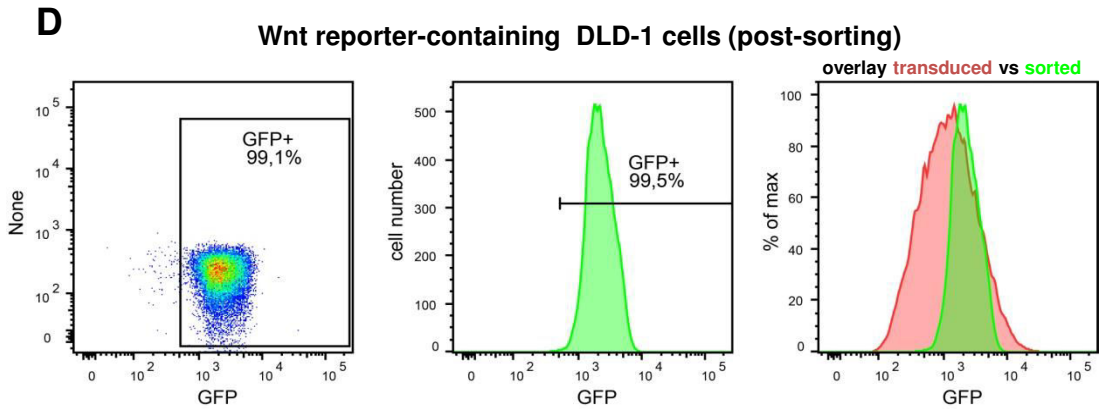
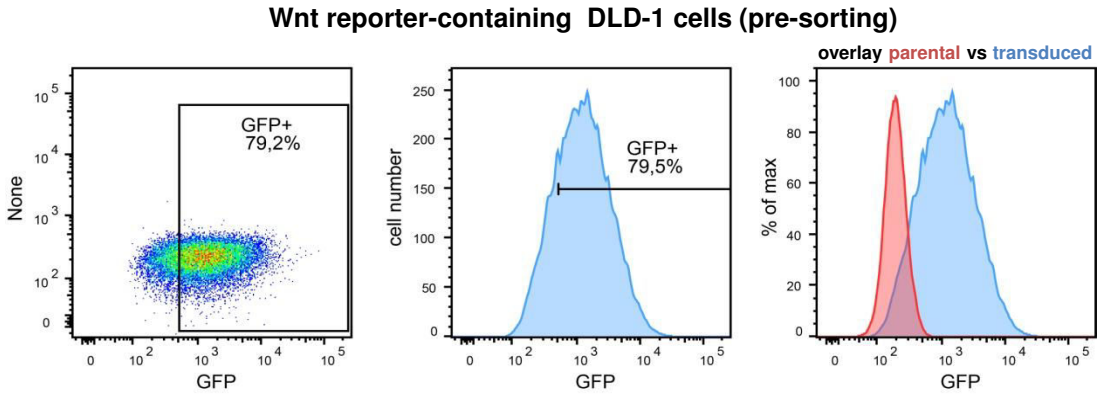
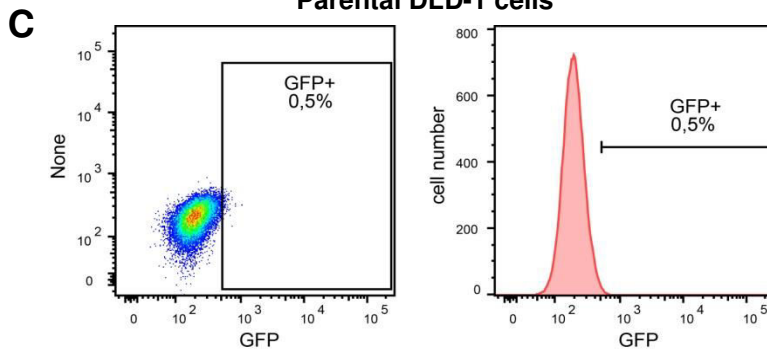
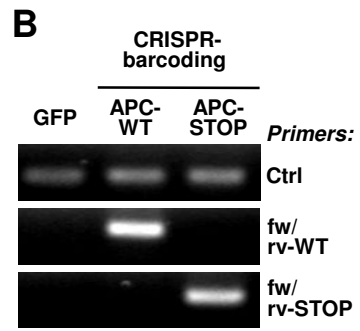
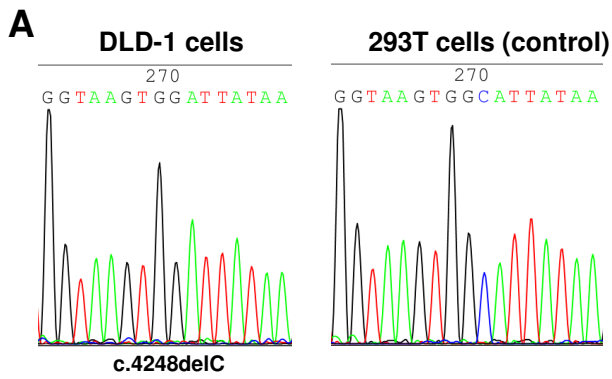


Figure S4, related to Figure 5. FACS Sorting of DLD-1 Cells Containing the Wnt Responsive Reporter

A) Sanger sequencing was performed to confirm the c.4248delC mutation on APC exon 15 in DLD-1 cells. gDNA from 293T cells was used as a control.

(B) The specificity of the qPCR primers for APC barcodes was assessed using gDNA from 293T cells transfected with a control GFP plasmid or the CRISPR/Cas9 vector with the APC-WT or the APC-STOP ssODN.

(C) Parental and Wnt reporter-containing DLD-1 cells were analyzed by flow cytometry before sorting.

(D) The cells in (C) were FACS sorted to isolate GFP⁺ cells expressing the Wnt reporter. Flow cytometric analysis of the cells before and after FACS sorting is shown.

(E) The data from Figure 5B are expressed as relative proportion of each barcode normalized to the amount of total DNA measured using qPCR primers amplifying an unmodified region of the TP53 gene. *p<0.05 (Student's test).

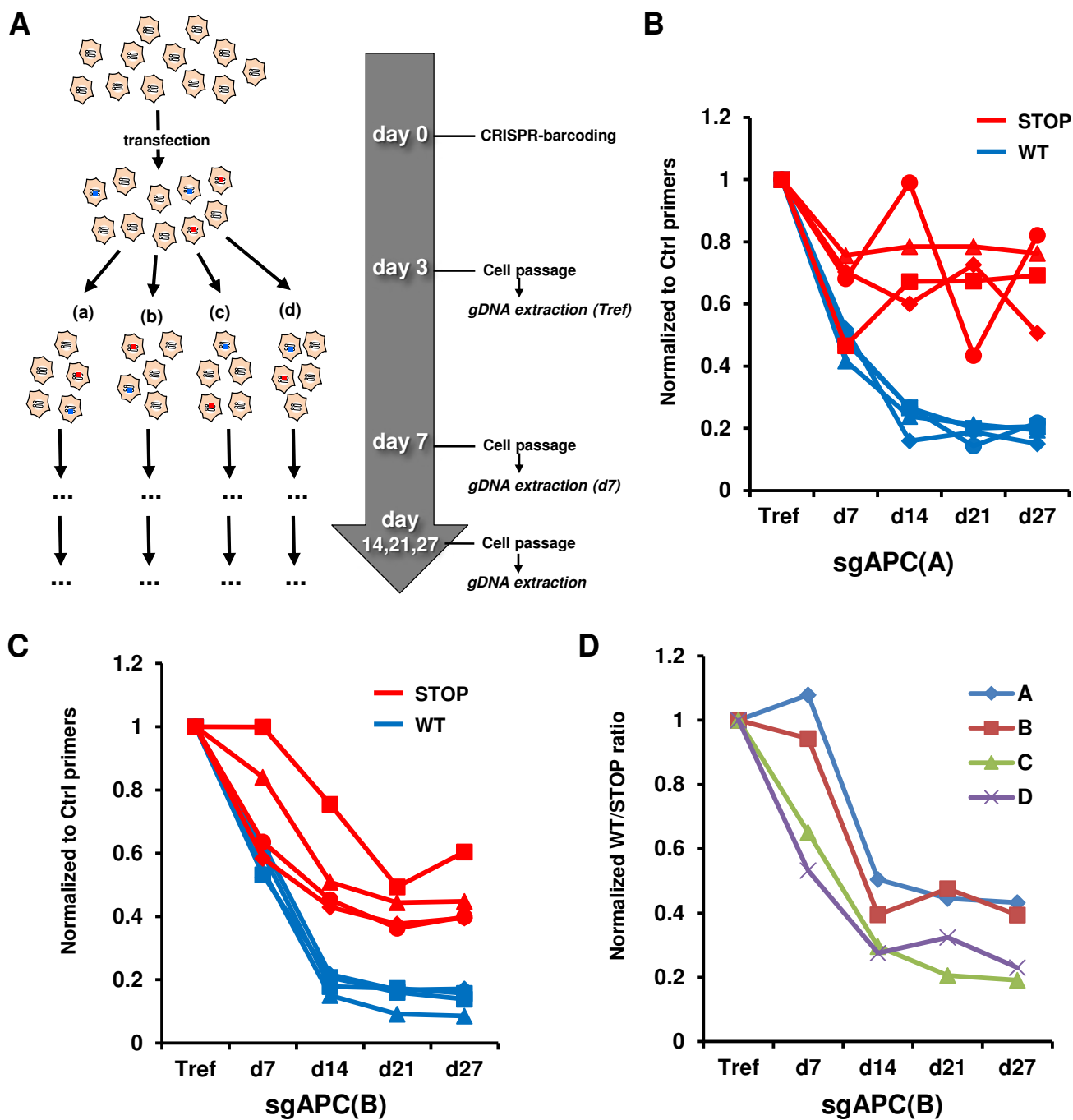
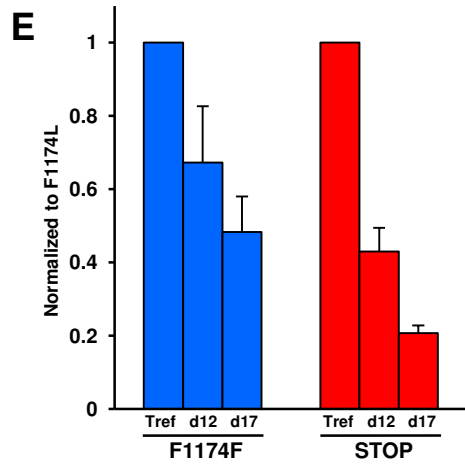
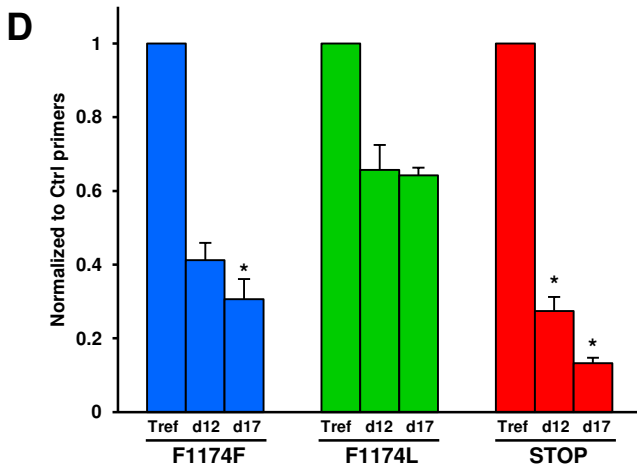
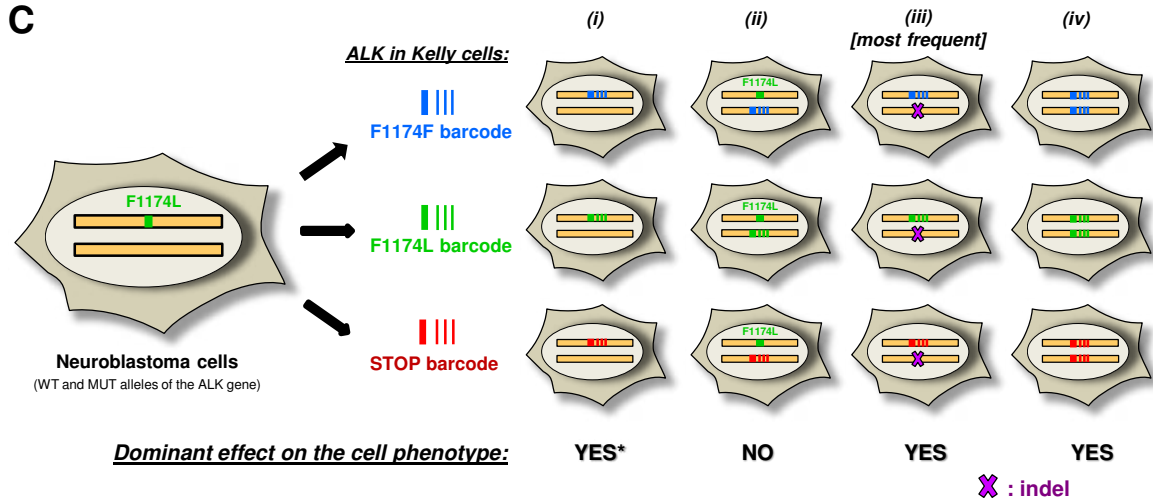
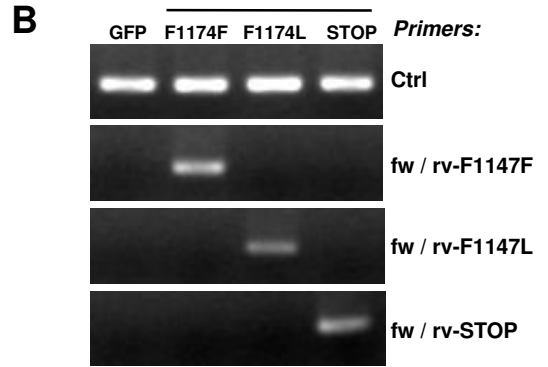
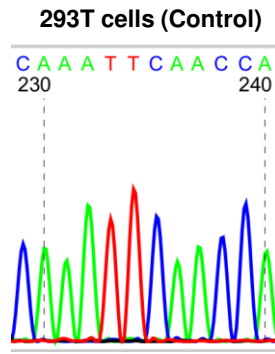
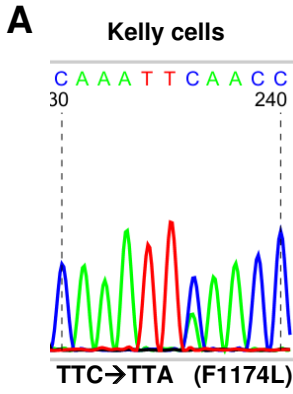


Figure S5, related to Figure 5. APC CRISPR-Barcoding in DLD-1 Cells

(A) Schematic representation of the CRISPR-barcoding approach to investigate the effects of *APC* gene repair on DLD-1 cell growth.

(B-C) APC CRISPR-barcoding was performed using two distinct sgRNA (sgAPC(A) and sgAPC(B)), and the APC-WT and APC-STOP qPCR values were normalized to the total amount of gDNA, measured using qPCR primers amplifying an unmodified region of the APC gene (Ctrl primers). The results correspond to the APC-WT to APC-STOP ratio of one representative of three independent experiments.

(D) The data in B are represented as APC-WT to APC-STOP barcode ratio.



F

	APC-WT	APC-MUT
mean	0.6198	1.365
sd	0.1855	0.4576

	p53-WT	p53-STOP
mean	50.919	13.637
sd	10.901	4.009

	ALK-F1174F	ALK-F1174L	ALK-STOP
mean	62.584	18.276	8.8747
sd	34.939	11.78	1.286

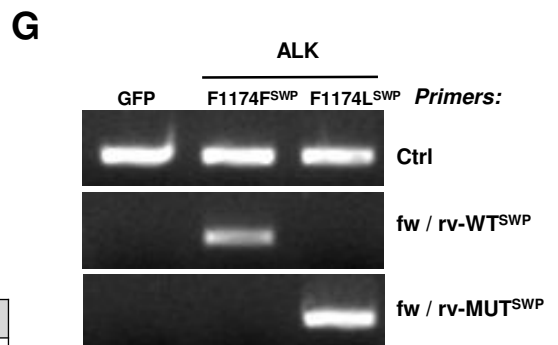


Figure S6, related to Figure 6. Repair of *ALK* Oncogenic Mutation in Kelly cells

(A) Sanger sequencing was performed to confirm the ALK-F1174L mutation in Kelly cells. gDNA from 293T cells was used as a control.

(B) PCR using control or barcode-specific primers was performed from gDNA extracted from 293T cells transfected with a control GFP plasmid or the sgALK vector, together with either the ALK-F1174F, ALK-F1174L or ALK-STOP ssODN.

(C) Diagram illustrating the different possible genotypes corresponding to the ALK-F1174 codon following CRISPR-barcoding transfection of Kelly cells. The predicted effects on the cell phenotype are indicated. (*): dominant effect referred to the F1174L mutation

(D) ALK CRISPR-barcoding in Kelly cells using sgALK(B). The relative proportion of the ALK-F1174F, ALK-F1174L and ALK-STOP barcodes was assessed by qPCR at the indicated time points. The results correspond to the mean values (\pm SEM, n=4), normalized to the amount of DNA measured using qPCR targeting an unmodified region of the ALK gene (Ctrl primers), of one representative of three independent experiments.

(E) The data shown in (D) were normalized to the ALK-F1174L barcode.

(F) To test for a potential effect of CRISPR/Cas9 induced mutations on mRNA stability, the relative proportions of APC, p53 and ALK barcodes were measured by qPCR in the gDNA and the cDNA of 293T cells transfected for CRISPR-barcoding. The mean values (arbitrary units; \pm s.d. ; n=4) of the different barcodes normalized to the total amount of gDNA or cDNA are represented.

(G) The specificity of the primers used to detect ALK “swapped” barcodes (SWP) was assessed as in B.

*p<0.05 compared to the F1174L control barcode. (Mann-Whitney’s test)

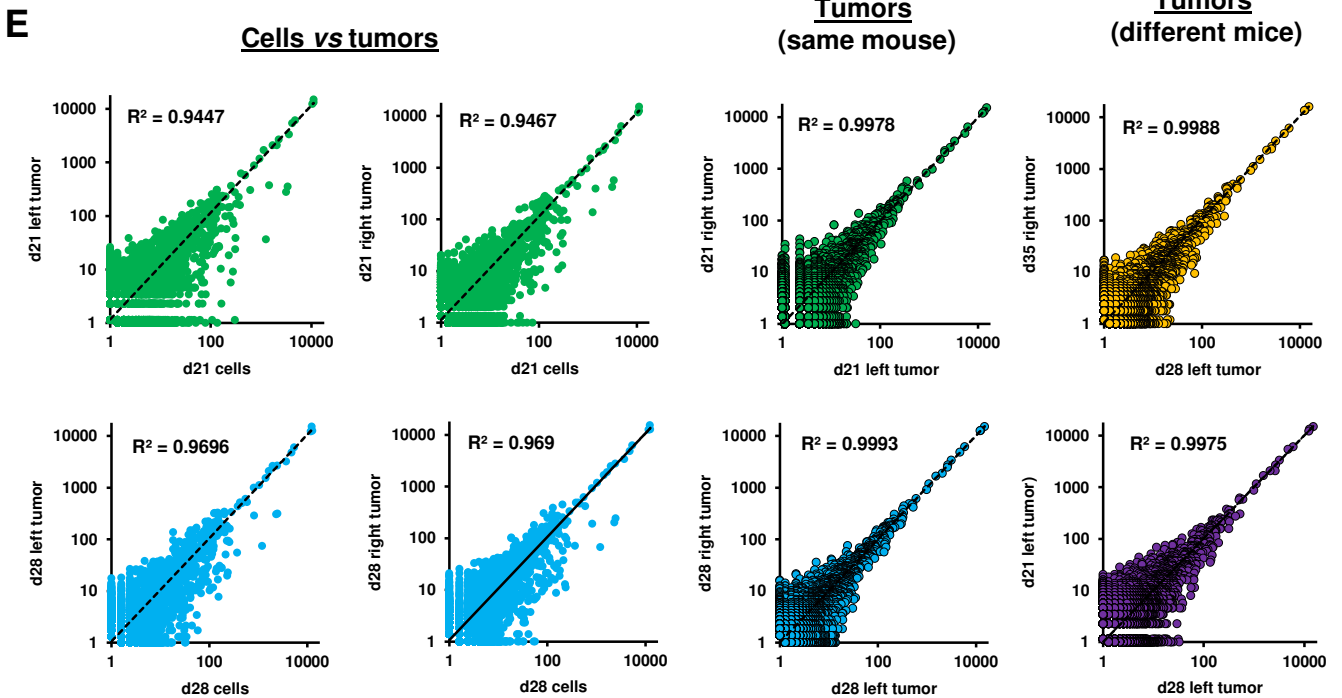
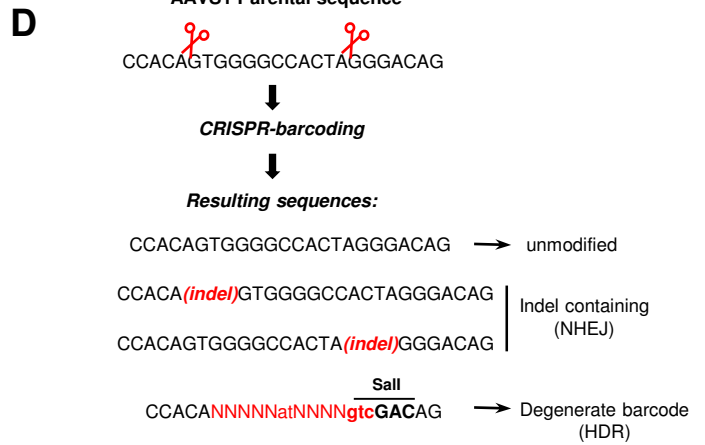
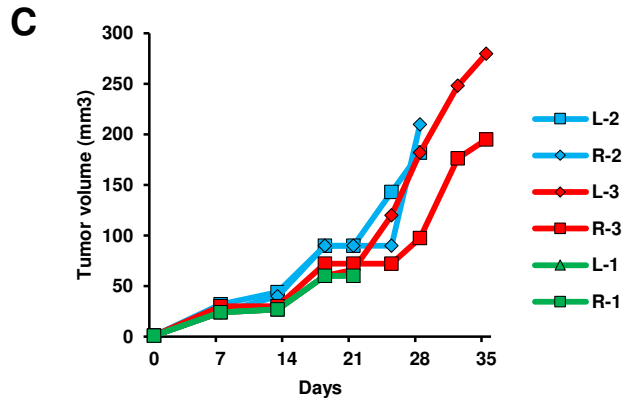
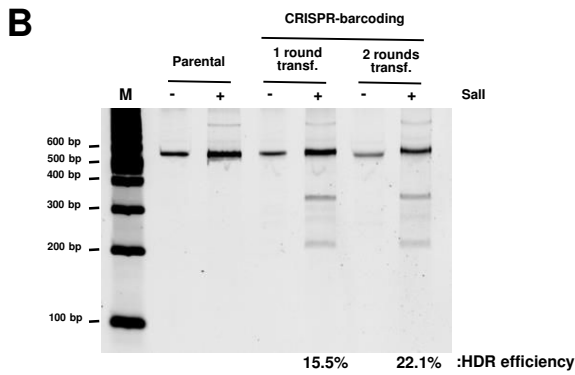
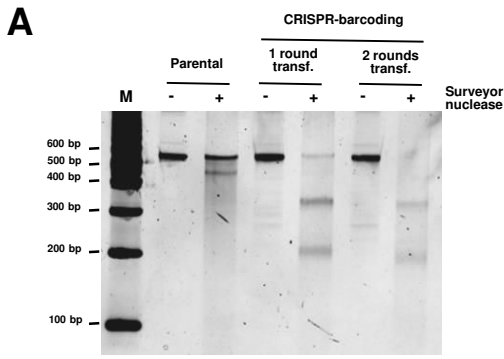


Figure S7, related to Figure 7. Analysis of Intratumor Heterogeneity through Highly Complex CRISPR-Barcodes

(A) AAVS1 surveyor assay from gDNA derived from BT474 cells before (parental) or after one or two rounds of CRISPR-barcoding transfection using two different sgRNAs.

(B) AAVS1 RFLP assay using SalI restriction enzyme and gDNA derived from BT474 cells before (parental) or after one or two rounds of CRISPR-barcoding transfection. HDR efficiency was assessed by comparing the intensity of the bands corresponding to the undigested and SalI-digested amplicons.

(C) Measurement of BT474 tumor volumes in SCID mice.

(D) Schematic representation of the different types of sequences resulting from AAVS1 CRISPR-barcoding.

(E) gDNA was derived from BT474 tumors or BT474 cells maintained in culture, and the AAVS1 barcodes were analyzed by deep sequencing. The plots show comparisons of the normalized frequency of each barcode from the indicated samples (BT474 tumors versus corresponding BT474 cells in culture; BT474 tumors originated in the same or different animals). Linear regressions were calculated, and the corresponding coefficients of determination (r -squared) are indicated.

Table S1, related to Figures 1, 4, 5, 6 and 7. HDR efficiency in the different cell models used in the study.

Cell line	Gene	HDR Efficiency
PC9	EGFR	~0.6 % (a)
PC9	KRAS	~1 % (a)
HCT-116	TP53	~3.5 % (a)
DLD-1	APC	~9 % (a)
Kelly	ALK	~0.4 % (a)
BT474	AAVS1	~22 % (b)
PC9	AAVS1	~8 % (b)

- (a) As assessed by deep sequencing (percentage of reads containing the barcodes vs the endogenous sequence).
- (b) As assessed by RFLP (to increase efficiency, the cells were subjected to 2-3 rounds of transfection using two different sgRNAs).

Table S2, related to the Experimental Procedures. List of sgRNA target sequences and ssODNs used for CRISPR-barcoding.

sgRNA target	
Name	Sequence 5'-3'
sgAPC(A)	AGGGCTATCTGGAAGATCAC
sgAPC(B)	CCATGCAGTGGAAATGGTAAG
sgTP53(A)	GGGCAGCTACGGTTTCCGTC
sgTP53(B)	CAGAATGCAAGAAGCCCAGA
sgTP53(R273)	ACTGGGACGGAACAGCTTTG
sgALK(A)	TGCAGCGAACAATGTTCTGG
sgALK(B)	AGAACATTGTTGCTGCATT
sgEGFR(A)	CTGCGTGATGAGCTGCACGG
sgEGFR(B)	GAGCTGCGTGATGAGCTGCA
sgKRAS(A)	GAATATAAACTTGTGGTAGT
sgKRAS(B)	CTTGTGGTAGTTGGAGCTGG
sgEML4	GACCTGAACAGCAAGTTTGT (a)
sgALK(Chr. Inv.)	GGCCTTGCTGAAACTTCCTT (a)
sgAAVS1(A)	GTCCCCTCCACCCACAGTG
sgAAVS1(B)	GTCACCAATCCTGTCCCTAG
ssODN	
Name	Sequence 5'-3' (b)
APC-WT	CACTTGATAGTTTTGAGAGTCGTTTCGATTGCCAGCTCCGTTCCAGAGTGA ACCATGCAGTGGAAATGGTAAGTgcgATcAtAGtCCgAGcGAcCTTCCAGAT AGCCCTGGACAAACCATGCCACCAAGCAGAAGTAAAACACCTCCACCAC CTCCTCAAACAG
APC-MUT	CACTTGATAGTTTTGAGAGTCGTTTCGATTGCCAGCTCCGTTCCAGAGTGA ACCATGCAGTGGAAATGGTAAGTtGacTcgagAattcAGTGATCTTCCAGATAG CCCTGGACAAACCATGCCACCAAGCAGAAGTAAAACACCTCCACCACCT CCTCAAACAG
TP53-WT	CTCCTACACCGGCGGCCCTGCACCAGCCCCCTCCTGGCCCCTGTCAT CTTCTGTCCCTTCCCAGAAAACCTACCAGGGCAGCTAtGGcTTtCGcCTcG GaTTCTTGcATTCTGGGACAGCCAAGTCTGTGACTTGCACGGTCAGTTG CCCTGAGGGGCTGGCTTCCATGAGACTTCAATGCCTGG
TP53-STOP	CTCCTACACCGGCGGCCCTGCACCAGCCCCCTCCTGGCCCCTGTCAT CTTCTGTCCCTTCCCAGAAAACCTACCAGGGCAGCTACGGTtgagaattcT CTTCTTGcATTCTGGGACAGCCAAGTCTGTGACTTGCACGGTCAGTTGC CCTGAGGGGCTGGCTTCCATGAGACTTCAATGCCTGG
TP53-R273H	TTTTAAATGGGACAGGTAGGACCTGATTTCCCTTACTGCCTCTTGCTTCTC TTTTCCATCCTGAGTAGTGGTAATCTACTGGGcCGaAAAtAGtTtCGAaGTa CaTGTtTGcCTGTcCTGGGAGAGACCGGCGCACAGAGGAAGAGAAT CTCCGCAAGAAAGGGGAGCCTCACCACGAGCTGCCCC
TP53-R273R	TTTTAAATGGGACAGGTAGGACCTGATTTCCCTTACTGCCTCTTGCTTCTC TTTTCCATCCTGAGTAGTGGTAATCTACTGGGACGcAAAtAGCTTTGAaGT cCGaGTaTgCGCaTgCcCTGGGAGAGACCGGCGCACAGAGGAAGAGAAT CTCCGCAAGAAAGGGGAGCCTCACCACGAGCTGCCCC

ALK-F1174F	GTAAC TTTGTATCCTG TTCCTCC CAGTTTAA GATTTGCC CAGACTCAGCT CAGTTAATTTTGGTTACATCCCTCTCTGCTCTGCAGCAAgtTtAAAtCAtCAa AACATaGTTTCGCTGCATTGGaGTGAGCCTGCAATCCCTGCCCCGGTTCA TCCTGCTGGAGCTCATGGCGGG
ALK-F1174L	GTAAC TTTGTATCCTG TTCCTCC CAGTTTAA GATTTGCC CAGACTCAGCT CAGTTAATTTTGGTTACATCCCTCTCTGCTCTGCAGCAAATTaAAAtCAtCA GAAtATcGTcCGcTGCATTGGaGTGAGCCTGCAATCCCTGCCCCGGTTCA TCCTGCTGGAGCTCATGGCGGG
ALK-STOP	GTAAC TTTGTATCCTG TTCCTCC CAGTTTAA GATTTGCC CAGACTCAGCT CAGTTAATTTTGGTTACATCCCTCTCTGCTCTGCAGCAAAtagtgaattatcGA ACATTGTTTCGCTGCATTGaGGTGAGCCTGCAATCCCTGCCCCGGTTTCAT CCTGCTGGAGCTCATGGCGGG
ALK-F1174L-SWP	GTAAC TTTGTATCCTG TTCCTCC CAGTTTAA GATTTGCC CAGACTCAGCT CAGTTAATTTTGGTTACATCCCTCTCTGCTCTGCAGCAAgtTtAAAtCAtCAa AACATaGTTTCGCTGCATTGGaGTGAGCCTGCAATCCCTGCCCCGGTTCA TCCTGCTGGAGCTCATGGCGGG
ALK-F1174F-SWP	GTAAC TTTGTATCCTG TTCCTCC CAGTTTAA GATTTGCC CAGACTCAGCT CAGTTAATTTTGGTTACATCcCTCTCTGCTCTGCAGCAAATTtAAAtCAtCAG AAtATcGTcCGtTGCATTGGaGTGAGCCTGCAATCCCTGCCCCGGTTTCATC CTGCTGGAGCTCATGGCGGG
EGFR-T790T	CTCCCTCCCTCCAGGAAGCCTACGTGATGGCCAGCGTGGACAACCCCC ACGTGTGCCGCCTGCTGGGCATCTGCCTCACCTtACaGTcCaACTgAtA CcCAGCTCATGCCCTTCGGCTGCCTCCTGGACTATGTCCGGGAACACAA AGACAATATTGGCTCCCAGTACCTGCTCAACTGGTGTGTG
EGFR-T790M	CTCCCTCCCTCCAGGAAGCCTACGTGATGGCCAGCGTGGACAACCCCC ACGTGTGCCGCCTGCTGGGCATCTGCCTCACCTtACtGTaCAGCTtAtaA tGCAaCTgATGCCCTTCGGCTGCCTCCTGGACTATGTCCGGGAACACAAA GACAATATTGGCTCCCAGTACCTGCTCAACTGGTGTGTG
KRAS-G12G	TAACCTTATGTGTGACATGTTCTAATATAGTCACATTTTCATTATTTTTATT ATAAGGCCTGCTGAAAATGACTGAATATAAgtTgGTaGTgGTTGGcGCaGG aGGCGTAGGCAAGAGTGCCTTGACGATACAGCTAATTCAGAATCATTTT GTGGACGAATATGATCCAACAATAGAGGTA
KRAS-G12D	TAACCTTATGTGTGACATGTTCTAATATAGTCACATTTTCATTATTTTTATT ATAAGGCCTGCTGAAAATGACTGAATATAAgtTaGTtGTcGTTGGtGCcGac GGCGTAGGCAAGAGTGCCTTGACGATACAGCTAATTCAGAATCATTTT TGGACGAATATGATCCAACAATAGAGGTA
EML4-ALK	AGTGAACACAGTTGTGTTGTTCAATTTTAAAGGATTTTATAGATGATAAAT ATTGATGTAAGTGGAGACAGTTGACCTGAACAGCAAGTTCTTAGGCTCC ATGGCACCCAGGGTGTTCACCCAACCTTCCCTCCCTCCCTCGTTTAC GTGGGGTTATACTTGCAACACAGTCTGCTGG
AAVS1-deg	GTGTCCCCGAGCTGGGACCACCTTATATTCCCAGGGCCGGTTAATGTG GCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCACANNNNNatN NNNgtcGACAGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCT TCCTAGTCTCCTGATATTGGGTCTAACCCACCTCCTGTTAGGCA

(a) Sequence from Choi and Meyerson (2014), Nat Commun 5, 3728

(b) Mutations compared to the endogenous sequence are indicated in lowercase letters

Table S3, related to the Experimental Procedures. List of primers used for qPCR, surveyor assay and deep sequencing.

qPCR primers	
Name	Sequence 5'-3'
APC_Ctrl_FW	GCTGAAGATCCTGTGAGCGA
APC_Ctrl_RV	TTGTGCCTGGCTGATTCTGA
APC_cbc_FW	AGACCCCACTCATGTTTAGCAG
APC_cbc_WT_RV	GGTCGCTCGGACTAATGATCG
APC_cbc_STOP_RV	GGAAGATCACTGAATTCTCGAGTCA
TP53_Ctrl_FW	GACCCAGGGTTGGAAGTGTC
TP53_Ctrl_RV	CTAGGATCTGACTGCGGCTC
TP53_cbc_FW	CAGGTCCAGATGAAGCTCCCA
TP53_cbc_WT_RV	GAATCCGAGGCGAAAGCCA
TP53_cbc_STOP_RV	TCCCAGAATGCAAGAAGAGAATTCTC
TP53_cbc_R273_FW	CAAGGGTGGTTGGGAGTAGAT
TP53_cbc_R273R_RV	GGGCATGCGCATACTCGG
TP53_cbc_R273H_RV	CAAACATGTACTTCGAAACTATTTCCG
ALK_Ctrl_FW/ ALK_cbc_FW	GCCCATGTTTACAGAATGCCTTT
ALK_Ctrl_RV	CCCCAATGCAGCGAACAATG
ALK_cbc_WT_RV	GCAGCGAACTATGTTTTGATGATTAAC
ALK_cbc_MUT_RV/ ALK_cbc_WT_SWP_RV	CAATGCAACGGACGATATTCTGA
ALK_cbc_STOP_RV	CAGCGAACAATGTTTCGATAATTCCT
ALK_cbc_MUT_SWP_RV	GCAGCGAACTATGTTTTGATGATTAAAC
EGFR_Ctrl_FW	TGCTTCCCCCATTGAGGACT
EGFR_Ctrl_RV	CTCCTTGACCTCCTCACTG
EGFR_cbc_RV	CCTTCCCTGATTACCTTTGCGA
EGFR_cbc_T790T_FW	CACCTCTACAGTCCAAGTATTACC
EGFR_cbc_T790M_FW	CTCTACTGTACAGCTTATAATGCAACTG
KRAS_Ctrl_FW	ATCCTTTGAGAGCCTTTAGCCG
KRAS_Ctrl_RV	CCAGTTGACTGCAGACGTGTA
KRAS_cbc_G12G_RV	CTGCGCCAACCACTACCAAC
KRAS_cbc_G12D_RV	GTCGGCACCAACGACAAC
KRAS_cbc_FW	TATTAAGGTTACTGGTGGAGTATTTGAT
EML4_ALK_FW	CTCGTGGTAACATCAGAACAGAGA
EML4_ALK_RV	TGCCATGGAGCCTAAGAACTTG
PCR primers for surveyor assay and RFLP	
Name	Sequence 5'-3'
APC fw	AGCTGAAGATCCTGTGAGCG
APC rv	AATGGCTCATCGAGGCTCAG
AAVS1 fw	CATCTTCCAGGGGTCCGAGA
AAVS1 rv	TCAGTGAAACGCACCAGACG
PCR primers for APC deep sequencing	
Name	Sequence 5'-3'
APC 1_FW	GATTGCCAGCTCCGTTGAG
APC 1_RV	ATTTACTGCAGCTTGCTTAGGT
APC 2_FW	TGCTCAGACACCCAAAAGT
APC 2_RV	GTTTTACTTCTGCTTGGTGGC

Other PCR primers used for deep sequencing	
Name	Sequence 5'-3'
AAVS1ds_FW	<i>AATGATACGGCGACCACCGAGATCTACA</i> <u>CTCTTTCCCTACACGACGCTCTTCCGATC</u> <u>TGTTAATGTGGCTCTGGTTCTGG*</u>
AAVS1ds_RV	<i>CAAGCAGAAGACGGCATACGAGATCGTG</i> AT <u>GTGACTGGAGTTCAGACGTGTGCTCT</u> <u>CCGATCTGGTGGGGTTAGACCCAATAT</u> C*
EGFR_FW	<i>AATGATACGGCGACCACCGAGATCTACA</i> <u>CTCTTTCCCTACACGACGCTCTTCCGATC</u> <u>TCACGTGTGCCGCTGCTGG</u>
EGFR_RV	<i>CAAGCAGAAGACGGCATACGAGATCGTA</i> CG <u>GTGACTGGAGTTCAGACGTGTGCTCT</u> <u>TCCGATCTCCTCCCTGATTACCTTTGCG</u> A
KRAS_FW	<i>AATGATACGGCGACCACCGAGATCTACA</i> <u>CTCTTTCCCTACACGACGCTCTTCCGATC</u> <u>TTTTATTATAAGGCCTGCTGAAAATG</u>
KRAS_RV	<i>CAAGCAGAAGACGGCATACGAGATCCAC</i> TC <u>GTGACTGGAGTTCAGACGTGTGCTCTT</u> <u>CCGATCTAGAATGGTCCTGCACCAGTAA</u>
TP53_FW	<i>AATGATACGGCGACCACCGAGATCTACA</i> <u>CTCTTTCCCTACACGACGCTCTTCCGATC</u> <u>TCCTGTCATCTTCTGTCCCTTCC</u>
TP53_RV	<i>CAAGCAGAAGACGGCATACGAGATGCTA</i> CC <u>GTGACTGGAGTTCAGACGTGTGCTCT</u> <u>TCCGATCTAGAAATGCAGGGGGATACGG</u>
ALK_FW	<i>AATGATACGGCGACCACCGAGATCTACA</i> <u>CTCTTTCCCTACACGACGCTCTTCCGATC</u> <u>TAATTTTGGTTACATCCCTCTCTGC</u>
ALK_RV	<i>CAAGCAGAAGACGGCATACGAGATATCA</i> GT <u>GTGACTGGAGTTCAGACGTGTGCTCTT</u> <u>CCGATCTAAGGACTTGAGGTCTCCCC</u>

The flowcell sequences are in italics, the Illumina sequencing primers are underlined and the sequence of the indexes used is in bold. For AAVS1, each sample was amplified using a RV primer containing a distinct Illumina index.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Invasion Assay

PC9 cells containing the EGFR-T790M and -T790T barcodes pre-treated or not with gefitinib (0.5 μ M) for 48 hr were seeded in matrigel 6-well invasion chambers (8 μ M pores; BD Biosciences) at a density of 1.5×10^6 cells per chamber, in the presence or the absence of gefitinib (0.5 μ M) and incubated for 24 hr. As a control, a fraction of the cells was seeded in parallel in regular wells. The cells inside and outside the chambers were then trypsinized and gDNA was extracted for qPCR.

