



HAL
open science

Edition à la base près de séquences d'ADN chez le riz et la vigne pour introduire des caractères d'intérêt agronomiques par CRISPR

Leo Herbert

► To cite this version:

Leo Herbert. Edition à la base près de séquences d'ADN chez le riz et la vigne pour introduire des caractères d'intérêt agronomiques par CRISPR. Amélioration des plantes. Montpellier SupAgro, 2021. Français. NNT : 2021NSAM0013 . tel-04006505

HAL Id: tel-04006505

<https://theses.hal.science/tel-04006505v1>

Submitted on 27 Feb 2023

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

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

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE MONTPELLIER SUPAGRO

En Génétique et Amélioration des Plantes

École doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau
Portée par

L'Unité de recherche AGAP (Amélioration Génétique et Adaptation des Plantes)

Edition à la base près de séquences d'ADN chez le riz et la vigne
pour introduire des caractères d'intérêt agronomiques par CRISPR

Présentée par Léo Herbert
Le 24 Mars 2021

Sous la direction de Patrice This
et Christophe Périn

Devant le jury composé de

Annabelle Déjardin (Chargée de Recherche, HDR, INRAe)

Pierre Barret (Ingénieur de Recherche, HDR, INRAe)

Oumaya BOU CHABKE-COUSSA (Ingénieure de Recherche, INRAe)

Fabien Nogué (Directeur de Recherche, HDR, INRAe)

Pascal Gantet (Professeur, HDR, Université de Montpellier)

Patrice This (Directeur de Recherche, HDR, INRAe)

Christophe Périn (Directeur de Recherche, HDR, CIRAD)

Olivier Zekri (Responsable de Recherche, Pépinières Mercier)

Rapporteur

Rapporteur

Examineur

Examineur

Examineur – Président du jury

Codirecteur de thèse

Codirecteur de thèse

Invité



UNIVERSITÉ
DE MONTPELLIER

SupAgro
Montpellier

Remerciements

Je ne sais pas mais je pense que ma bonne humeur, mon humour débile, le fait de toujours être cool, à l'écoute prêt à aider, la porte du bureau remplie de bêtises qui décrivaient bien le personnage, et mes gâteaux tous les lundis ont sûrement joué sur le fait que j'étais apprécié et que je bossais avec beaucoup de monde avec plaisir. Je m'excuse par avance si j'ai oublié des gens que j'ai côtoyé avec qui j'ai apprécié de travailler mais il y en a eu beaucoup il faut dire ! La liste est longue, et tant mieux non ? ce serait triste sinon ...

Tout d'abord j'aimerais commencer par ceux de mes deux directeurs de thèse. Christophe, pour toute l'aide apportée, le support mental pendant ces trois années, et aussi pour ces lointaines soirées au bar et au Rockstore. Il faut dire que tu as bien du courage de prendre des doctorants cascadeurs aimant faire des commotions cérébrales. C'est une marque de fabrique en soi. Et un énorme merci à Patrice pour l'ensemble des discussions, congrès, et toutes les aides que tu as pu me fournir, tant sur les recherches que sur le plan mental. Au final c'était une sacrée expérience et un sacré défi d'avoir deux directeurs mais j'ai survécu je vais dire.

Ensuite viennent les deux équipes pour lesquelles j'ai travaillé. DAR pour commencer. Depuis mon M1 vous m'avez eu, et je ne remercierai jamais assez toute l'équipe ... Un immense merci à Anne Cécile la CRISPR master, rien n'aurait pu être fait sans ton aide tout au long de la thèse, je te dois énormément ! Viens ensuite Aurore ma coéquipière de boxe, Murielle alias Mumu, Donaldo mon collègue de bureau insatiable, Martine qui m'aura vu grandir depuis mon M1, Sergi el català com jo, Fabrice que je remercie pour m'avoir donné l'occasion de donner des cours magistraux à la fac, Emmanuel le boss, Anne pour toutes les corrections et les avancées dans la réflexion scientifique des présentations, Corinne pour tout l'administratif dans lequel nous étions perdus pauvres doctorants que nous étions, et la team serriste Eve, Rémy alias Popeye, et Christian ! Je n'oublie pas les sorghotes Nancy pour toutes les aides apportées sur la vigne, Maëlle pour les bars et la bonne ambiance musicale au labo, ainsi que Fred et Angel. Ensuite vient DAAV, avec la mention spéciale à Valérie pour tous les protoplastes faits ensemble, et qui a réussi à me supporter trois ans à courir partout ... après deux hyperactifs ensemble c'est cool ; et Béatrice qui a survécu aux millions de repiquages d'embryons ! N'oublions pas le papa Loïc et Cédric l'ingénieur qui fait des analyses « d'un autre monde » pour les fous rires des congrès, les cannelés bordelais qui m'ont fait découvrir le vrai monde d'après en 2018, et toutes les aides apportées. Je n'oublie pas Thierry L. et Laurent A. non plus qui ont participé à la bonne ambiance au congrès de Bordeaux, vous m'avez tellement régalaé que je ne peux m'empêcher de vous citer.

Je souhaite aussi remercier la startup IAGE, sans qui beaucoup des analyses n'auraient pas pu être faites. On s'est vu grandir mutuellement on va dire, ça fait bizarre. Franz, que j'ai eu le plaisir de suivre pendant cinq ans et maintenant en tant que big boss dans IAGE, Antoine, Max, et Jawhar, merci pour m'avoir aidé avec cette sainte ddPCR sans laquelle peu de choses auraient été possibles. Bon courage à vous pour la suite de l'aventure.

Mes deux techniciens favoris : Théo et Dylan ... Alors vous deux que dire, à part merci pour tout ! La liste est longue mais bon, dès qu'on peut à nouveau retourner au bar il y a des bières à rattraper.

Des « Specials thanks » au viking Ian qui grimpe partout et qui m'a bien remonté le moral entre les bières, et Thibault le kitesurfeur maladroit (merci de m'avoir accompagné dans la commotion, je me sentais bien seul), merci pour ces trois années à vos côtés mes poulets. Mais il n'y a pas que ces deux lascars en doctorants, Kelly et Clara, merci toutes les deux plus particulièrement pour avoir été là et m'avoir aidé et supporté pendant trois ans.

Tous les stagiaires/anciens doctorants du labo en général : Dylan parce que ça a été drôle de te teindre les cheveux en (col)vert et surtout de bosser avec toi ; Mathilde, Chloé, Roxane, Antoine et Xavier avec qui il était bon le temps des bières au Shakespeare à refaire le monde devant un match de l'Atletico ou Tottenham, Adriaan pour les bières hollandaises et le foot ; les doctorants de Graines d'AGAP : Marion, Aurélien, Charlotte, Clément, Aurélie, Nicolas, Céline, Benjamin ... et tous les autres (je sais plus, je suis un poisson rouge).

Bien sûr que je remercie tout le labo au CIRAD, après il est vrai que Tic et tac (Fabien et Julien) méritent des remerciements plus poussés pour les interminables débats et blagues débiles ; Joan aussi pour ne pas m'avoir étranglé depuis 3 ans parce que CRISPR marche dans le riz ; et tous les autres Catherine, Caroline, Hélène, Marie, David pour toutes les aides sur les stats, Tanguy les expérimentations physio et l'aide pour les analyses ...

Remerciements aussi pour les personnes d'un autre laboratoire avec qui j'ai travaillé et qui m'ont apporté une aide précieuse pour les sujets que je ne maîtrisais pas : Laurence, et Valentin que j'ai eu le plaisir de suivre depuis 5 ans en Master BFP puis en concours puis en thèse, et un spécial thanks à Antoine et Fanny d'Intégration à BPMP sans qui je n'aurais jamais fait cette thèse il y a 3 ans. En fait c'est l'équipe entière d'Intégration que je remercie et qui m'a toujours accordé le café dans la bonne humeur quand je passais les voir.

Remerciement aux membres de mes comités Laurence (re), Marianne, Virginie, et Sébastien qui m'ont bien orienté et qui étaient plutôt contents de l'enthousiasme que je mettais dans ma thèse.

Maintenant viennent ceux de l'ombre que personne ne connaît au labo, mais que ce club de boxe française a été un défouloir pendant plus de trois ans ! Je ne remercierai jamais assez mes entraîneurs pour avoir désespérément essayé de m'assouplir, mais bon la grâce du tronc d'arbre restera légendaire ... tout comme la nullité de ma commotion, on va peut-être en rester à la version officielle. Merci aussi de m'avoir fait découvrir les compétitions certes amicales mais sur le ring, ça change un homme. Et là, le saint graal, mon équipe de Foot à 7, la TennisUnited... des moments inoubliables, des victoires, du beau football, et des coupes dont la Champions League de Montpellier... mais surtout une bande de potes depuis 4 ans ! L'important c'est l'essentiel comme dirait l'un d'eux.

On finira bien sûr par toute ma famille qui a toujours été là et m'a toujours soutenu, et mes trois meilleurs potos sans qui supporter cette thèse n'aurait pas été possible : Mathias, Paul, et Teddy. Promis, on fêtera ça comme il se doit.

En espérant que la queue du raton laveur qui déambule dans le B3 vous manque ... Merci

Ah et j'ai failli oublier, mais le pangolin d'AGAP c'était pas prémédité je le jure !

Sommaire

Sommaire	1
Sommaire des Figures	6
Abréviations — traductions.....	9
Introduction générale.....	13
Partie 1 : CRISPR la « New Breeding Technology » de ce début de siècle	15
1) Aperçu historique des technologies d'édition des génomes avant CRISPR-CAS9	15
2) La technologie CRISPR-CAS9 : un bref rappel historique d'une technologie d'édition révolutionnaire.....	19
3) La technologie CRISPR-CAS9 « pour les nuls »	21
4) Avantages de la technologie.....	25
5) Limites à l'utilisation de la technologie CRISPR/CAS9	27
6) Nouvelles technologies dérivées de CRISPR : Multiplexing, KI, remplacement, Base Editing, et modulations transcriptionnelles ou chromatiniques	31
7) Les objectifs d'amélioration variétale par la technologie CRISPR.....	37
a) Modification rapide et ciblée de gènes ou d'allèles	37
b) Élargissement de la diversité génétique et prébreeding	39
c) Concept de « Rewilding » et de domestication <i>de novo</i>	41
d) Réarrangements Chromosomiques de génomes végétaux à l'aide du système CRISPR/CAS9.....	41
8) Des enjeux juridiques complexes et très différents selon les pays.....	41
Conclusions sur la technologie CRISPR	45
Partie 2 : La vigne et le riz, deux espèces candidates pour l'amélioration via les NBT.....	47
I] La vigne : les NBTs comme outils de validation fonctionnelle et d'amélioration génétique, intérêts et difficultés.....	47
1) Espèces et domestication	49
2) Morphologie et reproduction	51
3) Génomique et amélioration.....	53
4) La vigne et les biotechnologies.....	55
5) La microvigne, un outil pour accélérer les cycles de sélection.....	57

II] Le riz, une plante modèle pour le développement de l'utilisation des NBTs en amélioration des plantes.....	63
1) Morphologie et espèces du riz.....	63
2) Amélioration du riz : <i>Chhomrong dhan</i> (CHHD) et le riz d'altitude malgache.....	65
3) L'efficacité de l'utilisation de l'azote comme cible d'amélioration par les NBT	67
 Problématiques et objectifs	 73
 Chapitre 1 : Mise au point des technologies.....	 77
I] Production et transformation de protoplastes	79
1) Protocoles développés à partir des bases de tige de riz.....	79
2) Protocoles développés à partir de cultures cellulaires de vigne	83
II] Mise au point de la technologie ddPCR et analyses	87
 III] Transformation transitoire à l'aide de RiboNucléoProtéines (RNP) chez le riz : résultats préliminaires.....	 93
 Chapitre 2 : Application de l'édition des génomes au riz et à la vigne	 97
 I] Les technologies CRISPR : ou comment générer une série allélique à un locus pour faciliter l'analyse fonctionnelle et l'utilisation des construits d'édition comme NBT (New Breeding Technologies).....	 99
 II] Amélioration des traits agronomiques de la NUE chez le riz à l'aide des technologies CRISPR : vers leur utilisation comme NBT	 107
1) Kitaaké : ARN, influx ¹⁵ N, et Physiologie	109
a) ARN et influx ¹⁵ N.....	109
b) Physiologie.....	115
2) Chhomrong dhan (CHHD) : Physiologie au champ en Colombie	119
 III] Amélioration de la Vigne via CRISPR : transformations stables et analyse des mutants potentiels pour le nanisme	 123
1) Premier test avec les plasmides monocotylédones.....	123
2) Second test : amélioration des plasmides et tests préalables par ddPCR.....	123
a) Constructions CRISPR Vigne	123
b) Détection des mutations en protoplastes par la méthode PCR ou T7EI.....	125

c) Détection en ddPCR pour évaluer l'efficacité de mutation globale du construit cr265	127
3) transformation stable de la vigne avec les construits validés	131
4) Optimisation des construits CRISPR/CAS9 chez la vigne	131
Discussion générale	137
I] Réflexions sur la mise au point des différentes techniques	139
1) Les protoplastes comme système de prototypage de construits CRISPR : avantages et limitations	139
2) La ddPCR pour quantifier le taux de mutations global et spécifique comparé au séquençage NGS : combiner le meilleur des deux mondes	141
3) Quelle fiabilité des résultats obtenus en transformation transitoire pour prédire l'édition et la spécificité de construits CRISPR/CAS9 en transformation stable ?	143
4) Comment améliorer la sécurité des technologies CRISPR/CAS9 ? mutations aléatoires, off-Targets et in-targets	145
a) Off-Targets	145
b) In-Targets	147
II] Réflexions sur les technologies dérivées de CRISPR et l'utilisation sur des gènes d'intérêt agronomique	149
1) La technologie CRISPR/CAS9 et ses dérivés : un outil puissant et sans équivalent pour générer rapidement et simplement des séries alléliques et faciliter l'analyse fonctionnelle et la sélection .	149
2) Bien que la transformation génétique soit bien maîtrisée chez le riz, améliorer et rendre les protocoles de transformation les plus génériques possibles reste une nécessité chez la plupart des espèces végétales pour lesquelles peu de génotypes sont transformables avec une efficacité minimale	149
3) Générer et tester de nouveaux allèles pour l'efficacité d'utilisation de l'azote chez le riz	151
a) Exploration de la fonction de <i>OsNRT1.1b</i> et de <i>OsBT</i> dans le transport d'azote en plantules de riz	151
b) Impact des mutations BT KO, NRT1.1b KO, et BE sur la NUE en serre dans un fond Kitaaké : analyse des raisons de l'échec du phénotypage en serre et recommandations pour des expérimentations futures	153
c) Impact des mutations NRT1.1b KO, BE et BT KO sur la NUE au champ dans un fond CHHD : les lignées bases éditées, possédant un allèle <i>indica</i> C>T ou C>G semblent avoir un meilleur rendement et une meilleure NUE.	155
4) Recommandations pour le futur : combiner conditions contrôlées, conditions de champ en utilisant une source unique d'azote proche des conditions agronomiques	159
5) Vers l'utilisation de la technologie CRISPR/CAS9 chez la vigne : outils de prototypage, constructions CAS9 optimisées et production de plantes éditées pour le gène <i>VvGAI</i>	159
III] Discussion globale	163

Matériel et Méthodes	171
I] Matériel végétal et culture	173
1) Stérilisation des semences et des explants	173
2) Conditions en chambre de culture	175
3) Serre de confinement	175
II] Méthodes de transformations	179
1) Transformation transitoire du riz	179
2) Transformation transitoire de la vigne	181
3) Transformation stable du riz	183
4) Transformation stable de la vigne	187
III] Analyses moléculaires : ADN, ARN, plasmides, influx	191
1) Clonages en plasmides CRISPR	191
2) Extraction d'ADN végétal	195
3) Extraction d'ARN racinaire total et RT-PCR	197
4) PCR et purifications de produits PCR	199
5) Tests T7E1	201
IV] ddPCR	203
1) Design TaqMan FAM/HEX	203
2) Réaction de ddPCR	203
3) Design Evagreen ARN	205
4) Tests d'optimisation	205
5) Extraction et purifications après amplification par ddPCR	207
V] Génotypage	209
1) Détermination du nombre de copies du T-DNA par qPCR	209
2) Identification des mutations induites par CRISPR/CAS9	209
3) Test de résistance à l'hygromycine sur feuille	211
VI] Physiologie et NUE	213
1) NUE en faible et forte concentration d'urée — Champ en Colombie (Chhomrong Dhan)	213

2) NUE en faible et forte concentration en KNO_3 — Serre de confinement (Kitaaké).....	215
3) Influx ^{15}N en racines et ARN.....	215
Annexes.....	219
Bibliographie.....	257
Résumé.....	271

Sommaire des Figures

Figure	Titre	Page
1	Principe des réparations de l'ADN par Bortesi et Fisher, 2015	14
2	Différence des prix des technologies par F. Nogué	14
3	Fonctionnement du système CRISPR CAS9 d'après Jinek et al., 2012	20
4	Structure 3D du complexe CRISPR CAS9 d'après Nishimasu et al., 2014	22
5	Table des différentes CAS9 et leurs PAM associées d'après Li et al., 2019	26
6	Listing des nouveaux NBT CRISPR par Adli, 2018	30
7	Technologie CRISPR Base-Editor nCas9 d'après Hess et al., 2017	32
8	Répartition des cultures dans le monde d'après Cantu et Walker., 2019	46
9	Surface de culture des vignes mondiales et production	46
10	Arbre phylogénétique partiel de la vigne	48
11	Classification des variétés de vigne les plus cultivées	50
12	Vignes sur le domaine de Pech Rouge	50
13	Amélioration chronologique de la vigne	56
14	Mécanismes moléculaires de réponse aux GA et fonction du gène <i>VvGAI</i>	58
15	Carte mondiale des imports, exports, et consommation de riz en 2010, d'après Muthayya et al., 2014	62
16	Présentation des systèmes de cultures en riz irrigués à Madagascar	62
17	Amélioration et évolution de la production du riz de 1960 à 2010 d'après l'IRRI	64
18	Croissance et rendement sur faible et forte concentration d'azote chez le riz avec le gène <i>OsNRT1.1b japonica</i> ou d' <i>indica</i> d'après Hu et al., 2015	68
19	Croissance et NUE chez une lignée <i>japonica cv Donjing</i> WT par rapport à une lignée KO <i>Osbt</i> d'après Araus et al., 2016	68
20	Pipeline des transformations CRISPR en riz et vigne	72
21	Diagramme de la production et détection des riz et vignes issus de NBT	74
22	Test de transformation de protoplastes de cultures cellulaires de Merlot	82
23	Principe de la ddPCR sur protoplastes	86
24	Transformation de protoplastes de riz avec la eGFP	92
25	Résultats du test T7E1 sur les protoplastes transformés par Miao.0 CAO1 et par RNP CAS9-sgRNACAO1	94
26	Principe simplifié de l'iSTOP	100
27	Production de cals primaires et secondaires pour la transformation des lignées du programme Generice	106
28	Présentation des 8 génotypes sur fond Kitaaké utilisés dans l'étude sur la NUE: Code et génotypes	108
29	Présentation des génotypes Chhomrong dhan sélectionnés : Code et génotypes	108
30	Résultats d'influx pour les différentes lignées éditées	110

31	Etude du niveau d'expression du gène <i>OsBT</i> dans les différentes lignées éditées dans les deux conditions de culture à 5 mM NH₄NO₃ et 5 mM KNO₃	112
32	Etude du niveau d'expression du gène <i>OsNRT1.1b</i> dans les différentes lignées éditées dans les deux conditions de culture à 5 mM NH₄NO₃ et 5 mM KNO₃	112
33	Alignement des acides aminés de la protéine <i>BT</i> WT et de la protéine mutée par KO CRISPR	112
34	Principaux paramètres mesurés sur Kitaaké en serre et Chhomrong dhan en Colombie	114
35	Mesures de la NUE, du HI, et du Rendement en graines et chez Kitaaké réalisés en juin en serre	116
36	Mesures du HI, du rendement, et de la NUE des lignées Chhomrong Dhan au champ en Colombie	118
37	Radar des lignées Base Edités de CHHD	120
38	Transformation et régénération chez la vigne (2018)	122
39	Position des sites CRISPR simplex ou multiplexes sur <i>VvGAI</i> (vigne)	122
40	Résultat du test délétion des CRISPR vigne	124
41	Résultat de la T7E1 sur le CRISPR simplex vigne	124
42	Résultat de la T7E1 sur CRISPR simplex ou Base Editing en riz	124
43	ddPCR pour le CRISPR simplex sur <i>VvGAI</i>, par DropOff et Coamplification	126
44	Design de la ddPCR et du CRISPR Multiplex sur <i>VvGAI</i>	128
45	Transformation et régénération chez la vigne (2019) : stade embryonnaire	130
46	« Matrice » des combinaisons des différentes protéines CAS et des promoteurs non codants	132
47	ddPCR des combinaisons plasmidiques — Résultats préliminaires	132
48	Analyse préliminaire des Off-Targets du crNRT1.1b	144
49	Roue du rendement entre WT CHHD et les Base-Edités CHHD en Colombie	156
50	Diagramme d'inscription d'une nouvelle variété	166
51	Diagramme théorique des avantages du « Benefit-Sharing »	168
52	Stérilisation des grains de riz	172
53	Stérilisation des anthères de vigne	172
54	Workflow de la production et transformation de protoplastes de riz	178
55	Workflow de la production et transformation de protoplastes de vigne	180
56	Transformation stable du riz par <i>Agrobacterium</i>	184
57	Transformation stable de la vigne par <i>Agrobacterium</i> — tiré de Bouquet et al., 2006	186
58	Principe du clonage des CRISPR et recombinaisons	190
59	Fonctionnement du test T7E1	200
60	Présentations des données de culture en Colombie (CIAT-champ)	212

61	Planches représentant les systèmes de culture LN et HN en semi-hydroponie	214
62	Prélèvements racinaires en hydroponie pour l'expression et l'influx	216

Abréviations — traductions

A, T, G, C, U : nucléotides

AA : Acide Aminé

ABE : Adenine Base Editor

ADN : Acide désoxyribonucléique

ALS : Acetolactate syntase

ARN : Acide ribonucléique

At : *Arabidopsis thaliana*

BE : Base Editing

BER : Base Editing Repair

CAS9 : protéine CAS9

CBE : Cytosine Base Editor

CHHD : Chhomrong dhan

CMV : cytomégalovirus

CRISPR : Clustered Regulatory Interspaced short Palindromic Repeats associated protein

crRNA : CRISPR RNA

cv : cultivar

dCAS9 : dead-CAS9

ddPCR : Droplet Digital PCR

DHS : Distincte Homogène et Stable

DO : densité optique

DSB : Double Strand Breaks, cassures double brin

F/R : Forward / Reverse primers

FAO : Food and Agriculture Organization of the United Nations (<http://www.fao.org/home/en/>)

Floral deep : transformation par trempage des fleurs avec *Agrobacterium*

Frameshift : décalage du cadre de lecture

GA : Gibbéreline

GAI : Gibbereline Insensitive

GFP : green fluorescent protein

Ha : Hectare

HI : Harvest Index, indice de récolte

HDR : Homologous directed repair

HR : Homologous recombinaison

Indels : insertions-délétions (nucléotides, AA...)

kDa : Kilo Dalton

KI : Knock-In

KO : Knock-Out

Mb : millions de bases (ADN)

NbGrPan : Nombre de grains par panicule

NbPan : Nombre de panicules

NBT : New Breeding Technologies

nCAS9 : nickase-CAS9

NGG : nucléotide (A, T, G, ou C) suivi d'un G et d'un autre G

NHEJ : Non-homologous end joining

NLS : Nuclear Localization Signal

NPBT : New Plant Breeding Technology

NUE : Nitrogen Use Efficiency

Off-Targets : coupures hors site

OGM : Organisme Génétiquement Modifié

Os : *Oryza sativa* (riz)

PAM : Protospacer Adjacent Motif

Pc : *Petroselinum crispum* (persil)

PDS : Phytoène Desaturase

PEG : polyéthylène glycol

PI : PAM interacting domain

QTL : Quantitative Trait Loci

rAPOBEC1 : apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like

Rdt : Rendement

RNP : Ribo-Nucléo-Protéiques

SDN : Site-directed nuclease

sgRNA : single guide RNA

SNP : Single Nucleotide Polymorphism

ssp : sous-espèce, subspecies

T-DNA : Transfert DNA, ADN de transfert

TA : Température ambiante

TALE : Transcription Activator-Like

TALEN : Transcription Activator-Like Effector Nucleases

tracrRNA : trans-activating crisprRNA

Ubi : Ubiquitine

UGI : uracile glycosylase inhibitor

Vv : *Vitis vinifera* (vigne)

WT : Wild Type, sauvage

ZFN : Zinc Finger Nucleases

Zm : *Zea mays* (maïs)

Introduction générale

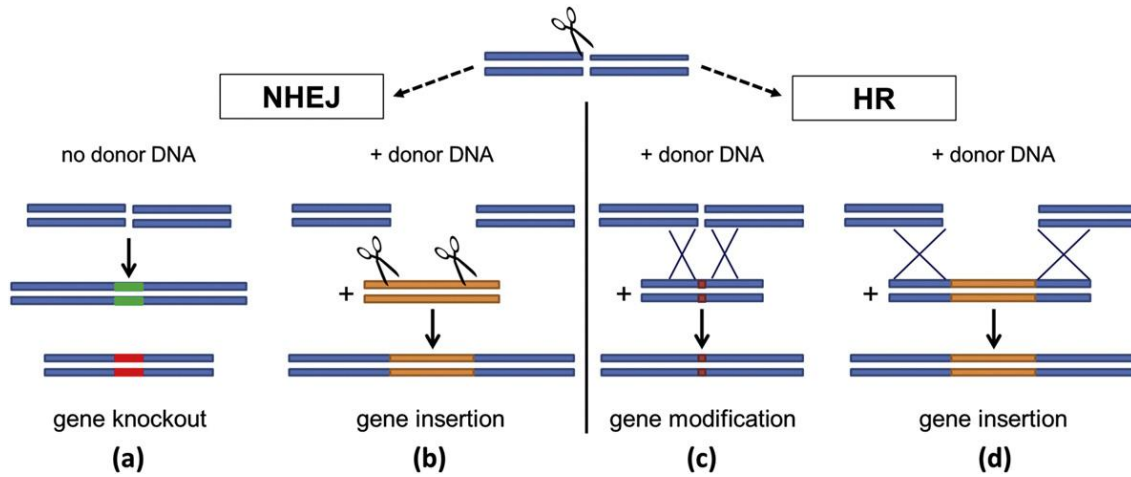
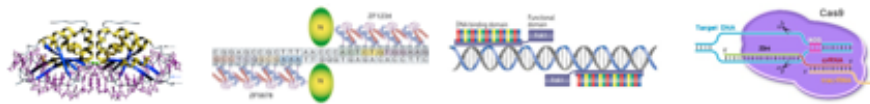


Figure 1 : Principe des réparations de l'ADN par Bortesi et Fisher, 2015

Il existe deux voies pour la réparation d'une cassure double brin sur l'ADN, par NHEJ (Non-Homologous End Joining) ou par HR (Homologous Recombinaison).

La réparation par NHEJ, voie majeure, consiste à rabouter les deux fragments d'ADN ce qui conduit fréquemment à l'apparition de mutations par ajout/suppression de nucléotides (a et b).

La réparation par HR (Homologous Recombinaison), voie mineure, utilise un ADN matrice présentant une homologie avec la séquence ciblée ; elle permet en général d'obtenir une réparation conforme de l'ADN. Cette voie de réparation peut également être utilisée pour insérer une nouvelle séquence (Knock-In, KI) en utilisant des zones d'homologies bordant la zone à insérer (c et d).



	Meganucleases	ZFNs	TALENs	CRISPRs
Nb of proteins required	1	2	2	1 + gRNA
Off-target	Low (target = 18-24 bp)	Rather high (target = 2 x 9 bp)	Low (target = 2 x 10-30 bp)	Rather high (?) (target = 19-23 bp)
Diversity of target sequences	One cutting site every 250 bp	One cutting site every 500 bp	One cutting site every 35 bp	One cutting site every 8 pb
Difficulty of production	Heavy (modification of the peptid chain and selection)	Uneasy (interaction effects on specificity)	Relatively easy (modular assembly)	Very easy (synthesis of one oligonucleotide)
Approximate price	50 000 €	5000 €	1000 €	10 €

Courtesy of F. Nogué, INRA, IJPB

Figure 2 : Différence des prix des technologies par F. Nogué

Dr Fabien Nogué (INRAe Paris) a élaboré un tableau comparatif des prix de développement des quatre NBT précédentes. La technologie CRISPR ne coûte qu'une dizaine d'euros à développer et produire, alors qu'une méganucléase en coûtait plusieurs dizaines de milliers il y a quelques années.

Partie 1 : CRISPR la « New Breeding Technology » de ce début de siècle

L'introduction comprend trois parties. La première partie, comprend un résumé de l'origine et de l'histoire de la technologie CRISPR. Dans une deuxième partie, je décrirai des aspects généraux en biologie, en amélioration génétique, et en agronomie, des deux plantes modèles étudiées : le riz et la vigne. Je décrirai dans la dernière partie les objectifs de la thèse, et en particulier les caractères agronomiques qui seront ciblés pour démontrer l'efficacité de cette technologie pour accélérer l'amélioration génétique.

1) Aperçu historique des technologies d'édition des génomes avant CRISPR-CAS9

La capacité de pouvoir insérer un ADN étranger dans un génome est apparue dans les années 1970 chez les procaryotes. La possibilité d'isoler par restriction enzymatique et de cloner des fragments d'ADN dans des vecteurs bactériens ont posé les bases de la biologie moléculaire actuelle (Hsu et al., 2014). Au même moment, la redécouverte de la capacité de transformation de l'ADN en bactérie a été mise à jour (Mandel and Higa, 1970). Les pionniers de la culture *in-vitro* ont établi, dans les années 60, les premiers protocoles de régénération de plantes à partir de cellules dédifférenciées (Meynard et al., 2020 ; Nishi et al., 1968). Il a fallu attendre les années 80 afin de pouvoir insérer de façon stable un ADN cloné dans le génome de plantes par transformation génétique à l'aide de la bactérie *Agrobacterium* et son plasmide Ti (Zambryski et al., 1983). Il était dorénavant possible d'étudier plus en détail la fonction de gènes d'intérêts par exemple, ou produire des banques de mutants d'insertion T-DNA chez *A. thaliana* ou le riz.

Cette insertion de T-DNA se produit au niveau de cassures doubles brins de l'ADN. Au cours de ces cassures, les mécanismes de réparation cellulaires peuvent emprunter deux voies (**Figure 1**) :

- la réparation non homologue (NHEJ), qui utilise une ligase pour rassembler les deux bouts, est la voie majoritaire chez les plantes (Puchta, 2004), et elle génère des erreurs en retirant ou rajoutant des nucléotides (indels) (Bortesi and Fischer, 2015)
- la réparation homologue (HR), qui utilise une matrice ADN similaire déjà existante pour réparer le fragment (Bortesi and Fischer, 2015).

La HR est une voie dominante chez certaines espèces comme la levure ou la mousse *Physcomitrella patens* (Meynard et al., 2020). Les DSB stimulant la voie HR par un facteur 1000, de nombreuses études qui ont porté sur la possibilité de modifier précisément l'ADN ont continué d'émerger dans les années 90 suite aux travaux sur la capacité de certaines protéines à effectuer des cassures double brin au niveau de l'ADN (Hsu et al., 2014).

L'utilisation de méganucléases a été une des premières méthodes d'édition des génomes par SDN (Site-directed nuclease) dans les années 1990 (Puchta et al., 1993). Les sites de reconnaissance à l'ADN devaient être testés *in-vitro*. Par exemple, la méganucléase I-SceI à un site de reconnaissance de 18 nucléotides, à comparer à une enzyme de restriction dont le site de reconnaissance fait 6 nucléotides. Cependant, le design de nouvelles nucléases reconnaissant un site spécifique était long et coûteux à produire (Meynard et al., 2020).

Une autre technologie développée dans les années 2000 était basée sur l'utilisation des conformations protéiques 3D dites doigts de zinc (ou zinc fingers) couplées à une endonucléase FokI pour former les Zinc Finger Nucleases (ZFN) (Kim et al., 1997). Plusieurs doigts de zinc consécutifs sont produits par le dédoublement d'acides aminés Cystéine et Histidine (Cys₂-His₂), et chaque doigt de zinc peut reconnaître 3 nucléotides. La chaîne qu'ils vont former sera fusionnée à une endonucléase FokI, ce qui permettra de la guider jusqu'à une séquence spécifique de 9 à 12 nucléotides sur l'ADN pour y effectuer une cassure simple brin (Bortesi and Fischer, 2015; Hsu et al., 2014 ; Meynard et al., 2020). Il est nécessaire de dimériser le ZFN pour pouvoir réaliser un DSB. Malheureusement, la spécificité des ZFN est difficile à prédire et nécessite de réaliser un grand nombre de constructions ZFN et de tests *in-vitro* (Bortesi and Fischer, 2015).

Une des seules technologies mature développée peu avant l'avènement des CRISPR a été les « TALEN ». Les « TALE ou TAL (Transcription Activator-Like) effectors » sont des protéines responsables de la pathogénicité de la bactérie *Xanthomonas* sur le riz, ainsi que sur d'autres espèces (poivre, tomate, soja, citron, coton, prune, cerise ...). Les TALE ont la capacité de se fixer et d'activer l'expression de gènes des canaux d'efflux de sucre *SWEET* de la plante (Boch et al., 2014; Bortesi and Fischer, 2015; Hsu et al., 2014; Meynard et al., 2020), favorisant l'infection. Le TALE se compose d'un domaine de fixation à l'ADN formé de répétitions de 33 à 35 AA, avec pour chaque répétition, deux AA hypervariables, les n° 12 et 13 (ou repeat variable diresidue ; RVD). La combinaison de deux AA particuliers, permet d'assurer la spécificité de reconnaissance d'une base de l'ADN (A, T, G, ou C). La répétition de plusieurs TALs avec des AA 12 et 13 différents permet de cibler des séquences de 15-20 nucléotides

(Meynard et al., 2020). Les Transcription Activator-Like Effector Nucleases (TALENs) reprennent le principe des TALE de *Xanthomonas* fusionnés à une nucléase FokI (Bortesi and Fischer, 2015; Hsu et al., 2014). Malgré la très haute spécificité de la construction TALs-FokI (TALEN) et la démonstration de leur efficacité pour induire des DSB, cette technologie nécessite d'utiliser des dimères, comme les ZFN. Le clonage des TALEN reste laborieux, bien que moins coûteux que les ZFN (Bortesi and Fischer, 2015; Hsu et al., 2014) (**Figure 2**). Enfin, les TALEN ont fait la preuve de leur efficacité et de leur spécificité.

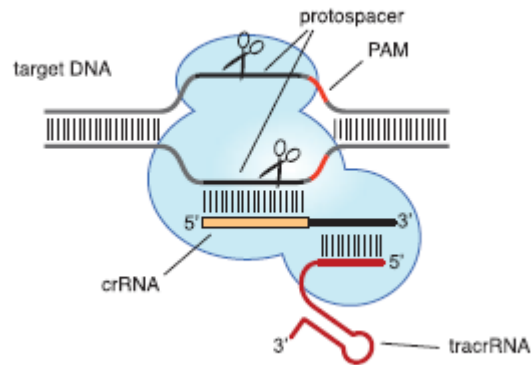
2) La technologie CRISPR-CAS9 : un bref rappel historique d'une technologie d'édition révolutionnaire

La technologie CRISPR-CAS9, pour « Clustered Regulatory Interspaced Short Palindromic Repeats associated protein CAS9 », est dérivée d'un système de défense bactérien contre les phages et les ADN invasifs.

En 1987, les 1^{res} répétitions palindromiques dans l'ADN ont été observées chez *E. coli*. Ces répétitions forment des structures stables en tige-boucle une fois transcrites (Ishino et al., 1987). Ce n'est qu'en 2002 que le nom de locus CRISPR (Clustered Regulatory Interspaced Short Palindromic Repeats) leur sera attribué (Jansen et al., 2002). S'en suivra la découverte des gènes *CAS*, situés à proximité des séquences CRISPR sur l'ADN (Hsu et al., 2014 ; Jansen et al., 2002). La fonction d'immunité bactérienne a été mise en évidence en 2005, accompagnée de la fonction de mémoire des infections contre les bactériophages (Mojica et al., 2005). Les bactéries étaient capables d'insérer des fragments de séquences virales (crRNA ou crispr RNA) à l'intérieur de ces séquences palindromiques. L'ensemble de ces séquences forme une mémoire des invasions passées. Les séquences CRISPR permettent de cibler les ADNs viraux lors d'une nouvelle infection virale une fois associées à une endonucléase, la CAS9 (Barrangou et al., 2007 ; Bolotin et al., 2005).

