

## UNIVERSITÉ TOULOUSE III PAUL SABATIER FACULTÉ DES SCIENCES PHARMACEUTIQUES

ANNÉE 2017

THESE 2017/TOU3/2010

## **THÈSE**

#### POUR LE DIPLÔME D'ÉTAT DE DOCTEUR EN PHARMACIE

Présentée et soutenue publiquement

par

Jean-Baptiste ROUBLIN

CRISPR/Cas9: HISTOIRE, METHODE, POTENTIEL & IMPACT

Le 24 Mars 2017

Directeur de thèse : Madame le Professeur Bettina COUDERC

**JURY** 

Président : Pr.Angelo PARINI

Assesseur: MCF. Isabelle LAJOIE-MAZENC

Assesseur: Dr. Pierre POUJOL



#### PERSONNEL ENSEIGNANT de la Faculté des Sciences Pharmaceutiques de l'Université Paul Sabatier au 1<sup>er</sup> octobre 2015

#### Professeurs Émérites

M. BASTIDE R M. BERNADOU J

M. CAMPISTRON G M. CHAVANT L Mme FOURASTÉ I M. MOULIS C

M. ROUGE P

Pharmacie Clinique Chimie Thérapeutique

Physiologie Mycologie Pharmacognosie Pharmacognosie Biologie Cellulaire

#### Professeurs des Universités

Mme BARRE A

#### Hospitalo-Universitaires

M. CHATELUT E
M. FAVRE G
M. HOUIN G
M. PARINI A
M. PASQUIER C (Doyen)
Mme ROQUES C
Mme ROUSSIN A
Mme SALLERIN B
M. SJÉ P
M. VALENTIN A

Pharmacologie
Biochimie
Pharmacologie
Physiologie
Bactériologie - Virologie
Bactériologie - Virologie
Pharmacologie
Pharmacologie
Pharmacie Clinique
Hématologie
Parasitologie

#### Universitaires

Mme BAZIARD G
Mme BENDERBOUS S
M. BENOIST H
Mme BERNARDES-GÉNISSON V
Mme COUDERC B
M. CUSSAC D (Vice-Doyen)
Mme DOISNEAU-SIXOU S
M. FABRE N
M. GAIRIN J-E
Mme MULLER-STAUMONT C
Mme NEPVEU F
M. SALLES B
M. SÉGUI B
M. SOUCHARD J-P
Mme TABOULET F

M. VERHAEGHE P

Biologie Chimie pharmaceutique Mathématiques - Biostat. Immunologie Chimie thérapeutique Biochimie Physiologie Biochimie Pharmacognosie Pharmacologie Toxicologie - Sémiologie Chimie analytique Toxicologie Biologie Cellulaire Chimie analytique Droit Pharmaceutique Chimie Thérapeutique

#### Maîtres de Conférences des Universités

#### Hospitalo-Universitaires

# M. CESTAC P Mme GANDIA-MAILLY P (\*) Mme JUILLARD-CONDAT B M. PUISSET F Mme SÉRONIE-VIVIEN S Mme THOMAS F

Pharmacie Clinique Pharmacologie Droit Pharmaceutique Pharmacie Clinique Biochimie Pharmacologie

#### Universitaires

Mme ARELLANO C. (*)	Chimie Thérapeutique
Mme AUTHIER H	Parasitologie
M. BERGĖ M. (*)	Bactériologie - Virologie
Mme BON C	Biophysique
M. BOUAJILA J (*)	Chimie analytique
Mme BOUTET E	Toxicologie - Sémiologie
M. BROUILLET F	Pharmacie Galénique
Mme CABOU C	Physiologie
Mme CAZALBOU S (*)	Pharmacie Galénique
Mme CHAPUY-REGAUD S	Bactériologie - Virologie
Mme COSTE A (*)	Parasitologie
M. DELCOURT N	Biochimie
Mme DERAEVE C	Chimie Thérapeutique
Mme ÉCHINARD-DOUIN V	Physiologie
Mme EL GARAH F	Chimie Pharmaceutique
Mme EL HAGE S	Chimie Pharmaceutique
Mme FALLONE F	Toxicologie
Mme FERNANDEZ-VIDAL A	Toxicologie
Mme GIROD-FULLANA S (*)	Pharmacie Galénique
Mme HALOVA-LAJOIE B	Chimie Pharmaceutique
Mme JOUANJUS E	Pharmacologie
Mme LAJOIE-MAZENC I	Biochimie
Mme LEFEVRE L	Physiologie
Mme LE LAMER A-C	Pharmacognosie
M. LEMARIE A	Biochimie
M. MARTI G	Pharmacognosie
Mme MIREY G (*)	Toxicologie
Mme MONTFERRAN S	Biochimie
M. OLICHON A	Biochimie
M. PERE D	Pharmacognosie
Mme PORTHE G	Immunologie
Mme REYBIER-VUATTOUX K (*)	Chimie Analytique
M. SAINTE-MARIE Y	Physiologie
M. STIGLIANI J-L	Chimie Pharmaceutique
M. SUDOR J	Chimie Analytique
Mme TERRISSE A-D	Hématologie
Mme TOURRETTE A	Pharmacie Galénique
Mme VANSTEELANDT M	Pharmacognosie
Mme WHITE-KONING M	Mathématiques

<sup>(\*)</sup> titulaire de l'habilitation à diriger des recherches (HDR)

#### Enseignants non titulaires

#### Assistants Hospitalo-Universitaires

Mme COOL C Physiologie
Mme FONTAN C Biophysique
Mme KELLER L Biochimie
Mme PALUDETTO M.N (\*\*) Chimie thérapeutique
M. PÉRES M. Immunologie

M. PERES M. Immunologie

Mme ROUCH L Pharmacie Clinique

Mme ROUZAUD-LABORDE C Pharmacie Clinique

<sup>(\*\*)</sup> Nomination au 1er novembre 2015

#### REMERCIEMENTS

À Mme Bettina Couderc pour m'avoir fait l'honneur de diriger ce travail, pour vos conseils avisés et la revue minutieuse de ce texte. Veuillez trouver ici le témoignage de mon plus profond respect et de ma plus grande reconnaissance

À Mr Angelo Parini pour avoir accepté la présidence du Jury ainsi que pour son enseignement. Veuillez trouver ici l'expression de ma sincère reconnaissance

À Mme Isabelle Lajoie-Mazenc et Mr Pierre Poujol, je vous remercie de me faire l'honneur de siéger à cette thèse et d'accepter de juger ce travail. Soyez ainsi assurés de ma profonde gratitude et de ma considération la plus sincère

À mes grands-parents Jacques, Yvette, Robert et Annette, pour les sentiments de curiosité, de magie et d'émerveillement que vous avez apporté à mon enfance et qui perdurent aujourd'hui

À mes parents, pour leur soutien et leur amour, à tous nos bons moments. À ma sœur, pour tous les rires, toutes les bêtises et aussi toutes les chaussettes que nous avons partagé et que nous partagerons. T'es la meilleure petite patate du monde

#### À mes cousins et au reste de ma famille

**Au Club des Copains & Associés**. Des études supérieures sont un long parcours, avec des passages désagréables mais aussi une multitude d'occasions de rire et de s'amuser. Vous avez rendu les mauvais moments surmontables et les bons, inoubliables.

Je souhaite donc remercier sans ordre précis : Pépito (le chien le plus stylé), ses heureux propriétaires : Romain S. et Marion, Diquy, Marc, Louisa, Pim's, Sylvain, la tante à Sylvain pour le prêt de son super appart, Pierre, Caroline, Rocky et Flo, Pauline, Romain F., Sarah et leur super nouvelle venue Hortense, Anthony, Brice, Askold, Camille, Arnaud et évidemment : Terrine, la meilleure mascotte qu'une soirée puisse avoir

#### À ceux qui m'ont accompagné pendant les périodes de ma scolarité hors de Toulouse :

Les membres de l' IFMT, les autres stagiaires, Bo, Bee et leur famille à Vientiane, ma douzaine de colocataires de Notelaarstraat et mes collègues belges, ma promo de Master et mes co-équipiers et co-stagiaires à Paris. Je n'aurai jamais cru que trois destinations aussi pluvieuses pourraient autant me plaire, c'est en grande partie grâce à vous

## INDEX

INTRODUCTION	
PARTIE I : HISTOIRE ET MÉTHODE	19
I HISTORIQUE	21
1. Introduction: CRISPR, un historique complexe	21
2. Génie Génétique : l'histoire parallèle	22
3. Découverte de répétitions palindromiques chez les bactéries et archées	23
3. Baptême	24
4. Découverte des gènes Cas	25
5. Provenance des Spacers (espaceurs): le déclic	25
5.1 Équipe de Francisco Mojica à l'Université d'Alicante	26
5.2 Équipe d'Alexander Bolotin à l'INRA	
5.3 Équipe de Christine Pourcel à l'Université Paris XI	27
6. Preuves expérimentales du S.I.A conféré par CRISPR-Cas	28
7. Classification initiale et regroupement des gènes Cas en familles	
8. Découverte des crARNs et de leur maturation	30
9. CRISPR cible l'ADN	31
10. La protéine Cas9, guidée par les crARN clive le double brin d'ADN cible	32
11. Une limite naturelle : le PAM ou Protospacer Adjacent Motif	33
12. Découvertes des tracrARN du système CRISPR-Cas9	
13. Simplicité du système CRISPR-Cas9 et rôle du tracrARN	
14. Résumé des connaissances microbiologiques	
15. Reconstitution de CRISPR-Cas9 dans un organisme étranger	
16. Étude de la Cas9 in vitro	
18. Du procaryote à l'eucaryote	40
19. Vers une commercialisation	41
20. CRISPR : popularité explosive	42
II MÉCANISME DÉTAILLÉ	43
1. Introduction	
2. Les acteurs	43
3. Déroulement de l'action au niveau macromoléculaire	44
4. Relation Structure-Activité	44
4.1 : Le gène Cas9	44
4.2 : La protéine Cas9	
4.3 : Liaison de la protéine à l'ARN guide puis à l'ADN	46
4.4 : Structure du complexe ARN guide : ADN cible	
4.5 : Techniques d'observations	47
5. Liaison Hors-Cible et spécificité	48
III OPTIMISATION, VARIANTES et ASTUCES	49
1. Introduction	
1.1 Utilisation simple en édition génomique	49
1.2 Rôle du cycle cellulaire	
1.3 Utilisations « sophistiquées »	
2. modification de l'ARN guide	
2.1 le sgARN	
2.2 CRISPR Multiplex	
3. Modifications de la Cas9	
3.1 Introduction	51

3.2 nCas9 ou Cas9 nickase	51
3.3 dCas9 ou dead Cas9 : Crispr Interference	52
3.4 dCas9 Fusion	52
3.4.1 dCas9 Fusion régulatrice	
3.4.2 dCas9 Fusion en imagerie	
3.5 dcas9 fusion et purification des acides nucléiques	54
3.6 Split Cas9	54
4. Cibler l'ARN : la RCas9	54
5. Plusieurs Cas9 pour plusieurs actions	55
6. S'émanciper de l'HDR: Homology Independent Targeted Integration	56
7. Conclusions	
7.1 Amélioration continue	56
7.2 Vers de nouvelles protéines effectrices	56
IV CRISPR EN PRATIQUE	
1. Introduction	59
2. Étude, sélection, synthèse, vectorisation, validation	59
2.1 Étude de la cible	
2.2 Sélection du gène à manipuler	
2.3 Sélection des gARN basée sur l'activité "on target" et "off target"	
2.4 Synthèse et clonage des gARNs	
2.5 Vecteurs	
2.6 Validation	
PARTIE II : POTENTIEL ET IMPACT	65
Avant-Propos	
I Thérapie génique	69
1. introduction	69
2. Techniques : du transfert de gènes à l'édition génomique	69
3. Approches	
4. Maladies monogéniques	
4.1 Annotation des gènes	
4.2 Gènes et maladies génétiques	
4.3 CRISPR-Cas9 et maladies monogéniques	
4.3.1 Exemple de la mucoviscidose : organoïdes et iPSCs	72
4.3.2 Exemple de la tyrosinémie : modèle in vivo, animal adulte	
4.3.3 Exemple de la cataracte : modification de zygotes de souris	73
4.3.4 Première modification d'embryons humains : β-Thalassémie	73
4.4 Aperçu d'autre travaux	73
4.5 Conclusion	74
5. CRISPR et Latence Virale	75
5.1 introduction	75
5.2 HBV : Hepatitis B Virus	75
5.2.1 Introduction	75
5.2.2 Cycle de vie	76
5.2.3 CRISPR comme solution thérapeutique	
5.2.4 Conclusions	
5.3. Herpesviridae	
5.3.1 Introduction	
5.3.2 CRISPR et Herpesviridae	
5.3.3 Conclusions	
5.4. Soigner HIV	79

5.4.1 Présentation	79
5.4.2 Cycle viral	80
5.4.3 Traitement et latence	80
5.4.4 Cibles : effacer le génome HIV1	81
5.4.6 Première étude in vivo	82
5.4.7 Limites	82
5.4.7.1. liées à l'outil d'édition génomique	82
5.4.7.2. liées au HIV	83
5.4.8 Conclusions	84
6. Cancérologie	84
6.1 Introduction	
6.2 immunomodulation anticancéreure et édition génomique	85
6.3 débuts cliniques : immunothérapie antitumorale et CRISPR-Cas9	85
6.4 Conclusions	86
6.5 Perspective: Thérapie génique CRISPR, production et commercialisation	86
II Des antibiotiques CRISPR	89
1. Introduction	89
2. La technique CRISPR-Cas9: espoir d'un nouveau type d'antibiotiques	90
3. exemples	90
3.1 Travaux de l'équipe de Citorik	91
3.2 Travaux de l'équipe de Bikard	
4. Conclusions	93
4.1 liés au développement d'antibiotiques CRISPR	93
4.2 Perspective : la résistance aux antibiotiques	
4.2.1 Impact	
4.2.3 Modèle économique et Innovation	
4.2.4 Solution à long terme : usage raisonné	94
4.2.3 Un problème mondialisé	95
IV Contrôle biologique : exemple du paludisme	97
1. Gene Drive	97
1.1 Origine	97
1.2 Le gene drive	97
1.3 Renouveau de la génétique dirigée	98
2. Gene drive et CRISPR: exemple de l'éradication du paludisme	99
2.1 Le paludisme	99
2.2 Gene drive : une alternative au traitement du paludisme	
2.3 Conclusions	
3. Autres applications de la génétique dirigée	101
3.1 Santé Publique	101
3.2 Agriculture	101
3.3 Recherche fondamentale	101
4. Risques	101
4.1 Écosystèmes et dissémination	101
4.2 Risque d'usage double	102
4.3 Risques de santé publique	102
5. Gene Drive : Des problématiques nouvelles	103
5.1 Introduction	103
5.2 Notion d'équité	103
6. Garde fous : des règles et des techniques	103
6.1 le Rapport NIH	103

6.2 Méthodes biotechnologiques	104
7. Conclusions	
V Plantes et agronomie	105
1.introduction	105
2.Exemples	106
2.1 plantes	106
2.2 Animaux	106
3. Plantations et acteurs industriels	106
4. Régulation de CRISPR en biotechnologie agricole	107
4.1 Procédé vs. Produit	
4.2 Le cas du champignon de paris qui ne brunissait pas	107
4.3 Vers une évolution des régulations	
5. Une action théorique	110
6. Conclusions	110
VI Outils de Recherche	111
1. Introduction	111
2. Screening / criblage	111
2.1 Présentation	111
2.2 Futur	112
3. Création de modèles	112
4. GWAS et neuroscience	113
VI PANORAMA	115
1.Introduction	115
2. Propriété de la technique CRISPR-Cas9	115
2.1 Introduction	
2.2 Les découvertes multiples	115
2.3 Combat catégorie Poids Lourds : Patent Interference No. 106,048 330	
2.3.1 Contexte	
2.3.3 Perspective: innovation et Cpf1	118
3. Émergence d'un marché de l'édition génomique	118
4. émergence de réflexions sur l'usage clinique de l'édition génomique	
4.1 Introduction	119
4.2 Une discussion globale	119
4.3 Perspective et approche culturelle	121
5. Au sujet de notre compréhension des mécanismes génétiques	122
5.1 Introduction	122
5.2 Exemple du lien gène/organisme/environnement	122
5.3 Au sujet de l'édition génomique et des maladies multigéniques	
5.4 Au sujet de l'édition génomique et des avancées futures	
CONCLUSION	
ANNEXES	
BIBLIOGRAPHIE	137

#### Liste des abréviations :

a.a: acides aminés

AAV: adeno associated virus

ADA-DICS: déficit immunitaire combiné sévère lié à la

déficience en adénosine déaminase ADN : acide désoxyribonucléique ARN : acide ribonucléique

**ARNi** : ARN interférence

**BLAST**: basic local alignment search tool **BPFs**: bonnes pratiques de fabrication **CAR**: Chimeric Antigen receptor **Cas**: CRISPR associated gene

 $\textbf{Cascade}: \textbf{CRISPR-} associated \ complex \ for \ antiviral$ 

defense

**CccADN**: covalently closed circular ADN

CCR5: cc chemokine receptor 5

CFRB: coordinated framework for regulation of

biotechnology

**CFTR**: cystic fibrosis transmembrane regulator

**CFU**: colony forming unit

**ChIP**: chromatin immunoprecipitation

Crarn: Crispr Arn

**CRISPR**: Clustered Regularly Interspaced Short

**Palindromic Repats** 

Csn1: nom donné à Cas9 sous l'ancienne

nomenclature

CTD: C terminal domain

**DB**: double brin **DCas9**: dead Cas9

DISC: déficit immunitaire combiné sévère

**Dsb**: double stranded break **EBV**: epstein-Barr virus

EMA: european medicine agency

EnCHIP: engineered binding molecule mediated

chromatin immunoprecipitation **ESC**: embryonic stem cells

**ESCRT**: endosomal sorting complex required for

transport

FACS: fluorescence activated cell sorting

gARN: ARN guide

GFP: green fluorescent protein

**GWAS**: genome wide association studies

**HBV**: virus de l'hépatite B **HCMV**: cytomégalovirus humain **HDR**: homology direct repair **HEG**: homing endonuclease gene

Hiti: Homology independent targeted integration

HIV : virus de l'immunodéficience humaine

HMM: modèle de markov caché
HR: homologous recombination
HSV: herpes simplex virus
Indels: insertions/deletions

INRA: institut national de la recherche agronomique

**INSTI**: integrase strand transfer inhibitor **IPSC**: induced pluripotent stem cell

ITOP: (induced transduction by osmocytosis and

propanebetaine)

Kb: kilobases

**LCTR**: large cluster of twenty nucleotides repeats

sequences

**LEDGF**: lens epithelium derived growth factor

**LTR**: long terminal repeats **MCR**: multiple chain reaction

Mcr-1: mechanism of colistin resistance 1

**MDR**: multi drug resistance **MGE**: mobile genetic elements

MIT: massachusetts institute of technology

NcARN: ARN non codant

NCBI: national center for biotechnology information

**NGS**: next generation sequencing **NHEJ**: non homologous end joining **NIH**: National Institute of Health

**NNRTI**: non nucleoside reverse transcriptase inhibitor

**NPBT**: new plants breeding techniques

**NRTI** : nucleoside reverse transcriptase inhibitor **OGM** : organisme génétiquement modifié

**PAM**: protospacer adjacent motif **PCR**: polymerase chain reaction **PD1**: programmed cell death 1

**Pdb**: paires de bases

**PIC**: pre integration complex

PNA FISH: peptide nucleic acid fluorescent in situ

hybridization

**R&D**: recherche et dévelopement

**RAAV**: adeno associated virus recombinant **RAAV**: adeno associated virus recombinant **Rcas9**: cas9 détournée pour couper l'ARN

RGN: ARN guided nuclease Rnase III: ribonucléase III RPE: Retinal Pigment Epithelium SaCas9; Staphylococcus aureus Cas9

**SgARN**: single guide ARN **ShARN**: short ARN

SIA: système immunitaire adaptatif

SiARN: silencing ARN

**SpCas9**: staphylococcus pyogenes Cas9 **SPIDR**: spacers interspaced direct repeats **SSODN**: single stranded DNA oligonucleotides

**STST**: short regularly spaced repats

TALENS: transcriptor activation-like effector

nucleases

TracrARN: trans activating CRISPR ARN

**UPSTO**: united states patent and trademarks office **USDA**: united state department of agriculture **USPTO**: united states patend and trademarks office

VZV : virus de la varicelle et du zona

**WT**: wild type

**ZFN**: zinc finger nucleases

## Index des figures

<b>fig.0</b> Un bactériophage en train d'infecter une bactérie est observé par ME	20
fig.1 Dr . Francisco Juan Martinez Mojica, à l'université d'alicante	23
fig.2 Représentation schématique d'un locus CR ISPR	24
fig.3 Extrait de la correspondance entre Ruud Jansen et Francisco Mojica au sujet du nom CRISP.	R 24
fig.4 Dr. Ruud Jansen	25
fig.5 Christine Pourcel	27
<b>fig.6</b> Illustration de la provenance des séquences <i>spacers</i>	28
<b>fig.7</b> Exemple de deux systèmes CRISPR distincts.	30
fig.8 John Van der Oost au centre et Stan Brouns tout à fait à droite.	30
fig 9 Illustration de la production des crARN à partir de la maturation des transcrits CRISPR	31
fig.10 llustration du mécanisme de défense de manière simplifiée	32
fig. 11 Emmanuelle Charpentier et l'une de ses collègues au Max Planck Institute de Berlin	34
fig.12 Les étapes du système immunitaire adaptatif CRISPR-Cas	36
<b>fig.13</b> L'équipe de Virginijus Siksnys : Giedrius Gasiunas, Tomas Sinkunas et Tautvydas Karvelis	37
fig.14 Emmanuelle Charpentier, Jennifer Doudna et leur collègues	38
fig .15 Illustration du détournement du S .I.A bactérien au profit de l'édition génomique	39
f <b>ig.16</b> Feng Zhang	40
fig.17 Évolution des recherches google pour les acronymes des techniques d'édition génomique	42
fig.18 Nombre de manuscrits par an contenant les mots clefs CRISPR ou TALENs ou ZFN	42
<b>fig.19</b> Schéma du complexe Cas9 : sgARN : ADN cible	43
<b>fig.20</b> Gène codant spCas9	44
<b>fig .21</b> Structure tertiaire représentée en ruban de la spCas9	45
<b>fig.22</b> Représentation de la surface de SpCas9	45
fig.23 Illustration des réarrangements de la Cas9 après contact avec le guide ARN	46
<b>fig.24</b> Structure du complexe formé par la protéine Cas9, l'ARN guide (en orange) et l'ADN cible	46
<b>fig.25</b> Structure cristalline du complexe formé par sgARN et l'ADN cible contenant le motif PAM	47
fig.26 Mécanismes impliqués dans la réparation d'une coupure double brin de l'ADN	49
<b>fig 27</b> Représentation des différents guide ARN : le duplex tracrARN:crARN et le sgARN	50
fig.28 Représentation schématique de la Cas9 nickases et du système double nickase	52
fig.29 Illustration des construction basée sur la création de dCas9 chimériques	52
fig.30 L'imagerie par CRISPR permet de détecter la longueur des télomères	53
<b>fig.31</b> Potentielles applications de la Rcas9	55
fig.32 Exemple de diagnostic opéré par un outil de conception de gARN	60
fig.33 exemple d'une cassette d'expression d'un sgARN	61
fig. 34 Approches Ex Vivo vs. In Vivo	70
fig.35 Illustration du protocole dévellopé par Schwant et al.	72
<b>fig.36</b> Cycle de vie de HBV et cibles thérapeutiques	76
fig.37 Schéma de la stratégie de suppression d'HBV par CRISPR-Cas9	77
<b>fig.38</b> Cycle viral de HIV-1 divisé en étapes	80
fig.39 Timothy Brown en 2011	81
fig.40 Répartition des indications en thérapie génique	84
fig.41 Assemblée Générale des Nations Unies, 13 Septembre 2016	89
fig.42 Mécanisme de l'antibiotique CRISPR conçu par l'équipe de Rob Citorik	91
<b>fig.43</b> Démonstration de l'activité souche dépendante de l'antibiotique CRISPR	92
<b>fig.44</b> détection de mcr-1 32 pays en juin 2016	95
fig.45 Mécanisme simplifié d'un Homing Endonucléase Gene	97
<b>fig.46</b> mécanisme d'introduction d'un <i>gene drive</i> via CRISPR-Cas9	98
fig.47 : Génétique dirigée par CRISPR-Cas9	100
<b>fig.48</b> Pourcentage des principales semences OGMs par rapport aux cultures conventionnelles	105
<b>fig.49</b> évolution du nombre de végétaux non régulés par l'USDA	108
<b>fig.50</b> Le premier repas <i>CRISPR</i> : Stefan Jansson de l'université Umeå en Suède	109
fig.51 Illustration du genome wide screening via la technique CRISPR/cas9	111
<b>fig.52</b> Utilisation de Cas9 pour générer des animaux génétiquement modifiés	113
fig 53 applications de la Cas9 associées aux cellules iPSCs humaines	114
<b>fig.54</b> Trois modèles d'ampoules à incandescence parmi les 22 développés	115
<b>fig.55</b> Extrait des conclusions des <i>Guidelines</i> ébauchées à Asilomar en 1975	119

Un gène est l'unité de base de l'information héréditaire. Il porte l'information nécessaire pour construire, maintenir et réparer un organisme. Les gènes interagissent avec d'autre gènes, avec la contribution de leur environnement, avec des événements déclencheurs, et avec une part de hasard pour produire la forme et la fonction ultime d'un organisme 1
Le génome est un terrain d'essai pour nos failles et nos désirs, bien que sa lecture ne nécessite pas la compréhension d'allégories ou de métaphores. Ce que nous lisons et écrivons dans notre génome, ce sont nos failles, nos désirs et nos ambitions. C'est la nature humaine <sup>1</sup>
Le sentiment de mystère est le plus beau qu'il nous soit donné d'éprouver, il est à l'origine de tout art et de toute science <sup>2</sup>

#### INTRODUCTION

La manière la plus simple d'expliquer ce qui se trouve derrière l'acronyme difficilement prononçable CRISPR-Cas9 est : un ciseau génétique. Ce dernier a la possibilité de couper in vivo ou in vitro la séquence nucléique que l'on souhaite et donc d'éditer les gènes. Mais ce ciseau génétique est plus complexe encore que son nom ne le laisse supposer. Son origine se trouve dans l'étude de micro-organismes présents dans des zones chaudes de la méditérranée espagnole. Son potentiel est à la fois immense et terrifiant, il aurait le pouvoir d'effacer des espèces entières tout en éradiquant en même temps un fléau de l'humanité qu'est le paludisme. Ce ciseau serait capable de découper le genome du virus HIV intégré dans le genome d'une cellule, de choisir la bactérie à laquelle il s'attaque, d'offrir des espoirs de thérapie pour des maladies monogéniques, d'accélérer nos outils de recherche et de modifier les semences agricoles en contournant la réglementation actuelle sur les OGMs. La technique progresse à une vitesse telle qu'elle bouleverse les cadres règlementaires, questionne les paradigmes éthiques et fais engager des sommes financières immenses pour sa propriété. Elle est utilisée pour la première fois sur l'embryon humain et en clinique en 2016. Elle est à la fois ce que l'homme désire et craint, constituée de manière si simpliste qu'elle peut sembler évidente : une protéine qui guidée par un ARN va être capable de découper, à souhait l'acide nucléique devenant ainsi l'outil d'édition le plus simple et le moins coûteux sur le marché. Le tout est empaquetable dans le vecteur de votre choix pour être introduit dans la cellule qui vous intéresse, animale, végétale, bactérienne. Malgré son potentiel et les adjectifs gratifiants dont il fait l'objet dans la presse spécialisée ou non, l'outil n'est pas encore idéal. Des défauts, et des lacunes en terme de maîtrise et de compréhension persistent. Il représente cependant une avancée inédite en recherche fondamentale et appliquée.

La première partie retracera l'histoire de la découverte des CRISPR-Cas, et l'évolution des connaissances emmenant à son détournement comme outil d'édition génomique. Elle comprendra ensuite une étude des mécanismes moléculaires et une revue des astucieuses variations qui étendent son potentiel bien au delà du «copier/coller». Le dernier chapitre consistera en une explication simplifiée de son usage en laboratoire.

La seconde partie aura pour objectif d'explorer le potentiel de cette technique et d'envisager son futur, son impact et ses limites. Elle se basera sur une série d'exemples majoritairement thérapeutiques, chacun répondant à un contexte précis. Ils s'étendent de la thérapie génique à la cancérologie et l'infectiologie en passant par l'agronomie, la génétique dirigée et la recherche. La conclusion de cette partie se fera par un panorama, approfondissant certains thèmes abordés dans les exemples précédents, et évoquant l'innovation, les brevets, la régulation, la place de la technique sur le marché pharmaceutique et enfin ses limites et celles de notre propre compréhension.

PARTIE I : HISTOIRE ET MÉTHODE



**fig.0** Un bactériophage en train d'infecter une bactérie est observé par microscopie électronique. L'interface moléculaire entre un hôte et son parasite est une zone où l'évolution est accélérée, stimulant l'innovation biologique. C'est le cas de l'interface entre hôte bactérien et virus. Les gènes impliqués dans ce conflit se sont révélés être une mine d'or codant pour des enzymes dont l'activité peut être détournée par les humains au profit de la biotechnologie. La Cas9 est l'une d'entre elles. <sup>3</sup>

### **I HISTORIQUE**

#### 1. Introduction: CRISPR, un historique complexe

La découverte du système CRISPR, est un mélange de hasard, de curiosité scientifique et de volonté d'innovation. Elle est le fruit de la convergence entre des travaux ayant pour objectif l'édition génomique et des recherches plus exploratoires dont le but était de percer les mystères d'une étrange structure génomique présente dans la nature.

Les humains ont voulu modifier la vie depuis des milliers d'années. Ils l'ont d'abord fait par reproduction sélective dans le but de renforcer certains traits de plantes et d'animaux. Mais nous n'avons que récemment compris les mécanismes moléculaires de ces modifications. Cette compréhension a débuté par l'élucidation des mécanismes de l'hérédité avec les travaux de Mendel et Morgan<sup>4</sup>, puis il y a environ 60 ans elle s'est poursuivie avec la découverte de la structure en double hélice de l'ADN<sup>5</sup>. En identifiant

les éléments de base du code génétique, nous avons eu accès à une étude plus poussée du vivant. La génétique est devenue un nouveau champ d'étude, au début des années 1960s les scientifiques bombardaient des plantes avec des radiations dans l'espoir de causer des mutations aléatoires dans le code génétique et d'étudier leur impact. Puis au fur et à mesure les méthodes et les outils à disposition sont devenus plus précis.

À la fin des années 1970s les premiers séquençages génétiques apparaissaient. L'avancée et l'accessibilité de ces techniques de séquençages<sup>6,7</sup> permit peu à peu la réalisation de bases de données regroupant une grande partie des génomes de nombreux organismes, incluant l'homme. En parallèle l'utilisation et la découverte d'enzymes (ligases, polymérases et enzymes de restrictions) et de la PCR (*Polymerase Chain Reaction*)<sup>8</sup> ont permis d'isoler les gènes et d'y introduire des mutations, que ce soit *in vitro* au sein des cellules, ou chez des modèles animaux. Grâce à ces techniques, la compréhension des mécanismes génétiques s'est approfondie et le génie génétique s'est développé. Le système CRISPR et plus particulièrement le système CRISPR-Cas9 représente aujourd'hui l'outil le plus simple d'utilisation en génie génétique et il promet d'être l'un des plus spécifiques.

« Aujourd'hui l'édition génomique ressemble à une découpe au scalpel, dans les années 1960s par comparaison, ils tapaient sur les gènes à coups de marteau » 9 .

Qu'est ce que CRISPR-Cas? Les procaryotes (bactéries et archées) sont constamment sujets aux attaques virales. Ces virus appelés bactériophages évoluent rapidement et leur biomasse surpasserait celle des bactéries d'un facteur dix 10,111. Les procaryotes sont donc soumis à une pression sélective immense, leur évolution s'est donc logiquement accompagnée d'un ensemble de défenses innées pour repousser ces prédateurs. On a longtemps pensé qu'ils étaient dépourvus de mécanisme immunitaire adaptatif. Ce n'est que récemment qu'un nouveau système de défense a été découvert, il fut appelé le système CRISPR-Cas<sup>12</sup> (pour *Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR associated genes*). Le détournement de ce système se révélera être un excellent outil biotechnologique, mais son implication future dans le domaine du génie génétique ne fut pas évidente dès sa découverte.

#### 2. Génie Génétique : l'histoire parallèle

Sans les décennies de recherches pour aboutir à des outils de génie génétique, il est possible que la communauté scientifique soit passée à côté de l'utilisation de CRISPR-Cas9 pour modifier le génome d'espèces animales.

Lorsqu'à la fin des années 1980s les scientifiques émirent l'hypothèse de modifier les génomes de cellules vivantes de mammifères cela marquait une nouvelle ramification de la recherche biomédicale. Ils imaginèrent rendre possible la modification du génome en un endroit précis de cellules souches embryonnaires murines<sup>13</sup>. En découlerait selon eux la naissance de souris portant la modification génique souhaitée. Bien que révolutionnaires, ces processus étaient inefficaces, ils nécessitaient la sélection et le criblage de millions de cellules afin de trouver celles chez qui une recombinaison homologue aurait eu lieu, échangeant ainsi le gène présent avec la version modifiée apportée par l'expérimentateur.

Au milieu des années 1990s, des biologistes, découvrirent qu'en introduisant une cassure double brin au niveau d'un locus du génome (en utilisant une méganucléase, une endonucléase dotée d'un site de reconnaissance extrêmement peu fréquent), ils augmentaient énormément la fréquence de recombinaisons homologues, ainsi que son mécanisme alternatif de réparation le NHEJ (non homologous end joining)<sup>14</sup> Ils venaient de réaliser que pour obtenir une édition génomique reproductible il leur fallait une méthode efficace pour produire une cassure double brin à l'endroit voulu.

La première stratégie pour obtenir ces DSB (*Double Strand Break ou Coupure double brin*) fut d'utiliser des ZFN ou *Zinc-finger Nucleases*. Ces protéines sont composées de domaines de liaison à l'ADN à doigts de zinc fusionnées avec un domaine de clivage de l'ADN provenant d'une enzyme de restriction<sup>15</sup>. Des publications montrèrent ensuite que les ZFNs étaient utilisables pour réaliser des éditions par recombinaison homologue chez l'animal notamment chez la drosophile et la souris <sup>16,17</sup>. En 2005, c'est un groupe de Sangamo Biosciences qui rapporta la première modifications chez des cellules humaines. Ils étaient parvenus à corriger une mutation du gène IL2Ry porté par le chromosome X, une mutation responsable de DICS (déficit immunitaire combiné sévère). Avec un taux de 18% de cellules « *corrigées* » dont 7% pour lesquelles la correction eut lieu au niveau des deux chromosomes X <sup>68</sup> l'expérience était un succès.

Ces résultats encourageants portaient les espoirs d'une application clinique des ZFN. Cependant, la production de ZFN capable de reconnaître fidèlement des sites spécifiques s'avéra lente et pointilleuse. Une solution technique plus fiable vit le jour en 2009 après la découverte d'une nouvelle classe de protéines trans activatrices : les TALENs. Provenant de l'association artificielle entre une protéine d'une bactérie pathogène du végétal, le Xanthomonas<sup>19,20</sup> et un domaine nucléase, les *Transcriptor Activation-Like Effector Nucleases*. Ces protéines utilisent un assemblage précis de séquences modulables pour cibler des séquences ADN choisies. Mais cette approche nécessite encore un travail considérable. Elle impose de créer une nouvelle protéine et de la valider avant chaque nouvelle expérience. Bien qu'encore utilisé, ce processus limite aujourd'hui leur vaste adoption. L'outil qui détrône ces précédentes techniques, qu'on nomme CRISPR-Cas9 est lui aussi emprunté par la biotechnologie au règne bactérien. (la comparaison entre les différentes techniques est présentée en annexe 1)

#### 3. Découverte de répétitions palindromiques chez les bactéries et archées

Ce que l'on connaît aujourd'hui sous le nom de système CRISPR à été observé à la fin des années 1980 par des microbiologistes qui s'intéressaient à l'exploration du génome de bactéries et archées. Ce n'est pas son activité immunitaire qui fut mise en évidence en premier, mais plutôt l'intrigante structure génétique qui code ce système : les bactéries étudiées possédaient dans leur génome d'étranges répétitions<sup>21,22</sup>.

En 1987, Yoshizumi Ishino de l'université d'Osaka étudiait le gène *iap* d'*Escherichia coli*. Durant leurs efforts pour parvenir à le séquencer (En 1987 le séquençage n'avait pas atteint l'efficacité qu'on lui connaît aujourd'hui) ils trouvèrent des mystérieuses séquences de 29 nucléotides répétées espacées par 32 nucléotides. La dernière phrase de leur publication :

« Jusqu'à aujourd'hui, jamais aucune séquence similaire à celle-ci n'avait été observée chez les procaryotes, sa signification biologique est inconnue »<sup>23</sup>

En 1993, la publication de Francisco Mojica et ses collègues fut la première à relier les pièces du puzzle qui démontrerait l'existence d'un système CRISPR. Ils réalisaient des recherches en Espagne sur l'archée halophile *Haloferax mediterranei* présente dans les marais salants et les côtes proches de la ville d'Alicante. Ils constatèrent que l'enzyme de restriction PstI découpait le génome bactérien de manière différente en fonction de la concentration en sels du milieu de croissance. En étudiant les fragments découpés, ils observèrent pour la seconde fois de l'histoire scientifique des répétitions quasi identiques et presque palindromiques de séquences de 30 bases, séparées à chaque fois par environ 36 bases.<sup>24</sup>



fig.1 Francisco Juan Martinez Mojica, à l'université d'alicante 25

Ils enquêtèrent pour savoir si ces étranges structures étaient un cas isolé et découvrirent des répétitions similaires chez une espèce proche : *Haloferax Volcanii*. Après une revue de littérature ils tombèrent sur la publication d'Ishino et de ses collègues <sup>23,26</sup>. L'équipe répertoria alors toutes les publications faisant état de leur présence ainsi que leurs propres découvertes dans un document publié en 2000 <sup>27</sup>. Ils recensèrent ces répétitions ou leur transcrits chez une vingtaine de micro-organismes distincts, incluant *Escherichia coli, Clostridium difficile, Mycobacterium tuberculosis* ou encore *Yersinia pestis*<sup>26,28-31</sup>. Deux années plus tard le nombre d'espèce répertoriées doubla, grâce aux travaux de Ruud Jansen et de ses collègues de l'université d'Utrecht<sup>21</sup>. Ces scientifiques s'assurèrent également que ces répétitions étaient uniques aux bactéries et archées et ne se retrouvaient ni chez les eucaryotes, ni chez les virus.



Fig.2 Représentation schématique d'un locus CR ISPR.

Les cubes bleus représentent les séquences palindromiques similaires ou « *repeats* » . Elles sont séparées par les diamants colorés : les « *spacers* » qui représentent chacuns une séquence différente de longueur proche.<sup>32</sup>

Au fur et à mesure de leur recensement ils se rendirent compte que ces répétitions n'étaient pas un cas isolé mais une structure courante dans le génome de plusieurs

micro-organismes. On leur donna d'abord le nom de SPIDR pour *Spacers Interspersed Direct Repeats*. Le terme « *Locus* » fut utilisé pour nommer les zones où ces séquences étaient regroupées, et « *Spacer* » désigna les séquences séparant les courts enchaînements palindromiques. Leurs principales caractéristiques furent relevées :

« Au sein d'une espèce unique telle que Mycobacterium tuberculosis, on observe un unique locus SPIDR, tandis que d'autres espèces, comme M.Jannaschii, contiennent 20 loci différents. Au sein d'un locus, le nombre de répétitions varie grandement, allant de deux à plusieurs douzaines de répétitions. Les loci SPIDR sont encadrés par une séquence leader de 300-500pdb qui est conservée au sein d'une même espèce, mais ne l'est pas entre espèces différentes »<sup>33</sup>

Mais quel était le rôle de ces séquences? La curiosité scientifique des chercheurs impliqués était piquée à vif, c'était l'étincelle qui permettrait de déclencher de multiples recherches <sup>27</sup>.

#### 3. Baptême

Asunto: Re: Acronym
Fecha: Wed, 21 Nov 2001 16:39:06 +0100
De: "Ruud Jansen" <R.Jansen@vet.uu.nl>
Empresa: Diergeneeskunde
A: "Francisco J. Martinez Mojica" <fmojica@ua.es>

Dear Francis

What a great acronym is CRISPR.

I feel that every letter that was removed in the alternatives made it less crispy so I prefer the snappy CRISPR over SRSR and SPIDR. Also not unimportant is the fact that in MedLine CRISPR is a unique entry, which is not true for some of the other shorter acronyms.

fig.3 : Extrait de la correspondance entre Ruud Jansen et Francisco Mojica au sujet du nom «CRISPR» 34

En 2001, plusieurs publications commençaient à paraître au sujet de ces répétitions. Mais elles les nommaient toutes de manière différentes. SPIDR (Spacers Interspaced Direct Repeats) pour Jansen et ses collègues<sup>33</sup>, STST (Short Regularly Spaced Repeats) par Mojica et son équipe<sup>35</sup>, ou encore LCTR (large cluster of twenty nucleotides repeats sequences) pour She et associés<sup>36</sup>. Pour éviter toute confusion, c'est l'acronyme CRISPR qui a été choisi. Comme Ruud Jansen le fit remarquer par *e-mail* à Francisco Mojica:

« il avait l'avantage de ne pas être déjà présent dans la base de donnée Medline<sup>37</sup> et d'avoir une sonorité sympathique. »

#### 4. Découverte des gènes Cas

Dès 2002, à l'Université d'Utrecht, Jansen et son équipe firent une découverte qui permettrait de mieux comprendre la biologie des séquences CRISPR. En étudiant leur environnement ils observèrent qu'elles étaient toujours accompagnées par des familles de gènes qui étaient retrouvées seulement si les séquences CRISPRs étaient aussi présentes. Ces gènes seraient appelés Cas, pour *CRISPR Associated*.



fig.4 Ruud Jansen<sup>41</sup>

Initialement ils identifièrent quatre gènes (Cas 1 à 4), un gène qu'ils nommèrent Cas1 était toujours associé aux CRISPRs. L'étude des séquences des gènes Cas montra qu'ils codaient probablement pour des domaines hélicases et nucléases <sup>38-40</sup>. La protéine codée par le gène cas3 montrait des motifs caractéristiques des hélicases de la superfamille 2, tandis que la protéine Cas4 comportait des motifs de la famille RecB des exonucléases (des domaines destinés à délier et à couper les acides nucléiques). Ces observations suggéraient que les protéines codées pouvaient être impliquées dans le métabolisme de l'ADN ou l'expression génique. La proximité géographique des séquences Cas avec les séquences CRISPR et la nature de de leur produit amena à penser qu'il existait un lien entre les deux. Ils constitueraient le système CRISPR-Cas.

Une autre observation importante était la mobilité des éléments CRISPR. Ils notèrent

que dans les espèces comportant deux loci CRISPRs ou plus, ceux-ci sont flanqués par une séquence leader de 300 à 500pdb sur l'un des côtés. Les répétitions directes et la séquences leader sont conservées au sein d'une espèce bactérienne, mais différent entre espèces. Pour eux la présence de multiples locus chromosomiques CRISPR était un signe de la mobilité de ces éléments géniques.

Bien que des analyses bioinformatiques initiales prévoyaient l'implication du système Crispr-Cas dans le métabolisme des acides nucléiques <sup>33,42</sup>, le premier indice quant à leur véritable fonction viendra de la découverte de l'origine des *spacers*.

#### 5. Provenance des Spacers (espaceurs): le déclic

En 2005 trois articles vinrent élucider le rôle naturel du système CRISPR-Cas, chacun apportant des informations complémentaires. Francisco Mojica, Christine Pourcel, Alexander Bolotin et leurs équipes tentaient de comprendre l'origine de ces séquences. En tapant une à une les séquences dans le logiciel Word, puis en les soumettant à l'outil BLAST<sup>43</sup> (pour *Basic Local Alignment Search Tool*), ils purent les comparer à la bibliothèque des séquences enregistrées dans l'outil. Ils se rendirent compte que loin d'être toutes inconnues ces séquences correspondaient parfois à des éléments avec lesquelles les bactéries avaient des contact plus que fréquents.

#### 5.1 Équipe de Francisco Mojica à l'Université d'Alicante

L'équipe de Mojica se rendit compte de l'étrange similarité entre les séquences spacers et celles de virus connus pour avoir un tropisme bactérien. Ils constatèrent que la séquence d'un spacer qu'ils avaient découvert chez une souche d' Escherichia coli était identique à celle du phage P1 (la comparaison des séquences est présentée en Annexe 2). Le Phage P1 est un virus bactériophage connu pour infecter de multiples souches d'Escherichia coli. Un détail allait avoir une importance majeure : la souche bactérienne sur laquelle ils enquêtaient était connue pour être résistante à ce phage. Intrigués ils vérifièrent s'il ne s'agissait pas d'un hasard. Après avoir répété cette procédure pour 4500 séquences spacers, 2% étaient similaires à des séquences connues, parmi elles les deux tiers correspondaient à des séquences virales ou à celles de plasmides. Les virus connus pour infecter les bactéries testées étaient ceux pour lesquels les spacers avaient la correspondance la plus complète<sup>44</sup>. Selon eux :

« ces résultats pourraient impliquer une relation entre CRISPR et une immunité dirigée contre un élément envahisseur »  $^{45}$ 

Ils venaient d'identifier la première preuve d'un système immunitaire adaptatif procaryote. L'hypothèse d'une interférence, similaire à celle de l'ARN interférence, par association complémentaire entre transcrits CRISPR et éléments viraux fut proposée<sup>44</sup>, mais le mécanisme exact s'avérerait plus complexe.

#### 5.2 Équipe d'Alexander Bolotin à l'INRA

L'équipe de Bolotin réalisa quelques mois plus tard des observations chez

Streptococcus spp qui les conduisirent à des conclusions similaires<sup>38</sup>. Ils montrèrent par ailleurs que la longueur des loci CRISPR est inversement proportionnelle au nombre de phages infectant la bactérie qui les contient . En résumé, plus la zone CRISPR contient de spacers, plus la bactérie est protégée contre un grand nombre de virus. L'hypothèse reliant CRISPR-Cas et défense immunitaire se confirmait<sup>44</sup>. Cette équipe identifia deux nouveautés qui allaient par la suite prendre une importance majeure :

- La première était la structure inédite du système CRISPR-Cas chez *Streptococcus thermophilus*. Bien que ressemblant à ceux déjà observés, il ne possédait pas la plupart des multiples gènes Cas connus. À la place ils en comportait un nombre inférieur, l'un codant pour une grosse protéine dont l'activité prévue était une activité nucléase. Ils la nommèrent Cas9. Ils avaient découvert ce qui serait nommé dans la classification actuelle un système Crispr de type II, le plus simple connu aujourd'hui. C'est à dire celui ayant le moins d'éléments nécessaires à son fonctionnement, par opposition à ce que nous appelons un système de type I dont l'activité repose sur un grand nombre de gènes et protéines Cas.
- La seconde nouveauté était l'observation d'une séquence conservée présente dans le génome des éléments envahisseurs. Cette séquence serait plus tard nommée PAM pour *Protospacer Adjacent Motif* et son rôle se révélerait indispensable au mécanisme du système.