Mouse Xenografts

PC9 cells containing the EGFR-T790, KRAS-G12 and EML4-ALK barcodes were resuspended in PBS and subcutaneously inoculated in the left and right flank (2×10^6 cells per site) of male SCID mice (NCI). The size of the tumors was measured by caliper every 3-4 days. Sixteen days after injection, when the tumors were palpable, half of the mice were treated 5 days a week with gefitinib (25 mg/kg/day) by gavage. When the size of at least one of the tumors reached the arbitrary volume of 550 mm³, the mice were sacrificed and the tumors were dissected. Each tumor was cut in two halves, one for gDNA extraction and one for storing. BT474 cells (5×10^6 cells per site) were mixed with matrigel and inoculated bilaterally in the mammary fat pad of female SCID mice previously implanted with 17 β -estradiol pellets (Innovative Research of America). Mice were sacrificed 3, 4 and 5 weeks after injection and gDNA was extracted from the tumors for deep sequencing analysis. At each time point, gDNA was extracted from the same batch of BT474 cells kept in culture. All animal experiments were approved and performed according to the relevant regulatory standards set by Mount Sinai's Animal Care and Use Committee.

Lentiviral Production and Inhibitors

The lentiviral Wnt reporter was generated by cloning 14 Wnt-responsive elements upstream of a minimal promoter (Grumolato et al., 2013), followed by a destabilized GFP (Matsuda and Cepko, 2007). The construct also contains the puromycin resistance gene driven by a different promoter. Lentivirus was produced as previously described (Grumolato et al., 2010). Briefly, 293T cells were cotransfected with the Wnt-responsive lentiviral vector, pCMV Δ 8.91 and pMD VSV-G plasmids. Two days after transfection, the conditioned medium was collected, purified by centrifugation, supplemented with 8 μ g/mL polybrene, and added for overnight incubation to freshly plated DLD-1 cells. Two days after infection, the cells were selected in 2 mg/mL puromycin.

Cell Sorting by FACS

DLD-1 cells transduced with the Wnt-responsive reporter were sorted based on the level of GFP expression. Briefly, harvested cells were resuspended in sterile FACS-sorting buffer (D-PBS medium, 2 mM EDTA, 0.5% BSA), filtered on 30 μ m cell-strainer (BD Biosciences) to remove cell aggregates and isolated by FACS using a FACSAria-III Cytometer (BD Biosciences). DLD-1 cells containing the Wnt reporter were first FACS sorted to remove the fraction of GFP-negative cells, and the resulting cell population was transfected for APC CRISPR-barcoding. Five days after transfection, GFP^{hi} (defined as the 10% of the cells displaying the highest fluorescence) and GFP^{lo} cells (defined as the 10% of the cells displaying the lowest fluorescence) were isolated by FACS, and gDNA was extracted from sorted and un-sorted cells for qPCR analysis. Cells were maintained at 4°C during the entire sorting procedure. Reanalysis of sorted fractions consistently showed that more than 95% of the sorted cells were viable and corresponded to the defined GFP^{hi} or GFP^{lo} gates. Analysis and graphical representation of the data were performed using FlowJo software (Tree stars).

Deep Sequencing

PCR was performed from DLD-1 cell gDNA samples using Herculase II Fusion DNA Polymerase and two couples of primers (Table S3), designed to amplify two overlapping regions of the APC gene spanning 245 and 220 bp, respectively, both including the CRISPR/Cas9 targeted sequence. The samples were prepared from 1 μ g of purified amplicons using SPRIworks System I automaton (Beckman). Each couple of amplicons was indexed with the same adapter sequence (Craig et al., 2008), and the enriched libraries were sequenced on a Genome Analyzer IIx (Illumina) with 76 bp paired-end reads. Image analysis and base calling were performed by Real-Time Analysis (RTA 1.9) and CASAVA software (v1.8.2, Illumina). The mean read count passing filters for each sample was at least 160,000, and 94% of the sequenced bases have a QScore above 30. Motifs of interest were counted in the FastQ files.

For AAVS1 barcode sequencing, gDNA was isolated from BT474 cells and tumors tissues using the Epicenter protocol (MC85200) (Jayaprakash et al., 2015). DNA was eluted in 100 μ l of TE buffer and checked for quality and quantity using a 1% agarose and qubit fluorometer respectively (Qubit dsDNA BR Assay Kit). The eluted DNA was further heated at 70°C for 30 min to inactivate any leftover proteinase K that was

introduced in the DNA isolation protocol. One microgram of DNA was used for PCR amplifications. 25 cycle-PCR was performed using primers containing the flow cell sequences and an index for in silico segregation of the reads. The total volume of PCR reaction was 100 μ l using OneTaq® 2X Master Mix (NEB M0482S). For PC9 cells, DNA was extracted using the NucleoSpin Tissue kit and 25 cycle PCR amplification was performed from 0.8 μ g of DNA using Herculase II Fusion DNA Polymerase (Agilent Technologies). The PCR products were run on a 2% low melt agarose and a 300 bp product was size selected and eluted from the gel. The PCR product was then quantified using an Agilent bioanalyzer and qubit fluorometer. Deep sequencing was then performed using the Illumina platform (75 bp SE sequencing), and each sample receiving >250,000 reads.

Surveyor and RFLP Assays

The sequence of the primers used to amplify a ~800 bp region of the APC and AAVS1 genes encompassing the CRISPR/Cas9 targeted sites is provided in Table S3. PCR was performed using Herculase II Fusion DNA Polymerase and the amplicon was purified using PCR clean-up kit (Macherey-Nagel) and eluted with 30 μ L of TE buffer. Surveyor assay was performed as previously described (Ran et al., 2013). Briefly, 15 μ L of purified-PCR products were then mixed with Taq polymerase PCR buffer and subjected to a re-annealing process to enable DNA heteroduplex formation. After re-annealing, products were treated with SURVEYOR nuclease and SURVEYOR enhancer S (Transgenomic). The remaining 15 μ L of the purified amplicon was used as a control. For RFLP assay, the AAVS1 amplicons were digested in the presence or in the absence of SalI (Promega). For both surveyor and RFLP assays, the reaction products were analyzed on a 10% Acrylamide-TBE gel stained with SYBR-Gold (Life Technologies) and imaged with a Chemidoc gel imaging system (Bio-Rad).

RNA Extraction and Reverse Transcription

Total RNA was isolated using the Tri-Reagent (Sigma-Aldrich) and the RNeasy kit (Qiagen), DNase digested and reverse-transcribed using the Improm-II Reverse Transcription System (Promega).

Immunoblot

Cells were washed once with phosphate-buffered saline (PBS) and lysed on ice in lysis buffer containing 50 mM Hepes pH 7.6, 150 mM NaCl, 5 mM EDTA, 1% Nonidet P-40, 20 mM NAF, 2 mM sodium orthovanadate, supplemented with protease inhibitor mini tablets (Thermo Scientific). Lysates were cleared by centrifugation at 14,000 g during 15 min at 4°C and protein concentrations were determined by using the Bio-Rad protein assay (Bio-Rad). Sodium dodecyl sulfate (SDS) loading buffer was added to equal amounts of lysate, followed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and transfer to polyvinylidene fluoride (PVDF) membrane (Thermo Scientific). The membranes were probed using Mouse anti-TP53 (Cell Signaling) and mouse anti- β -catenin (BD Biosciences) antibodies and analyzed by chemiluminescence using ECL Western Blotting Substrate (Thermo Scientific).

Supplemental references

Craig, D.W., Pearson, J.V., Szelinger, S., Sekar, A., Redman, M., Corneveaux, J.J., Pawlowski, T.L., Laub, T., Nunn, G., Stephan, D.A., *et al.* (2008). Identification of genetic variants using bar-coded multiplexed sequencing. *Nat. Methods* 5, 887-893.

Grumolato, L., Liu, G., Harembaki, T., Mungamuri, S.K., Mong, P., Akiri, G., Lopez-Bergami, P., Arita, A., Anouar, Y., Mlodzik, M., *et al.* (2013). beta-Catenin-independent activation of TCF1/LEF1 in human hematopoietic tumor cells through interaction with ATF2 transcription factors. *PLoS Genet.* 9, e1003603.

Grumolato, L., Liu, G., Mong, P., Mudbhary, R., Biswas, R., Arroyave, R., Vijayakumar, S., Economides, A.N., and Aaronson, S.A. (2010). Canonical and noncanonical Wnts use a common mechanism to activate completely unrelated coreceptors. *Genes Dev.* 24, 2517-2530.

Jayaprakash, A.D., Benson, E.K., Gone, S., Liang, R., Shim, J., Lambertini, L., Toloue, M.M., Wigler, M., Aaronson, S.A., and Sachidanandam, R. (2015). Stable heteroplasmy at the single-cell level is facilitated by intercellular exchange of mtDNA. *Nucleic Acids Res.* 43, 2177-2187.

Matsuda, T., and Cepko, C.L. (2007). Controlled expression of transgenes introduced by in vivo electroporation. *Proc. Natl. Acad. Sci. USA* 104, 1027-1032.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281-2308.

Modeling intratumor heterogeneity through CRISPR-barcodes


Alexis Guernet, Stuart A. Aaronson, Youssef Anouar and Luca Grumolato

Molecular & Cellular Oncology (2016), VOL 3, NO. 6, e1227894

Dans ce commentaire, nous résumons le principe et l'intérêt de notre approche CRISPR-barcoding.

COMMENTARY

Modeling intratumor heterogeneity through CRISPR-barcodes

Alexis Guernet^{a,b}, Stuart A. Aaronson^c, Youssef Anouar^{a,b}, and Luca Grumolato ^{a,b}

^aNormandie Univ, UNIROUEN, INSERM, DC2N, Rouen, France; ^bInstitute for Research and Innovation in Biomedicine, Rouen, France; ^cDepartment of Oncological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA

ABSTRACT

We have devised a barcoding strategy to recapitulate cancer evolution through the emergence of subclonal mutations of interest, whose effects can be monitored in a dynamic manner. This approach can be easily adapted for a variety of applications, including combined modeling of multiple mechanisms of drug resistance or repair of oncogenic driver mutations in addicted cancer cells.

ARTICLE HISTORY

Received 17 August 2016
Revised 18 August 2016
Accepted 19 August 2016

KEYWORDS

ALK; APC; CRISPR/Cas9; EGFR; genetic barcoding; non-small cell lung cancer; resistance; TP53; tumor heterogeneity

Cancer is an evolutionary process in which the stepwise accumulation of genetic alterations is shaped by Darwinian selection. As a result, each tumor is composed of a complex mixture of clonal cell subpopulations containing a partially overlapping, but distinct, pattern of driver and passenger mutations. Such intratumor heterogeneity has dramatic consequences, not only for cancer progression and metastatic spread, but also for resistance to therapy.^{1,2} While snapshots of the clonal architecture of a particular tumor at a given stage can now be obtained through deep sequencing and mathematical modeling, such complexity is generally not taken into account when investigating the effects of a particular oncogenic mutation. We recently devised a novel approach based on new technologies for DNA editing to recapitulate and trace the emergence of a new mutation in a subset of cancer cells, thus enabling functional studies on a gene of interest in a context that mimics intratumor heterogeneity.³

Originally an adaptive immune system in prokaryotes, CRISPR (clustered regularly interspaced short palindromic repeats) has been engineered into a new powerful tool for genome editing.^{4,5} This system is composed of the Cas9 nuclease from *S. Pyogenes* and a short RNA sequence, the single-guide RNA (sgRNA). When co-expressed in cells, Cas9 and the sgRNA form a complex that specifically recognizes a particular DNA sequence through Watson-Crick pairing and promotes its cleavage. The double-strand DNA break induced by CRISPR/Cas9 can trigger two distinct cellular mechanisms for DNA repair that can be exploited for DNA editing: error-prone non-homologous end-joining (NHEJ) and high-fidelity homology-directed repair (HDR). DNA repair through NHEJ frequently generates insertions or deletions (indels), which can alter the frame of a coding sequence and result in gene inactivation. In HDR, a donor DNA co-introduced into the cells functions as a template for precise repair; through appropriate design of the donor DNA this mechanism can be used to generate a wide

range of genetic modifications, including specific point mutations or the insertion of an entire gene. Depending on the extent of the desired modification, a single-stranded DNA oligonucleotide (ssODN) or a double-stranded DNA targeting construct can be used as donor DNA for HDR.^{4,5}

Despite the undeniable potential of this new technology, two major intrinsic limitations must be considered when applying CRISPR/Cas9. First, in certain contexts this system can tolerate a few mismatches between the sgRNA and its target sequence, which can result in off-target DNA cleavage.⁴⁻⁷ The second limitation is related to the fact that the efficiency of CRISPR/Cas9-mediated DNA editing can vary substantially depending on the cell model, the sgRNA, and the targeted DNA. Thus, in a given experiment only a fraction of cells within a population will contain the desired genetic modification. Hence, derivation and subsequent analysis of a certain number of individual clones constitute an almost inevitable step⁸ that is obviously not compatible with modifications that have a negative impact on cell growth.

We used CRISPR/Cas9 technology to insert a potentially functional modification in the sequence of a gene of interest coupled to a series of silent point mutations, serving as a genetic label for cell tracing. Because of the low efficiency of HDR, within a mass population the modified allele would be introduced only in a small fraction of tumor cells, thus mimicking the emergence of a new mutant clone. In parallel, a distinct subpopulation was tagged with a second barcode consisting of different silent mutations, which provided an internal control for possible CRISPR off-target effects. After exposing the cells to a given selective condition, we assessed the effects of the mutation of interest by measuring the proportion of each barcode in genomic DNA by real-time quantitative PCR or deep sequencing (Fig. 1). We applied this strategy to recapitulate different mechanisms of resistance to inhibitors of the epidermal growth factor receptor (EGFR) in non-small cell lung cancer

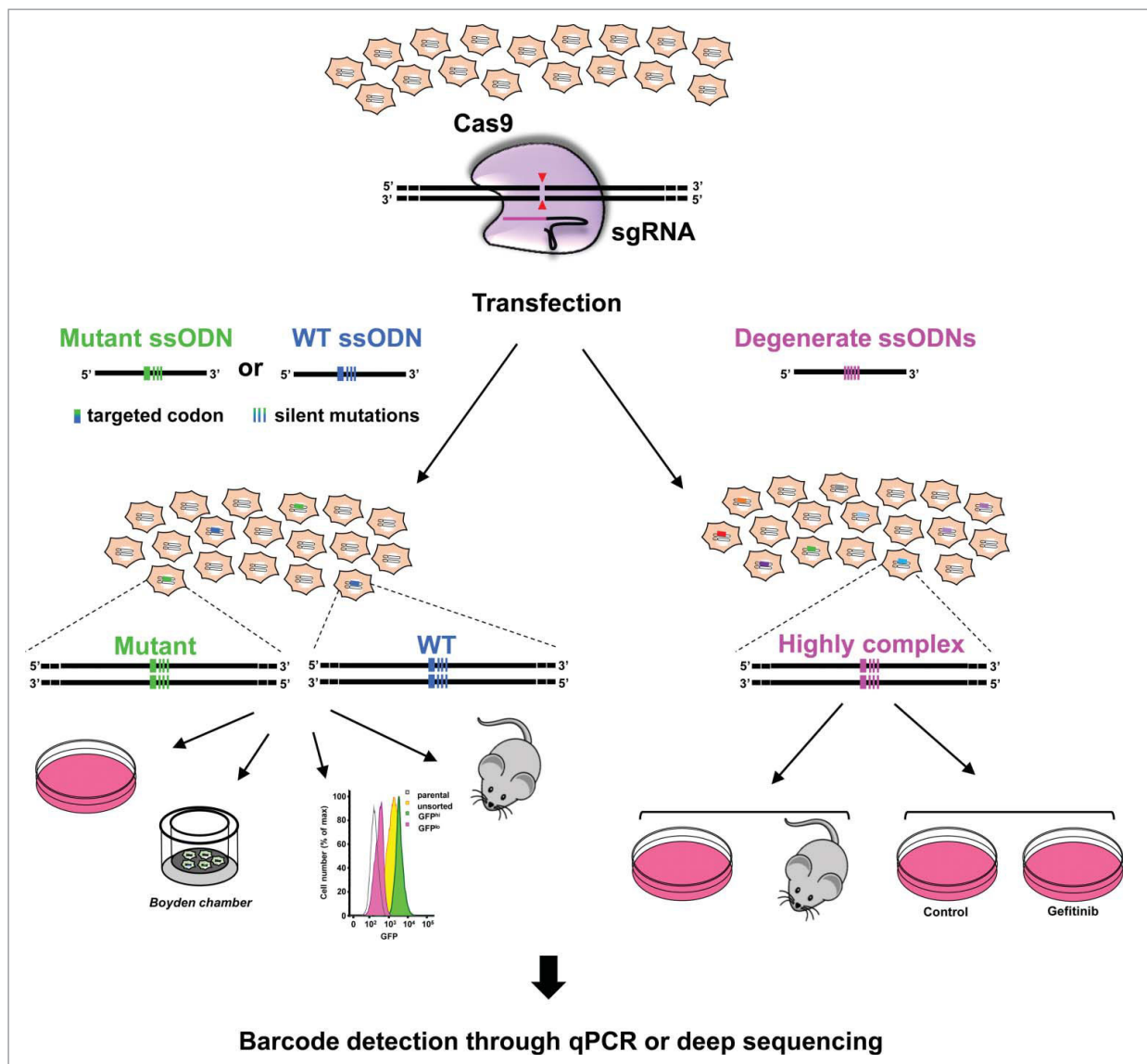


Figure 1. CRISPR-barcoding strategy. Cells are co-transfected with a vector for Cas9 and sgRNA expression, together with a single-stranded DNA oligonucleotide (ssODN) containing the mutation of interest (Mutant) coupled with a series of silent mutations. In parallel, a ssODN containing only silent mutations (WT) is used as an internal control for off-target DNA cleavage by Cas9. The cells are then exposed to a given selective condition and the effects of the mutation on cell growth, invasion, signaling activation, or tumorigenicity can be assessed by quantitative PCR (qPCR) or deep sequencing. A similar approach was used to introduce highly complex barcodes through a degenerate ssODN in the different clones that constitute a mass population of cancer cells. We used this strategy to compare the relative fitness of breast cancer cells to grow *in vivo* after injection in immunodeficient mice and to investigate the effects of the EGFR inhibitor gefitinib in lung cancer cells.

cells, and generated a multiplex model for *in vitro* and *in vivo* investigation of new therapeutic strategies to prevent or delay propagation of resistant cells. Using CRISPR-barcoding, we also manipulated the endogenous sequence of different oncogenes and tumor suppressors, including adenomatous polyposis coli (*APC*), tumor protein p53 (*TP53*) and anaplastic lymphoma kinase (*ALK*), and showed that this approach can be used to inactivate a gene through insertion of a STOP codon or to repair oncogenic mutations in addicted cancer cells. Finally, through insertion of a degenerate sequence at a specific genome location, we simultaneously traced the fate of several thousand different clones within a mass population of cancer cells upon inoculation in immunodeficient mice or in the presence of a therapeutic agent (Fig. 1), demonstrating that CRISPR-barcoding can represent a convenient alternative

strategy of randomly integrating lentiviral libraries^{9,10} to investigate intratumor heterogeneity.

In conclusion, our study establishes a novel and unique means to recapitulate genetic heterogeneity in virtually all types of cancer cells that can be easily implemented to assess effects on cell growth/survival, invasion, transcriptional activity, resistance/sensitivity to a particular agent, and tumorigenicity. Compared with other CRISPR/Cas9-based approaches, our strategy provides several major advantages, including (i) the ability to investigate genome modifications that negatively affect cell growth; (ii) an internal control for potential off-target cleavage by the sgRNA-Cas9 complex; (iii) the use of a cell mass-population, thus preventing potential bias due to clonal selection; (iv) the fact that the system does not require high efficiency of DNA editing, which substantially broadens the choice

of cell models that can be used; (v) ease of multiplexing; (vi) better cost- and time-effectiveness.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Funding

This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM), Université de Rouen Normandie, Ligue contre le Cancer de Haute Normandie, the National Cancer Institute (P01CA080058), and the Breast Cancer Research Foundation. A.G is recipient of a doctoral fellowship from the Normandy Region. L.G. was supported by a Chair of Excellence program from INSERM and Université de Rouen Normandie.

ORCID

Luca Grumolato  <http://orcid.org/0000-0001-8231-3032>

References

1. Kleppe M, Levine RL. Tumor heterogeneity confounds and illuminates: assessing the implications. *Nat Med* 2014; 20:342-4; PMID:24710377; <http://dx.doi.org/10.1038/nm.3522>
2. McGranahan N, Swanton C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell* 2015; 27:15-26; PMID:25584892; <http://dx.doi.org/10.1016/j.ccell.2014.12.001>
3. Guernet A, Mungamuri SK, Cartier D, Sachidanandam R, Jayaprakash A, Adriouch S, Vezain M, Charbonnier F, Rohkin G, Coutant S, et al. CRISPR-Barcoding for Intratumor Genetic Heterogeneity Modeling and Functional Analysis of Oncogenic Driver Mutations. *Mol Cell* 2016; 63:526-38; PMID:27453044; <http://dx.doi.org/10.1016/j.molcel.2016.06.017>
4. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 2014; 346:1258096; PMID:25430774; <http://dx.doi.org/10.1126/science.1258096>
5. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 2014; 157:1262-78; PMID:24906146; <http://dx.doi.org/10.1016/j.cell.2014.05.010>
6. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* 2013; 31:822-6; PMID:23792628; <http://dx.doi.org/10.1038/nbt.2623>
7. Kuscu C, Arslan S, Singh R, Thorpe J, Adli M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Bio* 2014; 32:677-83; PMID:24837660; <http://dx.doi.org/10.1038/nbt.2916>
8. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protocols* 2013; 8:2281-308; PMID:24157548; <http://dx.doi.org/10.1038/nprot.2013.143>
9. Bhang HE, Ruddy DA, Krishnamurthy Radhakrishna V, Caushi JX, Zhao R, Hims MM, Singh AP, Kao I, Rakiec D, Shaw P, et al. Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. *Nat Med* 2015; 21:440-8; PMID:25849130; <http://dx.doi.org/10.1038/nm.3841>
10. Nguyen LV, Pellacani D, Lefort S, Kannan N, Osako T, Makarem M, Cox CL, Kennedy W, Beer P, Carles A, et al. Barcoding reveals complex clonal dynamics of de novo transformed human mammary cells. *Nature* 2015; 528:267-71

CRISPR/Cas9 editing of the genome for cancer modeling

Alexis Guernet and Luca Grumolato

Methods (2017), 121-122; 130-137

Dans cette revue, nous examinons les données de la littérature concernant les diverses stratégies CRISPR/Cas9 utilisées pour générer des modèles pour l'étude du cancer.

En particulier, nous avons abordé les sujets suivants :

- L'étude de mutations oncogéniques *in vitro*
- L'utilisation de CRISPR/Cas9 pour récapituler l'évolution tumorale
- La génération de nouveaux modèles *in vivo* de cancer
- L'étude de l'hétérogénéité intra-tumorale par CRISPR-barcoding



CRISPR/Cas9 editing of the genome for cancer modeling



Alexis Guernet, Luca Grumolato*

Normandie Univ, UNIROUEN, INSERM, DC2N, Rouen, France

Institute for Research and Innovation in Biomedicine (IRIB), Rouen, France

ARTICLE INFO

Article history:

Received 1 December 2016

Received in revised form 9 February 2017

Accepted 3 March 2017

Available online 10 March 2017

Keywords:

CRISPR/Cas9

Cancer model

Tumor suppressor

Oncogene

Tumor heterogeneity

CRISPR-barcoding

ABSTRACT

The CRISPR/Cas9 revolution has democratized access to genome editing in many biological fields, including cancer research. Cancer results from the multistep accumulation of mutations that confer to the transformed cells certain biological hallmarks typical of the malignant phenotype. One of the major goals in cancer research is to characterize such mutations and assess their implication in the oncogenic process. Through CRISPR/Cas9 technology, genetic aberrations identified in a patient's tumor can now be easily recreated in experimental models, which can then be used for basic research or for more translational applications. Here we review the different CRISPR/Cas9 strategies that have been implemented to recapitulate oncogenic mutations in both *in vitro* and *in vivo* systems, including novel strategies to model tumor evolution and genetic heterogeneity.

© 2017 Elsevier Inc. All rights reserved.

Contents

1. Introduction	130
2. CRISPR/Cas9 strategies to recapitulate oncogenic alterations in cultured cells	131
3. Modeling tumor evolution through CRISPR/Cas9	132
4. New mouse cancer models	134
5. CRISPR-barcoding	135
6. Concluding remarks	136
Acknowledgments	137
References	137

1. Introduction

In 2012, two papers [1,2] showed how the basic components of an adaptive immune system in prokaryote can be harnessed to specifically cleave DNA *in vitro*, paving the way for a new era in genome editing. A few months later, it was demonstrated in mammalian cells that different types of genetic modifications can be obtained through simple transfection using expression vectors containing the Cas9 endonuclease and a single-guide RNA (sgRNA) [3,4]. In this system, the Cas9-sgRNA complex specifically recognizes a particular DNA sequence through Watson-Crick pairing and promotes its cleavage, thus triggering two distinct cellular

mechanisms for DNA repair. In the first case, the two DNA blunt ends are put together by nonhomologous end-joining (NHEJ), an error prone process that frequently generates insertions or deletions (indels) of a few nucleotides, which can be used to disrupt the frame of a coding sequence. The second mechanism is called homology directed repair (HDR) and it involves a donor DNA as a template for homologous recombination. Through appropriate design of the donor DNA, this mechanism can be exploited to generate a wide range of genetic modifications, including specific point mutations or the insertion of an entire gene [5–7] (Fig. 1).

Cancer is a genetic disease, resulting from the sequential accumulation of mutations capable of driving transformation and, subsequently, progression. A given tumor, depending on the type, can contain several thousand mutations, most of which are defined as passenger, since they do not participate in the oncogenic process, but they are instead a collateral consequence of cancer genetic

* Corresponding author at: Normandie Univ, UNIROUEN, INSERM, DC2N, Rouen, France.

E-mail address: luca.grumolato@univ-rouen.fr (L. Grumolato).

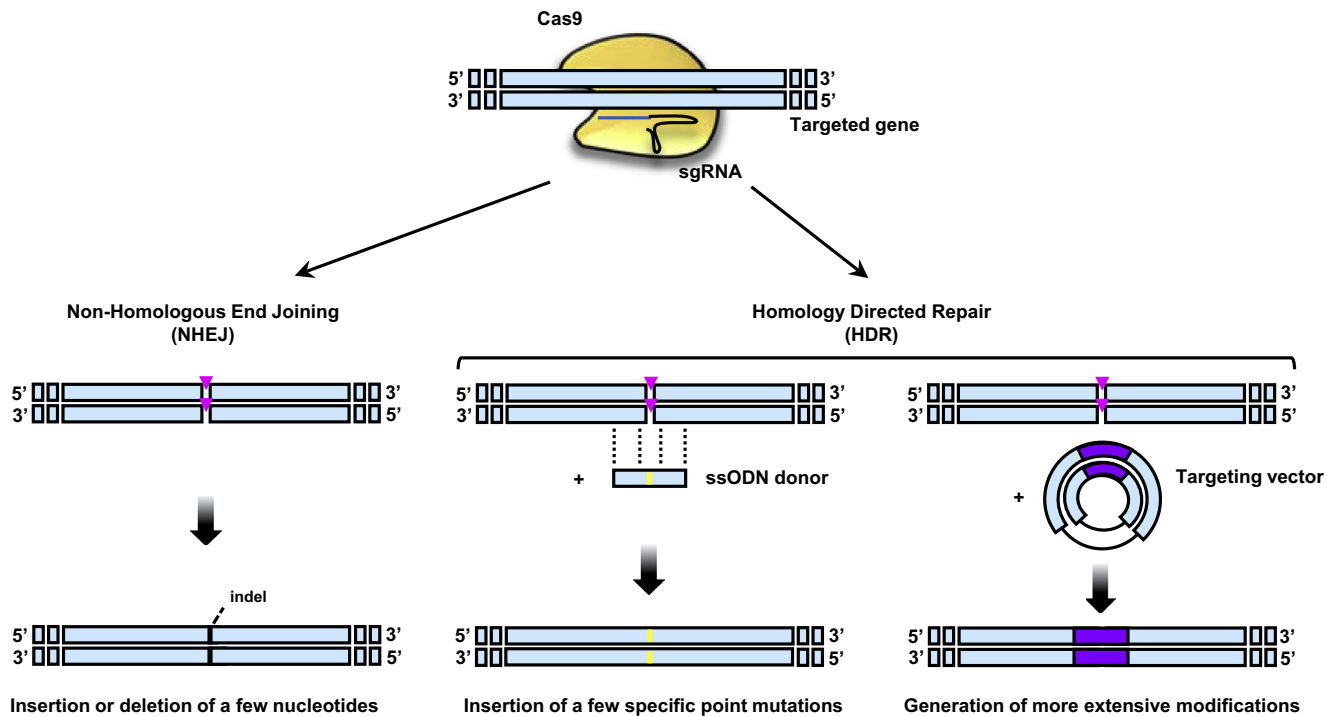


Fig. 1. Cas9 nuclease induces a specific double-strand DNA break resulting in the activation of two cellular mechanisms of DNA repair. Non-homologous end joining (NHEJ) frequently provokes the insertion or deletion of a few nucleotides (indels), which can be used to disrupt the function of a gene of interest through frameshift mutations. In the presence of a template DNA displaying sequence homology to the targeted locus, the DNA break can be fixed through homology directed repair (HDR). By using as a donor DNA a single-strand oligonucleotide (ssODN) or a targeting plasmid, HDR can be applied to introduce few point mutations or more extensive modifications of the genome, respectively.

instability [8,9]. Thanks to its efficiency and accessibility, CRISPR/Cas9 technology now provides the means to easily recreate mutations identified in patients' tumors, thus enabling their functional characterization through both *in vitro* and *in vivo* experimental models.