Le complexe CAS9-guide ne doit couper que l'ADN viral et ne pas couper son propre site crRNA lors d'une attaque : il a été mis en évidence un protospacer adjacent motif (PAM) en aval des 20 pb du protospacer sur l'ADN viral. Le PAM est représenté par 3 nucléotides consécutifs : NGG. Il n'est pas présent dans les locus CRISPR, ni dans aucune séquence bactérienne, mais uniquement sur l'ADN viral, formant un système permettant de différencier le soi du non-soi (Deveau et al., 2008). C'est le PAM qui va permettre l'activation des deux

Cas9 programmed by crRNA:tracrRNA duplex



Cas9 programmed by single chimeric RNA

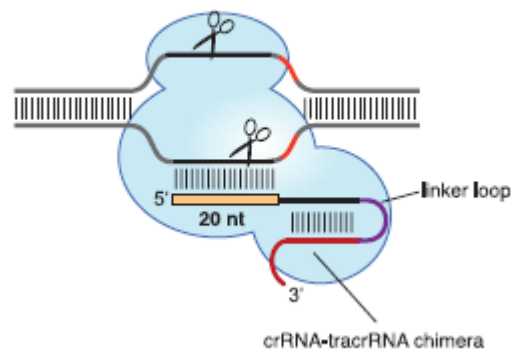


Figure 3 : Fonctionnement du système CRISPR CAS9 d'après Jinek et al., 2012

Le système CRISPR-CAS9 se compose de deux acteurs : le sgRNA et la CAS9. L'ARN guide (sgRNA) se compose d'une partie fixe en 3' (tracrRNA) et une partie variable de 20 nucléotides en 5' (crRNA) qui permet de cibler une zone précise de l'ADN (protospacer). Lorsque le complexe CAS9-sgRNA est formé il scanne l'ADN jusqu'à trouver la cible, qui doit être suivie d'une séquence PAM (protospacer adjacent motif, NGG). Les domaines RuvC et HNH de la CAS9 vont alors induire la cassure en 5' et 3' de l'ADN cible par changement de conformation du complexe.

sites nucléases, RuvC et NHN de la CAS9 pour la coupure double brin (Nishimasu et al., 2014).

La mécanistique et le rôle de chacun des acteurs du système CRISPR/CAS9 ont été décrits au début de la décade 2010-2020. Il a d'abord été démontré que la CAS9 était la nucléase responsable de la coupure double brin de l'ADN (Garneau et al., 2010). Puis, un an après, le rôle de l'hybridation du crRNA (crispr RNA) avec un autre ARN, le tracrRNA (trans-activating crisprRNA) pour guider la CAS9 vers sa cible (Deltcheva et al., 2011) (**Figure 3**, haut). Enfin, les 1^{ers} tests pour transférer le système CAS9 entre bactéries ont été concluants. La fusion du crRNA du tracrRNA en un seul ARN, appelé sgRNA (pour Single Guide RNA) a permis de simplifier le système et ouvert la porte aux premières applications (Jinek et al., 2012) (**Figure 3**, bas).

3) La technologie CRISPR-CAS9 « pour les nuls »

La technologie est basée sur l'utilisation d'une nucléase CAS9 qui va se lier à un ARN guide, le sgRNA, grâce auquel elle sera guidée vers sa séquence cible (Doudna, 2017) (**Figure 3 et 4**). L'ARN guide est composé de deux parties, le tracrRNA qui permet de lier la nucléase CAS9 et le crRNA (Bortesi and Fischer, 2015; Doudna, 2017). Le crRNA possède quant à lui une partie variable de 20 pb complémentaire de la zone ciblée. Aujourd'hui, c'est une des méthodes les plus simples et rapides à mettre en œuvre pour induire des mutations à un site spécifique dans le génome.

Grâce aux 20 premiers nucléotides en 5' de l'ARN guide (la partie variable de son crRNA), la CAS9 va être guidée sur le protospacer, qui est la séquence d'ADN ciblée. En aval du protospacer, une séquence NGG (le PAM) est nécessaire pour que le complexe CAS9/sgRNA/protospacer devienne actif (Nishimasu et al., 2014) lors de la reconnaissance par le crRNA de la séquence ciblée. Les sites catalytiques RuvC et HNH de la CAS9 vont induire des cassures simples brins, respectivement en 5' et 3' de l'ADN ciblé, et leur activation conjointe va conduire à une cassure double brin de l'ADN, trois paires de bases avant le PAM. Les mécanismes de réparation de la cellule vont alors entrer en jeu. Si la réparation est identique à la séquence de départ, le complexe CAS9 sgRNA va alors de nouveau induire un DSB au même locus jusqu'au moment où les systèmes de réparation, qui ont un taux d'erreur non nul, introduisent une mutation par délétion, insertion, et/ou substitution. Si ces mutations se produisent dans une séquence codante, elles peuvent entraîner des mutations perte de fonction (Bortesi and Fischer, 2015). Si elles se produisent suffisamment tôt, la plupart de

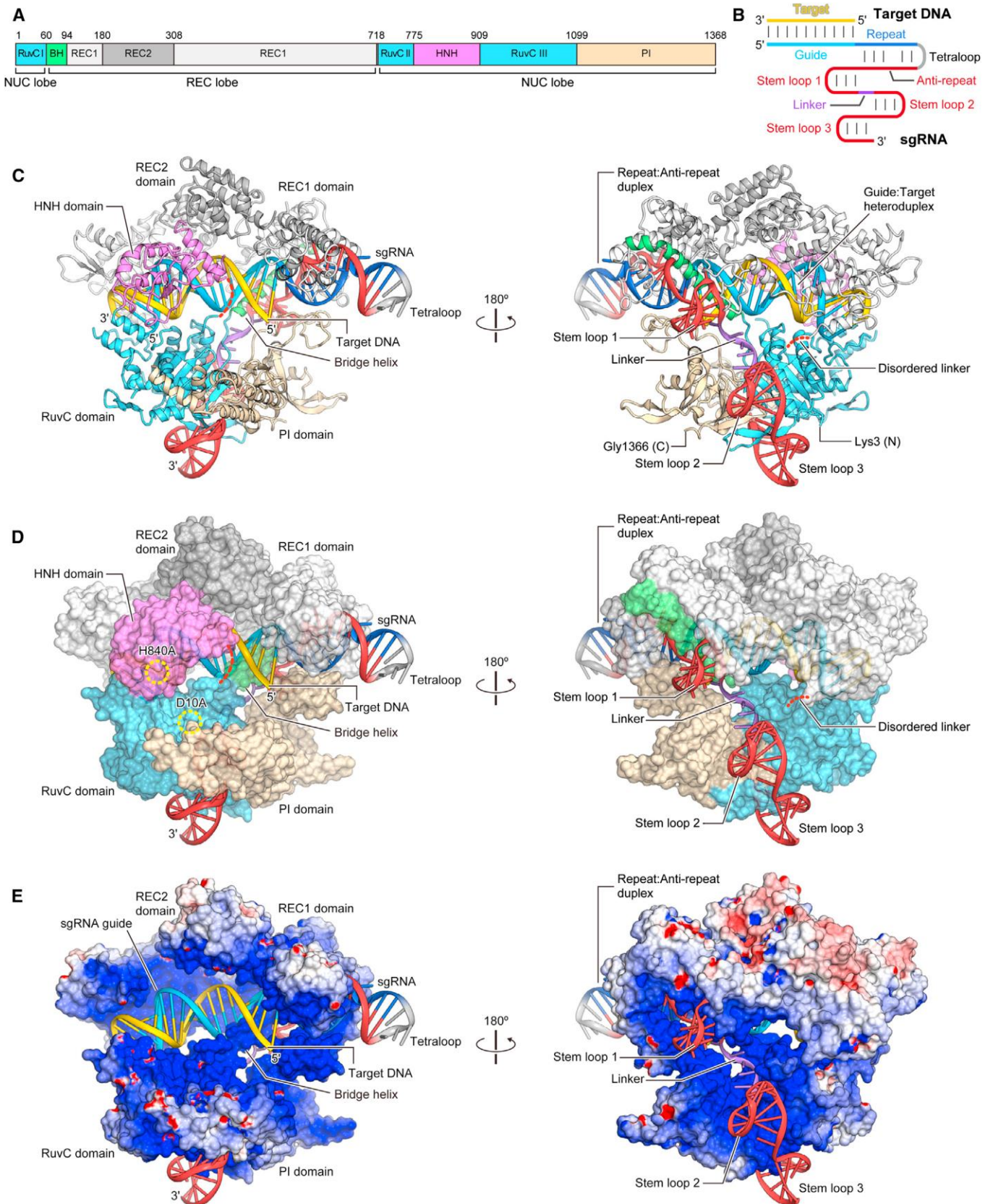


Figure 4 : Structure 3D du complexe CRISPR CAS9 d'après Nishimasu et al., 2014
 Nishimasu et al. ont décrit la structure 3D de la CAS9 complexée avec son guide. La séquence ADN présente les 2 parties protéiques de la CAS9 et la position des deux enzymes de coupure (A), et le sgRNA présente ses 3 parties pour le ciblage et recrutement (B). Les deux côtés de la CAS9-sgRNA sont représentés en feuillets (C), ou par structure en 3D (D, E).

ces mutations peuvent se transmettre au travers des générations lorsque les cellules germinales sont ciblées (Woo et al., 2015). L'utilisation de la technologie CAS9 permet donc d'induire des mutations et des insertions à théoriquement n'importe quel locus et a permis de faciliter grandement l'analyse fonctionnelle chez de nombreux organismes. Les études sur l'insertion par KI ou le remplacement de séquences à l'aide d'un ADN de réparation ont été ensuite grandement augmentées et facilitées avec l'arrivée de CRISPR, peu après l'avènement des premières génération de SDN comme les TALEN (Meynard et al., 2020). La CAS9 est une protéine procaryotique (Hsu et al., 2014) et pour son utilisation en cellules eucaryotes plusieurs modifications ont été introduites. Des séquences d'adressage au noyau NLS (Nuclear Localization Signal) ont été ajoutées à la CAS9. La CAS9 est également sous le contrôle d'un promoteur constitutif comme le promoteur CMV en cellule humaine, le promoteur CaMV 35S, un promoteur de type ubiquitine pZmUbi chez les céréales (Meynard et al., 2020 ; Miao et al., 2013) ou PcUbi4.2 chez les dicotylédones (Fauser et al., 2014). Les sgRNA sont exprimés en général grâce à des promoteurs snRNA (Small-Nuclear-Ribonucleo-protein-Particules) d'ARN non codant comme U3 et U6 chez les monocotylédones (Miao et al., 2013), et les dicotylédones (Fauser et al., 2014). Ces promoteurs permettent d'exprimer des snRNA, qui sont des petits ARNs nucléaires non codants transcrits par la Polymérase III (Meynard et al., 2020 ; Miao et al., 2013) qui jouent un rôle dans la maturation et le clivage des ARN ribosomiques.

Pour permettre la transformation stable *in-planta* ou en cellules animales, plusieurs systèmes peuvent être utilisés. Au préalable, il est aussi possible de faire de la transfection (protéines) ou transformation (d'ADN) en protoplastes (cellules végétales sans parois) en transitoire. Cela permet de tester l'efficacité de beaucoup de constructions différentes en seulement quelques jours avant de les transformer stablement (Bès et al., 2021 ; Hichri et al., 2010 ; Woo et al., 2015). En cellules animales, il est possible d'utiliser les virus (*Lentivirus*), les micro-injections ou les électroporations, pour ne citer qu'eux (Liu et al., 2017). En plante, la biolistique ou la transformation par *Agrobacterium tumefaciens* peuvent être utilisées sur des cals embryogènes ou des embryons immatures (Svitashev et al., 2016, Sathishkumar et al., 2019). La biolistique permet de travailler avec de l'ADN et/ou des protéines (complexes Ribo-Nucléo-Protéiques (RNP) par exemple) complexés à des microprojectiles (billes d'or) pour les transformer. La biolistique a cependant plusieurs problèmes. Elle induit de nombreuses cassures d'ADN génomique et crée des mutations hors site (Harkess, 2019), et elle peut être responsable de l'insertion de copies multiples de l'ADN dans le génome, ou de parties tronquées des plasmides (Meynard et al., 2020).

La transformation par agrobactéries est la technologie la plus populaire, soit par mise en contact des agrobactéries avec des embryons ou des cals (Sallaud et al., 2003, 2004), ou par « floral dipping » (Clough and Bent, 1998). En plante, les premiers essais d'utilisation de la technologie CRISPR/CAS9 ont été réalisés en ciblant des gènes donnant un phénotype facile à mettre en évidence *in-vitro*. Le gène *PDS* codant pour l'enzyme phytoène désaturase, conduit à un phénotype albinos chez les plantes mutées. Il a donc été le premier ciblé chez *Arabidopsis* (Li et al., 2013) et sur la vigne (Nakajima et al., 2017). Sur le riz *OsCAOI* (Chlorophylle A oxygénase 1) a été ciblé car il affecte la photosynthèse (phénotype jaune pâle) mais contrairement à la mutation *PDS*, il n'est ni létal, ni stérile (Herbert et al., 2020 ; Woo et al., 2015).

Une autre cible très étudiée chez les plantes pour la mise en place de stratégies d'édition est le gène *ALS* (Acetolactate syntase). C'est une cible des herbicides de type sulfonyleurea. Une modification précise d'un acide aminé (tryptophane 548 en Leucine, ou serine 627 en isoleucine) induit une résistance de la plante aux herbicides, un système intéressant pour tester l'efficacité de construits et de systèmes de réparation par HDR (Homologous Directed Repair) (Meynard et al., 2020) en facilitant la sélection de mutants édités devenus résistants à l'herbicide (Meynard et al., 2020).

4) Avantages de la technologie

La technologie CRISPR présente de nombreux avantages comparativement aux ZFN, TALEN, ou Méganucléases. Tout d'abord son coût de revient ; en effet les plasmides CRISPR pour monocotylédones ou pour dicotylédones sont disponibles sur des sites comme addgene (<https://www.addgene.org/>) à faible coût. Le clonage des 20 pb du crRNA dans le sgRNA, puis le transfert du guide et de son promoteur dans le vecteur binaire qui est utilisé pour la transformation (Bès et al., 2021), sont des étapes à la portée de tous les laboratoires de biologie moléculaire en partant d'un couple d'oligonucléotides et d'enzymes de restriction/ligases/recombinases. Le système CRISPR/CAS9 est un système modulaire, facilement réutilisable comparativement, par exemple, aux méganucléases (Bortesi and Fischer, 2015; Hsu et al., 2014). Un second avantage est la facilité de multiplexer le système CRISPR, c'est-à-dire d'avoir plusieurs guides ciblant plusieurs locus portés par une même construction. En effet il est possible d'intégrer plusieurs sgRNA sur le même plasmide qui seront pris en charge par une seule CAS9 (Bortesi and Fischer, 2015; Hsu et al., 2014; Ran et al., 2013). Par

Nuclease	Feature of improvements	PAM	Major purpose of improvements
SpCas9	NLS-SpCas9-NLS structure, <i>Streptococcus pyogenes</i> derived Cas9 protein	NGG	Classical Cas9
SaCas9	<i>Staphylococcus aureus</i> derived Cas9 protein replacement	NNGRRT	Reduce size
SaKKHCas9	E782K/N968K/R1015H in SaCas9 protein	NNNRRT	Reduce size and extend PAM
StI Cas9	<i>Streptococcus thermophiles</i> derived Cas9 protein replacement	NNAGAAW	Improve specificity
NmCas9	<i>Neisseria meningitidis</i> derived Cas9 protein replacement	NNNNGMTT	Improve specificity
AsCpf1	<i>Acidaminococcus</i> derived nuclease	TTTN	Increase the efficiency of HDR
LbCpf1	<i>Lachnospiraceae</i> derived nuclease	TTTN	Increase the efficiency of HDR
VQR-SpCas9	D1135V/R1335Q/T1337R in SpCas9	NGA	Extend PAM
EQR-SpCas9	D1135E/R1335Q/T1337R in SpCas9	NGAG	Extend PAM
VRER-SpCas9	D1135V/G1218R/R1335E/T1337R in SpCas9	NGCG	Extend PAM
SpCas9-HF1	N497A/R661A/Q695A/Q926A in SpCas9	NGG	Improve specificity
eSpCas9	K848A/K1003A/R1060A in SpCas9	NGG	Improve specificity
HeFSpCas9	N497A/R661A/Q695A/K848A/Q926A/K1003A/R1060A in SpCas9	NGG	Improve specificity
EvoSpCas9	M495V/Y515N/K526E/R661Q in SpCas9	NGG	Improve specificity
HypaCas9	N692A/M694A/Q695A/H698A in SpCas9	NGG	Improve specificity
Sniper-Cas9	F539S/M763I/K890N in SpCas9	NGG	Improve specificity
xCas9(3.7)	A262T/R324L/S409I/E480K/E543D/M694I/E1219V in SpCas9	NG/GAA/GAT	Extend PAM
SpCas9-NG	R1335A/L1111R/D1135V/G1218R/E1219F/A1322R/T1337R in SpCas9	NG	Extend PAM

Figure 5 : Table des différentes CAS9 et leurs PAM associées d'après Li et al., 2019

Différentes CAS9 ont été mises au point par de simples mutations sur des domaines de la protéine favorisant ainsi la réduction des OFF-Targets ou la spécificité de ciblage en modifiant le PAM. Parmi elles on retrouve en particulier les Enhanced CAS9 (eSpCAS9, HF1, HeFS ...).

La CPF1 est présente dans le tableau par ses capacités de coupure à l'aide d'un PAM et un « sgRNA » différent. D'après Li et al, 2019.

un système naturel de « processing » des ARN guides, c'est-à-dire auto-séparation naturelle par clivage des différents sgRNA transcrits, comme en bactérie. Le record actuel de mutations obtenues simultanément est de 13200 par la technique de l'édition à la base près (Base Editing, cf plus loin) en cellule humaine (Smith et al., 2019).

Il est également possible de travailler avec des complexes Ribo-Nucléo-Protéiques préassemblés de CAS9-sgRNA purs, au lieu de construits ADN, principalement sous forme de plasmides. Ces complexes peuvent être transformés en protoplastes à l'aide du PEG (polyéthylène glycol), chez les monocotylédones ou les dicotylédones (Malnoy et al., 2016 ; Woo et al., 2015), ou bien par bombardement (Svitashev et al., 2016).

Un avantage non négligeable, dérivant de la capacité à multiplexer les cibles, de la technologie CRISPR est l'édition simultanée des différents copies d'un gène chez les diploïdes et polyploïdes. En effet, en utilisant un guide commun à tous les hétéro-ou homo-allèles ou en utilisant une construction multiplexe, la technologie CRISPR/CAS9 permet de faciliter l'analyse fonctionnelle de gènes chez des espèces polyploïdes, ce qui n'était pas possible avec les méthodes précédentes (Sathishkumar et al., 2019) ouvrant la voie à l'analyse fonctionnelle chez des espèces d'intérêt agronomiques aux génomes complexes comme le blé ou la canne à sucre.

5) Limites à l'utilisation de la technologie CRISPR/CAS9

Le 1^{er} point important est la nécessité de la présence d'un PAM adjacent aux 20 pb ciblés en 3' pour que la CAS9 puisse induire des DSB (Bortesi and Fischer, 2015; Jinek et al., 2012). Ce PAM n'est pas présent dans la séquence du guide, mais doit impérativement être adjacent en 3' de la séquence ciblée. Il existe différentes nucléases de type CAS qui reconnaissent différents PAM ; chez la CAS9 ce dernier est lu par le PI (PAM interacting domain) (Hsu et al., 2014 ; Li et al., 2019). Plusieurs équipes ont muté ce PI afin d'obtenir des protéines CAS9 avec des spécificités différentes, capables de reconnaître d'autres PAM. On retrouve ainsi des protéines CAS9 avec des domaines PI modifiés qui reconnaissent des PAM comme NGA, NGCG, ou bien plus beaucoup plus longs comme la SaKKHCAS9 avec NNNNGMTT (M= A ou C) (**Figure 5**) (Li et al., 2019 ; Nishimasu et al., 2014).

La structure du sgRNA elle-même est un élément clef de l'efficacité du complexe CAS9. Les 12 premiers nucléotides avant le PAM sont les plus importants (Pattanayak et al., 2013) pour la

spécificité de reconnaissance, on parle de séquence « seed ». En règle générale, toute différence dans ces douze premiers nucléotides affecte fortement l'interaction CAS9-sgRNA-séquence cible (Jinek et al., 2012). Ensuite le guide est plus stable si la partie tracrRNA est allongée, par exemple en la faisant passer de 48 à 84 nucléotides, et/ou si des structures secondaires formant des boucles dans l'ARN sont ajoutées à la suite en 3' pour augmenter sa stabilité (Hsu et al., 2014).

Les mutations hors site (ou off-Targets) sont des mutations induites à d'autres endroits du génome que la séquence ciblée. La CAS9 peut tolérer jusqu'à 5 mismatches dans son crRNA au niveau du protospacer à condition qu'ils soient éloignés du PAM (Hsu et al., 2014). La probabilité d'induire des cassures double brin hors site est donc élevée, si la cible n'est pas choisie judicieusement (Bortesi and Fischer, 2015). Des « enhanced » CAS9 ont vu le jour (eSpCAS9, HF1, HeFS) (Kulcsár et al., 2017) avec des capacités réduites de off-Targets en jouant sur certains acides aminés (**Figure 5**) (Li et al., 2019).

Une première approche consiste à utiliser des spacers les plus spécifiques possibles. De nombreux algorithmes et bases de données ont été développés dans ce but (Li et al., 2019). Li et al (2019) ont recensé plusieurs logiciels comme Target Finder (<http://crispr.mit.edu/>), CasFinder (<http://arep.med.harvard.edu/CasFinder/>), et Cas-OFFinder (<http://www.rgenome.net/cas-offfinder/>) (Doench et al., 2016). Cependant cette utilisation « virtuelle » peut-être biaisée ; Li et al (2019) proposent donc plusieurs technologies pour aller expérimentalement étudier ces off-Targets. Il est possible de faire un séquençage complet (WGS), de la CHIPseq (Chromatine immunoprécipitation puis séquençage) sur la CAS9 sur l'ADN, ou entre autres du CIRCLE-seq après avoir circularisé le génome et provoqué des coupures off-Targets avec des RNP (Akcakaya et al., 2018 ; Li et al., 2019).

Un autre problème enfin, qui affecte en partie l'efficacité de la technologie CAS9, est dû à l'efficacité des technologies de transfection, qui sont variables en fonction des espèces, et la formation de mutations mosaïques. Il arrive donc que lors de la régénération ou de la croissance, toutes les cellules ne soient pas mutées ou ne présentent pas les mêmes mutations. Chez les animaux, en particulier lors de l'utilisation de la technologie à partir d'embryons, cela reste un frein important à l'utilisation de cette technologie (Hsu et al., 2014).

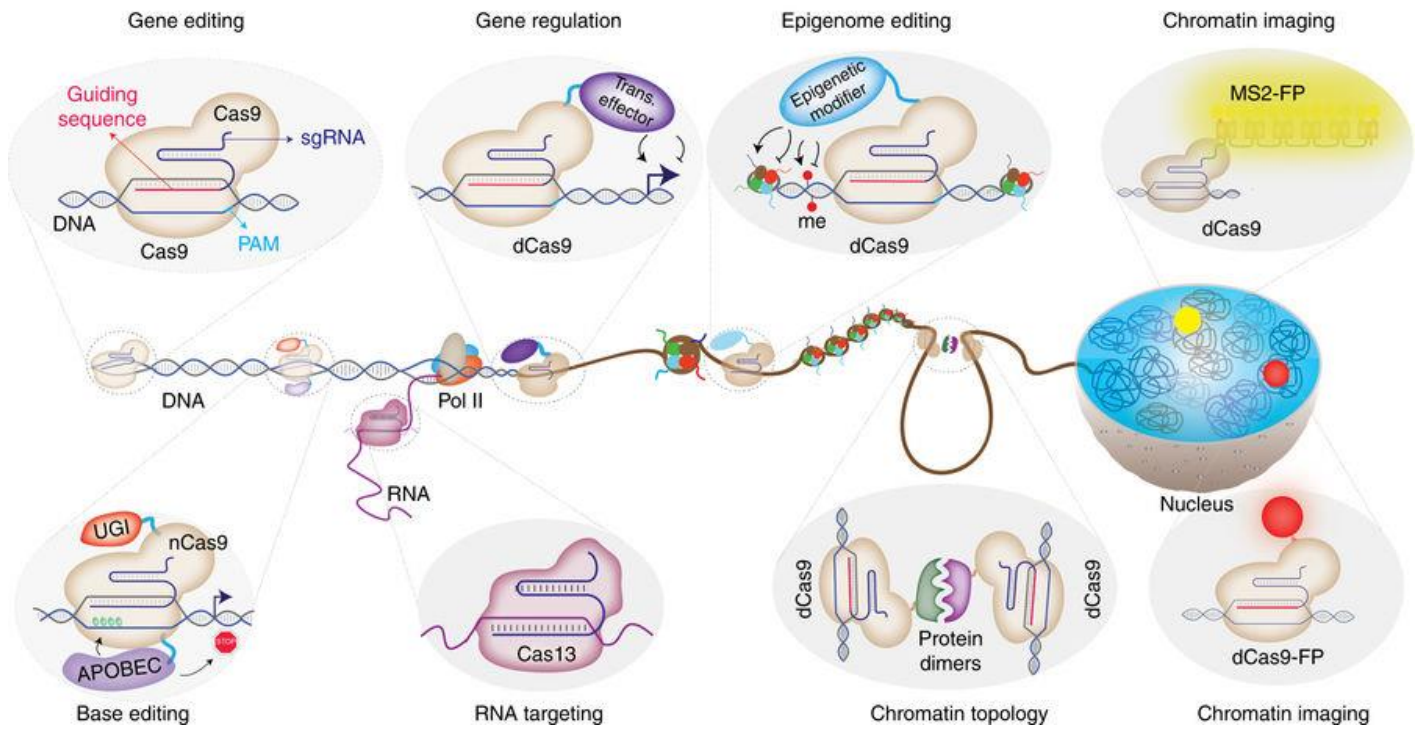


Figure 6 : Listing des nouveaux NBT CRISPR par Adli, 2018

Plusieurs nouvelles technologies dérivées de CRISPR-CAS9 ont émergées. Adli en a représenté quelques-unes. A partir de possibilité de se servir de la CAS9 comme cargo, il est possible d'activer par effecteur un gène, de jouer sur les marques épigénétiques, d'aller marquer des conformation de l'ADN, de faire des hybridations in vivo, cibler l'ARN avec une autre endonucléase, ou aller faire du Base Editing.

6) Nouvelles technologies dérivées de CRISPR : Multiplexing, KI, remplacement, Base Editing, et modulations transcriptionnelles ou chromatiniques

L'idée d'utiliser la capacité de reconnaissance de la CAS9 pour amener toute une série d'effecteurs à un locus donné a vite émergé grâce à sa capacité de cargo (**Figure 6**).

Nous avons brièvement parlé du cas du multiplexing plus haut. Non seulement ce dernier permet d'étudier plusieurs gènes à la fois, mais il permet aussi de faire des délétions en utilisant deux cibles afin de plus facilement étudier une région ou un domaine plus ou moins étendu. Par exemple, chez le riz des délétions jusqu'à 245 kb (Sathishkumar et al., 2019) ont été obtenues, en protoplastes ; mais dans la plupart des cas on obtient des délétions plus courtes et l'efficacité de délétion est inversement proportionnelle à la taille de la délétion (Saika et al., 2019 ; Zhou et al., 2014).

L'idée de faire du Knock-In (KI) pour insérer un ADN à un locus précis à l'aide des nucléases dirigées a été envisagée dès la mise au point des ZFN, en raison de la stimulation de HR induite au niveau d'un site de cassure double brin (Shukla et al., 2009). Le mécanisme utilisé dans ce cas lors de la réparation de la coupure se fait par NHEJ ou HDR à l'aide d'une matrice ADN, dite « de réparation », qui porte en 5' et 3' des brins homologues à la séquence. L'insertion par NHEJ est plus fréquente, mais peut introduire des mutations non désirées au locus coupé ou également conduire à l'insertion complète de l'ADN de réparation (Svitashev et al., 2016). L'insertion par HDR est quant à elle beaucoup plus rare (Svitashev et al., 2016). Par exemple, par bombardement de RNP et/ou un ADN de réparation il a été possible de remplacer précisément le gène *ALS2* chez le maïs en modifiant les deux acides aminés responsables de la résistance à l'herbicide sur l'ADN de réparation (Svitashev et al., 2016). A l'exception de la mousse *Physcomitrella patens*, chez les plantes, la réparation par NHEJ reste majoritaire, ce qui limite la probabilité d'obtenir des KI (Nakamura et al., 2021). Les taux d'efficacité restent inférieurs à 2 % en moyenne chez la plupart des espèces de plantes où des remplacements par HDR ont été décrits. De nombreux travaux et recherches portent sur l'impact de différents paramètres sur le taux de KI, comme la taille des zones d'homologies, la disponibilité de l'ADN, le type de nucléase CAS, et la spécificité du sgRNA, liste non exhaustive (Collonnier et al., 2017). Le remplacement d'un gène déjà présent par un nouveau est lié à la méthodologie du KI, mais avec la difficulté supplémentaire qu'il faut retirer le gène sauvage et remplacer par le gène muté en même temps, ou plus simplement couper-insérer. Peu de résultats ont été publiés et les taux restent extrêmement bas (Li et al., 2018).

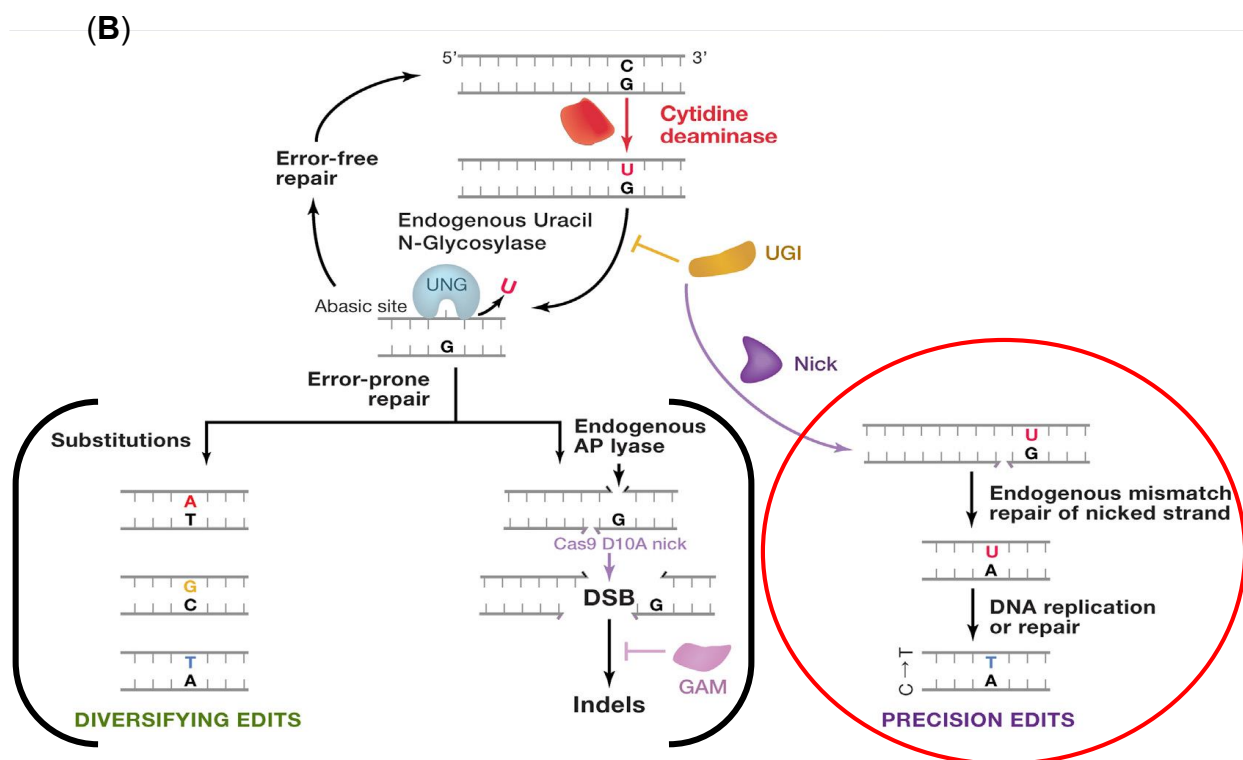
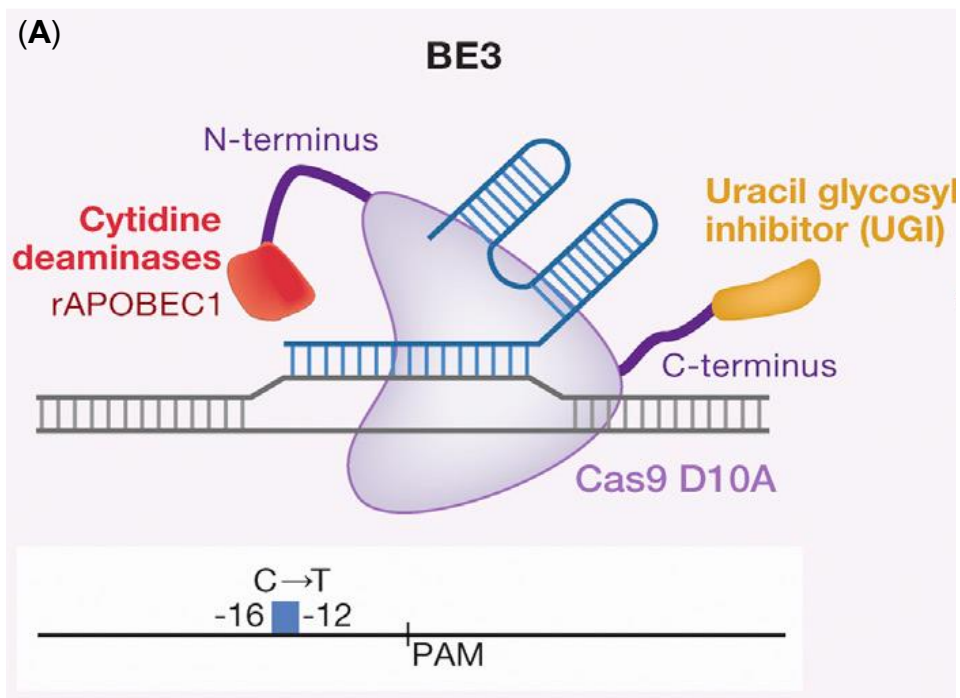


Figure 7 : Technologie CRISPR Base-Editor nCas9 d'après Hess et al., 2017

La technologie Base-Editor utilise une protéine nickase nCAS9 (site catalytique RuvC inactivé d'une CAS9). Elle est couplée à une cytidine deaminase rAPOBEC1 (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like) en N-Terminal, et une uracile glycosylase inhibitor (UGI) en C-Terminal bloquant l'excision de l'uridine et favorisant le Base Editing Repair (BER). Cette version est dénommée BE3.

La mutation par substitution C en T aura lieu non plus côté 3' mais dans la fenêtre du 4^{ème} au 8^{ème} nucléotide côté 5' du crRNA (A). La conversion du C-G en T-A consiste à l'induction de l'excision et réparation d'une base pour produire un mésappariement U-G transformé en U-A puis T-A (à droite, entouré en rouge) (B). Cependant il peut y avoir des erreurs de réparation amenant à une substitution non désirée ou une délétion (à gauche) (B).

Le « Base Editing (BE) », ou édition à la base près, est une des technologies dérivées de CRISPR/CAS9 sur laquelle de nombreux travaux ont porté ces dernières années. Elle se base sur la faculté à convertir une base en une autre de manière irréversible. La CAS9 a un de ses sites RuvC inactivés. Cette dernière se trouve alors transformée en nickase (nCAS9) capable de n'effectuer plus qu'une coupure simple brin. Cela permet toujours à la nCAS9 de continuer à stimuler les systèmes de réparation de l'ADN sans cassures double brins. Elle est couplée à une cytidine deaminase rAPOBEC1 (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like) en N-Terminal, et dans la plupart des versions récentes, à une uracile glycosylase inhibitor (UGI). Originellement, cette technologie a permis de convertir une Cytosine en Thymidine (Komor et al., 2016). Lorsque le complexe arrive sur site, la cytidine deaminase rAPOBEC va conduire à la conversion d'une cytosine en une uridine (base ARN). La glycosylase inhibitor UGI va bloquer l'excision de l'uridine par les mécanismes endogènes d'Uracyl-N-Glycosylase en favorisant le Base Editing Repair (BER). L'uridine sera remplacée par une thymidine durant la phase de réplication de l'ADN (**Figure 7**). On parle alors de transversion C en T. Cette version du Base Editor, qui est dénommée BE3, fonctionne chez les animaux et les plantes à condition d'avoir été codon optimisée (Hess et al., 2017 ; Komor et al., 2016 ; Zong et al., 2017). Le sgRNA utilisé est donc le même que pour la CAS9 « classique », cependant la « mutation » aura lieu non plus côté 3', mais dans la fenêtre du 4^e au 8^e nucléotide côté 5' du crRNA. Cette technologie a un champ d'application thérapeutique considérable, comme pour la maladie d'Alzheimer (Komor et al., 2016), l'amélioration de traits d'intérêt agronomique (Lu and Zhu, 2017), la création de codons STOP par iSTOP pour effectuer des mutations « propres » à la base près en affectant seulement l'étape de la traduction des ARN chez les animaux (Billon et al., 2017) ou chez les plantes (Herbert et al., 2020), et l'évolution de la fonction de protéines par modifications ciblées de leurs structures et des interactions protéine-protéine ou protéine-ADN (Hess et al., 2017). Les premières expérimentations conduites chez le riz ont permis d'obtenir des taux de transversion C en T autour de 7 % et des taux de transversion de plusieurs C en T autour de 0.31 % à 12 % en position 4 à 8 du protospacer (Zong et al., 2017). Dorénavant, suivant la cible, les taux de substitution peuvent monter à plus de 50 % chez le riz en fonction du % en GC du crRNA et de l'utilisation de plasmides codons optimisés (Herbert et al., 2020).

Le BE a récemment été modifié pour permettre de convertir des Adénines. L'avantage de cette méthode ABE (Yan et al., 2018) par rapport à l'utilisation du système C-T est la plus grande spécificité qui évite les conversions off-Targets de C en T. En effet le système ABE fonctionne

uniquement sur des duplexes ARN, ou ADN/ARN, tandis que le BE C-T convertira dès qu'il y aura coupure simple brin (Hua et al., 2019 ; Yan et al., 2018).