#### 5.3 Équipe de Christine Pourcel à l'Université Paris XI

La même année l'équipe de Christine Pourcel compléta leur hypothèse en se concentrant sur deux micro-organismes, *Yersinia pestis* et *Streptococcus pyogénes*. Leur idée était d'utiliser les très variables séquences CRISPRs afin d'obtenir un nouvel outil d'identification des différentes souches de ces organismes pathogènes. Au sein du génome de *Yersinia pestis*, trois locus CRISPR distincts peuvent être identifiés. L'un d'eux est très polymorphique (un génome d'une souche de Yersinia pestis utilisé dans leur étude est représenté en Annexe 3). Les auteurs séquencèrent au total 109 allèles des 3 locus CRISPRs de *Yersinia pestis*.

Celà leur permis d'identifier une évolution dynamique des séquences. Ils identifièrent de «vieux» spacers ainsi que des nouveaux. Un portrait de l'évolution de ces CRISPRs pouvait alors se dessiner. Ils montrèrent que l'addition de nouveaux motifs spacers se faisait de manière polarisée, c'est à dire proche d'une séquence leader (précédemment décrite par Jansen et son équipe)<sup>33</sup>. En terme d'activité, c'est à dire de nombre de nouveaux spacers acquis, le locus CRISPR le plus actif était majoritairement celui le plus proche des gènes Cas. Cette observation confirmait le rôle très probable des produits de ces gènes dans l'acquisition de nouveaux spacers et leur insertion.

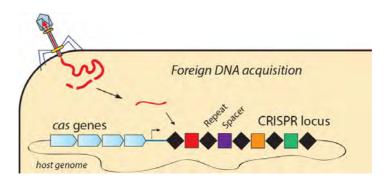
fig.5 Christine Pourcel



En reliant les séquences des *spacers* à celles de prophages, les auteurs suggéraient que ces système faisait parti d'un mécanisme de défense capable de s'approprier des séquences d'éléments génétiques étrangers. Pour eux :

« CRISPRs pourrait représenter la mémoire d'agressions génétiques passées »39

Ils montraient que ces séquences étaient collectées dans le but de lutter contre les virus dont elles provenaient en cas de réinfection. C'était un peu comme garder l'image de son ennemi dans sa poche, ou plus précisément dans le cas de la bactérie, le négatif de cette photographie. Leur publication dévoilait donc un ensemble d'échanges dynamiques entre génomes viraux et les systèmes CRISPRs des génomes bactériens. Il sera ultérieurement montré que d'anciens *spacers* sont perdus, et de nouveaux acquis, permettant ainsi à l'hôte de s'adapter dynamiquement aux infections par des virus nouveaux ou mutés. <sup>12,45</sup>



**fig.6** Illustration de la provenance des séquences *spacers*, la séquence virale correspondant à celle des *spacers* serait appelée : *Protospacer*<sup>46</sup>

Il s'agissait de découvertes surprenantes et excitantes, mais comme de nombreuse idées novatrices, l'obtention d'une reconnaissance par la communauté scientifique s'avéra longue et complexe. Les deux publications furent soumises à 3 ou 4 revues majeures, au même moment, en 2003. Mais elles déclenchèrent une vague de scepticisme de la part des rapporteurs. Lorsqu'ils étaient favorable aux idées démontrées, c'est le petit nombre (2%) de correspondances entre séquences virales et bactériennes qui empêchaient de les convaincre. Finalement après deux années, ces études furent publiées à quelques semaines d'écart en 2005. <sup>34</sup>

#### 6. Preuves expérimentales du S.I.A conféré par CRISPR-Cas

Les trois publications précédentes traçaient une carte qui menait à l'existence d'un système immunitaire adaptatif (ou S.I.A) chez les procaryotes, mais ce dernier n'avait pas encore été observé en activité. La première preuve *in vitro* de ce S.I.A vint d'un lieu qui peut sembler inattendu : des laboratoires de l'industrie des produits laitiers.

Employés par Rhodia Food, une filiale de grands groupes de l'industrie alimentaire produisant des cultures destinées à l'élaboration de yaourts, Phillipe Horvath et son équipe travaillaient sur la bactérie lacto-acidique *Streptococcus thermophilus*. Rodolphe Barrangou et lui n'avaient « *initialement aucune idée du rôle du système CRISPR* »<sup>47</sup>. Mais ils étaient curieux et se renseignaient activement sur les techniques capables de protéger leurs cultures des infections virales.

Se basant sur les publications citées précédemment ils voulaient utiliser cette défense naturelle pour immuniser leurs cultures. Leur technique consista à exposer la bactérie sélectionnée à un virus et à récupérer les souches qui survivaient à l'attaque. Puis toute l'astuce consista à identifier par séquençage si ces bactéries avaient acquis de nouveaux *spacers*. D'après Horvath:

«CRISPR n'est pas le seul mécanisme capable de conférer aux cultures une résistance aux phages mais il est celui qui peut être utilisé de manière efficiente »<sup>47,14</sup>

Ils observèrent l'immunité acquise en action. Ils constatèrent que CRISPR, en association avec les gènes Cas, conférait une résistance aux phages, la spécificité de celle-ci étant déterminée par la similarité entre les séquences *spacers* et celles des phages. Le retrait ou l'ajout d'un *spacer* pouvait modifier l'identité de cette résistance<sup>14</sup>. (Un exemple de leur travaux est présenté en Annexe 4)

Ils étudièrent également le rôle de deux gènes : Cas 9 et Cas 7. Les bactéries ont besoin de Cas7 pour acquérir une résistance, mais dès qu'elles possèdent dans leur système CRISPR un *spacer* provenant d'un phage, elles n'en ont plus besoin pour rester résistantes. Cela suggère que Cas 7 est impliqué dans la création de nouveaux *spacers* et répétitions, mais pas dans l'immunité elle-même. Par contre le gène Cas9, dont la séquence code deux motifs nucléases (HNH et RuvC) et dont l'aboutissement est donc de coder pour une protéine Cas9 catalysant la coupure de liaisons nucléiques, est nécessaire pour résister aux phages et représente donc un composant actif de l'immunité bactérienne. <sup>38,48</sup>

Finalement ils découvrirent que, si certains isolats de phages pouvaient surmonter cette immunité bactérienne, c'était que les séquences de leur génome qui correspondaient aux *spacers* (nommées : *protospacers*) variaient sensiblement , même d'une seule base. Ce système immunitaire dépend donc d'une correspondance précise entre les séquences des *spacers* et de la cible virale.

#### 7. Classification initiale et regroupement des gènes Cas en familles

En parallèle les recherches portant sur les gènes Cas se poursuivaient, au fur et à mesure que leur nombre grandissait, ils furent d'abord classifiés en familles. Dès 2002, en identifiant des similarités entre les séquences Cas, Jansen et ses collègues posèrent les bases de la classification permettant de les répartir en différents groupes. En 2002 la fonction exacte de ces séquences restait inconnue. Leur classification serait donc fondée sur la présence de gènes signatures<sup>33</sup>. Ils décrivirent en fonction de leur similarités 4 familles de gènes codant pour les protéines Cas. Ils montrent des régions hautement conservées codant pour des domaines fonctionnels. A ces quatre familles vinrent se rajouter de nouvelles. On se rendit compte à posteriori que certains gènes précédemment étudiés étaient en fait des gènes associés aux CRISPR. L'équipe de Kira Makarova au NCBI (*National Center for Biotechnology Information*) s'intéressant aux éléments génétiques conservés avait identifié dès 2002 une vingtaine de familles de gènes qui se révélèrent être des Cas.<sup>491</sup>

En 2005, le rôle des CRISPRs et la fonction des Cas commence à être comprise. C'est l'équipe de Daniel Haft à l'*Institute of Genomic Research* dans le Maryland qui, inspirée par les précédents travaux, identifie un total de 45 familles de gènes. Ils examinèrent toutes les séquences pouvant coder pour des protéines non référencées dont les gènes étaient au voisinage des CRISPR et décrivirent ces gènes comme :

« appartenant à différentes classes, comportant différents motifs, différents assortiments et différents tropismes d'espèces procaryotes».

Afin de donner un sens à l'hétérogénéité de ces séquences ils réalisèrent de multiples comparaisons par alignements de séquences, et utilisèrent des HMM, acronyme signifiant en français : modèle de Markov caché (Ces modèles mathématiques servant à identifier les familles avec une grande sensibilité et spécificité).<sup>42</sup>

Ils mirent alors la classification à jour et montrèrent qu'on pouvait décrire les familles de protéines Cas comme des « Guildes », c'est à dire un ensemble d'entités qui réalisent une tâche similaire. Ils rajoutaient un élément à la classification : celui de la fonction des protéines. Ces guildes seraient tantôt impliquées dans la maintenance des répétitions<sup>33</sup>, la capture de nouveaux *spacers*<sup>39,44</sup> ou l'expansion ou la contraction des ensembles CRISPR. En décrivant la multitude de guildes connues, ils s'aperçurent que les systèmes CRISPR-Cas existaient chez les procaryotes sous une très grande diversité de formes et de fonctions. L'association de protéines assurant leur fonction variant intensément en nombre et en taille. Voici un exemple adapté de la publication de Haft <sup>42</sup>



fig.7 Exemple de deux systèmes CRISPR distincts.

Streptococcus pyogenes et Escherichia coli possèdent chacun des groupes de gènes Cas différents associés aux séquences CRISPR (ici schématisés en noir sous le nom RPT). Ces gènes codent pour des protéines ayant des fonctions différentes. Ces protéines peuvent être rassemblées en guildes (partageant des rôles similaires). Ici c'est l'ancienne nomenclature qui est utilisé, Csn1 par exemple est désormais appelé Cas9.

En 2005, le rôle naturel de CRISPR-Cas en tant que système immunitaire adaptatif était établi. Les recherches réalisées avaient permis de montrer qu'il pouvait prendre plusieurs formes chez différentes espèces de procaryotes. L'étude des protéines qui le composent avait montré qu'elles réalisaient des fonctions variées mais un élément manquait pour comprendre son fonctionnement et parvenir à la réalisation de la classification que nous connaissons aujourd'hui : le mécanisme d'action.

#### 8. Découverte des crARNs et de leur maturation

En 2008 Stan Brouns, John Van der Oost et leurs collègues de l'université néerlandaise Wageningen réalisèrent une expérience intéressante<sup>50</sup>. Ils insérèrent le système CRISPR d'une souche d'*Escherichia coli* dans une autre souche d'*Escherichia coli* qui en était dépourvue. Par purification d'affinité et co-précipitation, ils caractérisèrent un complexe de 5 protéines Cas, aussitôt appelé Cascade (en terme de classification E.coli sera décrit comme le plus complexe des systèmes CRISPR de type I et de classe 1, le rôle de la cas9 est assuré par le complexe Cascade et Cas3)

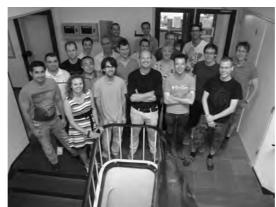


fig.8 John Van der Oost au centre et Stan Brouns tout à fait à droite.

En inhibant sélectivement chacun de ces composants, ils montrèrent que le groupe protéique Cascade est nécessaire au clivage du long précurseur d'ARN issu de la transcription du locus CRISPR en un ARN de 61 nucléotides. Les ARNs maturés obtenus seraient appelés les crARNs (pour CRISPR ARN). Ils montrèrent que ces crARNs commençaient tous par les 8 dernières bases des séquences répétées suivies par la séquence complète du *spacer* puis le début de la prochaine région répétée.<sup>50</sup>

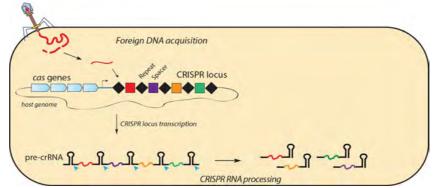


fig 9 Illustration de la production des crARN à partir de la maturation des transcrits CRISPR

Afin de prouver l'implication des crARNs dans l'immunité liée au système CRISPR, ils construisirent le premier système CRISPR partiellement artificiel. En créant un locus artificiel contenant 4 *spacers* de leur choix, chacun correspondant à une séquence connue du phage lambda. Comme prévu, les souches portant ces nouvelles séquences CRISPR développaient une résistance au phage lambda.

Les auteurs créèrent deux versions de leurs systèmes CRISPR artificiels, l'une anti-sens (complémentaire des ARNm et du brin codant), l'autre sens (complémentaire seulement du brin non codant). La version sens fonctionna de manière plus efficace, ce qui leur permis d'émettre une hypothèse nouvelle : la cible du système CRISPR pourrait être l'ADN.

#### 9. CRISPR cible l'ADN

La plupart des hypothèses initiales prévoyaient un mécanisme de défense par interférence ARN. Si elles étaient fausses, cela présageait l'existence d'un fonctionnement inédit. Dans une publication parue fin 2008, afin d'identifier avec certitude la cible du système CRISPR, Sontheimer et Marraffini utilisèrent une astucieuse méthode. Ils introduisirent un intron capable de s'auto-épisser au sein du gène cible du système CRISPR-Cas de *Staphylococcus epidermidis*. Cet ADN plasmidique n'était donc plus complémentaire de la séquence *spacer* portée par le génome bactérien. Par contre l'ARNm le serait car la séquence intronique serait éliminée par épissage et sa complémentarité serait restaurée. Cette astuce permettait de savoir lequel des deux acides nucléiques était la cible du système CRISR, leur résultats penchèrent en faveur de l'ADN<sup>51</sup>. Ils comprirent d'emblée l'importance de leur découverte :

« La capacité de provoquer une destruction guidée d'un ADN qui contient une séquence cible donnée pourrait avoir des applications considérables si le système est capable de fonctionner en dehors de son contexte bactérien ou archéen natif »

Comme nous le verrons, de multiples obstacles seraient à franchir avant d'avoir la possibilité d'isoler un système aussi complexe et de pouvoir l'appliquer ailleurs qu'au sein d'un micro-organisme. Il restait notamment à identifier un acteur principal de ce

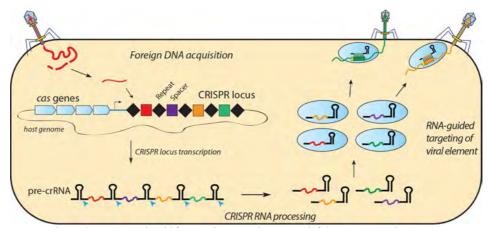
système de défense, celui qui empêchait les éléments envahisseurs d'agir. Cet acteur le ferait par un mécanisme jamais observé dans la nature : il allait couper l'ADN guidé par un court fragment d'ARN.

#### 10. La protéine Cas9, guidée par les crARN clive le double brin d'ADN cible

En 2010, au sein de son laboratoire de l'université de Laval Josiane Garneau collaborait avec l'entreprise agroalimentaire Danisco dont faisaient parti P.Horvath et R.Barrangou. Pour comprendre le fonctionnement de l'immunité CRISPR ils étudièrent *in vivo* l'interférence plasmidique chez *Streptococcus Thermophilus*. Leurs expériences reposaient sur l'introduction dans la bactérie de plasmides codant une résistance à un antibiotique suivie de l'exposition de ces colonies aux mêmes antibiotiques. Si elles perdaient leurs résistances c'était que le plasmide avait été détruit. En analysant le contenu des cellules ils se rendirent compte de la présence inattendue du plasmide sous forme linéarisée. En reproduisant l'expérience après inactivation des gènes Cas, seule la forme circulaire était présente. Lorsqu'ils séquencèrent les plasmides linéarisés, ils observèrent une coupure à bouts francs. Ils venaient de découvrir la clef du fonctionnement du système CRISPR, ils conclurent :

« l'ADN double brin est ciblé puis clivé à l'intérieur du protospacer par la machinerie CRISPR/Cas. »

Poursuivant leurs recherches, ils observèrent des résultats similaires concernant la coupure de l'ADN viral. De plus le nombre de *spacers* distincts correspondait au nombre de coupures relevées. Il serait montré que *S.Thermophilus* possède un système CRISPR de type II, où l'unique protéine Cas9 est responsable de la découpe de la cible. D'autre types CRISPR ont différentes manières d'organiser cette attaque, et associent souvent l'activité de plusieurs protéines Cas pour y parvenir. Par contre quelque soit le type de système, ils nécessitent tous une molécule d'ARN pour les guider : le CRISPR ARN ou crARN.



**fig.10** Illustration du mécanisme de défense de manière simplifiée: Durant l'acquisition une séquence d'un élément génique envahisseur est ajoutée aux locus CRISPR. Elle est enregistrée sous forme d'un *spacer* borné par des motifs répétés. Ces locus constituent ainsi une mémoire génétique des agressions passées. Après transcription, celui ci donnera un long transcrit (le crARN) qui une fois découpé donnera autant de crARN que de *spacers*. Chacun des crARN est complémentaire de la séquence virale ou plasmidique qui à été gardée en mémoire. Lorsque cet élément génique pénétrera à nouveau dans la cellule, il sera reconnu par les nucléases Cas (représentées par un cercle bleu) qui guidées par le crARN complémentaire de la séquence virale catalysera le clivage.

#### 11. Une limite naturelle : le PAM ou Protospacer Adjacent Motif

Les possibilités de découpe semblaient illimitées, à un détail près : le PAM. Il s'agit d'une limite naturelle à laquelle est soumise l'activité de découpe du système CRISPR. Cette séquence n'appartient pas au génome procaryote mais à celui de l'envahisseur (viral ou plasmidique). C'est cette limite qui fait que les protéines Cas seront par exemple capable de couper une séquence virale tout en épargnant la séquence *spacer* du génome bactérien, Ils permettent la différentiation entre le *soi* et le *non-soi* Il s'agit d'un court motif présent à proximité du protospacer <sup>71,72</sup>. Ces divers motifs seront appelés PAM pour *Protospacer Adjacent Motif* suite à leur mise en évidence chez de nombreux micro-organismes quelquesoit le type CRISPR<sup>52</sup>. Ces motifs PAMS sont indispensables à l'activité CRISPR-Cas<sup>19,21,22</sup> et représenteront donc une limite des futures applications biotechnologiques.

#### 12. Découvertes des tracrARN du système CRISPR-Cas9

Dans les années 2000s, les CRISPR étaient encore un domaine discret, seulement une poignée de microbiologistes s'intéressaient à la récente découverte de ces fragments répétés du génome bactérien. Emmanuelle Charpentier, était l'une d'entre eux, biologiste au département microbiologie et génétique de l'université de Vienne elle était spécialisée dans le domaine des ncARN régulateurs (de petits ARN non codants). Elle s'intéressait à l'identification des sites du génome de *Streptococcus pyogenes* qui codaient pour ces ARNs régulateurs. Son équipe et elle utilisaient un outil bioinformatique pour prédire la localisation de ces séquences, mais pas leur contenu. Elle collabora donc avec le microbiologiste Jörg Vogel du *Max Planck Institute for Infection Biology*, qui développait des méthodes de cartographie ARN à large échelle. En 2008 il parvint à séquencer l'ensemble des petits ARNs générés par *Streptococcus pyogenes*. <sup>53</sup>

Streptococcus pyogenes est tristement célèbre pour son rôle dans les angines à streptocoque et la fasciite nécrosante. Dans le domaine des CRISPR il était devenu connu par le fait qu'il possède un système CRISPR de type II, et donc la plus simple itération recensée de ces systèmes avec sa protéine effectrice : Cas9. Le mécanisme précis de ce système échappait encore aux chercheurs. En étudiant le transcriptome de la bactérie la première chose qu'ils observèrent était une sur-abondance d'un petit ARN nouveau qu'ils nommèrent *trans-activating CRISPR RNA* (tracrARN). Ce nom provenant du fait que la position de la séquence prédite par bioinformatique se trouvait proche du locus CRISPR-Cas9. Charpentier et ses collègues espéraient qu'il s'agissait d'un élément impliqué dans le système CRISPR de type II.

#### 13. Simplicité du système CRISPR-Cas9 et rôle du tracrARN

Comme nous l'avons vu la maturation du crARN à partir de son transcrit issu du locus CRISPR est l'une des étapes capitales du fonctionnement de ce système<sup>50</sup>. Jusqu'alors seulement trois protéines cas (Cse3, Cas6 et Csy4) dotées d'activité endoribonuclease et pouvant donc couper le précurseur ARN avaient été identifiées. <sup>50,54</sup> On ne connaissait pas encore leurs homologues capable d'effectuer la maturation du pre-CrARN pour de nombreux sous types du systèmes CRISPR Cas. Cette lacune

laissait penser qu'il existait certainement au moins une voie alternative impliquant des protéines Cas encore inconnues ou des mécanismes différents<sup>53</sup>. La découverte du tracrARN viendrait combler ce vide pour le système CRISPR de type II et plus encore.

Emmanuelle Charpentier s'était relocalisée au *Umeå Centre for Microbial Research* dans le nord de la Suède. Porté par Elitza Deltcheva, étudiante en Master résidente de l'ancien laboratoire viennois, l'enquête sur le rôle du tracrARN se poursuivait. Ils débutèrent une série d'expériences et dévoilèrent une nouvelle voie d'activation du systèmes CRISPR dans laquelle le petit tracrARN l'endoribonucléase III de l'hôte et la protéine Cas9 participent à l'obtention de crARNs actifs. Ils se rendirent également compte que non seulement le tracrARN était impliqué dans l'activité de ce système CRISPR mais que pour couper l'ADN le système CRISPR-Cas9 n'avait besoin que de trois composants :

- le tracrARN
- le crARN
- la protéine Cas9.

Ce système CRISPR-Cas9 était surprenant de simplicité comparé aux « autres systèmes qui n'utilisaient qu'un ARN et une multitude de protéines ».<sup>53</sup>

Emmanuelle Charpentier et ses collègues soupçonnaient les deux ARNs d'interagir entre eux pour guider la protéine Cas9 vers la séquence ADN ciblée du génome viral. Ce concept était inédit, un tel travail en équipe était routinier entre protéines mais n'avait jamais été observé pour des ARNs. Plus tard d'autres études confirmeraient leur intuition et montreraient que le tracrARN est essentiel à la protéine Cas9 pendant l'étape de clivage de l'ADN. 55,56

Jörg Vogel déclara *a posteriori* qu'il s'agissait d'une « période très intense ». Ils savaient que leurs découvertes allaient changer la donne, mais ils étaient aussi effrayés par la possibilité d'être doublés si leurs travaux fuitaient et qu'un autre groupe publiait avant eux<sup>53</sup>. Pour s'assurer que leur publication passerait sans encombre les demandes des arbitres, ils travaillèrent en silence pendant un an pour couvrir tous les angles de leur publication avant de la soumettre à la revue Nature.<sup>57</sup>



**fig. 11** Emmanuelle Charpentier et l'une de ses collègues au Max Planck Institute de Berlin, son actuel lieu de travail.<sup>58</sup>

Ils regroupèrent leurs conclusions dans une publication parue en 2011 <sup>57</sup>. (La figure en Annexe 5 représente leurs conclusions relatives au mécanisme de maturation du crARN).

#### 14. Résumé des connaissances microbiologiques

L'histoire de CRISPR en tant que système immunitaire étudié en microbiologie allait entrer en collision avec le domaine de l'édition génomique. Au fur et à mesure, l'accumulation des études sur le système CRISPR-Cas permit d'avoir une idée plus complète de son fonctionnement en tant que système immunitaire. C'est cette même base de connaissances qui permettrait aux scientifiques impliqués dans leurs découvertes d'identifier le potentiel qu'avait ce fantastique système naturel. Il pourrait être détourné de son activité initiale pour servir d'outil d'édition du génome et il avait le potentiel de faire partie des outils les plus performants de ce domaine.

Avant de quitter l'histoire microbiologique de CRISPR, pour nous intéresser seulement à l'ensemble formé par la protéine Cas9 et ses guides ARN, voici une description de l'action du système CRISPR-Cas9 en tant que système immunitaire adaptatif :

Le système CRISPR-Cas déclenche une immunité vis à vis d'éléments génétiques étrangers par le biais de 3 étapes : l'adaptation, l'expression et enfin l'interférence. Ces 3 étapes peuvent être réparties en deux groupes distincts : tout d'abord la « *prise d'information* » qui comprend l'étape d'adaptation puis « *l'action* » qui regroupe l'expression et l'interférence. Les protéines impliquées dans la prise d'information (Casl et Cas2) ont tendance à être hautement conservées entre les différents types CRISPR, tandis que celles des deux dernières étapes varient grandement d'un micro-organismes à un autre. <sup>59-61</sup>

Pendant l'étape d'adaptation, de courts morceaux d'ADN, homologues aux séquences virales ou plasmidiques envahisseurs, sont intégrés dans les loci CRISPRs. Par exemple, lors d'une attaque virale, cela déclenchera l'insertion d'un *spacer* dérivant de l'ADN viral. C'est la protéine Cas1, conservée quelque soit le système CRISPR qui semble impliquée dans l'acquisition. Typiquement dans le cas de virus, ce n'est qu'un seul *spacer*, ayant une longueur proche des 30pdb qui est inséré au niveau du côté leader du locus CRISPR. Ce *spacer* conférera à la bactérie une résistance dirigée contre le phage dont il est issu. Chaque intégration est accompagnée de la duplication d'une répétition et crée ainsi une nouvelle unité « *spacer-repeat* »<sup>10,52,53</sup>. La sélection de l'élément à capturer dans l'ADN envahisseur (le *protospacer*) semble être déterminée par la reconnaissance de PAMs. Les PAMs sont de courtes séquences de quelques nucléotides qui diffèrent entre les sous types de CRISPR-Cas<sup>45,52</sup>. Barrangou et ses collègues furent les premier à observer l'acquisition des *spacers* au sein des système CRISPR de *Streptococcus thermophilus*<sup>12</sup>.(Annexe 6: figure représentant l'acquisition d'un nouveau *spacer* au sein d'un locus CRISPR)

La seconde étape est l'expression. Durant celle-ci un long transcrit primaire du locus CRISPR est généré, c'est le pre-crARN. Il est ensuite découpé afin d'obtenir les courts crARN. Ce découpage est catalysé par une endoribonuclease qui opère soit en tant que sous-unité d'un complexe enzymatique de grande taille (par exemple le complexe Cascade ou *CRISPR associated complex for antiviral defence* chez E.coli) ou comme une unique enzyme (comme la Cas6 de *Pyrococcus furiosus*). Une variante existe avec le tracrARN qui agit comme un guide pendant la maturation du pre-crARN, qui est catalysée par la RNASE III en présence de la Cas9<sup>57</sup>.Dans le cas du complexe Cascade des systèmes de type I<sup>50,64</sup> le crARN mature reste associé au complexe protéique après le clivage initial par l'endonucléase. Par contre dans le cas de *Pyrococcus furiosus*, le

crARN, maturé par cas6, est pris en charge par un complexe protéique distinct, par lequel il sera maturé à nouveau au niveau 3′ 54,65,66

Une fois les crARN matures obtenus, l'ultime étape est l'interférence. Durant celle-ci l'ADN étranger est ciblé et clivé au sein de la séquence correspondant au *spacer* : la séquence proto-spacer.. Le crARN guide les différents complexes de protéines cas (le Complexe cascade chez *Escherichia coli*, Cas9 chez *S.pyogenes* ou Cas6 pour *P.Furiosus*) par appariement aux séquence complémentaires cibles du virus ou du plasmide qui correspondent au *spacer* (qui a servi de matrice au crARN)<sup>50</sup>. Les PAMs jouent aussi un rôle important pendant le processus d'interférence,<sup>14,67</sup> ils assurent que la coupure est limitée aux fragments qui la contiennent.

La connaissance de ces mécanismes permit la classification des systèmes CRISPR en types, en fonction de leurs protéines effectrices. Chacun est représenté par une bactérie qui représente son archétype : Type I avec *Escherichia coli*, Type II : *S.pyogenes* et Type III pour *P.Furiosus*. Parmi les systèmes découverts à ce moment là, le système CRISPR de type II et sa Cas9 est celui qui par sa simplicité et sa taille, avait le plus d'atouts pour être détourné au profit de l'édition génomique. (la classification des gènes Cas en 2011 est représentée en Annexe 7)

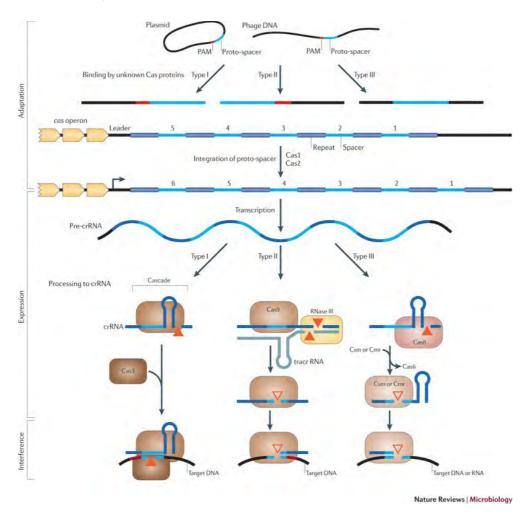


fig.12 Les étapes du système immunitaire adaptatif CRISPR-Cas 68

#### 15. Reconstitution de CRISPR-Cas9 dans un organisme étranger

L'une des interrogations sur CRISPR-Cas9 portait sur son autonomie, pour la vérifier Rimantas Sapranauskas, Virginijus Siksnys et leurs collègues décidèrent de déterminer si le système CRISPR de *Streptococcus thermophilus* pouvait être reconstitué et fonctionner au sein d'une espèce microbienne distante : *Escherichia coli*. Leur intuition fut la bonne, transférer un locus CRISPR-Cas entier s'avéra suffisant pour provoquer une interférence au niveau des éléments génétiques envahisseurs <sup>69</sup> En utilisant ce système hétérologue, ils confirmèrent en 2011 que pour ce type de CRISPR tous les éléments requis pour assurer la découpe d'une séquence cible étaient déjà connus. À partir de cette constatation, les travaux portant sur le détournement de ce système naturel dans un but de modification génétique pouvaient débuter.

#### 16. Étude de la Cas9 in vitro

Virginijus Siksnys est un biologiste moléculaire de l'université de Vilnius, en Lituanie. Ses collègues et lui bénéficiaient d'une expertise conséquente dans le domaine des enzymes de restrictions, c'est donc naturellement qu'ils débutèrent des travaux sur CRISPR-Cas. Cet intérêt les emmena à collaborer avec l'équipe du groupe DuPont Agroalimentaire où on retrouvait Rodolphe Barrangou et Philippe Horvath. En 2012 ils purifièrent le complexe Cas9-crARN de Streptococcus thermophilus en utilisant un tag streptavidin sur Cas9. Ils étudièrent ainsi son activité dans des tubes à essai 45. Ils montrèrent que ce complexe pouvait cliver un ADN cible in vitro, créant ainsi une cassure double brin, à 3 nucléotides de la séquence PAM. Celà confirmait les observations in vivo réalisées jusqu'à présent. Ils allèrent plus loin et montrèrent qu'ils pouvaient reprogrammer Cas9 en ajoutant des spacers artificiels au sein de l'ensemble CRISPR pour cibler in vitro les séquences qu'ils souhaitaient. En mutant les domaines nucléases HNH et RuVC, ils montrèrent que le premier est impliqué dans le clivage du brin complémentaire au crARN tandis que le second clive le brin opposé. Ils découvrirent également que le crARN pouvait être réduit à une 20 aine de nucléotides en permettant encore un clivage efficace. Conscient du caractère innovant de leur travaux ils mentionnèrent que :

« ces découvertes tracent la voie vers la conception d'endonucléases universelles guidées par ARN ».<sup>55</sup>

L'outil rêvé d'édition génomique apparaissait à portée de main.



**fig.13** L'équipe de Virginijus Siksnys (à gauche) : Giedrius Gasiunas, Tomas Sinkunas et Tautvydas Karvelis à l'*institute of biotechnology* de l'Université de Vilnius

Virginijus Siksnys et son équipe soumirent leur publication à la revue Cell en Avril 2012. Elle fut rejetée sans revue par les éditeurs de Cell, qui reconnaîtraient *a posteriori* l'importance de leur publication. Ils se contentèrent donc d'envoyer leur manuscrit à la *National Academy of Sciences* qui après de nombreux aller-retours de questions publia leur résultat en Septembre<sup>69</sup>. Entre le dépôt et la publication une autre étude était parue, cette dernière parvenait à des conclusions similaires mais avait un avantage notable : elle décrivait l'invention d'une technique qui ne nécessite que deux éléments pour fonctionner : un seul guide ARN et la Cas9.

#### 17. Création du SgRNA

À la même période Emmanuelle Charpentier commença à travailler avec ses collègues sur la caractérisation biochimique du systèmes CRISPR-Cas9. Lors d'une conférence sur les tracrARN à l'American society for Microbiology à Puerto Rico en mars 2011, elle rencontra Jennifer Doudna. Celle-ci était une biologiste structurelle réputée et experte en ARN à l'University California Berkeley. À ce moment là de l'histoire des CRISPR, Doudna et son équipe tentaient de percer les mystères de la structure des complexes Cascade des systèmes CRISPR de type I.

Les deux équipes scientifiques débutèrent une collaboration transatlantique pour percer les mystères du fonctionnement des CRISPR-Cas9. Elles utilisèrent la Cas9 recombinante (Provenant de *Streptococcus pyogenes* et exprimée chez *Escherichia coli*), un crARN et un tracARN transcrits *in vitro*. Comme Siksnys, ils dévoilèrent que la Cas9 pouvait couper de l'ADN purifié *in vitro*, qu'elle pouvait être reprogrammée en utilisant des crARNs sur-mesures et que les deux domaines nucléases coupaient chacun un fragment opposé. Mais l'originalité de leur étude résidait dans la démonstration que, *in vitro*, ces deux ARNs pouvaient fonctionner sous forme fusionnée. Cet unique ARN guide serait appelé sgARN pour single-guide ARN. Ce concept de sgARN se révélera largement utilisé en laboratoire de par sa praticité par rapport à l'instable association tracrARN:crARN.

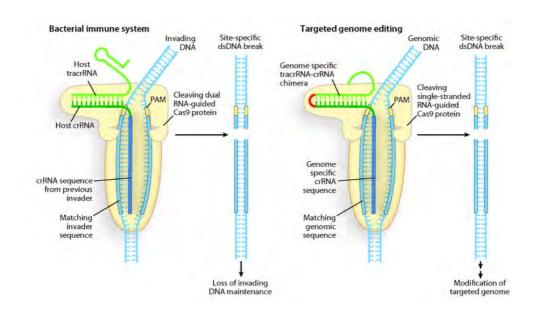


**fig.14** De gauche à droite : Emmanuelle Charpentier, Jennifer Doudna, Martin Jinek, Krzysztof Chylinski, Ines Fonfara et Michael Hauler (absent de la photographie) : l'équipe à l'origine de la création du sgARN.

Leur publication contenant le nouveau sgARN fut soumise à la revue Science et publiée en ligne en Juin, depuis elle connut une popularité croissante. Elle proposait une : « methodologie alternative (aux ZFN et Talens) basée sur la Cas9 programmée par un guide ARN qui pourrait avoir un potentiel considérable pour le ciblage et l'édition génomique»

Jennifer Doudna qualifierait plus tard la technologie de démocratique :

« Elle ouvre les portes du génie génétique à n'importe quelle personne qui possède des compétences de base en biologie moléculaire et qui souhaite éditer le génome ».



**fig .15** Illustration du détournement du système immunitaire bactérien au profit de l'édition génomique. Chez la bactérie le duo crARN:TracrARN de l'hôte va guider la cas9 grâce à la complémentarité entre séquence envahisseur et la séquence complémentaire du crARN issue d'une attaque virale antérieure entraînant la coupure site spécifique et donc la perte du maintien de l'ADN viral. Cette propriété peut être détournée pour réaliser une coupure précise qui permettra de modifier le génome cible. <sup>70</sup>

Après le succès de leur publication, Emmanuelle Charpentier pris ensuite deux

décisions. La première fut en accord avec son ambition originale d'appliquer son travail à des fins médicales, elle cofonda une entreprise afin d'exploiter leur méthodologie dans le domaine de la thérapie génique. CRISPR Therapeutics, basée à Cambridge aux Etats-Unis et Bâle en Suisse naquit donc en novembre 2013. Sa seconde décision fut d'obtenir un poste permanent doté d'un important support institutionnel; elle s'installa alors au *Helmholtz Centre for Infection Research* à Hannovre où elle occupa un poste de directrice d'un département et d'enseignante à la faculté de médecine.<sup>53</sup>

Jennifer Doudna déposa via UC Berkeley une demande de brevet le 25 Mai 2012 encadrant l'utilisation du duo Cas9-sgARN à des fins de modification génétique<sup>71</sup>. Cette demande créditait Emmanuelle Charpentier, Martin Jinek, Krzysztof Chylinski et ellemême comme inventeurs. Cela peut souvent prendre des années pour qu'un nouvel outil ne soit adopté en biotechnologie. Mais avant même la fin de l'année 2012, au moins six travaux portant sur différents usages de la technique CRISPR Cas9 furent soumis pour publication. L'un d'eux allait déclencher l'une des plus grandes bataille entre scientifiques de notre siècle.

#### 18. Du procaryote à l'eucaryote

Depuis la publication de l'étude de L.Marraffini et E.Sontheimer en 2009, décrivant CRISPR-Cas9 comme étant une enzyme de restriction programmable, nombreux sont les chercheurs qui envisagèrent son utilisation en tant qu'outil d'édition génomique. L'un des facteurs limitant l'utilisation de cet outil pouvait être sa capacité à agir dans les cellules de mammifères. Celles-ci ont une organisation interne bien plus complexe, leur génome est 1000 fois plus grand, empaqueté dans le noyau et enroulé dans une structure élaborée qu'est la chromatine. Des tentatives précédentes de « transfert de compétences » entre bactéries et eucaryotes avaient échouées, par exemple celle des introns auto épisseurs (self splicing introns). Fin 2012, la plupart des experts restaient prudents au sujet de l'application de CRISPR aux cellules eucaryotes. <sup>34</sup>



fig.16 Feng Zhang 315

Feng Zhang du *Massachusetts Institute of Technology* était déjà célèbre pour ses travaux sur le développement de l'optogénétique, un récent et excitant domaine des neuroscience qui consiste à insérer dans des neurones des gènes codant des protéines opsine. Ces protéines sont sensibles à certaines longueurs d'ondes. En éclairant ces neurones à une certaine fréquence il est alors possible de les activer ou de les inhiber et ainsi de cartographier les réseaux neuronaux. Mais l'un des principaux défis de l'optogénétique est d'obtenir des neurones contenant le gène opsine. Lorsque le fonctionnement des TALENs fut connu, F.Zhang intégra pendant son postdoc le laboratoire de George Church pour réaliser des travaux explorant l'utilisation de la technique aux cellules neuronales. G.Church est un influant et médiatique professeur

sénior à Harvard, spécialisé dans la génétique et la biologie synthétique et appartient comme F.Zhang au *Broad Institute*, initiative commune du MIT et Harvard.

Deux groupes feraient parti des pionniers de la transposition des TALENs aux cellules de mammifères. Le premier groupe travaillait pour Sangamo Bioscience, l'autre était celui composéde Feng Zhang, Paola Arlotta et Georges Church. Leurs publications parurent en ligne sur *Nature Biotechnology* à presque un mois d'intervalle, c'est celle de Miller et ses collègues de Sangamo Biosciences qui aurait la primeur.<sup>21</sup>

Avec l'émulation scientifique autour de CRISPR ils se lancèrent donc logiquement dans la transposition de la technologie aux cellules humaines, avec pour objectif cette fois d'être les premier à y parvenir. Cependant en Janvier 2013 ce sont 4 publications qui montrèrent la réussite d'une transposition de la technique CRISPR-Cas9 aux cellules murines et humaines :

- Cong et Zhang parue le 3 Janvier : Multiplex Genome Engineering Using CRISPR/Cas Systems<sup>72</sup>
- Mali et Church le 3 Janvier: RNA Guided Human Genome Engineering via Cas973
- Jinek et Doudna le 29 Janvier: RNA programmed genome editing in human cells<sup>74</sup>
- Cho et Kim le 29 janvier: Targeted genome engineering in human cells with the Cas9 RNA guided endonuclease<sup>75</sup>

Les mêmes groupes de recherche élucidèrent la structure cristalline des complexes protéines-ARN l'année suivante.

- Jinek et Doudna le 6 Février 2014 76
- Nureki et Zhang le 13Février 2014 77
- Anders et Jinek le 27 Juillet 2014 78

L'apparition de découvertes simultanées dans l'histoire scientifique n'est pas aussi rare qu'on pourrait le penser, elle correspond souvent à un moment où l'ensemble des connaissances nécessaires à la réalisation d'une nouvelle observation sont réunies. Mais dans le cadre de demande de brevets c'est à ce moment précis que tout se complique.

#### 19. Vers une commercialisation

Nombreuses sont les équipes qui avaient soumis des brevets concernant la technique d'édition génomique qui dérivait des CRISPR. Au fur et à mesure des publications, les applications qui pouvaient en découler apparaissaient de plus en plus extraordinaires. L'intérêt commercial et lucratif allait de pair. Les demandes de brevet déposées par Siksnys<sup>79</sup>, Sontheimer & Marrafini<sup>81</sup> et Kim<sup>82</sup> avaient été refusées par l'USPTO (*United States Patent and Trademarks Office*) car elles ne parvenaient pas à « décrire le caractère inventif ».

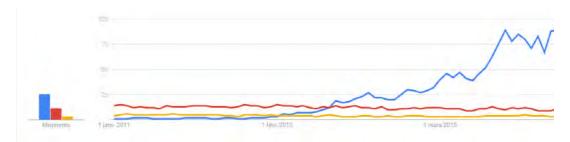
Deux demandes de brevets pouvaient se voir attribuer l'invention de techniques d'édition génomique basée sur l'utilisation de la Cas9 et son guide ARN. La première avait été déposée en Mai 2012 par Jennifer Doudna, *UC Berkeley* et l'Université de

Vienne<sup>71</sup>, la seconde en Décembre 2013 par Feng Zhang et le *Broad Institute*<sup>82</sup>. À la surprise générale c'est celle déposée par Feng Zhang à qui un brevet fut attribué en Avril 2014, déclenchant une agressive bataille entre équipes juridiques qui continue encore plus de 2 ans après l'attribution<sup>83</sup>. Le chapitre 8 de cette thèse y est en partie consacré.

Les disputes entre institutions académiques au sujet de brevet ne sont pas rares, mais un tel niveau de rancœur et de médiatisation est inédit. Souvent ces institutions parviennent à une entente, mais dans ce cas les intérêts commerciaux et la liste des industriels licenciés derrière chacune des compagnies, cotées en bourse, dont les équipes rivales font parties des comités scientifiques\* laisse peu d'espoir au compromis.

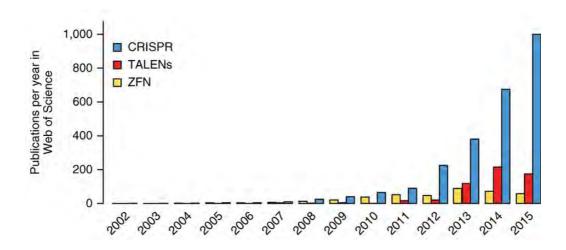
# 20. CRISPR: popularité explosive

Au début de l'année 2013 plusieurs écrits annonçaient la déferlante de publications à venir. Les applications : l'utilisation de la technique pour éditer des cellules souches humaines ou encore la modification d'un organisme entier (le poisson tigre ou zebrafish) <sup>86</sup>. Cette même année le nombre de recherches Google pour le terme CRISPR décolla, et encore aujourd'hui l'intérêt médiatique ne cesse de croître. L'intérêt scientifique et commercial alla de pair, avec en tête de file les potentielles applications thérapeutiques chez l'homme et l'utilisation dans le domaine de l'agriculture commerciale. Dans ces deux cas cela soulèvera de nombreuses préoccupations et réflexions éthiques.



**fig.17** Évolution des recherches sur le moteur de recherche google pour les acronymes des techniques d'édition génomique : ZFN en jaune, TALENs en rouge et CRISPR en bleu. Source Google Trends <sup>87</sup>.

<sup>\*</sup>Les scientific advisory board en 2016 comprennent: pour Intellia, Jennifer Doudna, pour Crispr Therapeutics, Emmanuelle Charpentier et pour Editas, George Church et Feng Zhang. Jennifer Doudna faisait à l'origine partie du comité scientifique et des membres fondateurs d'Editas 84, qu'elle quitta pour fonder Caribou/Intellia en Novembre 2014, quelque mois après l'attribution du premier brevet CRISPR à Feng Zhang 85.



**fig.18** Nombre de manuscrits par an déposés depuis 2002 contenant les mots clefs CRISPR ou TALENs ou ZFN d'après Thomson Reuters Web of Science.

Fin 2014, des équipes scientifiques avaient, entre autre, utilisé la technique CRISPR-Cas9 pour augmenter la résistance du blé aux insectes ou encore reproduire des effets carcinogènes ou corriger des mutations chez la souris adulte 88-90. Les pionniers des CRISPRs continuèrent d'explorer le potentiel de la technique, mais ils n'étaient plus seul, de nombreux scientifiques à travers le monde les avaient rejoints. Ce système immunitaire microbien dont l'homme ne soupçonnait pas l'existence s'était révélé en quelques années un outil exceptionnel pour les scientifiques.

# II MÉCANISME DÉTAILLÉ

### 1. Introduction

En microbiologie CRISPR-Cas désigne la structure du génome qui supporte l'immunité adaptative de la bactérie ou de l'archée. En biotechnologie c'est donc le même mot qui a été utilisé pour décrire la technique d'édition génomique par le duo composé de l'enzyme Cas9 et son guide d'ARN artificiel ( soit le sgARN unique, soit le doublet tarcrARN : crARN). La technologie CRISPR-Cas9 désigne donc la méthode qui permet d'éditer cellules, tissus et organismes entier de manière universelle : chez les plantes, animaux ou bactéries.

Les bactéries ont acquis les Cas9 pendant leur évolution. L'hypothèse majeure relative à leur origine est l'intégration d'un élément génétique parasite dans leur génomes, il provenait de casposons. Les casposons (transposons) sont des MGE, ou *Mobile Genetic Elements*. Tous les organismes vivants évoluent et persistent tout en étant soumis aux perpétuelles attaques d'éléments génétiques mobiles. C'est la « *domestication* » par une bactérie des casposons qui aurait permis l'apparition de la Cas1. Cette Cas1 serait l'ancêtre commun des autres espèces de protéines Cas apparues au cours de l'évolution. 91

La Cas9 est l'endonucléase guidée par l'ARN des systèmes CRISPR de Type II. Cas9 peut aussi désigner les gènes Cas9 qui codent pour cette protéine. Bien que des protéines homologues existent, c'est aujourd'hui celle-ci qui apparaît la plus pratique car elle est capable d'agir seule par rapport à certains complexes protéiques qui

remplissent le même rôle chez la bactérie. Elle est aussi de petite taille. La saCas9 de *Staphylococcus aureus* étant celle dont le rapport taille/spécificité semble le plus optimal<sup>89</sup>. Bien que de taille supérieure, la spCas9 de *Streptococcus pyogenes* est également largement répandue en laboratoire car elle fut adoptée la première.