2. CRISPR/Cas9 strategies to recapitulate oncogenic alterations in cultured cells

Gene inactivation through NHEJ-mediated frameshift mutations is probably the most straightforward application of CRISPR/Cas9 technology. As an example of this type of strategy, Malina et al. used a single lentiviral vector for dual expression of Cas9 and sgRNA to successfully inactivate the *Trp53* tumor suppressor gene in mouse embryonic fibroblasts (MEFs). The selective advantage conferred by the loss of Trp53 function resulted over time in an increased proportion of mutant cells in the mass population of transduced MEFs, while a further enrichment could be obtained through artificial induction of Trp53 by Nutlin-3a treatment. Among the nine loci that were sequenced because of their partial homology with the *Trp53* region recognized by the sgRNA, only one showed detectable levels of off-target cleavage, indicating a high degree of DNA editing specificity [10]. Using a similar approach, another study showed that CRISPR/Cas9-mediated inactivation of *Trp53* in MEFs can promote *in vitro* growth, while concurrent overexpression of mutant H-Ras was required for full transformation, measured as the capacity to form tumors in a xenograft model [11]. Of note, the authors developed a polycistronic lentiviral system allowing expression of several sgRNAs from the same vector, which they used to simultaneously inactivate three different tumor suppressor genes in MEFs, *i.e.* *Trp53*, *Pten* and *Vhl* [11]. Besides disruption of strong tumor suppressors,

gene inactivation through NHEJ-mediated repair has also been applied in the context of more subtle phenotypes. For instance, it was shown that CRISPR/Cas9 knockout of the mixed lineage leukemia 3 (*MLL3*) gene could synergize with other oncogenic events to promote tumorigenesis in mouse hematopoietic stem cells. Consistent with human acute myeloid leukemia data, genetic analysis revealed that the mouse leukemia cells almost invariably contained both mutant and wild-type alleles, implying that *MLL3* acts as a haploinsufficient tumor suppressor in this type of malignancy [12].

Besides the inactivation of a tumor suppressor, CRISPR/Cas9 technology can be implemented for more precise genome editing, including point mutations of a proto-oncogene. For this type of application, Cas9 and the sgRNA are used in combination with a donor DNA containing the modification of interest, flanked by sequences homologous to the targeted locus, serving as a template for HDR. As an example, Harrod et al. generated an estrogen independent, *anti*-estrogen resistant strain of MCF7 breast cancer cells through insertion of the Y537S mutation in the *ERS1* gene, encoding the estrogen receptor [13]. This residue, located in the ligand-binding domain of the receptor, is frequently mutated in advanced breast cancers that become resistant to hormone therapy. As a donor DNA, the authors of this study used a 1.8 kb region of *ERS1* encompassing the Y537 codon, which was cloned into a targeting plasmid and engineered by site-directed mutagenesis to include the Y537S mutation, as well as additional silent mutations to avoid recognition of the template by the sgRNA/Cas9 complex [13]. As a convenient alternative to double strand donor DNAs, a single strand oligonucleotide (ssODN) can be used as a template for HDR when the modification of interest is limited to a few, clustered point mutations [5,6] (Fig. 1). For instance, this strategy was used to introduce a specific mutation in the *XPO1* gene capable of conferring to the encoded protein, exportin-1, resistance to a small

molecule inhibitor that showed promising anticancer properties in preclinical studies [14]. Despite the tremendous potential of HDR-mediated DNA editing, a major drawback of these approaches derives from their relatively low efficiency, which implies the need to select for individual clones containing the mutation of interest.

The capacity of cancer cells to evade therapy through acquisition of novel genetic aberrations is a hallmark of tumor progression, which can be modeled using CRISPR/Cas9 technology. Shalem et al. generated a sgRNA lentiviral library to identify genes whose inactivation can lead to resistance to the BRAF inhibitor vemurafenib in melanoma cells [15]. In a later study, the same group devised a complementary approach to establish vemurafenib resistance through gene overexpression, using a new system based on a catalytically inactive form of Cas9 that, in combination with a specific sgRNA, acts as a transcription factor to promote expression from a genomic sequence of interest [16]. In both the loss-of-function and the gain-of-function strategies, deep sequencing analysis of the pools of sgRNA enriched in the presence of vemurafenib enabled unbiased identification of genes potentially involved in drug resistance. As a control for specificity, the authors selected genes for which at least three distinct targeting sgRNAs were significantly enriched in the presence of the inhibitor [15,16]. Based on the data obtained during their first two screens in melanoma cells, the same group recently generated a different sgRNA lentiviral library targeting the noncoding sequence surrounding three genes previously shown to participate in vemurafenib resistance [17]. This study demonstrates how CRISPR/Cas9 technology can be also applied for high-resolution analysis of DNA functional elements involved in the regulation of gene expression during cancer progression.

While DNA cleavage by Cas9 and a single sgRNA is used to generate indels within a particular sequence, expression of two distinct sgRNAs can induce more substantial alterations of the genome, ranging from the deletion of a particular locus to a chromosomal translocation (Fig. 2). To recapitulate a somatic deletion on chromosome 6 frequently observed in different types of cancer, Sibrashvili et al. transfected different melanoma cell lines with plasmids containing Cas9 and two sgRNAs targeting both ends of a locus encoding two small nucleolar RNAs (snoRNAs). Consistent with a tumor suppressor role of the two snoRNAs associated with their capacity of inhibiting RAS signaling, excision of this genomic region resulted in clones displaying stronger MAPK activation and higher tumorigenic potential [18]. Through a similar approach, Kataoka et al. disrupted the 3'-UTR sequence of the *PD-L1* gene in different mouse and human cell lines, thus mimicking the frequent structural aberrations involving this genomic region identified in human T-cell and B-cell lymphomas, as well as stomach adenocarcinomas. The authors found that PD-L1 levels were strongly increased in cells expressing Cas9 and two 3'-UTR targeting sgRNAs, while in the same conditions separate transfection of each sgRNA had no effect, confirming the specificity of their strategy. They also demonstrated that PD-L1 up-regulation induced by disruption of its 3'-UTR could promote tumor growth and reduce activation of CD8⁺ cytotoxic T lymphocytes, thus establishing a new genetic mechanism of tumor immune escape [19].

When the two sgRNAs are designed to target distant genomic locations, DNA repair can generate more profound structural abnormalities, including chromosomal inversions (if the two sgRNAs target the same chromosome) or translocations (if the two sgRNAs target distinct chromosomes). This strategy has been implemented to model frequent rearrangements in human tumors, which result in the expression of fusion proteins displaying oncogenic properties. For example, Choi and Meyerson used sgRNAs targeting two regions separated by 13 Mb on the short arm of chromosome 2 to recreate the inversion that originates the *EML4-ALK* fusion gene, which is found in a significant fraction of human lung

cancers. In the same study, the authors mimicked two other lung cancer fusion oncogenes: *KIF5B-RET*, resulting from a pericentric inversion on chromosome 10, and *CD74-ROS1*, originated from a translocation between chromosome 5 and 6 [20]. As another example of this type of approach, Torres et al. induced a translocation between chromosome 8 and 21 in human hematopoietic stem cells to model the *RUNX1-ETO* fusion, an oncogene that drives a particular subtype of acute myeloid leukemia [21].

3. Modeling tumor evolution through CRISPR/Cas9

Transformation of a normal cell into a metastatic tumor requires the multistep accumulation of a series of mutations, resulting in the acquisition of distinct biological hallmarks that define the malignant phenotype. Taking advantage of specific functional properties conferred by certain oncogenic mutations, two recent studies [22,23] used CRISPR/Cas9 to experimentally recapitulate the simplified model of intestinal tumorigenesis proposed by Fearon and Vogelstein more than 25 years ago, based on the most frequent aberrations identified in human tumors [24] (Fig. 3A). The same authors previously developed a three-dimensional culture system based on a defined cocktail of factors, in which a single intestinal stem cell can give rise to a “mini-gut”, an organoid structure displaying the typical morphological and functional characteristics of the intestinal epithelium [25]. They used CRISPR/Cas9 technology to truncate the *APC* tumor suppressor gene in intestinal organoids, thus mimicking a well-characterized early event in colorectal cancer development, which drives the formation of small adenomas from the normal intestinal epithelium [9]. By withdrawing activators of Wnt signaling from the culture medium, Matano et al. and Drost et al. positively selected human intestinal stem cells in which *APC* loss provoked β -catenin stabilization, and consequently Wnt pathway up-regulation. They then used a similar approach of selective pressure to derive cells containing an activating mutation of the *KRAS* oncogene and a loss-of-function mutation of *SMAD4*, by withdrawing from the culture media the epidermal growth factor (EGF) or an inhibitor of the transforming growth factor β signaling, respectively. Finally, they added Nutlin-3 to select for cells in which the *TP53* gene was inactivated by CRISPR/Cas9 (Fig. 3B). At each step, acquisition of the new desired mutation was confirmed by sequencing analysis. Through this strategy, they could obtain cultures with different combinations of these mutations, thus recapitulating different stages of tumorigenesis. Consistent with a multistep model of tumor evolution, quadruple mutations were necessary to transform normal organoids into invasive carcinomas upon inoculation in immunodeficient mice [22,23], while *APC* and *TP53* loss was sufficient to induce chromosomal instability, a hallmark of colorectal tumor progression [23]. Though oversimplified, this model provided for the first time the means to directly assess whether the accumulation of different mutations occurs randomly, or following a precise chronological order.

As another example of application of CRISPR/Cas9 technology to model the combination of multiple genetic aberrations during tumorigenesis, Heckl et al. transduced mouse hematopoietic stem cells with a pool of lentiviral sgRNA vectors targeting six different putative tumor suppressor genes, followed by transplantation into lethally irradiated mice. As a control for potential non-specific effects, the authors sequenced the top five predicted off-target sites for each gene, none of which showed mutations. In this case, selection was not based on independence from a particular signaling pathway, but instead on the capacity of the mutant cells to out-compete wild-type cells in repopulating the hematopoietic system of recipient mice [26]. Based on a similar approach of *in vivo* selection, Chen et al. transduced a non-metastatic mouse lung cancer

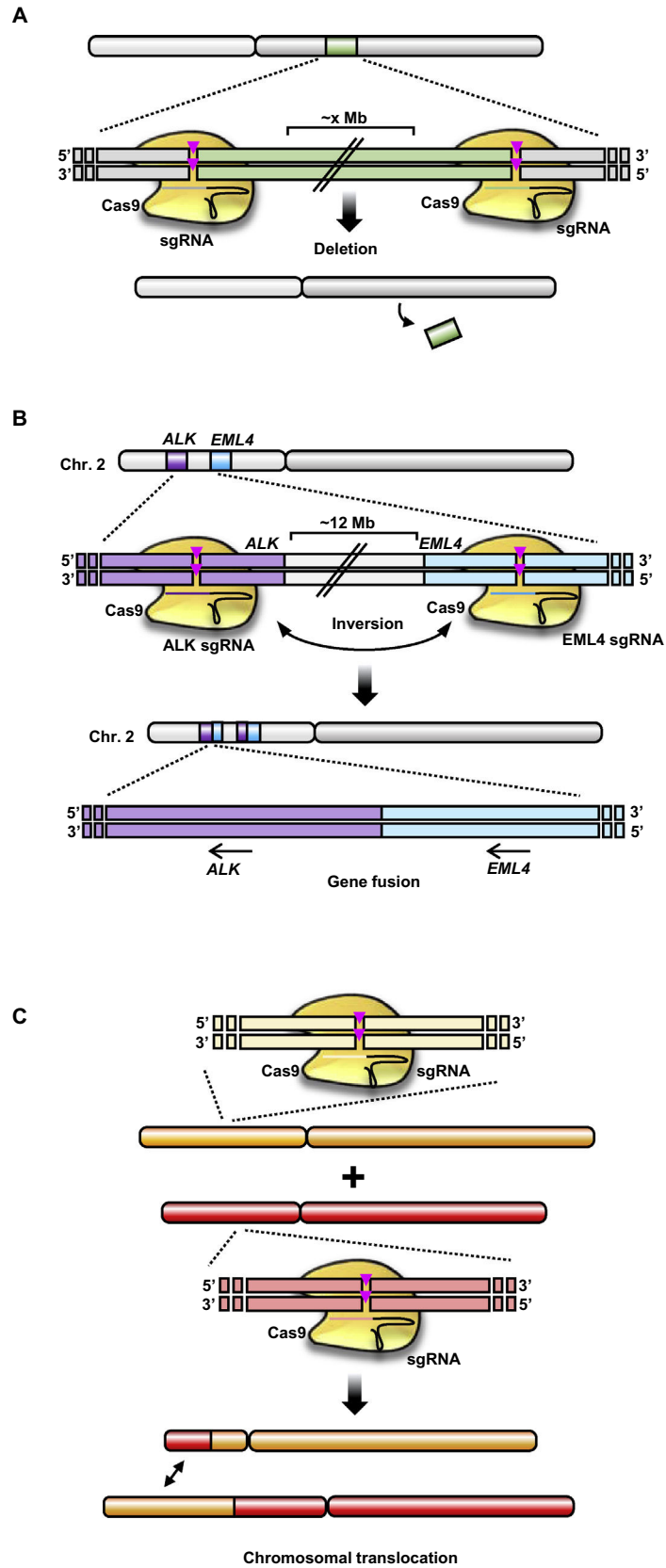


Fig. 2. CRISPR/Cas9 strategies based on the concurrent utilization of two distinct sgRNAs. (A) Coexpression of Cas9 and two sgRNAs can be used to precisely delete a specific genomic region. (B) Representation of chromosomal inversion involving the *ALK* and *EML4* loci using Cas9 protein and two distinct sgRNAs targeting either *ALK* or *EML4* gene. The chromosomal inversion results in the expression of the EML4-ALK fusion protein. (C) sgRNA targeting of two different chromosomes can generate a chromosomal translocation.

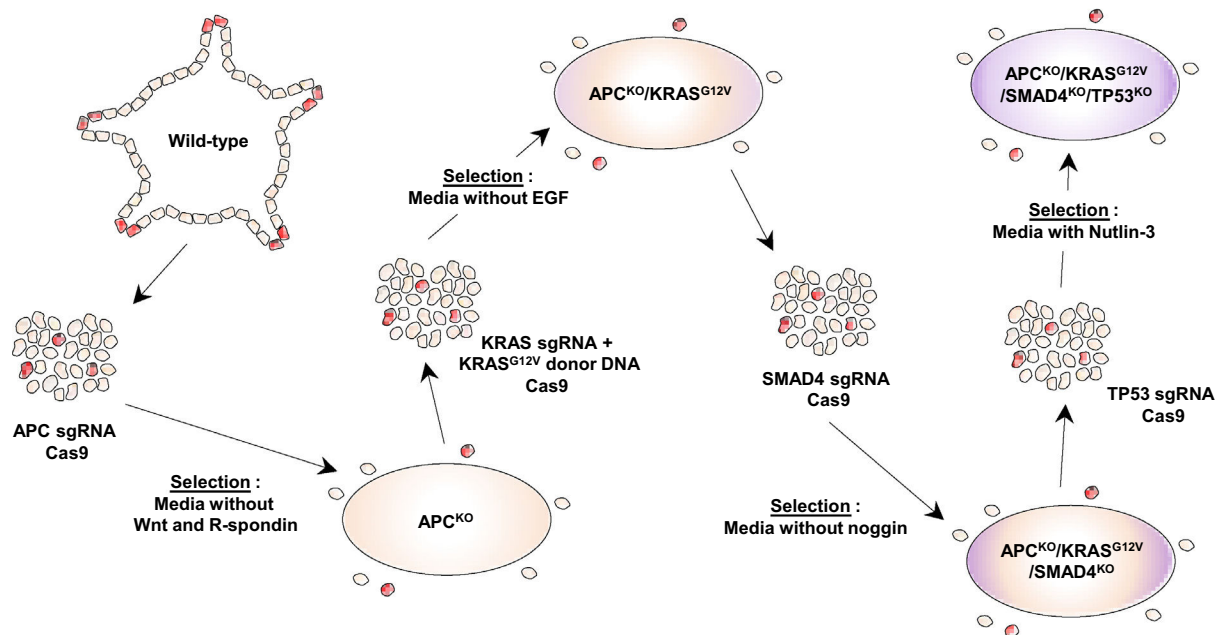
A Fearon-Vogelstein model of intestinal tumorigenesis:**B** CRISPR/Cas9-mediated sequential accumulation of mutations in intestinal organoids:

Fig. 3. Modeling intestinal tumorigenesis through CRISPR/Cas9. (A) Tumor evolution in human intestine as initially proposed by E. Fearon and B. Vogelstein. The sequential accumulation of *APC* loss, *KRAS* mutation, inactivation of *SMAD4* and *TP53* progressively transform normal epithelium into an invasive carcinoma. (B) Schematic diagram illustrating the strategy used by Matano et al. and Drost et al. to model intestinal tumorigenesis in human organoids through CRISPR/Cas9 technology (adapted from [23]).

cell line with a lentiviral CRISPR/Cas9 library targeting thousands of protein-coding genes and hundreds of microRNA precursors. Upon subcutaneous inoculation in immunodeficient mice, the cells containing the library formed fast growing tumors and lung metastases, and deep sequencing analysis of the enriched sgRNAs enabled the identification of several genes whose inactivation could contribute to tumor growth and invasion [27]. A major advantage of the strategy used in these two latter papers is that selection does not imply previous knowledge of the different functional hallmarks acquired throughout tumorigenesis. However, simultaneous targeting of several tumor suppressors in the same cell, as it was the case in the hematopoietic stem cell model [26], may not fully recapitulate the process of tumor evolution, where mutations are acquired and selected over time, in some cases in a hierarchical manner.

4. New mouse cancer models

Genetically-modified mice have been widely used to model different types of human malignancies through expression of active oncogenes or inactivation of tumor suppressors, either alone or in combination. To overcome potential perturbations of normal

mouse development, systems have been developed to enable conditional and/or inducible control of these oncogenic events. While extremely powerful, these strategies generally require precise genome editing in the germline, including insertion of specific recombination sites, such as *LoxP* sites, and derivation of mouse colonies containing the desired genotype. Because of its high efficiency and flexibility, CRISPR/Cas9 technology now offers the opportunity to directly modify the genome of somatic cells in the adult. This strategy was used for the first time by Tyler Jacks' laboratory to target *Pten* and *Trp53* tumor suppressor genes in the liver of wild-type mice through hydrodynamic tail vein injection of a plasmid encoding Cas9 and sgRNAs. CRISPR/Cas9 mediated inactivation of *Pten* phenocopied its deletion through Cre-*LoxP* technology, and combined *Pten* and *Trp53* mutations resulted in the formation of liver tumors. As a control for specificity, in the same conditions a sgRNA targeting GFP had no effect. By co-injecting the CRISPR/Cas9 vector with a donor DNA for HDR, the authors were also able to introduce a specific point mutation in the gene encoding β -catenin that is frequently observed in human hepatocellular carcinomas, thus demonstrating that this strategy can also be applied for more precise genome editing [28]. As another example of somatic transfection of CRISPR/Cas9 plasmids for cancer modeling, Zuckermann et al. used polyethylenimine and in utero electroporation to target

different tumor suppressors, including *Ptch1*, *Trp53*, *Pten* and *Nf1*, in the developing mouse brain. Depending on the set of mutations, the mice developed medulloblastomas and/or glioblastomas [29].

While transfection can only be used for a limited number of organs and tissues, somatic delivery through viral vectors provides a wider range of applications. Sanchez-Rivera et al. combined Cre-LoxP and CRISPR/Cas9 technologies to simultaneously induce the expression of mutant *Kras* and inactivate different tumor suppressors in mouse lung, by delivering intratracheally a lentivirus containing Cre recombinase, Cas9 and specific sgRNAs. Depending on the combination of targeted genes, the mice developed lung tumors with distinct histopathological and molecular features, indicating that this strategy can be used to investigate the cooperation between different genetic aberrations during tumorigenesis. Of note, sequence analysis of predicted off-target sites revealed extremely low mutation rates, confirming that the effects observed on tumor formation derived from deletion of the intended target, and not from editing of other genes [30]. Intratracheal infection using either adenoviral or lentiviral vectors was also used to somatically deliver Cas9 and two distinct sgRNAs targeting the *Eml4* and *Alk* loci on mouse chromosome 17 [31,32]. As discussed above, DNA repair of the double cleavage can generate a chromosomal rearrangement, resulting in the expression of the oncogenic *Eml4-Alk* fusion protein. In both studies, the mice developed lung tumors, which responded to treatment with the ALK inhibitor crizotinib [31]. Mouse pancreas is another example of organ in which Cas9 and sgRNA targeting a tumor suppressor were successfully delivered through parenchymal injection of lentiviral particles [33].

Despite its proven efficacy in certain models, the major limitation of Cas9 somatic delivery relies on the relative large size of this protein compared to the packaging capacity of lentiviral and adeno-associated viral (AAV) vectors. As an alternative approach, the laboratories of Feng Zhang and Philip Sharp generated a knockin mouse in which Cas9 is inserted in the *Rosa26* locus, driven by an ubiquitous promoter that can be activated upon removal of a LoxP-STOP-LoxP cassette by Cre recombinase [34]. Through intranasal or intratracheal inoculation, the authors delivered a single AAV vector containing Cre recombinase, Renilla luciferase to monitor infection, three sgRNAs (targeting *Trp53*, *Lkb1* and *Kras*) and a *Kras*-G12D HDR template. One month post-injection, the mice developed several grade I bronchial alveolar adenomas, which progressed, a month later, to grade III lung adenocarcinomas and, occasionally, to invasive grade IV adenocarcinomas [34]. Using a similar mouse model for Cre-inducible expression of Cas9, Chiou et al. specifically inactivated *Lkb1* tumor suppressor in adult pancreas through retrograde injection in the bile duct of a lentivirus containing Cre recombinase and a *Lkb1* specific sgRNA, which resulted, two months later, in multiple pancreatic tumors [35]. While one cannot completely rule out the possibility of off-target effects potentially generated by the chosen sgRNAs, this type of model provides a useful screening platform for early-stage identification of candidate genes, which can then be validated using different sgRNAs. As a different means for tissue specific CRISPR/Cas9 genome editing, Annunziato et al. took advantage of a mouse model in which the Cre recombinase is driven by the *Wap* promoter, thus enabling specific expression of Cas9 in the mammary epithelium. Four months after mammary duct injection with a lentivirus containing a sgRNA targeting *Pten*, these mice, which also harbored conditional deletion of the *Cdh1* gene encoding E-cadherin, developed tumors displaying the typical histology of human lobular breast carcinomas. Of note, in a similar model in which Cas9 was directly delivered through a lentiviral vector, the authors observed highly infiltrated lesions containing both T- and B-cells, suggesting a specific immune response triggered by Cas9. As a possible explanation for these profoundly different

phenotypes, the authors argued that, in the *Wap*-CRE model, early activation of Cas9 during postnatal development due to ectopic expression of Cre could result in immune tolerance [36].

While further studies are needed to fully assess whether Cas9 can be immunogenic in mouse models, these observations raise concern over potential non specific effects of prolonged *in vivo* expression of this endonuclease. To overcome this issue, Dow et al. inserted in the mouse genome a cassette containing a doxycycline-inducible Cas9, together with sgRNAs targeting *Apc* or *Trp53* [37]. To minimize the risk of CRISPR/Cas9 off-target editing, the authors also generated a second inducible model based on the “nickase” strategy, in which a Cas9 variant that cleaves only one strand of DNA is used in combination with two closely spaced antiparallel sgRNAs, resulting in highly specific double strand breaks [38]. In both cases, doxycycline treatment rapidly induced hyperplastic overgrowth in the intestine, as a result of β -catenin stabilization provoked by CRISPR/Cas9-mediated inactivation of *Apc* [37]. Despite being more time-consuming compared to the other strategies discussed above, this system can be combined with tissue specific promoters for expression of the doxycycline transactivator, thus providing the means to tightly control both the site and the duration of Cas9 expression.

5. CRISPR-barcoding

As a direct consequence of the accumulation of multiple driver and passenger mutations during the oncogenic process, most tumors are composed of an intricate pattern of genetically heterogeneous subclonal populations [39–41]. Such diversity has a profound impact in cancer progression and response to therapy, since it confers extraordinary flexibility and capacity to adapt to very different environmental conditions [42–44]. Despite major clinical implications, few strategies enable modeling intratumor heterogeneity in a controlled experimental setting. Taking advantage of the relatively low efficiency of HDR genome editing, we used CRISPR/Cas9 to introduce an oncogenic mutation of interest in a small subset of cells within a mass population. As a means to track DNA-edited cells, we coupled the potentially functional modification with a series of silent mutations, serving as a genetic barcode that can be “read” by real-time qPCR or deep sequencing analysis of genomic DNA. In parallel, we labeled a distinct subpopulation within the same culture through a different set of silent mutations, functioning as an internal control for normalization. While sgRNAs are generally chosen using algorithms designed to identify potential off-target sequences to avoid non-specific effects, occasional minor shifts in the secondary structure of the RNA-DNA duplex [45,46] can make such predictions more problematic. Thus, by expressing the same sgRNA in both control and mutant cells, our approach provides a higher level of specificity compared to other CRISPR/Cas9 strategies [47]. To avoid insertion of distinct barcodes in the different alleles of the same cell, the ssODNs containing the two sets of mutations are transfected separately in combination with the plasmid encoding Cas9 and the sgRNA. Immediately after transfection, the cells are pooled to form a heterogeneous mass population composed of unmodified and barcoded cells (Fig. 4A). We applied this approach in different tumor cell lines to introduce nonsense or missense mutations in the *TP53* gene, resulting in an inactive or a dominant-negative form of this tumor suppressor, respectively. We showed that TP53 induction through Nutlin 3 treatment resulted in the amplification of the cells in which TP53 function was lost, and we demonstrated how this type of system can be conveniently used to test different combinations of cancer drugs. Using a similar strategy, we recapitulated different genetic mechanisms leading to resistance to inhibitors of the EGF receptor in a non-small cell lung cancer human

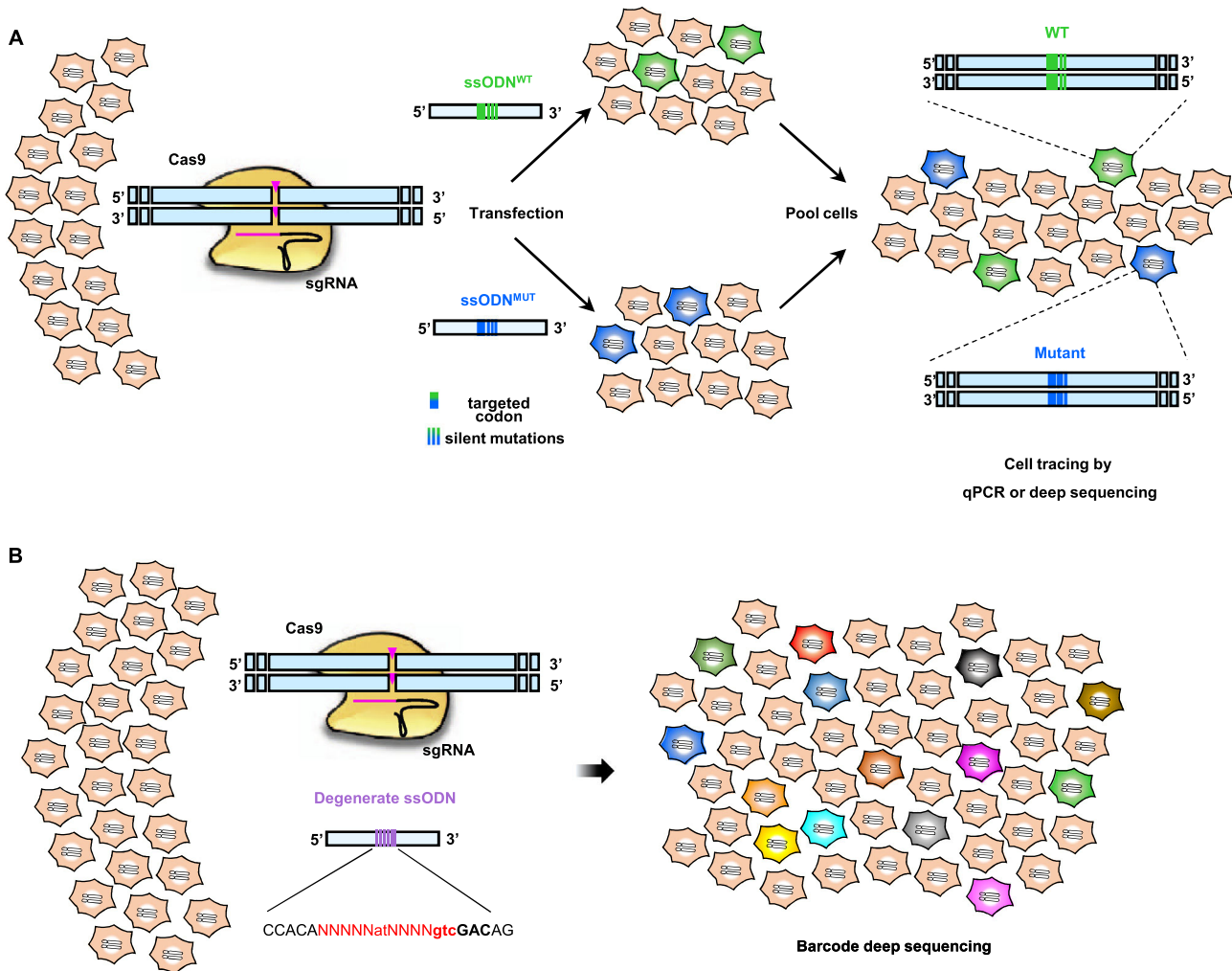


Fig. 4. CRISPR-barcoding to functionally characterize oncogenic mutations in a context of intratumor heterogeneity. (A) Diagram illustrating the CRISPR-barcoding strategy. Cells are co-transfected with a Cas9/sgRNA plasmid and a ssODN containing a potentially functional mutation of interest, coupled with a series of silent point mutations (ssODN^{MUT}). In parallel, a second batch of the same cells is transfected with the CRISPR/Cas9 vector and a ssODN containing a different set of silent mutations (ssODN^{WT}). Immediately after transfection, the two pools are mixed to form a heterogeneous mass population composed of unmodified and barcoded cells. The relative proportion of each barcode is assessed by qPCR or deep sequencing. (B) Through insertion of a highly complex series of degenerate sequences at a specific genomic location, CRISPR-barcoding can be used to trace several thousands of genetically labelled clones within a mass population of tumor cells.

cell line. Through this model, we also showed that CRISPR-barcoding can be implemented not only to investigate effects on proliferation/survival, but also on cell invasion, as well as tumorigenicity in immunodeficient mice [47].

As discussed above, CRISPR/Cas9 technology can be applied in different contexts to recreate oncogenic mutations, based on the selective pressure favoring the mutated cells. However, due to the relatively low efficiency of genome editing, current approaches do not allow generation of modifications that have a negative impact on cell growth, thus making it extremely difficult, if not impossible, to repair oncogenic driver mutations. By enabling early tracing of the modified cells, without the need for clonal expansion, CRISPR-barcoding overcomes this limitation, as illustrated by repairing addictive mutations of the *APC* and *ALK* genes in cell lines derived from colon cancer and neuroblastoma, respectively. Finally, through incorporation of a highly complex series of degenerate sequences at a given genome location, CRISPR-barcoding enables tracing of thousands of individual clones within an heterogeneous tumor cell population [47], thus representing a convenient alternative to the use of lentiviral DNA barcode libraries [48–51] (Fig. 4B).

6. Concluding remarks

It has been less than five years since the beginning of the CRISPR/Cas9 revolution, but the impact of this new technology on cancer research has been tremendous. Compared to previous methods for genome editing based on protein-DNA interactions, CRISPR/Cas9 is extremely more flexible and straightforward. Indeed, for a cancer scientist it is now relatively easy to experimentally recreate a mutation identified in a patient's tumor and assess its effects in a cell line or in the context of an entire organism. As discussed in this review, CRISPR/Cas9 has already been used to recapitulate a wide array of oncogenic events, from inactivation of a tumor suppressor or point mutation of a proto-oncogene, to deletion of a non-coding sequence or generation of a chromosomal rearrangement. Such modifications have been established both *in vitro* and *in vivo*, thus emphasizing the extraordinary potential of CRISPR/Cas9 in generating new cancer models. After an initial period where the efforts were mostly focused on exploring the broad possibilities of this technology, these models can now be used to elucidate new aspects of cancer biology. Indeed, it should become relatively easy to directly assess whether a given mutation

found in human tumors can actually participate in the tumorigenic process and, hence, be considered as a *bona fide* therapeutic target. Most importantly, CRISPR/Cas9 models of such mutations could subsequently be used to implement high-throughput screens to identify new anticancer drugs.

Acknowledgments

This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM), the Université de Rouen Normandie and the Ligue Contre le Cancer de Normandie. A.G is recipient of a doctoral fellowship from the Normandy Region. L. G. was supported by a Chair of Excellence program from INSERM and Université de Rouen Normandie.