Sur les mêmes principes de fusionner la CAS9 avec d'autres protéines comme pour le BE, il est possible de construire des facteurs de transcription artificiels en exploitant la capacité de la dead-CAS9 et en la fusionnant à des activateurs ou à des répresseurs (Bortesi and Fischer, 2015). Avec ses deux sites catalytiques mutés, la dCAS9 fusionnée va reconnaître un locus spécifique. La fixation du complexe sur sa cible n'occasionnera pas de coupure, mais va permettre l'activation ou la répression d'un gène, en fonction de la protéine effectrice fusionnée. Il est possible de multiplexer le système pour activer ou réprimer (l'un, mais pas l'autre, le complexe dCAS9 étant commun) plusieurs locus simultanément (Piatek et al., 2015).

Dans une approche similaire, il est possible de fusionner des déméthylases spécifiques à la Lysine 27 de l'histone 3 (H3K27) avec une dCAS9 pour forcer la décondensation chromatinienne et l'activation d'un gène (Bortesi and Fischer, 2015; Hsu et al., 2014). Cet outil permettrait de faciliter l'étude des mécanismes de la régulation chromatinienne comme les possibles bivalences H3K27me3-H3K4me3 pour certains gènes dont l'expression peut varier drastiquement (Bellegarde et al., 2018).

Une autre approche est l'utilisation d'une dCAS9-GFP, ou autres protéines fluorescentes, pour faire du FISH in vivo (Bortesi and Fischer, 2015). En utilisant plusieurs guides à la suite sur une séquence, il est possible de visualiser les complexes dans le noyau, et permettre l'apparition d'une forte fluorescence à certains endroits spécifiques du génome. Ainsi la conformation de l'ADN sur certains endroits dans le noyau peut être étudié suivant les conditions de culture des cellules par exemple (Chen et al., 2013).

Pour finir sur l'utilisation d'autres endonucléases, la CAS12a, plus connue sous le nom de CPF1 (CRISPR from *Prevotella* and *Francisella* 1) a été identifiée chez la bactérie *Lachnospiraceae bacterium* (LbCPF1) et représente une bonne alternative à la CAS9 (Herbert et al., 2020). Plus petite que la CAS9 (151 kDa contre 162 kDa), elle utilise un PAM riche en T, « TTTN », ce qui facilite le ciblage des régions riches en AT (Herbert et al., 2020 ; Meynard et al., 2020). Elle n'a pas besoin de tracrRNA et elle mature elle-même son crRNA de 22 à 24 pb, ce qui augmente sa spécificité de ciblage. Elle coupe en -21 et -17 du PAM en faisant des bouts cohésifs facilitant la réparation par HDR ou le retrait d'un triplet de nucléotides (Herbert et al., 2020 ; Meynard et al., 2020).

7) Les objectifs d'amélioration variétale par la technologie CRISPR

Les avancées dans ces technologies CRISPR peuvent dorénavant dépasser le cadre strict du laboratoire de recherche et pourraient être utilisées en agronomie, plus spécifiquement en amélioration génétique et en sélection pour introduire, plus rapidement, et de manière ciblée des polymorphismes d'intérêt agronomique, mais pas seulement. Par exemple, il est maintenant facile de créer des lignées végétales pouvant servir de modèles, ou des lignées animales disposant d'un type de mutation rare pour les études des cancers. Une fois ces lignées obtenues en une seule génération grâce à la technologie CRISPR, il est possible d'aller directement sur le gène endogène et d'utiliser différents outils supplémentaires pour étudier ce gène (Hsu et al., 2014).

Aujourd'hui les possibilités d'édition ou de modifications sont quasiment illimités et ne s'arrêtent pas au simple KO de gènes (Li et al., 2020a). Voici ici quelques exemples en plantes, et plus particulièrement en agronomie, comme la croissance, le rendement, la résistance aux stress biotiques et abiotiques.

a) Modification rapide et ciblée de gènes ou d'allèles

Il existe à présent de nombreux exemples de création de nouveaux traits dans les cultivars élites soit par Knock-out ou modification de gènes par BE. Pour la résistance aux stress biotiques, la désactivation du gène de susceptibilité *VvWRKY52* à *Botrytis cinerea* a conduit à la création de lignées résistantes chez la vigne (Wang et al., 2018) ; des tomates résistantes au Mildiou ont été produites en mutant le gène *SIM1o1* (Nekrasov et al., 2017). Autre cas de figure, une résistance aux herbicides a été obtenue en mutant le gène *ALS* en faisant de l'édition à la base près chez le maïs (Li et al., 2020 b).

Pour la qualité des produits, il a par exemple été possible de produire du blé à faible teneur en gluten (allergène) en mutant les gènes *a-GLIADIN* (Sánchez-León et al., 2018), ou de l'huile de la Monnoyère (*Thlapsi arvensis L.*) en réduisant la quantité de composés toxiques comme l'acide érucique en mutant le gène *FAEI* (McGinn et al., 2019).

Le gène *SIMYB12* chez la tomate a été muté pour changer la couleur du fruit qui passe d'une couleur rouge à une couleur rosé (Deng et al., 2018). L'allongement de la durée de conservation a pu être obtenu chez la tomate en mutant le gène *RIN* responsable de la maturation du fruit (Ito et al., 2015). Chez le riz, il est aussi possible d'augmenter la valeur nutritionnelle en augmentant

la quantité d'amidon et d'amylose en mutant les gènes *SBE* (Sun et al., 2017). Mieux encore, l'amélioration du rendement chez le riz a pu être faite en mutant 4 gènes simultanément, *Gn1a*, *DEP1*, *GS3*, et *IPA1* (Li et al., 2016). Enfin, il a été possible d'inactiver des gènes conférant la stérilité mâle chez le maïs (Svitashev et al., 2016), ou permettant la recombinaison méiotique chez le riz (Fayos et al., 2020 ; Mieulet et al., 2018).

Chez une espèce polyploïde, la technologie CRISPR/CAS9 a permis de modifier en une seule transformation l'ensemble des allèles du gène *GBSS* de la pomme de terre, espèce tétraploïde (Andersson et al., 2018).

Pour finir, CRISPR/CAS9 a permis de générer une large diversité allélique potentiellement exploitable en sélection pour des espèces orphelines proches d'espèces cultivées ; chez la tomate, CRISPR/CAS9 a même permis de réaliser une domestication *de novo* que nous traiterons un peu plus loin (Lemmon et al., 2018).

b) Élargissement de la diversité génétique et prébreeding

Différentes approches ont été utilisées pour accroître la diversité chez les plantes cultivées grâce à CRISPR/CAS9. Comme abordé brièvement précédemment, le multiplexing permet de muter plusieurs gènes en même temps, permettant de travailler sur plusieurs caractères d'intérêt d'un seul coup pour créer de nouveaux allèles et traits chez la même plante. Pour un gène donné, il est possible de faire un multiplexage de tous les sgRNA pour l'éditer à divers endroits (Jacob et al., 2018). Le multiplexing a permis de travailler sur des traits d'intérêts agronomiques affectant notamment le rendement ou la taille, qui sont souvent associés à des QTL (Quantitative Trait Loci). Il est maintenant possible d'accélérer le processus en testant tous les loci par les technologies CRISPR (Xu et al., 2016). Il a notamment été possible de muter plusieurs gènes affectant négativement le poids du grain chez le riz (*GW2*, *GW5*, *TGW6*) avec une construction CRISPR-Cas9 multiplexe et d'obtenir des lignées améliorées (Xu et al., 2016). Chez la tomate, une suite de mutations obtenues dans le promoteur de *SlCLV3*, affectent quantitativement le nombre de loges carpellaires et la taille du fruit (Rodríguez-Leal et al., 2017).

c) Concept de « Rewilding » et de domestication *de novo*

Brièvement, le principe du « Rewilding » consiste en un processus de réintroduction de caractères d'espèces sauvages pour améliorer la productivité d'espèces domestiquées. Cela peut se faire par inactivation des gènes de domestication comme imaginé par Andersen et al., 2015. Néanmoins c'est plutôt la domestication *de novo* qui est visée aujourd'hui. Cela correspond à l'accélération de la domestication de formes sauvages porteuses de gènes intéressants en modifiant directement les gènes de domestication (Pourkheirandish et al., 2020). Certains parlent même de seconde révolution verte en évoquant cette possibilité offerte par les NBT (Østerberg et al., 2017). Récemment, il a été possible, en partant d'une plante peu cultivée, une Solanacée (*Physalis pruinosa*) « orphan crop » (Lemmon et al., 2018), de la néo-domestiquer en mutant plusieurs gènes impliqués dans des traits d'intérêt agronomique préalablement identifiés chez la tomate (*Solanaceae lycopersicum*) en un temps record (Lemmon et al., 2018 ; Zsögön et al., 2018). Cette stratégie permettrait une possible évolution dirigée et accélérée intéressante pour certaines espèces sauvages (Zhang and Qi, 2019). C'est d'ailleurs ce qui a été fait dans la variété de riz Africaine « Kabre » en ciblant des locus de domestication (Lacchini et al., 2020).

d) Réarrangements Chromosomiques de génomes végétaux à l'aide du système

CRISPR/CAS9

Grâce aux technologies dérivées de CRISPR, il a été possible d'induire des délétions importantes dans les génomes des plantes ainsi que des réarrangements chromosomiques comme des translocations (Schindele et al., 2020).

8) Des enjeux juridiques complexes et très différents selon les pays

La réglementation concernant les NBT est encore très discutée aujourd'hui. Chaque continent et pays a ses propres restrictions et méthodes d'évaluations qui peuvent être contradictoires entre elles (Dalla Costa et al., 2019).

Initialement, il a été comparé une complémentation des techniques ancestrales de croisement et d'hybridation avec de nouvelles techniques qui peuvent amener à la présence d'un transgène (T-DNA) à un endroit inconnu dans le génome du produit final (Sprink et al., 2016). C'est

ainsi qu'est apparu en 2011 le terme de New Plant Breeding Technology (NPBT), suite à l'émergence des nouvelles biotechnologies comme CRISPR, qui ont directement été évaluées pour leurs risques et bénéfices potentiels (Sprink et al., 2016).

En Europe, c'est le procédé technologique d'obtention qui va déterminer si le produit final se classe parmi les OGM (Dalla Costa et al., 2019 ; Sprink et al., 2016). En 1990, le Conseil des Communautés Européennes avait déjà fait la distinction entre les méthodes traditionnelles de croisement et les technologies impliquant l'insertion d'acide nucléique étranger. Dans ce cas, le produit final était un OGM. Est qualifié comme OGM « tout organisme auquel le matériel génétique a été altéré de manière non naturelle par transformation et/ou recombinaison » (directive UE 2001/18/CE) (Sprink et al., 2016). Depuis le 25 juillet 2018, en Europe, tout organisme obtenu par mutagenèse non naturelle (dont les NBT) sont classés parmi les OGM (Dalla Costa et al., 2019). L'approche est totalement différente pour d'autres pays ou continents. S'il fallait résumer brièvement la stratégie des USA à ce sujet, des plantes ne sont pas considérées comme OGM s'il n'y a plus de traces du transgène dans le produit final (Dalla Costa et al., 2019). En 1989 il avait déjà été établi que seul le risque potentiel sur l'humain, l'animal, ou l'environnement par le produit final devait être évalué. Depuis 1992 ce n'est pas moins de 121 produits qui ont été dérégulés par l'USDA (US Department of Agriculture) (Sprink et al., 2016). Par exemple en 2018 des champignons non-brunissant modifiés par CRISPR CAS9 ont été commercialisés (Sathishkumar et al., 2019).

Au Canada, la législation est assez similaire à celle des USA, mais ne distingue plus naturel de non naturel. Dès 1990 on parlera de « plantes avec de nouveaux caractères », sans réelle régulation spécifique sur leur production. Cependant, le nouveau caractère introduit doit être évalué pour tous ses risques potentiels, NBT ou non. En 2016, 100 plantes NPBT ont déjà été approuvées (Sprink et al., 2016). Pour finir, en Argentine la commercialisation des NPBT se fait au cas par cas. Comme au Canada, il y a une évaluation du nouveau caractère pour ses risques potentiels (Sprink et al., 2016).

Les réflexions actuelles sur de nouveaux modèles pour évaluer ces plantes seraient la séparation des filières OGM/non OGM, des analyses de risques amont, l'absence d'ADN étranger et d'évènement off-Target etc... (Sprink et al., 2016). Pour finir, la non acceptation des produits édités par le consommateur reste le point critique pour la commercialisation future des plantes éditées (Sathishkumar et al., 2019).

Conclusions sur la technologie CRISPR

La puissance de cette technologie en raison de sa simplicité d'utilisation et de son efficacité a permis l'ouverture de nouvelles perspectives, révolutionnaires. Sa découverte inattendue, dans un domaine de recherche obscur illustre à merveille l'intérêt de mener des recherches fondamentales (Hsu et al., 2014). Les technologies CRISPR/CAS9 ont balayé les technologies d'analyse fonctionnelle qui les précédaient et sont en perpétuelle évolution. Pour son utilisation comme NBT, les conclusions sont plus mitigées. Les points de blocages restent nombreux, parmi lesquels on peut citer : l'absence d'un protocole efficace de transformation génétique chez de nombreuses espèces et/ou pour des cultivars élites, la non acceptabilité sociale de cette technologie, associée aux OGM dans l'esprit des consommateurs, et des réglementations à géométrie variable.

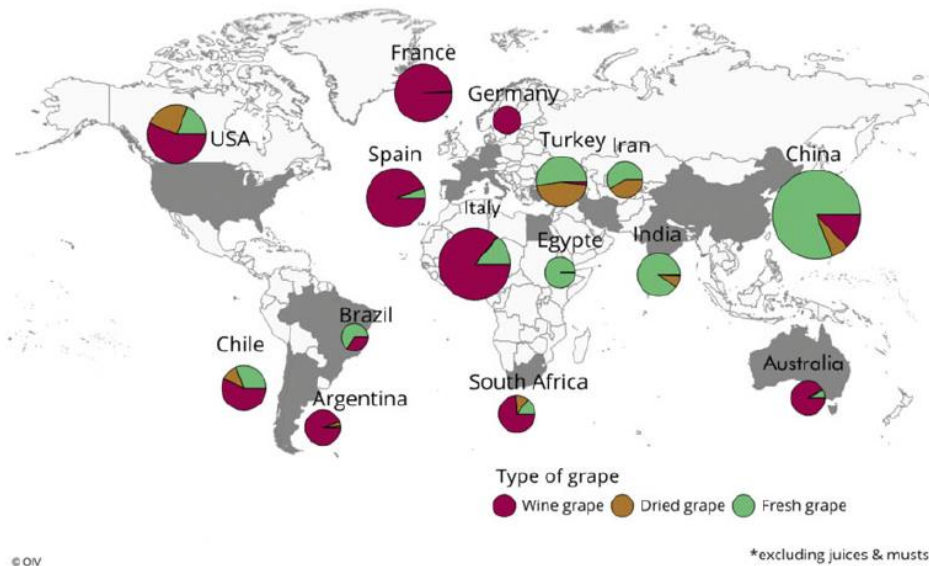


Fig. 1.3 Vine-growing areas and utilization of grape production in 2015. *Source* Provided courtesy of OIV (pers. comm. Nicolas Goldschmidt, July 2018)

Figure 8 : Répartition des cultures dans le monde d'après Cantu et Walker., 2019

Les vignes donnant du vin sont extrêmement présentes en Europe, tandis que les vignes pour le raisin de table sont présentes en majorité sur le continent Asiatique.

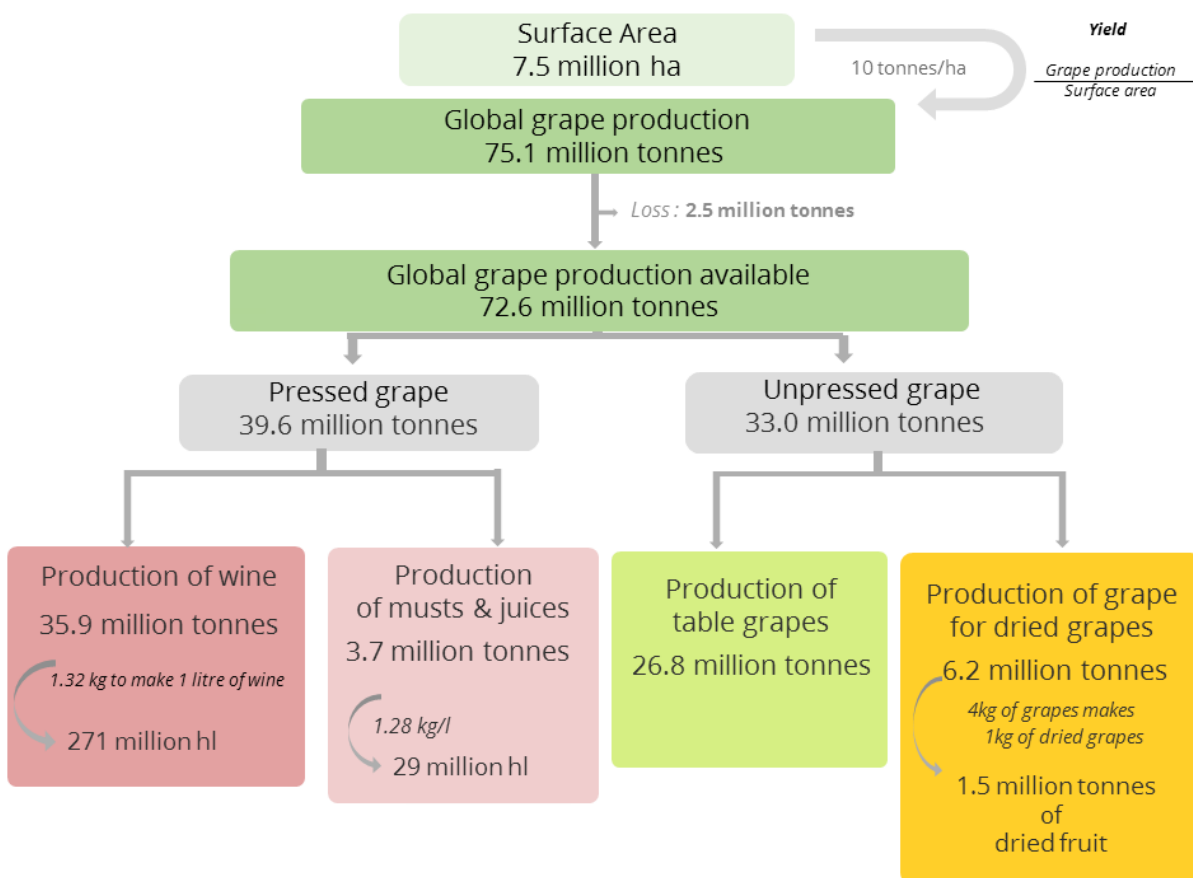


Figure 9 : Surface de culture des vignes mondiales et production

D'après l'OIV 2014, la production de raisin approchait les 75 millions de tonnes pour 7,5 millions d'hectares de cultures. La production de vin approchait les 271 million d'hectolitres, et une seconde partie moindre de 26,8 million de tonnes donnait les raisins de table.

Partie 2 : La vigne et le riz, deux espèces candidates pour l'amélioration via les NBT

De par leurs impacts commerciaux et agronomiques importants, la vigne et le riz représentent deux espèces stratégiques pour le développement et l'amélioration des NBT. Il est possible de réfléchir à une action sur les leviers de la production et du rendement par exemple afin de garantir la sécurité alimentaire, mais aussi d'agir en faveur de la recherche pour l'accélérer et faciliter le travail sur des caractères difficilement manipulables auparavant. Les cibles d'amélioration, et les contraintes liées à la biologie de ces deux espèces sont par contre complètement différentes.

I] La vigne : les NBTs comme outils de validation fonctionnelle et d'amélioration génétique, intérêts et difficultés

La vigne, *Vitis vinifera L.*, est une espèce fruitière cultivée importante dans le monde avec une surface de près de 7,5 millions d'hectares de vignobles, soit 0.5 % des terres arables et des cultures permanentes mondiales réparties sur tous les continents (**Figure 8**). L'Europe comptabilise près de 3,7 millions d'hectares (Ha) de cultures, l'Asie près de 2 millions d'hectares, et les Amériques plus d'un million (OIV 2019, <http://www.oiv.int/fr/statistiques/recherche>). En Europe, l'Italie, la France et l'Espagne sont les 3 principaux pays viticoles totalisant 2.4 M d'hectares. Selon les années, l'Italie et la France se disputent la place de premier producteur mondial de vin. En 2016, la France a produit 45,4 millions d'hectolitres, l'Italie 50 millions.

La production actuelle de raisins est de 78 millions de tonnes par an, réparties entre presque 50 millions pour la production de vin, 27 millions pour le raisin de table, et 1.3 million pour les raisins secs. Le vin est donc le principal produit de la vigne : sa production mondiale avoisine les 300 millions d'hectolitres par an, pour un marché de près de 31 milliards d'euros (FAO stats 2014 <http://www.fao.org/3/a-i7042e.pdf>, OIV 2019) (**Figure 9**).

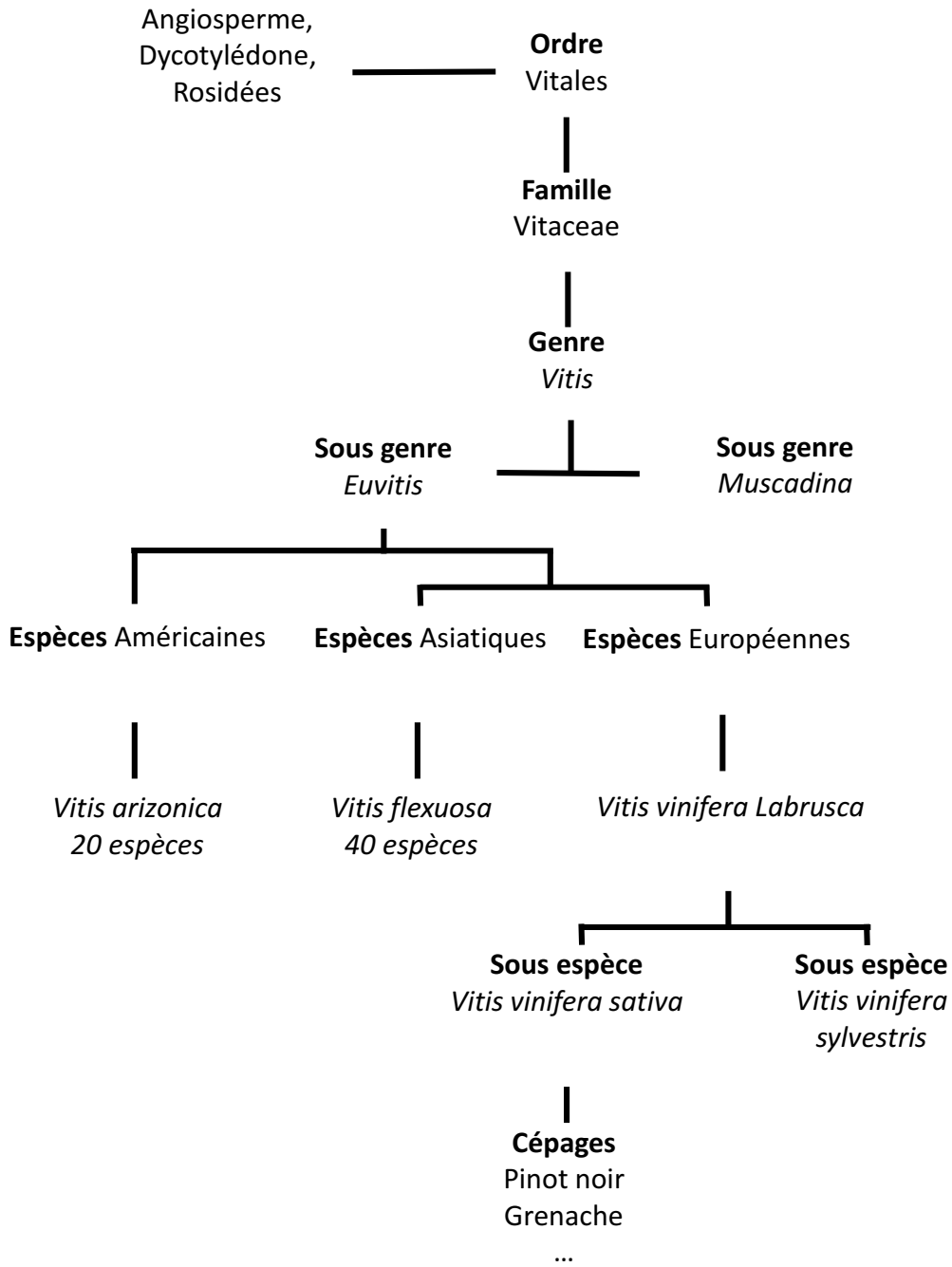


Figure 10 : Arbre phylogénétique partiel de la vigne

Description phylogénétique de la vigne à partir du clade des Angiospermes jusqu'aux cépages.

1) Espèces et domestication

La vigne appartient aux Angiospermes dans le groupe des Dicotylédones. Elles sont ensuite apparentées au groupe des *Rosidae*, à l'ordre des *Vitales*, puis à la famille des *Vitaceae*. Le genre *Vitis* donnera les plantes actuellement cultivées dont les fruits sont consommables.

La famille des *Vitaceae* est composée de plus d'un millier d'espèces, le genre *Vitis* en contenant environ 80 que l'on retrouve dans l'hémisphère nord divisées en deux sous-genres : *Muscadina*, et *Euvitis* qui est le plus représenté (Galet, 1988). Les espèces du sous-genre *Euvitis* peuvent être séparées en trois taxons selon leurs origines géographiques : américaine, asiatique et européenne (Péros et al., 2011). Le groupe américain regroupe une vingtaine d'espèces (dont *Vitis arizonica*), le groupe asiatique une quarantaine d'espèces parmi lesquelles on retrouve *Vitis flexuosa* en Chine, au Japon et au sud de l'Asie et une vigne à épines *Vitis spinosa* en Chine. Enfin, le groupe européen se compose d'une seule espèce, *Vitis vinifera*. Par une analyse des ADN chloroplastique et des marqueurs microsatellites, Péros et al. (2011) ont pu clarifier les relations entre ces différentes espèces, notamment une plus grande proximité de *Vitis vinifera* et des espèces asiatiques (Galet, 1988 ; Péros et al., 2011). L'espèce *Vitis vinifera* L. *subsp. sativa* (*Vv*), originaire du continent européen est l'espèce la plus cultivée, puisqu'elle représente près de 98 % des surfaces plantées dans le monde (Velasco et al., 2007). Elle co-existe encore avec la forme sauvage *Vitis vinifera subsp. sylvestris* qui a été domestiquée pour donner la vigne que nous cultivons actuellement (This et al, 2006) (**Figure 10**).

Les premières cultures de *Vitis vinifera* (*Vv*) datent de -8000 à -6000 et la production de breuvages apparentés au vin est apparue au Néolithique (Jaillon et al., 2007). Actuellement, on connaît près de 6000 cultivars (Lacombe, 2012), mais seuls 33 représentent plus de 50 % des surfaces cultivées (Sathishkumar et al., 2019) car ils permettent de faire des vins de qualité (Bouquet et al., 2006) (**Figure 11**). Parmi les cépages de cuve les plus cultivés, nous pouvons retrouver :

- Cabernet Sauvignon (340 000 Ha en 2017 ; OIV 2017) qui est un raisin noir de cuve, issu d'un croisement entre Cabernet franc et Sauvignon obtenu au 17^e siècle (Bowers and Meredith, 1997), sensible à l'oïdium (Doré and Varoquaux, 2006).
- Chardonnay B (210 000 Ha) qui est un raisin blanc de cuve français, issu d'un croisement entre le Pinot noir et le Gouais (Bowers, 1999), présent dans le plus grand nombre de pays (OIV, 2017)

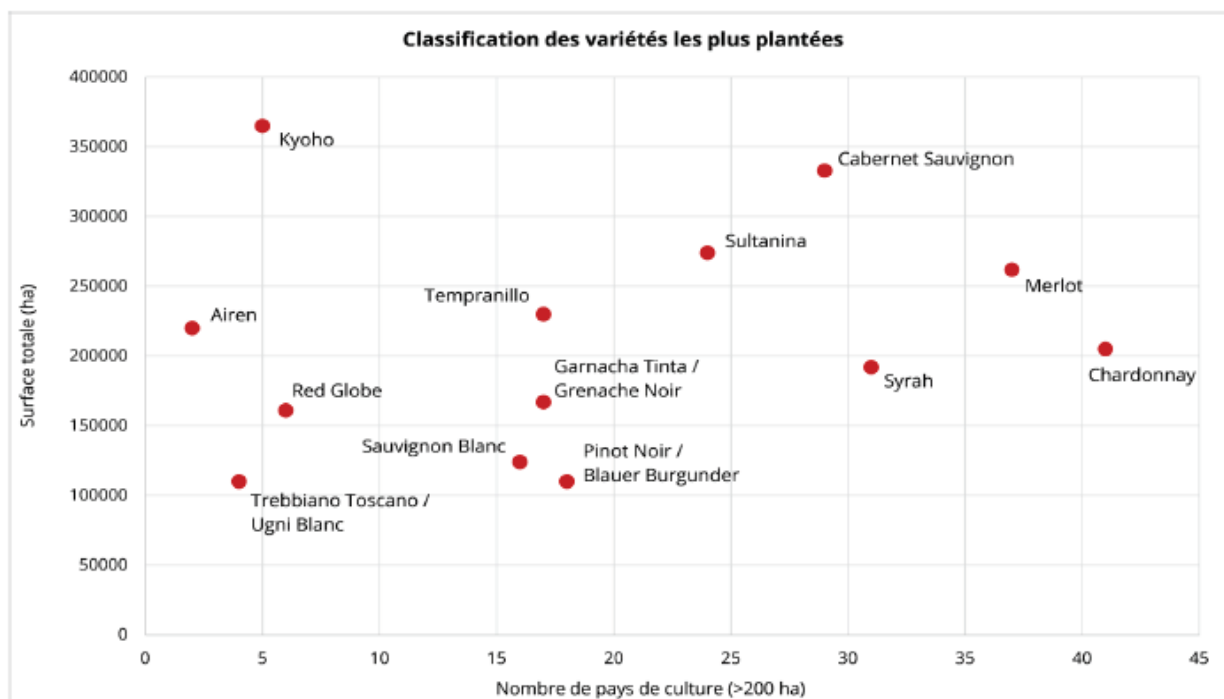


Figure 11 : Classification des variétés de vigne les plus cultivées

D'après l'OIV 2017, le Cabernet-Sauvignon, le Merlot et le Chardonnay représentent les vignes les plus cultivées dans plus d'une trentaine de pays. A contrario du Airen ou Kyoho, très cultivés mais dans seulement quelques pays.



Figure 12 : Vignes sur le domaine de Pech Rouge

Visite du domaine de Pech Rouge de l'INRAE en Septembre 2019. Observation des différentes parcelles et croisement effectuée par l'INRAE dont Mr Alain Bouquet.

- Grenache N. (163 000 Ha), raisin noir espagnol de cuve, utilisé à travers le monde principalement dans les régions chaudes
- Pinot N. (112 000 Ha) qui est la variété séquencée en 2007 par Velasco *et al* (2007), un raisin noir de cuve français entrant notamment dans la confection des vins de Champagne et de Bourgogne. (OIV 2017, <http://www.oiv.int/public/medias/5861/fr-distribution-vari-tale-du-vignoble-dans-le-monde.pdf>).

La plus grande partie de ces 6000 cépages se retrouvent cependant dans des collections de ressources génétiques, comme la collection du CRB Vigne du domaine INRAe de Vassal vers Marseillan (<https://www6.montpellier.inrae.fr/vassal/>), ou de l'unité expérimentale INRAe de Pech Rouge dans l'Aude (**Figure 12**).

2) Morphologie et reproduction

La vigne est une plante pérenne, caduc, ligneuse et sarmenteuse, et lianescente. Les tiges ligneuses représentent les sarments, que l'on appelle aussi rameaux. Ils ont des nœuds où apparaissent les feuilles, les fleurs, et les vrilles. Les vrilles servent à la vigne pour s'accrocher lors de son développement. Les feuilles sont palmées, composées de cinq nervures principales. Elles sont plus ou moins découpées, ce caractère servant notamment de critère d'identification. Les variétés cultivées, encore appelées cépages, ont différentes morphologies (Doré and Varoquaux, 2006). Les fleurs quant à elles sont petites, vertes à blanches, et forment des inflorescences en grappe qui peuvent compter jusqu'à plusieurs centaines de fleurs : elles sont le plus souvent hermaphrodites (Galet, 1988 ; Velasco et al., 2007) notamment chez *Vitis vinifera*, alors que de nombreuses espèces de *Vitis* présentent des plantes dioïques mâles ou femelles. Elles donneront les baies de raisins sous forme de fruits charnus (Galet, 1988).

La reproduction de la vigne peut se faire de deux façons :

- Par reproduction sexuée après une phase juvénile de 2 à 5 ans (Huglin and Schneider, 1998) qui est essentiellement de type anémophile : chez les espèces dioïques, l'allogamie est stricte, alors que chez les cépages hermaphrodites, l'autogamie est possible, bien que la vigne, très hétérozygote, soit affectée par la dépression de consanguinité (Lacombe, 2012). Ce mode de reproduction est notamment utilisé pour la création des nouvelles variétés.
- Par reproduction végétative, par marcottage en milieu naturel ou par greffage qui est actuellement le principal mode de reproduction des cépages cultivés dans le monde. Il

permet notamment le maintien de l'identité du cépage, même si une variabilité dite clonale peut apparaître au cours des cycles successifs de multiplication.

Néanmoins le cycle est long. Trois à six ans sont nécessaires entre la germination du pépin et une plante adulte pouvant donner des raisins (Chatbanyong and Torregrosa, 2015).

3) Génomique et amélioration

Vitis vinifera, est une espèce diploïde, et comme l'ensemble des espèces du sous-genre *Euvitis* présente $2n=38$ chromosomes, alors que les espèces du sous-genre *Muscadinia* présentent $2n=40$ chromosomes. Après *Arabidopsis*, le riz, et le peuplier, le génome de la vigne a été séquencé en 2007 (Jaillon et al., 2007). Deux séquences ont ainsi été publiées, la séquence de la lignée PN40024 (Jaillon et al., 2007) et la séquence du cépage Pinot Noir (Velasco et al., 2007). L'assemblage du génome de PN40024 a été facilité car cette lignée, issue de plusieurs autofécondations successives est hautement homozygote (93 %). Cela a aussi permis de démontrer que le génome diploïde de la vigne était issu de trois génomes ancestraux (Jaillon et al., 2007). Avec cette première version 8x, près de 30434 gènes ont été prédits pour 487 Mb, et près de 41,4 % du génome sont des éléments transposables ou répétés (Chaïb et al., 2010 ; Jaillon et al., 2007). Une version 12X.0 a été publiée en 2009 avec la version 0 de l'annotation des gènes (Adam-Blondon et al., 2011) qui décrit plus de 26 000 gènes. Enfin, une seconde version de la séquence 12x a été publiée (Canaguier et al., 2017) et est disponible sur le site de l'URGI (<https://urgi.versailles.inra.fr/Species/Vitis/Data-Sequences/Genome-sequences>).

Le génome de Pinot N. hautement hétérozygote, a quant à lui permis de révéler un fort polymorphisme entre les deux haplotypes avec plus de 1 751 176 de SNP, et des polymorphismes d'insertion ou délétions de séquences (Velasco et al., 2007). Le séquençage du génome a ensuite facilité le développement des cartes génétiques positionnant les marqueurs moléculaires précisément (Sathishkumar et al., 2019). Néanmoins, parmi tous les cultivars une grande diversité génétique est présente avec un haut taux de polymorphisme, et une grande hétérozygotie (Bouquet et al., 2006 ; Velasco et al., 2007). Il est donc essentiel de connaître les séquences des gènes d'intérêts dans les cultivars et les haplotypes sélectionnés afin de pouvoir travailler dessus en parallèle de la référence de PN40024. Une des priorités des axes de recherches de la communauté vigne a été les gènes de résistance (Doré and Varoquaux, 2006). En effet la vigne est hautement susceptible aux pathogènes comme le mildiou, l'oïdium,

le phylloxera, ou les nématodes. Néanmoins il est possible de faire des greffages, ou bien de croiser certains cultivars avec des espèces comme *Vitis riparia* ou *Vitis cinerea* qui sont connues comme résistantes à certains pathogènes (Jaillon et al., 2007). Velasco et collaborateurs ont souligné l'intérêt de la connaissance du génome de différentes espèces de vignes pour développer des variétés améliorées capables de tolérer l'attaque de ces pathogènes (Velasco et al., 2007). L'utilisation généralisée de porte-greffe de *Vitis* américains ou hybrides s'est généralisé. En effet ils sont résistants au phylloxera qui a ravagé une grande partie des vignobles Européens au début du siècle. Ils sont également résistants, aux nématodes, à un pH alcalin du sol, à la salinité et la sécheresse (Keller, 2010).

Cependant, la longueur de son cycle de développement et la forte hétérozygotie du génome de la vigne posent problème pour les croisements entre individus. Cela n'a néanmoins pas empêché la production de nouvelles variétés, comme le Regent qui est un cultivar allemand, issu d'un double croisement Silvaner par Müller, puis par Chamboursin (Sathishkumar et al., 2019). Néanmoins, ces croisements ne préservent pas le background génétique complet du cultivar initial (Dalla Costa et al., 2019). Il a donc fallu penser à des alternatives de modifications et cultures *in-vitro* afin de maintenir l'identité des variétés.