### 2. Les acteurs

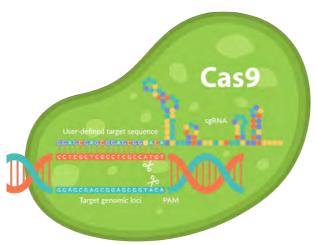


fig.19 Schéma du complexe Cas9 : sgARN : ADN cible

La protéine Cas9 est représentée en vert, le guide ARN, ici le sgRNA monobrin, est conçu pour contenir une séquence de 20 nucléotides complémentaires de la séquence ADN cible. L'ADN cible sera scindé par la Cas9 au niveau du brin complémentaire au sgARN (en orange), et du brin opposé en bleu. Cette coupure parallèle générera des bouts francs. La présence du motif PAM présent sur le brin d'ADN opposé est indispensable à l'activité enzymatique. Au niveau moléculaire c'est toute une machinerie qui permet la réalisation de ces coupures.<sup>93</sup>

## 3. Déroulement de l'action au niveau macromoléculaire

Au contact de l'ADN double brin qu'elle va ensuite découper, la Cas9 reconnaît d'abord une séquence conservée de 2 à 4 paire de bases : le PAM. Pour *Proto Adjacent motif* <sup>44,78,94</sup>. Après fixation au PAM, la Cas9 interroge la séquence ADN opposée à celuici à la recherche d'une complémentarité avec l'ARN guide<sup>95</sup>. Si il y à complémentarité, avec les 12 premières bases (la séquence appelée *seed*), l'invasion de l'ARN peut avoir lieu en même temps que l'ADN se déroule pour former une boucle R ( c'est à dire une boucle composé d'ADN simple brin formé par hybridation avec un ARN).

S'ensuit un clivage précis de chaque brin par les domaines HNH et RuvC de la Cas9 qui génèrent une cassure double brin à bout franc. La cassure a lieu 3 nucléotides en amont de l'extrémité 3' de la séquence *protospacer*, en mesurant à partir du PAM <sup>96</sup>.

Au delà des mécanismes précédents (la complémentarité avec l'ARN guide et la reconnaissance par le PAM), un troisième mécanismes plus complexe affecte la spécificité de la Cas9. Il s'agit du contrôle conformationnel du clivage de l'ADN. L'ADN nucléaire est balayé sous toutes les dimensions par le passage de la Cas9, ce qui entraîne un temps de passage différent en fonction des zones cibles et hors cible. Ce temps appelé le differential binding residence time, qu'on peut traduire par « durée du séjour au niveau des zones de liaison » est court pour les séquences hors-cible, et prolongé pour les séquences cibles. Ces trois mécanismes contribuent à la spécificité du ciblage des sites de clivage 97.

## 4. Relation Structure-Activité

# 4.1: Le gène Cas9

Voici le gène qui code la protéine Cas9 de Streptococcus pyogenes :

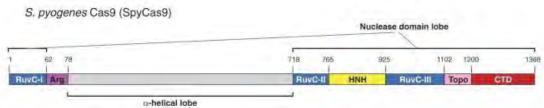


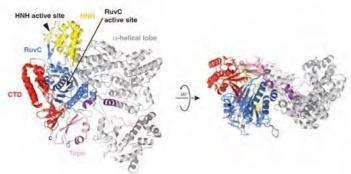
fig.20 Gène codant spCas9

Les deux domaines nucléases sont RuvC et HNH, le domaine RuvC en bleu est discontinu il est constitué de trois segments,une hélice α (en gris) relie le premier et le second segment. La zone Arg en violet représente une région riche en arginine chargée positivement. En jaune, le domaine HNH dont la séquence est située entre deux domaines RuvC. La séquence Topo en rose est un domaine présentant des homologies avec les topo-isomérases. Le CTD, *C Terminal Domain* représenté en rouge est impliqué dans la liaison du motif PAM.

Après transcription et traduction les propriétés physico-chimiques des acides aminés codés vont conditionner le repliement et la fonction de cette protéine.

## 4.2 : La protéine Cas9

Voici la représentation en trois dimensions de la protéine Cas9. C'est grâce à sa structure unique qu'elle va se lier à l'ARN guide (tracrARN:crARN ou sa forme fusionnée le sgARN). Ce duo protéine:ARN se liera ensuite à l'ADN.



**fig .21** Structure tertiaire représentée en ruban de la spCas9 Le domaine HNH est en jaune, le RuvC en bleu, la zone Topo en rose et le CTD capable de sécuriser la liaison au PAM en rouge, à noter la séparation spatiale entre le lobe dit nucléase (à gauche) et le lobe en hélice  $\alpha$  'à droite en gris.<sup>76</sup>

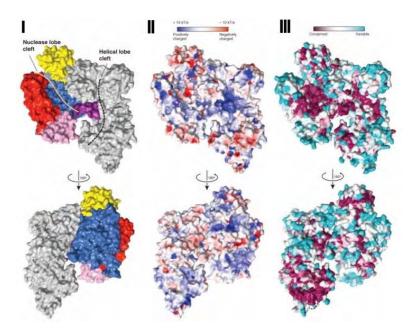


fig.22 Représentation de la surface de SpCas976

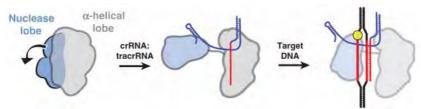
I Représentation illustrant la présence de deux sillons sur la surface moléculaire, ils délimiteront le canal qui renfermera le complexe sgARN:ADN cible

II Représentation des potentiels électrostatiques de surface allant de −10 kT/e (rouge) à +10 kT/e (bleu), les zones bleues chargées positivement se lieront préférentiellement à des éléments chargées négativement (telles que l'ADN)

**III** Représentation de la spCas9 en fonction de la conservation de ces zones au cours de l'évolution. En vert les zones variables entre espèces, en violet celles conservées. Cette représentation peut être générée à partir de l'outil ConSurf<sup>98</sup> qui se base sur de multiples alignements de séquences de Cas9 <sup>76</sup>

# 4.3 : Liaison de la protéine à l'ARN guide puis à l'ADN

L'analyse structurelle de la Cas9 de *Staphylococcus pyogenes* a permis de mieux comprendre la complexité de son mécanisme d'action. La cristallographie par rayon X et la microscopie électronique ont révélé l'architecture moléculaire de la Cas9 pendant son activité. La protéine subit d'intenses réarrangements conformationnels durant sa liaison à l'ARN guide. <sup>64,76,99</sup>



**fig.23** Illustration des réarrangements de la Cas9 après contact avec le guide ARN En rouge : séquence ARN complémentaire à l'ADN cible et portion du double brin d'ADN contenant cette séquence complémentaire . En jaune : séquence PAM, en Noir double brin d'ADN non ciblé.

Comme nous l'avons vu une hélice riche en arginine unit les deux lobes de la cas9. Elle semble agir comme une charnière et joue un rôle clef dans l'association au guide ARN.

Lorsqu'ils se lient au guide crARN : tracARN, les deux lobes structurels de la Cas9 se réorientent et les deux fentes contenant des zones de liaison à l'ADN (riches en acides aminés chargés positivement) se retrouvent face à face. Un canal central est ainsi formé. C'est cette structure contenant la séquence ARN complémentaire qui permettra la liaison à l'ADN cible avec une grande spécificité.

L'ADN cible se dénature au niveau de la zone où est reconnu le motif PAM. Les motifs Arginines de l'extrémité C terminale de la Cas9 vont interagir avec le PAM situé au niveau du sillon majeur du brin non complémentaire. Le groupe phosphodiester situé en position +1 sur le brin d'ADN cible interagit avec le sillon mineur au niveau du PAM. C'est cette interaction qui déclenche probablement la séparation locale des brins. Le brin d'ARN guide va s'intercaler en aval immédiat du PAM et l'hybride ARN-ADN (ou boucle R) va se former. 17,18

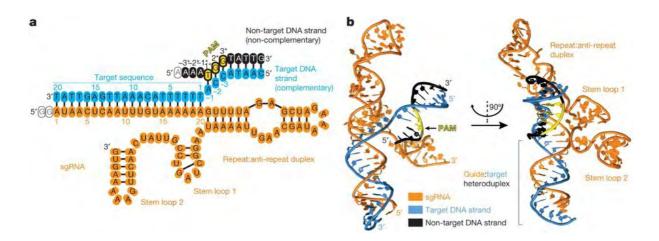
La création de cette structure permet la poursuite du réarrangement. Une rotation supplémentaire par rapport au lobe en hélice α a lieu, elle permet l'englobement complet de l'ADN cible, les deux domaines nucléase sont alors au niveau de chacun des deux brins d'ADN, en face de la séquence cible, le clivage peut avoir lieu. La découpe se fait 3 à 4 nucléotides en amont de la séquence PAM<sup>76</sup>. Le domaine HNH de la Cas9 va permettre de cliver le brin complémentaire. Le domaine RuvC de la Cas 9 lui va cliver le brin opposé au brin complémentaire.



**fig.24 : a.** Structure du complexe formé par la protéine Cas9, l'ARN guide (en orange) et l'ADN cible (en bleu pour la séquence complémentaire, en noir pour l'ADN non ciblé) **b.** modèle 3D en nylon de la Cas9<sup>100</sup>.

## 4.4 : Structure du complexe ARN guide : ADN cible

Si on pouvait observer ce qui se passe à l'intérieur de la protéine quand le complexe Cas9:ARN:ADN est formé, voici un exemple des acides nucléiques que nous pourrions apercevoir :



**a**. exemple de séquence d'ARN guide (ici sa forme fusionnée : sgARN) et d'acide nucléique cible. L'ARN guide est coloré en orange. Celui-ci est complémentaire de l'ADN cible sur une longueur de 20 nucléotides, c'est cette séquence qui sera artificiellement modifiée pour permettre de désigner la cible choisie. L'ADN cible est représenté en bleu ciel, l'ADN non cible en noir. Le trinucléotide 5'-NGG-3' constituant le *Protospacer Adjacent Motif* (PAM), il est surligné en jaune au sein de la séquence non cible (en noir).

**b**. vues sous plusieurs angles de la boucle R : la jonction 4 voies formée par le sgRNA et l'acide nucléique cible.<sup>78</sup>

Des études mécanistiques montrent le rôle clef de la séquence PAM dans l'initiation de la liaison avec les brins d'ADN. Sans ce motif même les séquences cibles parfaitement complémentaires avec l'ARN guide ne peuvent être reconnues par la Cas9 <sup>16</sup>. (La structure cristalline du site de liaison au PAM de spCas9 est présentée en Annexe 8)

## 4.5: Techniques d'observations

En observant ce qu'il se passait au niveau moléculaire la manière dont les acteurs interagissent entre eux a été décrite. C'est par des expériences appelées *single-molecules* ou de biochimie de masse (*bulk biochemical*) que furent réalisées ces observations. En faisant varier indépendamment chacun des paramètres au niveau moléculaire ou macromoléculaire le rôle de chacun des composants se révélait.

# 5. Liaison Hors-Cible et spécificité

Pour évaluer le comportement de la cas9 dans les cellules, les chercheurs utilisèrent des expériences d'immunoprécipitation et de séquençage à haut débit (ChIP-seq) pour déterminer le nombre et le type de sites de liaison de la Cas9 aux chromosomes. Leurs conclusions furent identiques quelque soient les cellules qu'ils observaient (cellules embryonnaires de rein humain (HEK293) ou de souris (mESCs). La Cas9 dont l'activité catalytique avait été inactivée pouvait se lier à plus de sites que ceux complémentaires au guide sgARN utilisé. Ces interactions hors-cible avec l'ADN ont lieu typiquement au niveau de sites dotés d'un motif PAM et des séquences partiellement complémentaires à l'ARN guide. Ces résultats sont cohérents avec le mode d'interrogation de l'ADN qui avais été décrit pour la Cas9 95,101,102

La Cas9 active découpe rarement l'ADN hors de sa cible, ce qui implique que pour un clivage efficace, une complémentarité proche de la perfection entre l'ARN guide et la zone cible est nécessaire. Ces observations sont conformes avec l'observation *in vitro* de la formation des complexes d'ARN <sup>95</sup>

# III OPTIMISATION, VARIANTES et ASTUCES

# 1. Introduction

# 1.1 Utilisation simple en édition génomique

La plus simple utilisation du duo Cas9:ARNguide consiste à générer une coupure double brin au niveau d'un site spécifique. De cette découpe peuvent découler plusieurs applications en fonction du mécanisme de réparation endogène impliqué.

La cellule hôte dispose de deux mécanismes majeurs pour réparer le génome au niveau du site de coupure double brin. Soit le NHEJ (pour *Non Homologous End Joining*) soit le HDR (pour *Homology Direct Repair*), chacun a des propriétés différentes et donc des applications bien particulières.

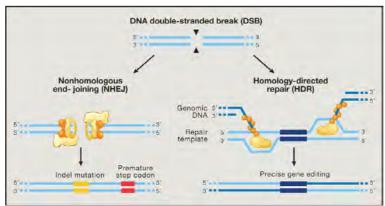


fig.26 Mécanismes impliqués dans la réparation d'une coupure double brin de l'ADN<sup>86</sup>

-NHEJ: Après coupure les hétérodimères de protéine Ku se lient à chaque extrémité du fragment coupé et serviront de support aux protéines de réparations qui leur sont associées, elles combleront l'espace entre les deux bouts de la coupure. La réparation par NHEJ est une voie qui commet fréquemment (5%) des erreurs, ces erreurs sont appelées *indels* pour *insertion/deletion* et sont susceptibles de décaler le cadre de lecture. Cette voie est donc la voie de choix pour générer des *knockout* mais son caractère aléatoire entraînera une population cellulaire expérimentale où les mutations apparues sont hétérogènes.<sup>103</sup>

-HDR: Contrairement au NHEJ cette voie utilise un brin donneur/modèle pour combler la coupure. Lorsque le modèle est présent, c'est cette voie qui sera prépondérante, elle repose sur les protéines Rad51 qui se lient aux extrémités de la coupure durant l'initiation de la réparation. Elles recrutent des facteurs accessoires qui dirigeront la recombinaison homologue avec les portions du fragment donneur. Ce fragment peut être apporté artificiellement sous forme de ssODN (pour single stranded DNA oligonucléotide), d'oligonucléotide double brin ou d'un plasmide donneur. <sup>101</sup> C'est cette voie qui sera préférée pour réaliser avec précision des mutations, insertions et délétions par le biais d'une séquence donneur: c'est le *knock in*. En prenant pour précaution de s'assurer de l'absence de PAM dans la séquence donneur, afin qu'elle ne soit pas à son tour découpée.

## 1.2 Rôle du cycle cellulaire

Bien que les cellules possèdent différents mécanismes de réparations (le NHEJ et l'HDR) c'est la phase du cycle cellulaire qui gouverne principalement le choix entre ces deux voies. Le NHEJ est dominant pendant les phases G1 , S et G2 tandis que l'HDR est limité à la fin des phases S et G2, lorsque la réplication de l'ADN est complète et les chromatides sœurs sont disponibles pour servir de matrice de réparation. En pratique, cela limite la réalisation de *knock in* par la Cas9 et ses guides ARNs aux cellules en division et restreint donc cette approche, surtout *in vivo*. Plusieurs artifices peuvent être déployés pour contourner cette limitation<sup>104-107</sup>. D'autre voies mineures de réparation de l'ADN tels que le single *strand annealing, alternative end joining, microhomology-mediated joining, mismatch base nucleotide* ou *excision repair* peuvent aussi être utilisées pour éditer le génome<sup>103,108-111</sup>.

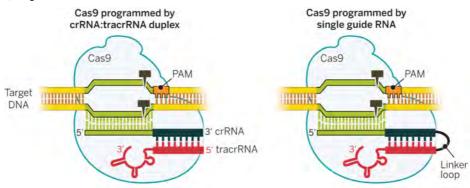
## 1.3 Utilisations « sophistiquées »

En jouant sur la fonctionnalité des trois éléments de base de l'outil, il peut être modifié pour ne pas être un simple « *ciseau génétique* ». Ces modifications ont pour but d'obtenir un outil plus sophistiqué et plus précis, elles permettent également d'ouvrir la porte à de nouvelles fonctions. On peut par exemple citer l'inactivation de la cas9 suivi de sa fusion à une protéine fluorescente. Le but : l'obtention d'un duo protéine:ARN qui va être capable de traquer spécifiquement une séquence d'intérêt, de s'y fixer et d'éclairer son parcours dans une cellule vivante. On peut aussi être capable de couper l'ARN par une astuce consistant à leurrer la Cas9 avec un motif PAM fourni artificiellement. Enfin par un astucieux cocktail nucléotidique, il est possible de s'affranchir de l'HDR pour réaliser un *knock in*, permettant ainsi réaliser cette approche sur des cellules hors division, facilitant ainsi sa mise en place *in vivo*. Voici une liste non exhaustive rassemblant en catégories ces modifications. Elles trouvent des applications tant en recherche fondamentale qu'appliquée.

# 2. modification de l'ARN guide.

# 2.1 le sgARN

La plus ancienne modification est celle réalisée par l'équipe de Jennifer Doudna et Emmanuelle Charpentier : la création d'un seul ARN guide le sgARN (pour *single guide ARN*) à partir de la fusion du tracrARN et du crARN.



 $\label{eq:cases} \textbf{fig 27} \ \ \text{Représentation des différents guide ARN: le duplex tracrARN: crARN à gauche et le sgARN au sein du complexe Cas9, ARN guide, ADN cible $^{112}$$ 

Le sgARN conserve les éléments critiques au fonctionnement de la forme originale mais ses deux fragments sont liés par une boucle pour n'en former qu'un seul. Il comporte :

- une séquence 5' de 20 nucléotides qui va pouvoir s'apparier à la séquence cible et permettre de programmer le système pour cibler n'importe quelle séquence d'intérêt, tant qu'elle est à proximité d'un PAM<sup>76</sup>
- une structure tertiaire en double brin à l'extrémité 3' qui permettra de se lier à la Cas9

L'avantage majeur du sgARN est sa synthèse : une seule séquence est plus simple à produire et à utiliser en laboratoire. Cela évite par exemple de devoir procéder au retempérage (ou *annealing* en anglais) permettant la liaison tracARN : crARN. Ce processus est chronophage, il utilise des tampons cytotoxiques et son rendement est inférieur à 100%. Un autre avantage est le fait que cette structure fusionnée est plus stable *in vivo*, et plus facile à cloner en grande quantité lorsque les laboratoires

souhaitent en commander pour réaliser un grand nombre d'expérimentations.93

# 2.2 CRISPR Multiplex

Une autre des modifications initiales est le multiplex. Cette technique décrite pour la première fois par Feng Zhang et son équipe dans leur publication regroupant leurs travaux sur la cellule humaine consiste à cibler non pas un, mais plusieurs sites distincts. Pour y parvenir il suffit de fournir à la Cas9 plusieurs sgARNs. En pratique cela se traduit par la transformation de bactéries à l'aide d'un plasmide contenant les séquences codant les multiples sgARNs<sup>73</sup>. Cette technique peut alors permettre de réaliser des délétions d'une large portion du génomes ou encore de cibler plusieurs sites distincts. Elle pourra être utilisée en synergie avec les modifications suivantes.

## 3. Modifications de la Cas9

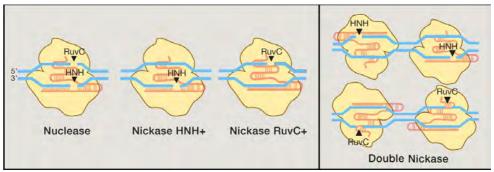
#### 3.1 Introduction

Il est possible de muter les régions nucléases HNH ou RUVc. Cela aboutira à un variant de la Cas9 qui ne pourra cliver qu'un seul des deux fragments. Cette activité de clivage monobrin est appelée une activité *nickase*. Si par contre on mute les deux domaines on obtient une protéine incapable de tout clivage, on la nomme la dCas9 ou deadCas9 (mutation Asp10  $\rightarrow$  Ala, His840  $\rightarrow$  Ala). Cette dCas9 est donc une protéine guidée par ARN et dotée d'une activité de liaison. <sup>5,7</sup>

### 3.2 nCas9 ou Cas9 nickase

Mis à part l'optimisation des gARNs durant leur conception, la spécificité du système CRISPR peut être augmentée par des modifications de la Cas9 elle-même. Les Cas9 nickases se lient à l'ADN basées sur la spécificité de la séquence gARN, mais elles ne génèrent qu'une entaille (nick en anglais) en ne coupant qu'un seul brin au lieu d'une coupure double brin. Ces coupures monobrin sont rapidement réparées par HDR en utilisant le brin complémentaires. Pour générer une coupure double brin, c'est l'utilisation de deux nickases qui sera nécessaire, chacune coupant un brin opposé et créant un brin découpé en quinconce.

Cette technique surnommée *double-nick* CRISPR permet d'améliorer la spécificité du ciblage.En effet ils est très peu probable que deux coupures monobrins soient à la fois hors-cible et assez proche pour générer une coupure double brin (un peu comme la croyance populaire d'un éclair qui ne peut frapper deux fois au même endroit). Cette technique peut être combinée avec l'HDR pour réaliser des éditions génomiques de manière très spécifique<sup>113-115</sup>



**fig.28** Représentation schématique du fonctionnement de la Cas9, de Cas9 *nickases* et du système double *nickase*<sup>86</sup>

# 3.3 dCas9 ou dead Cas9 : Crispr Interference

L'une des faculté clef de las Cas9 est sa capacité à se lier spécifiquement aux acides nucléiques lorsqu'elle est guidée par son ARN. Dépourvue d'activité nucléase la dCas9 peut être utilisée pour des expériences de régulation génique par interférence à l'échelle du génome entier. Ce processus est nommé CRISPRi pour *Crispr Interference*. En fonction du site reconnu par le complexe dCas9-ARN guide cette stratégie permettra par exemple de bloquer la phase d'élongation pendant la transcription ou encore de bloquer la liaison de facteurs de transcription. Un des intérêts majeurs de CRISPRi est son utilisation pour réprimer plusieurs gènes simultanément et le fait que la technique soit réversible. Cette technique offre une alternative intéressante à l'ARNi, notamment grâce à l'absence de système CRISPR chez les eucaryotes, évitant toute compétition avec des processus endogènes.

### 3.4 dCas9 Fusion

La possibilité de cibler une séquence via la dCas9 et de s'en servir comme plateforme pour emmener une protéine au contact de l'ADN cible permet de nombreuses possibilités voici quelques exemples qui seront détaillés par la suite :

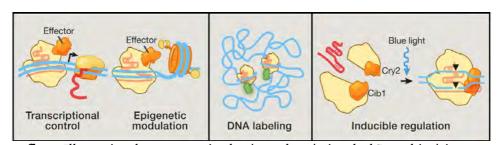


fig.29 Illustration des construction basée sur la création de dCas9 chimériques

À gauche la fusion avec des effecteurs permet la régulation génétique et épigénétique. Au centre la fusion avec la GFP permet le marquage dynamique et in vivo de l'ADN. À droite une construction permettant le contrôle de son activité. Ce contrôle se fait via la reconstitution d'une cas9 préalablement scindée dont chaque lobe porte une protéine qui, une fois induite, forme un hétérodimère(ici les protéines cib1/cry2).

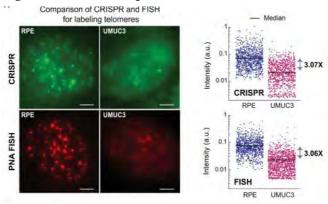
## 3.4.1 dCas9 Fusion régulatrice

En générant des versions chimériques de la dCas9 qui sont fusionnées à des domaines de régulation, il est possible d'utiliser la technique CRISPRi pour réaliser de la régulation d'expression génique au sein de cellules de mammifères. Par exemple la fusion dCas9 aux domaines VP64 ou KRAB transforme cette protéine en régulateur transcriptionnel et permet respectivement de déclencher (on parle alors de CRISPRa ou CRISPR activation) ou de réprimer la transcription dans des cellules humaines. La fusion avec des acyltransferases et des déméthylases permettrait de générer des modifications épigenetiques<sup>116,117</sup>.

## 3.4.2 dCas9 Fusion en imagerie

Lorsqu'elle est fusionnée à un fluorophore, la dcas9 devient un outil de visualisation spécifique des séquences d'ADN et permet une visualisation dynamique de la chromatine <sup>118</sup>. C'est le travail qu'a réalisé une équipe de l'Université de Californie <sup>118</sup>.

Pour savoir si les techniques CRISPR et PNA FISH (*Peptide nucleic acid Fluorescent in situ hybridization*, une variante du Quantitative FISH ou Q FISH) sont aussi efficaces l'une que l'autre ils les ont utilisées pour mesurer la longueur des télomères. En analysant le *median telomere puncta intensity*, c'est à dire l'intensité lumineuse médiane produite par chaque point (donc chaque télomère rendu fluorescents) ils ont été capable d'apprécier leur longueur. En comparant la fluorescence de deux lignées cellulaires, les cellules RPE (*Retinal pigment epithelium*) et les cellules UMUC3, lignée de cellules cancéreuses de vessie humaine, ils observent un taux médian d'intensité lumineuse environ 3 fois supérieur pour les cellules RPE quelque soit la méthode de marquage utilisée. De plus le nombre de télomères calculés est retrouvé identique qu'importe la méthode. La corrélation entre ces mesures suggère que l'imagerie par CRISPR permet d'obtenir des résultats comparable à celle par PNA FISH. Par contre elle possède un avantage immense : elle peut être utilisée sur des cellules vivantes.



**fig.30** L'imagerie par CRISPR permet de détecter la longueur des télomères : Comparaison de la longueur des télomères des cellules RPE et UMUC3 en utilisant l'imagerie par CRISPR (en haut) ou PNA FISH (en bas). Le diagramme à point montre l'intensité de chaque télomères identifiés sur une échelle logarithmique. En bleu pour RPE et violet pour UMUC3.

Cet outil d'imagerie via CRISPR a le potentiel d'améliorer les techniques antérieures pour étudier les réarrangements dynamiques des chromosomes dans leur état natif, au cœur des cellules vivantes. Des applications d'imageries multicolores peuvent même être créées en utilisant plusieurs dCAS9 distinctes. Par exemple il est désormais possible de visualiser simultanément jusqu'à 6 loci en utilisant la technique appelée CRISPR *rainbow* <sup>119</sup>.

# 3.5 dcas9 fusion et purification des acides nucléiques

En se basant sur le concept de chIP (*chromatin ImmunoPrecipitation*), les chercheurs ont développé le enChIP (*pour engineered DNA binding molecule-mediated chIP*) qui permet de purifier une séquence génomique spécifique en la ciblant par ARNg. Une dCas9 inactive fusionnée à une étiquette épitope (une courte séquence peptidique pour laquelle on possède un anticorps capable de s'y lier avec une grande affinité) peut être utilisée pour purifier l'ADN lié au gARN <sup>124</sup>.

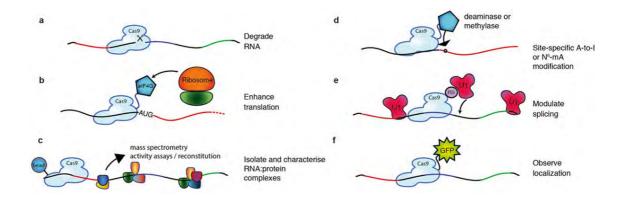
# 3.6 Split Cas9

La Cas9 a une structure bilobée, cela peut s'avérer être une propriété intéressante pour contrôler son activité à l'intérieur même des cellules. En la découpant, en deux sous-unité, on sépare les deux lobes. On peut alors contrôler son ré-assemblage en utilisant des domaines hétéro-dimères inductibles qui agirons alors comme des bandes velcro inductibles. La dimérisation de ces hétéro-dimères peut être déclenchée chimiquement par des petites molécules ( pour les domaines ABI/PYL) ou encore par certaines longueur d'onde ( domaines CIB1/CRY2). 121,122 Ils permettent une régulation spatiale et temporelle de l'activité Cas9 qui peut se révéler utile en clinique pour faciliter le contrôle de l'activité systémique de la Cas9.

## 4. Cibler l'ARN: la RCas9

La Cas9 a montré qu'elle pouvait être un outil versatile dans un large panel de cellules eucaryotes et procaryotes. Mais elle a longtemps été pensée incapable de cibler l'ARN.<sup>55</sup> En 2014 une astuce a été découverte, il est possible « d'induire en erreur » la Cas9 pour qu'elle soit capable de se lier à l'ARN simple brin complémentaire d'une séquence ARN guide. Pour y parvenir il faut utiliser comme artifice de courts oligonucléotides d'ADN contenant la séquence PAM nécessaire à l'activation de l'enzyme. Ces oligonucléotides présentateurs de PAM seront surnommés : les PAMmers.

En utilisant ces PAMmers sur mesure, la Cas9 peut être dirigée pour se lier ou découper des ARN cibles tout en évitant l'ADN qui pourrait correspondre à la séquence guide. Cette technique peut aussi servir à isoler un ARNm endogène spécifique. La connexion fondamentale entre la liaison du PAM et la sélection du substrat fais de la Cas9 un outil de choix pour réaliser des reconnaissances de transcrits de manière programmable (*transcript recognition*) sans avoir recours à des *tags* (étiquettes).



**fig.31** Potentielles applications de la Rcas9 : l'analyse de transcrits, leur détection et leur manipulations

- a. Clivage spécifique d'ARNsb par Rcas9
- **b**. dRcas9 fusionnée au facteur d'initiation eucaryote eIF4G ciblant la région 5' non traduite de l'ARNm pour activer la traduction.
- **c**. la dRCas9 attachée à des billes (magnétiques ou de haute densité) peut être utilisée pour cibler puis isoler un ARN spécifique ou un complexe ARN:protéine d'intérêt. Les applications sont multiples : étudier la structure d'un ARN lorsqu'il est lié de manière native à une protéine, identifier des complexes ARN:Proteines, Analyses downstream, enrichissement en transcrits en vue d'un séquençage.
- **d**. la dRCas9 attachée à une ARN deaminase ou un domaine N6-mA methylase pourrait permettre de réaliser respectivement une édition A vers I (Adenosine vers Inosine) ou une méthylation sur un site spécifique dans le but de générer des évenements épitranscriptomiques.
- **e.** dRCas9 fusionnée à un domaine de recrutement U1 (domaine riche en arginine et sérine RS) pourrait être programmée pour reconnaitre un *site splicing enhancer* et ainsi permettre de promouvoir l'inclusion d'un exon d'intérêt.
- **f.** dRCas9 attachée à une protéine fluorescente, comme la GFP, peut être utilisée pour observer la localisation d'un ARN et son transport dans une cellule vivante.

# 5. Plusieurs Cas9 pour plusieurs actions

L'utilisation de systèmes sgARN :Cas9 orthogonaux <sup>123-126</sup> permet aux chercheurs d'exploiter différents outils CRISPR pour des applications distinctes dans une même cellules, autorisant la manipulation de plusieurs voies de signalisation à la fois.

Bien qu'une Cas9, par exemple celle issue de *Streptococcus pyogenes* (la spCas9), puisse être utilisée pour réaliser une unique action sur plusieurs sites, elle ne peut réaliser plusieurs actions différentes sur différents locus. Par exemple pour une cellule dans laquelle on utilise une *dead* spCas9 couplée à un activateur; on ne peut pas y réaliser une édition génomique avec une spCas9 sur une séquence différente sans être obligé de couper la zone ciblée par l'activateur. Pour pouvoir utiliser la technique cas9 pour réaliser plusieurs actions au sein d'une même cellule, il faut développer une méthode pour relier chacune des actions à chacune des cibles. La solution pour obtenir ce niveau de contrôle est d'utiliser des protéines Cas9 orthogonales.

# 6. S'émanciper de l'HDR: Homology Independent Targeted Integration

Un défaut majeur du transfert de gène par la technique CRISPR-Cas9 est sa restriction aux cellules en division. Pour surmonter cette limite une équipe de chercheurs internationaux a développé une stratégie nommée HITI pour *homology-independent targeted integration*. Elle combine les éléments classiques, la cas9 et ses ARN guides avec un « *cocktail d'acides nucléiques* » destinés à tirer profit du NHEJ qui, au contraire de l'HDR, est une voie active que la cellule soit en division ou non. Leur publication rapporte le succès de cette approche à la fois *in vitro* et *in vivo*, en utilisant un comme preuve de concept la dégénérescence rétinienne par retinitis pigmentosa chez le rat. <sup>106,127</sup>

#### 7. Conclusions

#### 7.1 Amélioration continue

Comme souvent pour toute technologie naissante, les outils dérivés de CRISPR devront être améliorés. Par exemple la grande taille des Cas9 et des protéines associées impacte l'efficacité du packaging dans un vecteur viral. Le ciblage limité par le PAM et l'activité hors cible peut affecter la précision de l'édition, les sgARN sont soumis à l'activité des RNAses cellulaires influant le rendement. Il est très probable que ces approches combinées permettront de produire des variants CRISPR optimisés avec une spécificité et une activité améliorée. Réduisant ainsi les risques et limites d'une utilisation clinique *in vivo*.Mais d'ores et déjà ces protéines et leurs guides représentent une fantastique boîte à outils, applicable notamment en recherche.

## 7.2 Vers de nouvelles protéines effectrices

L'exploration des luxuriants systèmes CRISPR des bactéries et archées devrait accroître la gamme d'outils moléculaires à disposition des chercheurs. Certains pourraient rendre les précédents efforts obsolètes. L'analyses bioinformatique des souches natives de cas9 permet d'identifier de plus petits homologues de la cas9. Certains de ces variants possèdent des PAMS différents par rapport à la SpCas9, et offrent ainsi des alternatives pour cibler l'ADN, complétant la boîte à outil CRISPR. En 2016 c'est d'autres nucléases provenant de la classe II qui pourraient détrôner la cas9 : cpfl <sup>128,129</sup> et c2c2 <sup>130</sup>).

## - C2C2 cible naturellement l'ARN<sup>131</sup>

- Cpf1 découverte fin 2015<sup>132</sup>, ne nécessite qu'un seul ARNg (à la place du duo tracrARN:RNA), elle interagit avec des motifs PAMs différents (5'-TTN-3'), coupe donc l'ADN à une position différentes et au lieu de générer des coupures à bouts francs, il crée de pratiques bouts collants. Ces 3 particularités lui procurent plusieurs avantages :
  - La découpe en bout collant permet, à la manière d'un velcro, de réaliser des insertions directionnelles des gènes, analogue à celles réalisées par les enzymes de restrictions traditionnelles. Ces bouts collants pourraient s'avérer particulièrement pratiques lorsque ce sont des cellules qui ne se divisent pas dans lesquelles on souhaite insérer un gène, la voie HDR nécessaire au *knock in* par cas9 étant cycle dépendante.

- Le Protospacer Adjacent Motif riche en T pourrait étendre le nombre de sites ciblés par la technique CRISPR aux régions riches en A-T dépourvues des motifs 3'NGG que favorisent spCas9.
- Puisque Cpf1 ne nécéssite pas de tracrARN, l'ARN nécessaire pour le guider ne mesure qu'approximativement 42nts de long. La synthèse de ces ARNs sera donc moins onéreuse que celle des ARNg de la Cas9. Cpf1 et ses guides sont plus petits que leurs équivalents du système spCas9, ils seront aussi plus facile à transporter par des vecteurs de faibles capacité, comme les AAVs.

La protéine Cpf1 est une alternative à la Cas9 qui promet de multiples avantages, sa découverte révèle qu'il nous reste encore beaucoup à apprendre de la biologie des systèmes CRISPR et que de futures prospections pourraient permettre de dénicher de nouvelles nucléases plus adaptées à l'édition du génome.

# IV CRISPR EN PRATIQUE

## 1. Introduction

Une infinité de possibilitées semblent à la portée de la technique CRISPR-Cas9. Par contre en laboratoire des constantes demeurent, que l'on souhaite rendre fluorescent des gènes in vivo ou qu'on injecte CRISPR à un animal pour aller découper un génome viral intégré dans l'ADN de ses cellules. Qu'importe la technicité de l'expérience, ces étapes simples encadreront sa réalisation.

# 2. Étude, sélection, synthèse, vectorisation, validation

### 2.1 Étude de la cible

La cible va déterminer une variété de facteurs influençant l'expérimentation. La séquence utilisée pour créer les gARN dépendra du gène cible et de l'espèce dont proviennent les cellules. Lorsque cela est possible, la région ciblée devrait être séquencée avant la conception des gARN. La moindre variation entre l'ADN cible et le gARN pourrait influencer la spécificité. Le nombre d'allèles pour chaque gène peut varier en fonction de l'organisme, ce qui pourrait impacter l'efficacité des *knock-out* (inactivation d'un gène) ou *knock-in* (intégration d'un gène) par Cas9.

# 2.2 Sélection du gène à manipuler

Dans le but de manipuler un gène donné, l'identification du gène d'intérêt est nécessaire. Cependant, en fonction du résultat souhaité la région exacte du gène peut varier. Quelques exemples :

- Pour activer ou réprimer un gène cible en utilisant des dCas9 activateur ou dCas9 répresseur, les gARN devront cibler le promoteur qui contrôle l'expression du gène d'intérêt.
- Pour réaliser des *knock out* génétiques, les gARN devront de préférence cibler les exons exprimés de la région 5' pour réduire la probabilité que la région cible soit retirée des ARNm par épissage alternatif. Les Exons proches de la zone N-Terminale sont à cibler, car une mutation déplaçant le cadre de lecture augmentera la chance de production d'une protéine non fonctionnelle.
- Les gARNs peuvent être créés pour cibler des exons qui codent pour des domaines protéiques indispensables. L'objectif étant que la mutation (insertion ou délétion) qui résulte de la cassure double brin ait plus de chance d'altérer le fonctionnement de la protéine.
- Pour les expériences qui utilisent le processus d'HDR, il est essentiel que la séquence cible soit proche de la zone où on souhaite réaliser l'édition.

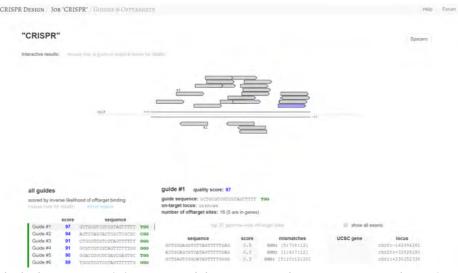
# 2.3 Sélection des gARN basée sur l'activité "on target" et "off target"

Une fois le gène cible identifié, il est nécessaire de vérifier l'existence de PAM dans sa séquence. La séquence PAM est absolument nécessaire à la Cas9 pour se lier à l'ADN. Ils est donc nécessaire d'identifier toute les séquences PAM au sein de la zone à cibler ( le PAM pour spCas9 est 5' NGG 3'). Si aucune séquence PAM n'est présente dans cette zone, on peut se tourner vers une Cas9 d'une espèce différente ou provenant d'un variant de la même espèce (le variant spCas9 EQR à par exemple 5' NGAG 3' comme séquence PAM).

Dès que toute les séquences PAM potentielles on été identifiées, c'est le moment de choisir quel site sera à même de permettre le clivage le plus spécifique. Le gARN doit être complémentaire de la séquence cible, mais il est aussi obligatoire de vérifier qu'une séquence identique n'existe pas, quelque part, ailleurs dans le génome. Dans un monde idéal, le gARN ne correspondrait uniquement qu'à la séquence voulue. En réalité il est possible qu'un gARN ait des sites additonnels pour lesquels une complémentarité partielle existe. Ces sites sont appelés "off-targets" ou hors-cible en français, et doivent être pris en compte pendant la construction des gARN expérimentaux.

Mis à part l'activité hors cible, il est également important de tenir compte des facteurs qui favoriseront un clivage efficace du site désiré, c'est l'activité dite *on-target*. Deux gARN ciblant des séquences pour lesquels ils possèdent une homologie de 100% n'auront pas forcément le même rendement. Cette efficacité peut être augmentée ou diminuée en fonction des nucléotides qui composent la séquence. Par exemple un gARN qui contiendrais un G en position 20 ( soit 1pb en aval du PAM) peut être plus efficace qu'une séquence contenant un C à cette position. Malgré le fait que ces deux séquences présentent une homologie parfaite avec leur cible.

Il est donc nécessaire de tenir compte des prédictions d'effets *on-target* et *off-target*. Plusieurs logiciels de conception de gARN ont été dévellopés<sup>133-136</sup> et sont capables d'identifier les motifs PAM et séquences cibles puis de hierarchiser les gARNs correspondants en se basant sur ces prédictions. La base des algorithmes est d'éviter les sgARN qui ont des complémentarités partielles hors cibles possédant moins de 3 mésappariements. Il est également possible d'acheter des gARNs préalablement validés.



**fig.32** Exemple de diagnostic opéré par un outil de conception de gARN : CRISPR design(crispr.mit.edu). Un total 23 séquences guides potentielles (en gris ont été trouvés pour cibler la séquence d'intérêt de 127 nts soumise à l'outil .

# 2.4 Synthèse et clonage des gARNs

Une fois la séquence complémentaire des gARNs choisie, il faut concevoir les oligonucléotides codant ces gARN et cloner ces oligonucléotides à l'intérieur de vecteurs d'expression. Dans la plupart des cas, les séquences guides complémentaires sont achetées synthétisées, remises en suspension (annealing) , puis insérées dans des plasmides qui contiennent la séquence codant le reste de la structure du gARN en utilisant des techniques de restriction-ligation. Une alternative consiste à concevoir et commander la séquence codant pour la cassette d'expression.



Ce fragment obtenu de 455pdb conçu *in silico* pourra être acheté à un fournisseur d'oligonucléotides (tel que *Integrated DNA technologies*<sup>138</sup> au prix de 89 USD pour 500ng en décembre 2016). Les fragments reçus pourront ensuite être amplifiés. Par exemple par PCR avec comme amorce :

gRNA-F:TGTACAAAAAAGCAGGCTTTAAAG; gRNAR:TAATGCCAACTTTGTACAAGAAAG)

ou à partir d'un kit de clonage plasmidique (par exemple pCR-Blunt II-TOPO thermofisher<sup>139</sup>). La Cas9 et ses différentes formes, pourront être achetées auprès de fournisseurs identiques à ceux des ARN guides (par exemple au prix de 128 euros en Janvier 2017 pour 10 microgrammes de protéine Cas9 chez le fournisseur Thermofisher<sup>140</sup>)

#### 2.5 Vecteurs

Pour introduire le système Cas9et gARN dans les cellules cibles, les approches qui avaient connu un sucés précédent ont à nouveau été testées. À la différence des techniques de thérapie géniques traditionnelles dont la réussite repose presque toujours sur une expression maintenue, une expression transitoire de la Cas9 peut suffire pour obtenir le résultat escompté. Les vecteurs viraux qui ont le potentiel de servir de véhicule aux composants nécessaires (Cas9, guides ARN, +/- séquences matrices d'un gène donneur en vue de réparation) sont nombreux : rétrovirus, lentivirus autoinactivant, adénovirus et AAV (*Adeno Associated Virus*). Les méthodes non virales peuvent transporter de l'ADN plasmidique, l'ARNm Cas9, les gARN transcrits ou synthétisés *in vitro*, les complexes ribonucléoprotéiques Cas9/gARN et les matrices d'acides nucléiques donneurs. Pour faire entrer dans la cellule ces cargaisons de

protéines et d'acides nucléiques, plusieurs techniques sont à la disposition des chercheurs, elles incluent : l'électroporation, la transfection (lipidique, polymère cationique, calcium phosphate), ou l'osmocytose induite par la récente technique iTOP).<sup>1,2,143</sup>Le choix de ces méthodes dépendra de leur efficacité en fonction du type de cellules visées et de la taille du contenu à insérer.

Bien que les expériences initiales utilisaient des plasmides ou des vecteurs viraux codant la Cas9 et les sgARN<sup>147</sup>. Les avancées en matière d'alternatives non ADN-dépendantes, telles que les complexe protéo-ARN pré-assemblés créent de nouvelles possibilités de systèmes de délivrance<sup>145</sup>. La délivrance directe de complexes ribonucléoprotéiques a montré de haut taux d'éditions génomiques dans les cellules T <sup>146</sup> Les Cas9 ribonucleoproteines délivrées en utilisant la transfection lipidique dans les cellules humaines et de souris142semblent prometteuses ont a été combinées avec des oligodeoxynucléotides simples brins pour introduire des mutations dans des cellules IPSCs humaines 147Les *nanoclews* d'ADN, un nouveau variant avancé des *nanocages* et nanoparticules à également été utilisé pour délivrer CRISPR.

Néanmoins, le premier pas pour les applications cliniques sera certainement de cibler les cellules et tissus pour lesquels des méthodes de délivrances sont déjà disponibles et validées, tel que le sang, le foie, les yeux et les muscles.<sup>112</sup>

#### 2.6 Validation

Une fois que gARN et Cas9 ont été délivrées avec succès à l'intérieur des cellules cibles, c'est le moment de valider les modifications apportées. Dans le cas de l'édition, la délivrance de la Cas9 et de gARN aux cellules sauvages créera plusieurs génotypes différents réparties dans la population cellulaire « mutante ». Un pourcentage de cette population peut conserver un phénotype sauvage à cause :

- d'un défaut d'expression de Cas9 et/ou de gARN.
- d'une absence de clivage efficace bien que la cas9 et le gARN aient été insérés dans la cellule.

Les cellules qui ont été mutées sont hétérogènes, elles peuvent être homozygotes ou hétérozygotes. De plus chez les cellules dont les deux allèles ont été mutés, ces deux allèles peuvent encore être différents à cause de la fréquence des erreurs liées au fonctionnement du NHEJ. Même pour les expériences favorisant le HDR, ce ne sera pas tous les allèles mutés qui contiendront l'édition souhaitée car un important pourcentage des coupures DB seront encore réparées par NHEJ. Le résultat final de la plupart des expériences sera donc une population cellulaire hétérogène de cellules qui contiendront une grande variétés de mutations/éditions au sein des gènes ciblés.