References

- [1] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, E. Charpentier, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity, *Science* 337 (6096) (2012) 816–821.
- [2] G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria, *Proc. Natl. Acad. Sci. U.S.A.* 109 (39) (2012) E2579–E2586.
- [3] L. Cong, F.A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, et al., Multiplex genome engineering using CRISPR/Cas systems, *Science* 339 (6121) (2013) 819–823.
- [4] P. Mali, L. Yang, K.M. Esvelt, J. Aach, M. Guell, J.E. DiCarlo, et al., RNA-guided human genome engineering via Cas9, *Science* 339 (6121) (2013) 823–826.
- [5] P.D. Hsu, E.S. Lander, F. Zhang, Development and applications of CRISPR-Cas9 for genome engineering, *Cell* 157 (6) (2014) 1262–1278.
- [6] J.A. Doudna, E. Charpentier, Genome editing. The new frontier of genome engineering using CRISPR-Cas9, *Science* 346 (6213) (2014) 1258096.
- [7] J.D. Sander, J.K. Joung, CRISPR-Cas systems for editing, regulating and targeting genomes, *Nat. Biotechnol.* 32 (4) (2014) 347–355.
- [8] C. Kandoth, M.D. McLellan, F. Vandini, K. Ye, B. Niu, C. Lu, et al., Mutational landscape and significance across 12 major cancer types, *Nature* 502 (7471) (2013) 333–339.
- [9] B. Vogelstein, N. Papadopoulos, V.E. Velculescu, S. Zhou, L.A. Diaz Jr., K.W. Kinzler, Cancer genome landscapes, *Science* 339 (6127) (2013) 1546–1558.
- [10] A. Malina, J.R. Mills, R. Cencic, Y. Yan, J. Fraser, L.M. Schippers, et al., Repurposing CRISPR/Cas9 for in situ functional assays, *Genes Dev.* 27 (23) (2013) 2602–2614.
- [11] J. Albers, C. Danzer, M. Rechsteiner, H. Lehmann, L.P. Brandt, T. Hejhal, et al., A versatile modular vector system for rapid combinatorial mammalian genetics, *J. Clin. Invest.* 125 (4) (2015) 1603–1619.
- [12] C. Chen, Y. Liu, A.R. Rappaport, T. Kitzing, N. Schultz, Z. Zhao, et al., MLL3 is a haploinsufficient 7q tumor suppressor in acute myeloid leukemia, *Cancer Cell* 25 (5) (2014) 652–665.
- [13] A. Harrod, J. Fulton, V.T. Nguyen, M. Periyasamy, L. Ramos-Garcia, C.F. Lai, et al., Genomic modelling of the ESR1 Y537S mutation for evaluating function and new therapeutic approaches for metastatic breast cancer, *Oncogene* (2016).
- [14] J.E. Neggers, T. Vercruyse, M. Jacquemyn, E. Vanstreels, E. Baloglu, S. Shacham, et al., Identifying drug-target selectivity of small-molecule CRM1/XPO1 inhibitors by CRISPR/Cas9 genome editing, *Chem. Biol.* 22 (1) (2015) 107–116.
- [15] O. Shalem, N.E. Sanjana, E. Hartenian, X. Shi, D.A. Scott, T.S. Mikkelson, et al., Genome-scale CRISPR-Cas9 knockout screening in human cells, *Science* 343 (6166) (2014) 84–87.
- [16] S. Konermann, M.D. Brigham, A.E. Trevino, J. Joung, O.O. Abudayyeh, C. Barceña, et al., Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex, *Nature* 517 (7536) (2015) 583–588.
- [17] N.E. Sanjana, J. Wright, K. Zheng, O. Shalem, P. Fontanillas, J. Joung, et al., High-resolution interrogation of functional elements in the noncoding genome, *Science* 353 (6307) (2016) 1545–1549.
- [18] Z. Siprashvili, D.E. Webster, D. Johnston, R.M. Shenoy, A.J. Ungewickell, A. Bhaduri, et al., The noncoding RNAs SNORD50A and SNORD50B bind K-Ras and are recurrently deleted in human cancer, *Nat. Genet.* 48 (1) (2016) 53–58.
- [19] K. Kataoka, Y. Shiraiishi, Y. Takeda, S. Sakata, M. Matsumoto, S. Nagano, et al., Aberrant PD-L1 expression through 3'-UTR disruption in multiple cancers, *Nature* 534 (7607) (2016) 402–406.
- [20] P.S. Choi, M. Meyerson, Targeted genomic rearrangements using CRISPR/Cas technology, *Nat. Commun.* 5 (2014) 3728.
- [21] R. Torres, M.C. Martin, A. Garcia, J.C. Cigudosa, J.C. Ramirez, S. Rodriguez-Perales, Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR-Cas9 system, *Nat. Commun.* 5 (2014) 3964.
- [22] M. Matano, S. Date, M. Shimokawa, A. Takano, M. Fujii, Y. Ohta, et al., Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids, *Nat. Med.* 21 (3) (2015) 256–262.
- [23] J. Drost, R.H. van Jaarsveld, B. Ponsioen, C. Zimmerman, R. van Bostel, A. Buijs, et al., Sequential cancer mutations in cultured human intestinal stem cells, *Nature* 521 (7550) (2015) 43–47.
- [24] E.R. Fearon, B. Vogelstein, A genetic model for colorectal tumorigenesis, *Cell* 61 (5) (1990) 759–767.
- [25] T. Sato, R.G. Vries, H.J. Snippert, M. van de Wetering, N. Barker, D.E. Stange, et al., Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche, *Nature* 459 (7244) (2009) 262–265.
- [26] D. Heckl, M.S. Kowalczyk, D. Yudovich, R. Belizaire, R.V. Puram, M.E. McConkey, et al., Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing, *Nat. Biotechnol.* 32 (9) (2014) 941–946.
- [27] S. Chen, N.E. Sanjana, K. Zheng, O. Shalem, K. Lee, X. Shi, et al., Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis, *Cell* 160 (6) (2015) 1246–1260.
- [28] W. Xue, S. Chen, H. Yin, T. Tammela, T. Papagiannakopoulos, N.S. Joshi, et al., CRISPR-mediated direct mutation of cancer genes in the mouse liver, *Nature* 514 (7522) (2014) 380–384.
- [29] M. Zuckermann, V. Hovestadt, C.B. Knobbe-Thomsen, M. Zapatka, P.A. Northcott, K. Schramm, et al., Somatic CRISPR/Cas9-mediated tumour suppressor disruption enables versatile brain tumour modelling, *Nat. Commun.* 6 (2015) 7391.
- [30] F.J. Sanchez-Rivera, T. Papagiannakopoulos, R. Romero, T. Tammela, M.R. Bauer, A. Bhutkar, et al., Rapid modelling of cooperating genetic events in cancer through somatic genome editing, *Nature* 516 (7531) (2014) 428–431.
- [31] D. Maddalo, E. Machado, C.P. Concepcion, C. Bonetti, J.A. Vidigal, Y.C. Han, et al., In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system, *Nature* 516 (7531) (2014) 423–427.
- [32] R.B. Blasco, E. Karaca, C. Ambrogio, T.C. Cheong, E. Karayol, V.G. Minerò, et al., Simple and rapid in vivo generation of chromosomal rearrangements using CRISPR/Cas9 technology, *Cell Rep.* 9 (4) (2014) 1219–1227.
- [33] P.K. Mazur, A. Herner, S.S. Mello, M. Wirth, S. Hausmann, F.J. Sanchez-Rivera, et al., Combined inhibition of BET family proteins and histone deacetylases as a potential epigenetics-based therapy for pancreatic ductal adenocarcinoma, *Nat. Med.* 21 (10) (2015) 1163–1171.
- [34] R.J. Platt, S. Chen, Y. Zhou, M.J. Yin, L. Swiech, H.R. Kempton, et al., CRISPR-Cas9 knockin mice for genome editing and cancer modeling, *Cell* 159 (2) (2014) 440–455.
- [35] S.H. Chiou, I.P. Winters, J. Wang, S. Naranjo, C. Dudgeon, F.B. Tamburini, et al., Pancreatic cancer modeling using retrograde viral vector delivery and in vivo CRISPR/Cas9-mediated somatic genome editing, *Genes Dev.* 29 (14) (2015) 1576–1585.
- [36] S. Annunziato, S.M. Kas, M. Nethe, H. Yucel, J. Del Bravo, C. Pritchard, et al., Modeling invasive lobular breast carcinoma by CRISPR/Cas9-mediated somatic genome editing of the mammary gland, *Genes Dev.* 30 (12) (2016) 1470–1480.
- [37] L.E. Dow, J. Fisher, K.P. O'Rourke, A. Muley, E.R. Kastenhuber, G. Livshits, et al., Inducible in vivo genome editing with CRISPR-Cas9, *Nat. Biotechnol.* 33 (4) (2015) 390–394.
- [38] F.A. Ran, P.D. Hsu, C.Y. Lin, J.S. Gootenberg, S. Konermann, A.E. Trevino, et al., Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity, *Cell* 154 (6) (2013) 1380–1389.
- [39] R.J. Gillies, D. Verduzco, R.A. Gatenby, Evolutionary dynamics of carcinogenesis and why targeted therapy does not work, *Nat. Rev. Cancer* 12 (7) (2012) 487–493.
- [40] M. Greaves, C.C. Maley, Clonal evolution in cancer, *Nature* 481 (7381) (2012) 306–313.
- [41] N. McGranahan, C. Swanton, Biological and therapeutic impact of intratumor heterogeneity in cancer evolution, *Cancer Cell* 27 (1) (2015) 15–26.
- [42] A. Marusyk, K. Polyak, Tumor heterogeneity: causes and consequences, *Biochim. Biophys. Acta* 1805 (1) (2010) 105–117.
- [43] M. Kleppe, R.L. Levine, Tumor heterogeneity confounds and illuminates: assessing the implications, *Nat. Med.* 20 (4) (2014) 342–344.
- [44] A.A. Alizadeh, V. Aranda, A. Bardelli, C. Blanpain, C. Bock, C. Borowski, et al., Toward understanding and exploiting tumor heterogeneity, *Nat. Med.* 21 (8) (2015) 846–853.
- [45] Y. Lin, T.J. Cradick, M.T. Brown, H. Deshmukh, P. Ranjan, N. Sarode, et al., CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences, *Nucleic Acids Res.* 42 (11) (2014) 7473–7485.
- [46] X. Wang, Y. Wang, X. Wu, J. Wang, Y. Wang, Z. Qiu, et al., Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors, *Nat. Biotechnol.* 33 (2) (2015) 175–178.
- [47] A. Guernet, S.K. Mungamuri, D. Cartier, R. Sachidanandam, A. Jayaprakash, S. Adriouch, et al., CRISPR-barcoding for intratumor genetic heterogeneity modeling and functional analysis of oncogenic driver mutations, *Mol. Cell* 63 (3) (2016) 526–538.
- [48] L.V. Nguyen, C.L. Cox, P. Eirew, D.J. Knapp, D. Pellacani, N. Kannan, et al., DNA barcoding reveals diverse growth kinetics of human breast tumour subclones in serially passaged xenografts, *Nat. Commun.* 5 (2014) 5871.
- [49] L.V. Nguyen, D. Pellacani, S. Lefort, N. Kannan, T. Osako, M. Makarem, et al., Barcoding reveals complex clonal dynamics of de novo transformed human mammary cells, *Nature* (2015).
- [50] H.E. Bhang, D.A. Ruddy, V. Krishnamurthy Radhakrishnan, J.X. Caushi, R. Zhao, M.M. Hims, et al., Studying clonal dynamics in response to cancer therapy using high-complexity barcoding, *Nat. Med.* 21 (5) (2015) 440–448.
- [51] A.N. Hata, M.J. Niederst, H.L. Archibald, M. Gomez-Caraballo, F.M. Siddiqui, H.E. Mulvey, et al., Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition, *Nat. Med.* 22 (3) (2016) 262–269.

2. Résultats non publiés

Nouvelle stratégie thérapeutique pour prévenir la résistance aux inhibiteurs d'EGFR dans le cancer bronchique

Dans cette étude, nous avons utilisé un modèle cellulaire de résistance du CBNPC à la thérapie ciblée pour identifier un composé capable de prévenir l'émergence de clones insensibles aux inhibiteurs d'EGFR.

Nos travaux ont permis :

- L'identification du sorafenib pour sa capacité à prévenir l'amplification des cellules de CBNPC contenant la mutation T790M ou C797S en présence de différents types d'inhibiteurs d'EGFR
- L'étude du mécanisme d'action responsable des effets du sorafenib dans les cellules de CBNPC
- La démonstration que le sorafenib peut prévenir la résistance aux inhibiteurs d'EGFR dans un modèle *in vivo* de CBNPC basé sur l'approche CRISPR-barcoding.

2.1. Introduction

Les cancers bronchiques non à petites cellules (CBNPC) représentent près de 85% des tumeurs du poumon. Des mutations activatrices au sein du domaine kinase de l'*epidermal growth factor receptor* (EGFR) ont été observées dans les CBNPCs, chez environ 10-15% des patients caucasiens et approximativement 30-40% de ceux d'origine est-asiatique. Ces mutations, dont les plus fréquentes sont la substitution de la leucine en arginine en position 858 (L858R) ou la délétion de quelques nucléotides dans l'exon 19, confèrent aux tumeurs une sensibilité aux inhibiteurs d'EGFR de 1^{ère} et de 2^{ème} génération, tels que le gefitinib (Iressa ; AstraZeneca), l'erlotinib (Tarceva ; Roche) ou l'afatinib (Giotrif ; Boehringer Ingelheim) (Lynch *et al.*, 2004 ; Paez *et al.*, 2004 ; Pao *et al.*, 2004). Malheureusement, la quasi-totalité des patients répondant favorablement à ces traitements voit leur tumeur progresser après 9 à 14 mois de thérapie. Cette rechute est due au développement d'une résistance, causée dans plus de la moitié des cas, par l'apparition d'une seconde mutation (T790M) au niveau du domaine catalytique du récepteur (Kobayashi *et al.*, 2005 ; Pao *et al.*, 2004).

Afin de surmonter ce mécanisme de résistance, une nouvelle classe d'inhibiteurs covalents spécifiques de la mutation T790M a été générée (Zhou *et al.*, 2009). L'une de ces molécules, l'osimertinib (Tagrisso ; AstraZeneca), a été récemment approuvée pour le traitement des CBNPCs ayant développé une résistance aux inhibiteurs d'EGFR de 1^{ère} génération (Cross *et al.*, 2014 ; Jänne *et al.*, 2015). Cependant, après quelques mois de traitement, les tumeurs traitées avec l'osimertinib acquièrent de nouvelles altérations génétiques qui les rendent insensibles à toute thérapie ciblée actuellement disponible. Parmi ces mutations, la plus fréquente est la substitution C797S dans le domaine catalytique de l'EGFR, qui empêche la fixation irréversible de l'osimertinib au récepteur (Thress *et al.*, 2015).

En utilisant notre approche de CRISPR-barcoding, nous avons développé de nouveaux modèles cellulaires pour étudier la résistance des CBNPCs à la thérapie ciblée (Guernet *et al.*, 2016). Grâce à un criblage de petites molécules, nous avons identifié le sorafenib (Nexavar ; Bayer) pour sa capacité à empêcher l'amplification des cellules présentant les mutations T790M et C797S, induites respectivement par le gefitinib et l'osimertinib. Cet inhibiteur multikinase, déjà approuvé pour le traitement des cancers du foie, du rein et de la thyroïde, pourrait ainsi représenter une nouvelle stratégie thérapeutique pour prévenir la résistance des CBNPCs aux inhibiteurs d'EGFR.

2.2. Matériels et méthodes

Cultures cellulaires et inhibiteurs

Les cellules PC9 (CBNPC) proviennent de l'*European Collection of Authenticated Cell Cultures* (distribuées par Sigma-Aldrich), les cellules HCC-4006 (CBNPC) ont été achetées chez ATCC et les cellules HCC-827 (CBNPC) ont été obtenues auprès du laboratoire du Dr S. Aaronson. L'ensemble de ces cellules ont été cultivées en milieu *Roswell Park Memorial Institute* (Life Technologies). Les milieux de culture sont supplémentés avec 10% de sérum de veau fœtal (Life Technologies) et 0,6% de pénicilline et streptomycine (Life Technologies). Les inhibiteurs gefitinib et sorafenib ont été obtenus de Santa Cruz Biotechnology et l'osimertinib de MedChemExpress.

Transfection et vecteurs lentiviraux

Les cellules PC9 ont été transfectées avec le dispositif Nucleofector II (Lonza) à l'aide du kit L Amaxa Nucleofector et du programme d'électroporation recommandé par le fabricant. L'efficacité des transfections a été vérifiée en parallèle avec un vecteur contenant la *Green Fluorescent Protein*.

Pour la production des lentivirus, les cellules HEK 293T sont co-transfectées avec le plasmide d'enveloppe PMD2.VSVG, le plasmide d'emballage PCMV dR 8.74 et le plasmide lentiviral inductible pTRIPZ vide ou contenant un short hairpin RNA (shRNA) ciblant CRAF (Dharmacon). Suite à la transfection, le milieu de culture contenant le virus est collecté une fois par jour pendant 3 jours, supplémenté avec du polybrene (8 µg/ml ; Santa Cruz Biotechnology) et ajouté pendant 12-14 heures à des cellules PC9 fraîchementensemencées pour la transduction. Deux jours après l'infection, les cellules sont sélectionnées avec 2 µg/mL de puromycine (Life Technologies). Des cellules PC9 non infectées sont traitées en parallèle en présence de puromycine et utilisées comme contrôle positif de sélection. L'expression des shRNA ciblant CRAF est induite par ajout de doxycycline (1 µg/mL ; Santa Cruz Biotechnology) au milieu de culture.

Western Blot

Les cellules sont lysées par une solution contenant 50 mM Hepes, 150 mM NaCl, 1% NP40, 5 mM EDTA, 0,2 mM orthovanadate de trisodium, 0,2 M de sodium fluoride et des inhibiteurs de protéases (Roche). Après centrifugation (14000 g, 15 min), les protéines présentes dans le surnageant sont dosées suivant la méthode de Bradford (Bio-Rad). Un tampon de charge

dénaturant composé de sodium dodécyl sulfate (SDS), de β -mercapto-éthanol et de bleu de bromophénol est ajouté aux lysats cellulaires. Les échantillons sont dénaturés à 95°C pendant 5 minutes, séparés sur gel de polyacrylamide en présence de SDS, puis transférés sur une membrane de fluorure de polyvinylidène (Millipore) grâce à un champ électrique.

La membrane est saturée avec 5% de lait dans un tampon *phosphate buffered saline* (PBS), rincée avec du PBS, et incubée avec un anticorps primaire dilué au 1/1000^{ème} dans une solution de PBS contenant 3% de sérum-albumine bovine, 0,5% de Tween et 0,02% d'azide de sodium. Les anticorps primaires utilisés dans cette étude sont : anti-phospho-STAT3, anti-STAT3, anti-phospho-ERK, anti-ERK, anti-EGFR, anti-CRAF et anti- α -tubuline (Cell Signaling Technology). Après trois lavages avec une solution de PBS Tween 0,5%, la membrane est incubée en présence d'un anticorps secondaire (Santa Cruz Biotechnology) dilué au 1/2000^{ème}, lavée trois fois avec du PBS Tween 0,5% et révélée par chimioluminescence à l'aide du substrat « *Clarity Western ECL Substrate* » (BioRad) et d'un appareil ChemiDoc XRS+ System (BioRad).

CRISPR barcoding

Le principe de CRISPR-barcoding est illustré dans la **Figure 21**. Les ARN guides (sgRNA) ont été dessinés grâce au *CRISPR Design Tool* (<http://crispr.mit.edu>) et clonés dans le vecteur pSpCas9(BB)-2A-Puro. Le Tableau 1 montre les séquences cibles des sgRNAs et des *single stranded oligodeoxynucleotide* (ssODN ; Integrated DNA Technologies), qui fonctionnent en tant que ADN donneurs lors de la réparation de l'ADN génomique (ADNg) clivé par CRISPR/Cas9. Les ssODN contiennent la mutation d'intérêt, ainsi que les mutations silencieuses constituant le code barre (Guernet *et al.*, 2016). Afin d'éviter l'incorporation de deux ADN donneurs distincts dans les différents allèles d'une même cellule, les cellules PC9 sont co-transfectées séparément avec les sgRNAs et les ADN donneurs muté ou contrôle. Les cellules sont ensuiteensemencées dans un même flasque (**Fig. 21**).

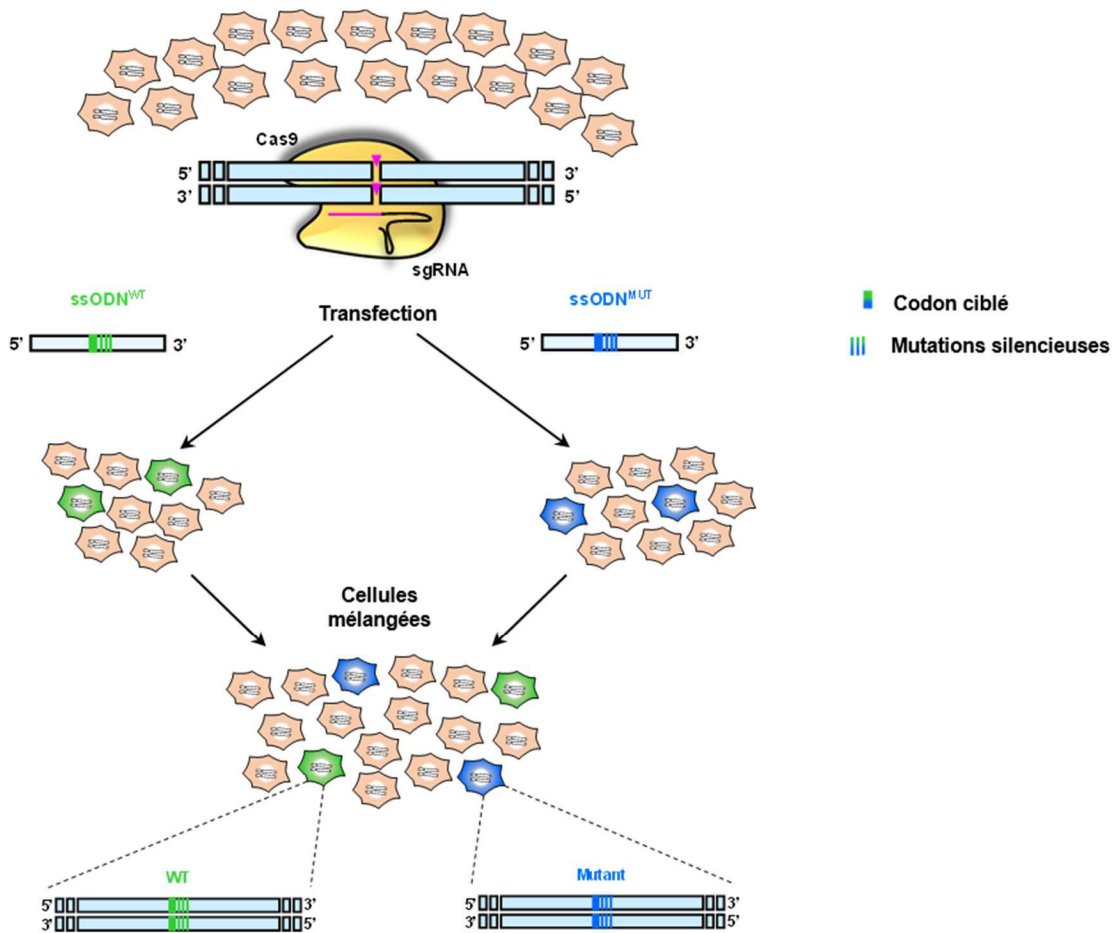


Figure 21. Principe de CRISPR-barcoding

Les cellules sont co-transfectées séparément avec un plasmide contenant un sgRNA et la Cas9, ainsi qu'avec un oligonucléotide simple brin contenant la séquence sauvage (WT) ou la séquence mutée (MUT), couplées à des mutations silencieuses constituant les codes-barres WT ou MUT. Les deux lots de cellules sont ensuite immédiatement mélangés afin de former une population hétérogène, composée de cellules non-modifiées et de cellules contenant l'un ou l'autre code-barres.

Tableau 1. Séquences des sgRNA et ssODN

sgRNA	
Noms	Séquences 5'-3'
sgEGFR	CTGCGTGATGAGCTGCACGG
ssODN	
Noms	Séquences 5'-3' (a)
EGFR-T790T	CTCCCTCCCTCCAGGAAGCCTACGTGATGGCCAGCGTGGACAACCCCCACGTGTG CCGCCTGCTGGGCATCTGCCTCACCTcAcGtCtCAaCTgATtACcCAGCTCATGCC TTCGGCTGCCTCCTGGACTATGTCCGGGAACACAAAGACAATATTGGCTCCCAGTA CCTGCTCAACTGGTGTGTG
EGFR-T790M	CTCCCTCCCTCCAGGAAGCCTACGTGATGGCCAGCGTGGACAACCCCCACGTGTG CCGCCTGCTGGGCATCTGCCTCACCTcAcTgTaCAGCTtATaAtGCAaCTgATGCCCT TCGGCTGCCTCCTGGACTATGTCCGGGAACACAAAGACAATATTGGCTCCCAGTAC CTGCTCAACTGGTGTGTG
EGFR-C797C	GATGGCCAGCGTGGACAACCCCCACGTGTGCCGCCTGCTGGGCATCTGCCT CACCTCCACCGTGCAGCTCATcAtGCAGCTCATGCCgTtTgGtTgTcTAcTcGAC TATGTCCGGGAACACAAAGACAATATTGGCTCCCAGTACCTGCTCAACTGG TGTGTGCAGATCGCAAAGGTAATCAGGGAAG
EGFR-C797S	GATGGCCAGCGTGGACAACCCCCACGTGTGCCGCCTGCTGGGCATCTGCCT CACCTCCACCGTGCAGCTCATcAtGCAGCTCATGCCCTTCGGgGaGtCtGcTtGAt TAcGTCCGGGAACACAAAGACAATATTGGCTCCCAGTACCTGCTCAACTGG TGTGTGCAGATCGCAAAGGTAATCAGGGAAG

(a) Les mutations silencieuses sont inscrites en lettres minuscules

PCR quantitative

L'ADNg est extrait grâce au kit NucleoSpin Tissue (Macherey Nagel). Les différentes amorces de PCR (Eurofins Genomics ; **Tableau 2**) ont été dessinées grâce à l'outil Primer-BLAST (NCBI). Pour éviter la possible amplification d'ADN donneurs non intégrés dans le locus d'intérêt, une des deux amorces cible la séquence génomique endogène bordant la région d'insertion du ssODN. La spécificité de chaque amorce a été testée. Les qPCR sont réalisées à partir de 100 ng d'ADNg et d'un master mix contenant du SYBR Green (Life Technologies) supplémenté des amorces à une concentration finale de 300 nM, en utilisant un appareil Quant Studio Flex PCR System (Life Technologies). L'analyse est effectuée en utilisant la méthode des courbes standards.

Tableau 2. Séquences des amorces de qPCR

qPCR primers	
Noms	Séquences 5'-3'
EGFR_Ctrl_FW	TGCTTCCCCCATTTCAGGACT
EGFR_Ctrl_RV	CTCCTTGACCTCCTCACTG
EGFR_cbc_T790_RV	CCTTCCCTGATTACCTTTGCGA
EGFR_cbc_T790T_FW	CACCTCTACAGTCCAAGTATTACC
EGFR_cbc_T790M_FW	CTCTACTGTACAGCTTATAATGCAACTG
EGFR_cbc_C797_RV	CCTTATCTCCCCTCCCCGTAT
EGFR_cbc_C797C_FW	CATGCCGTTTGGTTGTCTACTC
EGFR_cbc_C797S_FW	CCTTATCTCCCCTCCCCGTAT

Test clonogénique

Des cellules PC9 sontensemencées dans une plaque 6 puits à une densité de 50 000 cellules/puits. Le lendemain, les cellules sont traitées avec du gefitinib (1 μ M) ou de l'osimertinib (1 μ M) en présence ou en l'absence de sorafenib (5 μ M) pendant plusieurs jours, avec un changement de milieu tous les 3 jours. Les cellules sont ensuite rincées deux fois par du PBS et fixées avec une solution 10% méthanol, 10% acide acétique pendant 30 minutes. Les plaques sont séchées à l'air libre pendant la nuit, puis colorées pendant 3 heures en utilisant une solution 1% de crystal violet dans du méthanol. Après lavage à l'eau et séchage, une image des plaques est acquise en utilisant un appareil ChemiDoc XRS.

Xénogreffe chez des souris immunodéficientes

Ces expériences ont été effectuées dans le laboratoire du Pr S. Aaronson (Icahn School of Medicine at Mount Sinai, New York, USA). Les cellules PC9 résistantes au gefitinib (EGFR-T790M), parmi lesquelles une sous-population contenait le code-barres CRISPR EGFR-C797S, ont été mélangé à du matrigel et injectées en sous-cutané dans les flancs droit et gauche de souris SCID mâles (2 X 10⁶ cellules par site). La taille des tumeurs a été évaluée tous les 3-4

jours. Quand les tumeurs ont atteint un volume moyen d'environ 40-50 mm³, correspondant au jour 27 après injection, les souris ont été traitées par gavage, 5 jours par semaine, avec l'osimertinib (5 mg/kg), le sorafenib (60 mg/kg) ou une combinaison des deux. A partir du 39^{ème} jour, les animaux ont été traités 3 fois par semaine. Tous les protocoles utilisant des animaux inclus dans ces expériences ont été approuvés et ont été réalisés en accord avec la réglementation établie par le comité d'éthique et de soin animal de l'Icahn School of Medicine at Mont Sinai (New York, USA).

Analyses statistiques

Les analyses statistiques ont été réalisées avec le logiciel GraphPad Prism en utilisant le test de Mann-Whitney.

2.3. Identification d'un nouveau composé capable de prévenir la résistance au gefitinib des cellules de CBNPC présentant la mutation EGFR-T790M

En utilisant un modèle cellulaire de CBNPC généré pendant la première partie de ma thèse (Guernet *et al.*, 2016), nous avons réalisé un criblage, afin d'identifier des composés susceptibles de prévenir la résistance des cellules PC9 aux inhibiteurs d'EGFR. Parmi les différentes molécules testées, l'inhibiteur multikinase sorafenib (*nexavar*, Bayer Pharmaceuticals) s'est avéré capable d'empêcher l'amplification des cellules contenant la mutation T790M en présence du gefitinib. Nous avons confirmé ces résultats et montré que cet effet est dose-dépendant (**Fig. 22A**). Nous avons utilisé la même population de cellules PC9 contenant le code-barres EGFR-T790M dans un test clonogénique. En parallèle nous avons effectué un test de clonogénicité dans lequel nous avons appliqué différentes concentrations de gefitinib en présence ou en absence de sorafenib. Comme on peut le voir dans la **Figure 22B**,

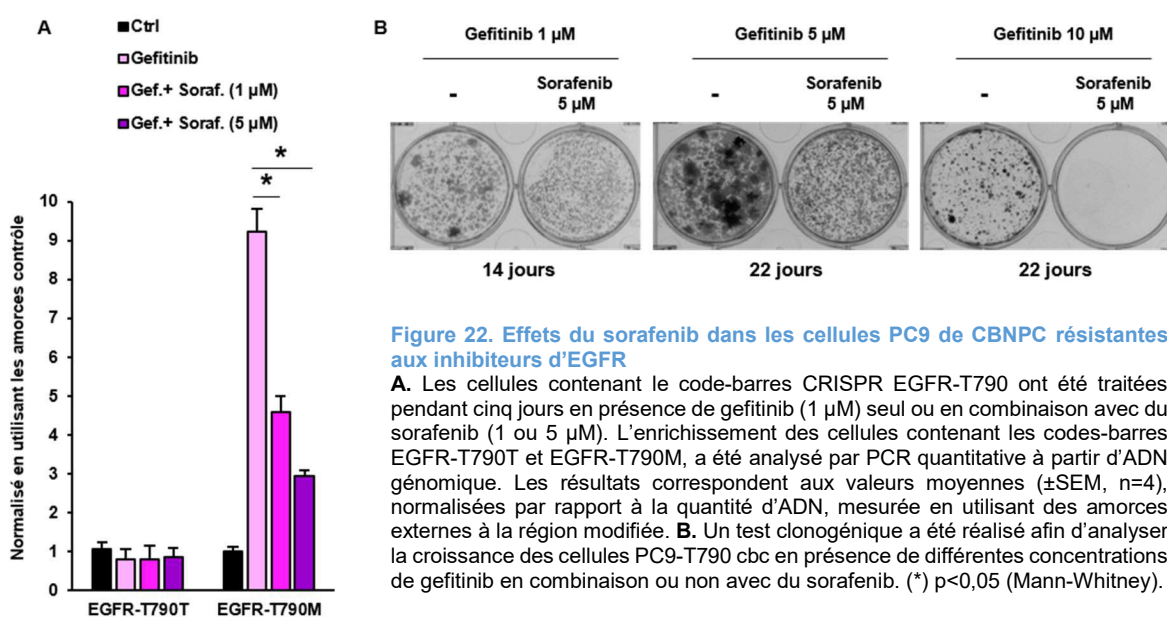


Figure 22. Effets du sorafenib dans les cellules PC9 de CBNPC résistantes aux inhibiteurs d'EGFR

A. Les cellules contenant le code-barres CRISPR EGFR-T790 ont été traitées pendant cinq jours en présence de gefitinib (1 μM) seul ou en combinaison avec du sorafenib (1 ou 5 μM). L'enrichissement des cellules contenant les codes-barres EGFR-T790T et EGFR-T790M, a été analysé par PCR quantitative à partir d'ADN génomique. Les résultats correspondent aux valeurs moyennes (\pm SEM, n=4), normalisées par rapport à la quantité d'ADN, mesurée en utilisant des amorces externes à la région modifiée. **B.** Un test clonogénique a été réalisé afin d'analyser la croissance des cellules PC9-T790M en présence de différentes concentrations de gefitinib en combinaison ou non avec du sorafenib. (*) $p < 0,05$ (Mann-Whitney).

le gefitinib entraîne la formation de colonies, correspondant à une amplification des cellules résistantes. Cet effet, déjà évident après deux semaines de traitement, devient plus marqué après trois semaines. En revanche, en présence de sorafenib la taille des colonies est fortement réduite (**Fig. 22B**), en accord avec la capacité de cet inhibiteur multikinase à prévenir l'émergence des clones contenant la mutation T790M (**Fig. 22A**).

En accord avec ces résultats, nous avons pu montrer par CRISPR-barcoding que le sorafenib prévient la résistance au gefitinib induite par la mutation T790M dans deux autres lignées cellulaires dérivées de CBNPC sensibles aux inhibiteurs d'EGFR, les cellules HCC4006 et HCC827 (**Fig. 23**).

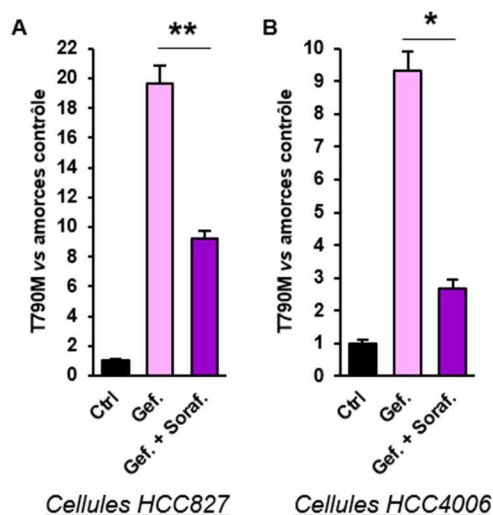


Figure 23. Effets du sorafenib sur la résistance de différentes lignées cellulaires de CBNPC aux inhibiteurs d'EGFR

Les cellules HCC827 (A) et HCC4006 (B) contenant le code-barres CRISPR EGFR-T790 ont été traitées pendant respectivement cinq et neuf jours en présence de gefitinib (50 nM pour les cellules HCC827, ou 1 μ M pour les cellules HCC4006) seul, ou en combinaison avec le sorafenib (5 μ M). L'enrichissement des cellules contenant le code-barres EGFR-T790M, a été analysé par PCR quantitative à partir d'ADN génomique. Les résultats correspondent aux valeurs moyennes (\pm SEM, n=5 (HCC827), n=4 (HCC4006)), normalisées par rapport à la quantité totale d'ADN, mesurée en utilisant des amorces externes à la région modifiée, et sont représentatives de 3 expériences indépendantes. (*) p<0,05 et (**) p<0,01 (Mann-Whitney).

2.4. Le sorafenib prévient la résistance des cellules de CBNPC aux inhibiteurs d'EGFR de 3^{ème} génération

Nous avons ensuite étudié les effets du sorafenib sur la résistance aux inhibiteurs d'EGFR de 3^{ème} génération induite par la mutation tertiaire C797S de ce récepteur. Pour cela, les cellules PC9 contenant le code-barres CRISPR EGFR-T790M décrites ci-dessus ont été sélectionnées en présence de gefitinib pendant plusieurs semaines, afin d'obtenir une population résistante aux inhibiteurs de 1^{ère} génération. A partir de ces cellules, nous avons réalisé une expérience de CRISPR-barcoding pour générer des clones présentant les codes-barres EGFR-C797S muté ou EGFR-C797C contrôle. La **Figure 24A et B** montre que, en présence des inhibiteurs d'EGFR de 3^{ème} génération WZ4002 et osimertinib, la proportion de cellules contenant la mutation C797S augmente. En revanche, un co-traitement avec le sorafenib prévient cet enrichissement (**Fig. 24A et B**). En accord avec ces résultats obtenus par qPCR, nous avons montré par un test de clonogénicité que le sorafenib permet également d'empêcher la formation de colonies résistantes en présence d'osimertinib (**Fig. 24C**). Ces résultats indiquent que le sorafenib est capable de prévenir la résistance des cellules de CBNPC à différentes classes d'inhibiteurs d'EGFR.