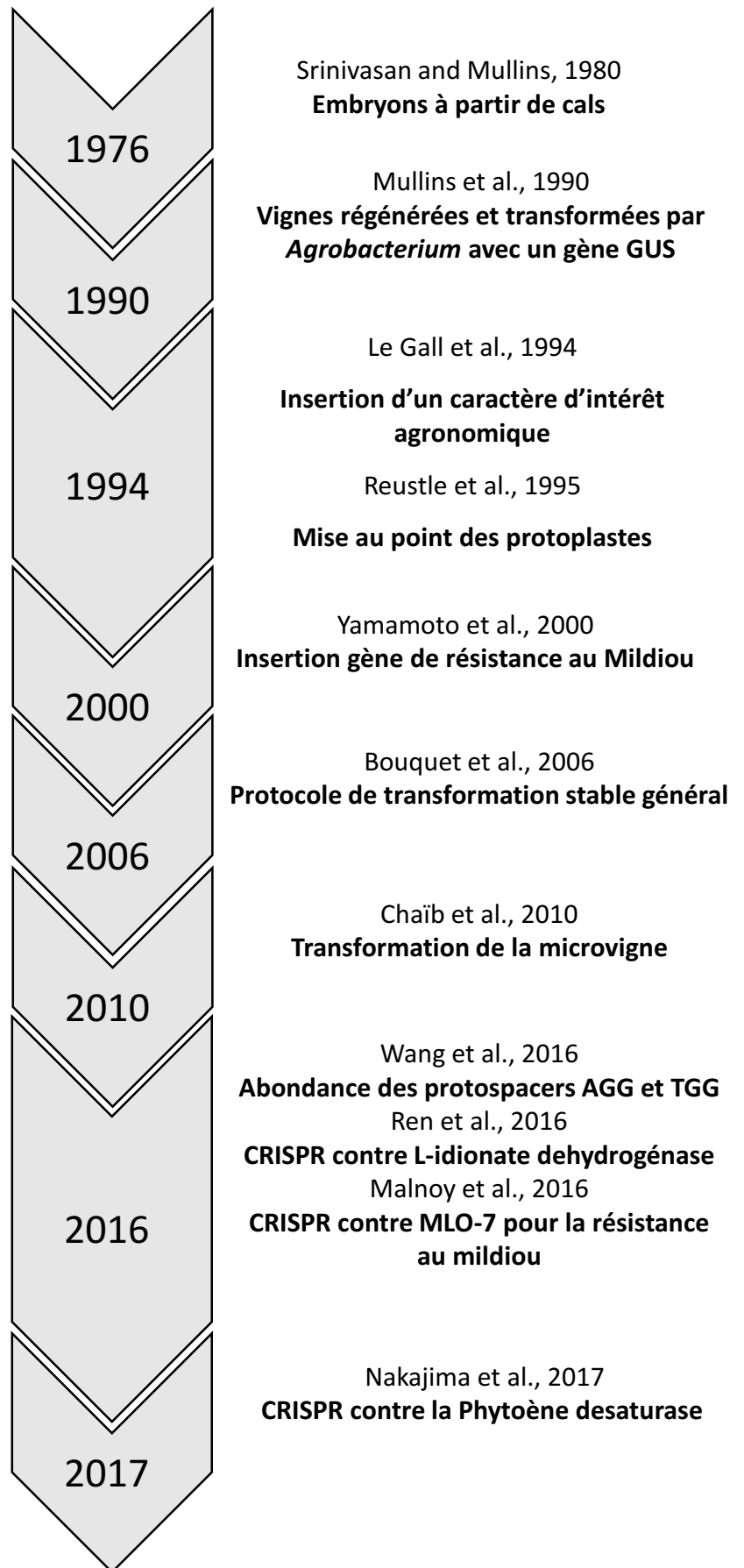
4) La vigne et les biotechnologies

L'apport des biotechnologies a été un facteur déterminant dans l'amélioration des plantes, dont la vigne. Pour Dalla Costa et collaborateurs, elles peuvent être classées en trois catégories : les techniques de conservation et de multiplication ; les techniques de sélection sanitaire ou génétique pour sélectionner des individus naturellement élites ; et les techniques pour modifier le génome et l'épigénome (Dalla Costa et al., 2019). Nous ne traiterons ici que des dernières techniques permettant une modification du génome.

Une des premières complexités de la vigne est l'influence du génotype sur la régénération d'embryons somatiques (Sathishkumar et al., 2019). Par exemple il a été montré que le Chardonnay ou le Grenache avaient une très bonne capacité à former des embryons somatiques régénérables, alors que le Cabernet-Sauvignon n'en avait presque pas (Gribaudo et al., 2017). Les premières expérimentations sur les embryons de vignes ont été menées dans les années 70 afin d'arriver à obtenir des milliers d'embryons à partir de cals issus de Cabernet-Sauvignon et Grenache (Srinivasan and Mullins, 1980). Mais ce n'est que dans les années 90 qu'il a été

Figure 13 : Amélioration chronologique de la vigne

Brève chronologie sur les différentes améliorations sur la culture et transformation de la vigne au cours des quarante dernières années.

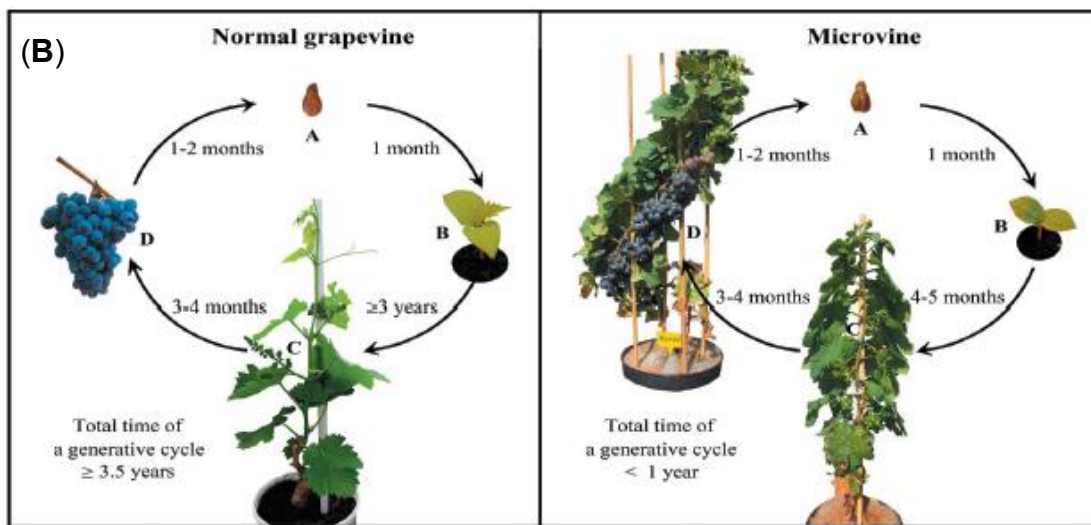
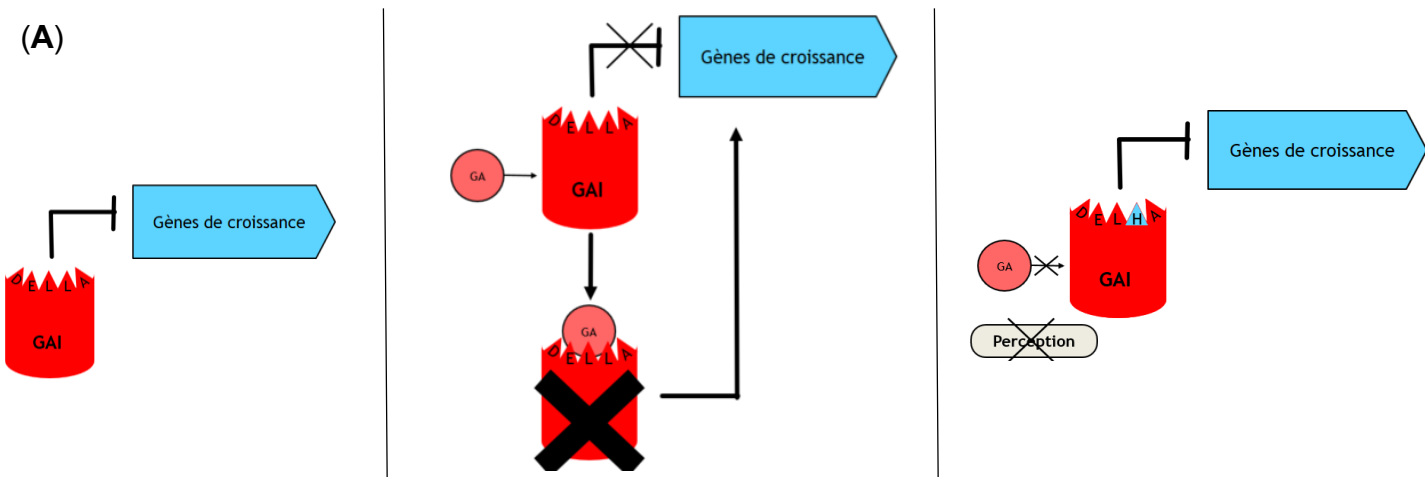


possible d'obtenir des vignes régénérées et transformées par *Agrobacterium* avec un gène GUS (Mullins et al., 1990). Un des tout premiers caractères d'intérêt agronomique travaillé a été la résistance aux virus. Un gène codant pour une protéine du Grapevine Chrome Mosaic Virus (Coat protein –GCMV-CP) a été inséré en 1994 par *Agrobacterium* dans des cals issus d'anthères du porte-greffe 110 Richter, afin de transférer la résistance aux porte-greffe qui sont le point d'entrée du virus (Le Gall et al., 1994). C'est cette même année que des protoplastes ont été isolés à partir d'embryons et d'embryoïdes dérivés de feuilles de la variété Seyval Blanc et ont permis la régénération de plantes *in-vitro* (Reustle et al., 1995). Les améliorations des techniques d'étude sur des maladies par la transformation génétique sont apparues dans les années 2000 grâce à l'ajout d'un gène de résistance contre le mildiou issu du riz dans le cultivar Neo Muscat, (Yamamoto et al., 2000). Quelques années plus tard, un protocole efficace de transformation et de régénération d'embryons a été mis au point afin de faciliter l'analyse fonctionnelle chez cette espèce (Bouquet et al., 2006). La technologie CRISPR-CAS9 a été utilisée avec succès chez la vigne pour la première fois en 2016. L'étude du génome de la vigne a permis de démontrer une abondance des protospacers AGG et TGG (Wang et al., 2016b) ; peu de temps après plusieurs gènes ont été ciblés et mutés avec succès comme la L-idionate déhydrogénase du métabolisme de l'acide tartrique (Ren et al., 2016) et MLO-7 pour la résistance au mildiou (Malnoy et al., 2016) chez le Chardonnay, et la Phytoène desaturase pour produire des vignes jaunes chez Neo Muscat (Nakajima et al., 2017) (**Figure 13**).

5) La microvigne, un outil pour accélérer les cycles de sélection

Afin d'accélérer le cycle de culture et les analyses, une vigne à cycle rapide, appelée microvigne, a été obtenue à la fin de la décennie passée (Chaïb et al., 2010). Ces microvignes disposent ainsi des avantages propres aux plantes annuelles : floraison au bout de 4 à 5 mois, obtention de fruits en 6 à 8 mois, taille réduite d'une plante adulte (Chaïb et al., 2010).

Ces vignes sont des mutants obtenus par culture *in-vitro* (régénération et séparation des couches chimériques du Pinot Meunier), présentant une mutation dans le gène *GAI*. Ce gène est historiquement connu car il est à l'origine de la « Révolution Verte » dans les années 60. Il s'agit d'un gène intervenant dans la perception des gibbérellines et à l'origine du semi-nanisme des plantes de la révolution verte : *GAI* (gibberellin insensitive). La protéine *GAI* intervient dans la réponse aux gibbérellines (GA), hormone impliquée dans l'élongation des entre-nœuds et dans les variations de croissance en réponse à différentes conditions environnementales, par



(C)

	cov	pid 1201	
1	Me13Tag16_GAI_like_protein1_AF378125	100.0%	100.0%
2	Me117Tag16_GAI_like_protein1_AF378125	100.0%	99.7%
3	Me14Tag16_GAI_like_protein1_AF378125	100.0%	99.5%
4	Me11Tag16_GAI_like_protein1_AF378125	100.0%	99.6%
5	Me14Tag13_GAI_like_protein1_AF378125	100.0%	99.7%
6	Me13Tag13_GAI_like_protein1_AF378125	100.0%	99.7%
7	Me14Tag13_GAI_like_protein1_AF378125	100.0%	99.7%
8	Me16Tag16_GAI_like_protein1_AF378125	100.0%	99.8%
9	Me15Tag13_GAI_like_protein1_AF378125	100.0%	99.8%
10	Me17Tag13_GAI_like_protein1_AF378125	100.0%	99.4%
11	Me15Tag16_GAI_like_protein1_AF378125	100.0%	99.3%
12	Me16Tag14_GAI_like_protein1_AF378125	100.0%	99.4%
13	Me17Tag14_GAI_like_protein1_AF378125	100.0%	99.3%
14	Me17Tag11_GAI_like_protein1_AF378125	100.0%	99.3%
15	Me13Tag9_GAI_like_protein1_AF378125	100.0%	99.4%
16	Me13Tag11_GAI_like_protein1_AF378125	100.0%	99.4%
17	Me117Tag10_GAI_like_protein1_AF378125	100.0%	99.4%
18	Me120Tag8_GAI_like_protein1_AF378125	100.0%	99.4%
19	Me15Tag10_GAI_like_protein1_AF378125	100.0%	99.4%
20	Me112Tag11_GAI_like_protein1_AF378125	100.0%	99.3%
21	Me117Tag8_GAI_like_protein1_AF378125	100.0%	99.4%
22	Me12Tag8_GAI_like_protein1_AF378125	100.0%	99.4%
23	Me117Tag10_GAI_like_protein1_AF378125	100.0%	99.5%
24	Me111Tag1_GAI_like_protein1_AF378125	100.0%	99.5%
25	Me118Tag9_GAI_like_protein1_AF378125	100.0%	99.6%
26	Me119Tag9_GAI_like_protein1_AF378125	100.0%	99.5%
27	Me116Tag14_GAI_like_protein1_AF378125	100.0%	99.5%
28	Me115Tag11_GAI_like_protein1_AF378125	100.0%	99.5%
29	Me117Tag13_GAI_like_protein1_AF378125	100.0%	99.5%
30	Me118Tag9_GAI_like_protein1_AF378125	100.0%	99.5%

Figure 14 : Mécanismes moléculaires de réponse aux GA et fonction du gène *VvGAI*

Le gène *GAI* est un répresseur de la réponses aux GA (**A, gauche**).

Les GA forment un complexe ternaire (GA-GID1-GAI1) avec la protéine GID1 (récepteur) et *GAI*, grâce au motif DELLA de *GAI*. Le complexe va ensuite être envoyée au protéasome où la protéine *GAI* sera dégradée levant l'inhibition de réponse aux GA. Les gènes de réponse à la croissance se retrouvent ainsi activés en présence de GA. En absence de GA, les gènes de réponse sont réprimés (**A, milieu**). La mutation du domaine DELLA de *VvGAI* en DELHA permet d'induire le nanisme (« microvine ») en empêchant la fixation des GA, ce qui est conceptuellement similaire à une absence constitutive de GA (**A, droite**).

Ce nanisme est accompagné d'une floraison continue, et un cycle court entre les générations en (**B**) (Chatabanyong et al. 2015). En multialignant des séquences homologues de *GAI* de cultivars de vigne, on constate que le gène *VvGAI* est très bien conservé en particulier la séquence du domaine DELLA (**C**).

exemple au cours du développement des inflorescences chez la vigne (Boss and Thomas, 2002; Zhong and Yang, 2012). GAI est un facteur de transcription répressif de la réponse aux GA. La balance entre l'inhibition ou l'activation de la biosynthèse des GA est donc finement régulée (Boss and Thomas, 2002). Au niveau moléculaire, les GA forment un complexe GA-GID1-GAI1 grâce au motif protéique DELLA (Acide aspartique — Acide glutamique — Leucine — Leucine — Alanine ; gat gag ctt ctc get) de GAI. La protéine GAI est ensuite prise en charge par un second complexe (GID2-SKP1-RBX1) et polyubiquitinée, pour être amenée au protéasome pour y être dégradée. Cela permet de lever sa répression sur les gènes de réponse à la croissance (**Figure 14-A**). Une mutation dans le gène *VvGAI* chez le cultivar Pinot Meunier provoque un phénotype de nanisme et une floraison accélérée continue de la vigne permettant de raccourcir le délai entre générations, et facilite le phénotypage des caractères de la qualité de la baie (Chaïb et al., 2010; Chatbanyong and Torregrosa, 2015). Chez le mutant *vvgai*, il n'y a plus aucune reconnaissance entre le complexe GID1-GA et le domaine DELLA de GAI, ce qui conduit à une répression permanente par la protéine GAI des gènes de réponse au GA. C'est plus précisément une mutation du domaine DELLA en DELHA (substitution de la seconde Leucine par une Histidine) qui est présente dans la mutation *vvgai* chez la micro vigne (Boss and Thomas, 2002; Zhong and Yang, 2012) (**Figure 14-B**). Les domaines N-ter, où se trouve le motif DELLA, et C-ter sont très conservés entre plantes, par exemple entre *Arabidopsis thaliana* et *Vitis vinifera* il y a 72 % de similarité (Boss and Thomas, 2002). Lors de l'introduction de mutations chez *Arabidopsis thaliana*, par exemple en supprimant le domaine DELLA de *AtGAI*, on retrouve un phénotype de nanisme. De plus la mutation de la 1^{re} leucine du motif a moins d'impact sur le phénotype que la mutation sur la seconde leucine (Zhong and Yang, 2012).

Disposer de microvignes pour d'autres cultivars que le Pinot Meunier permettrait d'accélérer la sélection chez cette espèce : on pourrait en effet envisager de créer des géniteurs intéressants (présentant des gènes favorables sous forme homozygote) dans un fond « MicroVigne », pour accélérer leur création, puis croiser ces géniteurs hétérozygotes pour GAI. Nous avons donc choisi de cibler ce gène pour deux raisons : le phénotype est rapidement visible, mais la production de microvignes par NBT serait utile pour la création de collections génétiques.

Chez la vigne, nous chercherons donc à mettre au point la mutation du gène *VvGAI* par délétion de son motif DELLA ou en faisant une mutation KO simple en utilisant la technologie CRISPR dans quatre cultivars : Chardonnay, Portant, Merlot, et Grenache. Ce gène est par ailleurs un des rares chez la vigne à être très conservé ce qui permettrait d'utiliser éventuellement la

technologie CRISPR pour d'autres cultivars/espèces de vigne dans le futur (**Figure 14-C**). Le phénotype sera facile à analyser, permettant éventuellement de valider la preuve de concept, et en même temps faciliter la production de plants de microvignes.

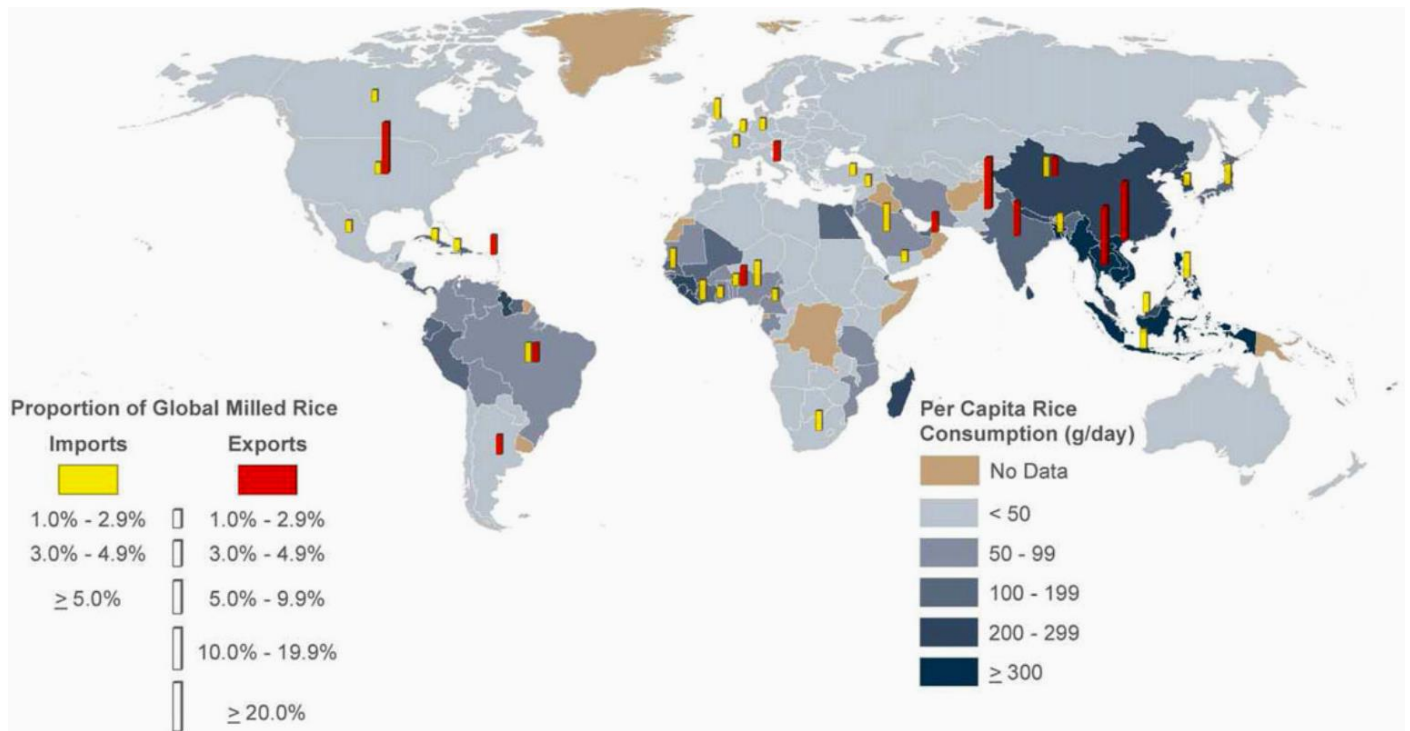


Figure 15 : Carte mondiale des imports, exports, et consommation de riz en 2010, d'après Muthayya et al., 2014

La Chine et l'Inde présentent la plus grosse consommation et exportation de riz dans le monde. L'Afrique et les pays Sud-Américains consomment aussi une grande partie de la production mondiale, figure issue de Muthayya et al., 2014.

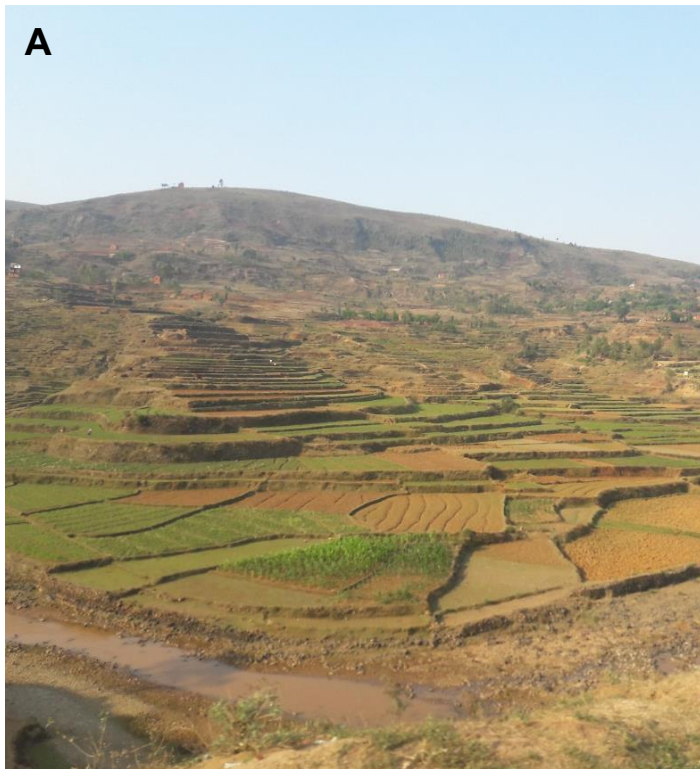


Figure 16 : Présentation des systèmes de cultures en riz irrigués à Madagascar

A Madagascar du côté de Antsirabe les riz sont cultivés de manière irrigués sur le flanc des montagnes modifiant le paysage (A). Du côté de la capitale Antananarivo, les riz y sont à même cultivés au milieu de la zone urbaine (B).

Photos : Léo Herbert



III) Le riz, une plante modèle pour le développement de l'utilisation des NBTs en amélioration des plantes

La culture et la domestication des céréales ont débuté il y a plus de 10000 ans, lors de la transition des premières sociétés humaines de chasseurs-cueilleurs en sédentaires-agriculteurs. Le riz fut une des premières céréales à être domestiquée dans les Indes et en Chine il y a -3500 et -7000 ans (Demol et al., 2002). Le riz est une monocotylédone qui appartient à la famille des *Poaceae*, à la sous-famille des *Oryzoïdeae*, dans la section des *Oryzeae*, et du genre *Oryza*. Ce n'est pas moins de 24 espèces que l'on retrouve surtout en Asie, en Afrique, en Amérique et en Océanie (Demol et al., 2002).

Le riz est la base de l'alimentation mondiale (Izawa and Shimamoto, 1996) en étant la source calorique de la moitié de l'humanité (FAO). En effet cette céréale est cultivée sur plus de 170 millions d'hectares, et l'Asie totalise près de 90 % des surfaces cultivées (FAOstat, (Muthayya et al., 2014)) (**Figure 15**). En moyenne son rendement mondial était de 4.6 tonnes par hectare (FAOstat 2017), mais ce dernier devra augmenter de 1 % par an jusqu'à 2050 afin de pallier à la demande croissante de la population (Meynard et al., 2020, FAO).

1) Morphologie et espèces du riz

Cette monocotylédone annuelle à cycle court a une morphologie variable pour une taille de 40 cm à 5 m de haut en fonction de l'écotype cultivé ; sa croissance se subdivise en trois phases : végétative, reproductive, et de maturation ; la dernière phase correspondant au remplissage des grains. Son système reproductif est situé sur des panicules portant 100 à 125 épillets de fleurs autogames (organes mâles et femelles sur la même fleur). La reproduction se fait par cléistogamie (autofécondation) pendant 5 à 9 jours. La maturation du grain dure 28 jours en moyenne (Demol et al., 2002). Le riz est cultivé dans différents systèmes de culture dans le monde. Il est majoritairement cultivé en conditions irriguées (**Figure 16**), mais également en conditions inondées, en conditions submergées (riz flottant), ou en conditions pluviales qui est la moins gourmande en eau et en intrants.

Oryza sativa (*Os*) est le riz le plus important que Demol et collaborateurs ont même qualifié de « cosmopolite » en raison de sa répartition mondiale (Demol et al., 2002). Il

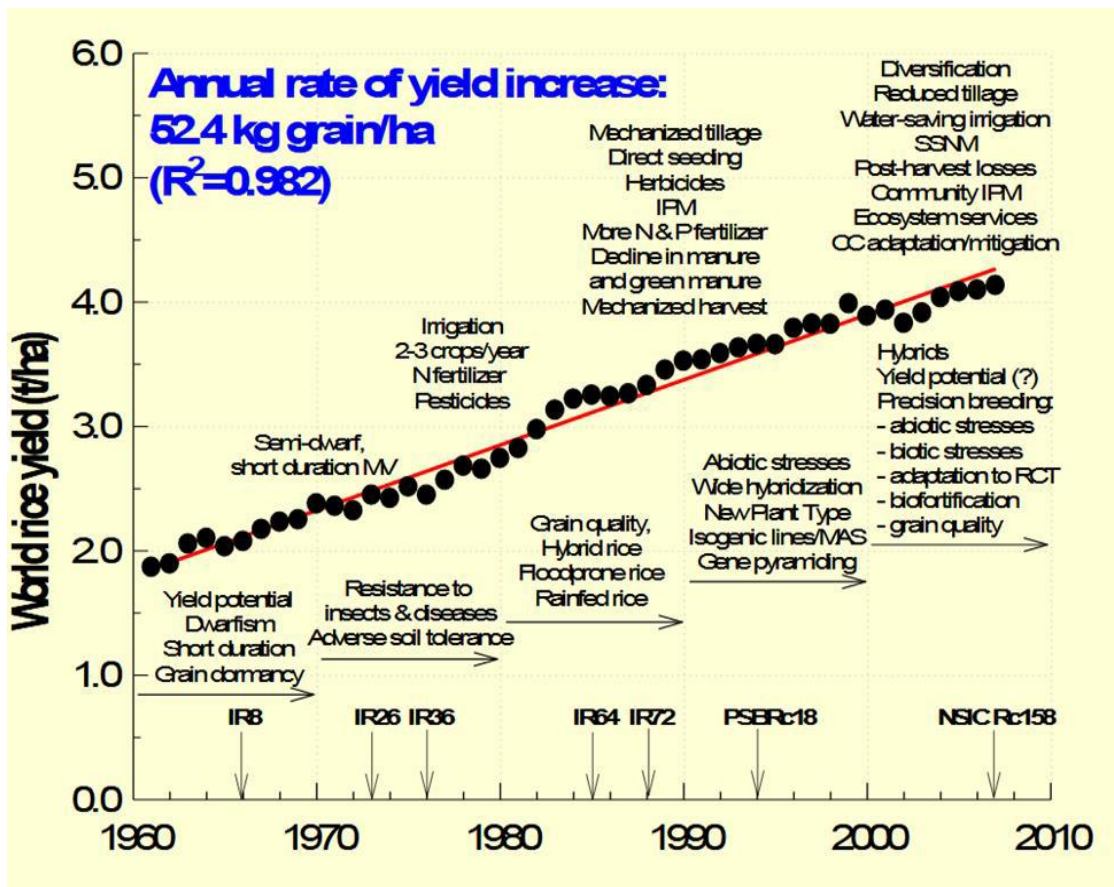


Figure 17 : Amélioration et évolution de la production du riz de 1960 à 2010 d'après l'IRRI

L'amélioration des variétés de riz a augmenté au cours des années lui conférant de meilleures résistances et caractéristiques agronomiques comme l'augmentation du rendement moyen (tonne/hectare).

existe deux sous-espèces de ces riz : *Os indica* et *Os japonica*. Dans des sous populations sauvages, avant d'être domestiquée, *Os indica* résulterait de l'introgression de gènes *japonica* (préalablement domestiqués) (Choi and Purugganan, 2018). Les *Os indica* ont en général des grains longs et minces. Les *Os japonica* ont des grains sont plutôt larges et ronds (Demol et al., 2002). Une autre espèce principalement cultivée en Afrique, *Oryza glaberrima*, résulte d'une domestication indépendante de *Os*.

Le riz est qualifié de plante modèle pour les céréales (Izawa and Shimamoto, 1996). *Oryza sativa*, est une plante diploïde ($2n=24$), avec 12 chromosomes, mais certaines espèces comme *Oryza latifolia* sont triploïdes ($2n=48$) (Demol et al., 2002). Son génome est petit, et varie de 390 Mb à 430 Mb. Le génome de la variété Nipponbare a été entièrement séquencé en 2005 (390 Mb), et a permis d'identifier plus de 38000 gènes dont la fonction de 2000 d'entre eux ont fait l'objet d'études expérimentales (International Rice Genome Sequencing, 2005 ; Meynard et al., 2020). En 2014 le projet 3000 génomes de riz a permis de séquencer 3010 variétés et de mettre au jour une importante diversité génétique, avec près de 4 millions de SNP identifiés entre *Nipponbare* et d'autres cultivars-espèces, avec plus de 45 % de nouveaux gènes absents du génome de référence *Nipponbare*, comme les gènes *Sub1A*, *Pisto11* (Li et al., 2014 ; Meynard et al., 2020). Les espèces sauvages de riz contiennent de nombreux gènes d'intérêt agronomique et sont classiquement utilisées en amélioration génétique mais leur exploitation reste délicate. Par exemple, après plus de 3000 hybridations entre *Oryza sativa* et *Oryza brachyantha*, un seul plant F1 a pu être obtenu (Demol et al., 2002). Néanmoins elle est possible entre *Os indica* et *japonica*, voir *Os glaberrima*.

2) Amélioration du riz : *Chhomrong dhan* (CHHD) et le riz d'altitude malgache

Les cibles d'améliorations sont nombreuses et portent par exemple sur la productivité, la résistance aux maladies, l'adaptation au milieu de culture, la teneur en protéines, et la qualité du grain (**Figure 17**). L'idéotype idéal serait un riz avec un système racinaire vigoureux, une taille moyenne de 1 m, peu de thalles sans panicules, et une récolte 100 à 130 jours après semis (Demol et al., 2002). Pour y parvenir l'idée de l'introduction de gènes provenant d'espèces ou de cultivars distants est privilégiée en raison de la grande diversité génétique existante au sein de cette espèce (Demol et al., 2002).

Oryza sativa japonica Chhomrong dhan (CHHD) est une lignée d'origine népalaise de riz irrigué, mais qui est également cultivée en conditions pluviales, naturellement résistante au

froid durant les stades de reproduction, au brunissement foliaire et racinaire prématuré, et avec un bon rendement (Steele et al., 2009). Elle mesure 1 m de haut, et dispose d'un rendement de 4.2 t/Ha après 160 jours de culture (Joshi and Gauchan, 2017). Elle s'est très bien acclimatée aux hauts plateaux malgaches qui ont une saison chaude et humide de novembre à avril, et une saison sèche et froide de mai à octobre. CHHD est cultivée en altitude, riz dit « upland », et malgré qu'elle soit originellement une variété irriguée, elle est cultivée en condition pluviale (Raboin et al., 2010). Elle a été originellement introduite à Madagascar par le CIRAD et le FOFIFA au milieu des années 80 pour y être étudiée et évaluée dans un projet d'amélioration du riz pluvial d'altitude malgache, avant d'y être officiellement déployée en 2006 (Raboin et al., 2010). Lors d'une mission à Madagascar en octobre 2019, nous avons appris qu'elle avait été adoptée rapidement par les agriculteurs locaux bien avant 2006 en raison du vol d'un sac de grains de cette variété qui a, par la suite, été vendue à d'autres agriculteurs. CHHD qui assure un bon rendement et une bonne qualité de grain et ne nécessite que peu d'intrants, la variété a vite été adoptée, et les semences ont progressivement été distribués par les agriculteurs eux-mêmes jusqu'à représenter plus de 80% des plants de riz cultivés sur les hauts plateaux Malgaches.

3) L'efficacité de l'utilisation de l'azote comme cible d'amélioration par les NBT

La transformation du riz date des années 80. Le premier riz transgénique a été obtenu en 1989, en transformant et en régénérant des protoplastes d'*Oryza sativa japonica* avec un ADN conférant une résistance à l'hygromycine (Shimamoto et al., 1989). Plus classiquement, *Agrobacterium tumefaciens* est maintenant utilisé en routine pour la transformation génétique du riz à partir de cals embryogènes (Sallaud et al., 2003).

Le choix, pour faire une démonstration de l'intérêt de la technologie CRISPR/CAS9 chez le riz, s'est porté sur la variété Chhomrong dhan, et sur l'amélioration de l'efficacité d'utilisation de l'azote. L'azote est un composé essentiel des acides aminés et protéines, représentant jusqu'à 5 % de la matière sèche totale de la plante. Il est principalement apporté sous forme inorganique (nitrate (NO_3^-) et ammonium (NH_4^+)), ou organique (urée, peptides, acides aminés) puis absorbé par les racines (Kiba and Krapp, 2016; Nacry et al., 2013). L'absorption du nitrate dépend de sa disponibilité dans le sol qui fluctue grandement. L'azote entraîne une reprogrammation du transcriptome conduisant à l'expression de transporteurs spécifiques de la réponse à de faibles ou de fortes concentrations azotées (Bellegarde et al., 2017).

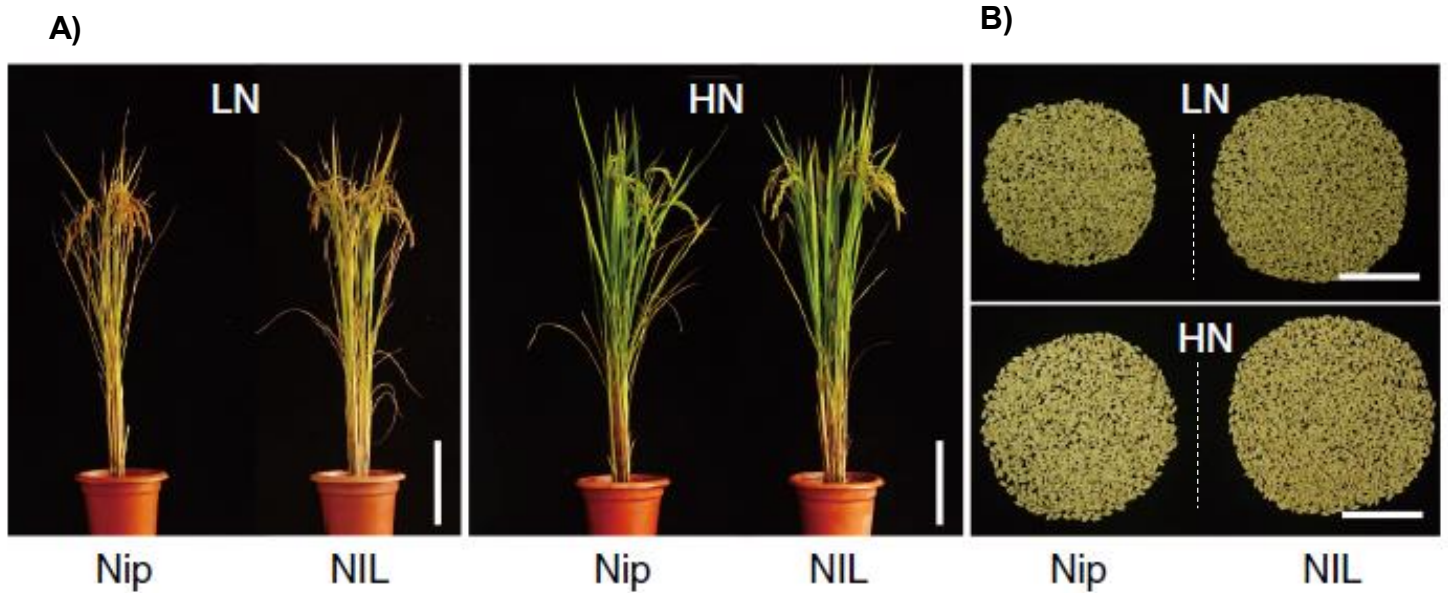


Figure 18 : Croissance et rendement sur faible et forte concentration d'azote chez le riz avec le gène *OsNRT1.1b japonica* ou d'*indica* d'après Hu et al., 2015
 Sur faible (LN) ou fort azote (HN), la lignée *japonica* *Nipponbare* (Nip) présente un nombre de thalles inférieur à la lignée *japonica* introgressée avec le gène *OsNRT1.1b* de *indica* (NIL) (A).
 Nip présente aussi un moins bon rendement que NIL en LN et HN (B), et une moins bonne NUE (rendement/kg d'azote utilisé) (C).