Comment déterminer si l'édition souhaitée s'est produite ? La méthode qui permettra cette validation dépendra de l'expérience, mais de nombreuses méthodes sont possibles, elles incluent entre autres :

- Mismatch-cleavage assay (pour les coupures DB réparées par NHEJ)
- PCR et digestion par enzymes de restriction pour les DSB réparées par HDR introduisant de nouveau sites de restriction
- L'amplification par PCR et l'électrophorèse (pour évaluer la réparation HDR ou NHEJ) pour des insertions ou délétions de taille importante la région d'intérêt peut être amplifiée à l'aide d'amorces qui l'encadrent (pour les délétions et petites insertions) ou qui recouvrent la frontière génome-insert ( pour les insertions).
- L' Amplification PCR, Subclonage et séquençage selon Sanger (pour évaluer la réparation HDR ou NHEJ) procurent une évaluation semi-quantitative de la fréquence de ciblage et la séquence exacte des allèles générés. Implique une amplification PCR des régions cibles de l'ADN, un subclonage dans un plasmide et un séquençage individuel de chaque clones.
- Amplification PCR et Next Generation Sequencing (pour évaluer la réparation HDR et NHEJ) procurent une évaluation quantitative des éditions génomiques au sein de la séquence cible et peut également évaluer les effets hors cible

# **PARTIE II: POTENTIEL ET IMPACT**

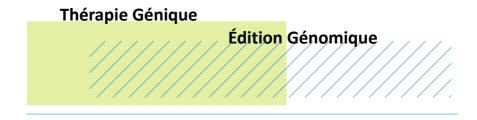
# **Avant-Propos**

Cette partie sera constituée d'une série d'exemples illustrant le potentiel actuel de la technique CRISPR-Cas9. Souvent les travaux présentés auront des limites, mais ils permettent tous d'entrevoir d'excitantes possibilités qui seront peut être un jour techniquement à notre portée. Chacune des présentations aura pour but d'apporter une perspective nouvelle sur la technique CRISPR et répond à un contexte particulier. Ce contexte scientifique, médical mais aussi législatif permet de mieux comprendre l'importance, les risques et l'impact de ces applications. Ces exemples montrent parfois que même si elle est la méthode qui peut paraître la plus avancée, CRISPR n'est pas forcément la meilleure solution à un problème donné. D'autre fois l'usage de la technique peut s'avérer si novatrice qu'elle surclasse les cadres réglementaires et éthiques actuels. Quelquefois elle montre que l'espoir de soigner des pathologies pensées incurables pourrait être à portée de main. Voici les exemples abordés :

- La thérapie génique: preuves de concept dans le cas de maladies monogénétiques; la première édition CRISPR de cellules embryonnaires, le premier essai clinique CRISPR mettant l'édition génomique au service de l'immunomodulation antitumorale et l'éradication in vitro et in vivo de latences virales avec les exemples d'HBV, HIV et des Herpesviridae.
- le développement d'antibactériens spécifiques dans un contexte d'apparition et de dissémination de bactéries multirésistantes.
- L'usage agronomique qui nous confronte aux limites des régulations actuelles sur les organismes génétiquement modifiés et entraîne une ré-évaluation de celles-ci.
- Les gene drives ou génétique dirigée, une technique au potentiel fulgurant qui permettrait par exemple d'éradiquer les espèces d'anophèles disséminant le paludisme, mais pour laquelle nous ne possédons pas d'outils capables d'évaluer l'impact environnemental, sanitaire et social.
- L'apport aux outils utilisés en recherche

Enfin une partie nommée Panorama abordera des thèmes importants liés à la technique CRISPR-Cas9 : sa propriété, la naissance d'une réflexion portant sur son usage clinique, l'émergence d'un marché basé sur sa découverte, et enfin les limites de notre propre compréhension.

Nous nous concentrerons sur les techniques d'édition génomique et en particulier la technique CRISPR, les autres techniques ne seront qu'évoquées. Pour faciliter une suite logique dans l'enchaînement des chapitres, la page suivante contient une infographie organisant les exemples du potentiel de CRISPR qui seront présentés répartis en catégories.



Usage thérapeutique	Maladies monogénétiques Cancérologie Antiviral (latence)	Antibactérien
		Génétique dirigée
Biotechnologie		Agronomie

	Création de modèles
Recherche	Criblage
	GWAS et Neuroscience

# I Thérapie génique

#### 1. introduction

Depuis sa conceptualisation par deux chercheurs dans une publication parue en 1972<sup>148</sup> la thérapie génique à connu d'importantes avancées médicales. En un court laps de temps la technique est passée du stade de concept aux recherches en laboratoires puis elle à franchit la barrière du passage en clinique. Cette histoire fut faite de succès et de revers. En 1990 la première patiente souffrant de ADA-DICS (Déficits Immunitaire Combiné Sévère lié à la déficience en Adenosine déaminase) était traitée par thérapie génique (par modification ex vivo) déclenchant une vague d'enthousiasme et d'études dans des indications variées allant des maladies génétiques, neurologiques, oculaires, infectieuses et vasculaires à la cancérologie. Mais des incidents dus aux vecteurs viraux (rétrovirus) utilisés survinrent, entraînant le décès de patients pendant les phases cliniques. D'abord en 1999 suite à une réaction inflammatoire déclenchée par le vecteur puis en 2002 suite à une mutation insertionnelle déclenchant une pathologie de type leucémie<sup>253</sup>. Suite à ces incidents une longue période d'introspection et de durcissement réglementaire s'établit durant laquelle des études se concentrèrent sur les mécanismes impliqués dans les effets « on target » et « off target » afin de garantir la sécurité de la technique et ultimement celle des patients.

Pendant cette période de nouvelles techniques émergèrent, développées en réponse au manque de spécificité des précédentes approches. Ces techniques sont capables d'édition génomique, et incluent les ZFNs, TALENs et désormais CRISPR. Bien que faisant partie des méthodes de thérapie génique, l'édition génomique n'est pas limitée à celle-ci et se révèle être un outil dont les applications ne cessent d'être découvertes dans de multiples domaines. En parallèle de nombreuses avancées ont été faites dans la conception des vecteurs utilisés en thérapie géniques, qu'il s'agisse de vecteurs viraux ou non.

Grâce aux avancées réalisées plusieurs études rapportèrent des progrès de la thérapie génique dans des applications diverses, que ce soit dans le cas de maladies génétiques, neurologiques, cancéreuses ou des protocoles antiviraux <sup>150–153</sup>. En 2012 c'est le glybéra corrigeant le déficit en lipoprotéine lipase, qui est le premier médicament de thérapie génique à être approuvé par l'EMA (*European Medicine Agency*) à la mise sur le marché<sup>154</sup>. Destiné à un très faible nombre de patients et possédant un lourd protocole de mise en œuvre, il est aussi l'un des médicaments les plus chers du marché européen. En 2016 c'est le Strimvelis, destiné au traitement des ADA DISC qui reçoit à son tour un avis positif <sup>155</sup>.

# 2. Techniques : du transfert de gènes à l'édition génomique

Initialement la thérapie génique était basée sur des vecteurs viraux et le transfert de gènes. La thérapie génique virale modifie des virus connus pour les utiliser comme un « cheval de Troie » destiné à insérer le gène d'intérêt dans les chromosomes <sup>156</sup>. La technique non virale utilise quant à elle des vecteurs non viraux pour délivrer le matériel génétique aux cellules mais son faible rendement la rend pour l'instant peu adaptée à un usage clinique.

Afin de s'affranchir du caractère semi aléatoire de l'insertion, des alternatives à la précision accrue ont été recherchées. CRISPR-Cas9 est la plus récente addition à l'arsenal d'édition génomique dont nous disposons. Même si elle présente encore des défauts, son faible coût et sa facilité d'utilisation font qu'elle est considérée comme un candidat de choix pour parvenir à de nouvelles applications cliniques et démocratiser l'édition génomique.

# 3. Approches

Les techniques d'édition génomique appliquées à la thérapie génique sont celles qui visent à éditer le génome de cellules humaines dans le but de réparer ou éliminer une mutation pathologique ou de conférer à ces cellules un avantage capable de participer à l'élimination d'une pathologie. Y parvenir requiert l'apport de ces modifications à un nombre suffisant de cellules pour parvenir à renverser la situation pathogène, la modification apportée devra perdurer pendant toute l'existence de la pathologie sous-jacente. Dans le cas des maladies génétiques cela signifie que la mutation devra durer indéfiniment et être exprimée par les cellules modifiées mais également par leur descendance.

Deux approches cliniques sont envisageables : in vivo et ex vivo

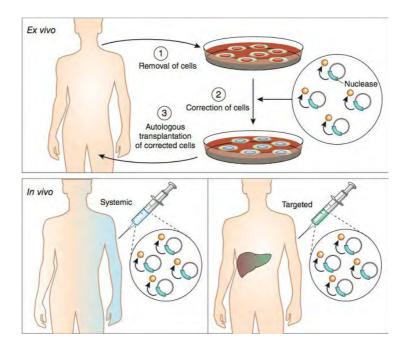


figure 34 Approches Ex Vivo vs. In Vivo

Pour l'édition génomique *ex vivo* les cellules du patient sont prélevées, éditées puis ré-introduites (illustration en haut). Pour que cette approche soit fructueuse, les cellules cibles doivent être capables de survivre hors du corps et de « *nicher* » à nouveau dans leur tissus d'origine après transplantation. L'approche *in vivo* implique l'édition génomique des cellules *in situ* (illustrations du bas). Pour la thérapie *in vivo* systémique, des vecteurs qui sont relativement indifférent de l'identité cellulaire ou de leur état doivent être utilisés pour permettre d'atteindre un large panel de type tissulaires. Par exemple la délivrance systémique d'AAV de type 8 a été utilisée sur des modèles précliniques pour cibler le tissu hépatique avec une grande efficacité. L'approche *in vivo* peut également être réalisée par injection locale des vecteur viraux aux tissus souhaités, tels que l'œil, le cerveau ou les muscles.

# 4. Maladies monogéniques

## 4.1 Annotation des gènes

L'annotation de l'ADN ou l'annotation du génome caractérise le processus qui consiste à identifier la localisation des gènes et de toutes les régions codantes au sein du génome puis à déterminer l'activité de ces gènes. Une fois que le génome est séquencé, il doit être annoté pour pouvoir lui donner un sens.

## 4.2 Gènes et maladies génétiques

En décembre 2016, c'est environ 25 000 gènes qui ont été déjà été annotés dans le génome humain, parmi ces gènes, la mutation de 3677 d'entre eux a été reliée avec l'apparition de phénotype pathologiques<sup>157</sup> ( d'après <a href="https://www.omim.org/">https://www.omim.org/</a>). Ce nombre évolue au fur et à mesure des découverte et connaît une croissance constante.

Aujourd'hui, grâce à la diminution des coûts du séquençage, achèvement du *human genome project* en 2003<sup>158</sup> et l'accumulation exponentielle de données provenant du séquençage du génome d'individus malades, le rôle de la génétique dans la santé humaine est devenu un axe majeur de la recherche et de la médecine clinique. Ces

avancées dans la compréhension des composantes génétiques des pathologies génétiques ont améliorées notre compréhension de leurs mécanismes et ont permis d'élaborer de nouvelles stratégies thérapeutiques.

Malgré les efforts déployés et des hypothèses thérapeutiques validées il n'y a eu qu'un nombre limité de réussites lorsqu'on a utilisé des petites molécules pour traiter des maladies à forte composante génétique. L'émergence de stratégies thérapeutiques basées sur l' atteinte des acides nucléiques au sein des cellules et tissus pathologiques a donc catalysé les espoirs. De telles stratégies sont particulièrement adaptées au traitement de maladies d'origine monogénique à forte pénétrance, telles que les déficit immunitaire combinés sévères (DICS), l'hémophilie ou certaine déficiences enzymatiques. Ces pathologies à la génétique bien définies manquent souvent de traitement sûr ou efficaces. Par contre l'édition génomique semble aujourd'hui incapable d'apporter une solution thérapeutique aux pathologies ayant une composante multigénique mais son utilisation en recherche permettra d'améliorer notre compréhension de celles-ci.

## 4.3 CRISPR-Cas9 et maladies monogéniques

Forte des expériences réalisées avec les techniques antérieures, c'est logiquement que la communauté scientifique s'est lancée dans l'utilisation de CRISPR-Cas9 pour évaluer ses applications dans le traitement de maladies génétiques. Les essais précliniques sur des modèles animaux sont nombreux, des années s'écouleront sûrement avant d'éventuellement voir des applications cliniques. Mais ces preuves de concept apportent de multiples espoirs et une meilleure compréhension des mécanismes pathogènes. Voici une série d'exemples, chacun pionniers d'un modèle d'étude : organoïdes, cellules souches induites, animal vivant et zygote.

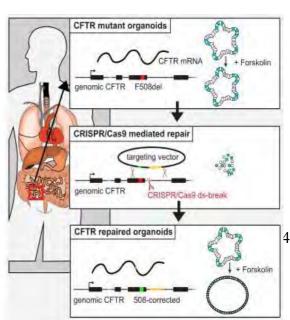
### 4.3.1 Exemple de la mucoviscidose : organoïdes et iPSCs

La mucoviscidose est la pathologie génétique pulmonaire la plus commune. Elle est causée par la mutation du gène CFTR présent sur le chromosome 7, cela entraîne la mutation de la protéine canal CFTR (*cystic fibrosis transmembrane conductance regulator*). Ce canal transporte les ions chlorures et thiocyanates, sa mutation entraîne la dysrégulation du transport des fluides épithéliaux dans les poumons, le pancréas et d'autres organes.En 2013, une preuve de concept utilisant la Cas9 guidée par sgARNs a permis de corriger la mutation du gène CFTR par recombinaison homologue.

**fig.35** Illustration du protocole développé par Schwant et al.

Les cellules souches intestinales humaines ou murines peuvent lorsqu'elles sont cultivées à long terme aboutir à la formation d'organoïdes génétiquement et phénotypiquement stables.

L'augmentation du taux d' AMPc du milieu de culture entraîne le gonflement



de ces organoïdes en déclenchant l'ouverture des canaux CFTR. Cette réponse est perdue pour les organoïdes qui proviennent de patients atteints de mucoviscidose. (L'augmentation du taux cellulaire en AMPc peut être déclenché artificiellement par l'ajout de forskoline, un activateur de l'adénylate cyclase). C'est ce modèle que Scwhank et ses collègues ont utilisé pour montrer que l'allèle qu'ils avaient corrigé était exprimé et fonctionnel, publiant l'une des premières études portant sur l'utilisation de la Cas9 pour corriger des pathologies monogéniques. <sup>159</sup>

En 2015 des travaux portant sur des cellules souches induites (iPSCs) provenant de patients porteur de la même mutation F508 du gène CFTR furent utilisées dans une nouvelle étude. Après correction ces iPSCs furent différentiées en cellules épithéliales pulmonaires matures chez lesquelles une expression et une fonction physiologique de CFTR furent démontrées<sup>160</sup>.

## 4.3.2 Exemple de la tyrosinémie : modèle in vivo, animal adulte

La restauration d'un phénotype sain chez des souris adultes atteintes de tyrosinémie héréditaire a été montrée en utilisant le duo Cas9:sgARN administré par injection hydrodynamique en 2013. Elles permettent de corriger la mutation Fah chez les hépatocytes aboutissant à l'expression de la protéine sauvage chez environ 1/250ème des hépatocytes. L'expansion de ces hépatocytes permettant de produire la protéine en quantité suffisante pour permettre un retour à l'état physiologique.<sup>90</sup>.

# 4.3.3 Exemple de la cataracte : modification de zygotes de souris.

La mutation du gène Crygc est associée avec de multiples formes de cataractes autosomiques dominantes. Une étude parue dans Cell en 2013 a exploré le potentiel d'une injection d'ARNm Cas9 et de sgARN ciblant l'allèle mutant. La correction déclenchée se faisant par HDR, soit en utilisant l'allèle sauvage endogène comme matrice soit un oligonucléotide exogène. Les zygotes transgéniques aboutirent à des souris adultes fertiles et capables de transmettre la modification à leur descendance mais un nombre faible de modifications hors cibles furent rapportées.

# 4.3.4 Première modification d'embryons humains : β-Thalassémie

Les avancées des travaux d'édition génomique via cas9 laissaient supposer que tôt ou tard, sa transition à l'embryon humain apparaîtrait. La technique a ses défenseurs et ses détracteurs, certains considèrent qu'elle aurait la capacité d'éradiquer de dévastatrices pathologies génétiques mais elle va de pair avec de nombreuses inquiétudes et dérives. Elle pourrait engendrer des effets imprévisibles sur les générations futures et soulève diverses questions éthiques.

Suites aux suppositions d'édition génomique utilisant CRISPR-Cas9 sur l'embryon humain, dès Mars 2015, des groupes industriels appelaient à un moratoire concernant l'usage de la technologie sur l'embryon. De manière concomitante, divers experts encourageaient la communauté scientifique à se rassembler autour d'un nouveau sommet sur l'édition génomique <sup>161,162</sup>. En mai 2015, concrétisant les rumeurs fut réalisée la première étude sur un embryon au stade précoce à l'université de Guangzhou, en Chine, déclenchant une vague de protestations. Cette publication dresse avec honnêteté

et précision la réalisation de la manipulation et ses nombreux obstacles. En utilisant des embryons tripronucléaires (non fertiles mais capables de se diviser) provenant d'un centre de fertilité local, ils montrèrent que la technologie pouvait être appliquée pour cliver le gène de la  $\beta$ -globine endogène. Remplacer ce gène par l'allèle sain apporté sous forme d'oligonucléotides par HDR s'avéra plus délicat, seule une faible portion des zygotes étant modifiés avec succès et certains zygotes étaient mosaïques (plusieurs de leurs cellules possédaient des génotypes différents). Un nombre élevé de coupures horscibles furent rapportées, plus élevé encore que celui observé pour les modèles animaux ou de cellules somatiques humaines. Bien que controversée, cette étude conclut sur l'immaturité de la technique, le manque flagrant de nos connaissance dans le domaine de la réparation de l'ADN embryonnaire et du fonctionnement de la cas9 sur ce même ADN. S'il était considéré un jour d'utiliser la technique sur l'embryon, l'importance d'accumuler de nouvelles connaissances avant d'entreprendre une telle application sera cruciale.  $^{163}$ 

## 4.4 Aperçu d'autre travaux

D'autres études ont depuis montré des modifications restaurant un phénotype cellulaire ou tissulaire physiologique dans des pathologies monogéniques dans des indications variées :

- la myopathie de duchenne (pathologie liée à la dystrophine)164-168
- <sup>-</sup> l'anémie de Fanconi (à gène FANC-C) <sup>169</sup>
- la maladie de huntington (gène HTT)170
- la rétinite pigmentaire (gène RPGR) 171
- la maladie de vaquez (gène JAK2) 172
- le DAAT ou déficit en α-1 antitrypsin (gène SERPINA1) 169
- la drépanocytose chez des cellules progénitrices humaines (gène β-globine) 173
- l'hémophilie B (facteur IX) 174

#### 4.5 Conclusion

Bien que présentant l'avantage de pouvoir cribler et sélectionner les cellules modifiées avant leur ré-injection, l'approche *ex vivo* n'est pas applicable à toutes les maladies monogéniques. Plusieurs pathologies comme les myopathies ou la mucoviscidose nécessitent une modification *in vivo*.

Avant que l'édition génomique *in vivo* basée sur le système CRISPR puisse être appliquée chez l'homme dans des essais cliniques, des inconvénients pratiques et des défis techniques seront à surmonter. Tout d'abord définir et atteindre des objectifs de précision en matière de clivage et de réparation. Ensuite parvenir à une délivrance efficace en ciblant un type particulier de tissus, de cellules ou d'organes. Nous devrons aussi comprendre comment contrôler les différentes voies physiologiques de réparation

et définir de manière prédictible le devenir mutationel de la réparation ADN après la création de cassures doubles brins.

Que ce soit pour l'édition *ex vivo* ou *in vivo*, les effets hors-cibles freinent les potentielles applications. Malgré l'adoption rapide de la Cas9 pour réaliser des perturbations génétiques et épigénétiques, et les progrès réalisés pour améliorer la spécificité de la protéine, nous connaissons encore mal son profil de liaison et de clivage hors-cible mais aussi ciblé. Comprendre la liaison de la Cas9 dans diverses situations d'accessibilité à la chromatine et différents états épigénétique permettra également une meilleure évaluation de son fonctionnement et permettra d'améliorer le choix des ARN guides

Les données générées par les études de ces liaisons permettront de compléter les modèles prédictifs et de minimiser l'activité hors-cible de la protéine, en thérapie génique ou pour d'autres applications qui requièrent un niveau élevé de précision. L'édition génomique semble avoir un long chemin à parcourir avant de pouvoir transposer les preuves de concepts de thérapie génique en clinique.

#### 5. CRISPR et Latence Virale

#### 5.1 introduction

La thérapie génique va au delà de la modification des gènes « *humains* », elle peut permettre d'éliminer les génomes viraux pathogènes présents dans nos cellules. Dans son état naturel, le système CRISPR participe à la lutte antivirale, c'est donc avec une certaine logique et non sans ironie qu'il est aujourd'hui « *recyclé* » dans le but d'étendre cet effet à certain virus humains.

La découverte de petites protéines Cas9 dans des espèces bactériennes telles que *Neisseria meningitidis* et *Staphyloccus aureus* a déclenché des espoirs de production industrielles. Elles étaient assez petites pour être contenues dans un AAV (*adeno-associated virus*) qui peut être rapidement produit à grands volumes. Ces AAV offrent la possibilité de cibler directement l'ADN viral qui se cache dans des zones reculées du corps humain, tels que HBV et HSV (respectivement *Hepatitis B Virus* et *Herpes Simplex Virus*).

En ce qui concerne l'édition d'un ADN viral, la probabilité que la Cas9 coupe le génome de la cellule hôte de manière hors-cible paraît faible tant les séquences exogènes virales et endogènes humaines sont peu homologues. Voici trois exemples, chacun représente un espoir de parvenir à une première thérapie curative. Ils partagent des problématiques communes : atteindre les cellules d'intérêt, faire face au repli du matériel génétique ciblé pendant la latence, identifier la cible génétique et les mécanismes d'échappement pour obtenir une répression suffisante dans le but de prévenir l'apparition de formes mutantes résistantes.

# 5.2 HBV : *Hepatitis B Virus*

#### 5.2.1 Introduction

HBV déclenche une infection virale qui attaque le foie et peut causer à la fois une pathologie aigüe ou chronique. Le virus qui se transmet par contact avec le sang ou d'autres fluides biologiques de l'individu infecté est présent chez environ 240 millions de personnes de façon chronique (définie par une positivité aux antigènes de surface HBV durant au moins 6 mois).

Il s'agit d'un problème majeur de santé publique, on estime que 686 000 personnes meurent chaque année des suites de complications de la maladie (cirrhose ou carcinome hépato-cellulaire). Malgré l'existence d'un vaccin prophylactique et le développement d'inhibiteurs efficaces de la polymérase virale (Ténofovir, Entecavir), la forme chronique du virus Hépatite B reste une cause majeure de morbimortalité, particulièrement en Afrique et en Asie, zones de forte endémie où l'accès au vaccin est souvent limité, notamment pour la vaccination néonatale.

## 5.2.2 Cycle de vie

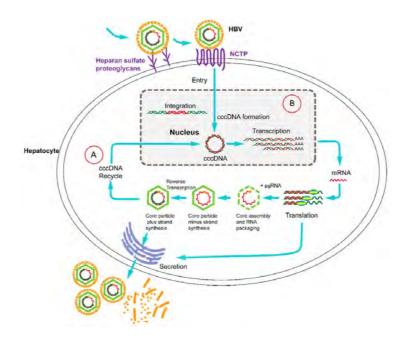


fig.36 Cycle de vie de HBV et cibles thérapeutiques175

Après fixation aux récepteurs de surface, HBV entre dans l'hépatocyte. La particule virale migre jusqu'au noyau, où le génome HBV sera converti en cccADN qui servira de matrice à la transcription virale. La traduction des ARNm dans le cytoplasme aboutit à la production de protéines de capside (C), polymérase (P), surface (S) et la protéine X du virus de l'hépatite B (HBx). Ensuite le génome viral ARN (pgARN) est emballé dans les progéniteurs de la capside virale. Après reverse transcription de son matériel génétique la nucléocapside obtenue sera soit encapsulée et sécrétée hors de l'hépatocyte ou encore re-importée au noyau pour être transformée à nouveau en cccADN. Le système CRISPR-Cas9 peut perturber le cycle HBV en atteignant le cccADN en A ou B. NCTP = le polypeptide co-transporteur Sodium Taurocholate nécessaire au cycle HBV, cccADN = covalently closed circular ADN.

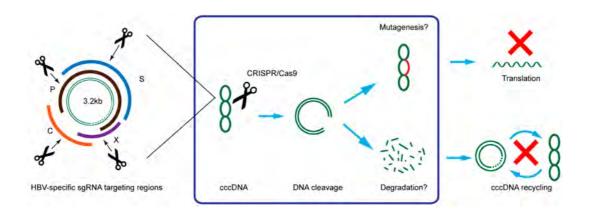
# 5.2.3 CRISPR comme solution thérapeutique

Les traitements actuels, les inhibiteurs de la polymérase virale, bien qu'ils empêchent la transcription inverse (la formation d'ADN génomique viraux à partir des transcrits d'ARN pré-génomique) ne parviennent pas à atteindre le cccDNA. Le cccDNA (pour *covalently closed circular DNA*) est la partie du génome viral qui persiste dans le noyau des cellules infectées. C'est ce très stable cccADN qui fait que le pourcentage d'infections par HBV complètement traitées par les inhibiteurs de polymérases est faible. Par conséquence il y à un grand intérêt pour la recherche d'une solution qui permettrait d'éliminer l'ADNccc dans le but d'obtenir une première thérapie curative.

C'est dont naturellement que le système CRISPR-Cas9 est apparu comme un candidat potentiel. Le génome HBV possède 4 cadres de lecture ouverts (Surface, Core, Polymérase et protéine X) qui seront traduits en seulement 7 protéines. Ces protéines sont hautement importantes pour la réplication du virus <sup>3</sup>. En théorie cibler l'une de ces

sept protéines à l'aide du duo Cas9/ARNg pourrait être suffisant pour supprimer les gènes viraux liés à l'expression et à la réplication.

Les premières expériences *in vivo* et *in vitro* montrant l'efficacité d'une approche CRISPR-Cas pour aller retirer définitivement HBV des cellules hépatiques furent réalisées en 2014. Lin et son équipe insérèrent l'ADN HBV et l'association Cas9/sgARN spécifique du cccADN dans des hépatocytes de souris par injection hydrodynamique. <sup>177</sup> Ils montrèrent que l'endonucléase guidée pouvait cliver HBV et permettre son élimination des cellules et faciliter sa clairance hors de l'organisme, entraînant la diminution des antigènes HBV de surface.



**fig.37** Schéma de la stratégie de suppression d'HBV par CRISPR-Cas9. Gènes  $\,$ C :core;  $\,$ P: polymérase;  $\,$ S: surface;  $\,$ X:  $\,$ HBx.  $^{175}$ 

### 5.2.4 Conclusions

Ces preuves de concept sont prometteuses mais pour être efficace la technique doit être capable d'éliminer tous les virus. L'un des obstacles est le fait que l'ADN HBV peut être retrouvé dans des tissus variés, autre que le foie. 178 Pour éradiquer l'infection, le CRISPR/ARNg devra être délivré à la totalité des cellules hépatiques atteintes et aux éventuels réservoirs extra-hépatiques. Dans le but d'y parvenir *in vivo* cela nécessiterait un vecteur doté de caractères particuliers. Les AAV recombinants pourraient répondre à ces critères. Ils sont peu immunisants, non intégratifs, ont un potentiel d'infection très élevé et possèdent naturellement un tropisme hépatique. 35

Hormis l'identification d'un vecteur, la prochaine étape sera d'évaluer cette technique sur un modèle murin dont le foie à été humanisé,c'est aujourd'hui le meilleur modèle préclinique pour évaluer la réplication HBV *in vivo* <sup>36</sup>. D'autres obstacles techniques restent à surmonter : bien qu'en théorie cibler une des protéines essentielles à HBV paraisse suffisant, c'est au moins 10 locus qu'il faudra modifier pour pallier aux mutations par modification du cadre de lecture qui pourraient restaurer un gène HBV actif malgré la coupure <sup>181</sup>.

En 2016 des obstacles persistent, mais les nucléases exogènes sont pour l'instant le seul moyen à notre disposition pour inactiver le cccADN, elles représentent par ailleurs un nouvel outil d'étude de la persistance du génome viral.

## 5.3. Herpesviridae.

## 5.3.1 Introduction

Les Herpesvirus sont de gros virus à ADN qui sont portés par presque la totalité de la population humaine. De nombreux pathogènes humains font partis de cette famille, voici quelques exemples :

- Herpes simplex virus (HSV type 1 & type 2 respectivement responsables de boutons de fièvre et d'herpes génital).
- Le cytomégalovirus humain ou HCMV est la cause virale la plus fréquente de malformations congénitales et le responsable de pathologies sévères de l'immunodéprimé.
- EBV ou Ebstein-Barr virus est associé à la mononucléose et un grand nombre d'autres pathologies telles que le carcinome nasopharyngé, la maladie de Hodgkin et le lymphome post transplantation.
- VZV ; virus de la varicelle et du zona.

Les Herpesvirus persistent à vie chez leur hôte. Ils y parviennent en établissant une infection latente, qui redevient périodiquement productive. Les antiviraux actuels tels que les analogues nucléosidiques Ganciclovir et Aciclovir agissent sur les manifestations cliniques de ces épisodes de réactivation en bloquant la réplication. Par contre leur impact est limité sur les formes latentes qui dépendent des polymérases cellulaires pour maintenir leur génome dans les cellules de l'hôte. Une stratégie pour éliminer la forme latente dépend souvent de leur réactivation par un agent chimiothérapeutique et le ciblage secondaire de ces cellules « *réveillées* » par Ganciclovir par exemple <sup>195</sup>. Bien qu'attrayante cette stratégie n'est pas toujours efficace et peut s'accompagner d'une mort de la cellule hôte, ce qui n'est pas tolérable dans certains cas (par exemple des neurones infectés par HSV1 et 2 ou VZV).

Il y a donc un besoin de thérapies alternatives et les recherches utilisent les nouvelles techniques découvertes pour répondre à cette nécessité. Publiée en Juin 2016, une étude néerlandaise<sup>183</sup> vise à évaluer la possibilité d'une thérapie curative grâce à la Cas9 en utilisant comme modèles des lignées cellulaires infectées par ces virus, elle va se heurter à des contraintes inattendues tout en apportant des espoirs.

## 5.3.2 CRISPR et Herpesviridae

Le but de cette étude est de combattre à la fois l'infection productive et latente grâce au système CrispR-Cas9, en ciblant des éléments génétiques viraux important pour leur survie. En se basant sur des cibles validées par de précédentes études ils parvinrent facilement à perturber la réplication de EBV, HCMV et HSV1 à l'aide de sgARN complémentaires de gènes viraux essentiels. Certains sgARN déclenchant une inhibition plus importante que d'autres, s'expliquant probablement par la cinétique d'expression de leur cibles.

Par contre stopper la latence s'avéra plus difficile. En ciblant EBNA1, un gène codant une protéine impliquée dans la maintenance d'EBV ils parvinrent à diminuer la latence chez seulement 50-60% des cellules en n'utilisant qu'un seul sgARN. En introduisant un second sgARN ciblant une zone différente de EBNA1, ils atteignirent le score encourageant de 95% de cellules « nettoyées ». La latence d'HSV-1 se révéla plus compliquée à perturber chez le modèle utilisé, le génome HSV-1 quiescent semble impénétrable aux sgARN utilisés.

Les mécanismes sur lesquels la latence repose sont différents en fonction des virus : Alors qu' EBV demeure sous la forme d'un épisome dans le noyau de cellules se divisant activement, HSV-1 est latent dans des neurones sensoriels où la réplication du génome n'a pas lieu184Cette différence pourrait expliquer l'écart d'efficacité observé. Durant la latence, le génome HSV-1 est hautement méthylé et l'expression du génome viral est restreinte. La Cas9 et son guide ARN seraient incapables d'accéder au génome d'HSV-1 alors que le génome de EBV, très exposé durant la réplication est accessible.

Mais cela ne reflète pas forcément l'action qu'aurait le système CRISPR-Cas9 *in vivo* sur HSV-1. Il sera nécessaire d'étudier ces mécanismes chez d'autres modèles *in vivo* et *in vitro* afin de mieux comprendre ces phénomènes et d'évaluer le potentiel curatif de la technique CRISPR-Cas9 dans le cas de de HSV-1.

#### 5.3.3 Conclusions

Il s'agit d'une première étude et elle porte à l'optimisme, bien que la mise en œuvre de la technique CRISPR-Cas9 n'est parfois pas aussi évidente que ce qui pourrait être imaginé. De nouvelles recherches utilisant de nouveaux modèles seront nécessaires pour explorer le potentiel de cette technique en ce qui concerne les herpesviridae. Chacune d'entre elles apportera de nouvelles découvertes emmenant à mieux comprendre des mécanismes complexes, que ce soient ceux des latences virales ou du fonctionnement des endonucléases ARN guidées.

## 5.4. Soigner HIV

## 5.4.1 Présentation

Le virus de l'immunodéficience humaine (VIH) est un rétrovirus qui s'attaque aux cellules du système immunitaire et les détruit ou les rend inefficaces. Aux premiers stades de l'infection, le sujet ne présente pas de symptômes. Cependant, l'évolution de l'infection entraîne un affaiblissement du système immunitaire et une vulnérabilité accrue aux infections opportunistes. Le syndrome d'immunodéficience acquise (SIDA) est le dernier stade de l'infection à VIH. Il peut se déclarer au bout de 10 à 15 ans. Les antirétroviraux permettent de ralentir son évolution. Le VIH se transmet à l'occasion de rapports sexuels non protégés, d'une transfusion de sang contaminé ou de l'échange de seringues contaminées. Il peut aussi se transmettre de la mère à l'enfant pendant la grossesse, l'accouchement ou l'allaitement au sein. Fin 2015,on estime que 36,7 millions de personnes vivent avec le HIV.

## 5.4.2 Cycle viral

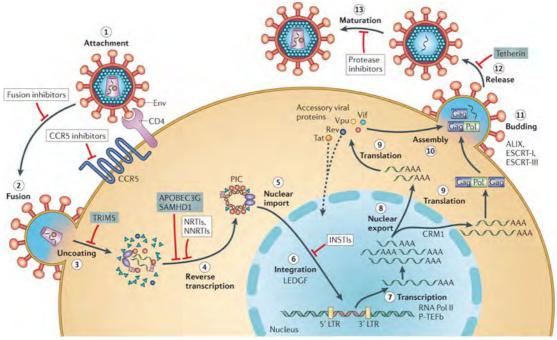


fig.38 Cycle viral de HIV-1 divisé en étapes 185

1 L'infection débute lorsque l'un des pics formés par la glycoprotéine d'enveloppe (Env) se lie au récepteur cellulaire CD4 et au récepteur trans-membranaire CCR5 (CC-Chemokine receptor 5). 2 Cela entraîne la fusion des membranes cellulaires et virales et l'entrée de la particule virale dans la cellule. **3&4** La perte partielle de l'enveloppe facilite la reverse transcription qui va produire le PIC ou preintegration complex. 5&6 Après transport dans le noyau les intégrases associées au PIC orchestrent la formation du provirus intégré au génome, aidé par les protéines de liaison à la chromatine de l'hôte et le LEDGF (lens epithelium derived growth factor).7 La transcription du provirus réalisée par l'ARN polymérase II de l'hôte (ARN Pol II) et le facteur d'élongation positif b (P-TEFb); produit des ARNm viraux de tailles différentes.8 Les plus longs quittent le noyau par le transporteur CRM1. 9&10 Les ARNm serviront de matrice à la production protéique et de longs ARN seront incorporés dans les particules virales avec des composés protéiques. 11&12 Le bourgeonnement de la particule virale et sa libération hors de la cellule est contrôlé par les complexes ESCRT (pour endosomal sorting complex required for transport) et ALIX et s'accompagne (13) d'une maturation par les protéases afin de créer une particule virale infectieuse. LTR: longues répétitions terminales. INSTI: integrase strand transfer inhibitor, NNRTI non nucleoside reverse transcriptase inhibitor, NRTI nucleoside reverse transcriptase inhibitor. Les sites d'action des inhibiteurs cliniques sont indiqués en blanc sur la fig.38 ceux des facteurs de restriction cellulaires sont en bleu.

#### 5.4.3 Traitement et latence

Chacune des étapes du cycle de vie de HIV-1 est une cible potentielle pour l'intervention d'antiviraux affectant la réplication. Le standard actuel est la thérapie antirétrovirale, débutée au milieu des années 90 et surnommée « *triple thérapie ou trithérapie* ». Elle régule la pathologie mais ne la traite pas . Son action permet de diminuer la réplication de HIV-1 à des niveaux inférieurs aux seuils de détection, mais le génome persiste sous forme d'infection latente chez un petit nombre de lymphocytes T CD4+ en phase de repos <sup>186</sup>. Le patient doit donc prendre à vie un traitement onéreux et souvent mal toléré, tout en étant le plus adhérent possible pour éviter la réactivation à partir des réservoirs de latence.

Pour espérer soigner HIV, il faut comprendre le phénomène de latence. Chez les LT

CD4+; cellules à longue durée de vie, les provirus HIV-1 persistent dans un état dit silencieux, dépourvus d'épisodes transcriptionnels. Cet état est réfractaire à la fois à la réponse immunitaire et aux médicaments actuels. Cependant les cellules T mémoires peuvent être réactivées par un antigène appelé « *antigène de rappel* ». Cette réactivation entraîne l'induction d'un cycle productif de réplication virale <sup>41</sup>. Si cette réactivation se fait hors traitement, le HIV-1 pourra se répandre au sein des T CD4+ disponibles et relancer un niveau similaire de réplication à celui observé avant traitement.

Les efforts pour éliminer ce réservoir de cellules contenant l'infection latente se concentrent sur deux stratégies. La première à été prénommé en anglais « *Shock and Kill* » / « *Choquer puis tuer* ». Elle consiste à ré-activer les cellules latentes (c'est le choc) dans le but de les rendre à nouveau sensibles (pour les « *tuer* »). Des composés comme les inhibiteurs de l'histone déacétylase et les agonistes PKC en sont capables <sup>187</sup>, cependant aucune stratégie n'a montré une efficacité chez un pourcentage suffisant de cellules infectées. Une alternative à pour but d'éliminer l'ADN proviral de chacune des cellules infectées. Ce concept a été testé avec les recombinases modifiées, les ZFN et les TALENS, avec des essais cliniques allant jusqu'en phase II¹88.La technique CRISPR-Cas9, plus précise, plus rapide et surtout plus simple,a catalysé les expérimentations dans ce domaine.

## 5.4.4 Cibles : effacer le génome HIV1

Le cas de Thimothy Brown, le « *Patient de Berlin* », a déclenché de nombreux espoirs et inspire encore aujourd'hui de nombreuses recherches. Atteint de HIV et de Leucémie Aiguë Myéloïde il reçut en 2009 deux greffes de cellules souches hématopoïétiques. Ces cellules issues d'un individu porteur de la mutation Δ32 homozygote sont dépourvues de récepteurs CCR5. Suite à cette intervention son taux de HIV sanguin est devenu indétectable et il n'est plus nécessaire pour lui de prendre une thérapie antirétrovirale.



**fig.39** Timothy Brown en 2011 194

La guérison de Timothy Brown accrut l'intérêt du génie génétique pour CCR5 en tant que cible thérapeutique. En théorie si l'expression de CCR5 pouvait être bloqué, l'entrée du HIV dans ses cellules hôtes deviendrait quasi-impossible. Les trois outils d'éditions génomiques ZFN, TALENS<sup>75,190,191</sup> et désormais CRISPR-Cas9 ont été mis en œuvre pour perturber CCR5 mais plusieurs inconvénients liés à CCR5 apparurent <sup>192</sup>. CCR5 n'est pas seulement un récepteur de HIV-1, il occupe dans la cellules plusieurs autre rôles. Il n'est pas indispensable au HIV-1, certains virus utilisent le co-récepteur alternatif CXCR4 rendant inefficace le ciblage de CCR5 dans leur cas. Un inconvénient supplémentaire est qu'il affecte seulement l'entrée du virus dans la cellule hôte et sa perturbation (ou celle de CXCR4) n'impacte pas les cellules déjà infectées par le virus.

Comme pour les infections par Herpesviridae, hépatite B ou autre virus latents, il est

donc nécessaire d'atteindre le génome intégratif de HIV-1 chez les cellules où l'infection est latente pour obtenir une thérapie permanente ou stérilisante. Pour atteindre l'ADN proviral, la destruction de gènes codant pour certaines protéines de HIV-1 est une des techniques explorées mais elle n'offre pas la possibilité d'éliminer l'intégralité du génome et les reliquats de LTR (*long terminal repeats*) persistants pourraient continuer à promouvoir, après le clivage, la transcription de protéines qui peuvent s'avérer toxique. 193,194

#### 5.4.5 Les LTRs

Les LTRs sont des séquences hautement conservées présentes en 5' et 3' du génome viral intégré. S'attaquer à ces LTRs apparaît donc comme l'approche idéale, cette méthode aurait un double effet, elle permettrait de supprimer l'intégralité du génome situé entre chacun des LTR et impacterait aussi l'expression des séquences LTR elles-mêmes. Cette hypothèse a été testée avec succès en utilisant la technologie Cas9/LTR-sgARN, dans de nombreux types cellulaires comprenant les lymphocytes T et les cellules de la lignée myéloïde qui sont des réservoirs majeurs de l'infection latente HIV-1. Ce système est aussi capable d'éliminer plusieurs copies de HIV-1 par cellules par approche multiplex (avec plusieurs sgARN), même si elles sont situées sur des chromosomes différents<sup>195,196</sup>

#### 5.4.6 Première étude in vivo

Dans une publication parue dans Nature en mai 2016, Rafal Kaminski et ses collègues décrivent comment ils sont parvenus à réaliser la première preuve de concept d'élimination du génome HIV *in vivo*. Ils ont travaillé sur des souris et des rats transgéniques dont le génome est doté de séquences HIV-1.

Ils choisirent un orthologue de la Cas9 de petite taille, la saCas9 ( Staphylococcus aureus Cas9) et une approche multiplex basée sur plusieurs guides sgARNs. L'injection de la queue des souris transgéniques avec un vecteur rAAV9 (un adeno-associated virus 9 recombinant) qui exprime saCas9 et ses ARN guides à permis de créer une coupure de l'ADN HIV-1 et d'exciser un fragment d'ADN de 978 pdb qui s'étend entre les gènes LTR et Gag. L'atteinte des organes est complète, on retrouve la perturbation du génome viral intégré à la fois dans le rein, le foie, le cœur, les poumons le rein et au niveau des lymphocytes circulants. L'injection rétro-orbitale chez le rat transgénique de rAAV9:saCas9/gARN élimine les séquences virales ciblées et diminue significativement le niveau d'expression virale dans les lymphocytes du sang circulant. Cette preuve de concept est la toute première fois qu'une suppression de HIV-1 est réalisée in vivo par CRISPR-Cas9 sur un panel de cellules et tissus qui portent une copie intégrée du génome viral.

#### 5.4.7 Limites

#### 5.4.7.1. liées à l'outil d'édition génomique

Une incertitude de taille à l'application clinique et en recherche fondamentale est la possibilité d'effets hors-cible. Par rapport au ciblage des séquences « humaines » le ciblage du génome viral présent dans les cellules humaines semble présenter un risque moindre. La spécificité des ZFN, TALEN et Cas9 à été étudiée en

détail par de nombreuses expériences *in vivo* et *in vitro*, par séquençage et *computational characterization*. Bien que des essais complémentaires et une amélioration continue de ces outils soient nécessaires, il n'est pas apparu de génotoxicité, de cytotoxicité ou d'effets hors cible dans la majorité des études de thérapie génique virale. 84,196-198

Nous n'avons pas encore le recul nécessaire pour évaluer la toxicité d'une expression à long terme de la Cas9/sgARN dans une cellule. Pour l'instant les animaux transgéniques qui expriment la Cas9 et ses guides ARNs complémentaires des génomes viraux de manière constitutive n'ont pas déclenché de toxicité <sup>200–202</sup>. Un suivi par des méthodes de séquençage du génome entier capables de s'assurer de la stabilité du génome et de la potentielle génotoxicité sera nécessaires. En ce qui concerne l'édition d'un ADN viral, la probabilité que la Cas9 coupe le génome de la cellule hôte de manière hors-cible paraît faible tant les séquences exogènes virales et endogènes humaines sont peu homologues.

Il existe aujourd'hui un seul exemple où une homologie entre la séquence ciblée et le génome cellulaire a été observé : les sites de liaison au facteur de transcriptions compris dans le LTR de HIV-1. Dans ce cas un ciblage imprécis peut être évité par sélection des gARN. Des études bio-informatiques des cibles et de leur homologie avec le génome de l'hôte permettront d'améliorer la spécificité de l'édition génomique et de faire basculer les attentes dans le domaine du possible.

#### 5.4.7.2. liées au HIV

Une question à laquelle toute thérapie ciblant HIV doit répondre se pose : HIV-1 pourrait il échapper à l'attaque programmée par CRISPR-Cas9? Si oui, comment?

Rapidement après la publication de Rafal Kaminski et ses collègues, deux publications  $^{203,204}$  ont apporté des réponses inattendues à cette interrogation. La capacité HIV-1 à acquérir des mutations qui lui permettent d'échapper aux attaques par interférence ARN (ARNi)  $^{205,206}$  était déjà connue. Comme on pouvait s'y attendre le virus utilise une technique similaire qui interfère avec la reconnaissance par les sgARN et la Cas9.

Mais dans le cas de CRISPR ce qui est surprenant est la concentration de ces mutations au niveau exact des sites où la Cas9 est supposée couper l'ADN viral, alors que la zone de complémentarité avec le sgARN est bien plus grande. Une autre particularité est la haute fréquence d'insertions et de délétions (*indels*) présents au niveau des séquences virales les moins conservées. Ces propriétés suggèrent que les mutations qui confèrent la résistance à la Cas9/sgARN ne sont pas le résultats d'erreurs fréquentes de la reverse transcriptase virale (qui à une tendance naturelle à la confusion).

L'origine de ces mutations serait très probablement liée à la machinerie NHEJ (non homologous end joining) qui répare l'ADN brisé au site de coupure. Voici une hypothèse du mécanisme de résistance à la Cas9 développée par HIV : La Cas9 dirigée par sgARN clive l'ADN cible au niveau d'un site situé à 3 nucléotides du PAM. Lorsque la machinerie NHEJ répare la cassure double brin l'insertion de courts nucléotides, la délétion et la substitution peuvent se produire au niveau du site de clivage. La majorité de ces mutations abolissent le fonctionnement du gène et à terme tuent le virus. Cependant, certaines mutations ne causent aucun dommage au virus et favorisent l'échappement viral.

#### 5.4.8 Conclusions

Bien que ces résultats paraissent décourageants, il est important de les remettre dans leur contexte. Les expériences ont été réalisées en culture cellulaires, ces cellules n'ont été infectées par HIV-1 qu'après expression stable de Cas9/sgARN. La production virale est bien supérieure dans ce modèle que dans celle d'un modèle d'infection latente par HIV-1. Cela facilite l'échappement du virus et sa propagation aux cellules adjacentes. Il sera important d'évaluer l'échappement à CRISPR-Cas9 dans un modèle animal. Si ces mutations échappatoires sont fréquentes, la stratégie « *shock & kill* » par dCas9 pourrait être une meilleure option qu'un clivage direct. Malgré les potentielles difficultés qui seront à surmonter pour traduire ces découvertes en thérapies, ces publications apportent avec elles l'idée que soigner HIV pourrait un jour être à notre portée.