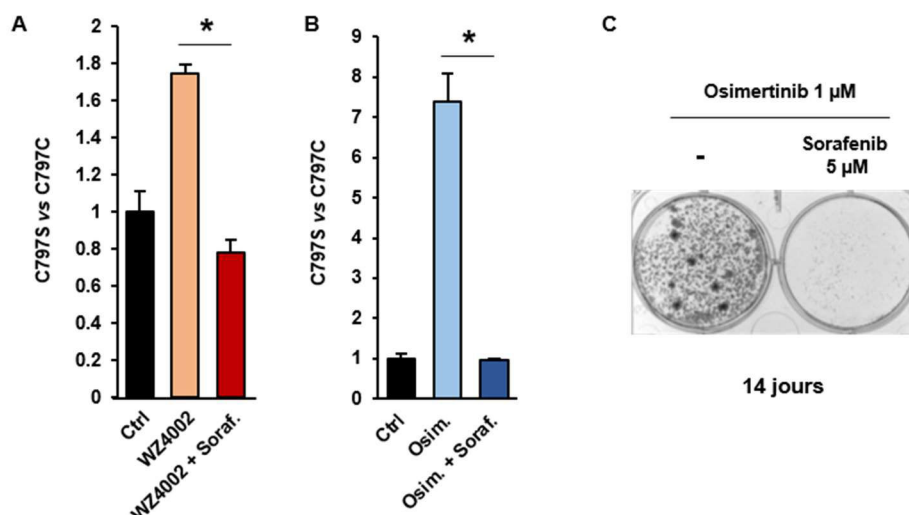


Figure 24. Le sorafenib prévient la résistance des CBNPCs à l'osimertinib induite par la mutation C797S

A-B. Les cellules PC9-T790M contenant le code-barres CRISPR EGFR-C797 ont été traitées avec les inhibiteurs d'EGFR de troisième génération WZ4002 (1 µM ; **A**) ou osimertinib (1 µM ; **B**) en présence ou en l'absence de sorafenib (5 µM), pendant respectivement cinq et neuf jours. L'enrichissement des cellules contenant les codes-barres EGFR-T790M a été analysé par PCR quantitative à partir d'ADN génomique. Les résultats correspondent aux valeurs moyennes (\pm SEM, n=4), normalisées par rapport à la quantité d'ADN, mesurée en utilisant des amorces externes à la région modifiée, et sont représentatives de 3 expériences indépendantes. **C.** Un test clonogénique a été réalisé en utilisant les mêmes cellules en présence d'osimertinib en combinaison ou non avec du sorafenib. (*) p<0,05 (Mann-Whitney).

2.5. Mécanisme d'action du sorafenib dans les cellules de CBNPC

Le sorafenib inhibe plusieurs RTKs, tels que VEGFR, PDGFR, c-kit et RET, mais également la sérine/thréonine kinase RAF (Wilhelm *et al.*, 2006). Nous avons étudié l'impact du sorafenib sur les voies intracellulaires activées par l'EGFR dans différentes lignées cellulaires issues de CBNPC sensibles aux inhibiteurs de ce récepteur. A l'aide d'une analyse par Western blot, nous avons observé que le sorafenib entraîne une diminution précoce de la phosphorylation de STAT3 dans les cellules PC9 parentales (**Fig. 25A**). En revanche, malgré sa capacité à inhiber RAF, aucun effet n'a été remarqué sur la voie des MAPK, contrairement au tramétinib, un inhibiteur de MEK, utilisé ici comme contrôle positif (**Fig. 25A**). Après trois jours de traitement, le sorafenib induit une forte diminution de l'expression de l'EGFR, qui est accompagnée par une diminution des taux de phosphorylations de ERK, AKT et de la protéine

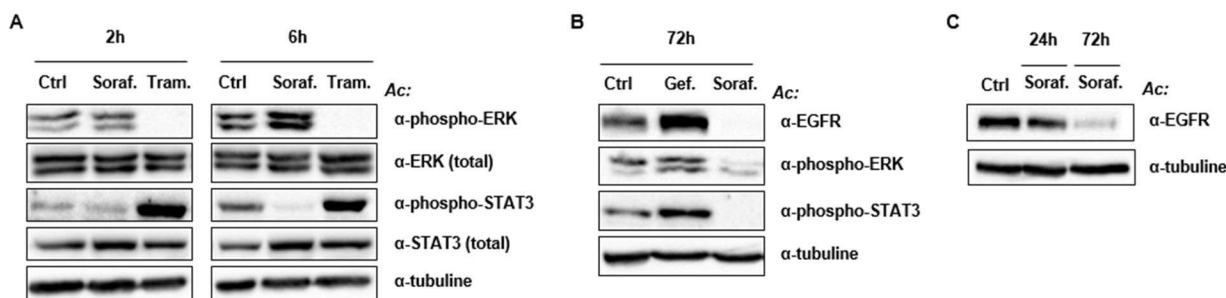


Figure 25. Effets du sorafenib dans les cellules PC9

A. Les cellules PC9 ont été traitées avec le sorafenib (5 µM) ou l'inhibiteur de MEK tramétinib (50 nM) pendant 2h et 6h. Les protéines ont ensuite été extraites pour réaliser un western blot en utilisant les anticorps indiqués. **B.** Les cellules PC9 ont été traitées pendant 72h avec le gefitinib (1 µM) ou le sorafenib (5 µM) et le lysat cellulaire a été extrait. Un western blot a ensuite été réalisé en utilisant les anticorps indiqués. **C.** Les cellules PC9 ont été traitées pendant 24h et 72h avec le sorafenib (5 µM). Le lysat cellulaire a ensuite été extrait pour réaliser un western blot en utilisant les anticorps indiqués.

ribosomale S6, correspondant, respectivement, à des composants des voies MAPK, AKT et mTOR (**Fig. 25B et C** et données non présentées).

Comme le montre la **Figure 26**, ces résultats ont été confirmés dans les cellules HCC827 et HCC4006. L'ensemble de ces données suggère que le sorafenib agirait en deux temps, tout d'abord *via* une inhibition précoce de la phosphorylation de STAT3, puis, plus tardivement, en diminuant le niveau d'expression de l'EGFR, perturbant ainsi les voies intracellulaires activées par ce récepteur.

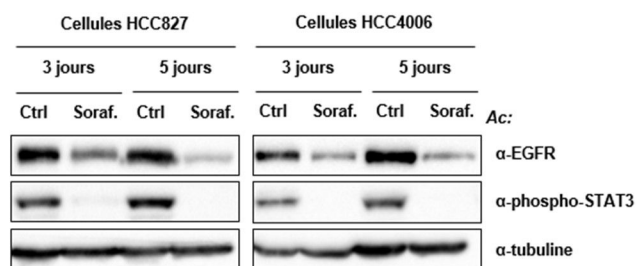


Figure 26. Effets du sorafenib dans les cellules HCC827 et HCC4006 dérivées de CBNPC

Les cellules ont été traitées en présence ou en absence du sorafenib (5 µM), et le lysat cellulaire a été extrait à différents temps, suivi d'un western blot en utilisant les anticorps indiqués.

D'autres inhibiteurs multikinases présentant un profil pharmacologique similaire à celui du sorafenib ont été approuvés en clinique pour le traitement de différents types de cancer. Parmi ces molécules, le sunitinib (*Sutent* ; Pfizer), utilisé pour le traitement des tumeurs stromales gastro-intestinales, du cancer du rein et du pancréas, inhibe le VEGFR, PDGFR, FLT-3, c-Kit, RET, mais n'affecte pas RAF (**Tableau 3**) (Faivre *et al.*, 2007). Nous avons comparé les effets de ces deux composés sur la résistance des cellules de CBNPC aux inhibiteurs d'EGFR. Dans des expériences de CRISPR-barcoding, le sunitinib, contrairement au sorafenib, ne parvient pas à prévenir l'amplification des cellules contenant la mutation EGFR-

Tableau 3. Profils pharmacologiques des inhibiteurs multi-kinases sorafenib, regorafenib et sunitinib

Sorafenib	Regorafenib	Sunitinib
Cibles pharmacologiques communes		
VEGFR-1	VEGFR-1	VEGFR-1
VEGFR-2	VEGFR-2	VEGFR-2
VEGFR-3	VEGFR-3	VEGFR-3
PDGFR-β	PDGFR-β	PDGFR-β
FGFR-1	FGFR1	
C-Raf	C-Raf	
B-RAF	B-RAF	
B-RAF V600E	B-RAF V600E	
c-Kit	c-Kit	c-Kit
RET	RET	RET
Flt-3		Flt-3
Autres cibles pharmacologiques		
p38		
	TIE2	
		PDGFR-α
		CSF1R

CSF1R, Colony stimulating factor 1 receptor ; EGFR, epidermal growth factor receptor ; FGFR, fibroblast growth factor receptor ; Flt-3, fms-related tyrosine kinase 3 ; Kit, stem-cell growth factor receptor ; p38, p38 mitogen-activated protein kinases ; PDGFR, platelet-derived growth factor receptor ; TIE2, angiopoietin receptor ; VEGFR, vascular endothelial growth factor receptor. Adapté de Faivre *et al.*, 2007 ; Wilhelm *et al.*, 2006, 2011.

T790M induite par le gefitinib (Fig. 27).

Ces résultats nous ont conduits à émettre l'hypothèse selon laquelle les effets du sorafenib dans les cellules de CBNPC pourraient dépendre de la sérine/thréonine kinase RAF. Par ailleurs, des études basées sur des modèles transgéniques ont révélé que l'une des trois isoformes de RAF, CRAF, serait impliquée dans la formation de tumeurs de la peau et du poumon, *via* un mécanisme pas complètement élucidé, mais indépendant de la voie MAPK (Ehrenreiter *et al.*, 2009 ; Blasco *et al.*, 2011 ; Karreth *et al.*, 2011). En particulier, il a été montré dans les cancers de la peau que CRAF peut

activer STAT3 *via* l'inhibition de la voie *Rho associated coiled-coil containing protein kinase* (Rock)-cofiline, participant ainsi au processus tumorigénique (Ehrenreiter *et al.*, 2009). Afin de tester l'implication de CRAF dans le mécanisme d'action du sorafenib dans les cellules de CBNPC, nous avons étudié les effets de sa sous-expression à l'aide de shRNA inductibles. Nous avons observé que l'inhibition de cette kinase n'affecte pas la croissance des cellules PC9, ni le taux de phosphorylation de STAT3 (Fig. 28).

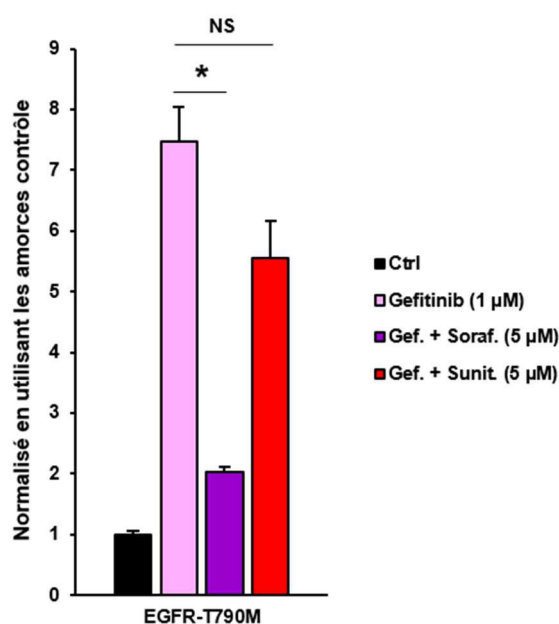


Figure 27. Le sunitinib ne prévient pas la résistance des cellules de CBNPC aux inhibiteurs d'EGFR
 Les cellules PC9 contenant le code-barres CRISPR EGFR-T790 ont été traitées pendant cinq jours en présence de gefitinib (1 µM) seul ou en combinaison avec sorafenib (5 µM) ou sunitinib (5 µM). L'enrichissement des cellules contenant le code-barres EGFR-T790M a été analysé par PCR quantitative à partir d'ADN génomique. Les résultats correspondent aux valeurs moyennes (±SEM, n=4), normalisées par rapport à la quantité totale d'ADN mesurée en utilisant des amorces contrôle ciblant une région non modifiée du gène *EGFR*. * p<0.05, NS – Non significatif (Test de Mann-Whitney).

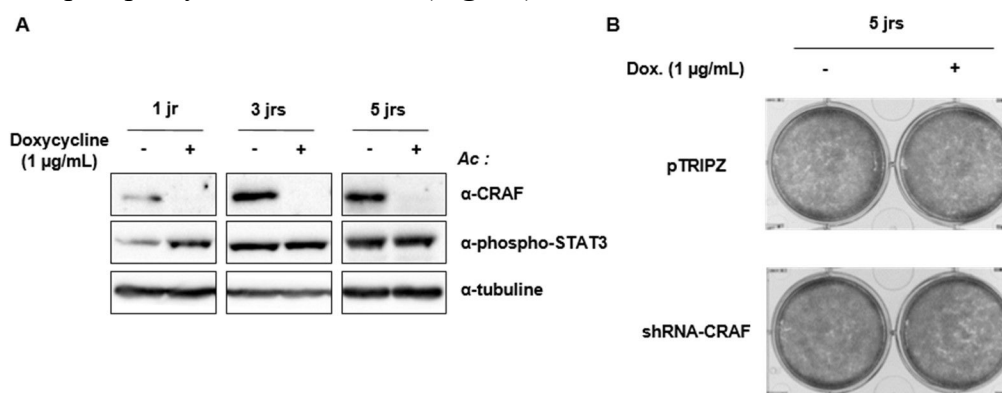


Figure 28. L'inhibition de CRAF n'a pas d'effet sur la phosphorylation de STAT3 et la croissance des cellules de CBNPC
A. Les cellules PC9, préalablement infectées avec un vecteur lentiviral permettant l'expression inductible d'un shRNA ciblant CRAF, ont été traitées pendant un, trois et cinq jours en présence de doxycycline (1 µg/mL), et l'état de phosphorylation de STAT3 a été analysés par western blot. **B.** Les mêmes cellules ont été utilisées dans un test de clonogénicité afin de tester les effets de l'inhibition de CRAF sur la croissance cellulaire. Le vecteur pTRIPZ vide a été utilisé comme contrôle négatif.

L'incapacité du sunitinib et des shRNA dirigés contre CRAF à reproduire les effets du sorafenib dans les cellules de cancer bronchique est en accord avec des données obtenues au laboratoire qui ont révélé que d'autres inhibiteurs spécifiques de RAF, tels que le dabrafenib ou CCT196969, n'affectent pas l'activation de STAT3 (données non illustrées). Ces données indiquent que l'inhibition de CRAF n'est pas suffisante pour récapituler l'effet du sorafenib dans les cellules de CBNPC

2.6. Le sorafenib prévient la résistance aux inhibiteurs d'EGFR *in vivo*

Dans le but d'étudier le potentiel thérapeutique du sorafenib, nous avons réalisé une expérience de xenogreffe de cellules de CBNPC, en collaboration avec le Pr S. Aaronson. Les cellules PC9 contenant la mutation T790M et le code-barres CRISPR C797S (voir paragraphe précédent) ont été injectées bilatéralement en sous-cutanées chez des souris SCID. Au bout de 27 jours, quand les tumeurs ont atteint un volume moyen de 40-50 mm³, les souris ont été traitées par gavage avec l'osimertinib (5 mg/kg), le sorafenib (60 mg/kg) ou une combinaison des deux composés. Contrairement aux souris contrôles, qui présentent un développement tumoral rapide, les souris traitées par l'osimertinib répondent dans un premier temps au traitement. Cependant, après quelques jours la croissance tumorale reprend, même si la vitesse de croissance est moins importante que celle observée dans le groupe contrôle. Moins efficace que l'osimertinib, le sorafenib permet également de retarder la progression des tumeurs (**Fig. 29A**). En revanche,

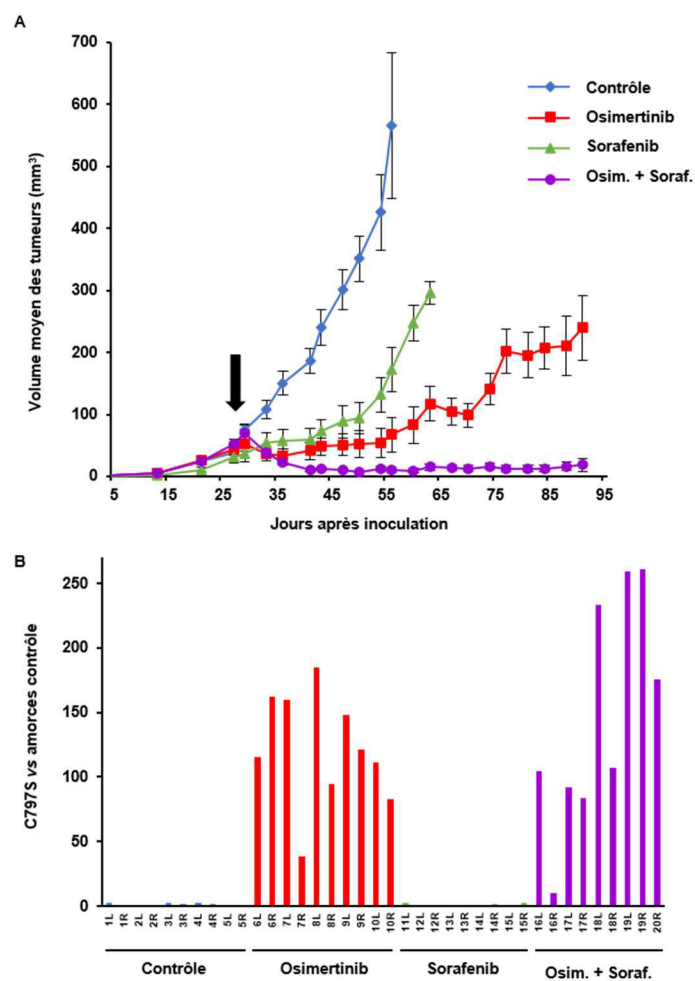


Figure 29. Effets du sorafenib sur la croissance tumorale

Les cellules PC9-T790M contenant le code-barres CRISPR EGFR-C797 ont été injectées en sous-cutanée dans les flancs des souris SCID mâles (2×10^6 cellules par site avec matrigel). Une fois que les tumeurs ont atteint un volume moyen de 50 mm³, les souris ont été réparties en quatre lots : contrôle, traitées par l'osimertinib (5 mg/kg), le sorafenib (60 mg/kg) ou les deux composés (flèche). **A**. La taille de la tumeur a été évaluée tous les trois-quatre jours. **B**. Niveau d'enrichissement du code-barres EGFR-C797S dans chacune des tumeurs issues des flancs droits "R" et gauches "L" des souris. Les résultats correspondent aux valeurs normalisées par rapport à la quantité d'ADN, mesurée en utilisant des amorces externes à la région modifiée.

la combinaison osimertinib-sorafenib permet de réduire la taille des tumeurs, qui reste stable pendant deux mois, jusqu'à la fin de l'expérience (**Fig. 29A**). Ces résultats sont en accord avec nos expériences *in vitro* et indiquent qu'un co-traitement avec le sorafenib peut prévenir la résistance *in vivo* des cellules de CBNPC aux inhibiteurs d'EGFR.

Nous avons ensuite évalué la proportion du code-barres EGFR-C797S dans l'ADN génomique extrait à partir de chaque tumeur. Ces analyses ont révélé un enrichissement de plus de 100 fois de l'étiquette C797S dans les groupes osimertinib et osimertinib-sorafenib (**Fig. 29B**). Considérant que l'efficacité de CRISPR-barcoding dans les cellules PC9 est d'environ 0,5-1% (Guernet *et al.*, 2016), ces résultats suggèrent que le traitement par l'osimertinib élimine la plupart des cellules sensibles, alors que la croissance des clones résistants contenant la mutation EGFR-C797S est arrêtée par le sorafenib.

DISCUSSION

1. CRISPR-barcoding

Le cancer est le résultat d'un processus évolutif au cours duquel les cellules accumulent différentes mutations *driver* et *passenger* (Nowell, 1976). Par conséquent, chaque tumeur est composée de sous-populations clonales génétiquement distinctes, ce qui est à la base de l'extraordinaire capacité des cellules cancéreuses à progresser et à s'adapter à leur environnement, et particulièrement aux thérapies (Gillies *et al.*, 2012 ; McGranahan et Swanton, 2015). Ces dernières années, les avancées fulgurantes des technologies de séquençage à haut débit ont permis l'identification dans les cancers de multiples aberrations génétiques potentiellement oncogéniques (Kamps *et al.*, 2017). La caractérisation de ces nouvelles mutations nécessite des méthodes efficaces pour évaluer leur rôle dans la formation et la progression des tumeurs. Nous avons mis au point une nouvelle stratégie, que nous avons appelée CRISPR-barcoding, qui permet d'étudier l'impact fonctionnel d'altérations génétiques dans un contexte qui récapitule l'hétérogénéité intratumorale.

1.1. Avantages et limitations de l'approche CRISPR-barcoding

1.1.1. Avantages de CRISPR-barcoding

CRISPR-barcoding est une stratégie rapide et flexible permettant d'étudier les conséquences fonctionnelles d'une modification génétique d'intérêt dans divers essais expérimentaux. Parmi les avantages majeurs, notre approche permet de s'affranchir des effets non-spécifiques dus aux clivages hors cibles potentiels par la Cas9, grâce aux codes-barres contrôle utilisés pour la normalisation. De plus, CRISPR-barcoding n'implique pas la sélection et l'amplification des cellules ayant incorporé la mutation, évitant ainsi le biais dû à la dérivation de clones. Contrairement aux autres stratégies basées sur le système CRISPR/Cas9, notre approche permet d'étudier différents types de mutations, y-compris celles qui affectent négativement la croissance cellulaire. Enfin, CRISPR-barcoding ne nécessite pas une efficacité élevée d'édition du génome, élargissant donc le choix de modèles cellulaires utilisables, et peut facilement être adaptée à une configuration multiplexe.

1.1.2. Limitations de CRISPR-barcoding

Dans notre étude publiée dans *Molecular Cell*, le taux d'efficacité pour l'incorporation des codes-barres variait entre 0.5 et 9%, en fonction du type cellulaire utilisé et de la séquence ciblée (Guernet *et al.*, 2016). En revanche, la proportion d'allèles réparés par NHEJ est généralement plus élevée (données non illustrées) (Maruyama *et al.*, 2015). Notre approche permet de suivre l'évolution des cellules ayant incorporé les codes-barres contrôle ou mutant,

mais pas celles contenant des *indels*, qui peuvent influencer le comportement de la population cellulaire globale. Afin de s'affranchir de ce biais potentiel, nous normalisons les résultats obtenus pour les codes-barres mutés avec ceux obtenus pour les codes-barres contrôles. Par ailleurs, bien que les mutations silencieuses qui composent les codes-barres soient susceptibles d'altérer l'épissage, la fidélité de transcription ou la stabilité de l'ARN messager (Sauna et Kimchi-Sarfaty, 2011), nous avons montré que des expériences consistant à inverser les mutations silencieuses entre les codes-barres mutants et contrôles peuvent être réalisées en tant que contrôle de spécificité (Guernet *et al.*, 2016).

Notre approche permet d'étudier rapidement et facilement l'effet de mutations oncogéniques dans les cellules. L'introduction de ce type de modification est relativement simple pour les oncogènes, puisque leur action est dominante, et donc il suffit qu'un seul allèle soit muté pour observer un effet sur les cellules. En revanche, généralement il faut que toutes les copies d'un oncosuppresseur soient modifiées pour qu'il y ait une perte de fonction de la protéine codée, et donc un effet oncogénique. Cela peut correspondre à l'introduction du code-barres sur les deux allèles ou bien à un code-barres associé à l'inactivation du second allèle par NHEJ. En ce qui concerne la réparation de mutations oncogéniques préexistantes, celles-ci sont plus simples pour les oncosuppresseurs, pour lesquels la correction d'un seul allèle est souvent suffisante à rétablir l'expression d'une protéine fonctionnelle. Nous avons montré que CRISPR-barcoding peut être également utilisé pour étudier les effets de la réparation d'un oncogène. A titre d'exemple, nous avons corrigé une mutation activatrice hétérozygote de l'oncogène ALK dans les cellules Kelly dérivées de neuroblastome. Comme le montre la **Figure 30**, l'introduction de codes-barres peut aboutir à quatre configurations différentes. Parmi ces possibilités, trois devraient permettre de rétablir l'expression d'un récepteur sauvage : i) la réparation de l'allèle mutant sans altération du second allèle sauvage ; ii) l'incorporation du code-barres réparé sur les deux allèles du gène ; iii) l'incorporation du code-barres sauvage associée à l'inactivation du second allèle par NHEJ (probablement la situation la plus fréquente). Le quatrième cas de figure correspond à l'insertion du code-barres réparé à la place de l'allèle sauvage, alors que la copie mutée du gène reste inaltérée. Ainsi, dans trois cas sur quatre, la forme phénotypiquement dominante de ALK exprimée par la cellule devrait être celle codée par l'allèle contenant le code-barres.

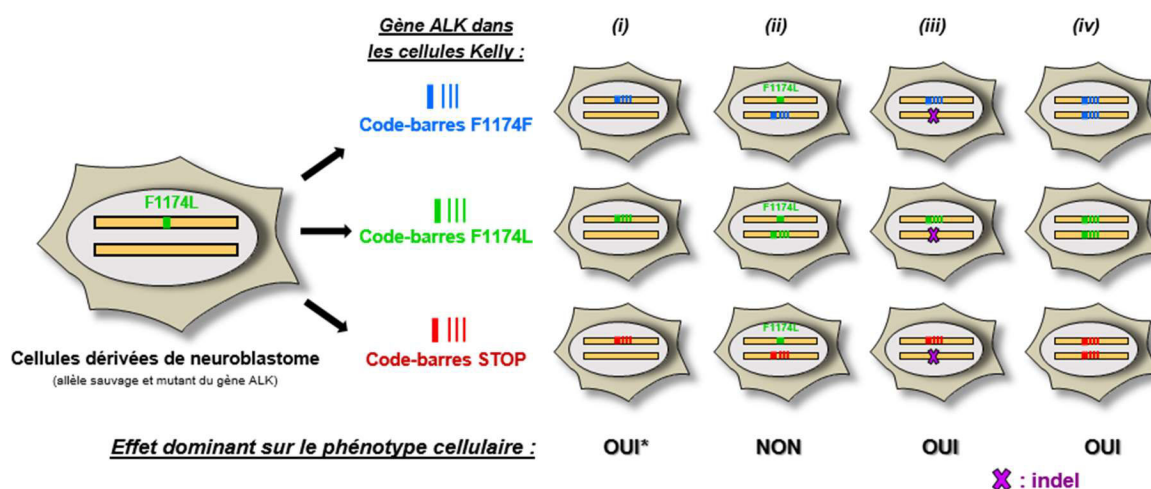


Figure 30. Réparation de la mutation oncogénique de ALK dans les cellules Kelly de neuroblastome

Schéma illustrant les différents génotypes possibles pour le codon ALK-F1174L à la suite de la modification par CRISPR-barcoding dans les cellules Kelly dérivées de neuroblastome. Les effets possibles sur le phénotype cellulaire sont indiqués. (*) : Effet dominant uniquement référé à la mutation F1174L.

Enfin, dans une expérience de CRISPR-barcoding il est important de tenir compte de la ploïdie des cellules étudiées et de l'éventuelle amplification du gène d'intérêt. Comme vu précédemment, le taux d'efficacité d'incorporation d'un code-barres est relativement faible. Si le gène est présent en plusieurs copies, il est possible qu'un nombre insuffisant d'allèles soit muté et, par conséquent, que l'insertion de l'étiquette génétique ne se traduise pas par un effet sur le phénotype des cellules. Afin de contrecarrer cette limitation potentielle, on pourrait imaginer d'augmenter l'efficacité d'incorporation des codes-barres, notamment *via* l'utilisation de différents sgRNA, ou bien en traitant les cellules avec des inhibiteurs de NHEJ, qui, selon certains auteurs, peuvent augmenter l'efficacité de HDR (Chu *et al.*, 2015 ; Maruyama *et al.*, 2015 ; Yu *et al.*, 2015).

Dans notre étude nous avons montré que CRISPR-barcoding est particulièrement bien adapté pour récapituler l'émergence de cellules contenant un mécanisme de résistance basé sur une mutation ponctuelle ou un remaniement chromosomique. Des approches d'étiquetage des cellules pourraient également être envisagées pour d'autres aberrations génétiques, telles que l'amplification d'un oncogène ou la délétion d'un oncosuppresseur. En revanche, l'utilisation de ce type de stratégie pourrait être plus difficile pour des mécanismes de résistance plus complexes et pas forcément associés à l'altération d'un seul gène, tels qu'une transition épithélio-mésenchymateuse ou une transformation histologique en CBPC.

1.2. Méthodes de détection des codes-barres

1.2.1. PCR quantitative

La méthode de détection que nous avons privilégiée pour la détection des codes-barres est la PCR quantitative (qPCR). A l'aide d'outils bioinformatiques, il est possible de générer facilement des amorces de qPCR permettant de détecter spécifiquement les fractions de cellules contenant l'un ou l'autre des codes-barres. Le principe de cette technique repose sur la capacité de suivre en temps réel le processus de PCR grâce à l'incorporation d'un composé fluorescent, le SYBR Green, dans les nouveaux fragments générés d'ADN double brin. Les données de fluorescence sont collectées à chaque cycle de la PCR, ce qui permet de déterminer le *threshold cycle* (Ct), c'est-à-dire le cycle pour lequel le signal fluorescent dû à l'amplification de l'ADN est significativement supérieur au bruit de fond. Puisqu'il est mesuré lors de la phase exponentielle de la PCR, le Ct dépend directement de la quantité de l'ADN cible dans l'échantillon.

Bien que cette approche soit relativement peu coûteuse, rapide et facile d'utilisation, elle présente néanmoins quelques limitations. La plus importante relève de la sensibilité de la technique. En effet, comme mentionné précédemment, nos études montrent que les cellules modifiées par CRISPR-barcoding ne représentent que 0.4 à 9% de la population totale. En fonction de l'efficacité de la PCR, qui dépend largement de la séquence des amorces, et qui est donc variable selon le gène étudié, une proportion faible de code-barres peut se révéler proche de la limite de détection de cette technique. Afin de surmonter cette limitation, il est possible d'optimiser les conditions de transfection et d'édition de l'ADN, ou bien d'améliorer l'efficacité de la PCR en modifiant la séquence des amorces et/ou en choisissant d'autres mutations silencieuses.

1.2.2. Autres technologies de PCR

La PCR TaqMan consiste en l'utilisation, en plus d'un couple d'amorces, d'une sonde d'hydrolyse spécifique de la séquence à amplifier. Cette sonde est couplée à un fluorochrome et à un quencher, capable de masquer le signal fluorescent. Lors de l'amplification, la polymérase, grâce à son activité exonucléase 5'-3', clive la sonde hybridée à la séquence cible, provoquant la séparation du fluorochrome et du quencher, et donc l'émission de fluorescence. Cette technologie permet ainsi d'améliorer la spécificité de la qPCR. Nous avons utilisé ce type d'approche pour détecter les codes-barres EGFR-T790T et EGFR-T790M lors du criblage effectué dans les cellules PC9 : en utilisant deux sondes contenant deux fluorochromes

différents, nous avons pu analyser la proportion des étiquettes sauvage et mutée simultanément à partir d'une même réaction de PCR.

Bien que coûteuse, la PCR digitale, également appelée PCR numérique, permet une quantification plus précise comparée aux autres techniques. Le principe de cette technologie consiste à diluer et à répartir un échantillon d'ADN en plusieurs milliers de chambres de réaction de PCR indépendantes, qui vont contenir ou non la séquence d'intérêt. Au cours de la réaction d'amplification, la présence de la molécule cible se traduira par un signal fluorescent. L'analyse de la fluorescence dans chaque chambre permettra ainsi d'évaluer la proportion de molécules cibles dans l'échantillon. Ce type d'approche pourrait être utilisé pour détecter les codes-barres faiblement représentés, notamment dans les modèles cellulaires présentant un taux modeste d'efficacité d'édition du génome.

1.2.3. Séquençage à haut débit

Le séquençage à haut débit, ou de nouvelle génération (NGS, *next generation sequencing*), est une approche qui, basée sur des technologies différentes, permet la lecture de plusieurs milliards de bases à la fois. Contrairement à la qPCR, qui repose sur la comparaison des fractions relatives d'un code-barres entre différents échantillons, le NGS peut être utilisé pour mesurer directement, dans un même échantillon, la quantité d'étiquettes CRISPR par rapport à la séquence endogène non modifiée. Le NGS peut également fournir des informations supplémentaires, notamment concernant la proportion de cellules contenant des *indels* dans le locus ciblé.

Un autre avantage du NGS comparé à la qPCR est lié à la possibilité de vérifier les éventuelles erreurs de séquence associées aux codes-barres. Par exemple, dans notre modèle cellulaire dérivé du cancer du côlon, nous avons montré une diminution au cours du temps de la fraction de cellules dans lesquelles l'oncosuppresseur *APC* avait été réparé par CRISPR-barcoding. Cependant, après quelques semaines nous avons remarqué que ces cellules ne disparaissaient pas, mais que, au contraire, leur proportion atteignait un plateau. Les analyses de NGS réalisées à partir de ces échantillons ont révélé que, un mois après l'insertion des étiquettes, la plupart des codes-barres réparés présentaient en réalité des mutations supplémentaires, qui empêchaient la restauration d'une protéine APC fonctionnelle. Ces altérations, probablement provoquées par une erreur lors de la réparation du génome ou dans la synthèse de l'ADN donneur, avaient donc généré des faux codes-barres réparés, qui n'auraient pas pu être mis en évidence par qPCR.

1.3. Utilisation de codes-barres complexes pour tracer différents clones au sein d'une population cellulaire

1.3.1. Approches antérieures

Dans les années 1980, le groupe de R. Mulligan utilisa pour la première fois un rétrovirus non réplicatif pour marquer différentes cellules au sein d'une population : grâce à l'analyse des sites d'intégration du virus dans le génome, ces expériences permirent d'étudier la dynamique et la différenciation des cellules souches hématopoïétiques (Lemischka *et al.*, 1986). Quelques années plus tard, Korczak et collaborateurs eurent recours au même type d'approche pour analyser l'hétérogénéité clonale de tumeurs et métastases après allogreffe chez la souris de cellules murines dérivées de cancer mammaire (Korczak *et al.*, 1988). Avec les progrès des modèles transgéniques et l'utilisation de nouveaux marqueurs fluorescents, il est maintenant possible de tracer l'évolution de cellules souches et progénitrices pendant le développement et chez l'adulte (Blanpain et Simons, 2013), mais également au cours de la tumorigenèse (Schepers *et al.*, 2012). Bien que puissantes, ces approches ne permettent pourtant pas de suivre le devenir et la dynamique de populations cellulaires extrêmement complexes. Pour contourner cette limitation technologique, le groupe de I. Weissman utilisa un vecteur lentiviral contenant une séquence partiellement dégénérée d'une trentaine de nucléotides, afin de marquer individuellement plusieurs milliers de cellules souches hématopoïétiques et ensuite étudier leur capacité à reconstituer les différents lignages myéloïdes et lymphoïdes après transplantation dans des souris irradiées (Lu *et al.*, 2011). Le même type de banque lentivirale de codes-barres a récemment été adapté pour étudier l'hétérogénéité intra-tumorale et suivre, au sein d'une population de cellules tumorales, le devenir de certains clones, tels que, par exemple, ceux capables de résister à une thérapie ciblée (Bhang *et al.*, 2015 ; Hata *et al.*, 2016 ; Nguyen *et al.*, 2014).

1.3.2. Introduction de codes-barres complexes *via* CRISPR-barcoding

Nous avons adapté CRISPR-barcoding pour marquer et suivre le devenir de plusieurs milliers de cellules cancéreuses au sein d'une même population *via* l'introduction d'une séquence d'ADN dégénérée (ou code-barres complexe) à un endroit précis du génome. En appliquant cette approche à une lignée dérivée de cancer mammaire humain, nous avons pu montrer que certains clones présentaient un potentiel tumorigénique plus important, indépendamment de leur taux de croissance *in vitro*. En insérant des codes-barres dégénérés dans les cellules de CBNPC, puis en les traitant en présence ou en l'absence d'un inhibiteur

d'EGFR, nous avons également révélé que des cellules résistantes pouvaient être présentes avant le début du traitement (Guernet *et al.*, 2016).