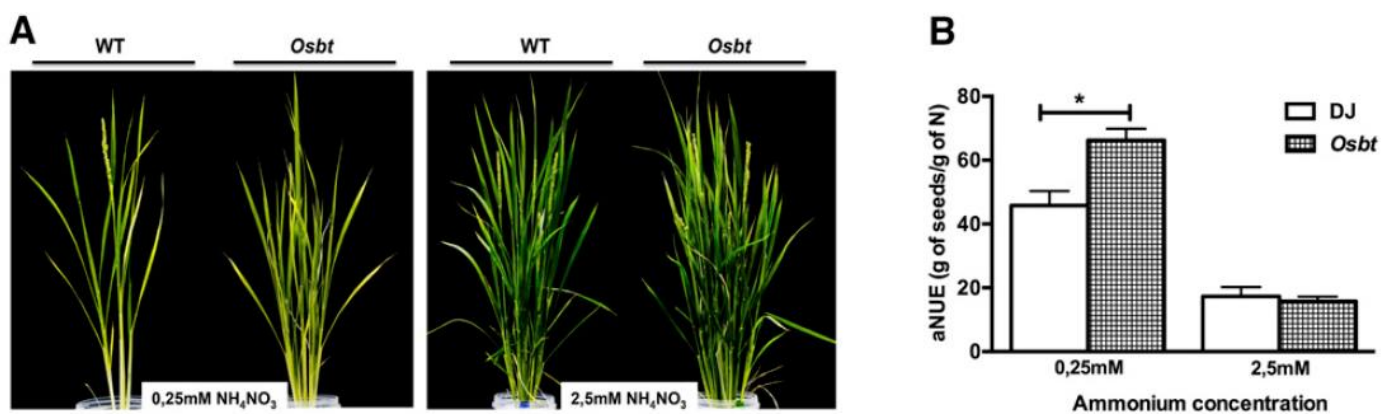


Figure 19 : Croissance et NUE chez une lignée *japonica* cv *Donjing* WT par rapport à une lignée KO *Osbt* d'après Araus et al., 2016
 Sur faible et forte concentration en nitrate d'ammonium les plants KO *Osbt* présentent un plus grand nombre de thalles (A), et, que sur faible concentration, une meilleure NUE (rendement (g) / g d'azote) (B).

Ces transporteurs de nitrate sont situés sur la membrane plasmique et appartiennent à la famille des NRT (Nitrate Transporter). Parmi eux, il y a 53 membres appartenant à la famille NPF (*NRT1.1*), dits transporteurs à faible affinité LATs (Low Affinity Transporteur), et il existe 7 membres dans la famille des *NRT2* (2.1, 2.2, 2.4, et 2.5), dits transporteurs à haute affinité du nitrate (HATS) (Nacry et al. 2013). Ils sont aussi impliqués dans la NUE (Nitrogen Use Efficiency), c'est-à-dire l'absorption de l'azote, et son utilisation pour la croissance et rendement (assimilation/internalisation/remobilisation) (Araus et al., 2016). *OsNRT1.1b* (Os10g40600.1), l'homologue de *AtNRT1.1* chez le riz est lui aussi un transcepteur (transporteur et senseur) racinaire d'influx du nitrate, qui est exprimé dans les poils racinaires, l'épiderme et les tissus vasculaires. C'est un transcepteur à double affinité : il est actif en fortes concentrations de nitrate (LATS), mais aussi en faibles s'il est phosphorylé devenant un HATS (Chen and Ma, 2015; Hu et al., 2015). Cependant, il existe deux variants alléliques pour le gène *OsNRT1.1b*, dont un qui permet une meilleure NUE et un meilleur rendement. Ce variant a été identifié dans les génotypes d'*Oryza sativa indica*, comparativement aux génotypes d'*Oryza sativa japonica*. En effet, les génotypes *indica* présentent en moyenne une meilleure NUE que les génotypes *japonica*, et cette différence n'est pas due à une différence dans le niveau d'expression de *OsNRT1.1b*. Une substitution d'un acide aminé (AA), la Thréonine 327 en Méthionine (Thr327Met), semble être impliquée dans cette différence entre les génotypes *japonica* et *indica*. Au niveau de l'ADN, cela correspond à la substitution de la Cytosine 980 en Thymine (980C>T) chez *Oryza sativa indica* comparativement à *japonica*. Ce changement d'AA intervient au niveau de la boucle cytoplasmique centrale du transcepteur (Hu et al., 2015). Lors du transfert par rétrocroisements de l'allèle *OsNRT1.1B indica* dans un fond génétique *japonica*, les lignées quasi-isogéniques (NIL) correspondantes ont une meilleure NUE (meilleure absorption, translocation de l'azote) leur permettant d'atteindre des rendements supérieurs aux lignées contrôles notamment en faibles concentrations de nitrate (Hu et al., 2015) (**Figure 18**). L'origine de l'allèle *OsNRT1.1b* semble remonter à l'époque de la domestication du riz (Hu et al., 2015).

Il existe aussi chez les plantes des régulateurs négatifs des gènes d'absorption du nitrate comme les gènes Bric-à-brac. *OsBT* est un gène de la famille des *BTB* (Bric-a-brac / Tramtrack/Broad gene) impliqué lui aussi dans la NUE. Il existe deux gènes *AtBT1* et *AtBT2* chez *Arabidopsis thaliana* et la surexpression de *BT2* induirait une baisse de la NUE ; réciproquement, chez le mutant perte de fonction *Atbt2*, les niveaux d'expression des transporteurs de nitrate à forte affinité *NRT2.1* et *NRT2.4* augmentent et entraînent une

augmentation de la NUE de 20 % en conditions de carence azotée (Araus et al., 2016). Chez le riz il existe un seul gène *OsBT*, et le mutant perte de fonction correspondant présente lui aussi une augmentation de la NUE et du nombre de thalles sous faible concentration de nitrate d'ammonium (**Figure 19**) (Araus et al., 2016).

Les gènes *OsBT* et *OsNRT1.1b* sont donc des candidats idéaux pour tenter d'améliorer la NUE en utilisant les technologies NBT. Il sera donc intéressant de transférer ces mutations, perte de fonction de *OsBT* et l'allèle indica d'*OsNRT1.1b*, dans le fond CHHD afin de tenter d'augmenter la NUE. Ces mutations seront également transférées dans le cultivar Kitaaké, cultivar utilisé en routine au sein du laboratoire pour la plupart des analyses fonctionnelles (Herbert et al., 2020). Kitaaké a un cycle court, inférieur à 100 jours, et n'est pas affecté par la photopériode. Enfin, il s'agit d'une variété semi-naine plus simple à cultiver et à manipuler.

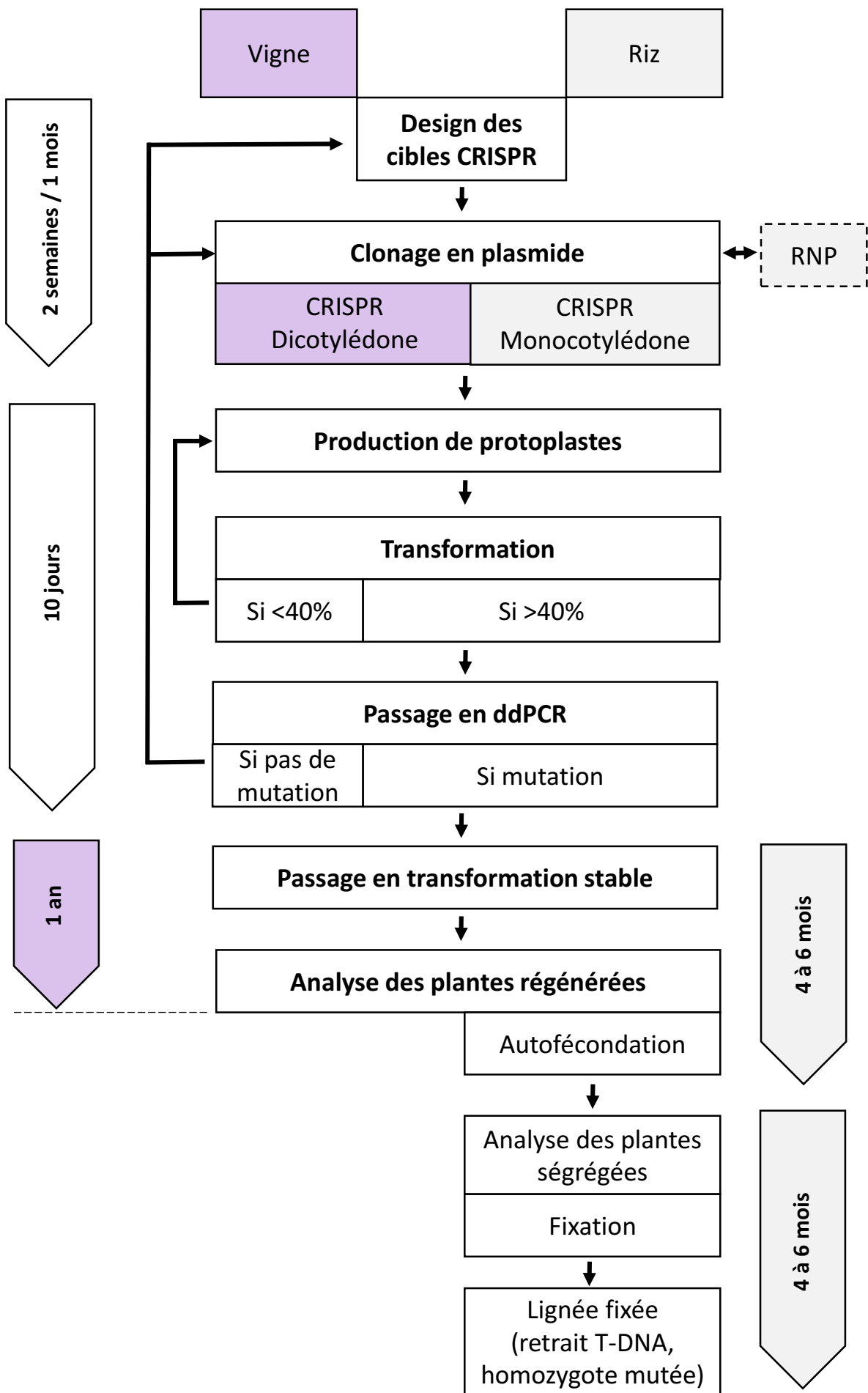


Figure 20 : Pipeline des transformations CRISPR en riz et vigne

Les étapes de design, clonages, transformations et analyses pour les deux espèces sont très similaires. Cependant la vigne aura un temps de régénération très important par rapport au riz ce qui empêche la ségrégation du T-DNA.

Problématiques et objectifs

L'utilisation des technologies CRISPR s'est démocratisée ces dernières années. Elles permettent une obtention des mutants plus rapide, et accélèrent l'analyse fonctionnelle. Leur utilisation chez la vigne se heurte à la lourdeur des protocoles de transformation et de régénération de l'espèce. Côté riz, la technologie a vite été adaptée et sa mise en place a été facilitée chez cette espèce qui disposait de protocoles de transformation robustes. Néanmoins, ces deux espèces bénéficient de l'amélioration et de la mise en place des technologies NBT.

Un des premiers objectifs de cette thèse est le développement et la validation d'une série de technologies de construits CRISPR en système protoplastes chez ces deux espèces. Dans un deuxième temps, l'objectif sera de faire la démonstration que les technologies NBT peuvent être utilisées pour améliorer des caractères d'intérêt chez le riz et la vigne (**Figure 20**).

La thèse est financée par deux projets :

Le projet NBT2 de l'institut Carnot Plant2Pro porte sur la partie vigne. Cette partie du travail se fait avec l'équipe DAAV de l'UMR AGAP (INRAe, en collaboration avec l'IFV, l'Institut Français de la Vigne et du Vin), qui maîtrise la culture de cellules ou la régénération d'embryons somatiques. Dans un premier temps, une fois le gène candidat idéal sélectionné pour toute la durée de la thèse (*VvGAI*), il sera nécessaire de mettre au point une méthode de transformation rapide de cellules (protoplastes) pour tester des constructions CRISPR (plasmides) développées spécifiquement pour la vigne. Ensuite je décrirai la mise au point de la transformation stable de ces CRISPR pour régénérer des vignes mutées pour le gène *VvGAI*. Pour finir, j'esquisserai quelques pistes afin de tirer des conclusions sur l'efficacité de la méthode et de l'intérêt du phénotype de la mutation.

Le projet GeneRice (generation and deployment of Genome edited Nitrogen use efficient rice varieties) finance le travail concernant la partie riz. La technologie CRISPR-CAS9, la transformation du riz, la production de protoplastes et leur transformation étaient déjà bien maîtrisées par l'équipe DAR de l'UMR AGAP (CIRAD) d'où l'intérêt, d'optimiser et profiter de ce projet pour transférer ces techniques sur la vigne. J'ai participé au développement d'une nouvelle technologie CRISPR « Base Editing » (du moins, au moment du démarrage de ma thèse, il s'agissait d'une toute nouvelle technique). Cette technologie m'a permis d'introduire un allèle indica pour le gène *OsNRT1.1b* dans deux fonds japonica. Nous allons aussi voir que cette technologie permet d'introduire de manière spécifique des codons stop (iSTOP).

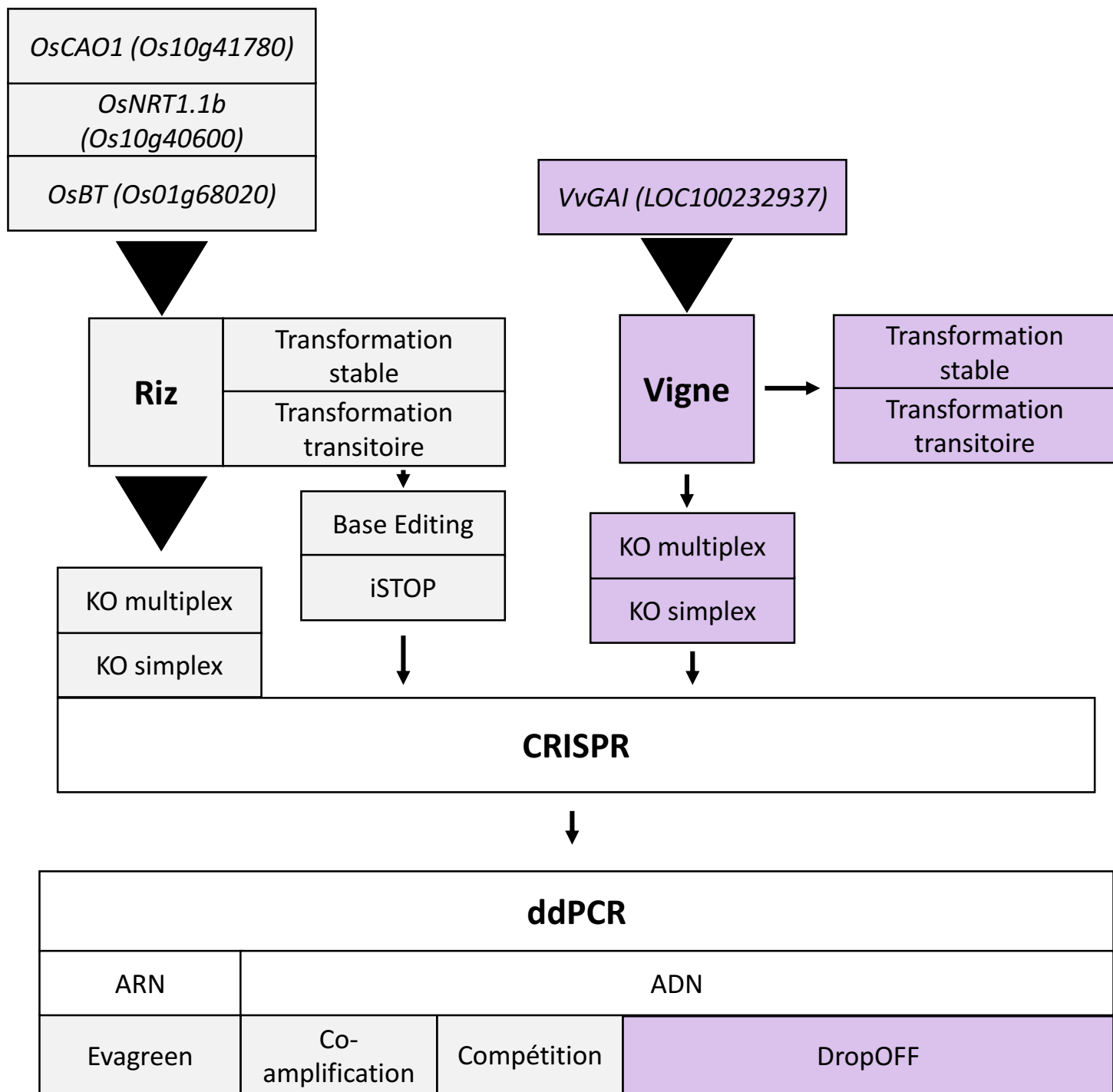


Figure 21 : Diagramme de la production et détection des riz et vignes issus de NBT

A partir des locus ciblés, un système de mutation ciblé par CRISPR va être développé. Un système d'analyse rapide par protoplastes et ddPCR permettra de confirmer l'intérêt de la faisabilité de la possible transformation stable qui suivra.

Comme prévu, j'ai généré des mutants pertes de fonction KO dans le gène *OsBT*. Pour finir j'ai analysé l'impact de ces mutations dans l'efficacité d'utilisation de l'azote.

Les résultats de cette thèse ont pu être obtenus par l'utilisation d'une nouvelle technologie capable de détecter des mutations spécifiques, la droplet digital PCR (ddPCR). Cette partie s'est faite en collaboration avec la société IAGE. Nous avons pu tester et comparer l'efficacité de nos constructions CRISPR chez le riz et la vigne en protoplastes par un protocole qui ne prend qu'une dizaine de jours (**Figure 21**).

Dans un premier chapitre, nous allons traiter de la mise au point d'une technologie de validation de construits CRISPR en protoplastes de riz et vigne. Nous allons voir plus en détail la transformation de protoplastes de riz et vigne, son optimisation, et l'analyse des mutations par la technologie ddPCR chez riz dans un premier temps, comme outil simple et rapide pour tester, comparer et valider des constructions CRISPR/CAS9.

Dans un second chapitre, nous traiterons de l'application des NBT chez le riz et la vigne. Nous verrons si le transfert chez la vigne des technologies de transformation transitoire et de détection par ddPCR des mutations a été possible. Enfin chez la vigne, nous analyserons les premiers résultats de la transformation stable. Puis sur le riz, nous comparerons l'intérêt des techniques CRISPR sur des applications dérivées (iSTOP) ainsi que sur l'utilisation de l'azote (NUE). Cette dernière se fait en utilisant une lignée malgache mise au champ et testée en Colombie (phénotype et rendement). Un « répliqua » a été réalisé en serre sur une lignée plus facile à maîtriser (Kitaaké) sur laquelle nous allons tester l'influx de nitrate, la co-expression des gènes ciblés, mais aussi phénotyper l'impact des mutations sur le rendement.

Chapitre 1 : Mise au point des technologies

Ce chapitre aborde l'ensemble des mises au point des technologies réalisées au cours de ma thèse et de mon stage Master I. Elles ont permis la création et la détection de mutations induites par des technologies CRISPR chez le riz et la vigne à l'aide d'un système transitoire. Le chapitre est divisé en trois parties :

- La production d'une part et la transformation de protoplastes de riz et de vigne d'autre part, qui auront pour but d'avoir des méthodologies rapides de test de constructions dans un système transitoire, et d'autre part d'améliorer le protocole d'édition chez la vigne.
- L'analyse des mutations induites et leur quantification qui se fera via la droplet digital PCR (ddPCR).
- Une alternative d'amélioration possible par la transformation transitoire de complexes CAS9-sgRNA purs, dit Ribo-Nucléo-Protéiques (RNP), chez le riz pour éviter l'insertion d'ADN.

I] Production et transformation de protoplastes

Afin de tester rapidement nos constructions CRISPR/CAS9 chez le riz et la vigne, nous avons développé une stratégie permettant à la fois d'analyser un grand nombre d'évènements d'édition en système transitoire, et ce quantitativement, afin de comparer entre elles l'efficacité de ces constructions. Les protoplastes, des cellules végétales vivantes débarrassées de leur paroi pectocellulosique, peuvent facilement être transformés, voire cotransformés pour tester l'effet d'un grand nombre de constructions. Chacun d'entre eux représente un évènement potentiel de transformation et d'édition. Des millions de protoplastes peuvent être obtenus en une journée à partir des plantules chez le riz, ou des cultures cellulaires chez la vigne, et utilisés pour tester nos constructions avant leur utilisation en transformation stable. Ce chapitre décrit les protocoles développés et optimisés pour la production et la transformation de protoplastes pour la transformation transitoire et pour tester l'efficacité de nos constructions CRISPR/CAS9.

1) Protocoles développés à partir des bases de tige de riz

Nous avons besoin de développer des protocoles de production de protoplastes, de transformation et de détection efficaces afin de tester rapidement et de façon reproductible les constructions CRISPR utilisées. L'article (Bès et al., 2021) qui est publié dans *Methods in Molecular Biology*, présente le protocole complet de production/transformation/détection que nous avons mis au point à partir de deux articles précédemment publiés (Woo et al., 2015 ; Zhang et al., 2011).



Chapter 11

Efficient Genome Editing in Rice Protoplasts Using CRISPR/CAS9 Construct

Martine Bes, Leo Herbert, Thibault Mounier, Anne-Cécile Meunier, Franz Durandet, Emmanuel Guiderdoni, and Christophe Périn

Abstract

Genome editing technologies, mainly CRISPR/CAS9, are revolutionizing plant biology and breeding. Since the demonstration of its effectiveness in eukaryotic cells, a very large number of derived technologies has emerged. Demonstrating and comparing the effectiveness of all these new technologies in entire plants is a long, tedious, and labor-intensive process that generally involves the production of transgenic plants and their analysis. Protoplasts, plant cells free of their walls, offer a simple, high-throughput system to test the efficiency of these editing technologies in a few weeks' time span. We have developed a routine protocol using protoplasts to test editing technologies in rice. Our protocol allows to test more than 30 constructs in protoplasts prepared from leaf tissues of 100, 9–11-day-old seedlings. CRISPR/CAS9 construct effectiveness can be clearly established within less than a week. We provide here a full protocol, from designing sgRNA to mutation analysis.

Key words Rice, CRISPR/CAS9, Protoplast, Genome editing

1 Introduction

The CRISPR/CAS9 technology became the dominant genome editing technology soon after the first demonstration of its eukaryotic cell efficacy in 2012 [1]. Since then, its use has increased exponentially, demonstrating its functionality in most species where it has been tested [2], including plants [3]. Beyond its obvious interest in biology, recent results obtained in rice [4–7] and tomato [8] have demonstrated its interest also as a plant breeding tool. Many CRISPR/CAS9-derived technologies have emerged and continue to appear in a regular basis. Among the most emblematic examples include the discovery of other nucleases like CPF1 [9, 10], the development of CAS9 variants capable of

Martine Bes, Leo Herbert and Thibault Mounier contributed equally with all other contributors.

Anindya Bandyopadhyay and Roger Thilmony (eds.), *Rice Genome Engineering and Gene Editing: Methods and Protocols*, Methods in Molecular Biology, vol. 2238, https://doi.org/10.1007/978-1-0716-1068-8_11, © Springer Science+Business Media, LLC, part of Springer Nature 2021

targeting a modified protospacer adjacent motif [11, 12], the use of CAS9 variants where the catalytic activity of the nuclease was inactivated and fused with various domains for transcriptional regulation of target [13, 14] or single nucleotide changes [15–18], and optimization of DNA of repair sequence to induce targeted modifications or replacement of targets [19, 20].

These new systems have often been tested initially in isolated cells, with human cells being one of the most popular prototyping systems (*see* for instance [10, 21, 22]). It is possible to have a similar approach in plants by using protoplasts, which are isolated cells with their walls removed. This approach has been used by some research teams to demonstrate the effectiveness of CRISPR/CAS9 constructions (*see* for instance [23–25]), but protoplasts are used much less systematically than human cells. One of the reasons for their underutilization is the supposed difficulty of producing and using protoplasts, which requires fine-tuning protocols. In response to these difficulties, we have therefore decided to develop a simple, detailed, and robust protocol for the production of protoplasts and its routine use to evaluate the effectiveness of CRISPR/CAS9 constructs in rice.

In this article we describe in detail the production, processing, DNA extraction of protoplasts, and molecular analyses for testing CRISPR/CAS9 technology in rice. This protocol makes it possible to carry out the transformation and evaluation of 20–30 constructions from one hundred 9–11-day-old rice seedlings in less than a week. In several places, we indicate the very critical points but also demonstrate the multipurpose of this protocol as quality points are also indicated at key steps to ensure the reproducibility of the protocol and its transferability to other laboratories. We believe this detailed protocol will be a useful resource for the rice and plant community and will popularize the use of protoplasts as a prototyping system for CRISPR technology evaluation.

2 Materials

For material details, *see* Table 1 (Fig. 1).

2.1 Mix for Digestion/Ligation and CRISPR Vector Construction

1. The pENTRY vector is digested using BsaI enzyme. Add sequentially 2 μg of pOs-sgRNA ENTRY Miniprep vector 2 μL of BsaI-HF restriction enzyme plus 5 μL of 10 \times BsaI-HF buffer QSP 50 μL H₂Odd.
2. The ligation mix is prepared by mixing together 6 μL of dsDNA primers at 30 ng/ μL , 2.5 μL of pENTRY BsaI digested vector (*see* below), and 1 μL of T4 Ligase (1 U/ μL) QSP 18 μL H₂Odd.

Table 1
Materials

Reagent or resource	Source	Identifier
Tube, 13 mL, 95 × 16.8 mm, round base, PS	Sarstedt	55468001
Cytometric tubes with a round bottom	Falcon	352054
Nylon mesh 70 μm BD Falcon pour tube 50 mL	Falcon	352350
Nylon mesh 30 μm	Sefar AG	03-37/24
Petri dish	Greiner	1608043
COREX 25 mL centrifuge round bottom tubes	Corex	8446
Cut ended 200 μL pipette tips	StarLab	S1111-1706
2 Funnels	Nalgene 80 mm	4252-0080
3 Erlenmeyers 250 mL	Schott Duran	1216 36
<i>Biological samples</i>		
DNA and leaf tissues from cv Nipponbare	This work	N/A
<i>Chemicals</i>		
Polyethylene glycol 4000	Fluka	81242
Macerozyme R10 Onozuka	Duchefa	651803
Cellulase RS Onozuka	Duchefa	578704
Percoll	Sigma	P1644
Fluorescein diacetate	Sigma	F7378
<i>Commercial assays</i>		
Endonuclease I T7	New England Biolabs	M0302L
PureLink PCR purification	Fisher Scientific	K31000
DNeasy plant extraction kit	Qiagen	69104
RQ1 RNase/DNase-free	Promega	M6101
KOD hot start polymerase	VWR	71086
HiSpeed plasmid maxi kit	Qiagen	12663
<i>Experimental models and organisms</i>		
Rice <i>Oryza sativa</i> cv Nipponbare	See Subheading 3	
Subcloning efficiency DH5a chemically competent cells	Fisher Scientific	11583117
<i>Oligonucleotides</i>		
CAO1-fw: CGGGACGCTGATGACGTGC	This work	N/A
CAO1-rw: GCGTGCAGCTTTCCCCCAGT	This work	N/A
CR-CAO1-F: GGCAACCGGACCACCTCAAGAGCT	This work	N/A

(continued)

Table 1
(continued)

Reagent or resource	Source	Identifier
CR-CAO1-R: AAACAGCTCTTGAGGTGGTCCGGT	This work	N/A
UBI-RAC1: TGCAATGCTCATTATCTCTAGAG	This work	N/A
M13-RP: CAGGAAACAGCTATGACC	This work	N/A
<i>Recombinant DNA</i>		
See also Fig. 1a, b		
pOs-sgRNA ENTRY	[26]	N/A
pOs-CAS9 hpt ccdB-2.0	[26]/This work	N/A
PHKN29	[27]	N/A
<i>Software and algorithms</i>		
TEFOR	[28]	
RGEN	[29]	
ORYGENESDB	Droc et al. (2006, 2009)	
<i>Other</i>		
dMicroscope ($\times 10$ or $\times 20$ objectives)	Leica	DM RXA2
Bioanalyzer Agilent 2100	Agilent Technologies	G2939A
DNA 1000 ScreenTape $\times 7$	Agilent	5067–5582
Fuchs and Rosenthal cell counting cell	VWR	631–1125
Stove 26 °C (with a stirrer in it)	Memmert	N/A
Centrifuge	Heraeus Multifuge 3SR	N/A
Swing buckets	Heraeus	75006441H

2.2 Solutions for Protoplast Digestion and Purification

1. A volume of 200 mL of mannitol solution is prepared with 21.86 g of mannitol with 76 mg of MES. The pH is adjusted to 5.7 and the solution is sterilized by filtration.
2. The first digestion solution is sterilized by filtration. A volume of 200 mL is prepared by adding 222 mg of $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 21.86 g mannitol, 390 mg MES, and pH adjusted to 5.7.
3. The second digestion solution is also sterilized by filtration. A volume of 20 mL is prepared by combining 150 mg of Macerzyme R10, 300 mg of Cellulase RS, 200 μL of a 100 \times stock solution of BSA 10%, and QSP 20 mL completed with first digestion solution.
4. 200 mL of W5 solution sterilized by filtration is prepared with 8.92 g of NaCl, 18.36 g of $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 372 mg of KCl, 388 mg of MES buffer, and pH adjusted to 5.7.

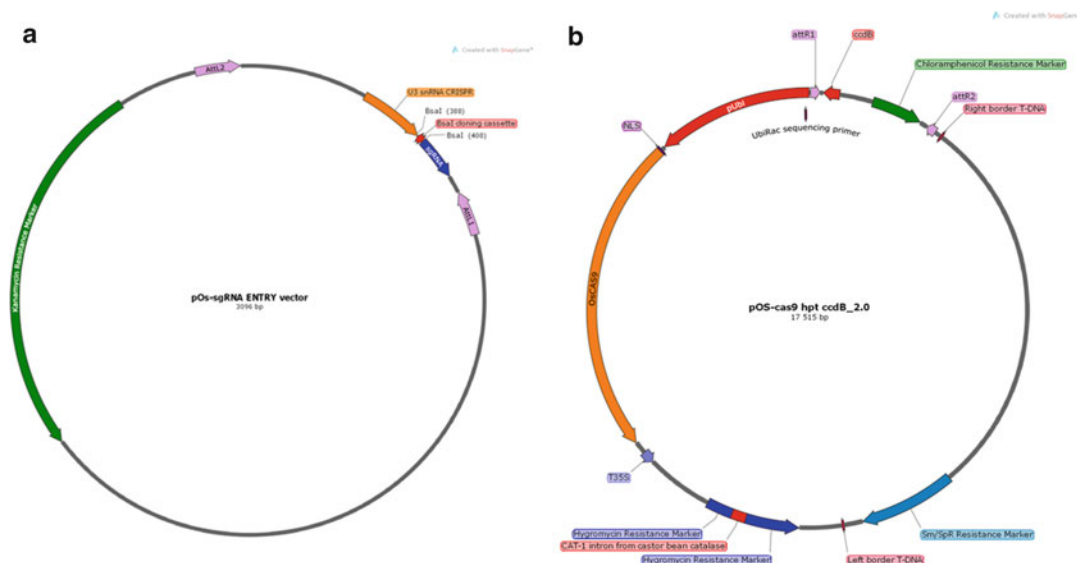


Fig. 1 Plasmid maps of pOs-sgRNA ENTRY (a) and pOs-CAS9 hpt ccd_2.0 BINARY vectors (b)

2.3 Solutions for Protoplast Transformation

1. Prepare 10 mL of PEG solution, sterilize by filtration, aliquot in 2 mL, and store to $-20\text{ }^{\circ}\text{C}$ until use. Add 364 mg of mannitol, 1 mL of $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ of stock solution 1 M, 4 g of PEG4000 Fluka[®], and pH adjusted to 5.7.

3 Methods

3.1 sgRNA Design (See Note 1)

1. Copy your targeted DNA sequence into the online TEFOR software (see Note 2) [28].
2. Select the corresponding species, *Oryza sativa*.
3. Warning! The genome sequence in the database is from the reference cultivar Nipponbare.
4. The software outputs a potential protospacer list and their position on your target genomic sequence.
5. Choose two different protospacers, e. g., close to the start codon, 50 base from, to generate knockouts.
6. Choose protospacers that have minimal potential off-target [29] (see Note 3).
7. Complementary oligonucleotides corresponding to the variable part of the sgRNA have to be designed for cloning in the right way.
8. 4 bp overhangs GGCA are added to the 5' end of the sequence followed directly by NGG (protospacer adjacent motif or PAM) on the targeted strand.

9. 4 bp overhangs AAAC are added to the 5' end of the reverse complementary strand following CCN on the complementary reverse strand.
10. Be careful not to add the PAM (NGG) in your oligonucleotide sequence.
11. For example, for CAO1, the protospacer is ACCGGAC CACCTCAAGAGCTNGG; therefore the CR-CAO1-F oligonucleotide is GGCAACCGGACCACCTCAAGAGCT and the reverse CAO1 oligonucleotide is CR-CAO1-R AAA CAGCTCTTGAGGTGGTCCGGT.
12. Order oligonucleotides, with no primer purification and modification step.

3.2 Cloning in pDest Vector

3.2.1 Annealing Oligonucleotides

1. The oligonucleotides are diluted to 100 μ M with distilled water.
2. 2 μ g of each oligonucleotide are added in a final volume of 50 μ L of distilled water.
3. DNA is denatured at 95 $^{\circ}$ C for 5 min in a dry block.
4. The block is switched off and the tube containing oligonucleotides is left in the dry bath for 1 h.
5. Dilute dsDNA to 100 ng/ μ L.

3.2.2 Digestion/Ligation in the pOs-sgRNA ENTRY Vector (Fig. 2)

1. Prepare a digestion mix with 2 μ g of the pENTRY vector (Tables 2 and 3) and do a negative control without dsDNA primers.
2. Incubate at 37 $^{\circ}$ C OVN.
3. Check on a 1.2% agarose gel the complete digestion of the vector by loading 3 μ L of the mix.
4. Prepare the ligation mix (Table 3) (*see Note 4*).
5. Incubate at 26 $^{\circ}$ C for 2–4 h or overnight at 16 $^{\circ}$ C.
6. Add 1 μ L of additional BsaI-HF and incubate one additional hour.

3.2.3 Transformation in DH5alpha

1. Follow your usual genetic transformation protocol (electroporation, heat shock) and select on kanamycin (50 mg/L).
2. Isolate 15 colonies.
3. Make the minipreps and check each plasmid using BsaI digestion; add a negative control (empty vector pOs-sgRNA ENTRY).
4. Positive clones, with integrated spacer, have lost their BsaI sites.
5. Sequence the spacer using M13 RP primer (Table 1).