Au delà de son usage dans la latence virale Crispr cas9 à été utilisé pour éradiquer des virus porcins endogènes (virus PERVs) de l'intégralité du génome porcin, afin de prévenir le risque de transmission durant une transplantation <sup>90</sup>. L'un des apports majeurs de la technique qui ne sera qu'évoqué ici est son utilisation en recherche fondamentale. Elle pourra renforcer la compréhension des mécanismes infectieux, à l'interface cellule virus, notamment par la réalisation de criblage dans le but de comprendre les facteurs cellulaires impliqués dans la réplication virale.

## 6. Cancérologie

#### 6.1 Introduction

Parmi les indications de la thérapie génique, la cancérologie est celle qui a le plus été traduite en études clinique. Bien qu'apportant des résultats mitigés elle a servi de tremplin clinique à la majorité des approches thérapeutiques <sup>207</sup>. C'est donc logiquement que la technique CRISPR-Cas9 réalisa ses débuts cliniques en cancérologie. Ce fort pourcentage d'études cliniques s'explique par le fait que la thérapie génique appliquée au cancer est moins sujette aux limitations techniques qui peuvent s'appliquer aux maladies monogéniques par exemple <sup>205</sup>.

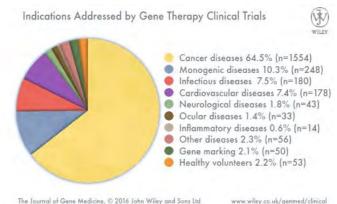


fig.40 Répartition des indications en thérapie génique<sup>209</sup>

Plusieurs approches existent, on peut citer le gène suicide qui permettent la transformation d'une prodrogue non toxique en drogue toxique (ex : HSV tk), les gènes suppresseur de tumeurs entraînant l'apoptose (ex : gène H19), les siARN (silencing ARN)<sup>210</sup>, la modification de gènes (par exemple l'introduction de gènes radiosensibilitants ou radioprotecteurs destinés aux cellules à proximité des tumeurs<sup>211,212</sup>) et enfin l'édition génomique et l'immunomodulation.

#### 6.2 immunomodulation anticancéreure et édition génomique

Parmi les stratégies thérapeutiques anticancéreuses basées sur la modulation du système immunitaire figure l'immunothérapie adoptive. Cette forme d'immunothérapie utilise les cellules du patients (lymphocytes T ou Cellules Présentatrices d'Antigènes dendritiques) puis les stimulent *ex vivo* avant de les réintroduire.

La thérapie cellulaire apporte des résultats prometteurs, mais nombreux sont les patients qui présentent une rechute. Les techniques d'éditions génomiques pourraient donc améliorer de tels traitements et réduire leur vulnérabilité. En 2015 la première réussite clinique des TALENs dans une étude pédiatrique de phase 1 ciblant la leucémie aiguë lymphoblastique à cellules B (B-ALL) était basée sur l'association entre édition génomique et thérapie cellulaire <sup>207,213</sup> lymphocytes CAR-T (des lymphocytes T artificiellement modifiées pour leur conférer un récepteur chimérique, ici capables de cibler l'antigène CD19 présent sur les cellules cancéreuses). Mais au lieu d'utiliser des lymphocytes autologues (du patient) cette étude avait pour ambition d'utiliser des lymphocytes hétérologues (provenant de donneurs). L'édition génomique par TALENs

était utilisée pour permettre de contourner le système HLA et prévenir le rejet lié à l'hôte. Les chercheurs parvinrent à des UCAR-T : des CAR-T universels qui pourraient grandement simplifier et standardiser la production de telles solutions thérapeutiques. En 2016 c'est deux équipes qui se livrent un duel pour parvenir à une première utilisation de CRISPR chez l'homme, elle concerne également l'association entre thérapie cellulaire et édition génomique.

## 6.3 débuts cliniques : immunothérapie antitumorale et CRISPR-Cas9

Une compétition biomédicale a lieu, certains nomment sputnik 2.0 cette course à l'application clinique de CRISPR-Cas9 <sup>214</sup>. D'un côté l'université de Pennsylvanie avec le soutien financier du *Parker Institute for Cancer Immunotherapy* (la fondation créé par Sean Parker, milliardaire fondateur de Napster et pionnier de Facebook). De l'autre des institutions chinoises qui ont déjà la primeur des premiers essais chez le singe et l'embryon <sup>160,214</sup>. Les deux équipes prévoient une approche similaire, prélever puis éditer le génome de lymphocytes T des patients pour y éteindre le gène PD-1 avant de les réintroduire dans la circulation sanguine. La voie anti-PD-1 est une voie validée depuis peu qui connaît un essor considérable avec l'accès au marché d'anticorps anti-PD-1 nivolumab et pembrolizumab<sup>216,217</sup>.

PD-1 ou *programmed cell death 1* est une protéine de surface cellulaire exprimée par les lymphocytes T, l'un de ses ligands PDL-1 exprimés par les cellules saines est un des *checkpoint* qui permet d'empêcher l'activité des LT et NK, permettant une régulation de l'immunité et une tolérance au *soi* <sup>218</sup>. C'est cette propriété qui est détournée par des cellules tumorales qui expriment PDL-1 et échappent alors à l'immunosurveillance. La création de lymphocytes T dont le gène codant PD-1 est inactivé par la cas9 guidée par sgARN permet d'obtenir une réponse immunitaire boostée, dirigée contre la tumeur<sup>219</sup>. S'agissant d'une manipulation ex-vivo, le criblage et la sélection des LT avant injection est possible et permet de s'affranchir des préoccupations d'effets hors cible de la cas9. Par contre les risques d'auto-immunité inhérente à la technique persistent.

Des deux équipes en lice après approbation de leur projets d'essais cliniques par leurs comités respectifs <sup>220</sup> c'est celle du *West China Hospita*l, hopital universitaire de la *Sichuan University* à Chengdu qui fut la première à avoir le feu vert pour débuter leur étude. Le 28 octobre 2016 fut débuté l'essai NCT02793856 <sup>220</sup> examinant l'efficacité et la sécurité d'une injection de lymphocyte T modifiés chez des patients atteints de cancer du poumon métastatique non à petites cellules (*non-small cell lung cancer*) réfractaires aux traitements traditionnels. Cet essai de phase I a recruté 10 patients qui recevront entre 2 et 4 injections successives puis seront suivies pendant 6 mois., au moment où cette thèse est écrite, leur résultats n'ont pas encore été publiés.

#### 6.4 Conclusions

Malgré les promesses des approches de thérapie génique en immunothérapie antitumorale, et même si elles parviennent à un résultat positif, elles restent aujourd'hui soumises à une contrainte de taille. Leur procédé implique aujourd'hui l'extraction, la modification, la sélection et la multiplication de cellules du patient, une lourde entreprise. À moins qu'elles démontrent un gain considérable d'efficacité, il est pour l'instant difficile de justifier la poursuite de tels travaux. Si on mesure le succès seulement par la facilité d'évoluer vers une production massive, les anticorps semblent

bien supérieurs, et peuvent être produits en quantités presque illimitées dans le but d'un usage clinique.

Mais comme nous l'avons vu dans les deux exemples précédents l'édition génomique pourrait permettre de révolutionner l'immunothérapie antitumorale en améliorant l'efficacité des traitements ou en leur permettant de s'affranchir des inconvénients de leur production. Cette approche pourrait alors devenir une alternative complémentaire aux autres stratégies thérapeutiques.

#### 6.5 Perspective: Thérapie génique CRISPR, production et commercialisation

De la même manière que les protéines/anticorps cliniquement actifs, des systèmes de production qui obéissent aux bonnes pratiques de fabrication (BPFs) sont nécessaires pour maintenir la sécurité du patient et limiter les risques liés à la production de composés destinés à l'édition génomique. Si elles voient le jour, les essais destinés à la sécurité et la production à large échelle de ces produits thérapeutiques devront satisfaire aux bonnes pratiques génomiques <sup>221</sup>. Les systèmes de délivrance choisis devront eux aussi satisfaire aux bonnes pratiques de fabrication actuelles.

Pour être rentable, l'ensemble de cette production devra être réalisable à un coût qui permet de démontrer un bénéfice pharmaco-économique afin d'obtenir le remboursement du produit par les mécanismes d'assurance maladie.

Les outils de production et de contrôle qualité actuels sont utilisables pour manufacturer des produits de thérapie génique aux standards BPFs. Pour l'édition de gènes humains par une approche le défi sera d'améliorer la compréhension du mécanisme d'action et d'augmenter suffisamment la spécificité de la coupure par cas9 afin de minimiser l'apparition d'effets hors-cible à un seuil acceptable. L'approche *ex vivo* est elle soumise aux mêmes contraintes que la thérapie cellulaire. La propriété intellectuelle jouera aussi un rôle clef dans la commercialisation de la technologie CRISPR. Les trois compagnies développant son usage commercial (*Editas Medicine, CRISPR Therapeutics*, et *Intellia*) devront prendre des décision légales majeures destinées à protéger chacun des aspects spécifiques à leur technique. L'incertitude liée à l'attribution de licences et les limites actuelles liées à l'utilisation de CRISPR en clinique ne semblent pas pour l'instant freiner les relations entre ces compagnies et l'industrie pharmaceutique (entre autres Novartis, Bayer, Juno, Vertex et Genethon).

# II Des antibiotiques CRISPR

## 1. Introduction

Les antibiotiques ont révolutionné la médecine, de nombreuses vies ont pu être sauvées, et leur découverte peut être considérée comme un tournant majeur de

l'histoire humaine. Malheureusement leur utilisation fut rapidement accompagnée par l'apparition de souches bactériennes résistantes.

Depuis leur introduction, des millions de tonnes d'antibiotiques ont été produits et employés dans des buts variés. L'amélioration des techniques de production a entraîné une diminution constante des prix des antibiotiques, encourageant leur usage hors prescription et hors indication. En conséquence, le monde est saturé par ces agents, ce qui contribue à la sélection de souches résistantes. Le développement de bactéries résistantes aux antibiotiques et leur distribution dans les populations microbiennes à travers l'ensemble de la biosphère est le résultat de nombreuses années de pression sélective continue, générée par l'homme à travers son mésusage. Il ne s'agit pas d'un événement naturel, mais d'une situation artificielle,« il n'y a peut être pas de meilleur exemple de notion darwinienne de sélection et de survie » 2222.



fig.41 Assemblée Générale des Nations Unies, 13 Septembre 2016

En Septembre 2016; pour débuter son discours devant l'assemblée générale des Nations Unies, Margaret Chan, directeur général de l'OMS qualifia la résistance aux antibiotiques de crise globale, elle décrit ce phénomène comme un « *Tsunami qui se rapproche de nous au ralenti* » <sup>223</sup>. Elle rajoute ensuite :

« La mauvaise utilisation des antibiotiques, que ce soit leur sous-utilisation et l'abus, est en train de causer l'échec de ces fragiles médicaments » et « l'émergence de résistances bactériennes est en train de prendre de cours notre capacité à découvrir des antibiotiques »

Les bactéries ont su tirer parti de la sur-utilisation des antibiotiques. Leur extraordinaire capacité génétique les a poussées à exploiter toutes les possibilités de résistances à leur disposition et tous les moyens de transmission génétique horizontale pour développer des mécanismes de résistances contre chacun des antibiotiques utilisés, que se soit en pratique clinique, en agriculture ou ailleurs. De nombreuses bactéries pathogènes responsables d'épidémies chez l'homme ont évoluées pour acquérir une multirésistance ( ou MDR : pour Multi Drug Resistance).

L'usage extensif des antibiotiques dans l'élevage animalier est un contributeur principal de l'apparition de résistances <sup>224–226</sup>. (Voir en Annexe 9 le schéma de dissémination de résistances aux antibiotiques entre élevage, hôpitaux, traitement des eaux usées et communauté). L'interdiction en Europe de l'antibiothérapie animale dans le seul but de favoriser la croissance des élevages fait partie de la stratégie de l'Union pour lutter contre l'émergence de résistances<sup>6</sup>. Jusqu'au début du XXIème siècle on avait observé des résistances d'

Escherichia coli à tous les antibiotiques connus sauf un, la Colistine, un antibiotique ancien de la famille des polymyxines, peu utilisé chez l'homme. L'usage extensif dans les fermes d'élevages porcin en Chine, a déclenché l'apparition de souches résistantes (conféré par le gène plasmidique mcr-1), conduisant fin 2015 à la détection d'une bactérie résistante à tous les antibiotiques connus <sup>228</sup>. (Différents mécanismes de résistance aux antibiotiques sont représentés en Annexe 10)

Les initiatives gouvernementales sont nombreuses pour lutter contre cette situation préoccupante<sup>229</sup>, et les partenariats public-privé se multiplient, l'un des axes majeurs : le développement d'antibiotiques innovants. Parmi les nouveaux antibiotiques et les nouvelles solutions thérapeutiques recherchées pour contrecarrer l'émergence de résistance, CRISPR-Cas est l'une des moins traditionnelles.

## 2. La technique CRISPR-Cas9: espoir d'un nouveau type d'antibiotiques

Les eucaryotes maintiennent l'homéostasie au sein des cellules et des tissus par le biais de systèmes de réparation d'ADN robustes. Contrairement à eux, les bactéries favorisent les mutations et l'acquisition de matériel génétique nouveau tels que les plasmides pour pouvoir assurer une flexibilité évolutionnaire.

La plupart des bactéries possèdent donc de faibles mécanismes de réparation (majoritairement par NEHJ). Une coupure double brin induite par CRISPR au sein du génome bactérien déclenchera donc souvent leur mort. Si cette coupure DB à lieu au niveau d'un plasmide, il sera éliminé de la bactérie (ce qui peut dans certain cas être une technique pour induire la mort de la bactérie). La technique CRISPR permet une grande spécificité, plutôt que de viser des familles bactériennes, le système CRISPR permettrait de cibler certains génotypes pour atteindre un genre ou une espèce bactérienne choisie. La technique peut aussi servir à cibler un gène de résistance à un antibiotique dans le but re-sensibiliser la cible<sup>230,231</sup>. Les antibiotiques séquences-spécifiques ainsi créés peuvent donc moduler sélectivement n'importe quelle population bactérienne et éliminer les pathogènes. Cette éradication précise pourrait permettre la survie de bactéries commensales bénéfiques, offrant un avantage par rapport aux antibiotiques à large spectre, moins sélectifs. Ils pourraient servir pour façonner à dessein le microbiome et il est imaginable que ces antimicrobiens pourraient aussi permettre d'atteindre des parasites eucaryotes.

#### 3. exemples

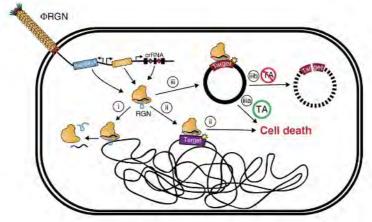
En 2014 Rob Citorik <sup>231</sup>, David Bikard<sup>230</sup> et leurs collaborateurs développèrent pour la première fois des antimicrobiens basés sur le le système CRISPR-Cas. Les deux équipes utilisèrent des gARN ciblant des séquences retrouvées seulement chez certaines bactéries pour guider la Cas9 :

- Si la séquence cible est absente, cela ne déclenche aucun effet.
- Si elle est chromosomique, le clivage est cytotoxique
- Si elle est épisomique, le clivage entraîne soit la mort cellulaire soit la perte du plasmide en fonction de la présence ou l'absence de système toxine-antitoxine

Rob Citorik et ses collègues choisirent les gènes de résistance aux beta-lactames et aux quinolones chez *Escherichia coli* comme cible et utilisèrent un plasmide ou un phagemid pour délivrer le système CRISPR dans la cellule bactérienne cible. (Les Phagemids sont des plasmides contenant une origine de réplication d'origine phagique qui peuvent être emballés dans une capside de bactériophage, ils sont incapables de se répliquer sans la présence d'un système *helper*). David Bikard et ses collègues utilisèrent un phagemid pour cibler *Staphylococcus aureus*, un pathogène hospitalier commun.

## 3.1 Travaux de l'équipe de Citorik

En emballant le système CRISPR-Cas dans un phagemid ou en exploitant le système plasmidique, ils créèrent des outils permettant de déclencher une létalité conditionnelle, modulaire et mutiplexable (pouvant avoir plusieurs cibles génétiques) dirigée contre des séquences d'intérêt. Pour que la cassure générée par l'endonucléase ARN guidée déclenche une cytotoxicité, ils la dirigèrent vers les sites plasmidiques blaSHV-18 ou blaNDM-1 qui codent respectivement pour une résistance large spectre et une pan-resistance aux β-lactamines²³²². Pour ne pas se limiter aux résistances enzymatiques (β-lactamases) et plasmidiques, ils étudièrent un cas de résistance portée par le génome bactérien lié à l'altération d'une protéine de l'hôte: la résistance aux quinolones. Les bactéries résistantes aux quinolones ont une mutation de la séquence codant la gyrase (gyrA), l'une des protéines ciblées par ces antibiotiques. Par rapport aux bactéries sauvages la séquence de ces bactéries ne diffère que d'un seul nucléotide, ce qui en fait un prototype de choix pour tester la spécificité de CRISPR-Cas en tant qu'antibiotique.



**fig.42**: Mécanisme de l'antibiotique CRISPR conçu par l'équipe de Rob Citorik
La construction portant la Cas9/tracrARN:crARN ici appelé RGN pour *RNA Guided Nuclease* est délivrée par un Phagemid. L'ensemble étant dénommé par l'acronyme ΦRGN. Ce ΦRGN démontre des effets antimicrobiens efficaces et spécifiques dirigés contre les souches portant le plasmide ou le chromosome contenant la séquence qu'il cible. Le ΦRGN affecte différemment la physiologie de la cellule hôte de manière séquence dépendante. Si la séquence cible est : **I** absente, le RGN ne déclenche aucun effet; **II** chromosomique, l'activité du RGN est cytotoxique; **III** épisomique, le RGN entraîne soit **IIIa** la mort cellulaire ou **IIIb** la perte du plasmide en fonction de la présence ou l'absence de système toxineantitoxine (TA).

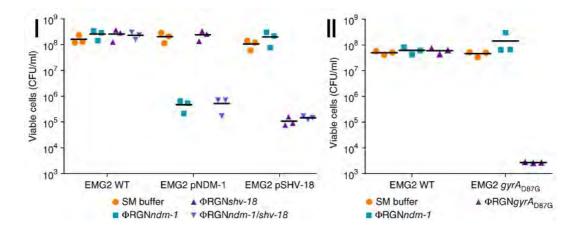


fig.43 Démonstration de l'activité souche dépendante de l'antibiotique CRISPR

I Trois souches de bactéries EMG2 sont traitées : soit la souche sauvage (WT; pour wild type) soit deux souches contenant des séquences actives de résistance plasmidique : pNDM-1 (codant blaNDM-1 conférant une résistance large spectre) ou pSHV-18 (codant blaSHV-18 coférant une résistance aux βlactamines). Ces différentes souches sont traitées soit par tampon SM, soit par PRGN: un phagemid contenand le dublex Cas9/ARN guides qui aura pour cible la séquence ndm-1 (on le nommera alors ΦRGNndm-1), ou la séquence shv18 (ΦRGNshv-18) ou les deux ΦRGNndm-1/shv-18 multiplex. Les souches bactériennes sont traitées à un MOI (multiplicity of Infection) ~20. Ces traitements entraînent une cytotoxicité séquence dépendante, prouvée par la réduction spécifique du nombre de cellules viables. (CFU = colony forming unit).La transformation d'E.coli avec le système CRISPR-Cas entraîne une diminution de la population bactérienne portant la séquence cible d'un facteur 1000 comparé aux bactéries qui en sont dépourvues. L'analyse séquentielles des 30 mutants (les cellules survivantes dotées de la séquence cible et ayant pourtant reçus et incorporés le plasmide) révéla que cette tolérance était exclusivement dus à une construction défectueuse du plasmide qui provient souvent de la délétion du spacer dans le locus CRISPR. Ces antibiotiques CRISPR ont re-sensibilisé avec succès des populations bactériennes résistantes aux béta-lactamines, malgré le fait que ces gènes de résistance soient codés par un plasmide à haut nombre de copies. (Un exemple de plasmide contenant le système destiné à cibler le gène ndm1 est représenté en annexe 11)

II Des populations *Eschirichia.coli* EMG2 WT ou son variant comportant la mutation gyrase : EMG2 gyrAD87G son traitées soit par Tampon SM, ΦRGNndm-1 ou ΦRGNgyrAD87G à MOI ~20. La résistance aux quinolones est codée par le génome bactérien, l'utilisation d'un CRISPR Phagemid dirigé contre la séquence mutée a été cytotoxique seulement chez les bactéries résistantes.

## 3.2 Travaux de l'équipe de Bikard

David Bikard et ses collègues étudièrent une approche *in vivo*, en montrant l'éfficacité de la méthode sur un modèle de colonisation de peau de souris par *Straphylococcus aureus*. Pour eux : « *cette nouvelle technologie offre l'opportunité de manipuler des populations complexes de bactéries, de manière séquence-spécifique* ». Pour y parvenir, le système CRISPR de type II devait être délivré à autant de cellules cibles que possible, sans nécessiter de sélection préalable et de manière à ce que l'expérience soit reproductible dans l'environnement naturel.

Les bactériophages transportent naturellement de l'ADN dans leur capsides qu'ils peuvent ensuite injecter dans l'hôte bactérien. Ils adoptèrent donc une technique de délivrance des gènes cas9 et leurs ARN guides en utilisant un phagemid<sup>233</sup>. Cette stratégie étant inspirée partiellement par la récente découverte de phages transportant leur propres systèmes CRISPR.<sup>234</sup>

Ils testèrent si cette technologie pouvait sélectivement tuer les souches de *Staphylococcus aureus* virulentes et antibiorésistantes. Les *Staphylococcus* sont à la fois des éléments prédominants du microbionte dermique<sup>13</sup> et l'une des causes d'infections nosocomiales les plus communes<sup>235,236</sup>. L'augmentation récente de la pathogénicité des staphylocoques est due en grande partie au transfert de gènes de résistance aux antibiotiques et de virulance via conjugaison plasmidique ou d'autres éléments mobiles. Cette évolution a entraîné l'augmentation des cas hospitaliers et communautaire de SARM et de SARV : soit *Staphylococcus aureus* résistant à la methicilline et *Staphylococcus aureus* résistant à la vancomycine, deux souches très difficile à traiter.<sup>237,238</sup>

Deux séries de traitement par phagemid suffirent à éradiquer la souche virulente. Les bactéries qui n'étaient pas tuées par la première série de traitements, y réchappèrent car elles n'avaient soit pas reçu le phagemid, soit l'avaient perdu ou en avaient reçu un défectueux. Ils n'observèrent aucune mutation du site de liaison qui auraient pu expliquer l'inefficacité perçue. *In vivo*, sur un modèle de colonisation de peau de souris, le traitement par phagemid diminua la proportion de *Staphylococcus aureus*, la faisant passer de 50% à 11% en seulement 24 heures.

#### 4. Conclusions

## 4.1 liés au développement d'antibiotiques CRISPR

Bien que prometteuse, la technique n'en est encore qu'à ses débuts, les deux études apportent de formidables preuves de concept dans le ciblage précis de population bactérienne et de leur résistance, mais de nombreuses problématiques restent sans réponse.

Un problème clef de cette technique est la réalisation d'un vecteur efficace. L'antibiotique doit atteindre chacune des cellules pour être efficace, sinon les populations de bactéries épargnées seraient capables de recoloniser l'environnement après le traitement. Le niveau de pénétrance dans une population bactérienne nécessite donc d'être étudié. Les deux études montrèrent des réductions de population bactérienne d'un facteur 1000 *in vitro*, mais *in vivo*, la réduction ne fut que d'un facteur 10 et l'effet protecteur d'un facteur  $100^{230,231}$ . D'après Luciano Marraffini, impliqué dans la première publication, cet effet modeste pourrait suffire : « *Il nous suffit d'aider le système immunitaire lorsqu'il est surchargé, des études ont montré que dans certains cas, réduire le nombre de bactéries pathogènes d'un seul log peut être suffisant, le système immunitaire prend le relais » <sup>15</sup>. Cette prédiction devra être confirmée pour les antibiotiques CRISPR.* 

Un inconvénient supplémentaire est le très restreint tropisme du vecteur phagemid. Le nombre de bactéries éligibles à ce procédé reste relativement faible. Les espoirs se tournent vers des phages synthétiques tels que le phage T7 qui serait un vecteur de choix, avec la possibilité de cibler spécifiquement un grand nombre d'hôtes.<sup>240</sup>

Finalement, la grande incertitude restante reste encore l'apparition de résistance, ce coup-ci dirigée contre ces nouveaux mécanismes antibiotiques. Les CRISPR antibiotiques déplacent la pression évolutive à laquelle l'homme soumet les bactéries au niveau génétique. La plupart des experts s'accordent sur le fait que tôt ou tard, les

bactéries utiliseront leur adaptabilité exceptionnelle pour contourner ce nouvel élément de l'arsenal thérapeutique. <sup>239</sup>

#### 4.2 Perspective: la résistance aux antibiotiques

#### 4.2.1 Impact

Deux récents rapports de la Rand Corporation et de la Chatham House sont dédiés à la résistance microbienne (des fondations non gouvernementales de conseil en affaires internationales la Chatham House est considérée comme le second groupe de réflexion les plus influents au Monde et le premier en Europe<sup>241</sup>). Ces rapports aux préoccupations délibérément économiques dressent les prédictions, coûts humains et financiers en cas d'absence d'action. 300 millions de décès prématurés et 100 Trillions de dollars de perte pour l'économie mondiale seraient à prévoir d'ici 2050<sup>242</sup>. L'utilité de tels rapport est primordiale :

« C'est lorsqu'une menace est comprise que les gouvernements y répondent avec des déploiements d'énergie et d'argent »<sup>243</sup>.

La dynamique de la réponse aux épidémies HIV et Ébola laisse espérer que nous parviendrons à répondre à la menace posée par la résistance aux antibiotiques. La forme que prendra cette réponse devra être multiple, les nouveaux antibiotiques en font partie.

#### 4.2.3 Modèle économique et Innovation

L'innovation en terme d'antibiotiques est pour l'industrie pharmaceutique un marché peu lucratif, comparés par exemple au potentiel revenus que générerait un nouveau médicament cardiovasculaire. Et ceci pour d'excellentes raisons de santé publique: si un nouvel antibiotique était développé et mis sur le marché, il ne serait réservé pendant les premières années qu'aux dernières lignes de la stratégie thérapeutique et donc sa consommation et son volume de vente seraient faibles. Certains rapports actuels prônant le développement d'antibiotiques innovants préconisent donc de changer le modèle économique en rompant le lien entre retour sur investissement et volume des ventes, sous forme d'incitations financières provenant des états ou des acteurs de santé<sup>244</sup>. Une autre approche consiste à utiliser des médicaments anciens, en repensant leur usage à l'aide des connaissances actuelles. C'est le cas du projet AIDA financé par l'union européenne<sup>245</sup>qui réalise des essais cliniques sur 5 antibiotiques développés avant les années 1980s.

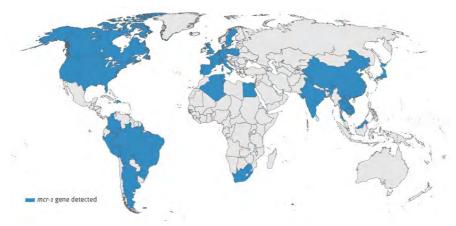
#### 4.2.4 Solution à long terme : usage raisonné

La course aux nouveaux antibiotiques ne doit être que l'un des nombreux axes à exploiter pour lutter contre les résistances bactériennes. Les dernières années ont montré que se concentrer sur un usage raisonné des antibiotiques, les mesures de santé publique, la prévention primaire et le contrôle des infections contribueront grandement à trouver l'harmonie entre lutte contre les infections bactériennes et l'apparition de résistances <sup>246</sup>. Y parvenir en rétablissant l'équilibre entre excès et accès en ce qui concerne l'usage des antibiotiques sera l'un des défis de santé de ce siècle, à l'échelle mondiale. Quant à la technique CRISPR-Cas appliquée chez la bactérie, elle permet une précision unique, ouvrant la possibilité d'étudier de nouveaux domaines

encore incompris tels que l'exploration du microbiome. Son usage naissant en tant qu'outil diagnostic, pourrait lui aussi être appliqué aux infections bactériennes, et être capable d'améliorer nos techniques actuelles<sup>247</sup>

## 4.2.3 Un problème mondialisé

Comme nous l'avons vu dans l'introduction de ce chapitre, la Colistine était jusqu'en 2015 le seul antibiotique pour lequel nous ne connaissions pas de résistance étendue, il représentait un antibiotique de dernier recours. L'identification d'une résistance à cet antibiotique portée par un plasmide représente un changement majeur. Antérieurement ces résistances étaient réduites à des mutations chromosomiques, qui se transmettaient verticalement. Depuis la détection de plasmides porteur du gène de résistance à la colistine (gène mcr-1 pour *mechanism of colistin resistance 1*) en Chine en 2015 <sup>228</sup> il a été identifié dans 32 pays, étalés sur les 5 continents <sup>248</sup>.



**fig.44** détection de mcr-1 dans des bactéries dans des échantillons d'origine animale, humaine ou environnementale dans 32 pays en Juillet 2016  $^{248}$ 

En Janvier 2017 le CDC (center for disease control and prevention) a rapporté le premier décès connu et médiatisé lié à une CRE (Carbapenem resistant Enterobacteriaceae) résistante « à 26 antibiotiques dont tous les Aminosides et polymixines testées et moyennement résistante à la tigecycline (un dérivé tétracycline récemment développé en réponse à l'émergence des resistances aux antibiotiques) »<sup>249</sup>. La détection de mcr-1 fut par contre négative. Face à l'étendue de la résistance bactérienne, un antibiotique de dernière ligne et des méthodes de re-sensibilisations semblent de plus en plus indispensable, ce rôle sera peut être occupé par l'édition génomique.

## IV Contrôle biologique : exemple du paludisme

## 1. Gene Drive

#### 1.1 Origine

Chaque gène autosomique a dans la majorité des cas 50% de chance d'être transmis à la descendance. Soit la descendance hérite d'une copie paternelle, soit d'une copie maternelle. Mais en 1957 des biologistes trouvèrent une des exceptions à cette règle, des gènes qui ne jouent pas selon les règles de l'évolution. Ces gènes manipulaient littéralement la division cellulaire et forçaient leur chance de transfert de sorte qu'ils se retrouvaient dans la descendance dans des proportions que le hasard seul ne pouvait expliquer. Ces gènes peuvent se répandre dans une population même si ils réduisent l'avantage évolutif de l'organisme qui les portent.Un tel comportement leur a valu le nom d' « éléments génétiques égoïstes », un exemple de ces éléments est le HEG : le Homing Endonuclease Gene.

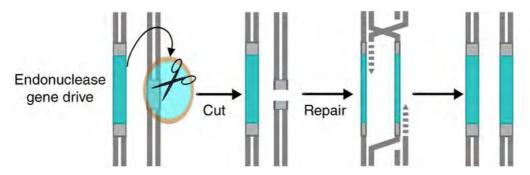


fig.45 Mécanisme simplifié d'un Homing Endonucléase Gene 250

Les HEG biaisent l'hérédité en codant pour une enzyme qui coupe le chromosome homologue, induisant la cellule à les utiliser comme modèle et à les copier lors de la réparation de la coupure. Il devient donc présent sur les deux copies chromosomiques. Les HEG codent pour des protéines qui sont capables de reconnaître et de cliver des séquences ADN de 15 à 30 pdb qu'on nomme donc séquence de reconnaissance. Les HEG sont eux même localisés au milieu de la séquence de reconnaissance, empêchant ainsi l'endonucléase de les cliver. Cela limite donc l'activité de l'enzyme aux séquences ne comportant pas encore le HEG. Lorsque la protéine rencontre un chromosome contenant la séquence de reconnaissance, elle la coupe en créant une cassure double brin. Celle-ci est souvent corrigée par recombinaison homologue en utilisant le chromosome non affecté comme modèle et c'est ce phénomène entraîne la conversion d'un hétérozygote en homozygote comportant deux brins contenant le HEG, un processus connu sous le nom de « homing ». Par le biais de ce mécanisme, la fréquence d'un HEG peut rapidement s'accroître dans la population.

### 1.2 Le gene drive

En 2002 un généticien, Austin Burt, proposa une technique ingénieuse pour répandre un gêne d'intérêt dans la descendance, il suggéra d'utiliser ces HEG pour propager un gène d'intérêt dans la population<sup>251</sup> permettait de créer une technique de *Gene Drive*, derrière le nom *gene drive* que l'on pourrait traduire par « *génétique dirigée* », se cache une technique de manipulation génétique qui permet d'augmenter artificiellement un génotype dans une population. En relâchant quelques individus qui possèdent une portion d'ADN élaborée par l'homme (appelée cassette *gene drive*) dans une population naturelle, on peut théoriquement obtenir en quelques dizaines de générations une population entièrement contaminée par la cassette *gene drive*. Mais l'utilisation des HEG dans la génétique dirigée se heurta rapidement à la difficulté de modifier les *Homing Endonucleases* pour couper de nouvelles cibles <sup>252,253</sup>. Les tentatives pour réaliser un *gene drive* avec les *Zinc Finger Nucleases* ou les TALENS, plus facilement dirigeables souffrirent d'instabilité. La raison évoquée de cette instabilité résiderait dans la nature répétitive des gènes qui codent ces systèmes, ils auraient une tendance à la recombinaison les empêchant d'être transmis avec fidélité. <sup>254,255</sup>

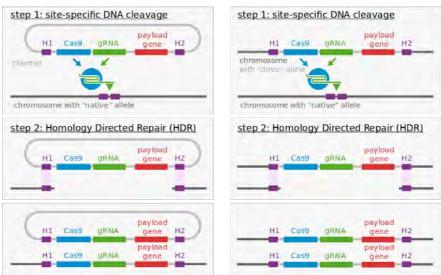
## 1.3 Renouveau de la génétique dirigée

Flaminia Catteruccia, directeur de l'Insecterie à la *Harvard Chan School of Public Health* est un des experts mondiaux en zoonoses transmises par les moustiques telles que la dengue, le chikungunya et le paludisme. C'est elle qui à été contactée pour réaliser l'analyse génétique du virus Zika apparue au Brésil au printemps 2015<sup>256</sup>. L'un des principaux atouts de son laboratoire par rapport aux autres entomologistes est son penchant pour la génétique <sup>257</sup>. Lorsque l'une de ses assistantes, Andrea Smidler démarra son Ph.D. Elle fut encouragée à effectuer une rotation dans d'autres laboratoires, l'un d'eux était le laboratoire de génétique de George Church, extrêmement actif dans le

domaine CRISPR-Cas9. Elle y rencontra un jeune postdoctorant, Kevin Esvelt<sup>257</sup>. Leur publication en 2014 fut la première à allier CRISPR-Cas9 et *Gene Drive*<sup>258</sup>. L'arrivée de la technique CRISPR-Cas9 bouleversa l'avenir de la génétique dirigée, la faisant passer de la théorie au domaine du techniquement possible en apportant la stabilité et la spécificité recherchée. En s'inspirant des HEG mais en utilisant la Cas9 et son guide à la place, les projets passés devenaient possible. L'application de la génétique dirigée à l'un de ses objectif les plus ambitieux, l'éradication du paludisme, allait connaître un renouveau.



#### propagation after sexual reproduction



**fig.46** mécanisme d'introduction d'un *gene drive* via CRISPR-Cas9 : le MCR ou *Multiple Chain Reaction* **étape 1** : Introduction exogène du plasmide contenant la cassette *gene drive* dans la cellule via le vecteur choisi. Le plasmide contient les gènes codant la Cas9, l'ARN guide et facultativement un gène cargo : gène d'intérêt que l'on souhaite voir représenté au sein d'une population rendue homozygote. En violet les séquences H1 et H2 flanquant la construction sont des séquences homologues des extrémités du gène cible, l'allèle natif. **Étape 2**. Codée par le plasmide, la protéine cas9 va couper le site déterminé par l'ARN guide. Par HDR à l'aide des séquences H1 et H2, la cassette contenant les gènes cas9, sgARN et Cargo est copiée dans le chromosome. On obtient alors un allèle modifié dit « *allele drive* ».

**étape 2: Propagation après reproduction sexuée**: L'expression du chromosome modifié permet la production de la Cas9 et de son guide qui vont cibler l'allèle natif présent sur le chromosome homologue. La cassure sera réparé en utilisant le chromosome compagnon préalablement modifié comme matrice. La mutation introduite sera alors homozygote<sup>259,260</sup>.

## 2. Gene drive et CRISPR: exemple de l'éradication du paludisme

#### 2.1 Le paludisme

Selon les estimations de l'OMS publiées à la fin de l'année 2015, 214 millions de cas de paludisme et 438 000 décès ont été dénombrés cette année là <sup>261</sup>. Le paludisme est une parasitose traitable, mais le manque d'accès au soin et à des structures fonctionnelles de santé publique font que l'épidémie perdure. En 2015, 80% des cas de paludisme et 78% des décès dus à cette maladie surviennent dans 15 pays, principalement en Afrique subsaharienne. L'impact de la maladie va au delà de l'épidémiologie, il est aussi économique et social. Le coût pour éradiquer définitivement la maladie d'ici 2030 grâce aux techniques « *traditionnelles* » est estimé à 101,8 Milliards de dollars par un récent rapport d'expertise<sup>262</sup>. Ce rapport commandité par l'initiative *Roll Back Malaria* qui dépends de l'OMS et des Nations Unies prévoit un bénéfice immense à l'éradication de la pathologie :

non seulement « économique en apportant le retour de la productivité, de la croissance et la réduction de la pauvreté des foyer. Mais cela favoriserait à terme l'équité, l'émancipation des femmes et renforcerait le système de soin » <sup>262</sup>.

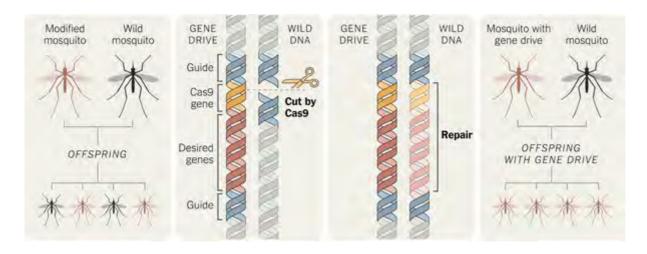
Le paludisme peut être traité chez l'homme en utilisant des combinaisons thérapeutiques à base d'artémisinine ( ou CTA); mais la thérapie nécessite une détection

du parasite et l'accès du patient au système de santé. Le parasite a désormais développé des résistances aux traitements de première ligne et échappe peu à peu au contrôle par insecticide. Ces préventions par épandage d'insecticide dans les zones de ponte jusqu'au domicile nécessitent par ailleurs la réalisation de campagnes organisées et de ressources. Un vaccin contre la malaria est en cours de développement mais prendra encore plusieurs années avant d'obtenir l'efficacité nécessaire et d'être prêt à être approuvé puis déployé à large échelle. <sup>262</sup>

## 2.2 Gene drive: une alternative au traitement du paludisme

Depuis Austin Burt, les projets visant à atteindre les vecteurs du paludisme via un *gene drive* furent parmi les premières tentatives d'application de la technique. De nombreux efforts furent déployés apportant des solutions innovantes mais techniquement difficiles à réaliser avec l'outil dont ils disposaient.<sup>263</sup>. Le but commun étant d'artificiellement modifier les populations de moustiques pour qu'elles transmettent elles-même un élément qui conduira à leur déclin ou préviendra la transmission du parasite ou son développement chez l'hôte.

Austin Burt avait imaginé une stratégie dévastatrice utilisant les gene drives pour perturber des gènes dans le but de causer l'infertilité ou la mort seulement si les deux copies des gènes sont supprimées. Ces « genetic load drives » se répandraient rapidement au sein des hétérozygotes qui ne seraient quasiment pas perturbés. A terme, elle entraînerait la chute de la population ou l'extinction, due à l'accumulation de mutations récessives<sup>219</sup>. Une seconde approche vise à perturber le sex ratio <sup>264–267</sup>. Pour cet exemple, le chromosome Y (ou son équivalent suivant le déterminisme sexuel de l'espèce) code une endonucléase qui coupe et détruit le chromosome X durant la méiose, s'assurant ainsi que la plupart des spermatozoïdes viables contiennent un chromosome Y 267-269. Ces deux techniques portent le nom de « suppression drive », car leur objectif est de supprimer l'intégralité de la population visée. Une alternative est le « resistance drive » visant à booster l'immunité d'Anopheles stephensi contre les sporozoïtes de Plasmodium falciparum, en leur apportant le matériel génétique (naturellement présent chez certains moustiques) codant pour des anticorps dirigés contre le parasite (gènes m1C3, m4B7, ou m2A10 )<sup>270</sup>. La technique gene drive ARN guidée, rendait applicables ces stratégies délaissées, les anciennes approches furent transposées<sup>254</sup>, plusieurs espèces porteuses ciblées<sup>258,259</sup>(Anopheles stephensi<sup>259</sup>, Anopheles gambiae<sup>222</sup>).



#### fig.47: Génétique dirigée par CRISPR-Cas9271

De gauche à droite : en pratique lors de la reproduction sexuée, une mutation autosomique à 50% de probabilité de se retrouver dans la descendance. La technique du gène drive qui ici utilise CRISPR-Cas9 permet artificiellement d'obtenir l'homozygotie. La zone guide sert de séquence de reconnaissance pour la coupure de l'ADN sauvage. La cassette génétique contient le gène d'intérêt, l'endonucléase utilisée est la Cas9 codée par le gène cas9 compris dans la cassette. L'expression du système permettra la transcription de la Cas9 qui guidée, coupera l'ADN sauvage au sein de sa séquence de reconnaissance. Le mécanisme de réparation cellulaire utilisera l'allèle intact comme matrice pour reconstruire le brin coupé. Lors de la reproduction on obtiendra donc artificiellement une population chez laquelle le gène d'intérêt est présent à près de 100%.

#### 2.3 Conclusions

Relâcher des anophèles stérilisés pour diminuer la transmission du paludisme n'est pas une approche nouvelle, elle à déjà été expérimentée en laboratoire et en pleine nature<sup>272</sup>, mais contrairement au *gene drive*, son impact n'est que local. Relâcher dans la nature quelques milliers de moustiques *gene drive* suffirait à éradiquer l'intégralité de l'espèce de manière peu coûteuse et quasi autonome. De nombreux investisseurs et organisations misent sur cette approche, tel que la fondation Bill et Melinda Gates via l'initiative TargetMalaria qui a déjà débloqué des investissements considérables<sup>273</sup>. Mais ce n'est pas parce que cette solution est techniquement réalisable qu'elle est pour autant celle qui doit être appliquée. La capacité d'éradiquer une espèce entière ou de transformer artificiellement l'intégralité d'une population sauvage soulève de nombreuses questions éthiques, écologiques et de gouvernance, elle représente un défi pour les paradigmes réglementaires actuels.

## 3. Autres applications de la génétique dirigée

#### 3.1 Santé Publique

Dans le domaine de la santé publique, plusieurs applications sont envisageables :

- Contrôle d'organismes transportant des agents responsables de maladies infectieuses chez l'homme tels que : la maladie de Chagas, Lyme, Dengue
- Contrôle ou modification d'organismes pouvant causer une infection : schistosome
- Contrôle ou altération d'organismes servant de réservoir à un pathogène (chauve souris, rongeurs)
- Conservation des écosystèmes, la génétique dirigée peut être utilisée pour contrôler ou modifier les organismes vecteur de pathogènes qui menacent la survie d'une autre espèce, éliminer les espèces invasives ou modifier les espèces en danger.

#### 3.2 Agriculture

En Agronomie cette technique est utilisable pour contrôler ou modifier les organismes vecteurs de pathogènes des plantes cultivées ou endommageant les plantations. L'élimination des « *mauvaises herbes* » qui sont en compétition avec les cultures ( via une sensibilisation ou une restauration de la sensibilisation à un herbicide par exemple) serait également possible.

#### 3.3 Recherche fondamentale

La modification d'un organisme peut permettre de réaliser des études portant sur le fonctionnement et les effets de la génétique dirigées elle-même mais peut aussi fournir des renseignement sur la biologie des espèces ou le mécanisme de maladies.

## 4. Risques

## 4.1 Écosystèmes et dissémination

Bien que de nombreuses bourses aient été allouées à l'étude de la sécurité liée à l'utilisation de *gene drives* chez le moustique, peu se sont penchées sur l'effet qu'une telle technique aurait sur d'autres espèces <sup>274–277</sup>.Un *suppression drive* comme celui d'Austin Burt comporte un risque en cas d'évasion ou de relâchement involontaire. L'éradication d'une espèce entière comporte des conséquences difficilement calculables sur son écosystème et la niche écologique laissée vide pourraient être remplie par une autre espèce.

Les études portant sur la biologie moléculaire des *gene drive* ont devancé la recherche sur la génétique populationnelle et la dynamique des écosystèmes. Ces deux domaines sont pourtant essentiels pour évaluer les bénéfices des *gene drives* et leurs conséquences biologiques et écologiques. Des lacunes persistent encore concernant les potentiels effets hors-cibles (au sein même de l'organisme modifié) et sur les autres espèces. De plus en terme d'écologie un *gene drive* pourrait causer des dommages considérables sur l'environnement si il venait par inadvertance à atteindre ce que l'on nomme une espèce *clé-de-voute*. C'est à dire une espèce qui à un apport disproportionné sur son environnement en comparaison à son abondance.

En terme de dissémination, des *gene drive* ne respecteraient pas les frontières internationales et contourneraient librement les traités internationaux, comme le Protocole sur la Biosécurité de Cartagene (traité gouvernant les mouvements transfrontaliers d'organismes génétiquement modifiés). Les pays relâchant des *gene drive* pourraient être accusés d'infraction à la Convention des Nations Unies sur les Armes Biologiques si des organismes porteurs de *gene drive* étaient ammenés à nuire à des espèces natives d'autres pays.

#### 4.2 Risque d'usage double

Les recherches qui peuvent être délibérément détournées dans le but d'un usage malveillant sont qualifiées de « *double usage ou dual use en anglais* ». <sup>278</sup>

Le cas des *gene drive* est différent des autres recherches en biologie synthétique. Alors que celles-ci permettraient de modifier des pathogènes pour les rendre plus dangereux, le gene drive en est incapable. Bactéries et virus ne peuvent être modifiés par cette technique car ils ne sont pas dotés de reproduction sexuée. La propagation chez les plantes agricoles et le bétail serait stoppée car leurs reproduction est fréquemment contrôlée. Enfin les inquiétudes liées à un usage chez l'homme sont infondées; nous sommes une espèce dont le temps de gestation est trop long pour que la génétique dirigée se répande dans une population humaine.