Parmi les avantages majeurs de CRISPR-barcoding, notre approche permet de maîtriser le site d'insertion de la séquence cible grâce au sgRNA qui détermine l'endroit exact du clivage de l'ADN. L'introduction de l'étiquette dans un endroit précis du génome réduit fortement la possibilité d'effets non-spécifiques. De plus, comparée à la génération des banques lentivirales, l'utilisation de CRISPR-barcoding est beaucoup plus simple et rapide, ce qui rend cette stratégie accessible à la plupart des laboratoires de recherche.

Concernant les limitations des codes-barres CRISPR hautement complexes, le taux de réparation de l'ADN par HDR est relativement faible, ce qui implique que seulement une fraction des cellules contiendra l'étiquette génétique. Bien que les *indels* générées par NHEJ puissent également être utilisées en tant que codes-barres, le fait de ne pas pouvoir marquer individuellement tous les clones constituant une population cellulaire peut se révéler un obstacle considérable pour certains types d'applications. Un autre inconvénient potentiel dépend de la possibilité qu'au sein d'une même cellule les deux allèles incorporent des codes-barres dégénérés différents, compliquant ainsi le suivi des cellules et l'analyse des résultats. Quand cela est possible, l'introduction des codes-barres dans une région haploïde du génome, telle qu'un locus localisé sur le chromosome X ou Y dans des cellules dérivées d'un individu de sexe masculin, garantirait un marquage avec une seule étiquette par cellule.

2. Sorafenib et CBNPC

Dans mes travaux de thèse, nous avons montré que le sorafenib est capable de prévenir l'amplification de cellules de CBNPC insensibles aux inhibiteurs de l'EGFR, y-compris ceux de 3^{ème} génération, tels que l'osimertinib. Cet effet a été confirmé *in vivo*, en utilisant un modèle de xéno greffe de cellules de CBNPC basé sur notre approche de CRISPR-barcoding. Le sorafenib agit *via* un mécanisme inédit indépendant de la voie des MAPK, qui implique l'inhibition de la phosphorylation de STAT3 lors des phases précoces du traitement et la sous-expression de l'EGFR après 3-4 jours.

2.1. Mécanisme d'action du sorafenib dans les cellules de CBNPC

Le sorafenib est employé depuis plusieurs années pour le traitement des cancers du rein, du foie et de la thyroïde. Développée dans un premier temps pour inhiber la sérine/thréonine kinase RAF, cette molécule s'est également révélée être un inhibiteur de différents RTKs, tels que VEGFR, PDGFR, FGFR1, FLT-3, c-Kit et RET (**Tableau 3**, page 147 des Résultats) (Wilhelm

et al., 2006). Eu égard à son action sur les récepteurs VEGFR et PDGFR, les effets anti-tumoraux du sorafenib ont été souvent associés à ses propriétés anti-angiogéniques, bien que les mécanismes moléculaires responsables de son activité dans certains types de cancer n'aient pas été complètement caractérisés.

Il a été récemment montré que, indépendamment de son activité inhibitrice de kinase, le sorafenib peut promouvoir l'activité enzymatique de la tyrosine phosphatase *Src homology region 2 domain-containing phosphatase-1* (SHP-1), provoquant la déphosphorylation de STAT3 dans des cellules dérivées de carcinomes hépatocellulaires (Tai *et al.*, 2011). Afin d'évaluer l'implication éventuelle de ce mécanisme dans la levée de la résistance des cellules de CBNPC aux inhibiteurs d'EGFR, nous avons étudié les effets du SC-1, un analogue du sorafenib dépourvu d'activité inhibitrice de RAF, mais capable de promouvoir la déphosphorylation de STAT3 *via* SHP-1 (Chen *et al.*, 2011). Comme le montrent nos expériences de CRISPR-barcoding, le SC-1 permet de réduire l'enrichissement des cellules PC9 contenant la mutation EGFR-T790M induit par le gefitinib de façon comparable à celle du sorafenib (**Fig. 31**). Ces données suggèrent ainsi que les effets de ces inhibiteurs dans les cellules de CBNPC pourraient dépendre de l'augmentation de l'activité phosphatasique de SHP-1. Cependant, en utilisant une forme constitutivement active de SHP-1, un dominant négatif ou un shRNA lentiviral, nous avons pu montrer que l'augmentation ou la réduction de l'activité de cette phosphatase ne permet pas de mimer ou bloquer, respectivement, les effets du sorafenib dans les cellules de CBNPC (données non illustrées).

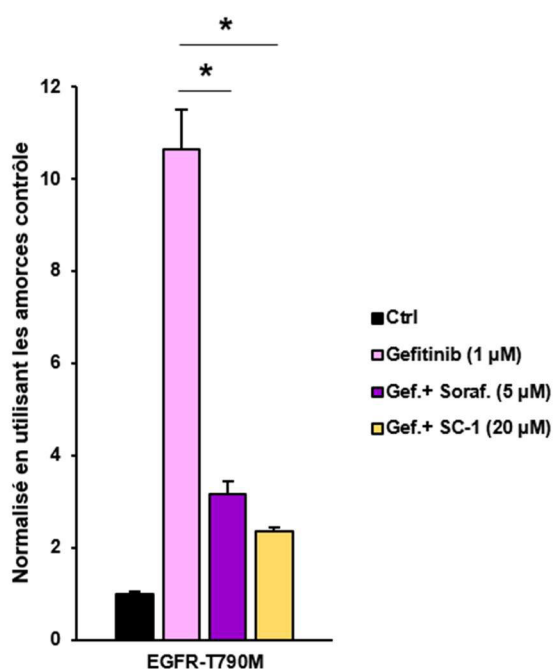


Figure 31. Comparaison des effets du sorafenib et du SC-1 sur la résistance aux inhibiteurs d'EGFR dans les cellules de CBNPC

Les cellules PC9 contenant le code-barres EGFR-T790 ont été traitées pendant cinq jours en présence du gefitinib (1 µM) seul ou en combinaison avec le sorafenib (5 µM) ou SC-1 (20 µM). L'enrichissement du code-barres EGFR-T790M a été analysé par PCR quantitative à partir d'ADN génomique. Les résultats correspondent aux valeurs moyennes (\pm SEM, n=4), normalisées par rapport à la quantité d'ADN mesurée en utilisant des amorces contrôle ciblant une région non modifiée du gène *EGFR*. *p<0.05 (Test de Mann-Whitney).

Le regorafenib (*Stivarga* ; Bayer), approuvé pour le traitement des cancers colorectaux et des tumeurs stromales gastro-intestinales, est un autre inhibiteur multikinase possédant un profil pharmacologique similaire à ceux du sorafenib et du sunitinib (Wilhelm *et al.*, 2011) (**Tableau 3**, page 147 des Résultats). Comme le sunitinib, ce composé s'est avéré incapable d'empêcher l'émergence de cellules de CBNPC résistantes aux inhibiteurs d'EGFR (données non présentées). Ces résultats suggèrent que les effets du sorafenib pourraient dépendre de l'action simultanée sur plusieurs cibles, telles que certains RTKs, RAF ou SHP-1. Cette combinaison serait spécifique du sorafenib, et la répression de CRAF ou l'activation de SHP-1 à elles seules ne seraient pas suffisantes pour récapituler les effets de ce composé. Comme alternative à cette hypothèse, il est également possible que le sorafenib agisse *via* une cible propre, non partagée par le sunitinib et le regorafenib, mais non encore identifiée.

L'identification de(s) cible(s) moléculaire(s) du sorafenib devrait permettre de mieux comprendre la série d'événements qui aboutit, après 3-4 jours de traitement, à la diminution de l'expression de l'EGFR et, par conséquent, à l'inhibition des différentes voies activées par ce récepteur, telles que MAPK, PI3K-AKT et mTOR. Des expériences préliminaires en utilisant un shRNA inductible ont révélé que la sous-expression de STAT3 n'est pas suffisante pour induire une réduction du taux d'EGFR (données non illustrées), indiquant que d'autres effecteurs précoces participent au mécanisme d'action du sorafenib. Afin d'obtenir une image globale des effets de ce composé dans les cellules de CBNPC, nous réaliserons une étude transcriptomique par RNA-seq après différents temps de traitement.

L'expérience de xéno greffe a révélé que, malgré la taille réduite des tumeurs, la combinaison osimertinib-sorafenib induit un fort enrichissement du code-barres EGFR-C797S, comparable à celui mesuré en présence de l'osimertinib seul (**Fig. 29B**). La courbe illustrée dans la **Figure 29A** montre que l'osimertinib, mais pas le sorafenib, induit une régression initiale du volume des tumeurs, suggérant que ces deux inhibiteurs pourraient provoquer respectivement la mort et l'arrêt de croissance des cellules de CBNPC. Ainsi, la combinaison osimertinib-sorafenib permettrait d'éliminer les cellules sensibles à l'inhibiteur d'EGFR, alors que les cellules contenant la mutation EGFR-C797S seraient toujours présentes, mais incapables de proliférer. Il est également possible que l'osimertinib et le sorafenib exercent un effet synergique sur les cellules avec un génotype EGFR-C797C, permettant donc l'enrichissement des cellules EGFR-C797S, sensibles uniquement à l'inhibiteur multi-kinase. Des expériences complémentaires vont être menées dans le but de caractériser les effets de la combinaison osimertinib-sorafenib dans les cellules de CBNPC (voir section "Conclusion et Perspectives").

2.2. Autres approches pour contrecarrer la résistance des CBNPCs aux inhibiteurs d'EGFR

Le groupe du Dr P. Jänne a récemment montré que la combinaison d'un inhibiteur d'EGFR avec le tramétinib (Mekinist, GlaxoSmithKline), un inhibiteur de MEK, peut prévenir l'apparition d'une résistance dans les cellules de CBNPC avec EGFR muté (Tricker *et al.*, 2015). Le sorafénib n'affecte pas directement les MAPKs, mais il entraîne, après 3-4 jours de traitement, une diminution de l'expression d'EGFR et, par conséquent, l'inhibition des différentes voies régulées par ce récepteur. Ces observations suggèrent que, comparé au tramétinib, le sorafénib pourrait exercer un effet plus large sur le phénotype de ces cellules et interférer avec différents processus cellulaires, tels que la prolifération (MAPK, STAT), la survie (PI3K), le métabolisme (PI3K/mTOR) et l'invasion (Src, STAT).

D'autres approches alternatives consistant en une inhibition directe du récepteur EGFR ont été développées pour contrecarrer la résistance des CBNPC. Il a ainsi été montré que la résistance à l'osimertinib causée par la triple mutation del19/T790M/C797S peut être contournée *in vitro* et *in vivo* en associant un anticorps monoclonal anti-EGFR, tel que le cetuximab (Erbivud, ImClone) ou le panitumumab (Vectibix, Amgen), avec l'inhibiteur d'EGFR et de ALK brigatinib (Alunbrig, Takeda Pharmaceutical Company Limited), approuvé en 2017 pour le traitement de CBNPC présentant des altérations de ALK (Uchibori *et al.*, 2017). En combinaison avec le cetuximab, un nouvel inhibiteur allostérique de l'EGFR, le EAI045, est également capable d'empêcher la croissance de tumeurs présentant les mutations EGFR-L858R/T790M et EGFR-L858R/T790M/C797S dans un modèle murin de cancer bronchique (Jia *et al.*, 2016). Contrairement à ces approches, qui visent à inhiber directement l'activité du récepteur, le sorafénib entraîne une diminution de l'expression de l'EGFR, suivie par une inhibition des voies en aval. Ce composé pourrait donc être utilisé pour contrecarrer les effets d'autres mutations d'EGFR, et notamment celles acquises après traitement avec un inhibiteur d'EGFR de 1^{ère} génération, telles que L747S, D761Y ou T854A. De plus, des études préliminaires semblent indiquer que le sorafénib, en combinaison avec le gefitinib, serait également capable de prévenir l'enrichissement de cellules contenant d'autres mécanismes de résistance, tels que la mutation KRAS-G12D ou la transition EML4-ALK (données non illustrées). Des expériences sont en cours pour confirmer ces résultats et caractériser le mécanisme responsable de ces effets.

2.3. Le sorafenib pour le traitement du CBNPC avec EGFR muté

Le sorafenib est utilisé en clinique depuis plus de 10 ans pour le traitement de différents types de cancers. Les effets secondaires générés par cette molécule varient en fonction des patients, et comprennent généralement des diarrhées, des réactions cutanées, une alopecie, l'infection, un «syndrome pieds-mains» (éruption cutanée et douleur sur les paumes et la plante des pieds) et la fatigue (EMA, 2014). Le sorafenib a été utilisé dans des essais cliniques de phase II pour étudier ses effets en monothérapie en 1^{ère} ou 2^{ème} ligne pour le traitement des CBNPCs au stade avancé (Blumenschein *et al.*, 2009 ; Dy *et al.*, 2010 ; Wakelee *et al.*, 2012). Ces études indiquent que le sorafenib permet d'augmenter la survie globale des patients de quelques mois, avec une meilleure survie sans progression, ainsi qu'un meilleur contrôle de la progression tumorale (Blumenschein *et al.*, 2009 ; Wakelee *et al.*, 2012). Ces résultats ont permis de justifier la réalisation d'un essai clinique de phase III, appelé MISSION (Monotherapy administration of Sorafenib in patients with non-small cell lung cancer). Cet essai a inclus 703 patients atteints de CBNPC non squameux de stade avancé, qui ont été randomisés pour recevoir en monothérapie le sorafenib ou un placebo, en troisième ou en quatrième ligne de traitement. Les résultats publiés en 2015 indiquent que, malgré une augmentation significative de la survie sans progression, le traitement par le sorafenib n'améliore pas de façon significative la survie globale des patients. En revanche, lorsque l'on tient compte du profil génétique des tumeurs, on observe que, en présence de mutations activatrices d'EGFR, les patients traités par le sorafenib affichent une survie globale et une survie sans progression significativement supérieures à celles des patients ayant reçu le placebo (respectivement 13.9 mois par rapport à 6.5 mois, $p=0.002$, et 2.7 mois par rapport à 1.4 mois, $p<0.001$) (Paz-Ares *et al.*, 2015). Ces données sont en accord avec nos résultats et suggèrent que le sorafenib pourrait représenter un outil thérapeutique pour le traitement des CBNPCs avec des mutations activatrices d'EGFR.

Dans un essai clinique de phase II paru en 2016, les chercheurs ont évalué la toxicité de la combinaison du sorafenib avec un inhibiteur d'EGFR de 1^{ère} génération, erlotinib, dans des CBNPCs de stade avancé. L'association de ces deux molécules était généralement bien tolérée, alors qu'environ un tiers des patients a dû interrompre momentanément le protocole thérapeutique et a nécessité par la suite une diminution des doses administrées (Lim *et al.*, 2016). Ces patients présentaient des taux d'enzymes hépatiques, des diarrhées ou des éruptions cutanées bien plus importants que ceux observés en monothérapie (Kim *et al.*, 2011).

D'après les résultats observés dans ces essais cliniques et dans nos expériences de xénogreffe menées en collaboration avec le Pr S. Aaronson, différents protocoles thérapeutiques peuvent être envisagés pour l'utilisation du sorafenib chez les patients de CBNPC présentant une mutation d'EGFR. Ce composé pourrait être administré en monothérapie après le développement d'une résistance envers l'osimertinib, afin de ralentir la progression tumorale. Le sorafenib pourrait également être utilisé en alternance avec un inhibiteur d'EGFR, dans le but de retarder l'apparition d'une résistance. Ce type de traitement séquentiel risquerait en revanche de favoriser l'émergence de clones présentant de nouvelles mutations en augmentant l'hétérogénéité intratumorale, ce qui pourrait avoir des conséquences néfastes pour le patient. Il a été établi par différents groupes que des clones résistants sont souvent présents avant le début du traitement (Turke *et al.*, 2010 ; Hata *et al.*, 2016). En utilisant un modèle de xénogreffe basé sur l'approche CRISPR-barcoding, nous avons montré que, contrairement au sorafenib seul, qui provoque un ralentissement de la progression tumorale, la combinaison avec l'osimertinib entraîne une diminution du volume des tumeurs, qui est maintenue sur plusieurs semaines. Ainsi, en ciblant simultanément les cellules sensibles et celles résistantes aux inhibiteurs d'EGFR, l'association de ces deux composés devrait permettre de réduire le risque de rechute. Si le niveau de toxicité acceptable observé chez la souris est confirmé par d'autres essais *in vivo*, la combinaison osimertinib-sorafenib pourra être testée chez les patients en tant que nouvelle stratégie thérapeutique pour prévenir ou ralentir la résistance des CBNPCs présentant des mutations activatrices d'EGFR.

CONCLUSION ET PERSPECTIVES

Les travaux présentés dans ce manuscrit ont tout d'abord permis la mise en place d'une nouvelle stratégie d'étude des altérations génétiques des cellules tumorales, basée sur le système CRISPR/Cas9. Cette approche consiste en l'identification et le suivi des cellules modifiées au sein d'une population globale grâce au système d'édition des gènes et à la détection quantitative de l'ADN édité. À l'aide de plusieurs modèles, nous avons montré que cette approche, appelée CRISPR-barcoding, est applicable à différents types d'études fonctionnelles et peut être employée pour déterminer les effets d'une mutation d'intérêt sur différents processus biologiques, tels que la prolifération, la migration ou encore l'invasion cellulaire.

Dans la seconde partie de ce travail, en utilisant CRISPR-barcoding nous avons révélé la capacité du sorafenib à empêcher l'amplification des cellules de CBNPC résistantes aux inhibiteurs d'EGFR. Nous avons montré que le mécanisme d'action du sorafenib implique une inhibition précoce de la phosphorylation de STAT3, suivie d'une diminution de l'expression de l'EGFR dans des phases plus tardives du traitement. Nos expériences n'ont pas permis d'identifier clairement la cible moléculaire de ce composé dans les cellules de CBNPC, suggérant que le sorafenib pourrait agir en modulant l'activité de différents effecteurs. Afin de tester cette hypothèse, nous étudierons les effets d'une combinaison de sunitinib ou regorafenib avec des inhibiteurs de RAF ou des activateurs de SHP-1. Le sorafenib pourrait également inhiber une cible encore non caractérisée à ce jour. Au vu de sa capacité à activer SHP-1 (Tai *et al.*, 2011), nous analyserons les effets de ce composé sur l'activité d'autres membres de la famille des protéine tyrosine phosphatases. En parallèle, afin de mieux comprendre les voies de signalisation régulées par ce composé, nous réaliserons une analyse transcriptomique par RNA-seq à partir de cellules de CBNPC traitées ou non avec le sorafenib. Dans la continuité de ce travail, nous allons également étudier les mécanismes impliqués dans la diminution de l'expression d'EGFR par cet inhibiteur multikinase. Des expériences effectuées au laboratoire indiquent que le sorafenib n'affecte pas le taux d'ARNm de ce récepteur (données non présentées), suggérant une régulation post-traductionnelle. Pour confirmer ces résultats, nous évaluerons la capacité de ce composé à inhiber l'expression d'une protéine EGFR exogène taguée, et nous analyserons les effets sur la stabilité du récepteur en présence ou en absence d'inhibiteurs de la synthèse protéique et du protéasome.

En collaboration avec le Pr S. Aaronson, nous avons montré que la combinaison osimertinib-sorafenib peut prévenir l'émergence d'une résistance *in vivo*. Afin de tester les effets du sorafenib dans un modèle plus proche de la réalité clinique, nous réaliserons des expériences utilisant des xénogreffes issues de patients (*patient-derived xenografts*) atteints de CBNPC avec EGFR muté. Après traitement, les tumeurs seront collectées afin d'étudier par immunohistochimie l'expression de différents effecteurs d'EGFR, ainsi que des marqueurs de la prolifération et de l'apoptose. A terme, ces travaux pourraient se traduire par la mise en place d'un essai clinique.

Eu égard au mécanisme d'action du sorafenib, ce composé pourrait également être utilisé pour d'autres types de cancer dont la résistance peut être provoquée par des mutations de l'EGFR. Le cetuximab et le panitumumab, anticorps monoclonaux dirigés contre ce récepteur, sont utilisés en clinique pour la prise en charge des patients atteints de cancers colorectaux métastatiques. Cependant, malgré une réponse initiale, ces patients peuvent développer une résistance provoquée par l'apparition de mutations dans le domaine extracellulaire d'EGFR, qui empêchent la fixation de ces anticorps (Montagut *et al.*, 2012). En collaboration avec le groupe du Dr A. Bardelli (Candiolo Cancer Institute-IRCCS, Université de Turin, Italie), nous utiliserons notre approche CRISPR-barcoding pour évaluer la capacité du sorafenib à prévenir la résistance des cellules de cancer colorectal au cetuximab.

Comme il est souvent le cas pour les thérapies ciblées, il est probable que les cellules de CBNPC soient capables de développer une résistance à la combinaison osimertinib-sorafenib. Nous réaliserons des expériences utilisant les cellules de CBNPC contenant les codes-barres CRISPR hautement complexes (Guernet *et al.*, 2016) afin de vérifier si des clones résistants peuvent être présents avant le début du traitement. Grâce à une dilution limite des cellules ainsi obtenues qui seraient insensibles à la combinaison osimertinib-sorafenib, il sera également possible d'isoler et amplifier des populations cellulaires homogènes pour étudier les mécanismes responsables de la résistance.

En clinique, l'émergence de clones capables d'échapper à la pression sélective représentée par la thérapie ciblée est probablement favorisée par des traitements anti-cancéreux délivrés à fortes doses et de manière ininterrompue. De nouvelles approches ont été proposées consistant à ajuster les doses et/ou la durée du traitement selon l'état de progression de la tumeur (Kaiser, 2017). En tirant profit de la capacité à tracer par CRISPR-barcoding les clones insensibles, nous pourrions ainsi optimiser le protocole d'administration du sorafenib et de l'osimertinib, afin de

prévenir l'émergence d'une résistance sur le long terme, tout en limitant les effets toxiques potentiels.

REFERENCES BIBLIOGRAPHIQUES

- Adjei, A.A., Mauer, A., Bruzek, L., Marks, R.S., Hillman, S., Geyer, S., Hanson, L.J., Wright, J.J., Erlichman, C., Kaufmann, S.H., *et al.* (2003). Phase II study of the farnesyl transferase inhibitor R115777 in patients with advanced non-small-cell lung cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* *21*, 1760–1766.
- Alamgeer, M., Ganju, V., and Watkins, D.N. (2013). Novel therapeutic targets in non-small cell lung cancer. *Curr. Opin. Pharmacol.* *13*, 394–401.
- Alfieri, R.R., Galetti, M., Tramonti, S., Andreoli, R., Mozzoni, P., Cavazzoni, A., Bonelli, M., Fumarola, C., La Monica, S., Galvani, E., *et al.* (2011). Metabolism of the EGFR tyrosin kinase inhibitor gefitinib by cytochrome P450 1A1 enzyme in EGFR-wild type non small cell lung cancer cell lines. *Mol. Cancer* *10*, 143.
- Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* *100*, 3983–3988.
- Altrock, P.M., Liu, L.L., and Michor, F. (2015). The mathematics of cancer: integrating quantitative models. *Nat. Rev. Cancer* *15*, 730–745.
- Anders, C., Niewoehner, O., Duerst, A., and Jinek, M. (2014). Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* *513*, 569–573.
- Arcila, M.E., Oxnard, G.R., Nafa, K., Riely, G.J., Solomon, S.B., Zakowski, M.F., Kris, M.G., Pao, W., Miller, V.A., and Ladanyi, M. (2011). Rebiopsy of lung cancer patients with acquired resistance to EGFR inhibitors and enhanced detection of the T790M mutation using a locked nucleic acid-based assay. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *17*, 1169–1180.
- Balak, M.N., Gong, Y., Riely, G.J., Somwar, R., Li, A.R., Zakowski, M.F., Chiang, A., Yang, G., Ouerfelli, O., Kris, M.G., *et al.* (2006). Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor-mutant lung adenocarcinomas with acquired resistance to kinase inhibitors. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *12*, 6494–6501.
- Barrangou, R., and Marraffini, L.A. (2014). CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol. Cell* *54*, 234–244.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* *315*, 1709–1712.
- Bashashati, A., Ha, G., Tone, A., Ding, J., Prentice, L.M., Roth, A., Rosner, J., Shumansky, K., Kalloger, S., Senz, J., *et al.* (2013). Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling. *J. Pathol.* *231*, 21–34.
- Bassett, A.R., Tibbit, C., Ponting, C.P., and Liu, J.-L. (2013). Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Rep.* *4*, 220–228.
- Bean, J., Riely, G.J., Balak, M., Marks, J.L., Ladanyi, M., Miller, V.A., and Pao, W. (2008). Acquired resistance to epidermal growth factor receptor kinase inhibitors associated with a novel T854A mutation in a patient with EGFR-mutant lung adenocarcinoma. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *14*, 7519–7525.

- Bhang, H.C., Ruddy, D.A., Krishnamurthy Radhakrishna, V., Caushi, J.X., Zhao, R., Hims, M.M., Singh, A.P., Kao, I., Rakiec, D., Shaw, P., *et al.* (2015). Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. *Nat. Med.* *21*, 440–448.
- Bhattacharya, S., Socinski, M.A., and Burns, T.F. (2015). KRAS mutant lung cancer: progress thus far on an elusive therapeutic target. *Clin. Transl. Med.* *4*, 35.
- Bibikova, M., Golic, M., Golic, K.G., and Carroll, D. (2002). Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics* *161*, 1169–1175.
- Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L.A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* *41*, 7429–7437.
- Bitinaite, J., Wah, D.A., Aggarwal, A.K., and Schildkraut, I. (1998). FokI dimerization is required for DNA cleavage. *Proc. Natl. Acad. Sci. U. S. A.* *95*, 10570–10575.
- Bivona, T.G., Hieronymus, H., Parker, J., Chang, K., Taron, M., Rosell, R., Moonsamy, P., Dahlman, K., Miller, V.A., Costa, C., *et al.* (2011). FAS and NF- κ B signalling modulate dependence of lung cancers on mutant EGFR. *Nature* *471*, 523–526.
- Blanpain, C., and Simons, B.D. (2013). Unravelling stem cell dynamics by lineage tracing. *Nat. Rev. Mol. Cell Biol.* *14*, 489–502.
- 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.
- Bloom, H.J., and Richardson, W.W. (1957). Histological grading and prognosis in breast cancer ; a study of 1409 cases of which 359 have been followed for 15 years. *Br. J. Cancer* *11*, 359–377.
- Blumenschein, G.R., Gatzemeier, U., Fossella, F., Stewart, D.J., Cupit, L., Cihon, F., O’Leary, J., and Reck, M. (2009). Phase II, multicenter, uncontrolled trial of single-agent sorafenib in patients with relapsed or refractory, advanced non-small-cell lung cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* *27*, 4274–4280.
- Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* *326*, 1509–1512.
- Bolderson, E., Richard, D.J., Zhou, B.-B.S., and Khanna, K.K. (2009). Recent advances in cancer therapy targeting proteins involved in DNA double-strand break repair. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *15*, 6314–6320.
- Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiol. Read. Engl.* *151*, 2551–2561.
- Brose, M.S., Volpe, P., Feldman, M., Kumar, M., Rishi, I., Gerrero, R., Einhorn, E., Herlyn, M., Minna, J., Nicholson, A., *et al.* (2002). BRAF and RAS mutations in human lung cancer and melanoma. *Cancer Res.* *62*, 6997–7000.

- Brouns, S.J.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J.H., Snijders, A.P.L., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960–964.
- de Bruin, E.C., McGranahan, N., Mitter, R., Salm, M., Wedge, D.C., Yates, L., Jamal-Hanjani, M., Shafi, S., Murugaesu, N., Rowan, A.J., *et al.* (2014). Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* 346, 251–256.
- Camidge, D.R., Ou, S.-H.I., Shapiro, G., Otterson, G.A., Villaruz, L.C., Villalona-Calero, M.A., Iafrate, A.J., Varella-Garcia, M., Dacic, S., Cardarella, S., *et al.* (2014). Efficacy and safety of crizotinib in patients with advanced c-MET-amplified non-small cell lung cancer (NSCLC). 2014 ASCO Annu. Meet. *J Clin Oncol* 32:5s, 2014 (suppl ; abstr 8001).
- Campbell, P.J., Yachida, S., Mudie, L.J., Stephens, P.J., Pleasance, E.D., Stebbings, L.A., Morsberger, L.A., Latimer, C., McLaren, S., Lin, M.-L., *et al.* (2010). The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* 467, 1109–1113.
- Capecchi, M.R. (1989). Altering the genome by homologous recombination. *Science* 244, 1288–1292.
- Carreira, S., Romanel, A., Goodall, J., Grist, E., Ferraldeschi, R., Miranda, S., Prandi, D., Lorente, D., Frenel, J.-S., Pezaro, C., *et al.* (2014). Tumor clone dynamics in lethal prostate cancer. *Sci. Transl. Med.* 6, 254ra125.
- Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J.-W., and Xi, J.J. (2013). Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res.* 23, 465–472.
- Chaumeil, J., Micsinai, M., and Skok, J.A. (2013). Combined immunofluorescence and DNA FISH on 3D-preserved interphase nuclei to study changes in 3D nuclear organization. *J. Vis. Exp. JoVE* e50087.
- Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S., *et al.* (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155, 1479–1491.
- Chen, K.-F., Tai, W.-T., Huang, J.-W., Hsu, C.-Y., Chen, W.-L., Cheng, A.-L., Chen, P.-J., and Shiau, C.-W. (2011). Sorafenib derivatives induce apoptosis through inhibition of STAT3 independent of Raf. *Eur. J. Med. Chem.* 46, 2845–2851.
- Chen, S., Sanjana, N.E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D.A., Song, J., Pan, J.Q., Weissleder, R., *et al.* (2015). Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell* 160, 1246–1260.
- Chen, Z., Fillmore, C.M., Hammerman, P.S., Kim, C.F., and Wong, K.-K. (2014). Non-small-cell lung cancers: a heterogeneous set of diseases. *Nat. Rev. Cancer* 14, 535–546.
- Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., and Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* 23, 1163–1171.

- Chevalier, B.S., Kortemme, T., Chadsey, M.S., Baker, D., Monnat, R.J., and Stoddard, B.L. (2002). Design, activity, and structure of a highly specific artificial endonuclease. *Mol. Cell* *10*, 895–905.
- Cho, S.W., Kim, S., Kim, J.M., and Kim, J.-S. (2013a). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* *31*, 230–232.
- Cho, S.W., Lee, J., Carroll, D., Kim, J.-S., and Lee, J. (2013b). Heritable gene knockout in *Caenorhabditis elegans* by direct injection of Cas9-sgRNA ribonucleoproteins. *Genetics* *195*, 1177–1180.
- Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J.-S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* *24*, 132–141.
- Choi, P.S., and Meyerson, M. (2014). Targeted genomic rearrangements using CRISPR/Cas technology. *Nat. Commun.* *5*, 3728.
- Chong, C.R., and Jänne, P.A. (2013). The quest to overcome resistance to EGFR-targeted therapies in cancer. *Nat. Med.* *19*, 1389–1400.
- Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J., and Voytas, D.F. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* *186*, 757–761.
- Chu, V.T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., and Kühn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* *33*, 543–548.
- Chylinski, K., Makarova, K.S., Charpentier, E., and Koonin, E.V. (2014). Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Res.* *42*, 6091–6105.
- Citri, A., and Yarden, Y. (2006). EGF-ERBB signalling: towards the systems level. *Nat. Rev. Mol. Cell Biol.* *7*, 505–516.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., *et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* *339*, 819–823.
- Conticello, S.G. (2008). The AID/APOBEC family of nucleic acid mutators. *Genome Biol.* *9*, 229.
- Cortot, A.B., and Jänne, P.A. (2014). Molecular mechanisms of resistance in epidermal growth factor receptor-mutant lung adenocarcinomas. *Eur. Respir. Rev. Off. J. Eur. Respir. Soc.* *23*, 356–366.
- Cortot, A.B., Kherrouche, Z., Descarpentries, C., Wislez, M., Baldacci, S., Furlan, A., and Tulasne, D. (2017). Exon 14 Deleted MET Receptor as a New Biomarker and Target in Cancers. *J. Natl. Cancer Inst.* *109*.
- Costa, D.B., Schumer, S.T., Tenen, D.G., and Kobayashi, S. (2008). Differential responses to erlotinib in epidermal growth factor receptor (EGFR)-mutated lung cancers with acquired

resistance to gefitinib carrying the L747S or T790M secondary mutations. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* *26*, 1182-1184 ; author reply 1184-1186.

Cross, D.A.E., Ashton, S.E., Ghiorghiu, S., Eberlein, C., Nebhan, C.A., Spitzler, P.J., Orme, J.P., Finlay, M.R.V., Ward, R.A., Mellor, M.J., *et al.* (2014). AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. *Cancer Discov.* *4*, 1046–1061.

Cui, J.J., Tran-Dubé, M., Shen, H., Nambu, M., Kung, P.-P., Pairish, M., Jia, L., Meng, J., Funk, L., Botrous, I., *et al.* (2011). Structure based drug design of crizotinib (PF-02341066), a potent and selective dual inhibitor of mesenchymal-epithelial transition factor (c-MET) kinase and anaplastic lymphoma kinase (ALK). *J. Med. Chem.* *54*, 6342–6363.

Dalerba, P., Dylla, S.J., Park, I.-K., Liu, R., Wang, X., Cho, R.W., Hoey, T., Gurney, A., Huang, E.H., Simeone, D.M., *et al.* (2007). Phenotypic characterization of human colorectal cancer stem cells. *Proc. Natl. Acad. Sci. U. S. A.* *104*, 10158–10163.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* *471*, 602–607.

Desai, A., and Adjei, A.A. (2016). FGFR Signaling as a Target for Lung Cancer Therapy. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* *11*, 9–20.

Desjarlais, J.R., and Berg, J.M. (1992). Redesigning the DNA-binding specificity of a zinc finger protein: a data base-guided approach. *Proteins* *12*, 101–104.

Dickinson, D.J., Ward, J.D., Reiner, D.J., and Goldstein, B. (2013). Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. *Nat. Methods* *10*, 1028–1034.

Ding, L., Ley, T.J., Larson, D.E., Miller, C.A., Koboldt, D.C., Welch, J.S., Ritchey, J.K., Young, M.A., Lamprecht, T., McLellan, M.D., *et al.* (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* *481*, 506–510.

Ding, L., Wendl, M.C., McMichael, J.F., and Raphael, B.J. (2014). Expanding the computational toolbox for mining cancer genomes. *Nat. Rev. Genet.* *15*, 556–570.

Dixit, A., and Verkhivker, G.M. (2014). Structure-functional prediction and analysis of cancer mutation effects in protein kinases. *Comput. Math. Methods Med.* *2014*, 653487.

Doyon, Y., McCammon, J.M., Miller, J.C., Faraji, F., Ngo, C., Katibah, G.E., Amora, R., Hocking, T.D., Zhang, L., Rebar, E.J., *et al.* (2008). Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat. Biotechnol.* *26*, 702–708.

Duchemann, B., Friboulet, L., and Besse, B. (2015). Therapeutic management of ALK+ nonsmall cell lung cancer patients. *Eur. Respir. J.* *46*, 230–242.

Dugar, G., Herbig, A., Förstner, K.U., Heidrich, N., Reinhardt, R., Nieselt, K., and Sharma, C.M. (2013). High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. *PLoS Genet.* *9*, e1003495.