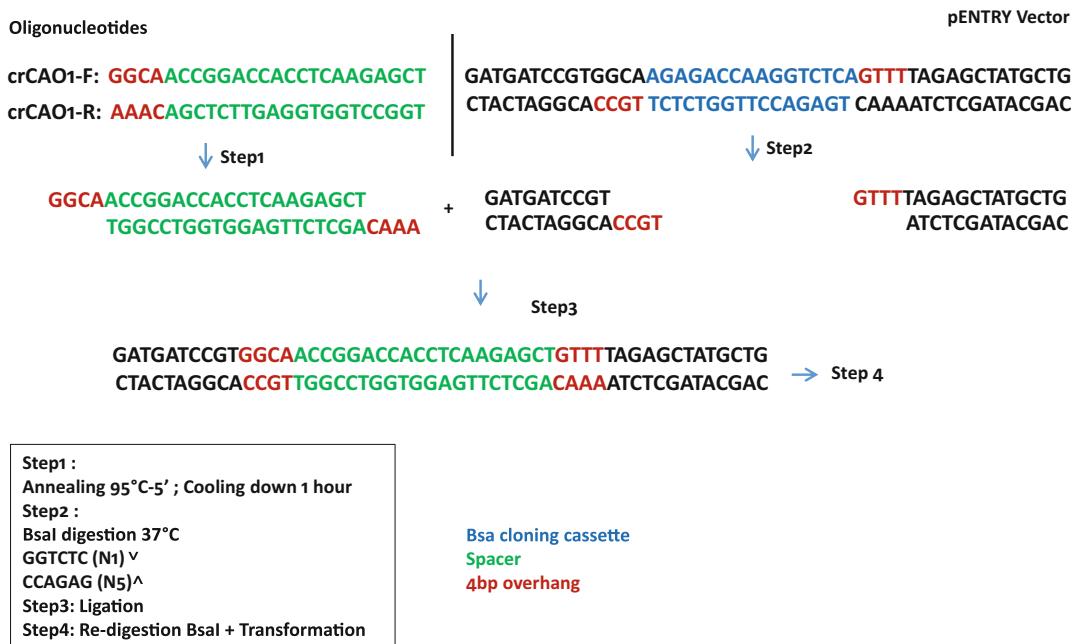


Fig. 2 Workflow of protospacer subcloning in pOs-sgRNA ENTRY using CR-CAO1-F and CR-CAO1-R oligonucleotides. Step 1, annealing of oligonucleotides (*see* Subheading 3.2.1); step 2, digesting pOs-sgRNA ENTRY vector with BsaI (*see* Subheading 3.2.2); and step 3, ligation of oligonucleotides and pOs-sgRNA ENTRY BsaI digested plasmid (*see* Subheading 3.2.2)

Table 2
 Digestion mix for pOs-sgRNA ENTRY

Reagent	Final concentration	Vol 1 ×
pOs-sgRNA ENTRY miniprep	40 ng/μL	Variable μL
BsaI-HF	1 ×	2 μL
Buffer BsaI-HF	U/μL	5 μL
H ₂ O dd	–	QSP 50 μL

Table 3
 Ligation mix for pOs-sgRNA ENTRY

Reagent	Final concentration	Vol 1 ×
dsDNA primers	30 ng/μL	6 μL
pOs-sgRNA ENTRY digested (Table 2)	25 ng/μL	2.5 μL
T4 ligase	U/μL	1 μL
H ₂ O dd	–	8.5 μL

**3.2.4 Subcloning
in Vector pOs-CAS9 hpt
ccdB-2.0**

1. After verification by sequencing, the construct containing pU3-sgRNA is subcloned by LR recombination following supplier's recommendations by adding 100 ng of pOs-sgRNA ENTRY plasmid and 150 ng of pOs-CAS9 hpt ccdB-2.0 plasmid.
2. Follow the classical bacterial genetic transformation protocol (electroporation, heat shock) and select on spectinomycin (50 µg/mL).
3. Isolate 15 colonies, if negative control is really negative (no colonies with the empty vector).
4. Make the minipreps and check each plasmid with a SpeI/StuI double digestion; add a positive control (empty vector, pOs-CAS9 hpt ccdB-2.0).
5. Positive vectors containing the insert should yield two bands at 9920 and 6527 bp.
6. Sequence two independent clones using UBI-RAC1 primer (Table 1) to check correct insertion of the pU3-sgRNA sequence.

**3.3 Plant Growth
for Protoplast
Preparation (See Note
5)**

1. Sterile 20 × 20 Petri dishes containing 250 mL of MS/2 medium with 4.5 g/L phytigel stored at 4 °C are left at room temperature for 40 min.
2. Meanwhile, 100 seeds are dehusked and then surface disinfected by immersing in 70 °C ethanol in a 50 mL Falcon tube for 1 min.
3. After rinsing with sterile distilled water, seeds are disinfected in 50 mL of a 9.6° sodium hypochlorite solution with two drops of Tween 20.
4. After 30 minutes of incubation under slow agitation (40 rpm), the seeds are rinsed three times in sterile distilled water.
5. The seeds are spread in a sterile Petri dish under the hood and then inoculated in the 20 × 20 square Petri dish containing medium.
6. Twenty-five seeds are inoculated per plate at equal distance from each other with an angle of 45 °C with respect to the medium surface, embryo upwards.
7. Four plates (100 seeds) are incubated in the growth room in the dark at 26 °C for 9–11 days.

**3.4 Protoplast
Extraction (See Note 6)**

1. Prepare the mannitol solution (Table 4) and the enzyme solution (Tables 5 and 6) for protoplast isolation.
2. Collect 5–7-cm-long shoot segments located just above the root from the seedlings; discard the rest of the shoot and root tissues.

Table 4
Mannitol solution

Reagent	dFinal concentration	dQte for 200 mL
Mannitol	600 mM	21.86 g
MES buffer	2 mM	76 mg
pH 5.7		

Table 5
Solution n°1 for digestion

Reagent	dFinal concentration	Qte for 200 mL
CaCl ₂ ·2H ₂ O	7.5 mM	222 mg
Mannitol	600 mM	21.86 g
MES buffer	10 mM	390 mg
pH 5.7		

Table 6
Solution n°2 for digestion

Reagent	Final concentration	Qte for 20 mL
Macerozyme R10	0.75%	150 mg
Cellulase RS	1.5%	300 mg
BSA 10%	0.1%	200 µL stock solution 100×
Solution n°1 digestion (Table 5)	–	QSP 20 mL

3. Using a scalpel, chop the shoot tissues into the smallest possible pieces (1–2 mm) in a Petri dish (Fig. 3); change the scalpel blade at least once during the process.
4. Chop the tissues in rotating the cutting angle by 60° twice to favor the subsequent release of protoplasts.
5. Add 20 mL of mannitol solution and incubate for 10 min in the dark at room temperature (25 °C).
6. Remove the mannitol solution using a pipet across a 70 µM sieve.
7. Add 20 mL of the enzyme solution (Table 6) into the Petri dish.
8. Incubate in the dark at 26 °C on a gyratory shaker under slow motion (15 rpm) for 5 h.

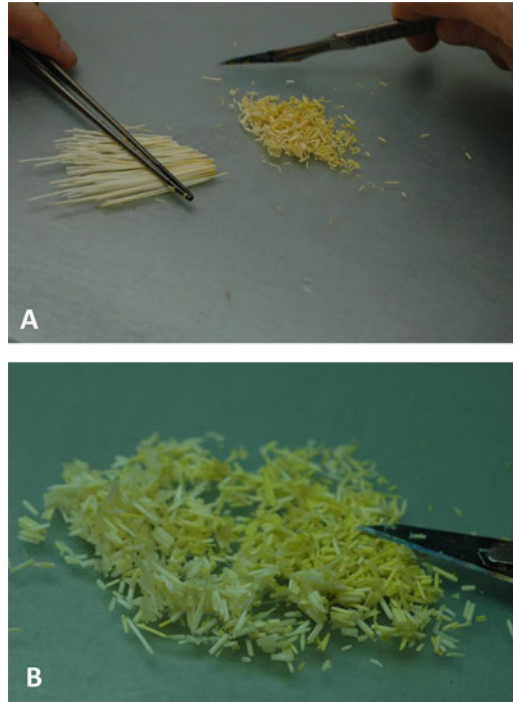


Fig. 3 Finely chopped 10-day-old rice seedling shoots

3.5 Protoplast Purification (See Note 7)

1. After 5 h of incubation, transfer the protoplast and debris suspension to a sterile, 50 mL Erlenmeyer.
2. Add 20 mL of W5 (Table 7) solution and gently shake the Erlenmeyer first clockwise during 10 s then counter clockwise for 10 more seconds.
3. Filter the suspension into a sterile Erlenmeyer through a 30 μm nylon sieve mesh.
4. Add twice 15 mL of medium W5 and gently squeeze the filter to wring it out.
5. Filter the suspension again through a new sieve into a new Erlenmeyer.
6. Transfer protoplast suspension to three Corning tubes of 25 mL in glass or sterile polystyrene with a cap.
7. Centrifuge at $250 \times g$ for 10 min at room temperature in swing buckets with acceleration and deceleration parameters setup at 2.
8. Discard the supernatant and leave a 1.5 mL volume at the bottom of each tube.
9. Gather the 3×1.5 mL tubes in a single tube.
10. Fill in the tube to a 25 mL final volume with W5 medium.

Table 7
W5 solution

Reagent	dFinal concentration	Qte for 200 mL
NaCl	154 mM	8.92 g
CaCl ₂ ·2H ₂ O	125 mM	18.36 g
KCl	5 mM	372 mg
MES buffer	2 mM	388 mg
pH 5.7		
Osmolarity 635 mosmol		

11. Centrifuge at $250 \times g$ for 10 min with acceleration and deceleration at 2 at room temperature in swing bucket rotors.
12. While centrifugation is operating, prepare 5 mL of the 60%, 40%, 20%, and 10% Percoll solutions with W5 solution.
13. Prepare the Percoll gradient by successively adding 2 mL of each solution to 30 mL tubes starting with the 60% solution by slowly pouring each solution along the tube, tilting the pipette.
14. Eliminate the supernatant after centrifugation.
15. Leave 2 mL of liquid containing the protoplast suspension in the bottom of the tube.
16. Pipet these 2 mL and place them at the top of the Percoll gradient slowly, by gravity by tilting the pipet without disturbing the gradient.
17. Centrifuge at $250 \times g$ for 10 min with acceleration and deceleration to 2 at room temperature in swinging buckets.
18. Remove the cellular debris by gently pipetting the debris layers (Fig. 4).
19. Gently mix the remaining layers by inversion.
20. Dispense the mix to 3×13 mL polystyrene tubes with caps.
21. Complete each tube to 13 mL with W5 medium.
22. Centrifuge at $250 \times g$ for 10 min with acceleration and deceleration at 2 at room temperature in swinging buckets.
23. Eliminate the supernatant as much as possible; be careful the pellets are often invisible.
24. Pellets are pooled in a single 13 mL centrifugation tube, filled with W5 solution.
25. Centrifuge at $250 \times g$ for 10 min at room temperature with acceleration and deceleration parameters setup at 2.
26. Eliminate as much as supernatant as possible; be careful pellet is invisible at this step.

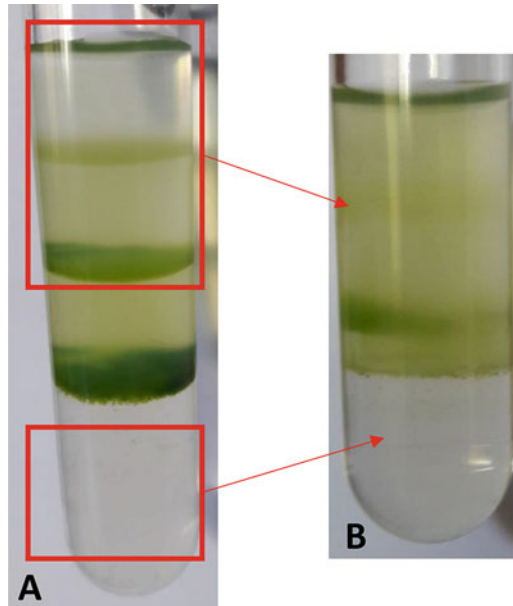


Fig. 4 Process to remove plant cell debris from Percoll centrifugated tubes. After centrifugation (a), purified protoplasts, up and down plant cell debris, were pooled together in a new tube (b) (red squares)

Table 8
MMg solution

Reagent	Final concentration	Qte for 200 mL
Mannitol	400 mM	14.56 g
MgCl ₂ ·6H ₂ O	15 mM	610 mg
CaCl ₂ ·2H ₂ O	5 mM	111 mg
MES buffer	4 mM	160 mg
pH 5.7		
Osmolarity 464 mosmol		

27. If you want to stop here and carry on the transformation of protoplasts the next day, then resuspend the pellet in 1 mL of W5 solution and store the tubes inclined in the dark at 26 °C until the next day, and then start again from **step 25**.
28. If you want to readily perform the transformation, resuspend the pellets in 500 µL MMg solution (Table 8).
29. Evaluate the protoplast viability using the FDA test.
30. Dilute the FDA stock solution by adding 450 µL of acetone to 5 µL of 5 mg/mL FDA stock solution.

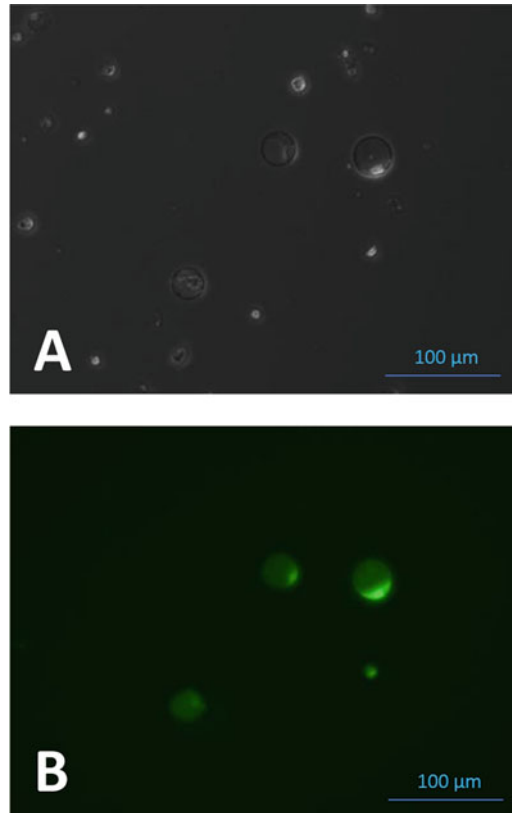


Fig. 5 FDA test of rice protoplasts, view on bright field (a) with exposure time of 0.010 s and on UV plus GFP filter (b). FDA positive and thus living protoplasts were fluorescent on UV plus GFP filter (b) with exposure time of 0.350 s

31. Add 2.5 μL of this solution to 50 μL of protoplast suspension in a 500 μL tube and incubate for 5 min.
32. Readily and rapidly (less than 10 min because fluorescence is rapidly fainting) estimate the number of fluorescent and non-fluorescent protoplasts per mL using a Mallasey's cell or a Rosenthal cell placed under UV light to visualize fluorescein (absorption at 494 nm and emission at 512 nm).
33. All fluorescent protoplasts are alive (Fig. 5). Usually, the viability percentage is close to 100%.
34. The minimum concentration required to perform a transformation is $2.10e^5$ protoplasts/mL and the optimal concentration is $2.10e^6$ protoplasts/mL.
35. Whether the protoplast suspension is too concentrated, adjust the concentration to $2.10e^6$ protoplasts/mL using MMg solution.
36. Let the protoplasts incubate for 1 h before transformation.

Table 9
PEG solution

Reagent	Final concentration	Qte for 10 mL
Mannitol	200 mM	364 mg
CaCl ₂ ·2H ₂ O	100 mM	1 mL stock solution 1 M
PEG4000 Fluka [®]	40%	4 g

Prepare 2 mL aliquot and store at -20°C until use

3.6 Protoplast Transformation (See Note 8)

1. Prepare a number of tubes corresponding to the number of transformations you wish to perform.
2. Add 10 μg of purified plasmid DNA (HiSpeed Plasmid Maxi Kit) in a maximum volume of 20 μL in each 5 mL cytometric tube.
3. Add 200 μL of protoplasts in MMg solution, ideally at a density of 2.10e^6 protoplasts/mL, in each tube.
4. Then add 200 μL of PEG 40% (*see Note 9*) (Table 9) by slowly pouring the PEG along the tube wall while achieving a long rotation movement inside the tube and mix gently by flicking.
5. Let incubate for 15 min in darkness at room temperature.
6. To stop the reaction, slowly add 900 μL of W5 solution stirring gently by inversion.
7. Centrifuge at $250 \times g$ for 10 min with acceleration and deceleration to 2 at room temperature in swinging buckets.
8. Remove the supernatant; be careful not to remove the pellet, which is invisible.
9. Add 500 μL of W5 medium and incubate at 26°C in the dark for 24 h and up to 48 h, tubes left horizontally.
10. Carry out a FDA viability test at 24 h after transformation.
11. Dilute the FDA stock solution by adding to 5 μL of 5 mg/mL stock solution a volume of 450 μL acetone.
12. Add 2.5 μL of this solution to 50 μL of protoplast solution in a 500 μL tube and incubate for 5 min.
13. Readily and rapidly (less than 10 min because fluorescence is rapidly fainting) estimate the number of fluorescent and non-fluorescent protoplasts per mL using a Mallasey's cell or a Rosenthal cell placed under UV light to visualize fluorescein (absorption at 494 nm and emission at 512 nm) (Fig. 6).

3.7 DNA Extraction

1. Add 45 μL of DNA RQ1 buffer and 45 μL of DNase to 450 μL of protoplast solution.
2. Incubate at 26°C for 30 min.

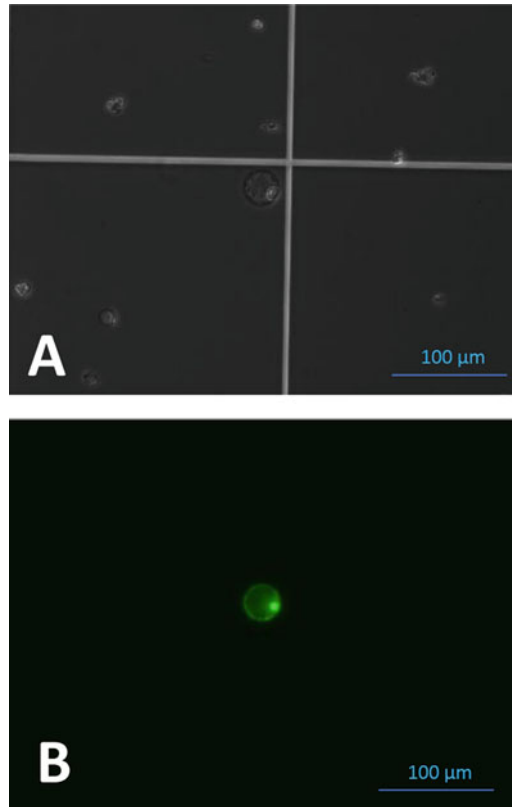


Fig. 6 Transformed protoplast with the pHKN29 GFP plasmid, view on bright field (a) exposure time of 0.2 s and on UV plus GFP filter (b) exposure time of 0.091 s. GFP positive transformed protoplasts were fluorescent on UV plus GFP filter (b)

3. Add 45 μL of STOP solution.
4. Centrifuge at $500 \times g$ for 3 min.
5. Eliminate the supernatant.
6. Add 400 μL of DNeasy plant extraction kit API buffer and incubate at 65°C for 10 min.
7. Extract the DNA by following the conditions of the DNeasy plant extraction kit.
8. Elute DNA in a volume of 100 μL of RNase and DNase-free water.

3.8 Mutation Analysis (See Note 10)

1. A PCR is carried out from the DNA isolated from protoplasts, using Nipponbare gDNA as control, exemplified here using the CAO1-fw and CAO1-rv primers and the following PCR conditions (Table 1).
2. After amplification control on a 1% agarose gel, the PCR products are purified with a PureLink PCR Purification column and dosed.

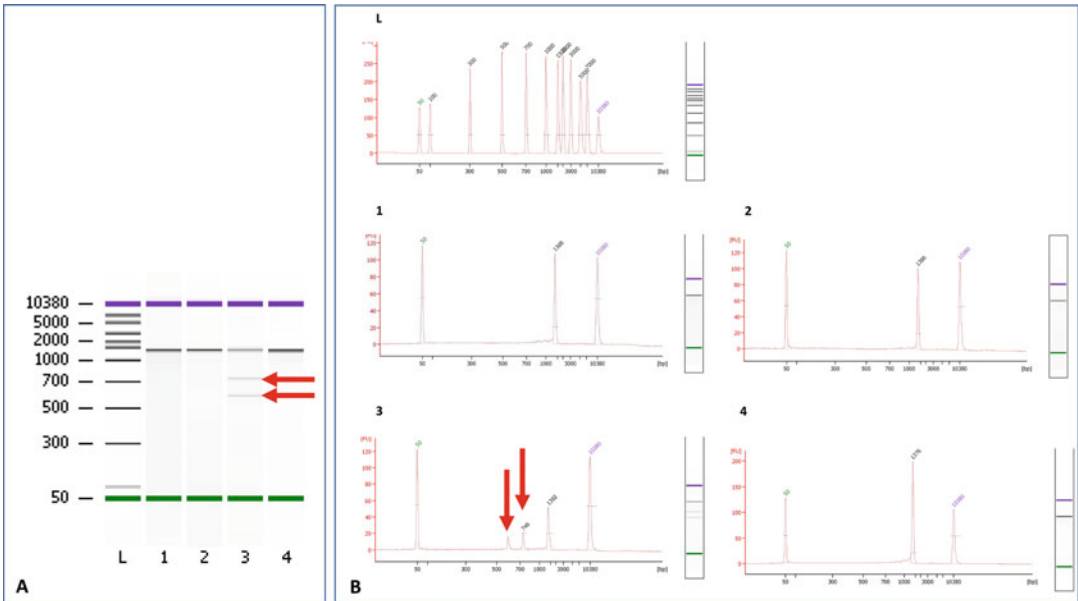


Fig. 7 Bioanalyzer electropherogram and gel of protoplast transformed with the pOs-CAS9 hpt ccd_2.0 BINARY vectors targeting the CAO1 gene. (a) Gel of T7-EI-treated PCR products. (b) Corresponding electropherograms. Red arrows indicate positive digested band from T7-EI reaction. L = ladder, 1 = Nipponbare protoplasts, 2 = pHKN29 GFP plasmid transformed in protoplasts, 3 = pOs-CAS9 hpt ccd_2.0 BINARY vectors targeting the CAO1 gene transformed in protoplasts, 4 = Nipponbare seedling DNA

3. 200 ng of purified PCR product are denatured with a mixture containing 2 μL of $10\times\text{NEB}$ buffer 2 and qsp 19 μL of water at 95 $^{\circ}\text{C}$ 5 min and then hybridized from 95 to 85 $^{\circ}\text{C}$ at 2 $^{\circ}\text{C}/\text{s}$, then from 85 to 25 $^{\circ}\text{C}$ at 0.1 $^{\circ}\text{C}/\text{s}$ in a thermocycler.
4. 1 μL of AT7E-I is added to the reaction and incubated at 37 $^{\circ}\text{C}$ for 15 min.
5. A volume of 1.5 μL of EDTA at 0.25 M is added to stop the reaction.
6. 10 μL are loaded on an agarose gel and 1 μL on the Agilent Nano Chip and then loaded to the Agilent 2100 Bioanalyzer. The presence of two bands at expected size after digestion of the PCR product is a confirmation of the activity of the CAS9 to induce mutations in your target. For instance, *see* the results obtained with a CRISPR/CAS9 construction targeting CAO1 in Fig. 7.
7. The PCR product can also be mass cloned in a PCR vector such as pGEMt following the supplier's recommendations pGEMt Vector System I Promega A3600 (Subheading 2).
8. For instance, 23 independent clones for CAO1 were randomly selected and their inserts sequenced, aligned to get an idea of the diversity of CAS9-induced mutations (Fig. 8).

Cut cones of 200 mL are used to transfer or mix protoplasts. Swing rotors must be used for all centrifugations.

7. If you wish to perform a transformation directly after extraction of protoplasts, thaw the PEG in ice 45 min before purification of protoplasts.
8. Do a control transformation using a construct allowing strong constitutive expression of a reporter gene. For instance, the pHK29-p35:GFP construct can be used as a positive control for protoplast transformation (Fig. 6).
9. When you defroze you PEG aliquot, it should be transparent.
10. The NEB T7-EI kit is used for the detection of CRISPR/CAS9-induced mutations.

Acknowledgments

This work was partly supported by the French government through the program “Investissements d’Avenir” and by the French National Research Agency, in the frame of the research project called “GENIUS” (Genome ENgineering Improvement for Useful plants of a Sustainable agriculture), ANR11-BTBR-0001-GENIUS. Thibault Mounier Ph.D. is partly funded by CRP-Rice and University of Montpellier. Léo Herbert is the recipient of a CIRAD and INRA Ph.D. fellowship. Martine Bes, Leo Herbert, and Thibault Mounier contributed equally to this work.

References

1. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–821
2. Hess GT, Tycko J, Yao D, Bassik MC (2017) Methods and applications of CRISPR-mediated base editing in eukaryotic genomes. *Mol Cell* 68:26–43
3. Yin K, Gao C, Qiu JL (2017) Progress and prospects in plant genome editing. *Nat Plants* 3:17107
4. Li X, Zhou W, Ren Y, Tian X, Lv T, Wang Z, Fang J, Chu C, Yang J, Bu Q (2017a) High-efficiency breeding of early-maturing rice cultivars via CRISPR/Cas9-mediated genome editing. *J Genet Genomics* 44:175–178
5. Mundt DJ, Kelsey JL, Golden AL, Panjabi MM, Pastides H, Berg AT, Sklar J, Hosea T (1993) An epidemiologic study of sports and weight lifting as possible risk factors for herniated lumbar and cervical discs. The Northeast Collaborative Group on Low Back Pain. *Am J Sports Med* 21:854–860
6. Shen L, Wang C, Fu Y, Wang J, Liu Q, Zhang X, Yan C, Qian Q, Wang K (2016) QTL editing confers opposing yield performance in different rice varieties. *J Integr Plant Biol* 60(2):89–93
7. Xu R, Yang Y, Qin R, Li H, Qiu C, Li L, Wei P, Yang J (2016) Rapid improvement of grain weight via highly efficient CRISPR/Cas9-mediated multiplex genome editing in rice. *J Genet Genomics* 43:529–532
8. Rodriguez-Leal D, Lemmon ZH, Man J, Bartlett ME, Lippman ZB (2017) Engineering quantitative trait variation for crop improvement by genome editing. *Cell* 171:470–480. e478
9. Moreno-Mateos MA, Fernandez JP, Rouet R, Vejnar CE, Lane MA, Mis E, Khokha MK, Doudna JA, Giraldez AJ (2017) CRISPR-Cpf1 mediates efficient homology-directed

- repair and temperature-controlled genome editing. *Nat Commun* 8:2024
10. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A et al (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163:759–771
 11. Hu X, Wang C, Fu Y, Liu Q, Jiao X, Wang K (2016) Expanding the range of CRISPR/Cas9 genome editing in rice. *Mol Plant* 9:943–945
 12. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales AP, Li Z, Peterson RT, Yeh JR et al (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523:481–485
 13. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA et al (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154:442–451
 14. Li Z, Zhang D, Xiong X, Yan B, Xie W, Sheen J, Li JF (2017b) A potent Cas9-derived gene activator for plant and mammalian cells. *Nat Plants* 3:930–936
 15. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR (2017) Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551:464–471
 16. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533:420–424
 17. Shimatani Z, Kashojiya S, Takayama M, Terada R, Arazoe T, Ishii H, Teramura H, Yamamoto T, Komatsu H, Miura K et al (2017) Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat Biotechnol* 35:441–443
 18. Zong Y, Wang Y, Li C, Zhang R, Chen K, Ran Y, Qiu JL, Wang D, Gao C (2017) Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nat Biotechnol* 35:438–440
 19. Gil-Humanes J, Wang Y, Liang Z, Shan Q, Ozuna CV, Sanchez-Leon S, Baltés NJ, Starker C, Barro F, Gao C et al (2017) High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *Plant J* 89:1251–1262
 20. Paix A, Folkmann A, Goldman DH, Kulaga H, Grzelak MJ, Rasoloson D, Paidemarry S, Green R, Reed RR, Seydoux G (2017) Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. *Proc Natl Acad Sci U S A* 114: E10745–E10754
 21. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA et al (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823
 22. Rahdar M, McMahon MA, Prakash TP, Swayze EE, Bennett CF, Cleveland DW (2015) Synthetic CRISPR RNA-Cas9-guided genome editing in human cells. *Proc Natl Acad Sci U S A* 112:E7110–E7117
 23. Collonnier C, Epert A, Mara K, Maclot F, Guyon-Debast A, Charlot F, White C, Schaefer DG, Nogue F (2017) CRISPR-Cas9-mediated efficient directed mutagenesis and RAD51-dependent and RAD51-independent gene targeting in the moss *Physcomitrella patens*. *Plant Biotechnol J* 15:122–131
 24. Lin CS, Hsu CT, Yang LH, Lee LY, Fu JY, Cheng QW, Wu FH, Hsiao HC, Zhang Y, Zhang R et al (2017) Application of protoplast technology to CRISPR/Cas9 mutagenesis: from single cell mutation detection to mutant plant regeneration. *Plant Biotechnol J* 16 (7):1295–1310
 25. Malnoy M, Viola R, Jung MH, Koo OJ, Kim S, Kim JS, Velasco R, Nagamangala Kanchiswamy C (2016) DNA-free genetically edited grapevine and apple protoplast using CRISPR/Cas9 ribonucleoproteins. *Front Plant Sci* 7:1904
 26. Miao J, Guo D, Zhang J, Huang Q, Qin G, Zhang X, Wan J, Gu H, Qu LJ (2013) Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res* 23:1233–1236
 27. Kumagai H, Kouchi H (2003) Gene silencing by expression of hairpin RNA in *Lotus japonicus* roots and root nodules. *Mol Plant Microbe Interact* 16:663–668
 28. Haeussler M, Schonig K, Eckert H, Eschstruth A, Mianne J, Renaud JB, Schneider-Maunoury S, Shkumatava A, Teboul L, Kent J et al (2016) Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* 17:148
 29. Bae S, Park J, Kim JS (2014) Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30:1473–1475

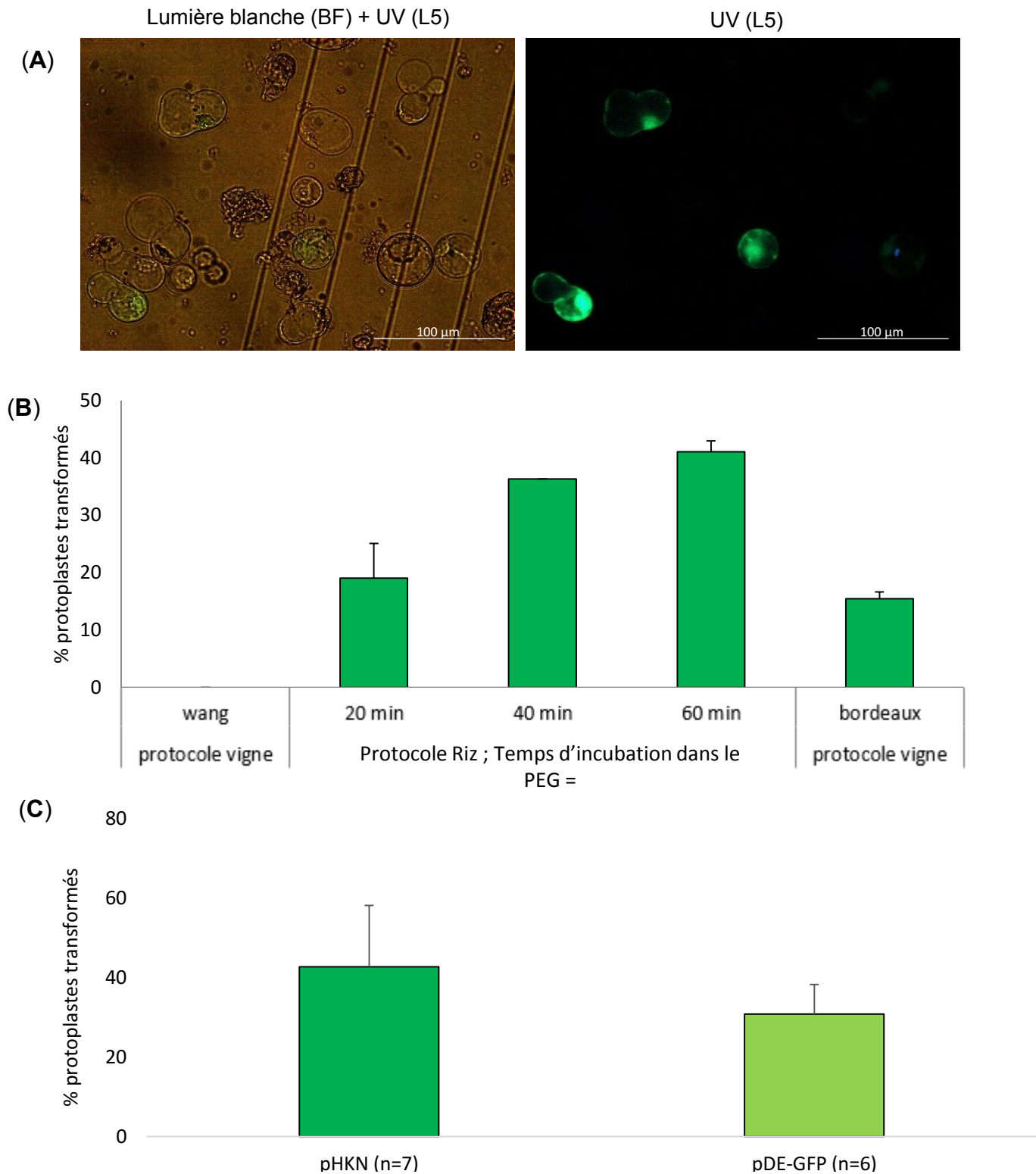


Figure 22 : Test de transformation de protoplastes de cultures cellulaires de Merlot

Photo de protoplastes de Merlot transformés transitoirement par le plasmide pHKN29 (GFP) après 48h au microscope à fluorescence avec un temps d'exposition de 2s en filtre L5 (excitation 494 nm, réception 512nm) avec et sans lumière blanche (A).

Comparaison de différents protocoles de transformation des protoplastes: Plusieurs protocoles ont été testés durant la mise au point de la transformation transitoire : Wang et al. 2015 (wang) ; Bès et al. 2020 (Riz); et Lauvergeat et al. 2020 (bordeaux). Les meilleurs taux ont été obtenus en appliquant le protocole modifié de Bès et al 2020 avec le MMG riz et une incubation de 1h avec le PEG 40% issu du protocole de Lauvergeat et al. 2020 à pH=5,7 (B).

Comparaison du taux de transformation selon la taille du plasmide: En moyenne, sur 6 ou 7 tests indépendants, le plasmide pHKN29 (8650pb) rentre plus facilement que le second contrôle pDE_CAS9_HPT_35S:GFP (environ 18000pb) (C).

Dans ce papier nous avons donc décrit en détail un protocole aisément reproductible de production de protoplastes. Les étapes de purification des protoplastes ont été optimisées au maximum pour éliminer les débris des parois cellulaires ou des protoplastes éclatés au cours de la lyse. L'étape de transformation a été améliorée pour avoir des taux de transformations proches de 90 %. La conservation des protoplastes a été rendue possible au-delà de 48 h pour tester l'expression si nécessaire (nous sommes allés jusqu'à 7 jours pour tester des produits PCR 35S:GFP, données non montrées). Pour finir, nous avons réussi à éliminer l'ADN des protoplastes morts qui interférait avec la détection des ADN des protoplastes encore vivants et édités.

2) Protocoles développés à partir de cultures cellulaires de vigne

Si la technologie a été maîtrisée assez rapidement chez le riz, il a fallu adapter le protocole chez la vigne. Nous avons choisi comme matériel de départ, des cultures cellulaires de vigne, dans la mesure où il était plus difficile et plus long d'obtenir des grosses quantités de feuilles de plants de vigne *in-vitro*. De plus, 2 à 3 vitro-plants après plus d'un mois de culture en tube avec cinq à sept feuilles apicales chacun (0.05 g) permettaient d'obtenir moins de 500000 protoplastes avec l'ajout de 0.01 % de pectolyase Y23, et seulement 12 % d'entre eux étaient transformés avec le plasmide GFP positif pHKN29 décrit plus bas (**Annexe 1-A**).

Le protocole a été adapté d'un protocole utilisé par Mme Virgine Lauvergeat de l'UMR EGFV à Bordeaux (lui-même adapté de Hichri et al., (2010)) à partir d'une culture cellulaire d'une dizaine de jours (50 mL) de Cabernet Sauvignon, ou de Merlot. Il est décrit en détail dans la partie « Matériel et Méthodes ». Brièvement, une digestion sur la nuit avec le même milieu que pour le riz est tout d'abord réalisée. Après digestion, une première séparation de la solution de protoplastes sur « coussin » de solution de lavage W5 est faite, puis trois lavages successifs sont réalisés par centrifugation à 80 g. Le protocole de production de protoplastes est répétable permettant d'obtenir entre 5×10^6 et 3×10^7 protoplastes vivants pour 50 mL de suspension cellulaire fraîche.

Pour évaluer l'efficacité de transformation, nous avons utilisé 10 μ g d'un plasmide fluorescent 35S:GFP (pHKN29) de 8 kb comme témoin (Kumagai and Kouchi, 2003). C'est un plasmide binaire dérivé du plasmide pCAMBIA1300, qui contient un promoteur 35S suivi de la séquence de la GFP (**Annexe 1-A**). Nous avons testé trois protocoles de transformation différents utilisant du PEG ; celui utilisé avec des protoplastes de riz (Bès et al., 2021), le protocole de Bordeaux (Hichri et al., 2010), et le protocole de Wang et al., (2015), tous avec un temps d'incubation de 20 min dans le PEG. Dans les trois cas les taux de transformation sont faibles comparés à ceux du riz (de 80% en moyenne) avec respectivement 19.1 %, 15.5 %, et 0 % (**Figure 22-A et B**), le protocole riz présentant les meilleurs taux de transformation.

Afin d'améliorer ce taux, nous avons testé et validé l'effet d'une augmentation du temps d'incubation du PEG avec les protoplastes et plasmides (40 ou 60 min) comme dans Wang et al., (2015). Le taux de transformation transitoire double en passant à 41.1 % en passant à 1h

d'incubation (**Figure 22-B**).