Le Bioterrorisme est une peur qui va de pair avec les gene drive. Pour l'instant il n'y à

sûrement pas lieu de s'inquiéter, il semble improbable qu'une organisation terroriste investissent des millions dans un programme de génie génétique de pointe.

« *J'ai passé de longs moments à réfléchir aux applications mal intentionnées que permettrait la génétique dirigée, et je n'en ai pas trouvé une seule qui aurait un franc succès »* à déclaré Ethan Bier, co-auteur du premier article réalisant une application animale du CRISPR *gene drive* <sup>259,279,280</sup>. Il ajoute :

« il y à tellement d'autre méfaits qui sont plus facile à réaliser ».

#### 4.3 Risques de santé publique

La technique peut également représenter une menace potentielle de santé publique. L'exemple théorique d'un moustique modifié pour ne plus servir d'hôte du virus de la dengue permet illustrer ces risques. Ce moustique pourrait par exemple devenir l'hôte d'autres arboviroses auxquelles il serait désormais plus susceptible. Le virus de la dengue, sous la pression évolutive du *gene drive* pourrait aussi développer un nouveau phénotype, résistant au *gene drive* et ayant un profil pathogène qui pourrait être différent. Le retrait d'une espèce entière pourrait libérer une niche écologique, celle-ci sera peut être remplie par un autre insecte vecteur de de zoonose.<sup>281-283</sup>

En ce qui concerne les applications de la génétique dirigée à l'agriculture, celà pourrait entraîner des effets indésirables sur le bien être humain. Le transfert d'un *suppression drive* à une espèce sauvage non ciblée pourrait nuire à la fois à l'environnement et au rendement agricole par exemple.

## 5. Gene Drive : Des problématiques nouvelles

#### 5.1 Introduction

L'homme au cours de son histoire a participé à l'éradication d'espèces pour se nourrir, s'enrichir ou encore par indifférence, mais la technique CRISPR-Cas9 apporte une catégorie nouvelle, le choix délibéré d'éteindre une espèce pour une raison humanitaire. Cette technique crée des problématiques que la société n'a que rarement eu à envisager. Quelle est la légitimité d'un individu, d'un pays ou d'une organisation lorsqu'il s'agit de modifier la nature d'une manière qui affecterait la planète entière? A quel point une incertitude est tolérable une fois encadrée par les règles qu'a érigée la société autour d'elle? Et à quel point le bénéfice apporté est considéré comme dépassant cette incertitude? Cela fait partie des questions auxquelles la société devra répondre.

« Je déteste le sujet du paludisme » admet Kevin Esvelt, «il rend l'utilisation de cette technologie terriblement attrayante » 284

## 5.2 Notion d'équité

Une prudence accrue et une évaluation globale seront nécessaires si les *gene drive* venaient à être utilisés. L'un des exemples est la répartition des conséquences entre pays à forts revenus et pays pauvres. La modification d'*Amaranthius palmeri*, une plante d'Amérique du nord l'illustre bien. Éteindre

l'espèce serait favorable aux États-Unis, où elle est considérée comme une plante invasive, nuisant au rendement des récoltes. Mais si il est un nuisible aux États-Unis, ce n'est pas le cas des espèces voisines du Mexique, d'Amérique du Sud, de Chine et d'Inde qui y sont cultivées à des fins alimentaires. La comparaison des bénéfices/risques ne doit donc pas impliquer seulement leur probabilité ou leur magnitude, mais doit prendre en compte l'aspect social et humain des changements entraînées, à l'échelle mondiale.

## 6. Garde fous : des règles et des techniques

## 6.1 le Rapport NIH

Le fait que ce domaine avance à une vitesse impressionnante est à la fois encourageant et préoccupant. Dans le cas de la génétique dirigée, la faisabilité technique a devancé les réflexions nécessaires à son encadrement.

Des instances gouvernementales tels que le *National Institute of Health* (NIH) ont requis la création d'un comité par la *National Academy of Sciences, Engeneering and Medicin*e. Ce comité est le premier dont le but est d'évaluer l'état de la science en matière de *gene drive* et d'établir des directives pour garantir un usage responsable de cette technologie émergente.

Le comité à considéré six cas d'études et conclut qu'en 2016, il n'y avais pour l'instant pas assez de preuves disponibles pour justifier le relâchement d'organismes modifiés par gene drive dans l'environnement. Mais les bénéfices et les espoirs portés par cette technique, tant dans le domaine de la recherche fondamentale qu'appliquée, sont significatifs. À eux seuls, ils justifient la poursuite de recherches en laboratoire et la réalisation d'essais de terrain, en situation hautement contrôlée. L'encadrement des recherches passe par la réalisation des études « pas à pas » et en adoptant des méthodes avancées de confinement 285. Au delà d'encadrer la pratique en laboratoire, il y a un consensus unanime sur l'importance d'engager le public dans la prise de décision liées aux gene drives. L'implication des populations concernées est essentielle pour organiser et définir les risques et participer aux décisions pratiques et politiques. Le rapport souligne donc qu'il est impératif que les futures recherches soient réalisées dans des conditions de totale transparence de manière à favoriser les discussions avec le public. Dans l'éventualité d'une décision visant à utiliser un gene drive, cette décision devrait être prise au cas par cas en évaluant spécifiquement le rapport bénéfice/risque de chaque drive.

## 6.2 Méthodes biotechnologiques

Dans le cas d'une dissémination d'un *gene drive* qui aurait des effets indésirables inattendus ou qui a été répandu sans le consentement préalable des autorités compétentes il existe une technique biotechnologique nommée *reverse gene drive*. Le fonctionnement du *reverse gene drive* est simple, il utilise la technique du *gene drive* pour venir effacer et recouvrir la modification précédente. La nouvelle séquence doit être recodée par rapport à l'originale pour empêcher le gene drive initial de la découper. Si nécessaire, un troisieme drive permet de restaurer l'exacte séquence sauvage, laissant

seulement la séquence codant l'ARN et la protéine Cas9 comme signature des éditions passées. Mais les chercheurs et régulateurs ne peuvent pas se reposer seulement sur ces techniques, encore imprécises en tant que seule stratégie d'atténuation des risques représentés par un autre *gene drive* .

#### 7. Conclusions

Parmi les applications de la technique CRISPR-Cas9, c'est à celle-ci que le néologisme « *disruptif* » s'applique le plus. Elle bouleverse la biotechnologie, la science, la médecine, mais aussi l'écologie, l'éthique, et le droit. La place perçue de l'homme dans l'écosystème et sa relation avec la nature aura un rôle important dans le débat à venir autour des *gene drives*. La décision d'utiliser un *gene drive* devra passer par une séries de questions : Si oui ou non, pourquoi et comment les recherches devraient être conduites? Quand, où et si oui ou non un organisme modifié par gene drive devrait être relâché ?

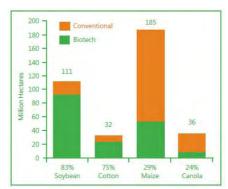
Il faut espérer que nous saurons structurer la réflexion autour des *gene drive* afin de rassembler et fédérer une communauté d'acteurs capables de parvenir aux décisions les plus multidisciplinaires possibles. L'implication des personnes qui seraient impactées par la mise en place d'un gene drive semble primodiale. Comme le souligne le rapport NIH, beaucoup reste à accomplir. Il existe de nombreuses lacunes dans notre compréhension de l'effet qu'aurait *gene drive*. Combler ces lacunes sera nécessaire à l'évaluation des risques sanitaires et environnementaux, à l'engagement du public, et ultimement à la décision de relâcher ou non un *gene drive*.

## V Plantes et agronomie

#### 1.introduction

Au delà des applications thérapeutiques de la technologie CRISPR, c'est dans le domaine de l'agronomie qu'elle risque de connaître un développement colossal dans un futur proche. L'agronomie a toujours été intimement liée à l'histoire de la génétique, les techniques biotechnologiques ont donc été adoptées rapidement, créant une nouvelle branche de cette industrie : la biotechnologie agricole. Les éleveurs de plantes et d'animaux ont sélectionné pendant des siècles des espèces pour leurs attributs et les ont combinés par des méthodes qu'on pourrait qualifier de traditionnelles : par croisements sélectifs. Mais ce processus pouvait prendre des décennies pour atteindre un but particulier. Jusqu'à récemment, les techniques de génie génétique utilisées pour manipuler l'ADN avaient été si imprécises que cela les rendait trop coûteuse et difficiles à utiliser. L'apparition des techniques d'édition génomique a été une avancée excitante pour l'agronomie, réduisant considérablement le temps et les coûts de R&D (recherche et développement). Elles se substituent peu à peu à la biolistique, la microinjection ou encore le transfert par agrobactérium dans le pipeline des producteurs de semences. CRISPR fais office de premier choix, et hormis sa précision accrue, la technique pourrait se révéler un outil unique pour les semenciers en n'étant pas soumise à la régulation actuelle sur les OGMs. 286

Vingt ans après leur autorisation sur le marché, les plantations génétiquement modifiées recouvraient en 2015 près de 180 millions d'hectares répartis dans 28 pays, récoltés par environ 18 millions de fermiers. Près de 115 millions d'hectares sont présents sur le seul continent américain. <sup>287</sup>



**fig.48** Pourcentage des principales semences génétiquement modifiées (en vert), par rapport aux cultures conventionnelles (en orange).<sup>287</sup>

Le fait que seulement 4 types de semences (Soja, Maïs, Coton et Colza) constituent la quasi-totalité des plantations génétiquement modifiées peut interpeller. Cette concentration s'explique par le fait que le développement d'une nouvelle plante génétiquement modifiée est aujourd'hui long et coûteux et nécessite l'accumulation de nombreuses connaissances. En moyenne 136 millions de dollars et 13 années seraient nécessaire pour créer une nouvelle semence et obtenir la conformité avec la régulation en vigueur.<sup>288</sup>

De multiples travaux sont également en cours concernant l'utilisation de l'édition génomique pour modifier le bétail. Ils ne seront qu'évoqués dans ce chapitre.

## 2.Exemples

## 2.1 plantes

Diverses modification via CRISPR ont été réalisées dans le but de produire des semences commerciales dotées de divers avantages. Elles sont toutes destinées à l'agriculture intensive. Leur objectif étant souvent d'accroître le rendement des cultures en améliorant par exemple la tolérance à la sécheresse ou en augmentant la croissance dans des conditions de privation en nutriment des semis<sup>285,286</sup>. L'utilisation récente de CRISPR chez le maïs <sup>287</sup> ou le soja <sup>290</sup> illustre la rapidité avec laquelle la technologie sort des laboratoires. Crispr peut également être utilisé pour combattre les pathogènes végétaux, comme montré dans le cas du virus *tomato yellow leaf curl* chez Nicotiana benthamiana <sup>291</sup>. Théoriquement, c'est l'ensemble des précédentes modifications envisagées qui pourraient être « *traduites* » en version CRISPR.

#### 2.2 Animaux

Dans l'ère moderne, les éleveurs d'animaux ont identifiés des marqueurs chromosomiques associés à certaines caractéristiques, ils sont connus sous le nom de *quantitative trait loci*. Les éleveurs dépistent désormais ces marqueurs pour sélectionner ces traits précieux. Ces processus seront accélérés en utilisant des

techniques d'édition génomique. Récemment la technique a été utilisée chez le porc la vache à lait<sup>293</sup> obtenir des animaux dépourvus de cornes. L'une des autres application chez l'animal est la production de médicaments ou de tissus. Par exemple le *knock in* d'ADNc au niveau du locus Alb porcin par CRISPR-Cas9 pourrait rendre possible la production d'albumine par cochons transgéniques.<sup>294</sup>

#### 3. Plantations et acteurs industriels

Différents développeurs de semences ont rapidement acquis des licences d'utilisation de CRISPR. Dupont Pioneer a investi dès octobre 2015 dans Caribou Bioscience, la *start-up* CRISPR cofondée par Jennifer Doudna. Ils expérimentent actuellement son utilisation sur le maïs, le blé, le soja et le riz et déclarent « *espérer vendre les graines modifiées par CRISPR aussitôt que possible, au maximum dans 5 ans* » <sup>295</sup>. D'autres acteurs tels que Calyxt, la division « *plante* » de Cellectis sont présents mais celui qui est le plus imposant est le colosse qui naîtra de l'acquisition-fusion entre l'américain Monsanto et l'allemand Bayer. L'importance de CRISPR dans cette fusion est loin d'être anodin.

Monsanto a acquis, dès septembre 2016, une licence d'exploitation de CRISPR provenant du *Broad Institute of Harvard and MIT* de Feng Zhang et George Church<sup>296</sup>. Cette licence non exclusive n'a que peu de limitations (l'interdiction de développer des *gene drive*, d'appliquer la technologie au tabac ou de créer des semences stériles surnommées « *terminator* » qui forcerait les fermiers à en acheter annuellement). Bayer s'est lancé dans un partenariat (ou *joint venture* en anglais) avec *Crispr Therapeutics*, l'autre *start-up*, celle d' Emmanuelle Charpentier, mais l'accord de licence qu'ils possèdent ne concernait que l'usage thérapeutique <sup>297</sup>.

L'acquisition-fusion, d'un montant de 58,8 milliard d'euros, créera une entité géante qui contrôlera plus de 25% des semences et pesticides mondiaux (Ce chiffre de 25 % représente la valeur marchande, mais si nous considérons la surface cultivée, il représente une proportion beaucoup plus élevée)<sup>298</sup>. Cet englobement de Monsanto catapultera Bayer dans le rôle de leader dans les domaines de la biotechnologie et de la production de nourriture à l'échelle mondiale, il signifiera également un rapprochement étroit entre l'industrie agricole et les pionniers de la technique CRISPR. L'étendue de ce partenariat sera globale, le marché dominant de Bayer étant l'Europe et celui de Monsanto le continent Américain.

Cette fusion fait suite à la vague d'acquisition que connaît ce secteur en réaction à son ralentissement depuis 2013 suite à la diminution des dépenses des exploitants agricoles qui entraînent peu à peu des prix de base inférieurs<sup>299</sup>. Les américains Dow Chemical et Dupont, Chemchina et le suisse Syngenta sont aussi en cours de fusion durant l'année 2016. Lorsqu'elles seront approuvées, 60% des semences et des agents chimiques utilisés en agriculture appartiendrons à seulement 3 compagnies. Cette situation où il y a peu de vendeurs pour un nombre important d'acheteurs est nommé un oligopole. L'implication de cette concentration du marché est l'un des éléments majeurs qui est évalué par les régulateurs gouvernementaux avant l'approbation finale de ces fusions. Malgré des débats engagés, début 2017, tout semble indiquer la réalisation future de ces transactions<sup>290</sup>.

CRISPR et les futures techniques d'édition génomique auront pour ces semenciers un rôle majeur. Mis à part la réduction des coûts de recherche et de production, ces techniques pourraient réduire drastiquement ceux du développement. En effet contrairement aux OGMs ils ne sont pas pris en compte par les régulations existantes <sup>300</sup> ce qui conduirait à des économies de temps et d'argent pour emmener un produit sur le marché estimées à 5,5 années et 35 millions de dollars, par produits, en moyenne <sup>301</sup>.

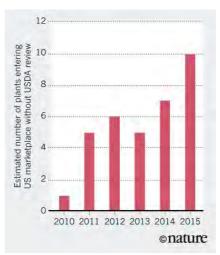
## 4. Régulation de CRISPR en biotechnologie agricole

#### 4.1 Procédé vs. Produit

Le débat portant sur la régulation des organismes modifiés a toujours été un débat houleux et politisé. Longtemps se sont opposés deux factions, celles souhaitant la régulation du procédé et celle privilégiant la régulation du produit. Pour l'édition génomique, en fonction de la position adoptée par les organismes régulateurs, cela pourrait signifier un allègement inouï des régulations portant sur les plantes modifiées par cette technique.

## 4.2 Le cas du champignon de paris qui ne brunissait pas

Concernant l'usage de CRISPR dans le domaine agricole, l'USDA, acronyme de United State Department of Agriculture a rendu le 13 Avril 2016 une décision très médiatisée. Ayant été encouragé à soumettre son dossier par des membres de ce département<sup>286</sup>, Yinong Yang un biologiste végétal de l'université de Pennsylvanie déposa un dossier auprès de l'institution gouvernementale. Son cas : une édition par CRISPR visant à réduire le brunissement de l'Agaricus bisporus, le champignon de Paris en ciblant la polyphénol oxydase (PPO). La réponse du département de l'agriculture américain fut favorable « il ne considère pas l'édition par CRISPR-Cas9 décrite ici comme devant être régulée ». La base de cette décision provenant du fait que pour cette modification génétique, aucun matériel étranger à la plante n'a été inséré dans son génome, c'est un knock out par Cas9 et ses guides ARN qui a été réalisé, la modification ne tombe donc pas dans les critères impliquant la régulation. Cela signifie que ces champignons pourront être cultivés puis vendus aux Etats-Unis sans passer par les longs et coûteux processus d'approbation requis par l'agence. Il est donc le premier organisme « CRISPR » a recevoir le feu vert du gouvernement américain et crée un précédent historique. L'USDA avait auparavant estimé qu'elle pratiquait une surrégulation des OGMs302 et suite à cette décision, nombreux sont ceux qui y voient le signe d'un fléchissement de leurs exigences réglementaires <sup>297</sup>.



**fig.49** Évolution du nombre de végétaux non régulés par l'USDA. De par l'évolution des techniques d'édition génomique, depuis 2010 un nombre croissant de plantes sont entrées sur le marché américain sans avoir été soumises à l'examen de l'USDA. Ce champignon est en 2016 le premier organisme à avoir été édité par CRISPR dans ce cas.<sup>302</sup>

## 4.3 Vers une évolution des régulations

CRISPR et les autres techniques d'édition génomique mettent les définitions réglementaires d'OGMs et du mot transgénique à l'épreuve. Lors de la rédaction des ébauches de la régulations des OGMs dans les années 1980s à 1990s, l'apport de séquences génétiques étrangères étaient essentielles à la fabrication de ces organismes. Avec l'arrivée des techniques d'édition génomique qui supplantent peu à peu les anciennes techniques ces régulations ont besoin d'être actualisées. La position des Etats-Unis, de la Chine ou de l'Europe sont vues comme des indicateurs de tendances, alors que les Etats-Unis semblent avoir assoupli leur attitude vis à vis des OGMs, l'Europe a quant à elle durci la sienne dans le passé.<sup>302</sup>

La mise à jour des cadres réglementaires pré-établis concerne divers groupes de travail dans le monde. Les États-Unis après plusieurs délais ont annoncé la mise en place un groupe de travail à l'USDA dans le but de mettre à jour leur régulation (le *Coordinated Framework for Regulation of Biotechnology* ou CFRB)<sup>303</sup>. De nombreuses opinions scientifiques parues dans les revues Science<sup>®</sup> ou encore Nature<sup>®</sup> qualifient l'absence de mise à jour du cadre juridique et l'état de dérégulation engendrée d'*opportunité manquée* <sup>294,295,298-300</sup> et craignent l'absence totale de régulation de ces plantes « *éditées* ». La Chine, seconde économie mondiale, à l'image des États-Unis est en train d'adopter une position plus favorable aux OGMs et à pour la première fois cité le développement de semences OGMs dans son 13ème *Zhōngguó Wǔnián Jihuà*, (soit « plan quinquennal de développement économique ») couvrant la période de 2016 à 2020. Bien que des recherches des laboratoires d'états portent sur des plantes modifiées par CRISPR<sup>89</sup> le gouvernement chinois n'as pas annoncé s'il autoriserait de tels végétaux à être cultivés <sup>295</sup>.



**fig.50** Le premier repas CRISPR: Stefan Jansson de l'université Umeå en Suède a réalisé en septembre 2016 le premier plat de pâtes accompagné de ses choux CRISPR  $^{308}$ 

Prônant toujours le principe de précaution et soulignant notre manque de compréhension des mécanismes génétiques, la régulation européenne historiquement hostile aux OGMs. Depuis Avril 2015, par la directive (EU) 2015/412 309 l'union confère aux états membres la possibilité d'interdire la culture des OGMs sur leur territoire et les 28 états ont adopté des mesures d'étiquetage des produits contenant des OGMs. En ce qui concerne NBPTs; les New Plant Breeding Techniques, le nom donné aux nouvelles plantes dont le génome à été altéré par édition génomique, la position de l'union reste encore indéterminée. En Europe, l'enjeu est de taille, si l'Union classifie ces NBPTs comme OGMs ils seraient par conséquent interdits de culture dans la majorité des pays membres, si aucune régulation n'a lieu ils pourraient en théorie être autorisés sur le territoire de ces même nations. Comme le déclare Jansson, le cuistot des premières plantes NBPTs à être officiellement dégustées :

«étant donné que les plantes que nous avons consommées ne sont pas, par définition, des OGMs en Suède, nous n'avons même pas eu à demander un permis ou informer les autorités (suédoises)»<sup>308</sup>

Pour l'instant les plantes modifiées par CRISPR ne sont techniquement pas prises en compte par l'actuelle législation Européenne. Les États-Unis encouragent l'Europe à calquer leur future régulation sur la leur, notamment au travers les négociations autour du TTIP (le partenariat transatlantique de commerce et d'investissement) avançant le fait « que des approches réglementaires différentes au sujet de la classification des NBPTs conduiraient à des perturbations significatives des échanges commerciaux » (Correspondance officielle entre l'Office of Agricultural Affairs et la Direction Générale de Santé et des Consommateurs de la Commission Européenne présentée en Annexe 12). La position de l'Union sera connue dans les années à venir, si elle ne régulait pas les NBPTs cela serait un retournement sans précédent.

## 5. Une action théorique

Même si la précision de la technique était idéale, l'édition génomique ne garantirait pas complètement l'obtention du trait escompté chez le végétal. La tolérance à la sécheresse par exemple, est non seulement associée à plusieurs gènes, mais elle est soumise à de complexes intéractions avec l'environnement. À quel point la fonction du gène dépendra des précipitations, de la chaleur, de la nature des sols, de leur profondeur? De plus l'ensemble des autres gènes auront une influence méconnue sur l'expression du gène modifié.

### 6. Conclusions

Aujourd'hui la manière dont nous produisons la quasi totalité de notre nourriture au niveau mondial passe par l'intensification des pratiques et des techniques de cultures et d'élevages. Les compagnies impliquées dans la conception des semences OGMs revendiquent l'innocuité de leurs créations, leur intérêt pour l'environnement et estiment qu'elles sont le seul moyen dont nous disposons pour nourrir un nombre toujours croissant d'êtres humains. À l'opposé les inquiétudes liées à la santé humaine, la brevetabilité des semences, l'impact sur l'environnement et la diversité des cultures, la remise en question des études favorables aux OGMs et une souveraineté alimentaire à deux vitesses en fonction des ressources économiques des nations sont mises en avant. La réalité se situe certainement quelque part entre ces deux extrêmes et est infiniment plus nuancée. La répartition et la définition des bénéfices et des risques varient au cas par cas, et concernent un réseau dense et varié d'acteurs : les développeurs, les nations, les agriculteurs, les consommateurs et l'environnement en font partie.

Au delà des considérations liées à l'agriculture intensive et les OGMs se pose la question de la régulation des NPBTs : les nouvelles techniques utilisées pour modifier les plantes que CRISPR rend industriellement applicables. Qu'elle semble à chacun trop sévère ou pas assez, une régulation des nouvelles techniques de production des plantes apparaît comme indispensable. Suite à l'évolution des techniques de production l'ancienne dichotomie Procédé vs. Produit apparaît inadéquate, cela entraîne peu à peu les régulateurs vers une approche au cas par cas, plus nuancée qui serait adaptée à la diversification qu'a connu la modification génétique.<sup>304</sup>

## VI Outils de Recherche

#### 1. Introduction

Comme nous l'avons vu dans le chapitre III (Optimisations, Variantes et Astuces) de la première partie et au cours des chapitres précédents, la protéine Cas9 et ses variantes représentent une incroyable boîte à outils au service de la recherche. Elle peut faire avancer notre compréhension des mécanismes biologiques, l'interaction hôte-virus par exemple, ou encore le suivi *in vivo* des chromosomes par la technique *CRISPR rainbow*. Elle est également complémentaire de différents outils de recherche, en voici trois nouveaux exemples : le criblage, la création de modèles et les *Genome Wide Association Studies*.

## 2. Screening / criblage

#### 2.1 Présentation

La Cas9 permet de modifier plusieurs cibles en parallèles, ce qui rend possible les criblages fonctionnels, capables d'identifier les gènes ayant un rôle dans l'obtention de phénotypes étudiés. Pouvoir concevoir de manière semi-automatisée et synthétiser des gARNS fait de la technique un candidat idéal pour les criblages à grande échelle et augmente le débit de ces expérimentations. Ces Cas9/gARNs seront capables de muter, activer ou réprimer presque tous les locus génétiques choisis par le chercheur, de manière individuelle ou simultanée. Ces criblages sont utiles par exemples pour l'étude de pathologies ou de phénotypes pour lesquels les influences génétiques sont inconnues. L'approche la plus courante ressemble à une enquête inversée. Elle consiste à créer une large population de cellules dont une grande variété de gènes ont été mutées, puis d'identifier la/les mutations coupables. En d'autre termes la comparaison entre les cellules mutantes créées et le phénotype étudié permettent d'identifier la/les modification(s) génétique(s) liée(s) à ce phénotype.

Antérieurement les criblages génétiques reposaient en grande partie sur l'ARN interférence par shARN (pour *Short ARN /* ARN court )<sup>310</sup> qui ont tendance à avoir des effets hors-cible et qui peuvent conduire à des faux négatifs (due à l'extinction incomplète des gènes cibles). En comparaison le système CRISPR est capable de réaliser avec une spécificité accrue des modifications permanentes des gènes ciblés<sup>307</sup>. Actuellement la méthode la plus populaire pour réaliser des *genome-wide screen* (criblage du génome entier) en utilisant la technique CRISPR/Cas9 implique l'utilisation de librairies CRISPR lentivirales (nommées *pooled lentiviral CRISPR libraries* en anglais).<sup>104,202,311–315</sup>

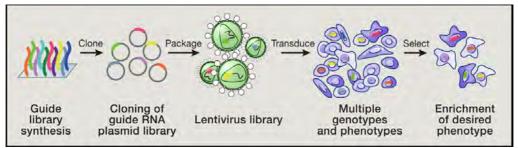


fig.51 Illustration du genome wide screening via la technique CRISPR/cas9

Chaque ARNg est conçu *in silico* à l'aide d'un des logiciels de conception<sup>133,134,316</sup> puis synthétisé. Cette librairie de guides sera ensuite amplifiée puis clonée dans des vecteurs lentiviraux. C'est la population hétérogène de vecteurs lentiviraux obtenus qui constituera la librairie CRISPR lentivirale. Des cellules exprimant la Cas9 seront ensuite transduites par les lentivirus provenant de la librairie dans le but d'activer, réprimer ou modifier en parallèle plusieurs gènes d'intérêt. Des cellules mutantes présentant de multiples génotypes et phénotypes seront obtenues. Le phénotype désiré pourra être enrichi en utilisant le tri cellulaire par cytométrie de flux en fluorescence (FACS: *fluorescence activated cell sorting*). Les séquences de ces cellules pourront ensuite être analysées par une technique de NGS (*Next Generation Sequencing*).

#### 2.2 Futur

Pour aller plus loin il sera nécessaire de développer des méthodes robustes qui apporteront des résultats cohérents quel que soient les lignées et les types cellulaires, les fabricants et les réactifs afin de générer des résultats valides indépendamment de l'utilisateur ou de la plate forme. Une autre considération pratique sera de rendre ces technologies de criblage bon marché et accessibles. La modification des sgARN pour les rendre plus courts, pourrait encore entraîner la diminution du coût des réactifs et permettre de plus largement démocratiser l'utilisation de ces technologies.

C'est que le criblage CRISPR à encore de multiples applications à explorer, la technique peut être utilisée pour identifier le rôle du génome dit « non codant » et de ses séquences régulatrices<sup>317,318</sup> multiplex pourra être utilisée pour disséquer la fonction des vastes étendues génomiques, en générant des microdélétions superposées. La dCas9 pourrait aussi contribuer à ces découvertes, fusionnée à des acteurs épigénétiques elle pourrait être utilisée pour étudier les effets de la méthylation ou de certains états de la chromatine sur la différentiation cellulaire ou l'origine de certaines pathologies.

#### 3. Création de modèles

L'édition génomique a été utilisée pour créer des modèles pathologiques chez le rongeur, le primate ou encore le porc et permet de diminuer l'écart entre ces animaux et les essais cliniques chez l'homme lorsqu'il s'agit de valider des preuves de concepts thérapeutiques<sup>211,319</sup>. En recréant chez un autre animal les mutations présentes chez la population de patients humains, l'édition génomique peut être utilisée pour modéliser une pathologie et se rapprocher au maximum de l'état physiologique humain. Elle permet de ne pas se reposer sur une phénocopie de la maladie. Le fait que la cas9 puisse être guidée pour cibler plusieurs séquences nous permettrait d'étudier des maladies communes comme le diabète ou certaines cardiomyopathies qui ont une composante

polygénique. La création d'animaux transgénique par Cas9 peut être réalisée en injectant la protéine et les transcrits sgARN directement dans les zygotes fertilisés elle aboutit à des modifications héritables d'un ou plusieurs allèles, par exemple chez le singe ou le rongeur<sup>316</sup>. En évitant les manipulations au stade des cellules souches embryonnaire, la durée du processus passe d'un peu plus d'une année à quelques semaines <sup>191,201,215,319</sup>. L'exemple qui suit compare l'efficience de la technique CRISPR-Cas9 à celle par recombinaison homologue « classique ».

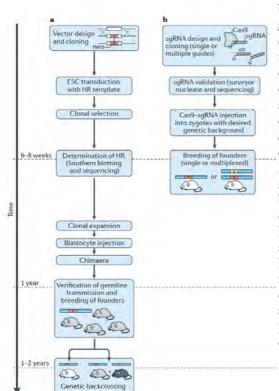


fig.52 Utilisation de Cas9 pour générer des animaux génétiquement modifiés<sup>170</sup>. a,b Comparaison des durées ciblage génique traditionnel utilisant recombinaison homologue (HR) « classique » dans des cellules souches embryonnaires (ESCs; partie a) et du ciblage génétique par CRISPR-Cas9 dans l'embryon au stade une cellule (partie b). Il y à deux phases chronophages et coûteuses dans l'approche par HR classique. La première de ces phases consiste en la conception et le clonage du vecteur, la transduction des ESCs, leur sélection et la création de chimères. La seconde comprend les croisement des animaux dans le but d'obtenir de multiples animaux ayant les traits d'intérêt. En revanche cloner les courts sgARN, vérifier la spécificité de leur guidage, injecter le duo Cas9sgARN dans le zygote par microinjection et identifier les animaux porteurs (nommés « fondateurs ») est relativement rapide 317. Puisque les embryons peuvent être obtenus à partir de n'importe quelle souche de souris et que l'édition simultanée de multiples gènes est possible, les croisements et rétro-croisements pour obtenir les lignées souhaitées sont inutiles dans le cas **b**.

### 4. GWAS et neuroscience

GWAS signifie *Genome Wide Association Study*, cette appellation définit des expériences qui examinent des variations génétiques chez différents individus, sur l'intégralité de leur génome, et les associent avec certains traits. Dans le cas des GWAS appliqués à l'homme, ces études comparent l'ADN de participants ayant différents phénotypes pour un trait ou une pathologie donnée. Les participants seront donc les personnes malades (les *cas*) et des individus similaire dépourvus de la pathologie (les *contrôles*). L'association de nouvelles approches: les iPSC (ou cellules souches pluripotentes induites, des cellules pluripotentes créées par reprogrammation de cellules différenciées adultes) et l'édition génomique par Cas9 vont étendre le champ de cette technique. L'application en neuroscience constitue un exemple frappant.<sup>320</sup>

Combinés iPSC, CRISPR et GWAS vont constituer un cocktail capable d'approcher et d'étudier *in vitro* les caractéristiques cellulaires et moléculaires associées aux troubles neurologiques (Alzheimer ou Parkinson, par exemple). Cette nouvelle approche possède plusieurs avantages, notamment le fait qu'une modification génétique peut être

étudiée dans différents contextes génétiques, les iPSCs conservant l'intégralité de l'information génétique des cellules de l'individu dont elles sont issues. Cela s'avère important pour les pathologies neurologiques complexes, car chaque variant génétique associé à ces maladies agit de concert avec de nombreux autres allèles. Un autre attrait est le fait que ces cellules modifiées peuvent être re-différenciées en n'importe quel type cellulaire y-compris les neurones difficiles d'accès et pratiquement dépourvus d'activité mitotiques.

Des modèles pathologiques basés sur les iPSCs ont été créés pour plusieurs pathologies neurologiques, incluant la maladie de Parkinson, Alzheimer et Huntington 322-325. Ils ont montré qu'ils pouvaient se rapprocher des caractéristiques cellulaires et moléculaires de ces pathologies. L'édition génomique appliquée à ces modèles peut être utilisée pour réaliser des analyses à haut débit des liens unissant les voies de signalisations impliquées dans la pathogenèse et ce que l'on nomme des *risk variants*. ( que l'on peut traduire par gène de prédisposition, désigne un allèle ou un haplotype pour lequel un risque plus élevé d'existence de la maladie serait présent). L'induction par édition génomique d'un phénotype ou sa correction, dans les cellules différenciées, pourrait également permettre de valider des cibles thérapeutiques potentielles.

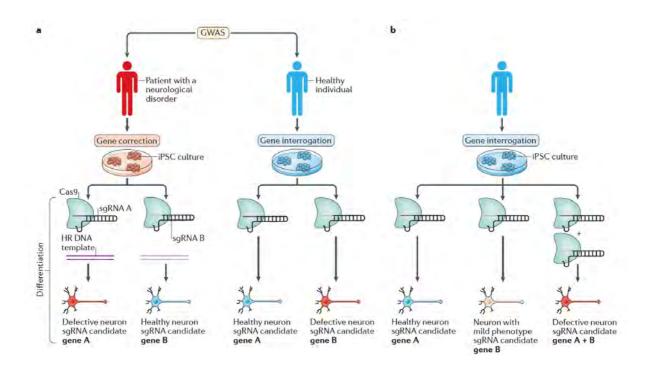


fig 53 : Applications de la Cas9 associées aux cellules iPSCs humaines 320

a Identification de gènes potentiellement associés à la pathogenèse dans le cadre d'une étude GWAS. Il est difficile de réaliser des cultures primaires (cellules issues directement de tissus) de neurones. En revanche des iPSCs dérivées de cellules somatiques (e.g. des fibroblastes) d'individus sains ou de patients atteints de troubles neurologiques peuvent se différencier en neurones et être cultivés *in vitro*. Les gènes associés aux pathologies peuvent être examinés de deux manières différentes : la *correction* et *l'interrogation*. Si la *correction* (exemple de gauche) site spécifique par HR et CRISPR du gène d'intérêt « secours »le phénotype pathologique (gène d'intérêt B dans l'exemple), l'hypothèse impliquant ce gène

dans la pathogenèse est validée. L'*interrogation* (à droite) s'apparente au criblage, elle consiste à muter le gène d'intérêt dans une cellule saine (à droite), dans le cas ou cela conduit au phénotype pathologique, l'implication du gène est confirmée (cas du **gène d'intérêt b**).

**b** La contribution de chaque élément génétique dans le cas de pathologies impliquant des variations multigéniques telles que les maladies d'Alzheimer et Parkinson peut aussi être évaluée en utilisant la méthode CRISPR multiplex. Cette approche permettrait de dénouer les possibles synergies (comme montré dans l'exemple avec les **gènes d'intérêts a et b**) et d'identifier les corrélations entre divers phénotypes d'une maladie et certains génotypes ou haplotypes.

### **VI PANORAMA**

#### 1.Introduction

Ce chapitre a pour but d'approfondir plusieurs thèmes évoqués précédemment :

- Le processus inventif et la propriété de la technique CRISPR-Cas9
- L'émergence d'un marché de l'édition génomique
- La régulation de l'usage clinique de l'édition génomique
- Notre compréhension incomplète des mécanismes génétiques

### 2. Propriété de la technique CRISPR-Cas9

#### 2.1 Introduction

Les spéculations pour le prix Nobel débutent chaque automne, lorsque l'agence de presse Thomson Reuters publie ses prédictions pour le prix scientifique. En automne 2016, l'outil CRISPR-CAS9 était parmi les grands favoris pour le prix Nobel de Chimie <sup>326</sup>. Il s'agit d'une nomination inhabituelle, tout d'abord car les publications les plus citées n'ont que 3 ans, un claquement de doigt à l'échelle de la science, ensuite car la technique est au centre d'un combat terrible concernant la propriété de son brevet. Finalement en 2016, le prix fut décerné à Jean-Pierre Sauvage, James Fraser Stoddart et Bernard L.Ferriga pour leur travaux sur la conception et la synthèse de machines moléculaires <sup>327</sup>. Thomson reuters base en partie ses prédictions sur le nombre de citations qu'une publication déclenche, depuis 14 ans l'agence a correctement nommé 39 futurs lauréats. Par contre il est intéressant de remarquer que la récompense fut décernée la même année que la prédiction de l'agence à seulement ½ des nominés.

Pour CRISPR les publications les plus citées sont celles d'Emmanuelle Charpentier et de Jennifer Doudna, respectivement au Max planck institute et à UC Berkeley. Un troisième candidat au titre est Feng Zhang, le biologiste moléculaire du Broad Institute d'Harvard et du MIT qui possède le brevet pour la technique CRISPR-Cas9 et déclare être parvenu à sa découverte de manière indépendante. Les deux groupes et leurs avocats spécialisés en propriété industrielle se livrent une longue et âpre bataille pour conserver la propriété de la découverte. L'enjeu commercial est majeur et leurs compagnies rivales ont chacune déjà enregistré des séries de brevets.

### 2.2 Les découvertes multiples

Au delà des colossales sommes d'argent investies et des subtilités judiciaires, attribuer le mérite d'une grande découverte scientifique dépasse le simple choix entre ces deux groupes rivaux. L'histoire des sciences est plus complexe, liée à la manière dont surgit la créativité et l'innovation. Il est fréquent que des groupes scientifiques parviennent à la même période à des découvertes similaires de manière indépendante ou non : ces événements sont appelés des *Découvertes Multiples*. Par exemple Isaac Newton et Gottfried Leibniz se disputèrent la primeur de la découverte du calcul infinitésimal. Scheele, Priestley et Lavoisier déclarent chacun avoir « découvert » l'oxygène de manière indépendante entre 1771 et 1774. <sup>328</sup>



**fig.54** Trois modèles d'ampoules à incandescence parmi les 22 développés. Parmi ces modèles créés de manières indépendantes par Edison, Swan et Maxim <sup>329</sup> c'est le modèle d'Edison qui fut le plus largement adopté, grâce à sa conception et le fait qu'il faisait partie d'un système complet d'éclairage, l'offre d'Edison comprenant entre autre une alimentation, un réseau électrique et un générateur.

La primeur n'a que peu de sens du point de vue de l'histoire des sciences. La recherche étant une œuvre collective et évolutive. L'invention du calcul infinitésimal par exemple inclue une multitude d'innovations apportées par des auteurs variés. Comme le sociologue des sciences Robert Merton le décrit, tous les domaines scientifiques sont construits sur ce qu'il appelle une « base culturelle cumulative », c'est à dire que les découvertes n'apparaissent pas au hasard mais sont plutôt des produits de leur temps. Les scientifiques s'attardent donc sur les défis de leur siècle, en utilisant les outils de leur époque, c'est donc peu surprenant qu'ils parviennent indépendamment à des observations proches.

Mais il ne peut cependant n'y avoir qu'un maximum de 3 personnes primées au Nobel pour une même découverte, et un brevet ne peut être attribué qu'à un seul groupe d'inventeurs.

2.3 Combat catégorie Poids Lourds: Patent Interference No. 106,048 330

#### 2.3.1 Contexte

Les composants du système CRISPR ne peuvent pas être brevetés, la Cas9 est une protéine présente dans la nature et elle fait partie de la physiologie naturelle de certaines bactéries. Mais ce système n'est pas utilisable tel quel dans les cellules eucaryotes pour servir d'outil biotechnologique. Ce qui est aujourd'hui breveté par le *Broad Institute*, ce sont les ajustements apportés pour transposer et optimiser cette

protéine et ses guides comme outil d'édition génomique, et c'est la propriété de ce brevet qui est remise en cause.

Deux demandes de brevets pouvaient se voir attribuer la propriété des techniques d'édition génomique basées sur l'utilisation de la Cas9 et de son guide ARN. La première avait été déposée en Mai 2012 par Jennifer Doudna, *UC* Berkeley et l'Université de Vienne, la seconde en Décembre 2012 par Feng Zhang et le *Broad Institute*<sup>331</sup>. Zhang et son institution avaient réalisé des demandes dites « accélérées » moyennant un prix légèrement plus élevé de leur enregistrement (un supplément de 70\$85). Un an plus tard Jennifer Doudna reçut un *e-mail* l'informant que Zhang, du *Broad institute* avait reçu l'attribution de brevets couvrant la technique d'édition génomique CRISPR-Cas9<sup>330</sup>. Cet incident fut l'élément déclencheur de la bataille rangée opposant les deux institutions. UC Berkeley, en conséquence soumis un « *interference proceeding* » à l'*United States Patent and Trademark Office* ou USPTO.

### 2.3.2 Interférences

Dans l'histoire de la commercialisation de CRISPR-Cas9 un détail a son importance : le timing. La législation américaine sur les brevets prévoyait que, dans le cas d'un contentieux entre inventeurs, le premier à l'avoir inventé (*first-to-invent*) se verrait attribué le brevet. Depuis Mars 2013, un système différent gouverne les brevets, le « *first-to-file* », qui récompense le premier à déposer le dossier de demande de brevet. Mais de nombreux brevets clefs pour CRISPR sont antérieurs à Mars 2013 et tombent donc sous l'ancienne juridiction.

Le PTAB330, patent trial and appeal board, a accepté la demande d'interférence de UC Berkeley, une audition devant un jury d'expert aura donc lieu pour déterminer qui fut le premier à inventer la technique utilisant la Cas9 et ses guides pour réaliser une édition génomique chez l'eucaryote (En annexe 13 est présenté le texte exact décrivant l'objet de l'interférence). Les deux institutions sont en train d'investir des sommes conséquentes ( plusieurs millions de dollars 332) pour défendre leur droit à détenir et licencier l'invention. L'avantage semblait au départ être pour l'équipe de Doudna, désignée en tant que Senior dans l'audience<sup>83</sup>. Le Broad institute est donc junior, cela signifie que jusqu'à qu'ils prouvent le contraire, Doudna, Charpentier leur co-dépositaires seraient potentiellement reconnus comme les pionniers de la technique d'édition génomique CRISPR-Cas9. Mais de nombreux facteurs entrent en compte, le 6 décembre 2016 a eu lieu l'audition préliminaire, première et unique occasion à laquelle les deux parties pourront communiquer avec le juge avant que le jugement sur le brevet ne soit rendu. Au centre des débats : l'indépendance des travaux réalisés l'intention inventive et le fait que les chercheurs du Broad Institute aient été les premiers à publier leurs expérimentations de la technique sur des cellules eucaryotes (suivis de près par ceux d'UC Berkeley) 75,76.

Plusieurs chercheurs ayant suivi le débat estiment que le comité pourrait décider de n'attribuer le brevet à aucun des partis et que les deux institutions pourraient « ressortir de la procédure chacune avec quelquechose » <sup>330</sup>. Même s'ils désignent un «vainqueur » il ne pourrait être que provisoire, les procédures d'appels sont nombreuses et complexes et de nouveaux acteurs font leur apparition (l'Université de Vilnius avec Virginijus ŠiksnyS <sup>79</sup>)\*. A l'image d'une autre interférence dans le domaine de la biotechnologie soumise pour l'invention du vaccin HPV la décision pourrait donc prendre des années

(pour le vaccin HPV le processus d'interférence s'était déroulé sur 6 années durant lesquelles de multiples procédures d'appel avaient eu lieu) 83,334.

\* Durant la rédaction de cette thèse le verdict de la procédure d'interférence a été rendu. Les juges de l'USPTO ont conclu à l'absence d'interférence. Ce qui signifie que le brevet déposé par le Broad Institute est maintenu et que les deux inventions sont considérées comme étant distinctes. Le Broad Institute détient donc un brevet relatif à l'utilisation de la technologie CRISPR-Cas9 chez l'eucaryote, son application la plus lucrative. Mais des incertitudes demeurent : leur « rivaux » d'U.C. Berkeley soutiennent que leur brevet serait, malgré cette décision , applicable à tous les types cellulaires. Si cette affirmation est juridiquement validée l'utilisation commerciale de CRISPR-Cas9 chez l'eucaryote nécessitera au minimum deux licences, l'une avec le Broad Institute, l'autre avec U.C.Berkeley.<sup>335</sup>

### 2.3.3 Perspective: innovation et Cpf1

Fin 2015, Feng Zhang et ses collègues ont publié la découverte et l'usage d'une nouvelle enzyme : Cpf1 et ont déposé une demande de brevet couvrant son usage biotechnologique en Janvier 2016 <sup>132,336</sup>. Faisant partie du système CRISPR, Cpf1 est une alternative à la Cas9 dotée de plusieurs avantages : elle est plus petite, est dirigée vers ses cibles par un unique ARN guide, a un PAM différent riche en T (5′ -TTN- 3′) et surtout, elle génère des bouts collants qui rendraient le transfert de gène de manière directionnelle possible. Elle a aussi un avantage non biologique de taille : elle n'entre pas dans le cadre des précédents brevets et des disputes portant sur leur attribution.

L'attribution de brevets peut sembler contraire à la dynamique des découvertes scientifiques comme illustré avec l'exemple des *découvertes multiples*. Malgré tout il est intéressant de constater que leur existence force à l'innovation et, à la manière d'Edison et de son ampoule à incandescence, il se pourrait que sur le marché de l'édition génomique l'invention la plus apte sera celle qui sera la plus adoptée.

# 3. Émergence d'un marché de l'édition génomique

Comme nous l'avons vu au fil des exemples abordés, les applications de l'édition génomique et celle de CRISPR s'étendent à travers toute les industries qui touchent au vivant. Danisco et ses yoghurts évoqués dans la première partie de cette thèse était sans doute parmi les pionniers d'une utilisation commerciale de CRISPR<sup>47</sup>, mais d'autres marchés ont rapidement émergés. (En annexe 14 est présenté un tableau illustrant l'intérêt industriel pour la technique CRISPR-Cas9 au travers des licences et brevets établis). Le marché de l'agriculture en est un exemple, celui de la santé humaine aussi. CRISPR a déclenché un renouveau de l'édition génomique dans le marché de la santé, avec des applications théoriques en thérapie génique, thérapie cellulaire, immunothérapie, infectiologie et récemment dans le domaine des diagnostics des maladies infectieuses 247,337. Il est difficile de prévoir la valeur de ce marché naissant et la place qu'aura la technique CRISPR-Cas9 dans celui-ci. Mais le marché des nouveaux outils d'édition génomique est en pleine expansion comme en témoigne les investissements initiaux dans les compagnies pionnières Caribou (11 millions de \$ dont 2.9 provenant de Novartis ), Editas (43 millions de \$), CRISPR Therapeutics (89 millions de \$) et Intellia (15 millions \$)3,336. Depuis début 2013, les entreprises basées sur la technologies CRISPR-Cas9 ont récolté plus d'1 Milliards de dollars en capital-risque et sur le marché public <sup>339</sup>. L'intérêt commercial de la technique n'a pas échappé aux grandes entreprises pharmaceutiques. Novartis, Pfizer, Bayer, et Astrazeneca ont développés des partenariats et contracté des accords de licences avec ces entreprises. De manière prophylactique, leur engagement se fait souvent avec des compagnies rivales, une forme d'assurance par rapport aux

jugements portant sur la propriété et l'étendue des brevets.