- Dürrenberger, F., and Rochaix, J.D. (1991). Chloroplast ribosomal intron of *Chlamydomonas reinhardtii*: in vitro self-splicing, DNA endonuclease activity and in vivo mobility. *EMBO J.* *10*, 3495–3501.
- Dutt, A., Ramos, A.H., Hammerman, P.S., Mermel, C., Cho, J., Sharifnia, T., Chande, A., Tanaka, K.E., Stransky, N., Greulich, H., *et al.* (2011). Inhibitor-sensitive FGFR1 amplification in human non-small cell lung cancer. *PloS One* *6*, e20351.
- Dy, G.K., Hillman, S.L., Rowland, K.M., Molina, J.R., Steen, P.D., Wender, D.B., Nair, S., Mandrekar, S., Schild, S.E., Adjei, A.A., *et al.* (2010). A front-line window of opportunity phase 2 study of sorafenib in patients with advanced nonsmall cell lung cancer: North Central Cancer Treatment Group Study N0326. *Cancer* *116*, 5686–5693.
- Eberlein, C.A., Stetson, D., Markovets, A.A., Al-Kadhimi, K.J., Lai, Z., Fisher, P.R., Meador, C.B., Spitzler, P., Ichihara, E., Ross, S.J., *et al.* (2015). Acquired Resistance to the Mutant-Selective EGFR Inhibitor AZD9291 Is Associated with Increased Dependence on RAS Signaling in Preclinical Models. *Cancer Res.* *75*, 2489–2500.
- Ehrenreiter, K., Kern, F., Velamoor, V., Meissl, K., Galabova-Kovacs, G., Sibia, M., and Baccarini, M. (2009). Raf-1 addiction in Ras-induced skin carcinogenesis. *Cancer Cell* *16*, 149–160.
- Engelman, J.A., Zejnullahu, K., Mitsudomi, T., Song, Y., Hyland, C., Park, J.O., Lindeman, N., Gale, C.-M., Zhao, X., Christensen, J., *et al.* (2007). MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* *316*, 1039–1043.
- Ercan, D., Xu, C., Yanagita, M., Monast, C.S., Pratilas, C.A., Montero, J., Butaney, M., Shimamura, T., Sholl, L., Ivanova, E.V., *et al.* (2012). Reactivation of ERK signaling causes resistance to EGFR kinase inhibitors. *Cancer Discov.* *2*, 934–947.
- Ercan, D., Choi, H.G., Yun, C.-H., Capelletti, M., Xie, T., Eck, M.J., Gray, N.S., and Jänne, P.A. (2015). EGFR Mutations and Resistance to Irreversible Pyrimidine-Based EGFR Inhibitors. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *21*, 3913–3923.
- Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J., and Church, G.M. (2013). Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* *10*, 1116–1121.
- Faivre, S., Demetri, G., Sargent, W., and Raymond, E. (2007). Molecular basis for sunitinib efficacy and future clinical development. *Nat. Rev. Drug Discov.* *6*, 734–745.
- Fearon, E.R., and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell* *61*, 759–767.
- Filosto, S., Khan, E.M., Tognon, E., Becker, C., Ashfaq, M., Ravid, T., and Goldkorn, T. (2011). EGF receptor exposed to oxidative stress acquires abnormal phosphorylation and aberrant activated conformation that impairs canonical dimerization. *PloS One* *6*, e23240.
- Filosto, S., Becker, C.R., and Goldkorn, T. (2012). Cigarette smoke induces aberrant EGF receptor activation that mediates lung cancer development and resistance to tyrosine kinase inhibitors. *Mol. Cancer Ther.* *11*, 795–804.

- Filosto, S., Baston, D.S., Chung, S., Becker, C.R., and Goldkorn, T. (2013). Src mediates cigarette smoke-induced resistance to tyrosine kinase inhibitors in NSCLC cells. *Mol. Cancer Ther.* *12*, 1579–1590.
- Fonfara, I., Le Rhun, A., Chylinski, K., Makarova, K.S., Lécrivain, A.-L., Bzdrenga, J., Koonin, E.V., and Charpentier, E. (2014). Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res.* *42*, 2577–2590.
- Friboulet, L., Li, N., Katayama, R., Lee, C.C., Gainor, J.F., Crystal, A.S., Michellys, P.-Y., Awad, M.M., Yanagitani, N., Kim, S., *et al.* (2014). The ALK inhibitor ceritinib overcomes crizotinib resistance in non-small cell lung cancer. *Cancer Discov.* *4*, 662–673.
- Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiácovo, M.P., Church, G.M., and Calarco, J.A. (2013). Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat. Methods* *10*, 741–743.
- Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* *31*, 822–826.
- Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* *32*, 279–284.
- Fu, Y., Rocha, P.P., Luo, V.M., Raviram, R., Deng, Y., Mazzone, E.O., and Skok, J.A. (2016). CRISPR-dCas9 and sgRNA scaffolds enable dual-colour live imaging of satellite sequences and repeat-enriched individual loci. *Nat. Commun.* *7*, 11707.
- Fukuoka, M., Yano, S., Giaccone, G., Tamura, T., Nakagawa, K., Douillard, J.-Y., Nishiwaki, Y., Vansteenkiste, J., Kudoh, S., Rischin, D., *et al.* (2003). Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer (The IDEAL 1 Trial) [corrected]. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* *21*, 2237–2246.
- Gainor, J.F., Varghese, A.M., Ou, S.-H.I., Kabraji, S., Awad, M.M., Katayama, R., Pawlak, A., Mino-Kenudson, M., Yeap, B.Y., Riely, G.J., *et al.* (2013). ALK rearrangements are mutually exclusive with mutations in EGFR or KRAS: an analysis of 1,683 patients with non-small cell lung cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *19*, 4273–4281.
- Gantz VM, Jasinskiene N, Tatarenkova O, Fazekas A, Macias VM, Bier E, James AA. (2015) Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. *Proc Natl Acad Sci U S A*, *112*, 49, 6736-6743
- Gao, H., Smith, J., Yang, M., Jones, S., Djukanovic, V., Nicholson, M.G., West, A., Bidney, D., Falco, S.C., Jantz, D., *et al.* (2010). Heritable targeted mutagenesis in maize using a designed endonuclease. *Plant J. Cell Mol. Biol.* *61*, 176–187.
- Garneau, J.E., Dupuis, M.-È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* *468*, 67–71.
- Garon, E.B., Finn, R.S., Hosmer, W., Dering, J., Ginther, C., Adhami, S., Kamranpour, N., Pitts, S., Desai, A., Elashoff, D., *et al.* (2010). Identification of common predictive markers of

- in vitro response to the Mek inhibitor selumetinib (AZD6244 ; ARRY-142886) in human breast cancer and non-small cell lung cancer cell lines. *Mol. Cancer Ther.* 9, 1985–1994.
- Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 109, E2579-2586.
- Gawad, C., Koh, W., and Quake, S.R. (2016). Single-cell genome sequencing: current state of the science. *Nat. Rev. Genet.* 17, 175–188.
- George, J., Lim, J.S., Jang, S.J., Cun, Y., Ozretić, L., Kong, G., Leenders, F., Lu, X., Fernández-Cuesta, L., Bosco, G., *et al.* (2015). Comprehensive genomic profiles of small cell lung cancer. *Nature* 524, 47–53.
- Gerlinger, M., Horswell, S., Larkin, J., Rowan, A.J., Salm, M.P., Varela, I., Fisher, R., McGranahan, N., Matthews, N., Santos, C.R., *et al.* (2014). Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat. Genet.* 46, 225–233.
- Geurts, A.M., Cost, G.J., Freyvert, Y., Zeitler, B., Miller, J.C., Choi, V.M., Jenkins, S.S., Wood, A., Cui, X., Meng, X., *et al.* (2009). Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* 325, 433.
- Gherardi, E., Birchmeier, W., Birchmeier, C., and Vande Woude, G. (2012). Targeting MET in cancer: rationale and progress. *Nat. Rev. Cancer* 12, 89–103.
- Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., *et al.* (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451.
- Gillies, R.J., Verduzco, D., and Gatenby, R.A. (2012). Evolutionary dynamics of carcinogenesis and why targeted therapy does not work. *Nat. Rev. Cancer* 12, 487–493.
- Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J., and O'Connor-Giles, K.M. (2013). Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* 194, 1029–1035.
- Guernet, A., Mungamuri, S.K., Cartier, D., Sachidanandam, R., Jayaprakash, A., Adriouch, S., Vezain, M., Charbonnier, F., Rohkin, G., Coutant, S., *et al.* (2016). CRISPR-Barcoding for Intratumor Genetic Heterogeneity Modeling and Functional Analysis of Oncogenic Driver Mutations. *Mol. Cell* 63, 526–538.
- Guilinger, J.P., Thompson, D.B., and Liu, D.R. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* 32, 577–582.
- Gürlevik, E., Schache, P., Goetz, A., Kloos, A., Woller, N., Armbrecht, N., Manns, M.P., Kubicka, S., and Kühnel, F. (2013). Meganuclease-mediated virus self-cleavage facilitates tumor-specific virus replication. *Mol. Ther. J. Am. Soc. Gene Ther.* 21, 1738–1748.
- Gysin, S., Salt, M., Young, A., and McCormick, F. (2011). Therapeutic strategies for targeting ras proteins. *Genes Cancer* 2, 359–372.

- Haffner, M.C., Mosbrugger, T., Esopi, D.M., Fedor, H., Heaphy, C.M., Walker, D.A., Adejola, N., Gürel, M., Hicks, J., Meeker, A.K., *et al.* (2013). Tracking the clonal origin of lethal prostate cancer. *J. Clin. Invest.* *123*, 4918–4922.
- Hai, T., Teng, F., Guo, R., Li, W., and Zhou, Q. (2014). One-step generation of knockout pigs by zygote injection of CRISPR/Cas system. *Cell Res.* *24*, 372–375.
- Ham, J.S., Kim, S., Kim, H.K., Byeon, S., Sun, J.-M., Lee, S.-H., Ahn, J.S., Park, K., Choi, Y.-L., Han, J., *et al.* (2016). Two Cases of Small Cell Lung Cancer Transformation from EGFR Mutant Adenocarcinoma During AZD9291 Treatment. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* *11*, e1-4.
- Hammond A, Galizi R, Kyrou K, Simoni A, Siniscalchi C, Katsanos D, Gribble M, Baker D, Marois E, Russell S, Burt A, Windbichler N, Crisanti A, Nolan T. (2016). A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. *Nat Biotechnol* *34*, 78-83
- Han, J.-Y., Park, K., Kim, S.-W., Lee, D.H., Kim, H.Y., Kim, H.T., Ahn, M.J., Yun, T., Ahn, J.S., Suh, C., *et al.* (2012). First-SIGNAL: first-line single-agent irressa versus gemcitabine and cisplatin trial in never-smokers with adenocarcinoma of the lung. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* *30*, 1122–1128.
- Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* *144*, 646–674.
- Harper, K.L., Sosa, M.S., Entenberg, D., Hosseini, H., Cheung, J.F., Nobre, R., Avivar-Valderas, A., Nagi, C., Girnius, N., Davis, R.J., *et al.* (2016). Mechanism of early dissemination and metastasis in Her2(+) mammary cancer. *Nature*.
- Hata, A.N., Niederst, M.J., Archibald, H.L., Gomez-Caraballo, M., Siddiqui, F.M., Mulvey, H.E., Maruvka, Y.E., Ji, F., Bhang, H.C., Krishnamurthy Radhakrishna, V., *et al.* (2016). Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. *Nat. Med.* *22*, 262–269.
- Heigwer, F., Kerr, G., and Boutros, M. (2014). E-CRISP: fast CRISPR target site identification. *Nat. Methods* *11*, 122–123.
- Heinrich, M.C., Maki, R.G., Corless, C.L., Antonescu, C.R., Harlow, A., Griffith, D., Town, A., McKinley, A., Ou, W.-B., Fletcher, J.A., *et al.* (2008). Primary and secondary kinase genotypes correlate with the biological and clinical activity of sunitinib in imatinib-resistant gastrointestinal stromal tumor. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* *26*, 5352–5359.
- Heler, R., Samai, P., Modell, J.W., Weiner, C., Goldberg, G.W., Bikard, D., and Marraffini, L.A. (2015). Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* *519*, 199–202.
- Heppner, G.H., and Miller, B.E. (1983). Tumor heterogeneity: biological implications and therapeutic consequences. *Cancer Metastasis Rev.* *2*, 5–23.
- Hilton, I.B., D’Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* *33*, 510–517.

- Honda, A., Hirose, M., Sankai, T., Yasmin, L., Yuzawa, K., Honsho, K., Izu, H., Iguchi, A., Ikawa, M., and Ogura, A. (2015). Single-step generation of rabbits carrying a targeted allele of the tyrosinase gene using CRISPR/Cas9. *Exp. Anim.* *64*, 31–37.
- Hosseini, H., Obradović, M.M.S., Hoffmann, M., Harper, K.L., Sosa, M.S., Werner-Klein, M., Nanduri, L.K., Werno, C., Ehrl, C., Maneck, M., *et al.* (2016). Early dissemination seeds metastasis in breast cancer. *Nature*.
- Hou, Z., Zhang, Y., Propson, N.E., Howden, S.E., Chu, L.-F., Sontheimer, E.J., and Thomson, J.A. (2013). Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc. Natl. Acad. Sci. U. S. A.* *110*, 15644–15649.
- Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., *et al.* (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* *31*, 827–832.
- Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* *157*, 1262–1278.
- Hu, X., Chang, N., Wang, X., Zhou, F., Zhou, X., Zhu, X., and Xiong, J.-W. (2013). Heritable gene-targeting with gRNA/Cas9 in rats. *Cell Res.* *23*, 1322–1325.
- Hunter, J.C., Gurbani, D., Ficarro, S.B., Carrasco, M.A., Lim, S.M., Choi, H.G., Xie, T., Marto, J.A., Chen, Z., Gray, N.S., *et al.* (2014). In situ selectivity profiling and crystal structure of SML-8-73-1, an active site inhibitor of oncogenic K-Ras G12C. *Proc. Natl. Acad. Sci. U. S. A.* *111*, 8895–8900.
- Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.-R.J., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* *31*, 227–229.
- 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.
- Infarinato, N.R., Park, J.H., Krytska, K., Ryles, H.T., Sano, R., Szigety, K.M., Li, Y., Zou, H.Y., Lee, N.V., Smeal, T., *et al.* (2016). The ALK/ROS1 Inhibitor PF-06463922 Overcomes Primary Resistance to Crizotinib in ALK-Driven Neuroblastoma. *Cancer Discov.* *6*, 96–107.
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* *169*, 5429–5433.
- Iwama, E., Goto, Y., Murakami, H., Harada, T., Tsumura, S., Sakashita, H., Mori, Y., Nakagaki, N., Fujita, Y., Seike, M., *et al.* (2017). Alectinib for patients with ALK rearrangement-positive non-small cell lung cancer and a poor performance status (Lung Oncology Group in Kyushu 1401). *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer*.
- Iyer, V., Shen, B., Zhang, W., Hodgkins, A., Keane, T., Huang, X., and Skarnes, W.C. (2015). Off-target mutations are rare in Cas9-modified mice. *Nat. Methods* *12*, 479.

- Iyevleva, A.G., Raskin, G.A., Tiurin, V.I., Sokolenko, A.P., Mitiushkina, N.V., Aleksakhina, S.N., Garifullina, A.R., Strelkova, T.N., Merkulov, V.O., Ivantsov, A.O., *et al.* (2015). Novel ALK fusion partners in lung cancer. *Cancer Lett.* *362*, 116–121.
- Jänne, P.A., Yang, J.C.-H., Kim, D.-W., Planchard, D., Ohe, Y., Ramalingam, S.S., Ahn, M.-J., Kim, S.-W., Su, W.-C., Horn, L., *et al.* (2015). AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. *N. Engl. J. Med.* *372*, 1689–1699.
- Jänne, P.A., van den Heuvel, M.M., Barlesi, F., Cobo, M., Mazieres, J., Crinò, L., Orlov, S., Blackhall, F., Wolf, J., Garrido, P., *et al.* (2017). Selumetinib Plus Docetaxel Compared With Docetaxel Alone and Progression-Free Survival in Patients With KRAS-Mutant Advanced Non-Small Cell Lung Cancer: The SELECT-1 Randomized Clinical Trial. *JAMA* *317*, 1844–1853.
- Jansen, R., Embden, J.D.A. van, Gaastra, W., and Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* *43*, 1565–1575.
- 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.
- Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., and Weeks, D.P. (2013). Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res.* *41*, e188.
- Jin, G., Kim, M.J., Jeon, H.-S., Choi, J.E., Kim, D.S., Lee, E.B., Cha, S.I., Yoon, G.S., Kim, C.H., Jung, T.H., *et al.* (2010). PTEN mutations and relationship to EGFR, ERBB2, KRAS, and TP53 mutations in non-small cell lung cancers. *Lung Cancer Amst. Neth.* *69*, 279–283.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* *337*, 816–821.
- Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., *et al.* (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* *343*, 1247997.
- Kaiser, J. (2017). When less is more. *Science* *355*, 1144–1146.
- Kamps, R., Brandão, R.D., Bosch, B.J. van den, Paulussen, A.D.C., Xanthoulea, S., Blok, M.J., and Romano, A. (2017). Next-Generation Sequencing in Oncology: Genetic Diagnosis, Risk Prediction and Cancer Classification. *Int. J. Mol. Sci.* *18*.
- Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., *et al.* (2013). Mutational landscape and significance across 12 major cancer types. *Nature* *502*, 333–339.
- Karreth, F.A., Frese, K.K., DeNicola, G.M., Baccarini, 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.

- Katayama, R., Shaw, A.T., Khan, T.M., Mino-Kenudson, M., Solomon, B.J., Halmos, B., Jessop, N.A., Wain, J.C., Yeo, A.T., Benes, C., *et al.* (2012). Mechanisms of acquired crizotinib resistance in ALK-rearranged lung Cancers. *Sci. Transl. Med.* 4, 120ra17.
- Kawano, O., Sasaki, H., Endo, K., Suzuki, E., Haneda, H., Yukiue, H., Kobayashi, Y., Yano, M., and Fujii, Y. (2006). PIK3CA mutation status in Japanese lung cancer patients. *Lung Cancer Amst. Neth.* 54, 209–215.
- Kearns, N.A., Pham, H., Tabak, B., Genga, R.M., Silverstein, N.J., Garber, M., and Maehr, R. (2015). Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat. Methods* 12, 401–403.
- Kim, J.-S. (2016). Genome editing comes of age. *Nat. Protoc.* 11, 1573–1578.
- Kim, Y.G., and Chandrasegaran, S. (1994). Chimeric restriction endonuclease. *Proc. Natl. Acad. Sci. U. S. A.* 91, 883–887.
- Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H.R., Hwang, J., Kim, J.-I., and Kim, J.-S. (2015). Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods* 12, 237–243, 1 p following 243.
- Kim, E.S., Herbst, R.S., Wistuba, I.I., Lee, J.J., Blumenschein, G.R., Tsao, A., Stewart, D.J., Hicks, M.E., Erasmus, J., Gupta, S., *et al.* (2011). The BATTLE trial: personalizing therapy for lung cancer. *Cancer Discov.* 1, 44–53.
- Kim, S., Kim, D., Cho, S.W., Kim, J., and Kim, J.-S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 24, 1012–1019.
- Kim, Y.G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U. S. A.* 93, 1156–1160.
- Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V.V., Nguyen, N.T., Zheng, Z., Gonzales, A.P.W., Li, Z., Peterson, R.T., Yeh, J.-R.J., *et al.* (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523, 481–485.
- Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., and Joung, J.K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529, 490–495.
- Kobayashi, S., Boggon, T.J., Dayaram, T., Jänne, P.A., Kocher, O., Meyerson, M., Johnson, B.E., Eck, M.J., Tenen, D.G., and Halmos, B. (2005). EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* 352, 786–792.
- Kohno, T., Takahashi, M., Manda, R., and Yokota, J. (1998). Inactivation of the PTEN/MMAC1/TEP1 gene in human lung cancers. *Genes. Chromosomes Cancer* 22, 152–156.
- Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424.

- Konermann, S., Brigham, M.D., Trevino, A.E., Hsu, P.D., Heidenreich, M., Cong, L., Platt, R.J., Scott, D.A., Church, G.M., and Zhang, F. (2013). Optical control of mammalian endogenous transcription and epigenetic states. *Nature* 500, 472–476.
- Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., *et al.* (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517, 583–588.
- Kong-Beltran, M., Seshagiri, S., Zha, J., Zhu, W., Bhawe, K., Mendoza, N., Holcomb, T., Pujara, K., Stinson, J., Fu, L., *et al.* (2006). Somatic mutations lead to an oncogenic deletion of met in lung cancer. *Cancer Res.* 66, 283–289.
- Korcak, B., Robson, I.B., Lamarche, C., Bernstein, A., and Kerbel, R.S. (1988). Genetic tagging of tumor cells with retrovirus vectors: clonal analysis of tumor growth and metastasis in vivo. *Mol. Cell. Biol.* 8, 3143–3149.
- Krause, D.S., and Van Etten, R.A. (2005). Tyrosine kinases as targets for cancer therapy. *N. Engl. J. Med.* 353, 172–187.
- Kris, M.G., Natale, R.B., Herbst, R.S., Lynch, T.J., Prager, D., Belani, C.P., Schiller, J.H., Kelly, K., Spiridonidis, H., Sandler, A., *et al.* (2003). Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: a randomized trial. *JAMA* 290, 2149–2158.
- Krishnaswamy, S., Kanteti, R., Duke-Cohan, J.S., Loganathan, S., Liu, W., Ma, P.C., Sattler, M., Singleton, P.A., Ramnath, N., Innocenti, F., *et al.* (2009). Ethnic differences and functional analysis of MET mutations in lung cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 15, 5714–5723.
- Kunin, V., Sorek, R., and Hugenholtz, P. (2007). Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol.* 8, R61.
- Kwak, E.L., Bang, Y.-J., Camidge, D.R., Shaw, A.T., Solomon, B., Maki, R.G., Ou, S.-H.I., DeZube, B.J., Jänne, P.A., Costa, D.B., *et al.* (2010). Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N. Engl. J. Med.* 363, 1693–1703.
- La Monica, S., Caffarra, C., Sacconi, F., Galvani, E., Galetti, M., Fumarola, C., Bonelli, M., Cavazzoni, A., Cretella, D., Sirangelo, R., *et al.* (2013). Gefitinib inhibits invasive phenotype and epithelial-mesenchymal transition in drug-resistant NSCLC cells with MET amplification. *PloS One* 8, e78656.
- Landau, D.A., Carter, S.L., Stojanov, P., McKenna, A., Stevenson, K., Lawrence, M.S., Sougnez, C., Stewart, C., Sivachenko, A., Wang, L., *et al.* (2013). Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* 152, 714–726.
- Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M.A., and Dick, J.E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* 367, 645–648.
- Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., and Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 505, 495–501.

- Lee, S.Y., Kim, M.J., Jin, G., Yoo, S.S., Park, J.Y., Choi, J.E., Jeon, H.S., Cho, S., Lee, E.B., Cha, S.I., *et al.* (2010). Somatic mutations in epidermal growth factor receptor signaling pathway genes in non-small cell lung cancers. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 5, 1734–1740.
- Lemischka, I.R., Raulet, D.H., and Mulligan, R.C. (1986). Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell* 45, 917–927.
- Li, C., Heidt, D.G., Dalerba, P., Burant, C.F., Zhang, L., Adsay, V., Wicha, M., Clarke, M.F., and Simeone, D.M. (2007). Identification of pancreatic cancer stem cells. *Cancer Res.* 67, 1030–1037.
- Li, D., Qiu, Z., Shao, Y., Chen, Y., Guan, Y., Liu, M., Li, Y., Gao, N., Wang, L., Lu, X., *et al.* (2013a). Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 681–683.
- Li, J.-F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M., and Sheen, J. (2013b). Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* 31, 688–691.
- Li, T., Huang, S., Jiang, W.Z., Wright, D., Spalding, M.H., Weeks, D.P., and Yang, B. (2011). TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.* 39, 359–372.
- Liang P, Xu Y, Zhang X, Ding C, Huang R, Zhang Z, Lv J, Xie X, Chen Y, Li Y, Sun Y, Bai Y, Songyang Z, Ma W, Zhou C, Huang J. (2015). *Protein Cell* 6(5), 363-372.
- Liao, H.-K., Gu, Y., Diaz, A., Marlett, J., Takahashi, Y., Li, M., Suzuki, K., Xu, R., Hishida, T., Chang, C.-J., *et al.* (2015). Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nat. Commun.* 6, 6413.
- Lim, S.H., Sun, J.-M., Choi, Y.-L., Kim, H.R., Ahn, S., Lee, J.Y., Lee, S.-H., Ahn, J.S., Park, K., Kim, J.H., *et al.* (2016a). Efficacy and safety of dovitinib in pretreated patients with advanced squamous non-small cell lung cancer with FGFR1 amplification: A single-arm, phase 2 study. *Cancer* 122, 3024–3031.
- Lim, S.M., Cho, B.C., Kim, S.-W., Kang, S.Y., Heo, D.S., Kim, H.T., Lee, D.H., Kim, D.-W., Jung, M., Choi, J.-H., *et al.* (2016b). A multicenter phase II study of sorafenib in combination with erlotinib in patients with advanced non-small cell lung cancer (KCSG-0806). *Lung Cancer Amst. Neth.* 93, 1–8.
- Linardou, H., Dahabreh, I.J., Kanaloupiti, D., Siannis, F., Bafaloukos, D., Kosmidis, P., Papadimitriou, C.A., and Murray, S. (2008). Assessment of somatic k-RAS mutations as a mechanism associated with resistance to EGFR-targeted agents: a systematic review and meta-analysis of studies in advanced non-small-cell lung cancer and metastatic colorectal cancer. *Lancet Oncol.* 9, 962–972.
- 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.

- Lo, T.-W., Pickle, C.S., Lin, S., Ralston, E.J., Gurling, M., Schartner, C.M., Bian, Q., Doudna, J.A., and Meyer, B.J. (2013). Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. *Genetics* 195, 331–348.
- Long, C., McAnally, J.R., Shelton, J.M., Mireault, A.A., Bassel-Duby, R., and Olson, E.N. (2014). Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. *Science* 345, 1184–1188.
- Long, C., Amoasii, L., Mireault, A.A., McAnally, J.R., Li, H., Sanchez-Ortiz, E., Bhattacharyya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2016). Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* 351, 400–403.
- Longati, P., Bardelli, A., Ponzetto, C., Naldini, L., and Comoglio, P.M. (1994). Tyrosines1234-1235 are critical for activation of the tyrosine kinase encoded by the MET proto-oncogene (HGF receptor). *Oncogene* 9, 49–57.
- Lovly, C.M., McDonald, N.T., Chen, H., Ortiz-Cuaran, S., Heukamp, L.C., Yan, Y., Florin, A., Ozretić, L., Lim, D., Wang, L., *et al.* (2014). Rationale for co-targeting IGF-1R and ALK in ALK fusion-positive lung cancer. *Nat. Med.* 20, 1027–1034.
- Lu, R., Neff, N.F., Quake, S.R., and Weissman, I.L. (2011). Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. *Nat. Biotechnol.* 29, 928–933.
- Lu, X., Peled, N., Greer, J., Wu, W., Choi, P., Berger, A.H., Wong, S., Jen, K.-Y., Seo, Y., Hann, B., *et al.* (2017). MET exon 14 mutation encodes an actionable therapeutic target in lung adenocarcinoma. *Cancer Res.*
- Luengo-Fernandez, R., Leal, J., Gray, A., and Sullivan, R. (2013). Economic burden of cancer across the European Union: a population-based cost analysis. *Lancet Oncol.* 14, 1165–1174.
- 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.
- Ma, H., Tu, L.-C., Naseri, A., Huisman, M., Zhang, S., Grunwald, D., and Pederson, T. (2016). Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat. Biotechnol.* 34, 528–530.
- Ma H, Marti-Gutierrez N, Park SW, Wu J, Lee Y, Suzuki K, Koski A, Ji D, Hayama T, Ahmed R, Darby H, Van Dyken C, Li Y, Kang E, Park AR, Kim D, Kim ST, Gong J, Gu Y, Xu X, Battaglia D, Krieg SA, Lee DM, Wu DH, Wolf DP, Heitner SB, Belmonte JCI, Amato P, Kim JS, Kaul S, Mitalipov S. (2017). Correction of a pathogenic gene mutation in human embryos. *Nature*, 548, 413-419
- Ma, Y., Zhang, X., Shen, B., Lu, Y., Chen, W., Ma, J., Bai, L., Huang, X., and Zhang, L. (2014). Generating rats with conditional alleles using CRISPR/Cas9. *Cell Res.* 24, 122–125.

- Maddalo, D., Manchado, E., Concepcion, C.P., Bonetti, C., Vidigal, J.A., Han, Y.-C., Ogradowski, P., Crippa, A., Rekhman, N., de Stanchina, E., *et al.* (2014). In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature* 516, 423–427.
- Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., and Joung, J.K. (2013). CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* 10, 977–979.
- Maemondo, M., Inoue, A., Kobayashi, K., Sugawara, S., Oizumi, S., Isobe, H., Gemma, A., Harada, M., Yoshizawa, H., Kinoshita, I., *et al.* (2010). Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. *N. Engl. J. Med.* 362, 2380–2388.
- Mak, A.N.-S., Bradley, P., Cernadas, R.A., Bogdanove, A.J., and Stoddard, B.L. (2012). The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science* 335, 716–719.
- Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J.J., Charpentier, E., Haft, D.H., *et al.* (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013a). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826.
- Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013b). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* 31, 833–838.
- Mali, P., Esvelt, K.M., and Church, G.M. (2013c). Cas9 as a versatile tool for engineering biology. *Nat. Methods* 10, 957–963.
- Mandegar, M.A., Huebsch, N., Frolov, E.B., Shin, E., Truong, A., Olvera, M.P., Chan, A.H., Miyaoka, Y., Holmes, K., Spencer, C.I., *et al.* (2016). CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. *Cell Stem Cell* 18, 541–553.
- Markham, A. (2017). Brigatinib: First Global Approval. *Drugs* 77, 1131–1135.
- Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843–1845.
- Marusyk, A., and Polyak, K. (2010). Tumor heterogeneity: causes and consequences. *Biochim. Biophys. Acta* 1805, 105–117.
- Maruyama, T., Dougan, S.K., Truttmann, M.C., Bilate, A.M., Ingram, J.R., and Ploegh, H.L. (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* 33, 538–542.
- Mashimo, T., Takizawa, A., Voigt, B., Yoshimi, K., Hiai, H., Kuramoto, T., and Serikawa, T. (2010). Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases. *PloS One* 5, e8870.
- McGranahan, N., and Swanton, C. (2015). Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell* 27, 15–26.

- Meacham, C.E., and Morrison, S.J. (2013). Tumour heterogeneity and cancer cell plasticity. *Nature* 501, 328–337.
- van Meerbeeck, J.P., Fennell, D.A., and De Ruyscher, D.K.M. (2011). Small-cell lung cancer. *Lancet Lond. Engl.* 378, 1741–1755.
- Mendes-Pereira, A.M., Martin, S.A., Brough, R., McCarthy, A., Taylor, J.R., Kim, J.-S., Waldman, T., Lord, C.J., and Ashworth, A. (2009). Synthetic lethal targeting of PTEN mutant cells with PARP inhibitors. *EMBO Mol. Med.* 1, 315–322.
- Mendes-Pereira, A.M., Lord, C.J., and Ashworth, A. (2012). NLK is a novel therapeutic target for PTEN deficient tumour cells. *PloS One* 7, e47249.
- Meng, X., Noyes, M.B., Zhu, L.J., Lawson, N.D., and Wolfe, S.A. (2008). Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat. Biotechnol.* 26, 695–701.
- Meyer, M., de Angelis, M.H., Wurst, W., and Kühn, R. (2010). Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. *Proc. Natl. Acad. Sci. U. S. A.* 107, 15022–15026.
- Michor, F., and Polyak, K. (2010). The origins and implications of intratumor heterogeneity. *Cancer Prev. Res. Phila. Pa* 3, 1361–1364.
- Miller, J., McLachlan, A.D., and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* 4, 1609–1614.
- Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., *et al.* (2011). A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* 29, 143–148.
- Miller, V.A., Kris, M.G., Shah, N., Patel, J., Azzoli, C., Gomez, J., Krug, L.M., Pao, W., Rizvi, N., Pizzo, B., *et al.* (2004). Bronchioloalveolar pathologic subtype and smoking history predict sensitivity to gefitinib in advanced non-small-cell lung cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 22, 1103–1109.
- 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.
- Mitsudomi, T., Morita, S., Yatabe, Y., Negoro, S., Okamoto, I., Tsurutani, J., Seto, T., Satouchi, M., Tada, H., Hirashima, T., *et al.* (2010). Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): an open label, randomised phase 3 trial. *Lancet Oncol.* 11, 121–128.
- Mojica, F.J., Díez-Villaseñor, C., Soria, E., and Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* 36, 244–246.
- Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60, 174–182.

- Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiol. Read. Engl.* *155*, 733–740.
- Mok, T.S., Wu, Y.-L., Thongprasert, S., Yang, C.-H., Chu, D.-T., Saijo, N., Sunpaweravong, P., Han, B., Margono, B., Ichinose, Y., *et al.* (2009). Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N. Engl. J. Med.* *361*, 947–957.
- Montagut, C., Dalmases, A., Bellosillo, B., Crespo, M., Pairet, S., Iglesias, M., Salido, M., Gallen, M., Marsters, S., Tsai, S.P., *et al.* (2012). Identification of a mutation in the extracellular domain of the Epidermal Growth Factor Receptor conferring cetuximab resistance in colorectal cancer. *Nat. Med.* *18*, 221–223.
- Moore, J.K., and Haber, J.E. (1996). Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* *16*, 2164–2173.
- Moreno-Mateos, M.A., Vejnar, C.E., Beaudoin, J.-D., Fernandez, J.P., Mis, E.K., Khokha, M.K., and Giraldez, A.J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods* *12*, 982–988.
- Moscou, M.J., and Bogdanove, A.J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science* *326*, 1501.
- Naoki, K., Chen, T.-H., Richards, W.G., Sugarbaker, D.J., and Meyerson, M. (2002). Missense mutations of the BRAF gene in human lung adenocarcinoma. *Cancer Res.* *62*, 7001–7003.
- Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., *et al.* (2011). Tumour evolution inferred by single-cell sequencing. *Nature* *472*, 90–94.
- Negri, T., Pavan, G.M., Viridis, E., Greco, A., Fermeglia, M., Sandri, M., Pricl, S., Pierotti, M.A., Pilotti, S., and Tamborini, E. (2009). T670X KIT mutations in gastrointestinal stromal tumors: making sense of missense. *J. Natl. Cancer Inst.* *101*, 194–204.
- Nelles, D.A., Fang, M.Y., O’Connell, M.R., Xu, J.L., Markmiller, S.J., Doudna, J.A., and Yeo, G.W. (2016). Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell* *165*, 488–496.
- Nelson, C.E., Hakim, C.H., Ousterout, D.G., Thakore, P.I., Moreb, E.A., Castellanos Rivera, R.M., Madhavan, S., Pan, X., Ran, F.A., Yan, W.X., *et al.* (2016). In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* *351*, 403–407.
- Ng, K.P., Hillmer, A.M., Chuah, C.T.H., Juan, W.C., Ko, T.K., Teo, A.S.M., Ariyaratne, P.N., Takahashi, N., Sawada, K., Fei, Y., *et al.* (2012). A common BIM deletion polymorphism mediates intrinsic resistance and inferior responses to tyrosine kinase inhibitors in cancer. *Nat. Med.* *18*, 521–528.
- Nguyen, L.V., Cox, C.L., Eirew, P., Knapp, D.J.H.F., Pellacani, D., Kannan, N., Carles, A., Moksa, M., Balani, S., Shah, S., *et al.* (2014). DNA barcoding reveals diverse growth kinetics of human breast tumour subclones in serially passaged xenografts. *Nat. Commun.* *5*, 5871.