Le protocole suivi pour toutes les transformations de protoplastes de vigne, suite à ces résultats a été modifié : les protoplastes sont incubés à 4 °C 2h dans une solution de MMG identique à celle du riz (400 mM mannitol, 15 mM MgCl₂ 6H₂O, 4 mM MES, 5mM CaCl₂ 2H₂O, pH=5.7) et la transformation se fait avec du PEG (40% PEG6000, 0.4M mannitol, 0.1M Ca(NO₃)₂, pH=6) à volume égal de protoplastes pendant 1h. Puis 900µL de solution W5 sont ajoutés pour diluer la solution de protoplastes durant 10 min, et les protoplastes sont centrifugés à 80 g sans accélération ni freins. En moyenne, nous obtenons 42.7 % de taux de transformation, avec des valeurs s'échelonnant de 28 % jusqu'à 72 % (**Figure 22-C**).

Pour vérifier si la taille du plasmide impacte le taux de transformation, nous avons construit un vecteur nommé pDE-GFP-NPTII, issu du plasmide pDE_Cas9_AttR, dont la taille est plus proche de celle des plasmides CRISPR qui seront testés en système « protoplaste » (20 kb). Pour ce faire, nous avons sous cloné une séquence 35 S:GFP:Ter qui a été insérée dans le plasmide pDE_Cas9_AttR (**Annexe 1-B**). Vingt microgrammes de ce plasmide sont utilisés par transformation. Les taux de transformation en moyenne atteignent 26.8 % estimés à l'aide de la fluorescence de la GFP (**Figure 22-C**). Cela démontre que la taille des plasmides a un impact sur des taux de transformation chez la vigne, comme déjà observé chez les orchidées (Ren et al., 2020).

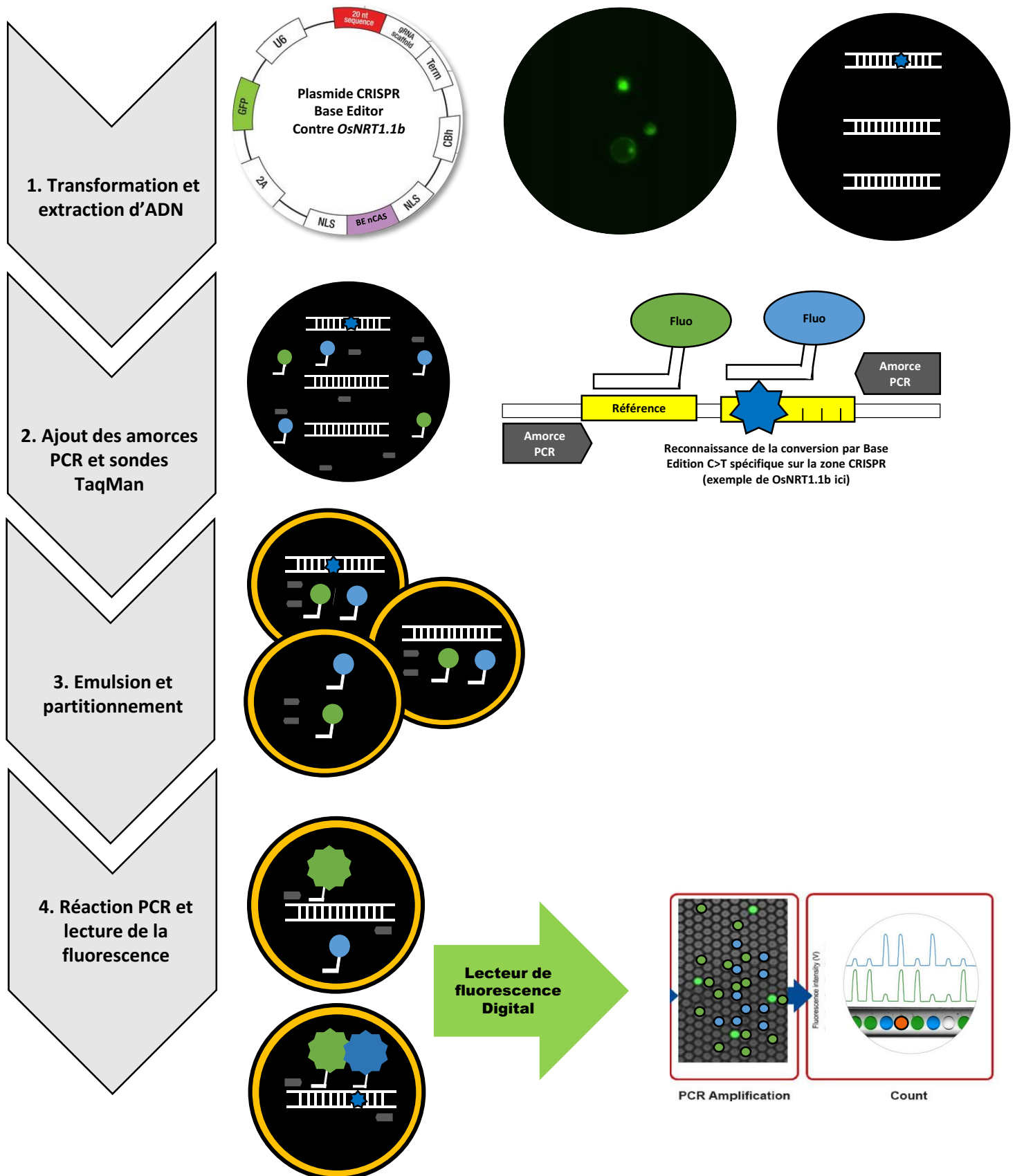


Figure 23 : Principe de la ddPCR sur protoplastes

Brièvement le pipeline de transformation transitoire et détection présente quatre étapes : **1.** la transformation des protoplastes et l'extraction d'ADN 48h après ; **2.** l'ajout dans ce pool d'ADN d'amorces PCR et de sondes TaqMan spécifiques (ici reconnaissance de la base édition sur *OsNRT1.1b* par exemple) ; **3.** l'émulsion qui va partitionner les ADN et le mix réactionnel dans des gouttelettes emprisonnées dans l'huile ; **4.** la réaction de PCR et décrochage des fluorochromes fixés sur la séquence spécifique si l'ADN est présent et/ou muté, suivi d'une lecture de cette fluorescence par gouttelette pour une quantification absolue: couleur orange si libération des deux fluorochromes ou verte ou bleu si un seul.

III) Mise au point de la technologie ddPCR et analyses

La technologie « droplet digital PCR » (ddPCR) permet la quantification absolue des événements de transformation induits par les plasmides CRISPR. Sa mise au point a été effectuée en collaboration avec une startup, IAGE (<https://www.iage-france.com/>).

Brièvement, la stratégie ddPCR est la suivante : les protoplastes transformés et potentiellement édités par des plasmides CRISPR forment une très large population d'évènements (édition, remplacement, insertion) que l'on peut considérer indépendants pour une cible donnée. À l'aide de sondes TaqMan méticuleusement choisies, la ddPCR permet de déterminer et de quantifier le nombre absolu d'évènements pour cette séquence cible, par exemple le nombre de séquences ayant subies une modification d'une seule base (via l'édition à la base près avec le Base Editor) par rapport au nombre de séquences sauvages (WT). Ce système *in-vitro* permet ainsi de tester et comparer l'effet de différents paramètres, cibles, différentes nucléases, etc. sur l'efficacité d'édition quantitativement, du moins théoriquement (Berman et al., 2015; BioRad ; Findlay et al., 2016). Combinée aux protoplastes, cette technologie permet en quelques semaines de vérifier l'efficacité d'une ou plusieurs constructions avant d'utiliser ces mêmes constructions en transformation stable.

La technologie ddPCR consiste à créer une émulsion d'une réaction de PCR, avec l'ADN de nos protoplastes, en 20 000 gouttelettes piégées dans de l'huile. Une réaction de PCR va alors se produire indépendamment dans chaque gouttelette. L'utilisation de deux sondes fluorescentes TaqMan (chromophore FAM et HEX) permet une quantification de la fluorescence terminale : la sonde HEX se fixe lorsque l'ADN est présent, et la sonde FAM si la séquence sauvage ou mutée est présente, selon le design. La lecture de la fluorescence est faite pour chaque gouttelette et permet de déterminer quelle gouttelette est fluorescente ou pas, d'où le terme de digital PCR (1 versus 0) (**Figure 23**).

Comme modèle, nous avons utilisé les ADN issus des protoplastes de riz afin de déterminer s'il est possible de détecter le taux de mutation à la base près sur *OsNRT1.1b*. De plus, nous disposions déjà de plantes avec la mutation C>T qui nous ont servi de contrôle pour la mise au point de cette technologie. Nous avons donc développé différentes stratégies basées sur la reconnaissance en protoplastes de la mutation à la base près, et de l'ADN sauvage, voire de l'ADN avec un indel. Ce développement fait l'objet de l'article « Léo Herbert, Anne-Cécile Meunier, Martine Bes, Dylan Pestel, Aurore Vernet, Murielle Portefaix, Remy Michel, Christian Chaine, Donaldo Meynard, Emmanuel Guiderdoni, Christophe Périn ; **Detection of CRISPR/ CAS9 mutations by Droplet Digital PCR: catch them all!** qui sera soumis pour publication à la revue Plant Method ou Rice. Le draft du papier décrit les stratégies, les calculs, et les résultats obtenus ; nous nous y référerons par moment. Il reste une expérimentation à effectuer avec une nouvelle technologie en ddPCR nommée QIAcuity Elle permettra de détecter en une seule expérimentation de ddPCR les 5 versions alléliques (A, T, G, C, ou indel) au locus *OsNRT1.1b* au niveau du nucléotide substitué, comme avec dans stratégie de DropON (décrite ci-dessous). Une version plus aboutie de l'article est en cours de rédaction.

TITLE PAGE

Title: Detection of CRISPR/ CAS9 mutations by Droplet Digital PCR: catch them all !

Running title: Detection of CRISPR/CAS9 mutations by ddPCR

Authors: Léo Herbert¹, Anne-Cécile Meunier¹, Martine Bes¹, Dylan Pestel¹, Aurore Vernet¹, Murielle¹ Portefaix, Remy Michel¹, Christian Chaine¹, Donaldo Meynard¹, Emmanuel Guiderdoni¹, Christophe Périn¹

Affiliations:

1 CIRAD, UMR AGAP, F-34398 Montpellier, France.

AGAP, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France

Contact: Christophe Périn: perin@cirad.fr

ABSTRACT

Genome editing technologies such as CRISPR/CAS9 or base editors are now mature and indispensable technologies for applied and fundamental research in plant biology. The limiting step for many plant species remains the genetic transformation, costly and time-consuming. As a result, testing GE and BE constructions upstream quickly and eventually comparing their efficiency and specificity is an important step to simplify and rationalize the use of these technologies. Classically, the global mutation rate is estimated by Dropoff using ddPCR, with two probe, a reference one and a probe that attaches only to the WT sequence at the site targeted by the GE construct; this approach, do not cover all GE detection situations. We then first compared Dropoff to alternative techniques, such as competition, Dropon, co-amplification for a BE construct targeting the *OsNRT1.1b* gene in rice. We were able to show that the mutation rates obtained were comparable in protoplasts and of the order of 5%. To illustrate the interest of co-amplification strategy for GE mutant quantification, we used it to compare the deletion rates obtained by LbCPF1 and SpCAS9 for the *OsSAP9* gene. We got with very close rates in protoplasts and plants, with rates of deletions of 1% and 20% for LbCPF1 and SpCAS9 respectively. We have also illustrated how it is possible to estimate specificity (expected mutations) and sensitivity (overall mutation rate) by combining DropOff and DropOn and determine whether a BE construct is both efficient and specific. Future evolution of ddPCR apparatus, will also simplify GE mutation detection by the number of simultaneously detectable fluorophores, such as the QIAcuity (Qiagen) which allows to read 5 fluorophores at the same time. It will thus be possible in the near future to quantify the global mutation rate and simultaneously quantify several specific alleles mutant rates in a single reaction.

BACKGROUND

Recently, genome editing tools and technologies derived from CRISPR/CAS9 have been developed for target mutagenesis in plants (Arora and Narula, 2017) and for studying gene function in many organisms (Bortesi and Fischer, 2015; Herbert et al., 2020). CRISPR/CAS9 induced double strand break repaired by the cell with the non-homologous end joining (NHEJ) or homology directed repair (HDR) generate mutations leading to frame shift and knock-out. Some CRISPR/CAS9 derived technologies, such as the Base Editing (BE), introduced more specific mutations. A first critical step for any genome editing experiment in plant, is the evaluation of efficiency of CRISPR/CAS9 genome editing construct prior stable transformation. Plant genetic transformation efficiency remains low for many plants, and introduction of sub-optimal genome editing construct will lead inevitably to the generation and screening of a large population of transformants, a costly and time-consuming step. Thus, precise and quantitative evaluation of mutation rates is needed for optimization and comparison of genome editing protocols and constructs.

A simple, inexpensive and popular way to test the efficiency of a new CRISPR/CAS9 construct/technology is to transitory transform protoplasts and evaluate mutations induced at the target locus before stable transformation in plants (Bes et al., 2021). The analysis of the mutations remains, however, the limiting step. Among the technologies used, T7E1 tests (Bes et al., 2021), HRM (Thomas et al., 2014) or RE-PCR (Liang et al., 2018) have been widely used to validate and compare CRISPR/CAS9 constructions but quantification of mutation is qualitative or at best semi-quantitative. Moreover, these technologies also require large amount of purified PCR product for mutation identification. Exhaustive identification of CRISPR/CAS9 induced mutations can be done by NGS sequencing of amplicons generated from transfected cells (Bell et al., 2014; Connelly and Pruett-Miller, 2019). Although always theoretically possible, this requires the creation and sequencing of a large number of NGS libraries when the number of comparisons becomes large.

The droplet digital PCR (ddPCR) is a technology commercially released in 2011. ddPCR use a microfluidic apparatus, relying on the partitioning of samples into more than 20 000 droplets by a water-oil emulsion. PCR reaction is performed independently in each of them. Amplicon generation is revealed by fluorescent TaqMan probes (for instance FAM, HEX

for the most common) or DNA intercalant (Evagreen) (Berman et al., 2015; Manoj, 2016; Quan et al., 2018). Instead of following reaction progress in real time PCR, the fluorescence is collected at the end of the reaction. The number of positive and negative droplets are then evaluated. Compared to qPCR, ddPCR provides absolute quantification with no need of standard curves and no reliance on amplification efficiency while sensitivity is also higher (Hindson et al., 2013; Taylor et al., 2017). ddPCR is used to detect the presence of very rare point mutations (Berman et al., 2015), copy number (Collier et al., 2017) or to quantify the expression of weakly expressed genes (Taylor et al., 2017). Advantages of the technology include its sensitivity, robustness and repeatability and the ability to simply multiplex (meaning amplify multiple sequences) in a single sample (Quan et al., 2018).

Conventional ddPCR for mutation detection used a reference probe and multiple other probes targeting the sequence or mutations sought. Hence, ddPCR assays can be used to detect specific mutation (Decraene et al., 2018) such as single base mutation, introduced by Base editor. One of the most used BE, and also the one we used in this study is composed of a nCAS9 with its RuvC site inactive, coupled to a rAPOBEC1 cytidine deaminase (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like enzyme) at the N-Terminus, and to a uracil glycosylase inhibitor (UGI). This technology make it possible to convert a Cytosine into Thymine (Komor et al., 2016). Cytidine deaminase will lead to the excision of a cytidine and its replacement by a uracil (RNA base). Finally, the glycosylase inhibitor will block the excision of uracil by favouring Base Editing Repair (BER) which will replace uracil by thymine. Some other version for C->A, C->G, A>G were recently developed (Gaudelli et al., 2017; Hua et al., 2019).

Droplet digital PCR has therefore naturally been used to detect CRISPR-induced mutations (Berman et al., 2015; Findlay et al., 2016) and seems perfectly well adapted for mono-base specific mutations generated by a Base Editor. Indeed, efficiency of Base editor is not absolute and quantification of unwanted 'on-target' mutations at the target site is also needed to estimate the reliability of a BE construct. If outcome of a perfect mutation is predictable, and detection possible using specifically designed probes, on another side, random mutation induced by NHEJ-based repairs generate a large variability of indel that cannot be captured with a single probe. Detecting and distinguishing expected mutations from non-specific mutations at a given locus is a classical problem when testing the efficiency and specificity of CRISPR constructs such as the BE. A way to quantify *en masse* these NHEJ-

induced mutations is to use the DropOFF strategy (Berman et al., 2015; Decraene et al., 2018; Findlay et al., 2016) also called GEF-dPCR (gene-editing frequency digital PCR) (Mock et al., 2016). A first reference probe (HEX), similar to the one for conventional ddPCR, is used to count all copies of the amplicon. A second probe (FAM), called drop-off probe, bind to WT DNA at the predicted cut site. The HEX probe will not bind, or will “drop off,” when the allele has an insertion or deletion at this site. Although DropOFF is the golden standard for CRISPR/CAS9 overall mutation detection by ddPCR, it cannot cover all situations. It is limited by the size of the amplicon, i.e., 250 bp and is not suitable to detect e.g. large insertions or deletions and the specific mutation rate produced by BE constructs.

We performed a series of ddPCR experiments in protoplasts to evaluate the sensitivity of different constructs (expected edition) and specificity (editions other than expected) using different strategies combining two probes and a conventional ddPCR. We compared our results with mutations rates obtained *in planta* to demonstrate predictability of ddPCR estimations using transfected protoplast. Lastly, we have also demonstrated that it is now possible, in a single tube, to evaluate the sensitivity and specificity of a base edition construct, using a combination of five fluorophores and a state-of-the-art ddPCR machine. This study illustrates, through concrete examples, different methods that can be used to quantify edits caused by genome editing tools. The comparison of conventional and alternative methods offers the reader the possibility of choosing in full knowledge of the facts the test most suited to his case

RESULTS

Gene models: BE of *OsNRT1.1b* and deletion of *OsSAP9* gene

Introduction of the *indica* allele of *OsNRT1.1b* gene by BE (C-T transversion) in a *japonica* background improve significantly the NUE (Nitrogen Use Efficiency) (Chen and Ma, 2015; Hu et al., 2015). Single base modification by BE in the *OsNRT1.1b* gene, coding for a nitrate tranceptor, is correlated with improved nitrate utilization and soil microbiome modification in rice (Hu et al., 2015; Zhang et al., 2019). An BE of *OsNRT1.1b* is thus a good model to test sensitivity and specificity at a single target base by ddPCR and an attempt to introduce this C-T transversion at position 980 how it was made by (Lu and Zhu, 2017).

The *OsSAP9* gene is thought to play a role in salinity, thermal and abiotic tolerance (Kang et al., 2017) but its molecular function remain elusive. In a first step to lead the allelic replacement at this locus and to promote its functional analysis, we sought to delete the coding sequence (CDS) from *OsSAP9* using LbCpf1 and SpCAS9. This is a good model to test the specificity and sensitivity of ddPCR detection for a large modification.

Four different detection systems to test the specificity and sensitivity of the BE construct for NRT1.1B

**In all the cases described below, HEX and FAM fluorophores can be interchanged but it is advisable to measure the rare event with the fluorophore FAM.*

DropOff (**Figure 1A**). A single amplicon and two probes. One HEX probe as reference and one FAM probe (also called DropOff probe) that hybridizes only to the WT sequence at the site of the targeted mutation. The DropOff allows to quantify overall all the mutations induced at the target locus. Hence the formula $\text{HEX}/((\text{HEX and FAM}) + \text{HEX})$ give the total mutation rate, *ie* specific and unspecific mutations and is useful to quantify the overall mutation rate for a CRISPR/CAS9 generating random NHEJ mutations.

DropOn (**Figure 1B**). A single amplicon and two probes. Conceptually, the DropOn is very similar to the DropOff except that the FAM probe is only fixed when the desired mutation, here a C-T transversion, is present. Conventional DropOn or ddPCR allows to quantify the number

of expected/specific mutations among wt and all other mutations (in this case, distinct of C-T transversions). This assay is useful to quantify the number of specific mutations at a locus. Therefore, it is classically used to detect several specific mutations in oncogenes by combining a single reference probe and several specific probes, one *per* mutation. The formula $(FAM \text{ and } HEX) / ((FAM \text{ and } HEX) + HEX)$ gives the specific mutations rate, *ie* expected mutations.

The sensitivity and specificity of the BE can therefore be theoretically deduced by combining the DropOff and DropOn assays ; the sensitivity is computed by global mutation rate (DropOff) and the specificity by the specific mutation rate/total mutation rate (DropOn/DropOff).

A third system, called rare mutation detection (RMD) or competition (**Figure 1C**), is used to estimate the specific mutation rate in relation to the number of non-mutated sequences in a single reaction. A single amplicon is amplified with two probes that compete for the same region. The first probe (HEX) binds to WT sequences and the second one (FAM) binds to the sequence with the desired mutation. This ddPCR assay is classically used in the field of oncology. It's perfectly adapted because it allows to detect rare and specific cancer mutations among a large number of WT sequence in tumor sample. All sequences will then be exhaustively detected. In the genome editing field, the context is different. It should be keep in mind that the total of detected events does not include all the mutations events, but only the expected ones. It does, however, provide an rough estimation of the specificity of a construct. The formula $(FAM + (FAM \text{ and } HEX)) / (HEX + (FAM \text{ and } HEX))$ can be compute to estimate the specific sensitivity rate of a construct.

A fourth system (**Figure 1D**), called co-amplification, consists in co-amplifying two amplicons. Indeed, the size limit for ddPCR amplification is around 250 bp and co-amplification strategy is useful in context where it is not possible to combine two probes in a single amplicon. This system is useful for quantifying the allelic insertion, deletion or replacement rates for large fragments. The first amplicon is used to quantify the total number of DNA copies through the reference probe (HEX). The second amplicon is used to evaluate the mutations sought using a FAM probe. We decided to use this system in a context where we had the possibility to compare it to the reference system (conventional Dropoff and DropOn) and estimate the potential biases. The formula $(FAM + (FAM + HEX)) / (HEX + (FAM + HEX))$ gives the specific mutation total rate.

Four TaqMan probes have been defined for all the experimentations concerning *OsNRT1.1B* (**Figure 2** and **Table 1**). One FAM probe that hybridize to the BE sequence at *OsNRT1.1B-T* (*indica* allele). A HEX probe that hybridize to the WT sequence at *OsNRT1.1B*, and HEX control probe that hybridize *OsNRT1.1B* 83 pb in 5' of the BE targeted region. Finally, a qPCR33 (Actin LOC_Os03g50885) HEX probe, which serves as an external reference for co-amplification experiments. A single primer pair is used for the DropOff, DropOn, and competition experiments (dp-F1 NRT1.1 and dp-R1-NRT1.1) (**Figure 2** and **Table 1**). The gRNA used for BE is shown in blue, the base that is targeted in red and the corresponding PAM is highlighted in black (**Figure 2** and **Table 1**).

Generation of rice plant BE mutants using the BE pnCAS9 construction

We produced C-T mutants using the BE construct in two different *japonica* genetic backgrounds, Kitaake and Chhomrong dhan (CHHD) (**Table 3**). The mutation rates obtained were very similar with, respectively 57% and 56% C-T transversion mutants obtained. Among regenerated plants, 14.3% for CHHD and 18.6% for Kitaake had other mutations than the C-T transversions for CHHD and Kitaake respectively (Supplemental Table 1).

Optimization C-T mutation detection using mutant plant DNA

First, using mutant plant DNA from a *OsNRT1.1b-T CHHD* homozygous plant, we performed a DropOn experiment at increasing temperatures (**Supplemental Figure 1**). Criteria used to determine the optimal temperature were i) the best possible separation between positive and negative droplets for both probes (ii) a C-T mutation detection rate as close as possible of 100%. Based on these two criteria, a temperature of 61°C was chosen for all ddPCR experiments on *OsNRT1.1b*. We performed a series of dilutions to estimate detection efficiency with DropOn. Mutant DNA were diluted from 100% (4 ng) to 0.25% (0.01 ng) with WT DNA (**Figure 3**). As expected, we successfully detect C-T mutations even at the highest dilutions (**Figure 3C**) that correspond to a proportion of 0.25% mutant alleles.

Detection BE and on target mutations

DropOff and DropON: the Golden path

The DropOff probe will then only bind to WT DNA and will no longer bind when a mutation is introduced and the mutations are therefore estimated from a loss of fluorescence from the FAM probe (**Figure 4B**). The non-template control (NTC) was negative, as expected

without any low fluorescent FAM droplet (**Figure 4A**). Mutation rate estimated by DropOff after transformation of the BE construct into protoplasts was around 3.89% (**Figure 4B**).

The DropOn probe will only bind to C-T mutated allele but not on WT allele. As expected, the NTC contains only a single FAM false-positive droplet, probably due to a problem during PCR (**Figure 4C**). Mutation rate is then estimated by the positive FAM droplet number. Mutation rate estimated by DropOn after transformation of the BE construct into protoplast is around 6.21% (**Figure 4D**).

If the specificity of BE is high, DropOff and DropOn should be theoretically identical. If the specificity of BE is low, the DropOff rate should be significantly higher than DropOn. Mutation rate estimations by DropOff and DropOn in our case are close. This is expected, if the specificity of the BE is high. Mutation rate estimated by DropON is nevertheless slightly higher than DropOff. Since the same DNA was used for the DropON and DropOff experiments, the simplest explanation is that the Dropoff and Dropon probes do not have exactly the same sensitivity and are not completely equivalent. To be more stable for the hybridation with the T, the probe NRT1.1-T has 1 more nucleotide (+C) in 3' to respect the standard recommendation. **Figure 4B** shows a higher rainfall for the fluorophore FAM than **Figure 4A**. This rainfall may correspond to a number of mutations that were not captured by the DropOFF probe (false negatives) which lead to a slight underestimation of the actual rate of BE-induced mutations. Indeed, the DropOFF probe is not optimized to detect very specific substitution events in our case.

In plants, mutation rate are 9 times higher (**Table 3**), than in protoplasts. Mutation rate estimated in protoplasts are usually lower than in plants and this underestimation is true in a wide range of plant species (Bortesi and Fischer, 2015). In rice, during genetic transformation by *agrobacterium tumefaciens*, T-DNA insertion that carries the BE takes place very early and probably before or during the first cycles of division. Nuclease act then very early and its expression is constitutive, during regeneration and plant growth (Sallaud et al., 2003). In comparison, the protoplasts used for transient tests do not divide and transitory expression of CRISPR/CAS9 is limited to 48h after transformation, limiting the number of possible mutations produced and detectable. Rate of 'on-target' mutation, *ie* mutations at the BE target regions that correspond to mutations that differ to C-T transversion, was estimated to 14 to 18% of regenerated plants including small deletions and C-G transversions. This rate represent a 10

fold lower value than the expected C-T mutations. These results are consistent with those obtained in protoplasts. The mutation rates detected using the DropON and DropOFF probes were close (3.89% and 6.21%), which already seemed to indicate that the majority of mutations generated by the BE were C-T mutations in protoplasts.

RMD and co-amplification: possible alternatives to estimate mutation rate when binding two probes in a single amplicon is not possible

We also compared mutation rate detected by DropOn, RMD (or competition) and co-amplification (respectively, **Figure 4C-D**, **5A-B** and **5C-D**). After applying the same thresholding for all the experiment, NTC controls have no positive FAM droplets. (**5A**, **5C**). After transformation with the BE construct, mutation rate estimations were very close to each other, 6.24% for RMD, and 4.32% for co-amplification. Biological repeats were similar (ie 5.08% and 5.21% see supplemental **Figure 2**). We confirmed co-amplification, RMD can be alternatively used to detect specific mutation rate.

As mentioned in the description of **Figure 1C**, the case of RMD does not take into account events other than “not mutated” and “mutated as expected (C>T)”. Thus, the calculated rate related to a sub-population and not to the global protoplast population. When the specific mutation rate (C>T) is very high compared to the non-specific mutation rate (other than C>T), which is the case with our BE construction (see **Supplemental Table 1**), then the mutation rates estimated by RMD, co-amplification, DropOn are expected to be quasi-identical. This is what we have obtained here, but if the non-specific mutation rate becomes important, then the rates estimated will diverge (see discussion and **Figure xxxx**).

Quantification of deletion rate by ddPCR using co-amplification

To illustrate the interest of the co-amplification approach, we compared the deletion rates in the OsSAP9 gene (LOC_Os07g07350) obtained using SpCAS9 and LbCPF1 (**Figure 6**). The reference gene is the rice actin gene (qPCR33 or LOC_0s03g50885) and the qPCR33 FAM probe is used to quantify the total amount of DNA. Whenever it was possible, we have designated CAS9 (sgRNA) and CPF1 (crRNA) spacer to target the same region (**Figure 6**). The HEX SAP9-2 probe detect the amplicon only when deletion occurs as the non-deleted fragment is much larger than 250 bp (759bp) and cannot be amplified during ddPCR reaction.

If *OsSAP9* is deleted (513 bp with SpCAS9 and 534 and LbCPF1), then a $\approx 246\text{pb}/225\text{pb}$ fragment is amplified and detected with the HEX SAP9-2 probe.

For comparison and to illustrate the interest and sensitivity of the ddPCR technology compared to a qualitative test, we amplified by PCR the deletion zone and run a 1.2% agarose gel (**Supplemental Figure 3**). Based on agarose gel, we can only conclude that only the SpCAS9 is capable of deleting the *OsSAP9* fragment but not the LbCPF1. For ddPCR experiment, two biological replicates were performed and we got about the same deletion efficiencies. Firstly, in NTC control (**Figure 7A and 7D**) as expected there is no HEX positive droplets. In protoplasts, deletion rates of 17.2% and 19.2% were obtained with SpCAS9 (**Figure 7B and 7E**) and 2.1% and 2.7% with LbCPF1 (**Figure 7C and 7F**). LbCPF1 can effectively delete *OsSAP9* fragment but with a 6 times lower rate than SpCAS9. Its efficacy, although much lower, was quite reproducibly detectable by ddPCR. Moreover, the efficacy deletion rates estimates in plants (**Table 3**) were very close to those estimated in protoplasts. Around 20.8% and 16.8% (two biological replicates) of the regenerated plants had deletions at the expected size with SpCAS9 while none and only one (0 and 1.03%) of the plants transformed by the LbCPF1 construct had deletion in *OsSAP9*.

Catch all mutations in single-tube CRISPR construct using 5 probes and state-of-the-art ddPCR machine.

The QiaCuity technology used multiple Taqman probes in a same well. Probes are described in **Figure 8** and **Table 4**.

DISCUSSION

Comparison of detection methods for GE and BE

Surveyor (T7EI), PCR or PCR-RE are technology relative easy to work with. But their sensitivity is quite limited and can be used only for qualitative comparison (see for instance **Supplemental Figure 3** and **Figure 7**) of GE efficiency. HRM is simple and efficient and rely on qPCR device; melting temperatures are different between individual amplicons and can be used to identify different mutant alleles. HRM is well adapted to discriminate between few alleles (few PCR products) but is more challenging when numerous mutants alleles need to be discriminated, for instance during NHEJ GE induced mutations. Amplicon sequencing (either by subcloning and sequencing PCR product or direct NGS amplicon sequencing) can be used to quantify and determine alleles classes generated by GE. NGS sequencing is laborious and expensive technology and its sensitivity depend on amount DNA and the sequencing depth. DNA concentration obtained following protoplast transformation is very limiting (≈ 5 ng) and when the number of mutations is low, the mutation rate can be confounded with sequencing errors (around 1-5% in NGS sequencing). Detection of GE mutation by ddPCR offer several advantages compared to alternative technologies. ddPCR can deal with very low DNA input and is well adapted to work directly with protoplast native genomic DNA without any step of pre-amplification as it is required for PCR-RE, surveyor or NGS technologies. Moreover, In our results, there is a good correlation between mutation rates obtained in protoplasts GE and plants with often a stronger efficiency of the constructs transformed in rice seed-mature callus, at least for constructs inducing deletion, NHEJ mutation and cytosine deaminase BE constructs (See results). ddPCR can be used to assay, in the same experiment, sensitivity and specificity of several GE constructs and to decide of the best one to transform stably in plant, saving time and money. First, the sensitivity of ddPCR, at least for DropOn, is as high as 0.25% (**Figure 3**). For a construct inducing NHEJ mutations, Dropoff ddPCR is the golden standard to estimate sensitivity and compare several constructs in a single experiment, whatever the number and type of mutations induced.

ddPCR to catch mutations rates of BE constructs in protoplasts ?

As illustrated in **Figure 9**, by combining DropOn and DropOff, or alternative strategy like co-amplification or competition, it is theoretically possible to estimate if a BE construct has a good sensitivity and specificity. When transforming protoplasts with a BE construct, it is possible to subdivide the total protoplast population transformed into three sub-populations (**pie**

charts in Figure 9). The largest sub-population has no mutation (**Figure 9 pie chart piece in blue, WT**), the other two sub-populations contain the expected mutation(s) (**Figure 9 pie chart piece in red, C-T**) and all other possible mutations (**Figure 9 pie chart in green, others**). It is therefore necessary to quantify and distinguish these three sub-populations. The DropOff allows the estimation of the total rate (sensitivity) of mutations (the sum of the two sub-populations),.

In all cases (**Figure 9**), DropOff makes it possible to measure the overall mutation rate. Here set at 5% in all three cases, close to the results we obtained experimentally. However, Dropon allows three different scenarios to be distinguished. In the case where a BE construct generates only random (non-specific) mutations or very few specific mutations (few C-T) (**Figure 9A**), then the rate estimated by DropOn will be close to zero. It will be equal to or close to that obtained by DropOff if the construct is very specific, *i.e.* generates mostly C-T mutations (**Figure 9C**). This is the case of the BE construct used in our work (see **Figures 4 and 5**). Finally, if a BE construct generates about half of the specific mutations, then the rate obtained in Dropoff will be double that obtained in DropOff (**Figure 9B**).

We also demonstrate and illustrate alternative to DropOff/DropOn using co-amplification strategy (**Figure 1D**), with an external reference gene to quantify the number of DNA copies (see results). After demonstration, co-amplification gave similar and close results to DropOff and Dropon (Figure 4 et 5), we illustrate interest of co-amplification strategy to quantify and compare efficiency of deletion of *OsSAP9* gene using LpCPF1 or SpCAS9 (**Figure 7**). Classical DropOff and DropOn should be used whenever possible, but we nevertheless demonstrated it is possible to use alternatives like co-amplification.

If a BE construct, for instance of BE editor without UGI, is supposed to generate several specific alleles, it is possible to combine a DropOff construct and several DropOn probes specific to different mutations (REF) to estimate proportion of each part subpopulations. Alternatively, it should be possible to use new ddPCR technologies like Qiagen that can analyze up to 5 fluorophores in a single reaction (REF) and give in a single reaction conclusion on specificity and sensitivity on a BE construct, combining the best worlds of Dropoff and DropOn constructs.

ddPCR can be used to assay, in the same experiment, sensitivity and specificity of several GE constructs and to decide of the best one to transform stably in plant, saving time and money.

MATERIAL AND METHODS

CRISPR design and primers for cloning sgRNA, crRNA

CRISPR Design

The CRISPOR-TEFOR software (<http://crispor.tefor.net>) was used to define spacers (SpCAS9) for the *OsNRT1.1b* (LOC_Os10g40600) and *OsSAP9* (LOC_Os07g07350) genes. The spacer identified for the *OsNRT1.1* was identical to that published in (Lu and Zhu, 2017). The spacer defined for the *OsNRT1.1b* Editing base was 14 base pairs ahead of the PAM. Two LbCPF1 crRNA were designed manually, in order to cut as close as possible to the SpCas9 cut-off sites. Two spacers and two crRNAs were defined to induce a 513 bp and 534 bp deletion of the *OsSAP9* gene using SpCAS9 and LbCPF1 respectively.

Primers for cloning

The cloning of spacer was done using two oligonucleotides containing two Bsa I restriction site on 5' end. The forward and reverse primers were denatured (100 uM) in 50 uL at 95°C for 5 minutes in a dry bath and hybridized by switching off the dry bath until it reaches room temperature, i.e. for about one hour. See also **Table 2** for the list of spacers and crRNAs used.

Constructions and cloning (cf also (Bes et al., 2021))

Cloning in ENTRY vectors

Cloning in pOs sgRNA ENTRY simplex

The cloning of the spacer targeting *OsNRT1.1* is performed using two oligonucleotides to which Bsa I sites are added (see Table 2). Two ug of the input vector (Miao et al., 2013) are digested by Bsa I (40 U) overnight in a final volume of 50uL. Ring primers are added (6uL) to 100 ng of digested input vector (2.5uL), 2uL of ligase buffer (10x), and T4 ligase (400 IU) in a total volume of 20uL. Ligation is performed at 16°C overnight and 4uL of each ligation product is then transformed into 50uL of chemocompetent DH5 alpha bacteria. The plasmid DNA of about 10 colonies is extracted, checked by restriction with Bsa I. A positive clone, which has lost the Bsa I sites, is used for the next step of cloning sgRNA into the destination vector.

Cloning into the pENTR4sgRNA4 (Zhou et al., 2014) Multiplex

Primers

The cloning of the two spacers targeting the deletion of OsSAP9 using the SpCAS9 is carried out using two pairs of oligonucleotides and *_Btg ZI_* and *_Bsa I_* sites (multiplex) to which cloning sites are added (see Table 2). For dephosphorylating the primers and hybridizing the primer pairs, 4uL of each oligo pair is added to 1.5uL of 10x T4 PNK (Euromedex) buffer, 1uL ATP (20mM), 1uL T4 PNK (10 UI) in a total volume of 15uL and incubated for 1h at 37°C. The forward and reverse primer pairs are then immediately denatured at 95°C for 5 minutes in the dry bath and hybridized by switching off the dry bath until it reaches room temperature, i.e. for about 1 hour.