Le dernier marché est celui des outils de recherche, Sigma-Aldrich a d'ores et déjà établi des licences dans le but de distribuer des outils destinés à la création de plantes modifiées, de modèles animaux, de création de lignées cellulaires sur mesure et de criblage génétique <sup>3</sup>. C'est peut être celui-ci qui connaîtra le développement le plus florissant dans les années à venir, avant que les connaissances nécessaires soient accumulées pour éviter les risques et surmonter les limites inhérentes à l'édition génomique dans le cadre des deux domaines précédents.

### 4. émergence de réflexions sur l'usage clinique de l'édition génomique

#### 4.1 Introduction

1975. Le contexte de guerre froide, du scandale du Watergate, et le fait que certains des premiers généticiens/biologistes moléculaires aient participé au projet Manhattan, font redouter à la communauté scientifique que le génie génétique ne serve un jour à concevoir de nouvelles armes. En réponse à ces craintes la conférence d'Asilomar fut organisée<sup>340</sup>. Elle avait pour but d'établir ce qu'impliquait notre capacité à décrypter et réorganiser les gènes. La conférence californienne réunissait 150 scientifiques internationaux pour discuter du moratoire sur une technique capable de grandes avancées médicales, mais soulevant également de nombreuses questions éthiques. Les scientifiques conclurent à une absence de moratoire, mais avec mise en place de limites, des conditions de précaution et de sécurité renforcées. Bien qu'elle ne prenait pas en compte les aspects économiques et environnementaux, rassemblait des participants peu diversifiés et n'apportait que les ébauches d'un cadre réglementaire, son influence est encore présente dans notre manière d'encadrer le génie génétique. Voici un extrait de leurs conclusions :

#### V. CONCLUSIONS

We believe that considerable benefits are likely to result from experiments involving the genetic alteration of microorganisms. The range of possible benefits extends from the use of these techniques to add to our knowledge of basic biological phenomena, to possible practical applications in the areas of agriculture and medicine.

We believe also that a scale of risks exists in the construction of genetically altered microorganisms, and we are uncomfortable about our inability to assess precisely the extent of such risks for many types of experiments. However, we believe that the containment procedures described in this proposal will reduce any risk to laboratory workers and to the environment to a level that is acceptably low and which will allow investigators to carry out research in this area. We believe that certain experiments should presently not be carried out under any circumstances (i.e. Class VI), but that most experiments can be done if containment facilities appropriate to the risk are utilized.

**fig.55** Extrait des conclusions des *Guidelines* ébauchées à Asilomar en 1975, l'intégralité est présenté en Annexe 15.

### 4.2 Une discussion globale

Nous sommes à un moment similaire, une accélération du génie génétique, qui

met à notre portée des possibilités dont l'implication sociale et éthique doit être évaluée. Rappelant celui d'Asilomar le *Comittee for the International Summit on Human Gene Editing* a été organisé au début du mois de Décembre 2015 à Washington à l'initiative de la National Academy of Science américaine, l'européenne Royal Society et l'académie scientifique chinoise <sup>339</sup>. Parmi les 500 participants on retrouve les membres fondateurs du comité d'Asilomar, et des scientifiques évoqués au cours des chapitres précédents dont : Jennifer Doudna, Emmanuelle Charpentier, Feng Zhang, George Church, Jin soo Kim, ou encore le directeur de l'institut de biomédecine et santé de Guangzhou ( où les premières recherches académiques sur l'embryon humain ont été réalisées). Ils rassemblent leurs conclusions dans un communiqué en 4 points <sup>342</sup>, ils sont présentés (traduits en français) à la page suivante <sup>342</sup>.

- 1 Au sujet des recherches fondamentales et précliniques . La nécessité de telles recherches dans un cadre légal et soumises à une réflexion éthique est évidente. Ces recherches auront pour but d'évaluer le rapport bénéfice/risque d'éventuelles applications cliniques et de mieux comprendre la biologie des embryons humains et cellules germinales. Si pendant ces processus de recherches un embryon ou une cellules germinale humaine était éditée, la cellule modifiée ne devrait pas aboutir à une grossesse.
- 2. Au sujet de l'usage clinique des cellules somatiques. Le comité souligne le besoin de comprendre les risques, tels que l'édition imprécise, et les potentiels bénéfices de chaque modification génétique proposées. Étant donné que l'usage clinique de ces modifications n'affecteraient que les individus traités, elles peuvent être évaluées de manière appropriée par le système réglementaire actuel qui sera emmené à évoluer. Ces mêmes régulateurs peuvent approuver les essais cliniques et l'usage de ces thérapies en prenant en compte le rapport entre leur bénéfices et risques
- Au sujet de l'usage clinique des cellules germinales. L'édition des cellules germinales 3. soulève de nombreux problèmes : a le risque d'édition imprécise (telles que les mutations horscibles) et l'édition incomplète de cellules à un stade embryonnaire précoce (déclenchant un mosaïcisme cellulaire) ;b la difficulté de prédire les effets délétères qu'une modification génétique peut engendrer soumise aux multiples circonstances que connaît la population humaine, par exemple l'interaction avec d'autre variants génétiques et l'environnement; c L'obligation d'envisager les implications à la fois pour l'individu traité mais aussi pour les générations futures qui porterons ces altérations génétiques; d le fait qu'une fois introduite dans la population humaine, il sera difficile d'y retirer la modification apportée, celle-ci ne connaîtra pas de frontière nationales: e la possibilité qu'une « amélioration génétique » permanente apportée à une fraction de la population puisse exacerber des inégalités sociales ou puisse être utilisée de manière cohercitive; f les considérations éthiques et morales liées au fait d'altérer volontairement l'évolution humaine en utilisant la technologie. Il serait irresponsable de procéder à toute utilisation clinique de l'édition de la lignée germinale à moins que (i) les problèmes pertinents en matière d'innocuité et d'efficacité n'aient été résolus, basés sur une compréhension et une évaluation appropriée des risques, des avantages potentiels et des alternatives; (ii) Un vaste consensus social au sujet de la pertinence de l'application proposée n'ait été établi. De plus, toute utilisation clinique ne devrait se faire que sous la surveillance réglementaire appropriée. À l'heure actuelle, ces critères n'ont pas été rassemblés pour toute utilisation clinique proposées : les questions de sécurité n'ont pas encore été suffisamment explorées; Les cas qui apporteraient un bénéfice conséquent sont limités; Et de nombreux pays ont des interdictions législatives ou réglementaires de la modification de la lignée germinale. Cependant, au fur et à mesure que les connaissances scientifiques évoluent et que les opinions de la société évoluent, l'utilisation clinique de l'édition de la lignée germinale devrait être réévaluée régulièrement.

#### 4. Le besoin d'un « Forum Continu ».

Alors que chaque nation possède l'autorité de réguler les activités placées sous sa juridiction, le génome humain est partagé par tous les hommes, quelque soit leurs nations. Le comité estime que la communauté internationale devrait s'efforcer d'harmoniser les régulations, dans le but de décourager des activités inacceptables tout en permettant l'avancée de la santé et du bien-être humain .

Ce forum devrait être inclusif, s'étendant entre les nations et englobant plusieurs perspectives et expertises. Incluant entre autres des scientifiques biomédicaux, sociaux, des acteurs de santé, des patients et leur familles, des invalides, des responsables politiques, des régulateurs, des investisseurs, des chefs spirituels, des défenseurs de l'intérêt public, des représentants industriels et des membres du public.

### 4.3 Perspective et approche culturelle

Une telle réunion est un témoignage des progrès réalisés et de la manière dont l'édition génomique est intimement liée à la société. En 1975 c'était un groupe composé principalement de scientifiques américains qui se réunissait à Asilomar, et qui ébauchait les lignes directrices qui encadreraient les recherches portant sur l'ADN recombinant. Quarante ans plus tard, c'est un groupe à l'expertise diverse, international et multiculturel qui s'est rassemblé pour conclure à une recommandation. Celle de ne pas stopper les recherches portant sur l'édition génomique humaine, et de s'abstenir d'appliquer la technique aux cellules embryonnaires humaines à des fins reproductives.

Les spécialistes en éthiques ont longtemps considérés les implications liées à la modification de la lignée germinale chez l'homme. Pendant des années ces inquiétudes n'étaient que théoriques, jusqu'aux récentes avancées dans le domaine de l'édition génomique<sup>338</sup>. L'outil CRISPR-Cas9 n'est que le premier des outils d'édition génomique dont l'usage simple ne demande que peu d'investissements. Les inquiétudes portant sur les lieux où la juridiction est plus laxistes sont grandes. Les récents débats espèrent que les scientifiques éviteront toutes tentatives de modification du génome de la lignée germinale dans les zones moins régulées, tant que les implications techniques et morales de telles expériences ne sont pas collectivement considérées. Bien qu'un tel moratoire sera difficile à appliquer et peut sembler incapable d'exercer une influence globale, on peut voir un espoir dans le précédent d'Asilomar qui avait freiné l'utilisation de l'ADN recombinant tant que sa régulation n'était pas établie.

Au sujet de la recherche fondamentale portant sur les cellules embryonnaires humaines, à présent, 4 nations (Suède, Royaume-uni, Japon et Chine) ont approuvé l'application de la technique CRISPR-Cas9 à ces expériences 343,344. Le cadre de ces recherches est strict, les zygotes modifiés ne doivent en aucun cas être capables d'aboutir à la naissance d'un individu. De telles applications de l'édition génomique permettront de mieux comprendre les mécanismes complexes du développement embryonnaire 345. L'aspect culturel à également une importance capitale, les réactions qu'avaient suscité la première utilisation embryonnaire de la technique CRISPR-Cas l'ont montré. Ce n'est pas un hasard si la première modification embryonnaire à eu lieu en Chine et si elle a suscité autant de réactions, y compris dans la communauté scientifique. Contrairement à l'influence judéo-chrétienne, la pensée confucéenne considère que l'homme n'est une « personne » qu'après sa « naissance » ; la régulation chinoise autorise l'expérimentation

### 5. Au sujet de notre compréhension des mécanismes génétiques

#### 5.1 Introduction

L'approche scientifique moderne repose grandement sur l'assimilation des organismes à des machines, métaphore débutée par Descartes en 1635. les gènes sont souvent considérés comme des « *plans* » et l'ADN comme une « *information* »,

Le premier séquençage génétique complet eut lieu à l'université de Ghent en 1972, celui d'un gène codant une protéine de la capside du bactériophage MS2<sup>346</sup>. Il y à 40 ans, la génétique était un domaine nouveau, à ce moment là les gènes étaient considérés à eux seuls comme la clef des mécanismes naturels. Puis des imperfections dans notre compréhension sont apparues : l'importance de l'environnement, du contexte et de l'agencement de ces gènes. Aujourd'hui nous commençons à peine à entrevoir les complexes interactions entre les gènes, leur environnement et leur écosystème.

Le message central de l'oeuvre de Richard Lewontin, ancien titulaire de la chaire de biologie de l'université d'Harvard, est que nous ne parviendrons pas à comprendre le vivant tant que nous penserons aux gènes, aux organismes et à l'environnement comme des entités séparées. 347

« Le gène, en réalité, définit des potentialités que les infinies variations des conditions - internes et externes - des organismes réaliseront ou non. Un organisme ne se développe donc pas à partir du programme génétique : il est façonné par quantité de phénomènes aléatoires. L'environnement, par exemple, qui n'est pas une sorte de toile de fond immuable sur laquelle s'agitent les organismes, mais qui participe pleinement au développement de chacun d'eux » 347

Au schéma classique gène/organisme Lewontin suggère de substituer la triade « gène/organisme/environnement » plus complexe mais beaucoup plus riche de perspectives nouvelles. Nous devons considérer le fait que les systèmes biologiques sont caractérisés par une immense hétérogénéité interne, physique et chimique. Ils subissent des échanges dynamiques intenses entre les différents processus internes qui les animent mais aussi avec le monde extérieur.

### 5.2 Exemple du lien gène/organisme/environnement

Pendant la transcription, le code ADN est copié dans un ARN, l'ARN messager entre ensuite dans le processus de production des protéines. La séquence nucléotidique précise qui constitue ce message influence la vitesse avec laquelle chaque petits bouts seront traduits en portion de séquences d'acides-aminés pendant la construction des protéines. Ces variations de vitesses influencent le repliement de la protéine au fur et à

mesure qu'elle est produite. Le repliement protéique dépend aussi de l'environnement cellulaire. L'insuline humaine par exemple est aujourd'hui produite par des bactéries recombinantes qui contiennent le gène humain de l'insuline. Quand le gène humain a été introduit pour la première fois dans les bactéries, la protéine produite était dépourvue d'activité physiologique, bien que sa séquence en acide-aminé restait identique. Il s'avéra qu'elle était repliée de manière incorrecte par la cellule bactérienne, un problème qui fut résolu en adaptant le milieu de culture.

### 5.3 Au sujet de l'édition génomique et des maladies multigéniques.

L'évolution du domaine de l'édition génomique passera par la prise en compte des limites de notre propre compréhension. Contrairement aux rares maladies causées par un changement qui ne touche qu'un seul gène, des pathologies communes comme le cancer ou le diabète sont liées à une multitude de variations génétiques différentes pour chaque patient. Deux hypothèses encadrent ces facteurs de risque génétiques :

L'hypothèse du CDCV : *Common disease / Common Variant* : c'est à dire des variations courantes qui favorisent l'apparition d'une pathologie courante et celle du CDRV : *Common Disease / Rare Variants* : des multitudes de variantions peu fréquentes qui associées participent à l'apparition de pathologies courantes<sup>348</sup>

Une approche curative basée sur l'édition génomique nécessitera d'appréhender la complexité de leur pathogenèse. Elle est liée à des variations génétiques portées par de multiples locus, chacune de ces variations ayant un faible effet mais collectivement et associées à l'interaction avec l'environnement, elles aboutissent au développement pathologique.

### 5.4 Au sujet de l'édition génomique et des avancées futures

La recherche utilise les outils à sa disposition, avec comme facteur le temps et les ressources financières. La création de nouvelles techniques expérimentales est en partie liée aux questions qui animent la communauté scientifique et il y à souvent un lien intime entre méthode et problématiques. L'invention du séquençage automatique de l'ADN était une réponse à la demande croissante de séquences génétiques, la disponibilité de ces machines et la facilité avec laquelle l'ADN a pu être séquencé a déclenché une multitudes de travaux portant sur l'étude de ces séquences. 347

Les progrès futurs de la biologie ne dépendront peut être pas de la découverte de concepts révolutionnaires, mais plutôt de l'existence de nouvelles méthodes d'exploration du vivant, dont le coût et la simplicité correspondront à nos ressources. Si notre compréhension des mécanismes génétiques est une limite à l'application actuelle de l'édition génomique, ces mêmes techniques pourraient servir à sonder le vivant en quête de réponses. La véritable révolution CRISPR pourrait être celle de la recherche.

# **CONCLUSION**

Dans le tout premier chapitre, la première citation est la suivante :

« Aujourd'hui l'édition génomique ressemble à une découpe au scalpel, dans les années 1960s par comparaison, ils tapaient sur les gènes à coups de marteau » <sup>7</sup>

Mais il est probable que dans quelques années, à la relecture de cette thèse, j'esquisse un sourire amusé en me rendant compte à quel point les techniques que j'y ai décrites sont devenues obsolètes. Et il est aussi possible que les concepts biotechnologiques actuels apparaissent aussi archaïques que le bombardement radioactif des plantes pratiqué dans les années 1960s pour introduire des modifications génétiques. Toujours est-il que CRISPR-Cas9 comme ces radiations constituent chacun, à leur période respective, de formidables outils d'exploration scientifique.

En ce qui concerne l'usage thérapeutique (maladies monogénétiques, latence virale, immunomodulation anticancéreuse), de prometteuses preuves de concepts sont réalisées. Face aux alternatives l'édition génomique ne sera pas toujours la solution la plus adapté ou la plus accessibles, mais parfois elle laisse entrevoir des possibilités de thérapies futures pour des pathologies que nous pensions incurables. Elles montrent les limites de l'outil (effets hors-cible orientant la technique vers une approche *ex vivo*) mais il se pourrait que dans certain cas l'urgence fasse passer les bénéfices devant nos incertitudes (e.g. antibiothérapie).

Comme nous l'avons vu au cours des chapitres précédents, la science est parfois une puissance amorale, c'est à la société de lui conférer cette dimension. CRISPR fait partie de ces outils dont l'usage est capable de redéfinir notre place au sein de la nature (e.g. génétique dirigée), notre définition de celle-ci (e.g usage agronomique) ou même notre propre évolution (e.g. édition embryonnaire). Le rôle que doit occuper la société dans le développement de l'édition génomique est multiple, il comprend sa régulation, sa gouvernance et donc la définition de son usage. L'édition génomique est une technologie naissante, mais elle évolue à un rythme effréné, dans un monde globalisé et multiculturel, les discussions internationales et multidisciplinaires autour de ces thèmes ont donc une importance majeure, celles-ci ont commencé à se structurer. Comme pour l'application thérapeutique le principe de précaution aura toujours autant

d'importance.

Si l'usage des CRISPR montre les limites de notre compréhension il nous permet aussi de l'approfondir (e.g animaux modèles, génomique fonctionnelle). C'est un catalyseur, il réduit la quantité de temps et de ressources nécessaires et son 'usage en recherche fondamentale sera peut être le déclic requis pour permettre son évolution en tant que solution thérapeutique.

Dans un monde où chaque centimètres de terre semble avoir été cartographié, le frisson des grandes découvertes pourrait bien se situer dans l'infiniment petit.

# **ANNEXES**

### Annexe 1:

	Zinc Finger Nuclease	TALENs	Cas9	Meganuclease
Site reconnu	9 à 18pdb par monomères, 18 à 36 pbp par paire de ZFN	14 à 20pdb par monomères, 28 à 40 pbp par paire de TALEN	22pbp (20bp séquence guide + 2pbp sequence PAM pour <i>S.</i> pyognes Cas9); jusqu'à 44 pbp par double nicking	entre 14 et 40 bp
Spécificité	Small number of positional mismatches tolerated	Small number of positional mismatches tolerated	Positional and multiple consecutive mismatches tolerated	Small number of positional mismatches tolerated
Contraintes	Difficulté à cibler des séquences non riches en G	La base en 5' de la cible doit être unT pour chaque monomère	La séquence cible doit précéder un PAM	Cibler de nouvelles séquences entraine une faible efficacité
Facilité de conception	Difficile, peut nécessiter de considérables efforts	Modéré, nécessite des méthodes complexes de clonage	Facile, désigner une nouvelle cible nécessite seulement des techniques standard de clonage et une synthèse d'oligonucleotides	Difficile, peut nécessiter de considérables efforts
Immunogenicité	Probablement faible, Les Zfs étand basé sur une structure de protéine humaine. Les domaines Fok1 dérivé de bactéries pourraient être imunogènes	Inconnus, protéines dérivées de Xanthamonas sp.	Inconnu, protéines dérivées de diverses espèces bactériennes	Inconnu, les meganucleases peuvent provenir de multiples organismes dont des eucaryotes
Facilité de délivrance aux cellules <i>ex vivo</i>		oar des méthodes telles q	que l'électroporation ou	la transduction virale
Facilité de délivrance <i>in vivo</i>	Relativement facile grâce à la faible taille des cassettes d'expression ZFN permettant l'utilisation de divers vecteurs viraux	Difficile à cause de la taille importante de chaque TALEN et de la nature répétitive de l'ADN codant les TALENs qui entrâine des évenements recombinatifs non souhaités lorsqu'il est emballé dans des vecteurs lentiviraux	Modéré, la spCas9 communément utilisée à une taille élevée ce qui peut causer des problèmes de vectorisation virale, mais de plus petits orthologues existent.	Relativement facile grâce à la faible taille des méganucléases ce qui permet l'utilisation de divers vecteurs viraux
Utilisation possible en multiplex	Faible	Faible	Élevée	Faible

propriétés comparées des méthodes d'édition génomiques<sup>349</sup>

Annexe 2 :

Features of the sequences most similar to CRISPR spacers from E. coli

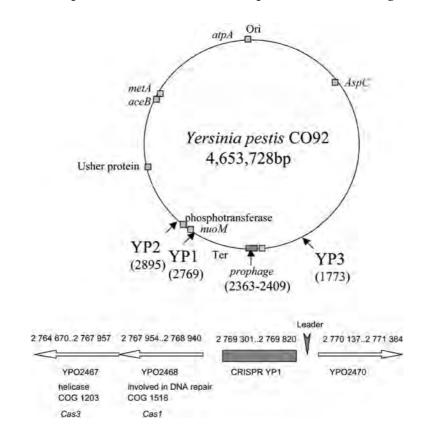
Strain	Gene	Element	Activity	Alignment
ECOR42	tral	Plasmid F	Helicase	gtttcccgtgcgtcgtatgaggcagaaaagag 
ECOR44	Unannotated	Phage P1	Unknown	ctgttggcaagccaggatctgaacaataccgt 
ECOR47	darB	Phage P1	Methylase	gctggtggcgcgggcaaacggaacaatcccgc !
ECOR49	resD	Plasmid F	Resolvase	atcgacttatgccccatcaggctctgcaatac

<sup>&</sup>lt;sup>a</sup>CRISPR-spacer sequence (top line) and best-match homologous sequence (bottom line).

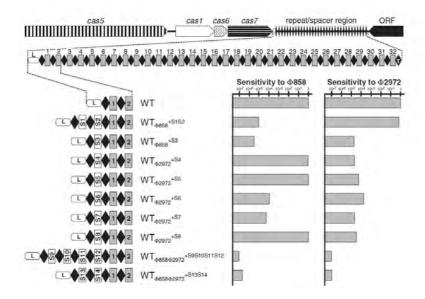
Similitudes entre les séquences spacers de locus CRISPR de certaines souches d'*Escherichia coli* et des séquences du phage P1  $^{44}$ 

#### Annexe 3:

Position des locus CRISPR et des spacers homologues au sein du génome de *Y.pestis* CO92. La position des trois locus CRISPRs sont indiqués en kb. Sont aussi indiquées : l'origine de réplication (ORI) de terminaison (Ter). Les gènes de chaque côté du locus CRISPR YP1 et leur positions exactes sont indiquées au bas de la figure.<sup>28</sup>



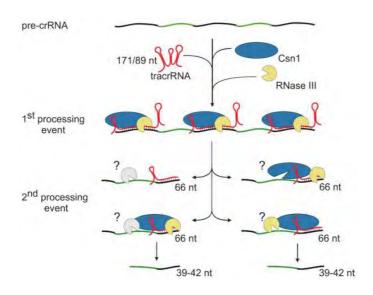
#### Annexe 4:



### Première preuve d'immunité acquise via Crispr 12

Locus CRISPR de *Streptococcus thermophilus*, mise en évidence de l'acquisition de nouveaux *spacers* chez les mutants résistants aux phages. Le locus CRISPR de la bactérie sauvage (WT) est représenté en haut. La région *repeat*-spacer WT représentée contiens: des segments répétés (symbolisés par des diamants noirs), les *spacers* (rectangles gris numérotés), une séquence leader (lettre L encadrée) et des répétitions terminales (diamant noir contenant la lettre T). En bas à gauche, les *spacers* présents chez les espèces résistantes sont détaillées, les nouveaux *spacers* sont représentés par les rectangles blancs numérotés de S1 à S14. En bas à droite la sensibilité de chaque souche (mesurée par EOP ou *efficiency of plaquing*, c'est à dire le ratio entre souches mutantes et sauvages comptées) aux phages 858 et 2972 est représentée sous la forme d'un histogramme.

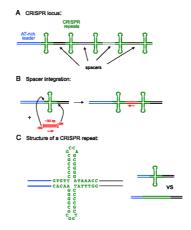
### Annexe 5:



Modèle de maturation du crARN impliquant le tracARN, la Rnase III et Csn1 (=Cas9)

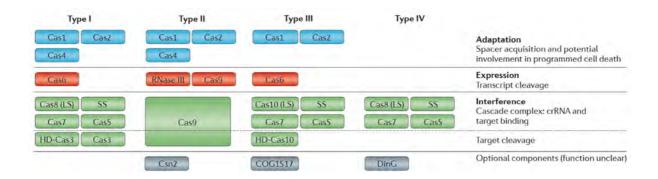
En noir : répétitions; en vert : *spacers*. Le tracrARN peut se lier avec une complémentarité quasi parfaite à chacune des séquences repeat du pre-crARN. Celà aboutit à la formation d'un duplex d'ARN qui pourra alors être découpé par la RnaseIII en présence de la protéine Cas9. En découle la libération de séquences *repeat-spacer-repeat* ( répétition-espaceur-répétition) individualisées. Ces éléments subissent alors de nouvelles modifications au niveau de la séquence *spacer* pour aboutir au crARN mature. Cet ARN matûré est constitué d'une séquence ne contenant qu'un seul exemplaire de répétition et d'espaceur. <sup>57</sup>

#### Annexe 6:



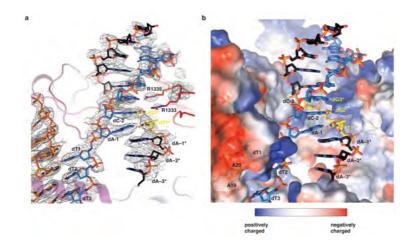
mécanisme d'acquisition des fragments *spacer*<sup>350</sup> **A**. les répétitions ou CRISPR (en vert) alternent avec les *spacers* (en dégradé de gris) de longueur similaire mais de séquences différentes. **B** intégration du protospacer par un mécanisme dit à double extrémité. Celà aboutit à la duplication de la répétition. **C**. séquence et structure cruciforme supposée du *repeat* CRISPR de 28nts chez *Escherichia coli*.

### Annexe 7



Classifications fonctionnelle systèmes des protéines Cas en 2011

#### Annexe 8:

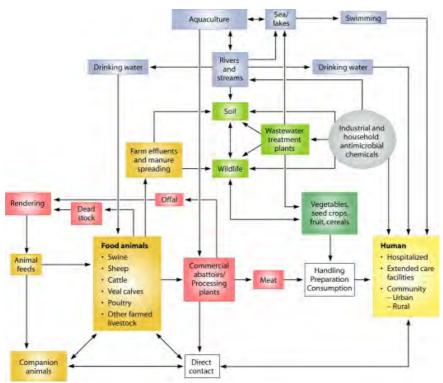


Détail du site de liaison au PAM chez la spCas9

Le duplex PAM se lie au sillon chargé positivement au niveau du domaine C terminal.

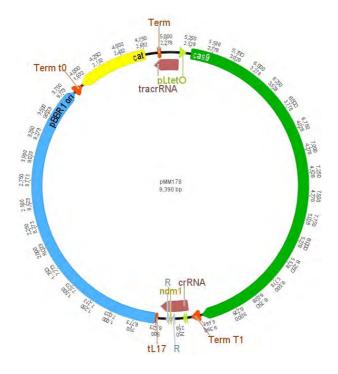
- **a.** Vue rapprochée du site de liaison au PAM De la protéine Cas9. Les acides nucléiques sont représentés sous formes de battonets et sont représentés recouvert d'une carte de densité electronique aplatie (maillage gris, deviation standard : 1**o**).
- **b.** Site de liaison au PAM, la surface moléculaire de Cas9 est colorée en fonction de son potentiel electrostatique. <sup>78</sup>

#### Annexe 9:



Dissémination des antibiotiques et de la résistance aux antibiotiques à travers le secteur agricole, les hôpitaux, le circuit de traitement des eaux usées, le réseau communautaire et les environnements associés.

### ANNEXE 10:



exemple de plasmide ciblant ndm 1 utilisé dans les expériences de l'équipe de Citorik<sup>231</sup> Le plasmide permet d'exprimer : la protéine Cas9, le tracrARN et un guide ARN ciblant le gène NMD-1. Il contiens une origine de réplication pBBR1 et une séquence de résistance au chloramphenicol.

Annexe 11:

Antibiotic class	Example(s)	Target	Mode(s) of resistance
β-Lactams	Penicillins (ampicillin), cephalosporins (cephamycin), penems (meropenem), monobactams (aztreonam)	Peptidoglycan biosynthesis	Hydrolysis, efflux, altered target
Aminoglycosides	Gentamicin, streptomycin, spectinomycin	Translation	Phosphorylation, acetylation, nucleotidylation, efflux, altered target
Glycopeptides	Vancomycin, teicoplanin	Peptidoglycan biosynthesis	Reprogramming peptidoglycan biosynthesis
Tetracyclines	Minocycline, tigecycline	Translation	Monooxygenation, efflux, altered target
Macrolides	Erythromycin, azithromicin	Translation	Hydrolysis, glycosylation, phosphorylation, efflux, altered target
Lincosamides	Clindamycin	Translation	Nucleotidylation, efflux, altered target
Streptogramins	Synercid	Translation	C-O lyase (type B streptogramins), acetylation (type A streptogramins) efflux, altered target
Oxazolidinones	Linezolid	Translation	Efflux, altered target
Phenicols	Chloramphenicol	Translation	Acetylation, efflux, altered target
Quinolones	Ciprofloxacin	DNA replication	Acetylation, efflux, altered target
Pyrimidines	Trimethoprim	C <sub>1</sub> metabolism	Efflux, altered target
Sulfonamides	Sulfamethoxazole	C <sub>1</sub> metabolism	Efflux, altered target
Rifamycins	Rifampin	Transcription	ADP-ribosylation, efflux, altered target
Lipopeptides	Daptomycin	Cell membrane	Altered target
Cationic peptides	Colistin	Cell membrane	Altered target, efflux

Mode d'action et résistances aux antibiotiques communs 222

Annexe 12 : communication entre l'Office of Agricultural Affairs et la Direction Générale de Santé à la commission Européenne au sujet des *New Plants Breeding Techniques* 

Ref. Ares(2015)4975884 - 10/11/2015

UNITED STATES MISSION

TO THE

EUROPEAN UNION

Office of Agricultural Affairs

Tel: (32)(2) , Fax: (32)(2)

November 3, 2015

Mr. Ladislav Miko Deputy Director-General Health and Food Safety Directorate-General (SANTE) European Commission Rue Breydel 4 B-1049 Brussels

Dear Ladislay,

When I met you both recently at your respective offices, I undertook to send you a letter to convey a few points regarding a set of tools for plant breeding, which the Commission refers to as "New Breeding Techniques" (NBTs).

You confirmed that the Commission intends to complete a legal analysis of the regulatory status of this set of NBTs by the end of this year. We would, of course, be interested to learn the details of the analysis when it is finalized. Within this context, we would also like to know how and to what extent EFSA's response of October 15, 2015 to DG SANTE's Biotechnology Unit's request for technical assistance on this matter will be used in formulating the analysis. More fundamentally, we would welcome your advice on whether or not the analysis will be legally binding on Member States.

Plant breeders in the EU and elsewhere wish to use these NBTs, as well as other precision tools under development in a rapidly changing field, as a means to increase the speed and precision of breeding new varieties that are beneficial to consumers and producers alike. The Commission's policy on regulating NBTs will clearly impact the degree to which this can be achieved not only in the EU but globally. Innovation in this area would be endangered if unjustified regulatory hurdles were to be implemented.

As you are aware, the government of the United States and governments of many other countries around the world are actively examining any potential risks associated with products developed using tools such as NBTs and considering appropriate oversight. As we know, different regulatory approaches between governments to NBT classification would lead to potentially significant trade disruptions. Finding commonalities in regulatory approaches in this area is of vital importance to promote trade, investment, research and innovation in plant breeding, and I would welcome further discussions in this regard.

Thank you for noting these points, and I look forward to receiving your response.

Yours sincerely,

#### Annexe 13:

« A method, in a eukaryotic cell, of cleaving or editing a target DNA molecule or modulating transcription of at least one gene encoded thereon, the method comprising:

contacting, in a eukaryotic cell, a target DNA molecule having a target sequence with an engineered and/or non-naturally-occurring Type II Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-CRISPR associated (Cas) (CRISPR-Cas) system comprising:

- a) a DNA-targeting RNA comprising
  - i) a targeter-RNA or guide sequence that hybridizes with the target sequence, and
- ii) an activator-RNA or tracr sequence that hybridizes with the targeter-RNA to form a double-stranded RNA duplex of a protein-binding segment, and
  - b) a Cas9 protein,

wherein the DNA-targeting RNA forms a complex with the Cas9 protein, thereby targeting the Cas9 protein to the target DNA molecule, whereby said target DNA molecule is cleaved or edited or transcription of at least one gene encoded by the target DNA molecule is modulated »

Description de l'objet de la demande d'interférence No. 106,048 (DK) opposant UC Berkeley et le Broad Institute pour la propriété du premier brevet englobant la technique CRISPR-Cas9 328

#### Annexe 14:

Industry sector	Product/application	Company	Intellectual property
Food	Yogurt, cheese Crops Livestock Crops	Danisco (DuPont) Dow Agrosciences  Recombinetics Cellectis Plant Sciences	7 919 277; 8 361 725; 13/722 539;11/990 885 PCT/US2013/039979 co-owned with Sangamo Biosciences PCT/US2014/0201857 Boston Children's Hosp., Institut Pasteur License
Laboratory	Research tools Expression systems Research tools Animal models Research tools Animal models	System Biosciences Sigma-Aldrich GE Healthcare Sage ThermoFisher Taconic	US 14/216 655 PCT/US2013/073307 Broad License Caribou, Broad License Cellectis Sublicense Broad License
Sublicensing	Ag, Industrial, Bio	Caribou	PCT/US2013/053287
Medical	Pharmaceuticals In vitro applications only Target validation Therapeutics Monogenic diseases Therapeutics	Novartis Cellectis AstraZeneca Crispr Therapeutics Sangamo Biosciences	Caribou License Boston Children's Hosp., Institut Pasteur License Open Innovation Model PCT/US2013/032559 PCT/US2013/032381; PCT/US2013/039979; PCT/US2013/028348 Caribou License
	Therapeutics	Editas	Broad, Duke, MGH Licenses

CRISPR-Cas9 : Propriété intellectuelle, licences et industrie <sup>3</sup>

19

#### V. CONCLUSIONS

We believe that considerable benefits are likely to result from experiments involving the genetic alteration of microorganisms. The range of possible benefits extends from the use of these techniques to add to our knowledge of basic biological phenomena, to possible practical applications in the areas of agriculture and medicine.

We believe also that a scale of risks exists in the construction of genetically altered microorganisms, and we are uncomfortable about our inability to assess precisely the extent of such risks for many types of experiments. However, we believe that the containment procedures described in this proposal will reduce any risk to laboratory workers and to the environment to a level that is acceptably low and which will allow investigators to carry out research in this area. We believe that certain experiments should presently not be carried out under any circumstances (i.e. Class VI), but that most experiments can be done if containment facilities appropriate to the risk are utilized.

We recommend that specific steps be taken as soon as possible to develop cloning vehicle-host systems which will further reduce biohazard potential, will minimize the necessity of elaborate containment facilities, and will obviate judgements which must necessarily be based on little or no data at the present time. Specifically, we recommend that special sponsored programs be instituted immediately for the development and testing of such systems. We recommend also the prompt establishment of experimental programs intended to evaluate more fully the potential hazards that may be involved in the genetic alteration of microorganisms.

We believe that perhaps the greatest potential for biohazards involving genetic alteration of microorganisms relates to possible military applications. We believe strongly that construction of genetically altered microorganisms for any military purpose should be expressly prohibited by international treaty, and we urge that such prohibition be agreed upon as expeditiously as possible.

Other recommendations for implementation of the guidelines proposed in this report are contained in Section  ${\tt IV}$ .

# **BIBLIOGRAPHIE**

- 1. Mukherjee, S. *The Gene: An Intimate History.* (Random House Children's Books, 2016).
- 2. Einstein, A. Living Philosophies. (Simon and Schuster, 1931).
- 3. van Erp, P. B., Bloomer, G., Wilkinson, R. & Wiedenheft, B. The history and market impact of CRISPR RNA-guided nucleases. *Curr. Opin. Virol.* **12**, 85–90 (2015).
- 4. Heberle-Bors, E. Génie génétique: une histoire, un défi. (Editions Quae, 2001).
- 5. Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).
- 6. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467 (1977).
- 7. Maxam, A. M. & Gilbert, W. A new method for sequencing DNA. *Proc. Natl. Acad. Sci.* 74, 560–564 (1977).
- 8. Mullis, K. B. Process for amplifying nucleic acid sequences. (1987).
- CultureLab: Atomic gardening: Day of the irradiated peanuts. Available at: https://www.newscientist.com/blogs/culturelab/2011/06/atomic-gardening-day-of-the-irradiated-peanuts.html. (Accessed: 27th November 2016)
- 10. Weinbauer, M. G. Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* **28,** 127–181 (2004).
- 11. Hatfull, G. F. Bacteriophage genomics. Curr. Opin. Microbiol. 11, 447–453 (2008).
- 12. Barrangou, R. *et al.* CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* **315**, 1709–1712 (2007).
- 13. Capecchi, M. R. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat. Rev. Genet.* **6,** 507–512 (2005).
- 14. Jasin, M. & Rothstein, R. Repair of strand breaks by homologous recombination. *Cold Spring Harb. Perspect. Biol.* **5**, a012740 (2013).
- 15. Bibikova, M. *et al.* Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol. Cell. Biol.* **21,** 289–297 (2001).

- 16. Porteus, M. H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**, 763 (2003).
- 17. Bibikova, M., Beumer, K., Trautman, J. K. & Carroll, D. Enhancing gene targeting with designed zinc finger nucleases. *Science* **300**, 764 (2003).
- 18. Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**, 636–646 (2010).
- 19. Moscou, M. J. & Bogdanove, A. J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501 (2009).
- 20. Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512 (2009).
- 21. Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–148 (2011).
- 22. Zhang, F. *et al.* Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* **29**, 149–153 (2011).
- 23. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. 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 (1987).
- Mojica, F. J., Juez, G. & Rodríguez-Valera, F. Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. *Mol. Microbiol.* 9, 613–621 (1993).
- 25. FRANCISCO JUAN MARTÍNEZ MOJICA. Available at: https://imem.ua.es/en/about-us/francisco-juan-martinez-mojica.html. (Accessed: 28th November 2016)
- 26. Nakata, A., Amemura, M. & Makino, K. Unusual nucleotide arrangement with repeated sequences in the Escherichia coli K-12 chromosome. *J. Bacteriol.* **171,** 3553–3556 (1989).
- 27. Mojica, F. j. m., Ferrer, C., Juez, G. & Rodríguez-Valera, F. Long stretches of short tandem repeats are present in the largest replicons of the Archaea Haloferax mediterranei and Haloferax volcanii and could be involved in replicon partitioning. *Mol. Microbiol.* 17, 85–93 (1995).

- 28. Herman, P. K., Stack, J. H. & Emr, S. D. A genetic and structural analysis of the yeast Vps15 protein kinase: evidence for a direct role of Vps15p in vacuolar protein delivery. *EMBO J.* **10**, 4049–4060 (1991).
- 29. Bult, C. J. *et al.* Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii. *Science* **273**, 1058–1073 (1996).
- 30. Nelson, K. E. *et al.* Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of Thermotoga maritima. *Nature* **399**, 323–329 (1999).
- 31. Flamand, M. C., Goblet, J. P., Duc, G., Briquet, M. & Boutry, M. Sequence and transcription analysis of mitochondrial plasmids isolated from cytoplasmic male-sterile lines of Vicia faba. *Plant Mol. Biol.* **19**, 913–923 (1992).
- 32. Gottesman, S. Microbiology: Dicing defence in bacteria. Nature 471, 588-589 (2011).
- 33. Jansen, R., van Embden, J. D. A., Gaastra, W. & Schouls, L. M. Identification of a novel family of sequence repeats among prokaryotes. *Omics J. Integr. Biol.* **6,** 23–33 (2002).
- 34. Barrangou, R. & Oost, J. van der. *CRISPR-Cas Systems: RNA-mediated Adaptive Immunity in Bacteria and Archaea*. (Springer Science & Business Media, 2012).
- 35. Mojica, F. J., Díez-Villaseñor, C., Soria, E. & Juez, G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* **36**, 244–246 (2000).
- 36. She, Q. *et al.* The complete genome of the crenarchaeon Sulfolobus solfataricus P2. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 7835–7840 (2001).
- 37. MEDLINE®/PubMed® Resources Guide. Available at: https://www.nlm.nih.gov/bsd/pmresources.html. (Accessed: 28th November 2016)
- 38. Ehrlich, S. D., Bolotin, A., Quinquis, B. & Sorokin, A. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin.

  \*Microbiology 151, 2551–2561 (2005).
- 39. Pourcel, C., Salvignol, G. & Vergnaud, G. 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 (2005).
- 40. Jansen, R., Embden, J. D. A. van, Gaastra, W. & Schouls, L. M. Identification of genes

- that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43,** 1565–1575 (2002).
- 41. Dr. Ruud Jansen. Available at: /node/302. (Accessed: 1st December 2016)
- 42. Haft, D. H., Selengut, J., Mongodin, E. F. & Nelson, K. E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **1**, e60 (2005).
- 43. BLAST: Basic Local Alignment Search Tool. Available at: https://blast.ncbi.nlm.nih.gov/Blast.cgi. (Accessed: 30th November 2016)
- 44. Mojica, F. J. M., Diez-Villasenor, C., Garcia-Martinez, J. & Soria, E. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *J. Mol. Evol.* **60**, 174–182 (2005).
- 45. Deveau, H. *et al.* Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. *J. Bacteriol.* **190**, 1390–1400 (2008).
- 46. The Doudna Lab | CRISPR research. Available at: http://rna.berkeley.edu/crispr.html. (Accessed: 28th November 2016)
- 47. There's CRISPR in Your Yogurt | The Scientist Magazine®. *The Scientist* Available at: http://www.the-scientist.com/?articles.view/articleNo/41676/title/There-s-CRISPR-in-Your-Yogurt/. (Accessed: 2nd December 2016)
- 48. Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* 1, 7 (2006).
- Makarova, K. S., Aravind, L., Grishin, N. V., Rogozin, I. B. & Koonin, E. V. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res.* 30, 482–496 (2002).
- 50. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008).
- 51. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**, 1843–1845 (2008).

- 52. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiol. Read. Engl.* **155**, 733–740 (2009).
- 53. The quiet revolutionary: How the co-discovery of CRISPR explosively changed Emmanuelle Charpentier's life. Alison Abbott . Nature News.
- 54. Carte, J., Wang, R., Li, H., Terns, R. M. & Terns, M. P. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **22**, 3489–3496 (2008).
- 55. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci.* **109**, E2579–E2586 (2012).
- 56. Jinek, M. *et al.* A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
- 57. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (2011).
- 58. Emmanuelle Charpentier. Available at: https://www.mpg.de/10729312/emmanuelle-charpentier. (Accessed: 7th December 2016)
- 59. Deveau, H., Garneau, J. E. & Moineau, S. CRISPR/Cas system and its role in phage-bacteria interactions. *Annu. Rev. Microbiol.* **64**, 475–493 (2010).
- 60. Al-Attar, S., Westra, E. R., van der Oost, J. & Brouns, S. J. J. Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biol. Chem.* **392**, 277–289 (2011).
- 61. van der Oost, J., Jore, M. M., Westra, E. R., Lundgren, M. & Brouns, S. J. J. CRISPRbased adaptive and heritable immunity in prokaryotes. *Trends Biochem. Sci.* **34**, 401–407 (2009).
- 62. Sontheimer, E. J. & Marraffini, L. A. Microbiology: slicer for DNA. *Nature* **468**, 45–46 (2010).
- 63. Carte, J., Pfister, N. T., Compton, M. M., Terns, R. M. & Terns, M. P. Binding and cleavage of CRISPR RNA by Cas6. *RNA* 16, 2181–2188 (2010).

- 64. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J. A. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329**, 1355–1358 (2010).
- 65. Wang, R., Preamplume, G., Terns, M. P., Terns, R. M. & Li, H. Interaction of the Cas6 riboendonuclease with CRISPR RNAs: recognition and cleavage. *Struct. Lond. Engl.* 1993 19, 257–264 (2011).
- 66. Hale, C. R. *et al.* RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**, 945–956 (2009).
- 67. Haber, J. E. Partners and pathways: repairing a double-strand break. *Trends Genet.* **16**, 259–264 (2000).
- 68. Makarova, K. S. *et al.* Evolution and classification of the CRISPR–Cas systems. *Nat. Rev. Microbiol.* **9,** 467–477 (2011).
- 69. Sapranauskas, R. *et al.* The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic Acids Res.* **39**, 9275–9282 (2011).
- 70. Yarris, L. Programmable DNA Scissors Found for Bacterial Immune System | Berkeley Lab. News Center (2012). Available at: http://newscenter.lbl.gov/2012/06/28/programmable-dna-scissors/. (Accessed: 13th December 2016)
- 71. Jinek, M. *et al.* Methods and compositions for rna-directed target dna modification and for rna-directed modulation of transcription. (2013).
- 72. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819–823 (2013).
- 73. Mali, P. et al. RNA-Guided Human Genome Engineering via Cas9. Science **339**, 823–826 (2013).
- 74. Jinek, M. *et al.* RNA-programmed genome editing in human cells. *Elife* **2**, e00471 (2013).
- 75. Cho, S. W., Kim, S., Kim, J. M. & Kim, J.-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**, 230–232 (2013).
- 76. Jinek, M. et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational

- activation. Science 343, 1247997 (2014).
- 77. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588 (2014).
- 78. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569–573 (2014).
- 79. Siksnys, V., GASIÜNAS, G. & Karvelis, T. RNA-DIRECTED DNA CLEAVAGE BY THE Cas9-crRNA COMPLEX. (2015).
- 80. Sontheimer, E. J. & Marraffini, L. A. TARGET DNA INTERFERENCE WITH crRNA. (2010).
- 81. Kim, J., Cho, S. W. & Kim, S. Composition for cleaving a target dna comprising a guide rna specific for the target dna and cas protein-encoding nucleic acid or cas protein, and use thereof. (2015).
- 82. Zhang, F. CRISPR-Cas systems and methods for altering expression of gene products. (2014).
- 83. School, S. L. The CRISPR Patent Interference Showdown Is On: How Did We Get Here and What Comes Next? *Stanford Law School* Available at: https://law.stanford.edu/2015/12/29/the-crispr-patent-interference-showdown-is-on-how-did-we-get-here-and-what-comes-next/. (Accessed: 21st September 2016)
- 84. Press Release | Investors & Media | Editas Medicine. Available at: http://ir.editasmedicine.com/phoenix.zhtml?c=254265&p=irolnewsArticle&ID=2125226. (Accessed: 18th January 2017)
- 85. This Gene-Editing Technology Will Change the World. But Who Gets the Credit? Bloomberg.com
- 86. Hsu, P. D., Lander, E. S. & Zhang, F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **157**, 1262–1278 (2014).
- 87. Google Trends. *Google Trends* Available at: /trends/explore. (Accessed: 6th October 2016)
- 88. Maddalo, D. *et al.* In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature* **516**, 423–427 (2014).