- Nihongaki, Y., Yamamoto, S., Kawano, F., Suzuki, H., and Sato, M. (2015a). CRISPR-Cas9-based photoactivatable transcription system. *Chem. Biol.* *22*, 169–174.
- Nihongaki, Y., Kawano, F., Nakajima, T., and Sato, M. (2015b). Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nat. Biotechnol.* *33*, 755–760.
- Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* *156*, 935–949.
- Nishimasu, H., Cong, L., Yan, W.X., Ran, F.A., Zetsche, B., Li, Y., Kurabayashi, A., Ishitani, R., Zhang, F., and Nureki, O. (2015). Crystal Structure of *Staphylococcus aureus* Cas9. *Cell* *162*, 1113–1126.
- Niu, Y., Shen, B., Cui, Y., Chen, Y., Wang, J., Wang, L., Kang, Y., Zhao, X., Si, W., Li, W., *et al.* (2014). Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell* *156*, 836–843.
- Nowell, P.C. (1976). The clonal evolution of tumor cell populations. *Science* *194*, 23–28.
- Nuñez, J.K., Kranzusch, P.J., Noeske, J., Wright, A.V., Davies, C.W., and Doudna, J.A. (2014). Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* *21*, 528–534.
- Nuñez, J.K., Lee, A.S.Y., Engelman, A., and Doudna, J.A. (2015). Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* *519*, 193–198.
- O’Brien, C.A., Pollett, A., Gallinger, S., and Dick, J.E. (2007). A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* *445*, 106–110.
- O’Connell, M.R., Oakes, B.L., Sternberg, S.H., East-Seletsky, A., Kaplan, M., and Doudna, J.A. (2014). Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* *516*, 263–266.
- Ohashi, K., Sequist, L.V., Arcila, M.E., Moran, T., Chmielecki, J., Lin, Y.-L., Pan, Y., Wang, L., de Stanchina, E., Shien, K., *et al.* (2012). Lung cancers with acquired resistance to EGFR inhibitors occasionally harbor BRAF gene mutations but lack mutations in KRAS, NRAS, or MEK1. *Proc. Natl. Acad. Sci. U. S. A.* *109*, E2127-2133.
- Onozato, R., Kosaka, T., Kuwano, H., Sekido, Y., Yatabe, Y., and Mitsudomi, T. (2009). Activation of MET by gene amplification or by splice mutations deleting the juxtamembrane domain in primary resected lung cancers. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* *4*, 5–11.
- van der Oost, J., Westra, E.R., Jackson, R.N., and Wiedenheft, B. (2014). Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.* *12*, 479–492.
- Ostrem, J.M., Peters, U., Sos, M.L., Wells, J.A., and Shokat, K.M. (2013). K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* *503*, 548–551.

- 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.
- Pao, W., and Chmielecki, J. (2010). Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. *Nat. Rev. Cancer* *10*, 760–774.
- Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., Singh, B., Heelan, R., Rusch, V., Fulton, L., *et al.* (2004). EGF receptor gene mutations are common in lung cancers from “never smokers” and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc. Natl. Acad. Sci. U. S. A.* *101*, 13306–13311.
- Pao, W., Wang, T.Y., Riely, G.J., Miller, V.A., Pan, Q., Ladanyi, M., Zakowski, M.F., Heelan, R.T., Kris, M.G., and Varmus, H.E. (2005a). KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS Med.* *2*, e17.
- Pao, W., Miller, V.A., Politi, K.A., Riely, G.J., Somwar, R., Zakowski, M.F., Kris, M.G., and Varmus, H. (2005b). Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med.* *2*, e73.
- Pardo, B., Gómez-González, B., and Aguilera, A. (2009). DNA repair in mammalian cells: DNA double-strand break repair: how to fix a broken relationship. *Cell. Mol. Life Sci. CMLS* *66*, 1039–1056.
- Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A., and Liu, D.R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* *31*, 839–843.
- Pavletich, N.P., and Pabo, C.O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* *252*, 809–817.
- Paz-Ares, L., Hirsh, V., Zhang, L., de Marinis, F., Yang, J.C.-H., Wakelee, H.A., Seto, T., Wu, Y.-L., Novello, S., Juhász, E., *et al.* (2015). Monotherapy Administration of Sorafenib in Patients With Non-Small Cell Lung Cancer (MISSION) Trial: A Phase III, Multicenter, Placebo-Controlled Trial of Sorafenib in Patients with Relapsed or Refractory Predominantly Nonsquamous Non-Small-Cell Lung Cancer after 2 or 3 Previous Treatment Regimens. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* *10*, 1745–1753.
- Pedone, P.V., Ghirlando, R., Clore, G.M., Gronenborn, A.M., Felsenfeld, G., and Omichinski, J.G. (1996). The single Cys2-His2 zinc finger domain of the GAGA protein flanked by basic residues is sufficient for high-affinity specific DNA binding. *Proc. Natl. Acad. Sci. U. S. A.* *93*, 2822–2826.
- Perez, C.A., Velez, M., Raez, L.E., and Santos, E.S. (2014). Overcoming the resistance to crizotinib in patients with non-small cell lung cancer harboring EML4/ALK translocation. *Lung Cancer Amst. Neth.* *84*, 110–115.
- Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., *et al.* (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* *10*, 973–976.

- Pinzon, C.A., Snyder, M., Pryor, J., Thompson, B., Golding, M., and Long, C. (2016). 207 EFFICIENT GENERATION OF MYOSTATIN PROMOTER MUTATIONS IN BOVINE EMBRYOS USING THE CRISPR/Cas9 SYSTEM. *Reprod. Fertil. Dev.* *29*, 212.
- Planchard, D., Lorient, Y., André, F., Gobert, A., Auger, N., Lacroix, L., and Soria, J.C. (2015). EGFR-independent mechanisms of acquired resistance to AZD9291 in EGFR T790M-positive NSCLC patients. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* *26*, 2073–2078.
- Polstein, L.R., and Gersbach, C.A. (2015). A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* *11*, 198–200.
- Popplewell, L., Koo, T., Leclerc, X., Duclert, A., Mamchaoui, K., Gouble, A., Mouly, V., Voit, T., Pâques, F., Cédronne, F., *et al.* (2013). Gene correction of a duchenne muscular dystrophy mutation by meganuclease-enhanced exon knock-in. *Hum. Gene Ther.* *24*, 692–701.
- Pourcel, C., Salvignol, G., and Vergnaud, G. (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiol. Read. Engl.* *151*, 653–663.
- Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* *152*, 1173–1183.
- Ramakrishna, S., Kwaku Dad, A.-B., Beloor, J., Gopalappa, R., Lee, S.-K., and Kim, H. (2014). Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Res.* *24*, 1020–1027.
- Ramanan, V., Shlomai, A., Cox, D.B.T., Schwartz, R.E., Michailidis, E., Bhatta, A., Scott, D.A., Zhang, F., Rice, C.M., and Bhatia, S.N. (2015). CRISPR/Cas9 cleavage of viral DNA efficiently suppresses hepatitis B virus. *Sci. Rep.* *5*, 10833.
- Ramalingam SS, Yang JC, Lee CK, Kurata T, Kim DW, John T, Nogami N, Ohe Y, Mann H, Rukazenzov Y, Ghiorghiu S, Stetson D, Markovets A, Barrett JC, Thress KS, Jänne PA. (2017). Osimertinib As First-Line Treatment of EGFR Mutation-Positive Advanced Non-Small-Cell Lung Cancer. *J Clin Oncol.*
- Ran, F.A., Hsu, P.D., Lin, C.-Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., *et al.* (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* *154*, 1380–1389.
- Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., *et al.* (2015). In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* *520*, 186–191.
- Rebar, E.J., and Pabo, C.O. (1994). Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science* *263*, 671–673.
- Richardson, C.D., Ray, G.J., DeWitt, M.A., Curie, G.L., and Corn, J.E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* *34*, 339–344.

Riely, G.J., Johnson, M.L., Medina, C., Rizvi, N.A., Miller, V.A., Kris, M.G., Pietanza, M.C., Azzoli, C.G., Krug, L.M., Pao, W., *et al.* (2011). A phase II trial of Salirasib in patients with lung adenocarcinomas with KRAS mutations. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 6, 1435–1437.

Rivière, J., Hauer, J., Poirot, L., Brochet, J., Souque, P., Mollier, K., Gouble, A., Charneau, P., Fischer, A., Pâques, F., *et al.* (2014). Variable correction of Artemis deficiency by I-SceI-meganuclease-assisted homologous recombination in murine hematopoietic stem cells. *Gene Ther.* 21, 529–532.

Robinson-Hamm, J.N., and Gersbach, C.A. (2016). Gene therapies that restore dystrophin expression for the treatment of Duchenne muscular dystrophy. *Hum. Genet.* 135, 1029–1040.

Rosell, R., Carcereny, E., Gervais, R., Vergnenegre, A., Massuti, B., Felip, E., Palmero, R., Garcia-Gomez, R., Pallares, C., Sanchez, J.M., *et al.* (2012). Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. *Lancet Oncol.* 13, 239–246.

Rouet, P., Smih, F., and Jasin, M. (1994). Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 91, 6064–6068.

Sargueil, B., Delahodde, A., Hatat, D., Tian, G.L., Lazowska, J., and Jacq, C. (1991). A new specific DNA endonuclease activity in yeast mitochondria. *Mol. Gen. Genet. MGG* 225, 340–341.

Sauna, Z.E., and Kimchi-Sarfaty, C. (2011). Understanding the contribution of synonymous mutations to human disease. *Nat. Rev. Genet.* 12, 683–691.

Scagliotti, G., von Pawel, J., Novello, S., Ramlau, R., Favaretto, A., Barlesi, F., Akerley, W., Orlov, S., Santoro, A., Spigel, D., *et al.* (2015). Phase III Multinational, Randomized, Double-Blind, Placebo-Controlled Study of Tivantinib (ARQ 197) Plus Erlotinib Versus Erlotinib Alone in Previously Treated Patients With Locally Advanced or Metastatic Nonsquamous Non-Small-Cell Lung Cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 33, 2667–2674.

Scagliotti, G.V., Novello, S., Schiller, J.H., Hirsh, V., Sequist, L.V., Soria, J.-C., von Pawel, J., Schwartz, B., Von Roemeling, R., and Sandler, A.B. (2012). Rationale and design of MARQUEE: a phase III, randomized, double-blind study of tivantinib plus erlotinib versus placebo plus erlotinib in previously treated patients with locally advanced or metastatic, nonsquamous, non-small-cell lung cancer. *Clin. Lung Cancer* 13, 391–395.

Schepers, A.G., Snippert, H.J., Stange, D.E., van den Born, M., van Es, J.H., van de Wetering, M., and Clevers, H. (2012). Lineage tracing reveals Lgr5⁺ stem cell activity in mouse intestinal adenomas. *Science* 337, 730–735.

Schrock, A.B., Frampton, G.M., Suh, J., Chalmers, Z.R., Rosenzweig, M., Erlich, R.L., Halmos, B., Goldman, J., Forde, P., Leuenberger, K., *et al.* (2016). Characterization of 298 Patients with Lung Cancer Harboring MET Exon 14 Skipping Alterations. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 11, 1493–1502.

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.

Sequist, L.V., Rolfe, L., and Allen, A.R. (2015). Rociletinib in EGFR-Mutated Non-Small-Cell Lung Cancer. *N. Engl. J. Med.* *373*, 578–579.

Seto, T., Kiura, K., Nishio, M., Nakagawa, K., Maemondo, M., Inoue, A., Hida, T., Yamamoto, N., Yoshioka, H., Harada, M., *et al.* (2013). 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*, 590–598.

Shah, S.P., Roth, A., Goya, R., Oloumi, A., Ha, G., Zhao, Y., Turashvili, G., Ding, J., Tse, K., Haffari, G., *et al.* (2012). The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* *486*, 395–399.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., *et al.* (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* *343*, 84–87.

Shan, Y., Eastwood, M.P., Zhang, X., Kim, E.T., Arkhipov, A., Dror, R.O., Jumper, J., Kuriyan, J., and Shaw, D.E. (2012). Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. *Cell* *149*, 860–870.

Shao, S., Zhang, W., Hu, H., Xue, B., Qin, J., Sun, C., Sun, Y., Wei, W., and Sun, Y. (2016). Long-term dual-color tracking of genomic loci by modified sgRNAs of the CRISPR/Cas9 system. *Nucleic Acids Res.* *44*, e86.

Shaw, A.T., and Engelman, J.A. (2014). Ceritinib in ALK-rearranged non-small-cell lung cancer. *N. Engl. J. Med.* *370*, 2537–2539.

Shaw, A.T., Yeap, B.Y., Mino-Kenudson, M., Digumarthy, S.R., Costa, D.B., Heist, R.S., Solomon, B., Stubbs, H., Admane, S., McDermott, U., *et al.* (2009). Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* *27*, 4247–4253.

Shaw, A.T., Ou, S.-H.I., Bang, Y.-J., Camidge, D.R., Solomon, B.J., Salgia, R., Riely, G.J., Varella-Garcia, M., Shapiro, G.I., Costa, D.B., *et al.* (2014). Crizotinib in ROS1-rearranged non-small-cell lung cancer. *N. Engl. J. Med.* *371*, 1963–1971.

Shen, W.H., Balajee, A.S., Wang, J., Wu, H., Eng, C., Pandolfi, P.P., and Yin, Y. (2007). Essential role for nuclear PTEN in maintaining chromosomal integrity. *Cell* *128*, 157–170.

Shimizu, Y., Bhakta, M.S., and Segal, D.J. (2009). Restricted spacer tolerance of a zinc finger nuclease with a six amino acid linker. *Bioorg. Med. Chem. Lett.* *19*, 3970–3972.

Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., *et al.* (2015). Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. *Mol. Cell* *60*, 385–397.

Singh, S.K., Hawkins, C., Clarke, I.D., Squire, J.A., Bayani, J., Hide, T., Henkelman, R.M., Cusimano, M.D., and Dirks, P.B. (2004). Identification of human brain tumour initiating cells. *Nature* *432*, 396–401.

- Slymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science* 351, 84–88.
- Smith, C., Gore, A., Yan, W., Abalde-Atristain, L., Li, Z., He, C., Wang, Y., Brodsky, R.A., Zhang, K., Cheng, L., *et al.* (2014). Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. *Cell Stem Cell* 15, 12–13.
- Smith, J., Bibikova, M., Whitby, F.G., Reddy, A.R., Chandrasegaran, S., and Carroll, D. (2000). Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res.* 28, 3361–3369.
- Smith, J., Grizot, S., Arnould, S., Duclert, A., Epinat, J.-C., Chames, P., Prieto, J., Redondo, P., Blanco, F.J., Bravo, J., *et al.* (2006). A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res.* 34, e149.
- Snuderl, M., Fazlollahi, L., Le, L.P., Nitta, M., Zhelyazkova, B.H., Davidson, C.J., Akhavanfard, S., Cahill, D.P., Aldape, K.D., Betensky, R.A., *et al.* (2011). Mosaic amplification of multiple receptor tyrosine kinase genes in glioblastoma. *Cancer Cell* 20, 810–817.
- 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.
- Sos, M.L., Koker, M., Weir, B.A., Heynck, S., Rabinovsky, R., Zander, T., Seeger, J.M., Weiss, J., Fischer, F., Frommolt, P., *et al.* (2009). PTEN loss contributes to erlotinib resistance in EGFR-mutant lung cancer by activation of Akt and EGFR. *Cancer Res.* 69, 3256–3261.
- Sottoriva, A., Spiteri, I., Piccirillo, S.G.M., Touloumis, A., Collins, V.P., Marioni, J.C., Curtis, C., Watts, C., and Tavaré, S. (2013). Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc. Natl. Acad. Sci. U. S. A.* 110, 4009–4014.
- Sousa, S.F., Fernandes, P.A., and Ramos, M.J. (2008). Farnesyltransferase inhibitors: a detailed chemical view on an elusive biological problem. *Curr. Med. Chem.* 15, 1478–1492.
- Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507, 62–67.
- Sternberg, S.H., LaFrance, B., Kaplan, M., and Doudna, J.A. (2015). Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature* 527, 110–113.
- Suda, K., Tomizawa, K., and Mitsudomi, T. (2010). Biological and clinical significance of KRAS mutations in lung cancer: an oncogenic driver that contrasts with EGFR mutation. *Cancer Metastasis Rev.* 29, 49–60.
- Sung, Y.H., Baek, I.-J., Kim, D.H., Jeon, J., Lee, J., Lee, K., Jeong, D., Kim, J.-S., and Lee, H.-W. (2013). Knockout mice created by TALEN-mediated gene targeting. *Nat. Biotechnol.* 31, 23–24.

- Sussman, D., Chadsey, M., Fauce, S., Engel, A., Bruett, A., Monnat, R., Stoddard, B.L., and Seligman, L.M. (2004). Isolation and characterization of new homing endonuclease specificities at individual target site positions. *J. Mol. Biol.* 342, 31–41.
- Swiech, L., Heidenreich, M., Banerjee, A., Habib, N., Li, Y., Trombetta, J., Sur, M., and Zhang, F. (2015). In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat. Biotechnol.* 33, 102–106.
- Tabebordbar, M., Zhu, K., Cheng, J.K.W., Chew, W.L., Widrick, J.J., Yan, W.X., Maesner, C., Wu, E.Y., Xiao, R., Ran, F.A., *et al.* (2016). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* 351, 407–411.
- Tai, W.-T., Cheng, A.-L., Shiau, C.-W., Huang, H.-P., Huang, J.-W., Chen, P.-J., and Chen, K.-F. (2011). Signal transducer and activator of transcription 3 is a major kinase-independent target of sorafenib in hepatocellular carcinoma. *J. Hepatol.* 55, 1041–1048.
- Takezawa, K., Pirazzoli, V., Arcila, M.E., Nebhan, C.A., Song, X., de Stanchina, E., Ohashi, K., Janjigian, Y.Y., Spitzler, P.J., Melnick, M.A., *et al.* (2012). HER2 amplification: a potential mechanism of acquired resistance to EGFR inhibition in EGFR-mutant lung cancers that lack the second-site EGFR T790M mutation. *Cancer Discov.* 2, 922–933.
- Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* 159, 635–646.
- Tesson, L., Usal, C., Ménoret, S., Leung, E., Niles, B.J., Remy, S., Santiago, Y., Vincent, A.I., Meng, X., Zhang, L., *et al.* (2011). Knockout rats generated by embryo microinjection of TALENs. *Nat. Biotechnol.* 29, 695–696.
- Thirlwell, C., Will, O.C.C., Domingo, E., Graham, T.A., McDonald, S.A.C., Oukrif, D., Jeffrey, R., Gorman, M., Rodriguez-Justo, M., Chin-Aleong, J., *et al.* (2010). Clonality assessment and clonal ordering of individual neoplastic crypts shows polyclonality of colorectal adenomas. *Gastroenterology* 138, 1441–1454, 1454.e1-7.
- 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.* (2015). Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. *Nat. Med.* 21, 560–562.
- Tomasetti, C., Li, L., and Vogelstein, B. (2017). Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science* 355, 1330–1334.
- Torres, R., Martin, M.C., Garcia, A., Cigudosa, J.C., Ramirez, J.C., and Rodriguez-Perales, S. (2014). Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR-Cas9 system. *Nat. Commun.* 5, 3964.
- Toyokawa, G., and Seto, T. (2015). Updated Evidence on the Mechanisms of Resistance to ALK Inhibitors and Strategies to Overcome Such Resistance: Clinical and Preclinical Data. *Oncol. Res. Treat.* 38, 291–298.
- Tran, P.N., and Klempner, S.J. (2016). Profile of rociletinib and its potential in the treatment of non-small-cell lung cancer. *Lung Cancer Auckl. NZ* 7, 91–97.

- Tricker, E.M., Xu, C., Uddin, S., Capelletti, M., Ercan, D., Ogino, A., Pratilas, C.A., Rosen, N., Gray, N.S., Wong, K.-K., *et al.* (2015). Combined EGFR/MEK Inhibition Prevents the Emergence of Resistance in EGFR-Mutant Lung Cancer. *Cancer Discov.* *5*, 960–971.
- Tsai, S.Q., and Joung, J.K. (2016). Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat. Rev. Genet.* *17*, 300–312.
- Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J., and Joung, J.K. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* *32*, 569–576.
- Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P., *et al.* (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* *33*, 187–197.
- Tsao, A.S., Scagliotti, G.V., Bunn, P.A., Carbone, D.P., Warren, G.W., Bai, C., de Koning, H.J., Yousaf-Khan, A.U., McWilliams, A., Tsao, M.S., *et al.* (2016). Scientific Advances in Lung Cancer 2015. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* *11*, 613–638.
- Tucker, E.R., Danielson, L.S., Innocenti, P., and Chesler, L. (2015). Tackling Crizotinib Resistance: The Pathway from Drug Discovery to the Pediatric Clinic. *Cancer Res.* *75*, 2770–2774.
- Turke, A.B., Zejnullahu, K., Wu, Y.-L., Song, Y., Dias-Santagata, D., Lifshits, E., Toschi, L., Rogers, A., Mok, T., Sequist, L., *et al.* (2010). Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC. *Cancer Cell* *17*, 77–88.
- Turner, N.C., and Seckl, M.J. (2010). A therapeutic target for smoking-associated lung cancer. *Sci. Transl. Med.* *2*, 62ps56.
- Tyner, J.W., Fletcher, L.B., Wang, E.Q., Yang, W.F., Rutenberg-Schoenberg, M.L., Beadling, C., Mori, M., Heinrich, M.C., Deininger, M.W., Druker, B.J., *et al.* (2010). MET receptor sequence variants R970C and T992I lack transforming capacity. *Cancer Res.* *70*, 6233–6237.
- Uchibori, K., Inase, N., Araki, M., Kamada, M., Sato, S., Okuno, Y., Fujita, N., and Katayama, R. (2017). Brigatinib combined with anti-EGFR antibody overcomes osimertinib resistance in EGFR-mutated non-small-cell lung cancer. *Nat. Commun.* *8*, 14768.
- Uramoto, H., Iwata, T., Onitsuka, T., Shimokawa, H., Hanagiri, T., and Oyama, T. (2010). Epithelial-mesenchymal transition in EGFR-TKI acquired resistant lung adenocarcinoma. *Anticancer Res.* *30*, 2513–2517.
- Veres, A., Gosis, B.S., Ding, Q., Collins, R., Ragavendran, A., Brand, H., Erdin, S., Cowan, C.A., Talkowski, M.E., and Musunuru, K. (2014). Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. *Cell Stem Cell* *15*, 27–30.
- Virchow, R., and Chance, F. (1860). Cellular pathology, as based upon physiological and pathological histology. Twenty lectures delivered in the Pathological institute of Berlin during the months of February, March and April, 1858 (Robert M. De Witt).

- Vogelstein, B., and Kinzler, K.W. (1993). The multistep nature of cancer. *Trends Genet. TIG* 9, 138–141.
- Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., and Kinzler, K.W. (2013). Cancer genome landscapes. *Science* 339, 1546–1558.
- Wakelee, H.A., Lee, J.-W., Hanna, N.H., Traynor, A.M., Carbone, D.P., and Schiller, J.H. (2012). A double-blind randomized discontinuation phase-II study of sorafenib (BAY 43-9006) in previously treated non-small-cell lung cancer patients: eastern cooperative oncology group study E2501. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 7, 1574–1582.
- Waltz, E. (2016). Gene-edited CRISPR mushroom escapes US regulation. *Nature* , 293.
- Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153, 910–918.
- Wang, J., Li, J., Zhao, H., Sheng, G., Wang, M., Yin, M., and Wang, Y. (2015a). Structural and Mechanistic Basis of PAM-Dependent Spacer Acquisition in CRISPR-Cas Systems. *Cell* 163, 840–853.
- Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., and Sabatini, D.M. (2015b). Identification and characterization of essential genes in the human genome. *Science* 350, 1096–1101.
- Weaver, D.T. (1995). What to do at an end: DNA double-strand-break repair. *Trends Genet. TIG* 11, 388–392.
- Wei, J., Wagner, S., Lu, D., Maclean, P., Carlson, D.F., Fahrenkrug, S.C., and Laible, G. (2015a). Efficient introgression of allelic variants by embryo-mediated editing of the bovine genome. *Sci. Rep.* 5, 11735.
- Wei, Y., Terns, R.M., and Terns, M.P. (2015b). Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. *Genes Dev.* 29, 356–361.
- Weiss, J., Sos, M.L., Seidel, D., Peifer, M., Zander, T., Heuckmann, J.M., Ullrich, R.T., Menon, R., Maier, S., Soltermann, A., *et al.* (2010). Frequent and focal FGFR1 amplification associates with therapeutically tractable FGFR1 dependency in squamous cell lung cancer. *Sci. Transl. Med.* 2, 62ra93.
- Whitworth, K.M., Lee, K., Benne, J.A., Beaton, B.P., Spate, L.D., Murphy, S.L., Samuel, M.S., Mao, J., O’Gorman, C., Walters, E.M., *et al.* (2014). Use of the CRISPR/Cas9 system to produce genetically engineered pigs from in vitro-derived oocytes and embryos. *Biol. Reprod.* 91, 78.
- Wilhelm, S., Carter, C., Lynch, M., Lowinger, T., Dumas, J., Smith, R.A., Schwartz, B., Simantov, R., and Kelley, S. (2006). Discovery and development of sorafenib: a multikinase inhibitor for treating cancer. *Nat. Rev. Drug Discov.* 5, 835–844.

- Wilhelm, S.M., Dumas, J., Adnane, L., Lynch, M., Carter, C.A., Schütz, G., Thierauch, K.-H., and Zopf, D. (2011). Regorafenib (BAY 73-4506): a new oral multikinase inhibitor of angiogenic, stromal and oncogenic receptor tyrosine kinases with potent preclinical antitumor activity. *Int. J. Cancer* *129*, 245–255.
- Wong, D.W.-S., Leung, E.L.-H., So, K.K.-T., Tam, I.Y.-S., Sihoe, A.D.-L., Cheng, L.-C., Ho, K.-K., Au, J.S.-K., Chung, L.-P., Pik Wong, M., *et al.* (2009). The EML4-ALK fusion gene is involved in various histologic types of lung cancers from nonsmokers with wild-type EGFR and KRAS. *Cancer* *115*, 1723–1733.
- Wood, K., Hensing, T., Malik, R., and Salgia, R. (2016). Prognostic and Predictive Value in KRAS in Non-Small-Cell Lung Cancer: A Review. *JAMA Oncol.* *2*, 805–812.
- Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., Yang, Z., Zu, Y., Li, W., Huang, P., Tong, X., *et al.* (2013). Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.* *41*, e141.
- Xie, M., Zhang, L., He, C., Xu, F., Liu, J., Hu, Z., Zhao, L., and Tian, Y. (2012). Activation of Notch-1 enhances epithelial-mesenchymal transition in gefitinib-acquired resistant lung cancer cells. *J. Cell. Biochem.* *113*, 1501–1513.
- Xue, W., Chen, S., Yin, H., Tammela, T., Papagiannakopoulos, T., Joshi, N.S., Cai, W., Yang, G., Bronson, R., Crowley, D.G., *et al.* (2014). CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature* *514*, 380–384.
- Yachida, S., Jones, S., Bozic, I., Antal, T., Leary, R., Fu, B., Kamiyama, M., Hruban, R.H., Eshleman, J.R., Nowak, M.A., *et al.* (2010). Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* *467*, 1114–1117.
- Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L., and Jaenisch, R. (2013). One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* *154*, 1370–1379.
- Yap, T.A., Gerlinger, M., Futreal, P.A., Pusztai, L., and Swanton, C. (2012). Intratumor heterogeneity: seeing the wood for the trees. *Sci. Transl. Med.* *4*, 127ps10.
- Yasuda, H., de Figueiredo-Pontes, L.L., Kobayashi, S., and Costa, D.B. (2012). Preclinical rationale for use of the clinically available multitargeted tyrosine kinase inhibitor crizotinib in ROS1-translocated lung cancer. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* *7*, 1086–1090.
- Yates, L.R., and Campbell, P.J. (2012). Evolution of the cancer genome. *Nat. Rev. Genet.* *13*, 795–806.
- Yin, H., Xue, W., Chen, S., Bogorad, R.L., Benedetti, E., Grompe, M., Kotliansky, V., Sharp, P.A., Jacks, T., and Anderson, D.G. (2014). Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat. Biotechnol.* *32*, 551–553.
- Young, J.J., Cherone, J.M., Doyon, Y., Ankoudinova, I., Faraji, F.M., Lee, A.H., Ngo, C., Guschin, D.Y., Paschon, D.E., Miller, J.C., *et al.* (2011). Efficient targeted gene disruption in the soma and germ line of the frog *Xenopus tropicalis* using engineered zinc-finger nucleases. *Proc. Natl. Acad. Sci. U. S. A.* *108*, 7052–7057.

- Yu, C., Liu, Y., Ma, T., Liu, K., Xu, S., Zhang, Y., Liu, H., La Russa, M., Xie, M., Ding, S., *et al.* (2015). Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* *16*, 142–147.
- Yu, H.A., Arcila, M.E., Rekhtman, N., Sima, C.S., Zakowski, M.F., Pao, W., Kris, M.G., Miller, V.A., Ladanyi, M., and Riely, G.J. (2013). Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *19*, 2240–2247.
- Yuan, L., Sui, T., Chen, M., Deng, J., Huang, Y., Zeng, J., Lv, Q., Song, Y., Li, Z., and Lai, L. (2016). CRISPR/Cas9-mediated GJA8 knockout in rabbits recapitulates human congenital cataracts. *Sci. Rep.* *6*, 22024.
- 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.
- Zhang, J., Fujimoto, J., Zhang, J., Wedge, D.C., Song, X., Zhang, J., Seth, S., Chow, C.-W., Cao, Y., Gumbs, C., *et al.* (2014). Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* *346*, 256–259.
- Zhang, S., Anjum, R., Squillace, R., Nadworny, S., Zhou, T., Keats, J., Ning, Y., Wardwell, S.D., Miller, D., Song, Y., *et al.* (2016). The Potent ALK Inhibitor Brigatinib (AP26113) Overcomes Mechanisms of Resistance to First- and Second-Generation ALK Inhibitors in Preclinical Models. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *22*, 5527–5538.
- Zhang, Y., Heidrich, N., Ampattu, B.J., Gunderson, C.W., Seifert, H.S., Schoen, C., Vogel, J., and Sontheimer, E.J. (2013). Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol. Cell* *50*, 488–503.
- Zhang, Z., Lee, J.C., Lin, L., Olivas, V., Au, V., LaFramboise, T., Abdel-Rahman, M., Wang, X., Levine, A.D., Rho, J.K., *et al.* (2012). Activation of the AXL kinase causes resistance to EGFR-targeted therapy in lung cancer. *Nat. Genet.* *44*, 852–860.
- Zhou, C., Wu, Y.-L., Chen, G., Feng, J., Liu, X.-Q., Wang, C., Zhang, S., Wang, J., Zhou, S., Ren, S., *et al.* (2011). Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. *Lancet Oncol.* *12*, 735–742.
- Zhou, W., Ercan, D., Chen, L., Yun, C.-H., Li, D., Capelletti, M., Cortot, A.B., Chirieac, L., Iacob, R.E., Padera, R., *et al.* (2009). Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature* *462*, 1070–1074.
- Zou, H.Y., Friboulet, L., Kodack, D.P., Engstrom, L.D., Li, Q., West, M., Tang, R.W., Wang, H., Tsaparikos, K., Wang, J., *et al.* (2015). PF-06463922, an ALK/ROS1 Inhibitor, Overcomes Resistance to First and Second Generation ALK Inhibitors in Preclinical Models. *Cancer Cell* *28*, 70–81.
- Zuckermann, M., Hovestadt, V., Knobbe-Thomsen, C.B., Zapatka, M., Northcott, P.A., Schramm, K., Belic, J., Jones, D.T.W., Tschida, B., Moriarity, B., *et al.* (2015). Somatic CRISPR/Cas9-mediated tumour suppressor disruption enables versatile brain tumour modelling. *Nat. Commun.* *6*, 7391.

Zuris, J.A., Thompson, D.B., Shu, Y., Guilinger, J.P., Bessen, J.L., Hu, J.H., Maeder, M.L., Joung, J.K., Chen, Z.-Y., and Liu, D.R. (2015). Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat. Biotechnol.* 33, 73–80.

Résumé de la thèse

Les tumeurs sont généralement constituées de différentes sous-populations de cellules cancéreuses génétiquement hétérogènes, responsables en grande partie de la capacité de la tumeur à évoluer rapidement et à s'adapter aux conditions environnementales. Cette diversité génétique a des conséquences majeures pour le patient, notamment au cours de la progression tumorale et pour l'acquisition d'une résistance aux traitements.

Nous avons développé une nouvelle stratégie basée sur la technologie CRISPR/Cas9 qui consiste à introduire, en plus d'une altération de séquence voulue d'un gène d'intérêt, une série de mutations silencieuses, constituant une sorte d'étiquette génétique qui peut être détectée par PCR quantitative ou séquençage de nouvelle génération. En parallèle, un code-barres constitué exclusivement de mutations silencieuses est utilisé comme contrôle interne pour les effets non-spécifiques potentiels qui peuvent être engendrés suite au clivage hors-cible par le système CRISPR/Cas9. Cette approche, que nous avons appelée CRISPR-barcoding, permet de générer et de suivre l'émergence d'un petit groupe de cellules cancéreuses contenant une mutation voulue au sein d'une population de cellules non modifiées, représentant ainsi un nouveau modèle expérimental d'hétérogénéité génétique intra-tumorale. Grâce à une série de preuves de concept, nous avons montré que CRISPR-barcoding est une nouvelle approche qui permet d'étudier de façon simple et rapide les conséquences fonctionnelles de différents types de modifications génétiques apportées directement au niveau de la séquence génomique.

Dans la deuxième partie de ma thèse, nous avons utilisé cette nouvelle approche pour l'étude de la résistance du cancer bronchique non à petites cellules (CBNPC) à la thérapie ciblée. Les patients de CBNPC dont la tumeur présente une mutation activatrice de l'epidermal growth factor receptor (EGFR) sont généralement traités avec des inhibiteurs de ce récepteur. Malheureusement, malgré une réponse initiale, presque invariablement ces patients rechutent, suite au développement d'une résistance causée, dans la majorité des cas, par l'apparition de mutations secondaires ou tertiaire de l'EGFR. Grâce à un criblage réalisé en utilisant un modèle de CBNPC basé sur la stratégie CRISPR-barcoding, nous avons pu identifier l'inhibiteur multikinase sorafenib pour sa capacité à prévenir la résistance de ces cellules aux inhibiteurs d'EGFR. Ce composé présente un mécanisme d'action original, impliquant une diminution précoce du niveau de phosphorylation de STAT3, suivie par une baisse considérable de l'expression de l'EGFR, aboutissant à une inhibition des voies intracellulaires en aval de ce récepteur, telles que RAS-MAPK et PI3K-AKT-mTOR. Ces données ont été confirmées in vivo en utilisant un modèle de xélogreffe de cellules de CBNPC modifiées par CRISPR-barcoding.

En conclusion, l'ensemble de nos résultats montre que le sorafenib peut prévenir l'émergence de cellules de CBNPC résistantes aux inhibiteurs d'EGFR, indiquant que ce composé pourrait représenter une nouvelle stratégie thérapeutique pour le traitement de ce type de tumeur.

Mots clés : CRISPR-barcoding - Hétérogénéité intratumorale - Cancer bronchique non à petites cellules - Résistance à la thérapie ciblée – EGFR - Sorafenib.