- 89. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* doi:10.1038/nbt.2969
- 90. Yin, H. *et al.* Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat. Biotechnol.* **32**, 551–553 (2014).
- 91. Koonin, E. V. & Krupovic, M. Evolution of adaptive immunity from transposable elements combined with innate immune systems. *Nat. Rev. Genet.* **16,** 184–192 (2015).
- 92. Ran, F. A. In vivo genome editing using Staphylococcus aureus Cas9. *Nature* **520,** 186–191 (2015).
- 93. Synthetic Guide RNA for CRISPR Genome Editing | Synthego.
- 94. Horvath, P. *et al.* Diversity, Activity, and Evolution of CRISPR Loci in Streptococcus thermophilus. *J. Bacteriol.* **190,** 1401–1412 (2008).
- 95. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).
- 96. Szczelkun, M. D. *et al.* Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 9798–9803 (2014).
- 97. Sternberg, S. H., LaFrance, B., Kaplan, M. & Doudna, J. A. Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature* **527**, 110–113 (2015).
- 98. Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, W529-533 (2010).
- 99. Nishimasu, H. Crystal structure of Staphylococcus aureus Cas9. *Cell* **162**, 1113–1126 (2015).
- 100. CRISPR/Cas9 Mini Model. Available at: http://www.3dmoleculardesigns.com/Education-Products/CRISPR-Cas9-Mini-Model.htm. (Accessed: 13th December 2016)
- 101. 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. Biotechnol.* **32**,

- 677-683 (2014).
- 102. Wu, X. *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* **32**, 670–676 (2014).
- 103. Gaj, T., Gersbach, C. A. & Barbas, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31,** 397–405 (2013).
- 104. Shalem, O. *et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science* **343**, 84–87 (2014).
- 105. Maruyama, T. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol* **33**, 538–542 (2015).
- 106. Suzuki, K. *et al.* In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **540**, 144–149 (2016).
- 107. Yu, C. Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* **16**, 142–147 (2015).
- 108. Decottignies, A. Alternative end-joining mechanisms: a historical perspective. *Front Genet* **4**, 48 (2013).
- 109. Lenhart, J. S., Schroeder, J. W., Walsh, B. W. & Simmons, L. A. DNA repair and genome maintenance in Bacillus subtilis. *Microbiol Mol Biol Rev* **76**, 530–564 (2012).
- 110. Bennardo, N., Cheng, A., Huang, N. & Stark, J. M. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet* 4, e1000110 (2008).
- 111. Sakuma, T., Nishikawa, A., Kume, S., Chayama, K. & Yamamoto, T. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci. Rep.* 4, (2014).
- 112. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096–1258096 (2014).
- 113. Mali, P. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* **31**, 833–838 (2013).
- 114. Ran, F. A. *et al.* Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell* **154**, 1380–1389 (2013).

- 115. Shen, B. *et al.* Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat. Methods* **11**, 399–402 (2014).
- 116. Hilton, I. B. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol* **33**, 510–517 (2015).
- 117. Kearns, N. A. *et al.* Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat. Methods* **12**, 401–403 (2015).
- 118. Chen, B. *et al.* Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491 (2013).
- 119. Ma, H. *et al.* Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat. Biotechnol.* **34,** 528–530 (2016).
- 120. Fujita, T. & Fujii, H. Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. *Biochem. Biophys. Res. Commun.* 439, 132–136 (2013).
- 121. Konermann, S. *et al.* Optical control of mammalian endogenous transcription and epigenetic states. *Nature* **500**, 472–476 (2013).
- 122. Kennedy, M. J. *et al.* Rapid blue light induction of protein interactions in living cells. *Nat. Methods* **7,** 973–975 (2010).
- 123. Briner, A. E. Guide RNA functional modules direct Cas9 activity and orthogonality. *Mol Cell* **56**, 333–339 (2014).
- 124. Esvelt, K. M. *et al.* Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* **10**, 1116–1121 (2013).
- 125. Fonfara, I. *et al.* Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res.* **42**, 2577–2590 (2014).
- 126. Dahlman, J. E. *et al.* Orthogonal gene knock out and activation with a catalytically active Cas9 nuclease. *Nat. Biotechnol.* **33**, 1159–1161 (2015).
- 127. El-Showk, S. Vision restoration in rats gives glimpse of gene therapy potential. *Nat. Middle East* (2016). doi:10.1038/nmiddleeast.2016.207

- 128. Yamano, T. *et al.* Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell* **165**, 949–962 (2016).
- 129. Dong, D. The crystal structure of Cpf1 in complex with CRISPR RNA. *Nature* **532**, 522–526 (2016).
- 130. Shmakov, S. Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Mol Cell* **60**, 385–397 (2015).
- 131. Abudayyeh, O. O. *et al.* C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573 (2016).
- 132. Zetsche, B. *et al.* Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).
- 133. Quilt Data. *Quilt Data* Available at: https://quiltdata.com/grna-search/. (Accessed: 14th December 2016)
- 134. sgRNA Scorer 1.0. Available at: https://crispr.med.harvard.edu/sgRNAScorer/.

  (Accessed: 14th December 2016)
- 135. CRISPR RGEN Tools. Available at: http://www.rgenome.net/cas-designer/. (Accessed: 15th December 2016)
- 136. Off-Spotter. Available at: https://cm.jefferson.edu/Off-Spotter/. (Accessed: 15th December 2016)
- 137. Addgene: CRISPR References and Information. Available at: https://www.addgene.org/crispr/reference/#protocols. (Accessed: 15th December 2016)
- 138. gBlocks Gene Fragments. Available at: http://www.idtdna.com/pages/products/genes/gblocks-gene-fragments. (Accessed: 15th December 2016)
- 139. Zero Blunt TOPO PCR Cloning Kit, with One Shot TOP10 Chemically Competent E. coli cells. Available at: http://www.thermofisher.com/order/catalog/product/K280020. (Accessed: 15th December 2016)
- 140. GeneArt Platinum Cas9 Nuclease (1 g/L) Thermo Fisher Scientific. Available at: https://www.thermofisher.com/order/catalog/product/B25642. (Accessed: 26th January 2017)

- 141. D'Astolfo, D. S. *et al.* Efficient Intracellular Delivery of Native Proteins. *Cell* **161,** 674–690 (2015).
- 142. Zuris, J. A. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat Biotechnol* **33**, 73–80 (2015).
- 143. Beisel, C. L., Gomaa, A. A. & Barrangou, R. A CRISPR design for next-generation antimicrobials. *Genome Biol.* **15,** 516 (2014).
- 144. Howes, R. & Schofield, C. Genome engineering using adeno-associated virus (AAV). *Methods Mol Biol* **1239**, 75–103 (2015).
- 145. Lin, S., Staahl, B. T., Alla, R. K. & Doudna, J. A. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* **3**, e04766 (2014).
- 146. Schumann, K. Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proc Natl Acad Sci USA* **112**, 10437–10442 (2015).
- 147. Hendriks, W. T., Jiang, X., Daheron, L. & Cowan, C. A. TALEN- and CRISPR/Cas9-mediated gene editing in human pluripotent stem cells using lipid-based transfection. *Curr Protoc Stem Cell Biol* 34, 5B 3.1-5B 3.25 (2015).
- 148. Friedmann, T. & Roblin, R. Gene therapy for human genetic disease? *Science* **175**, 949–955 (1972).
- 149. Check, E. Gene therapy: A tragic setback. *Nature* **420**, 116–118 (2002).
- 150. Gaspar, H. B. *et al.* Long-term persistence of a polyclonal T cell repertoire after gene therapy for X-linked severe combined immunodeficiency. *Sci. Transl. Med.* **3**, 97ra79 (2011).
- 151. Howe, S. J. *et al.* Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J. Clin. Invest.* **118**, 3143–3150 (2008).
- 152. Aiuti, A. *et al.* Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science* **341**, 1233151 (2013).
- 153. Biffi, A. *et al.* Lentiviral Hematopoietic Stem Cell Gene Therapy Benefits Metachromatic Leukodystrophy. *Science* **341**, 1233158 (2013).

- 154. European Medicines Agency Find medicine Glybera. Available at: http://www.ema.europa.eu/ema/index.jsp? curl=pages/medicines/human/medicines/002145/human\_med\_001480.jsp&mid=WC0b0 1ac058001d124. (Accessed: 23rd December 2016)
- 155. European Medicines Agency Find medicine Strimvelis. Available at: http://www.ema.europa.eu/ema/index.jsp? curl=pages/medicines/human/medicines/003854/human\_med\_001985.jsp&mid=WC0b0 1ac058001d124. (Accessed: 23rd December 2016)
- 156. Vannucci, L., Lai, M., Chiuppesi, F., Ceccherini-Nelli, L. & Pistello, M. Viral vectors: a look back and ahead on gene transfer technology. *New Microbiol.* **36,** 1–22 (2013).
- 157. OMIM Gene Map Statistics. Available at: http://www.omim.org/statistics/geneMap. (Accessed: 22nd December 2016)
- 158. All About The Human Genome Project (HGP). *National Human Genome Research Institute (NHGRI)* Available at: https://www.genome.gov/10001772/All-About-The-Human-Genome-Project-HGP. (Accessed: 22nd December 2016)
- 159. Schwank, G. *et al.* Functional Repair of CFTR by CRISPR/Cas9 in Intestinal Stem Cell Organoids of Cystic Fibrosis Patients. *Cell Stem Cell* **13**, 653–658 (2013).
- 160. Firth, A. L. Functional gene correction for cystic fibrosis in lung epithelial cells generated from patient iPSCs. *Cell Rep* **12**, 1385–1390 (2015).
- 161. Regalado, A. Scientists Call for Moratorium on Genetically Engineered Babies. MIT Technology Review Available at: https://www.technologyreview.com/s/536021/scientists-call-for-a-summit-on-gene-edited-babies/. (Accessed: 27th December 2016)
- 162. Regalado, A. Group Urges Halt on Editing Human DNA. *MIT Technology Review*Available at: https://www.technologyreview.com/s/535846/industry-body-calls-for-geneediting-moratorium/. (Accessed: 27th December 2016)
- 163. Liang, P. *et al.* CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. *Protein Cell* **6**, 363–372 (2015).
- 164. Li, H. L. et al. Precise Correction of the Dystrophin Gene in Duchenne Muscular Dystrophy Patient Induced Pluripotent Stem Cells by TALEN and CRISPR-Cas9. Stem

- Cell Rep. 4, 143–154 (2015).
- 165. Mendell, J. R. & Rodino-Klapac, L. R. Duchenne muscular dystrophy: CRISPR/Cas9 treatment. *Cell Res.* **26**, 513–514 (2016).
- 166. Nelson, C. E. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* **351**, 403–407 (2016).
- 167. Ousterout, D. G. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nat Commun* 6, 6244 (2015).
- 168. Long, C. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* **351**, 400–403 (2016).
- 169. Osborn, M. J. Fanconi anemia gene editing by the CRISPR/Cas9 system. *Hum Gene Ther* **26**, 114–126 (2015).
- 170. Malkki, H. Huntington disease: Selective deactivation of Huntington disease mutant allele by CRISPR-Cas9 gene editing. *Nat. Rev. Neurol.* **12**, 614–615 (2016).
- 171. Bassuk, A. G., Zheng, A., Li, Y., Tsang, S. H. & Mahajan, V. B. Precision Medicine: Genetic Repair of Retinitis Pigmentosa in Patient-Derived Stem Cells. *Sci. Rep.* **6**, 19969 (2016).
- 172. Smith, C. *et al.* Efficient and allele-specific genome editing of disease loci in human iPSCs. *Mol. Ther. J. Am. Soc. Gene Ther.* **23**, 570–577 (2015).
- 173. DeWitt, M. A. *et al.* Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. *Sci. Transl. Med.* **8,** 360ra134-360ra134 (2016).
- 174. Penn Scientists Use CRISPR for First Time to Correct Clotting in Newborn and Adult Mice PR News. Available at: https://www.pennmedicine.org/news/news-releases/2016/november/scientists-use-crispr-for-first-time-to-correct-clotting-in-newborn-and-adult-mice. (Accessed: 29th December 2016)
- Lin, G., Zhang, K. & Li, J. Application of CRISPR/Cas9 Technology to HBV. *Int. J. Mol. Sci.* 16, 26077–26086 (2015).
- 176. Ganem, D. & Prince, A. M. Hepatitis B virus infection--natural history and clinical consequences. *N. Engl. J. Med.* **350**, 1118–1129 (2004).

- 177. Lin, S.-R. *et al.* The CRISPR/Cas9 System Facilitates Clearance of the Intrahepatic HBV Templates In Vivo. *Mol. Ther. Nucleic Acids* **3**, e186 (2014).
- 178. Mason, A., Wick, M., White, H. & Perrillo, R. Hepatitis B virus replication in diverse cell types during chronic hepatitis B virus infection. *Hepatol. Baltim. Md* **18**, 781–789 (1993).
- 179. Lisowski, L. *et al.* Selection and evaluation of clinically relevant AAV variants in a xenograft liver model. *Nature* **506**, 382–386 (2014).
- 180. Dandri, M. & Lütgehetmann, M. Mouse models of hepatitis B and delta virus infection. *J. Immunol. Methods* **410**, 39–49 (2014).
- 181. Seeger, C. & Sohn, J. A. Complete Spectrum of CRISPR/Cas9-induced Mutations on HBV cccDNA. *Mol. Ther.* **24**, 1258–1266 (2016).
- 182. Feng, W., Hong, G., Delecluse, H.-J. & Kenney, S. C. Lytic induction therapy for Epstein-Barr virus-positive B-cell lymphomas. *J. Virol.* **78**, 1893–1902 (2004).
- 183. Diemen, F. R. van *et al.* CRISPR/Cas9-Mediated Genome Editing of Herpesviruses Limits Productive and Latent Infections. *PLOS Pathog* **12**, e1005701 (2016).
- 184. Nicoll, M. P., Proença, J. T. & Efstathiou, S. The molecular basis of herpes simplex virus latency. *FEMS Microbiol. Rev.* **36**, 684–705 (2012).
- 185. Engelman, A. & Cherepanov, P. The structural biology of HIV-1: mechanistic and therapeutic insights. *Nat. Rev. Microbiol.* **10,** 279–290 (2012).
- 186. Chun, T. W. *et al.* Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 13193–13197 (1997).
- 187. Xing, S. & Siliciano, R. F. Targeting HIV latency: pharmacologic strategies toward eradication. *Drug Discov. Today* **18**, 541–551 (2013).
- 188. Tebas, P. *et al.* Gene Editing of *CCR5* in Autologous CD4 T Cells of Persons Infected with HIV. *N. Engl. J. Med.* **370**, 901–910 (2014).
- 189. Pollack, A. New Hope of a Cure for H.I.V. The New York Times (2011).
- 190. Cradick, T. J., Fine, E. J., Antico, C. J. & Bao, G. CRISPR/Cas9 systems targeting -globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* **41**, 9584–9592 (2013).

- 191. Yang, H. *et al.* One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering. *Cell* **154**, 1370–1379 (2013).
- 192. Burke, B. P. *et al.* CCR5 as a Natural and Modulated Target for Inhibition of HIV. *Viruses* **6**, 54–68 (2013).
- 193. Wayengera, M. Proviral HIV-genome-wide and pol-gene specific zinc finger nucleases: usability for targeted HIV gene therapy. *Theor. Biol. Med. Model.* **8,** 26 (2011).
- 194. Aubert, M. *et al.* Successful targeting and disruption of an integrated reporter lentivirus using the engineered homing endonuclease Y2 I-AniI. *PloS One* **6**, e16825 (2011).
- 195. Ebina, H., Misawa, N., Kanemura, Y. & Koyanagi, Y. Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci. Rep.* **3**, (2013).
- 196. Hu, W. *et al.* RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proc. Natl. Acad. Sci.* **111**, 11461–11466 (2014).
- 197. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* **31**, 839–843 (2013).
- 198. Gabriel, R. *et al.* An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* **29**, 816–823 (2011).
- 199. Wu, Y. et al. Correction of a Genetic Disease in Mouse via Use of CRISPR-Cas9. *Cell Stem Cell* **13**, 659–662 (2013).
- 200. Zhou, Y. *et al.* High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* **509**, 487–491 (2014).
- 201. Platt, R. J. *et al.* CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. *Cell* **159**, 440–455 (2014).
- 202. Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M. D. C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat. Biotechnol.* **32**, 267–273 (2014).
- 203. Wang, G., Zhao, N., Berkhout, B. & Das, A. T. CRISPR-Cas9 Can Inhibit HIV-1 Replication but NHEJ Repair Facilitates Virus Escape. *Mol. Ther. J. Am. Soc. Gene Ther.* 24, 522–526 (2016).
- 204. Wang, Z. et al. CRISPR/Cas9-Derived Mutations Both Inhibit HIV-1 Replication and

- Accelerate Viral Escape. Cell Rep. 15, 481–489 (2016).
- 205. ter Brake, O., Konstantinova, P., Ceylan, M. & Berkhout, B. Silencing of HIV-1 with RNA interference: a multiple shRNA approach. *Mol. Ther. J. Am. Soc. Gene Ther.* **14**, 883–892 (2006).
- 206. Das, A. T. *et al.* Human immunodeficiency virus type 1 escapes from RNA interference-mediated inhibition. *J. Virol.* **78**, 2601–2605 (2004).
- 207. Reardon, S. Leukaemia success heralds wave of gene-editing therapies. *Nat. News* **527**, 146 (2015).
- 208. Galanis, E. & Russell, S. Cancer gene therapy clinical trials: lessons for the future. *Br. J. Cancer* **85**, 1432–1436 (2001).
- 209. The Journal of Gene Medicine Wiley Online Library. Available at: http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1521-2254. (Accessed: 27th December 2016)
- 210. Sato, Y. *et al.* Resolution of liver cirrhosis using vitamin A-coupled liposomes to deliver siRNA against a collagen-specific chaperone. *Nat. Biotechnol.* **26,** 431–442 (2008).
- 211. Pan, J. *et al.* Effect of recombinant adenovirus-p53 combined with radiotherapy on long-term prognosis of advanced nasopharyngeal carcinoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **27,** 799–804 (2009).
- 212. Greenberger, J. S. Radioprotection. Vivo Athens Greece 23, 323–336 (2009).
- 213. UCART19: First in Man Proof of Concept to be Presented at 2015 ASH Annual Meeting | cellectis. Available at: https://www.cellectis.com/en/content/ucart19-first-man-proof-concept-be-presented-2015-ash-annual-meeting-0. (Accessed: 29th December 2016)
- 214. Cyranoski, D. CRISPR gene-editing tested in a person for the first time. *Nat. News* **539**, 479 (2016).
- 215. Niu, Y. *et al.* Generation of Gene-Modified Cynomolgus Monkey via Cas9/RNA-Mediated Gene Targeting in One-Cell Embryos. *Cell* **156**, 836–843 (2014).
- 216. Haute Autorité de Santé KEYTRUDA (pembrolizumab), anticorps anti-PD1. Available at: http://www.has-sante.fr/portail/jcms/c\_2639673/fr/keytruda-pembrolizumab-anticorps-anti-pd1. (Accessed: 29th December 2016)

- 217. Haute Autorité de Santé OPDIVO (nivolumab), anticorps anti-PD1. Available at: http://www.has-sante.fr/portail/jcms/c\_2612055/fr/opdivo-nivolumab-anticorps-anti-pd1. (Accessed: 29th December 2016)
- 218. Metcalfe, W., Anderson, J., Trinh, V. & Hwu, W.-J. Anti-Programmed Cell Death-1 (PD-1) Monoclonal Antibodies in Treating Advanced Melanoma. *Discov. Med.* **19**, 393–401
- 219. Su, S. *et al.* CRISPR-Cas9 mediated efficient PD-1 disruption on human primary T cells from cancer patients. *Sci. Rep.* **6**, 20070 (2016).
- 220. Cyranoski, D. Chinese scientists to pioneer first human CRISPR trial. *Nat. News* **535**, 476 (2016).
- 221. Barker, R. W., Brindley, D. A. & Schuh, A. Establish good genomic practice to guide medicine forward. *Nat. Med.* **19**, 530–530 (2013).
- 222. Davies, J. & Davies, D. Origins and Evolution of Antibiotic Resistance. *Microbiol. Mol. Biol. Rev. MMBR* **74**, 417–433 (2010).
- 223. WHO | WHO Director-General addresses UN General Assembly on antimicrobial resistance. *WHO* Available at: http://www.who.int/dg/speeches/2016/unga-antimicrobial-resistance/en/. (Accessed: 1st November 2016)
- 224. Witte, W. Selective pressure by antibiotic use in livestock. *Int. J. Antimicrob. Agents* **16**, 19–24 (2000).
- 225. Zhu, Y.-G. *et al.* Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proc. Natl. Acad. Sci.* **110**, 3435–3440 (2013).
- 226. Landers, T. F., Cohen, B., Wittum, T. E. & Larson, E. L. A Review of Antibiotic Use in Food Animals: Perspective, Policy, and Potential. *Public Health Rep.* **127**, 4–22 (2012).
- 227. Commission Européenne COMMUNIQUES DE PRESSE Communiqué de presse Ban on antibiotics as growth promoters in animal feed enters into effect. Available at: http://europa.eu/rapid/press-release\_IP-05-1687\_en.htm. (Accessed: 4th November 2016)
- 228. Liu, Y.-Y. *et al.* Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect. Dis.* **16**, 161–168 (2016).

- 229. Kirby, T. Europe to boost development of new antimicrobial drugs. *The Lancet* **379**, 2229–2230 (2012).
- 230. Bikard, D. *et al.* Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nat. Biotechnol.* **32**, 1146–1150 (2014).
- 231. Citorik, R. J., Mimee, M. & Lu, T. K. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nat Biotechnol* **32**, 1141–1145 (2014).
- 232. Rasheed, J. K. *et al.* Characterization of the extended-spectrum beta-lactamase reference strain, Klebsiella pneumoniae K6 (ATCC 700603), which produces the novel enzyme SHV-18. *Antimicrob. Agents Chemother.* **44**, 2382–2388 (2000).
- 233. Melnikov, A. A., Tchernov, A. P., Fodor, I. & Bayev, A. A. Lambda phagemids and their transducing properties. *Gene* **28**, 29–35 (1984).
- 234. Seed, K. D., Lazinski, D. W., Calderwood, S. B. & Camilli, A. A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature* **494**, 489–491 (2013).
- 235. Lowy, F. D. Staphylococcus aureus Infections. N. Engl. J. Med. 339, 520–532 (1998).
- 236. Ea, G. *et al.* Topographical and temporal diversity of the human skin microbiome.,

  Topographical and Temporal Diversity of the Human Skin Microbiome. *Sci. N. Y. N Sci. N. Y. N* **324**, **324**, 1190, 1190–1192 (2009).
- 237. Ba, D. *et al.* Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant Staphylococcus aureus. *Lancet Lond. Engl.* **367**, 731–739 (2006).
- 238. Lm, W. *et al.* Genetic analysis of a high-level vancomycin-resistant isolate of Staphylococcus aureus. *Science* **302**, 1569–1571 (2003).
- 239. Parmley, S. Programmable sensitivity. SciBX Sci.-Bus. Exch. 7, (2014).
- 240. Ando, H., Lemire, S., Pires, D. P. & Lu, T. K. Engineering Modular Viral Scaffolds for Targeted Bacterial Population Editing. *Cell Syst.* **1,** 187–196 (2015).
- 241. TTCSP Global Go To Think Tank Index Reports | Think Tanks and Civil Societies Program (TTCSP) | University of Pennsylvania. Available at: http://repository.upenn.edu/think\_tanks/. (Accessed: 16th December 2016)

- 242. Antibiotic and Antimicrobial Resistance. *Chatham House* Available at: https://www.chathamhouse.org//node/13419. (Accessed: 16th December 2016)
- 243. King, Chemistry World, A. Antibiotic Resistance Will Kill 300 Million People by 2050. Scientific American Available at: https://www.scientificamerican.com/article/antibiotic-resistance-will-kill-300-million-people-by-2050/. (Accessed: 20th December 2016)
- 244. Estimating the Economic Costs of Antimicrobial Resistance | RAND. Available at: http://www.rand.org/randeurope/research/projects/antimicrobial-resistance-costs.html. (Accessed: 18th December 2016)
- 245. AIDA project General project information. Available at: http://www.aida-project.eu/project/general-project-information. (Accessed: 16th December 2016)
- 246. Das, P. & Horton, R. Antibiotics: achieving the balance between access and excess. *The Lancet* **387**, 102–104 (2016).
- 247. Compositions and methods of nucleic acid-targeting nucleic acids.
- 248. Xavier, B. B. *et al.* Identification of a novel plasmid-mediated colistin-resistance gene, mcr-2, in Escherichia coli, Belgium, June 2016. *Euro Surveill. Bull. Eur. Sur Mal. Transm. Eur. Commun. Dis. Bull.* **21,** (2016).
- 249. Chen, L. Notes from the Field: Pan-Resistant New Delhi Metallo-Beta-Lactamase-Producing Klebsiella pneumoniae Washoe County, Nevada, 2016. *MMWR Morb. Mortal. Wkly. Rep.* **66**, (2017).
- 250. DiCarlo, J. E., Chavez, A., Dietz, S. L., Esvelt, K. M. & Church, G. M. Safeguarding CRISPR-Cas9 gene drives in yeast. *Nat Biotechnol* **33**, 1250–1255 (2015).
- 251. Burt, A. Site-specific selfish genes as tools for the control and genetic engineering of natural populations. *Proc. R. Soc. Lond. B Biol. Sci.* **270**, 921–928 (2003).
- 252. Takeuchi, R., Choi, M. & Stoddard, B. L. Redesign of extensive protein–DNA interfaces of meganucleases using iterative cycles of in vitro compartmentalization. *Proc. Natl. Acad. Sci.* **111**, 4061–4066 (2014).
- 253. Thyme, S. B. *et al.* Reprogramming homing endonuclease specificity through computational design and directed evolution. *Nucleic Acids Res.* **42**, 2564–2576 (2014).
- 254. Hammond, A. et al. A CRISPR-Cas9 gene drive system targeting female reproduction in

- the malaria mosquito vector Anopheles gambiae. Nat. Biotechnol. 34, 78-83 (2015).
- 255. Koo, T., Lee, J. & Kim, J.-S. Measuring and Reducing Off-Target Activities of Programmable Nucleases Including CRISPR-Cas9. *Mol. Cells* **38**, 475–481 (2015).
- 256. Zika. *Institut Pasteur* Available at: http://www.pasteur.fr/fr/institut-pasteur/presse/fiches-info/zika. (Accessed: 30th October 2016)
- 257. Shaw, J. Editing an End to Malaria? *Harvard Magazine* (2016). Available at: http://harvardmagazine.com/2016/05/editing-an-end-to-malaria. (Accessed: 26th October 2016)
- 258. Esvelt, K. M., Smidler, A. L., Catteruccia, F. & Church, G. M. Concerning RNA-guided gene drives for the alteration of wild populations. *eLife* **3**, e03401 (2014).
- 259. Gantz, V. M. et al. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. Proc. Natl. Acad. Sci. U. S. A. 112, E6736-6743 (2015).
- 260. Gene drive. Wikipedia (2016).
- 261. OMS | Paludisme. *WHO* Available at: http://www.who.int/mediacentre/factsheets/fs094/fr/. (Accessed: 27th October 2016)
- 262. AIM 2016-2030 Roll Back Malaria. Available at: http://www.rollbackmalaria.org/about-rbm/aim-2016-2030. (Accessed: 27th October 2016)
- 263. Gabrieli, P., Smidler, A. & Catteruccia, F. Engineering the control of mosquito-borne infectious diseases. *Genome Biol.* **15,** 535 (2014).
- 264. Craig, G. B., Hickey, W. A. & VandeHey, R. C. An Inherited Male-Producing Factor in Aedes aegypti. *Science* **132**, 1887–1889 (1960).
- 265. Hamilton, W. D. Extraordinary Sex Ratios. *Science* **156**, 477–488 (1967).
- 266. Newton, M. E., Wood, R. J. & Southern, D. I. A cytogenetic analysis of meiotic drive in the mosquito, Aedes aegypti (L.). *Genetica* **46**, 297–318
- 267. Wood, R. J. & Newton, M. E. Sex-Ratio Distortion Caused by Meiotic Drive in Mosquitoes. *Am. Nat.* **137**, 379–391 (1991).
- 268. Galizi, R. et al. A synthetic sex ratio distortion system for the control of the human

- malaria mosquito. Nat. Commun. 5, (2014).
- 269. Windbichler, N. *et al.* Homing endonuclease mediated gene targeting in Anopheles gambiae cells and embryos. *Nucleic Acids Res.* **35**, 5922–5933 (2007).
- 270. Isaacs, A. T. *et al.* Transgenic Anopheles stephensi coexpressing single-chain antibodies resist Plasmodium falciparum development. *Proc. Natl. Acad. Sci. U. S. A.* **109,** E1922-1930 (2012).
- 271. Wade, N. Gene Drives Offer New Hope Against Diseases and Crop Pests. *The New York Times* (2015).
- 272. Lees, R. S., Gilles, J. R., Hendrichs, J., Vreysen, M. J. & Bourtzis, K. Back to the future: the sterile insect technique against mosquito disease vectors. *Curr. Opin. Insect Sci.* **10**, 156–162 (2015).
- 273. Who We Are » Target Malaria. Available at: http://targetmalaria.org/who-we-are/. (Accessed: 28th October 2016)
- 274. Scott, T. W., Takken, W., Knols, B. G. J. & Boëte, C. The Ecology of Genetically Modified Mosquitoes. *Science* **298**, 117–119 (2002).
- 275. Reeves, R. G., Denton, J. A., Santucci, F., Bryk, J. & Reed, F. A. Scientific Standards and the Regulation of Genetically Modified Insects. *PLOS Negl Trop Dis* **6**, e1502 (2012).
- 276. David, A. S., Kaser, J. M., Morey, A. C., Roth, A. M. & Andow, D. A. Release of genetically engineered insects: a framework to identify potential ecological effects. *Ecol. Evol.* 3, 4000–4015 (2013).
- 277. Alphey, L. Genetic Control of Mosquitoes. Annu. Rev. Entomol. 59, 205–224 (2014).
- 278. National Research Council (US) Committee on a New Government-University

  Partnership for Science and Security. Science and Security in a Post 9/11 World: A

  Report Based on Regional Discussions Between the Science and Security Communities.

  (National Academies Press (US), 2007).
- 279. Gantz, V. M. & Bier, E. Genome editing. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science* **348**, 442–444 (2015).

- *Technology Review* Available at: https://www.technologyreview.com/s/543721/with-thisgenetic-engineering-technology-theres-no-turning-back/. (Accessed: 30th October 2016)
- 281. Oye, K. A. et al. Regulating gene drives. Science 345, 626–628 (2014).
- 282. Caplan, A. L., Parent, B., Shen, M. & Plunkett, C. No time to waste--the ethical challenges created by CRISPR: CRISPR/Cas, being an efficient, simple, and cheap technology to edit the genome of any organism, raises many ethical and regulatory issues beyond the use to manipulate human germ line cells. *EMBO Rep.* **16**, 1421–1426 (2015).
- 283. Benedict, M. *et al.* Guidance for contained field trials of vector mosquitoes engineered to contain a gene drive system: recommendations of a scientific working group. *Vector Borne Zoonotic Dis. Larchmt. N* **8,** 127–166 (2008).
- 284. Regalado, A. Bill Gates sees CRISPR gene drive eradicating mosquitoes in Africa by 2029. *MIT Technology Review* Available at: https://www.technologyreview.com/s/601213/the-extinction-invention/. (Accessed: 26th October 2016)
- 285. TDR | The Guidance Framework for testing genetically modified mosquitoes. *WHO*Available at: http://www.who.int/tdr/publications/year/2014/guide-fmrk-gm-mosquit/en/.

  (Accessed: 30th October 2016)
- 286. Waltz, E. Gene-edited CRISPR mushroom escapes US regulation. *Nature* **532**, 293 (2016).
- 287. Biotech Crop Highlights in 2015 Pocket K | ISAAA.org. Available at: http://isaaa.org/resources/publications/pocketk/16/default.asp. (Accessed: 10th January 2017)
- 288. Bennett, D. J. & Jennings, R. C. Successful Agricultural Innovation in Emerging Economies: New Genetic Technologies for Global Food Production. (Cambridge University Press, 2013).
- 289. Svitashev, S. Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. *Plant Physiol* **169**, 931–945 (2015).
- 290. Li, Z. Cas9-guide RNA directed genome editing in soybean. *Plant Physiol* **169**, 960–970 (2015).

- 291. Ali, Z. Efficient virus-mediated genome editing in plants using the CRISPR/Cas9 system. *Mol Plant* **8,** 1288–1291 (2015).
- 292. Whitworth, K. M. Gene-edited pigs are protected from porcine reproductive and respiratory syndrome virus. *Nat Biotechnol* **34**, 20–22 (2016).
- 293. Carlson, D. F. Production of hornless dairy cattle from genome-edited cell lines. *Nat Biotechnol* **34**, 479–481 (2016).
- 294. Peng, J. Production of human albumin in pigs through CRISPR/Cas9-mediated knockin of human cDNA into swine albumin locus in the zygotes. *Sci Rep* **5**, 16705 (2015).
- 295. Talbot, D. 10 Breakthrough Technologies 2016: Precise Gene Editing in Plants. *MIT Technology Review* Available at: https://www.technologyreview.com/s/600765/10-breakthrough-technologies-2016-precise-gene-editing-in-plants/. (Accessed: 11th January 2017)
- 296. Monsanto Announces Global Licensing Agreement with Broad Institute on Key
  Genome-Editing Application | Monsanto Newsroom. Available at:
  http://news.monsanto.com/press-release/corporate/monsanto-announces-global-licensing-agreement-broad-institute-key-genome-edi. (Accessed: 11th January 2017)
- 297. Bayer et CRISPR Therapeutics AG signent une alliance portant sur la découverte, le développement et la commercialisation de traitements potentiels de maladies génétiques graves | Bayer France. Available at: https://www.bayer.fr/communique-bayer-CRISPR-Therapeutics. (Accessed: 11th January 2017)
- 298. Alessi, C. & Bunge, J. Bayer Makes \$62 Billion Bid for Monsanto. *Wall Street Journal* (2016).
- 299. Watchdogs weigh up pros and cons of seed mergers. *Financial Times* Available at: https://www.ft.com/content/4c5774a8-a69c-11e6-8898-79a99e2a4de6. (Accessed: 11th January 2017)
- 300. EFSA Panel on Genetically Modified Organisms (GMO). Scientific opinion addressing the safety assessment of plants developed using Zinc Finger Nuclease 3 and other Site-Directed Nucleases with similar function: Plants developed by ZFN-3. *EFSA J.* **10**, 2943 (2012).

- 301. Voytas, D. F. & Gao, C. Precision Genome Engineering and Agriculture: Opportunities and Regulatory Challenges. *PLOS Biol.* **12**, e1001877 (2014).
- 302. Ledford, H. Gene-editing surges as US rethinks regulations. Nat. News 532, 158 (2016).
- 303. Environmental Impact Statement; Introduction of the Products of Biotechnology. Federal Register (2016). Available at:

  https://www.federalregister.gov/documents/2016/03/07/2016-04992/environmental-impact-statement-introduction-of-the-products-of-biotechnology. (Accessed: 14th January 2017)
- 304. Kuzma, J. Policy: Reboot the debate on genetic engineering. *Nat. News* **531**, 165 (2016).
- 305. Breeding controls. Nat. News **532**, 147 (2016).
- 306. Ledford, H. US regulation misses some GM crops. Nat. News 500, 389 (2013).
- 307. Kuzma, J. A missed opportunity for U.S. biotechnology regulation. *Science* **353**, 1211–1213 (2016).
- 308. Umeå researcher served a world first (?) CRISPR meal Umeå University, Sweden. (2016). Available at: http://www.umu.se/english/about-umu/news-events/news/newsdetailpage/umea-researcher-served-a-world-first----crispr-meal.cid272955. (Accessed: 10th January 2017)
- 309. EUR-Lex 32015L0412 EN EUR-Lex. Available at: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AJOL 2015 068 R 0001. (Accessed: 11th January 2017)
- 310. Barrangou, R. Advances in CRISPR-Cas9 genome engineering: lessons learned from RNA interference. *Nucleic Acids Res* **43**, 3407–3419 (2015).
- 311. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science* **343**, 80–84 (2014).
- 312. Chen, S. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell* **160**, 1246–1260 (2015).
- 313. Kampmann, M. Next-generation libraries for robust RNA interference-based genomewide screens. *Proc Natl Acad Sci USA* **112**, E3384–E3391 (2015).
- 314. Kampmann, M., Bassik, M. C. & Weissman, J. S. Functional genomics platform for pooled screening and generation of mammalian genetic interaction maps. *Nat Protoc* 9,

- 1825–1847 (2014).
- 315. Malina, A. Adapting CRISPR/Cas9 for functional genomics screens. *Methods Enzym*. **546**, 193–213 (2014).
- 316. CRISPResso. Available at: http://crispresso.rocks/. (Accessed: 14th December 2016)
- 317. Korkmaz, G. Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat Biotechnol* **34**, 192–198 (2016).
- 318. Rajagopal, N. High-throughput mapping of regulatory DNA. *Nat Biotechnol* **34,** 167–174 (2016).
- 319. Klymiuk, N. Tailored pig models for preclinical efficacy and safety testing of targeted therapies. *Toxicol Pathol* **44,** 346–357 (2016).
- 320. Heidenreich, M. & Zhang, F. Applications of CRISPR-Cas systems in neuroscience. *Nat Rev Neurosci* **17**, 36–44 (2016).
- 321. Wang, H. *et al.* One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell* **153**, 910–918 (2013).
- 322. Soldner, F. *et al.* Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell* **146**, 318–331 (2011).
- 323. Ryan, S. D. *et al.* Isogenic human iPSC Parkinson's model shows nitrosative stress-induced dysfunction in MEF2-PGC1α transcription. *Cell* **155**, 1351–1364 (2013).
- 324. Jeon, I. *et al.* Neuronal properties, in vivo effects, and pathology of a Huntington's disease patient-derived induced pluripotent stem cells. *Stem Cells Dayt. Ohio* **30**, 2054–2062 (2012).
- 325. Yagi, T. *et al.* Modeling familial Alzheimer's disease with induced pluripotent stem cells. *Hum. Mol. Genet.* **20,** 4530–4539 (2011).
- 326. Web of Science Predicts 2016 Nobel Prize Winners.
- 327. The Nobel Prize in Chemistry 2016. Available at:

  https://www.nobelprize.org/nobel\_prizes/chemistry/laureates/2016/. (Accessed: 17th
  January 2017)
- 328. Merton, R. K. Resistance to the Systematic Study of Multiple Discoveries in Science. *Eur. J. Sociol. Arch. Eur. Sociol.* **4,** 237–282 (1963).

- 329. Kelly, K. What Technology Wants. (Penguin, 2010).
- 330. Patent Trial and Appeal Board (PTAB). Available at:

  https://acts.uspto.gov/ifiling/PublicView.jsp?

  identifier=106048&identifier2=null&tabSel=4&action=filecontent&replyTo=PublicView.jsp. (Accessed: 21st September 2016)
- 331. Zhang, F. United States Patent: 8697359 CRISPR-Cas systems and methods for altering expression of gene products. (2014).
- 332. Ledford, H. Bitter fight over CRISPR patent heats up. Nat. News 529, 265 (2016).
- 333. Reardon, S. CRISPR heavyweights battle in US patent court. *Nat. News* **540**, 326 (2016).
- 334. McNeil, C. Who Invented the VLP Cervical Cancer Vaccines? *J. Natl. Cancer Inst.* **98**, 433–433 (2006).
- 335. Ledford, H. Why the CRISPR patent verdict isn't the end of the story. *Nat. News* doi:10.1038/nature.2017.21510
- 336. Novel crispr enzymes and systems.
- 337. Pardee, K. *et al.* Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell* **165**, 1255–1266 (2016).
- 338. Ledford, H. CRISPR, the disruptor. *Nature* **522**, 20–24 (2015).
- 339. Total funding for CRISPR contenders tops \$1B as Crispr Therapeutics prices \$56M IPO. *BioWorld* Available at: http://www.bioworld.com/content/total-funding-crispr-contenders-tops-1b-crispr-therapeutics-prices-56m-ipo-0. (Accessed: 19th January 2017)
- 340. Berg, P. Meetings that changed the world: Asilomar 1975: DNA modification secured. *Nature* **455**, 290–291 (2008).
- 341. International Summit on Human Gene Editing. Available at: http://nationalacademies.org/gene-editing/Gene-Edit-Summit/. (Accessed: 19th January 2017)
- 342. On Human Gene Editing: International Summit Statement. Available at: http://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=12032015a. (Accessed: 19th January 2017)
- 343. Hotta, A. & Yamanaka, S. From genomics to gene therapy: induced pluripotent stem

- cells meet genome editing. Annu Rev Genet 49, 47-70 (2015).
- 344. Kamao, H. Characterization of human induced pluripotent stem cell-derived retinal pigment epithelium cell sheets aiming for clinical application. *Stem Cell Rep.* **2**, 205–218 (2014).
- 345. Callaway, E. UK scientists gain licence to edit genes in human embryos. *Nat. News* **530**, 18 (2016).
- 346. Jou, W. M., Haegeman, G., Ysebaert, M. & Fiers, W. Nucleotide Sequence of the Gene Coding for the Bacteriophage MS2 Coat Protein. *Nature* **237**, 82–88 (1972).
- 347. Lewontin, R. C. *The Triple Helix: Gene, Organism, and Environment*. (Harvard University Press, 2001).
- 348. Schork, N. J., Murray, S. S., Frazer, K. A. & Topol, E. J. Common vs. Rare Allele Hypotheses for Complex Diseases. *Curr. Opin. Genet. Dev.* **19**, 212–219 (2009).
- 349. Cox, D. B. T., Platt, R. J. & Zhang, F. Therapeutic Genome Editing: Prospects and Challenges. *Nat. Med.* **21**, 121–131 (2015).
- 350. Dyda, F. & Hickman, A. B. Mechanism of spacer integration links the CRISPR/Cas system to transposition as a form of mobile DNA. *Mob. DNA* 6, 9 (2015).
- 351. The Maxine Singer Papers: Risk, Regulation, and Scientific Citizenship: The Controversy over Recombinant DNA Research: Documents. Available at: https://profiles.nlm.nih.gov/ps/retrieve/narrative/dj/p-nid/218/p-docs/true. (Accessed: 19th January 2017)

**AUTEUR**: Jean-Baptiste Roublin

TITRE: CRISPR/Cas9: Histoire, Méthode, Potentiel & Impact

**DIRECTEUR DE THESE**: Professeur Bettina Couderc

Toulouse, le 24 Mars 2017

**RÉSUMÉ**: Les bactéries sont soumises à une immense pression évolutive par les attaques virales. CRISPR-Cas pour Clustered regularly interspaced short palindromic repeats - CRISPR associated genes désigne une structure génétique procaryote dont l'expression confère aux bactéries et archées une immunité adaptative dirigée contre les virus. La curiosité de diverses équipes scientifiques a permis la compréhension et le détournement de ce système au profit de l'édition génomique. L'outil obtenu est composé d'une protéine qui associée à un ARN guide sera capable de cibler et cliver n'importe quel acide nucléique de manière spécifique. La protéine Cas9 est en 2017 celle qui répond aux exigences de vectorisation, spécificité et production des laboratoires et de l'industrie. Elle est aujourd'hui considérée comme l'outil d'édition génomique le plus simple, le plus précis et le moins coûteux, entraînant une popularité sans précédent pour ce domaine. Ses applications sont multiples et de nombreuses preuves de concepts ont été développées. Chacune des applications correspond à un contexte scientifique mais aussi médical ou législatif. Ces contextes nous permettent de comprendre l'importance, les espoirs mais aussi les risques et les limites de ces applications. L'édition génomique pourrait être capable de redéfinir la place de l'humanité au sein de la nature (génétique dirigée), notre définition de celleci (usage agronomique) ou même notre propre évolution (édition embryonnaire). Elle peut aussi apporter des espoirs de traitements curatifs et une meilleure compréhension de pathologies, certaines pensées incurables (HIV, HBV, HSV, Mucoviscidose, Thalassémie, Cataracte, Tyrosinémie) et constituer l'un des derniers remparts face à la résistance bactérienne aux antibiotiques. Par ailleurs elle représente un outil de recherche et d'exploration des mécanismes pathogènes et génétiques qui pourrait nous permettre d'entrevoir les liens complexes qui unissent génome, organisme et environnement.

MOTS-CLES: Biotechnologie, CRISPR, Édition Génomique, Histoire des Sciences, Microbiologie.

**ABSTRACT**: Bacteria are placed under immense evolutionary pressure by viral attacks. The CRISPR-Cas (Clustered Regularly Interspaced Palindromic Repeats – CRISPR associated) system is a prokaryotic genomic structure that confers an adaptive immunity to bacteria and archaea. Its investigation by multiple scientific groups allowed a comprehension of this novel system, enabling the manipulation of its capabilities for the purpose of genome editing. This tool, made of a protein and a guideRNA, can specifically target and cleave any nucleic acid sequence. By its size, specificity, and ease of production, Cas9 is the protein that fits researchers' and industrial requirements in 2017. Today, it is considered the simplest, most precise and cost-effective tool in genome editing, garnering an unprecedented popularity in its field. Its applications are many, and several proofs of concept have been developed. Each of these corresponds to a specific scientific, medical, or regulatory context that allows us to understand its importance, its aspirations, its limits and its risks. Genome editing could allow us to redefine humanity's role within nature (gene drive), challenge our definition of it (agronomic application) and our own evolution (embryonic genome editing). It could also provide hopes of treatments and an enhanced understanding of several diseases traditionally thought incurable (HIV, HBV, HSV, Cystic Fibrosis, Thalassemia, Tyrosinemia, Cataracts), and become our last resort in the face of bacteria's antibiotic resistance. Moreover, it is an unmatched tool for genomic and pathogenic exploration that could give us a glimpse into the complex links between genes, organisms and environments.

**KEYWORDS**: Biotechnology, CRISPR, Genome editing, History of Sciences, Microbiology.

**DISCIPLINE ADMINISTRATIVE**: Pharmacie Industrielle

**INTITULE ET ADRESSE DE L'UFR OU DU LABORATOIRE** : Faculté des Sciences Pharmaceutiques de TOULOUSE. 35 Chemin des Maraîchers – 31400 Toulouse - France