Cloning the first sgRNA

1ug of pENTRY4 sgRNA4 is then digested using 2uL NEB Btg ZI (10U), 2.5uL 10x NEB Cutsmart Buffer, QSP 25uL H2Odd and incubated at 60°C for 3.5 hours. The linearized vector is then dephosphorylated by adding to the tube 2.7uL Antartik Phosphatase buffer and 1uL Antartik Phosphatase enzyme and incubated for 1h at 37°C. The enzyme is then heat-inactivated at 80°C for 2 minutes. Ligation is then performed by adding 2uL of linearized and dephosphorylated pENTRY4 sgRNA4 vector to 4uL of the first pair of annealed and dephosphorylated oligonucleotides, 2uL of 10x Ligation buffer, 1uL of T4 DNA ligase QSP 20ul H2O dd at 16°C overnight. The next day, 1uL of Btg ZI and 7uL of water are added to the ligation buffer and incubated at 60°C for 1h. Then 2uL of the redigested ligation product is transformed into 25uL of chemocompetent DH5 alpha bacteria. Mining is performed from twelve positive colonies. 5uL of miniprep are digested by Bam HI and then migrated to a 0.8% agarose gel. If the cloning has worked it substitutes a single Bam HI site and there is a band at 3223 bp against 2 bands at 2897 bp and 343 bp if the cloning has failed. Positive clones are verified by sequencing using the BtgZ primer (Table 2).

Cloning the second sgRNA

1 ug of the vector pENTRY4 sgRNA4 containing the first sgRNA is then digested with 2 uL Bsa I NEB (20U), 2.5uL Cutsmart 10x NEB buffer, QSP 25uL H2Odd and incubated at 60°C for 3.5 hours. The linearized vector is then dephosphorylated by adding to the tube 2.5uL Antartik Phosphatase buffer and 1uL Antartik Phosphatase enzyme and incubated for 1h at

37°C. The enzyme is then heat-inactivated at 80°C for 2 minutes. Ligation is then performed by adding 2uL of pENTRY4 sgRNA4 vector containing the first linearized and dephosphorylated sgRNA to 4uL of the second pair of annealed and dephosphorylated oligonucleotides, 2uL of 10x Ligation buffer, 1uL of T4 DNA ligase QSP 20ul H2O dd at 16°C overnight. The next day, 1uL of Bsa I and 7uL of water are added to the ligation buffer and incubated at 60°C for 1h. Then 2uL of the redigested ligation product is transformed into 25uL of chemocompetent DH5 alpha bacteria. Minipreps are made from twelve positive colonies. 5uL of miniprep are digested by Bsa I and then migrated to a 0.8% agarose gel. If the cloning has worked the BsaI site disappears. Negative vectors are linearized by Bsa I. Positive clones, which have lost the Bsa I site, are checked by sequencing using the Bsa I primer (Table 2).

Cloning crRNA for LbCPF1 in the pOs sgRNA ENTRY

The two crRNAs targeting the deletion of OsSAP9 using LbCPF1 were performed by synthesis. A 2206 bp fragment comprising pZmUbi-Hammerhead Ribozymes-crRNA-SAP9 crRNA-SAP9-HDV ribozyme bordered by the Asc I and Not I sites was synthesized by Genscript (**Supplemental Figure 4**).

Cloning by LR in destination binary vectors

The sg and crRNA constructs were transferred by LR recombination from entries vectors into the Binary destination vectors. The sgRNA used for the Base editor was transferred to the modified vector pBE_nCAS9-HPT_AttR-ccdB (pBE) (Zong et al., 2017) (Addgene #98163 plasmid where original cloning sites were replaced by AttR Gateway recombination sites). The sgRNAs for the OsSAP9 deletion were transferred to pZmUbi-OsCas9-HPT (Herbert et al 2020). The crRNAs targeting deletion in OsSAP9 have been transferred to pUbi_LbCPF1-destvect4.0 (Herbert et al 2020).

Briefly, 3uL (100 ng/uL) of each input vector was added to 1.5uL of the corresponding destination vector (at 100ng/uL), 3.5uL of TE buffer (tris-EDTA, pH 8), 2uL LR cloning II. Each tube was vortexed and then centrifuged briefly. After incubation for 3 h at 25°C, the reaction was stopped by the addition of 1uL Proteinase K for 10 min at 37°C. 2.5uL of the reaction was transformed in DH5 alpha chemocompetent bacteria. Selection was made on

LB/spectinomycin 250mg/mL. A dozen minipreps of positive colonies were analyzed per construct. 5uL of each miniprep was added to 0.2uL of Stu I enzyme (NEB 2U), 0.2uL of Bsp HI enzyme ((NEB 2U), 2uL of NEB Cutsmart buffer 10X QSP 20uL buffer then digested at 37°C overnight for selection of recombinant pCAS9-AttR1-ccdb-AttR2; enzymes KpnI and SpeI were used for selection of recombinant pBE-nCAS9-AttR1-ccdb-AttR2 and zzz, www for recombinant pLbCPF1-AttR1-ccdb-AttR2. A positive clone by construction was resequenced for verification with UbiRac primer (Table 2). The plasmids were then used directly for transformation into protoplasts (see below) or transferred by electroporation into *Agrobacterium tumefaciens* EHA105 strain for genetic transformation of mature grain callus. The plasmids were extracted a last time from *Agrobacterium tumefaciens* before genetic transformation and verified by enzymatic restriction.

ddPCR for quantification of protoplast mutations (QX200 Biorad)

TaqMan probes design

Oligonucleotides, probes (probes) are all described and grouped together in Appendix 11. The design of forward and reverse amplification primers and fluorescent probes (FAM, HEX) follows the recommendations of the Biorad supplier (BioRad). The primers are designed with the Primer3Plus software (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) with the thermodynamic and salt parameters set in SantaLucia 1998. The primers are 20 bp with a GC level between 50 and 60%. Their T_m must be between 50 and 65°C, avoiding secondary structures and dimers, without the presence of more than 3 consecutive Gs or 3 Cs, and when possible Gs and Cs in 3'. The size of the amplicon should be between 60 and 200 bp with a maximum size of 250 bp. The FAM or HEX probes should be between 13 and 30 bp and have a GC level of 30 to 80%, with a T_m greater than primers by 3 to 10°C (64°C<65°C optimum<70°C) and should not start with a G. These probes make it possible to distinguish two sequences which differ by a SNP (Single Nucleotide Polymorphism).

ddPCR reaction

The ddPCR reactions were performed in a QSP volume of 21uL distilled water containing: 8uL of transformed protoplast DNA, 11uL of ddPCR Super Mix No UTP (2x Bio Rad Cat #186-4021), 1uL of control primer pairs (900nM) plus FAM probe (250 nM), and 1uL of target primer pairs plus HEX probe (250nM). All probes and primers were designated using primer3 plus (<http://primer3plus.com>) and ordered at Biorad. Samples were then transferred to eight-

channel available droplet-generation cartridges (Biorad Cat#1864008 and #1863009) and a volume of 70uL of Droplet Generation Oil was added (Bio Rad cat #1863005). Each cartridge was loaded into the QX200 droplet generator. After droplet generation, 40uL of each sample was transferred into PCR 96 plates (Bio Rad Cat#12001925) and sealed with a pierceable Foil Heat Seal (Bio Rad Cat#1814040)the PX1 plate sealer. The PCR conditions were 10 minutes at 95°C, 40 cycles of 30 s at 95°C and 1 minute at 57°C-61°C (depending of the probe used see table 1, and a final denaturation step of 10 minutes at 98°C. All ddPCR reactions were analyzed using a QX200 droplet reader and Quantasoft software version 1.6.

Genetic transformation and mutations analysis in planta.

Genetic transformations were carried out in the Kitaake and Chhomrong dhan callus following the protocol using embryogenic secondary callus derived from the scutellum of mature seeds (Sallaud et al 2004). All selections were made on hygromycin. Regenerated plants were analyzed by PCR / SANGER sequencing; primers NRT-F3 and R3 were used for mutation analysis of the OsNRT1.1 target area (see Table 2); primers p53-F1 and R1 were used for mutation analysis of the expected deleted zone inside OsSAP9. All sequences were aligned with the reference wild type and the mutations were analyzed manually.

Production and transformation of protoplasts

Protoplasts were produced from Kitaake's in vitro plants following the protocol of (Bes et al 2020). Briefly, 100 dehusked and disinfected seeds were germinated in the dark on a MS/2 (sigma) medium containing phytagel on a petri dish for 9 days in the dark (26°C). The stems were chopped into 1-2 mm fragments and then the walls were digested with a mixture of Macerozyme R10 and cellulase Rs onozuka. The protoplasts were washed in W5 medium and purified; after centrifugation, protoplasts were diluted to a final concentration of $2 \cdot 10^6$ living protoplasts/mL. Ten ug of DNA were transfected into protoplasts by addition of PEG following the protocol of (Bes et al 2020). A 35:GFP vector was used as a positive transformation control (Kumagai et al 2003). Briefly, 200 uL of protoplasts in the MMG solution were added to 200 uL of PEG 40% and 10ug of DNA (20 uL) and incubated for 15 minutes at room temperature in the dark. The reaction was stopped by adding 900uL of W5; protoplasts were incubated in W5 medium for 48 hours before DNA extraction (DNeasy Plant Kit Qiagen ref 69104) and mutation analysis by ddPCR.

FIGURE LEGENDS

Figure 1: Strategies used to detect Base edition and on targets mutations.

A) DropOff. Single amplicon, two probes. The HEX probe (reference probe) attaches to an area outside the area targeted by the CRISPR. The FAM probe (experimental probe) attaches to the area targeted by the CRISPR in the absence of a mutation. The DropOff measures all the mutations theoretically induced by the Base editor, and all the others, unlike the DropOn. B) DropOn. A single amplicon, two probes. The HEX probe is attached to an area outside the area targeted by the CRISPR. The FAM probe attaches to the target area only if the correct mutation produced by the BE is present. C) Competition. Only one amplicon and two probes. The HEX probe only attaches to the WT (unmutated) locus and the FAM probe only to the mutant locus differing at a single base. Overestimation of “desired mutation” rate, because one part of the global protoplast population is missing (other mutations than C>T transversion). D) Co-amplification. Two amplicons. The first one is used to quantify the amount of DNA with the HEX probe on a constitutive reference gene, and the second one the amount of effective BE mutations. The FAM probe only binds if there is an effective mutation. The arrowheads indicate the position of the primers used. The red star indicates the position of the mutation induced by the Base editor. The TaqMan HEX probes are in green with a circle and the TaqMan FAM probes in blue with a star.

Figure 2: Primers, probes and guide used to target the *OsNRT1.1b* gene.

A) Actin gene used in the co-amplification strategy. B) The NRT1.1b gene, the primers used in DropOff, DropOn, detection probes and the spacer used. See also Table 1 and Table 2. The arrowheads indicate the position of the primers used. The red star indicates the position of the mutation induced by the Base editor. The TaqMan HEX probes are in Green with a circle and the TaqMan FAM probes are in Blue with a star. The position of the PAM in the sequence of *OsNRT1.1b* targeted by the NRT1.1 sgRNA (sgNRT1.1, in blue underline) is highlighted in bold (AGG). The mutation induced by the BE leads to the change of the C (in red) to a T 14 bases pair upstream of the PAM.

Figure 3: Sensitivity and specificity of the DropOn to detect BE-induced mutations.

A) 1D visualization of droplets detected for the FAM probe as a function of the *Oryza sativa* ssp. japonica Chhomrong dhan plant DNA dilution homozygous for the C-T mutation with wild DNA Chhomrong dhan. H₂O: negative control. The ratios “mutant DNA / wild DNA (WT plant)” range from 100% to 0.25%. B) 1D visualisation of droplets detected for the HEX probe as a function of the DNA dilution of plant T234.4.1 homozygous for the C-T mutation with wild DNA. H₂O: negative control. 100%-0.25% level of mutant DNA/wild DNA (WT plant). The presence of a few positive droplets in H₂O is due to slight spontaneous auto-hydrolysis of the HEX probe. C) Theoretical mutation rate (in red) detectable by DropOn vs detected as a function of the proportion of mutant DNA. In red, the mutated nucleotide (T) in T184 1.1b, T234.4.1 and T234.9.1 and in blue, the nucleotide present in the original WT plant (C). The DNA of plant T184 1.1b was diluted with WT DNA to perform the A-C assays.

Figure 4: Comparative efficiency of DropOff A-B) and DropOn C-D) for detecting BE-induced mutations.

A, B, C, D) 2D droplet scatter plots obtained for different protoplast transformation conditions: A) and C), for control experiments, protoplasts transformed with sgRNA-free constructs targeting NRT1.1b. The detected mutant rate is zero. B) and D) for protoplasts transformed with the CRISPR BE construct. The edit rates are 3.89% and 6.21% for DropOff and DropOn respectively.

Figure 5: Comparison of detection efficiency of BE-induced mutations by competition A-B) and co-amplification C-D).

A,B,C,D) 2D droplet scatter plots obtained for different protoplast transformation conditions: A), and C) control experiments. The detected mutant rate is zero. B), and D) with the CRISPR BE construct. The respective edition rates are 6.24% and 4.32% by competition and co-amplification. At the top of each column, a reminder of the designs for competition (RMD) and co-amplification.

Figure 6: Primers, probes and guides used to delete the OsSAP9 gene and its detection by co-amplification in rice.

The OsSAP9 gene, the primers used in co-amplification, the detection probes and the spacers used using SpCAS9 (sgRNA) or LbCPF1 (crRNA). See also Table 1 and Table 2. The positions of the spacers were chosen at approximately the same locus, modulating the PAM constraints.

The sgRNAs used are in blue and the crRNAs in red. The deletion of OsSAP9 by CAS9 is theoretically 513 bp and 534 bp by LbCpf1. The arrowheads indicate the position of the primers used. The red star indicates the position of the double strand break induced by the nuclease. The TaqMan HEX probes are in Green with a circle and the TaqMan FAM probes in blue with a star. The position of the PAMs in the target sequences are underlined and annotated (PAM).

Figure 7: Comparative efficiencies of SpCAS9 and LbCPF1 in deleting the OsSAP9 CDS measured by co-amplification.

Two biological replicates are used (A-B-C and D-E-F).

A,B,C,D,E and F) 2D droplet scatter plots obtained for different protoplast transformation conditions: A) and D) are control experiments, protoplasts transformed with sgRNA- and crRNA-free constructs. The mutant rate detected is nul. B) and E) result obtained for protoplasts transformed with the CRISPR/CAS9 construct. Estimated editing rates are 17.2% and 19.2%. C) and F) result obtained for protoplasts transformed with the LbCPF1 construct. Estimated editing rates are 2.1% and 2.7%. At the top of each column, a reminder of the design.

Figure 8: DroppALL on *OsNRT1.1b* Strategy with QIAcuity technology

The QIAcuity technology will be use with multiple Taqman probes in the same well. Five probes are chosen for this experiment related to DropOn or DropOff like Figure 1. One for the total number of amplicons (probe REFERENCE in FAM) ; 1 for the wild nucleotide (C) with Tamra fluorochrome ; 3 for single SNP mutations, including the expected C > T substitution (Hex-Vic for T ; Rox for A ; Crimson for G).

Figure 9: Theoretical rates of mutations estimated by Dropoff, Dropon, competition (RMD) or co-amplification on three different situations. for a specific rate of mutation equal to zero (A), 2.5% (B) or 5% (C). On left, pie charts with proportions of WT (blue), specific (C-T, red) and other mutations (green). Rate of mutations estimated in percent for each situation, by DropOff DropOn, Competition or co-amplification. For Competition, the percentage is computed by the ratio of C-T mutation on WT, so 2.6% (2.5%/95% WT) for B).

Supplementary Figure 1: Optimization of annealing temperatures for ddPCR for the detection by DropOn of the mutation introduced by BE for the NRT1.1b gene

A) 1D visualisation of droplets detected for the FAM probe (muted C>T DNA) as a function of annealing temperature (55, 57, 59, 61°C) on NRT1.1b gene. B) 1D visualisation of droplets

detected for the HEX probe (Reference) as a function of annealing temperature (55, 57, 59, 61°C) on NRT1.1b gene.

Supplemental Figure 2: Biological replication comparing detection efficiency of BE-induced mutations by dropON (A-B), competition C-D),co-amplification (E-F), and dropOff (G-H).

A), C), E) and G) are the control experiments in protoplasts. The detected mutant rate is zero. B), D), F), and H) with the CRISPR BE construct. The respective edition rates are 5.39%, 5.08%, 5.21%, and 3.15% by DropON, competition, co-amplification, and DropOff. At the top of each column, a reminder of the designs for drop on, competition and co-amplification.

Supplementary Figure 3: Detection of CRISPR/CAS9 and LbCPF1 induced deletions of the OsSAP9 gene in rice protoplasts by classical PCR.

The test is qualitative and only allows the conclusion that CAS9 induces detectable deletions and not LbCPF1.

Supplementary Figure 4: Schematic representation of the synthesized cassette for the expression of LpCPF1 crRNA.

crRNAs were under the control of a pZmUBI promoter. The LbCPF1 binary plasmid used is the same as described in Leo et al 2020. HDV: HDV ribozyme; HH: Hammerhead ribozyme

Supplementary Figure 5: Table of Kitaaké and Chhomrong dhan plants regenerated

TABLES

Table 1: List and sequences of primers and probes used for mutation detection by ddPCR

Table 2: List and sequences of primers used for cloning, sequencing and PCR mutation detection

Table 3: Mutation rates identified in rice for the base-edition OsNRT1.1 gene and the deletion of the OsSAP9 gene obtained by sequencing at target loci.

Table 4: List and sequences of primers and probes used for mutation detection by ddPCR QIAcuity

DECLARATIONS

- Ethical approval and consent to participate: Not applicable
- Consent for publication : Not applicable
- Availability of data and material: All the datasets, including videos and photographs, are included in the article and are also available from the corresponding author upon reasonable request. All the transgenic lines are also available through MTA from the corresponding author.
- Competing interests: The authors declare that they have no competing interests.
- Funding: Not applicable
- Authors' contributions: LH and ACM built the Base editor, CPF1 and CAS9 constructions. FD, ACM and LH analyzed the mutations generated. Genetic transformations and monitoring of greenhouse plants were done by AV, MP, RM and CC, DM. LH, ACM, CP and EG planned the experiments. CP wrote a first version of the paper. All authors reviewed and participated in the corrections and final version of the manuscript.
- Acknowledgments: This work was supported by several grants, including the Investissement d'Avenir program of the French National Agency of Research for the Project GENIUS (ANR-11-BTBR-0001_GENIUS), the CGIAR research program on RICE (CRP RICE), the Plant2pro project by Institut Carnot, and the Labex AGRO ANR 10-LABX-0001 project GENERICE. Léo Herbert is partially supported by a Montpellier SupAgro PhD fellowship. We would like to thank the company IAGE for its advice and the exchange of information on the digital droplet PCR technologies.

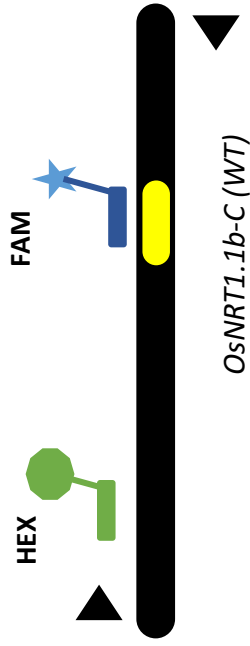
REFERENCES

- Arora, L., and Narula, A. (2017). Gene Editing and Crop Improvement Using CRISPR-Cas9 System. *Front Plant Sci* 8:1932.
- Bell, C.C., Magor, G.W., Gillinder, K.R., and Perkins, A.C. (2014). A high-throughput screening strategy for detecting CRISPR-Cas9 induced mutations using next-generation sequencing. *BMC Genomics* 15:1002.
- Berman, J.R., Cooper, S., Zhang, B., Karlin-Neumann, G., Litterst, C., and Jouvenot, Y. (2015). Ultra-sensitive quantification of genome editing events using droplet digital PCR. *BioRad Laboratories* 6712.
- Bes, M., Herbert, L., Mounier, T., Meunier, A.C., Durandet, F., Guiderdoni, E., and Périn, C. (2020). Efficient genome editing in rice protoplasts using CRISPR/CAS9 construct. *Methods in Molecular Biology*.
- Bortesi, L., and Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol Adv* 33:41-52.
- Collier, R., Dasgupta, K., Xing, Y.P., Hernandez, B.T., Shao, M., Rohozinski, D., Kovak, E., Lin, J., de Oliveira, M.L.P., Stover, E., et al. (2017). Accurate measurement of transgene copy number in crop plants using droplet digital PCR. *Plant J* 90:1014-1025.
- Connelly, J.P., and Pruett-Miller, S.M. (2019). CRIS.py: A Versatile and High-throughput Analysis Program for CRISPR-based Genome Editing. *Sci Rep* 9:4194.
- Decraene, C., Silveira, A.B., Bidard, F.C., Vallee, A., Michel, M., Melaabi, S., Vincent-Salomon, A., Saliou, A., Houy, A., Milder, M., et al. (2018). Multiple Hotspot Mutations Scanning by Single Droplet Digital PCR. *Clin Chem* 64:317-328.
- Findlay, S.D., Vincent, K.M., Berman, J.R., and Postovit, L.M. (2016). A Digital PCR-Based Method for Efficient and Highly Specific Screening of Genome Edited Cells. *PLoS One* 11:e0153901.
- Herbert, L., Meunier, A.C., Bes, M., Vernet, A., Portefaix, M., Durandet, F., Michel, R., Chaîne, C., This, P., Guiderdoni, E., et al. (2020). Beyond Seek and Destroy: how to Generate Allelic Series Using Genome Editing Tools. *Rice (N Y)* 13:5.
- Hindson, C.M., Chevillet, J.R., Briggs, H.A., Gallichotte, E.N., Ruf, I.K., Hindson, B.J., Vessella, R.L., and Tewari, M. (2013). Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods* 10:1003-1005.
- Hu, B., Wang, W., Ou, S., Tang, J., Li, H., Che, R., Zhang, Z., Chai, X., Wang, H., Wang, Y., et al. (2015). Variation in NRT1.1B contributes to nitrate-use divergence between rice subspecies. *Nat Genet* 47:834-838.
- Kang, M., Lee, S., Abdelmageed, H., Reichert, A., Lee, H.K., Fokar, M., Mysore, K.S., and Allen, R.D. (2017). Arabidopsis stress associated protein 9 mediates biotic and abiotic stress responsive ABA signaling via the proteasome pathway. *Plant Cell Environ* 40:702-716.
- Lu, Y., and Zhu, J.K. (2017). Precise Editing of a Target Base in the Rice Genome Using a Modified CRISPR/Cas9 System. *Mol Plant* 10:523-525.
- Manoj, P. (2016). Droplet digital PCR technology promises new applications and research areas. *Mitochondrial DNA A DNA Mapp Seq Anal* 27:742-746.
- Mock, U., Hauber, I., and Fehse, B. (2016). Digital PCR to assess gene-editing frequencies (GEF-dPCR) mediated by designer nucleases. *Nat Protoc* 11:598-615.

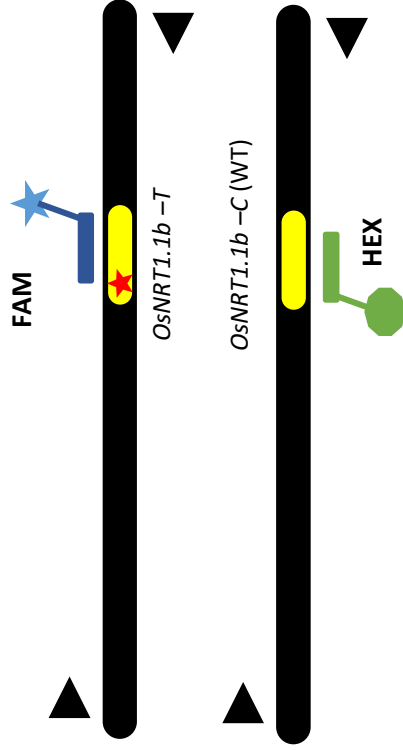
- Quan, P.L., Sauzade, M., and Brouzes, E. (2018). dPCR: A Technology Review. *Sensors (Basel)* 18.
- Taylor, S.C., Laperriere, G., and Germain, H. (2017). Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data. *Sci Rep* 7:2409.
- Zhang, J., Liu, Y.X., Zhang, N., Hu, B., Jin, T., Xu, H., Qin, Y., Yan, P., Zhang, X., Guo, X., et al. (2019). NRT1.1B is associated with root microbiota composition and nitrogen use in field-grown rice. *Nat Biotechnol* 37:676-684.
- Zhou, H., Liu, B., Weeks, D.P., Spalding, M.H., and Yang, B. (2014). Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res* 42:10903-10914.

BioRad Droplet Digital™ PCR Applications Guide.

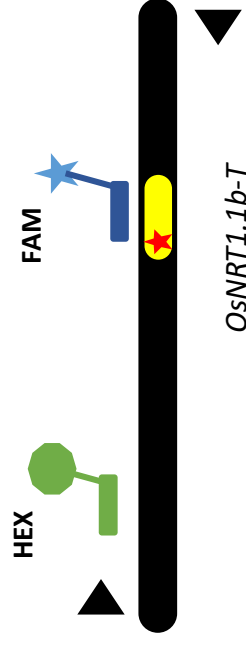
- Chen, Z.C., and Ma, J.F. (2015). Improving Nitrogen Use Efficiency in Rice through Enhancing Root Nitrate Uptake Mediated by a Nitrate Transporter, NRT1.1B. *Journal of Genetics and Genomics* 42, 463–465.
- Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 551, 464–471.
- Hu, B., Wang, W., Ou, S., Tang, J., Li, H., Che, R., Zhang, Z., Chai, X., Wang, H., Wang, Y., et al. (2015). Variation in NRT1.1B contributes to nitrate-use divergence between rice subspecies. *Nature Genetics* 47, 834–838.
- Hua, K., Tao, X., Liang, W., Zhang, Z., Gou, R., and Zhu, J. (2019). Simplified adenine base editors improve adenine base editing efficiency in rice. *Plant Biotechnology Journal*.
- Liang, Z., Chen, K., Zhang, Y., Liu, J., Yin, K., Qiu, J.-L., and Gao, C. (2018). Genome editing of bread wheat using biolistic delivery of CRISPR/Cas9 in vitro transcripts or ribonucleoproteins. *Nat Protoc* 13, 413–430.
- Lu, Y., and Zhu, J.-K. (2017). Precise Editing of a Target Base in the Rice Genome Using a Modified CRISPR/Cas9 System. *Molecular Plant* 10, 523–525.
- Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., Zhang, X., Wan, J., Gu, H., and Qu, L.-J. (2013). Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Research* 23, 1233–1236.
- Sallaud, C., Meynard, D., van Boxtel, J., Gay, C., Bès, M., Brizard, J.P., Larmande, P., Ortega, D., Raynal, M., Portefaix, M., et al. (2003). Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. *Theor Appl Genet* 106, 1396–1408.
- Thomas, H.R., Percival, S.M., Yoder, B.K., and Parant, J.M. (2014). High-Throughput Genome Editing and Phenotyping Facilitated by High Resolution Melting Curve Analysis. *PLoS ONE* 9, e114632.



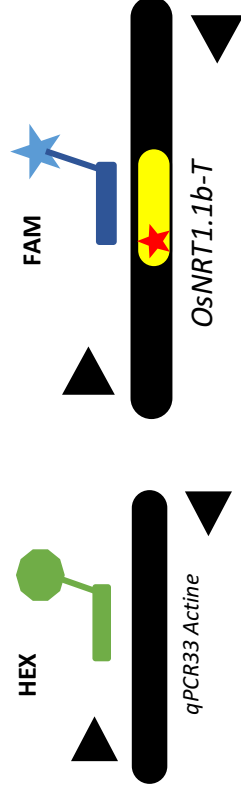
A)
$$\% \text{Mutants} = \frac{\text{HEX}}{[\text{FAM and HEX}] + \text{HEX}}$$



C)
$$\% \text{Mutants} = \frac{\text{FAM} + (\text{FAM and HEX})}{(\text{FAM and HEX}) + \text{HEX}}$$



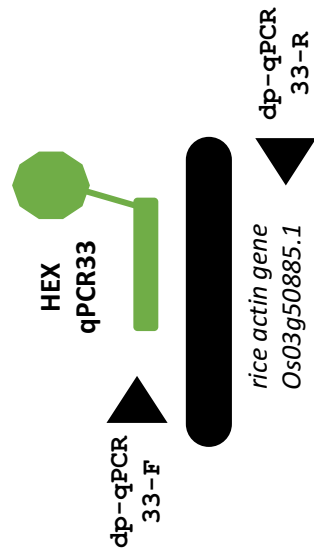
B)
$$\% \text{Mutants} = \frac{[\text{FAM and HEX}]}{[\text{FAM and HEX}] + \text{HEX}}$$



D)
$$\% \text{Mutants} = \frac{\text{FAM} + (\text{FAM and HEX})}{\text{HEX} + (\text{FAM and HEX})}$$

Figure 1

A)



B)

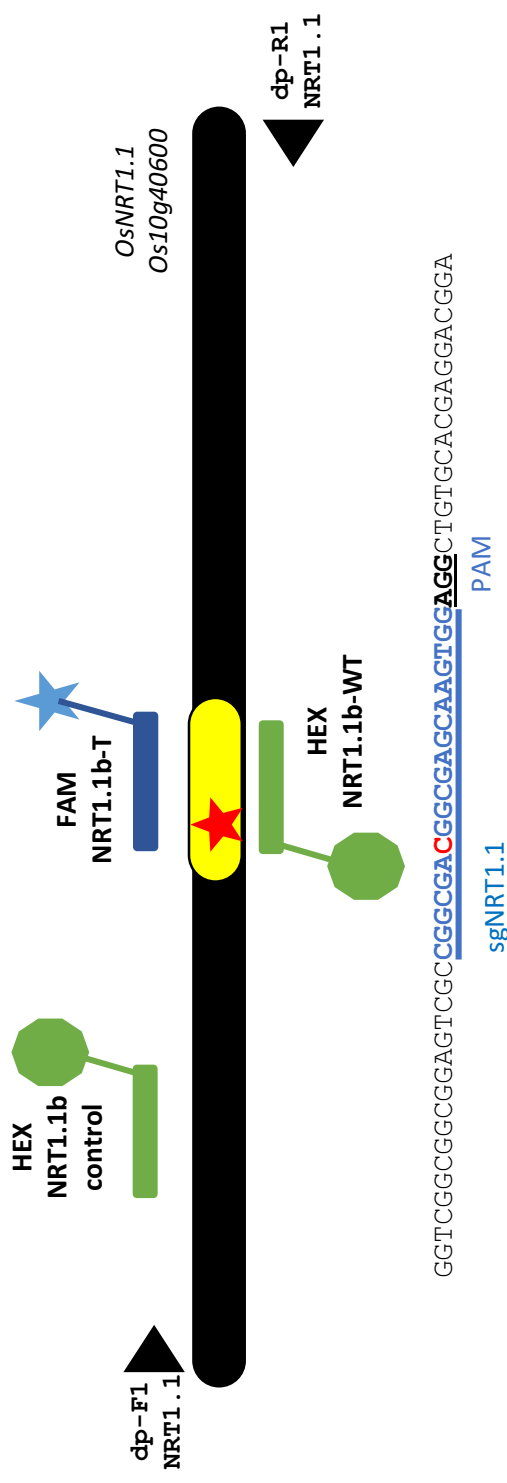


Figure 2

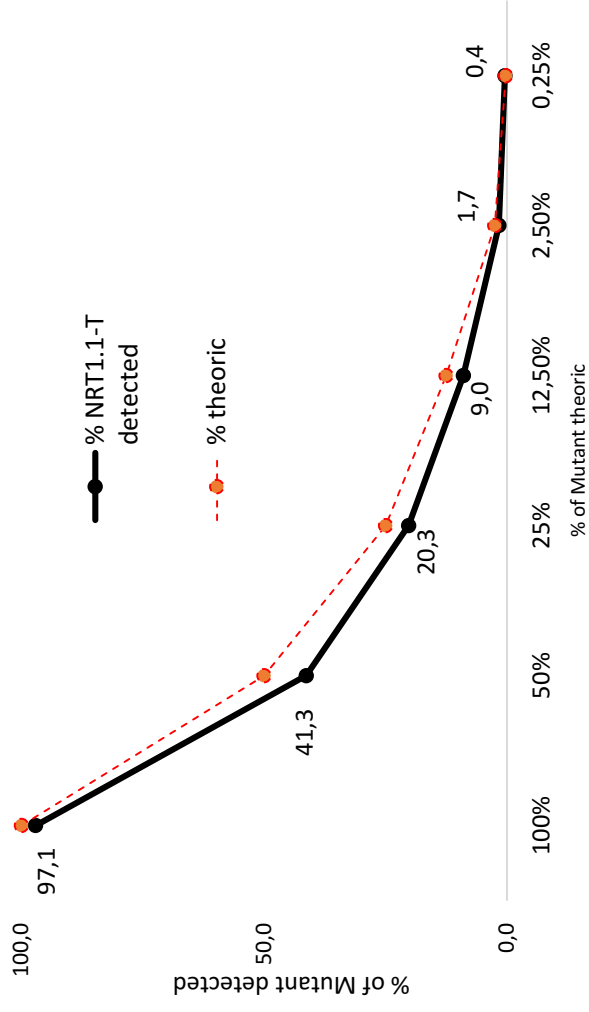
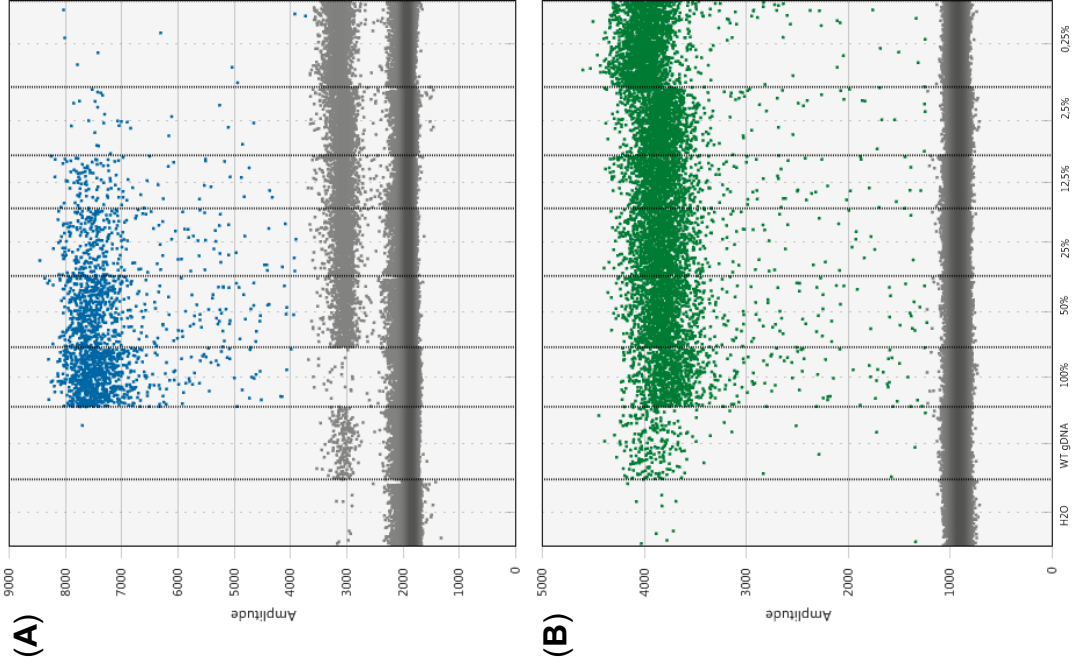
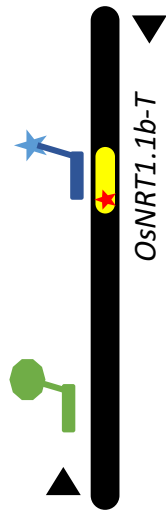
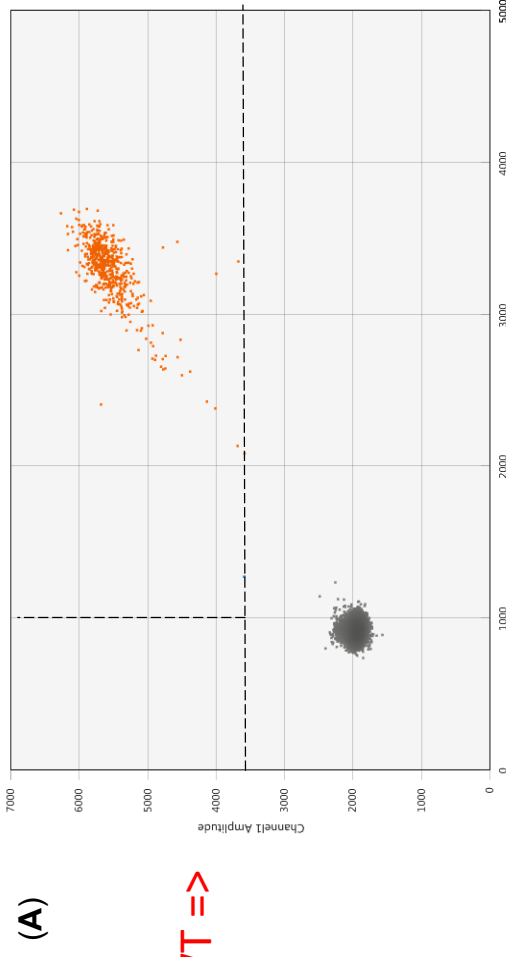
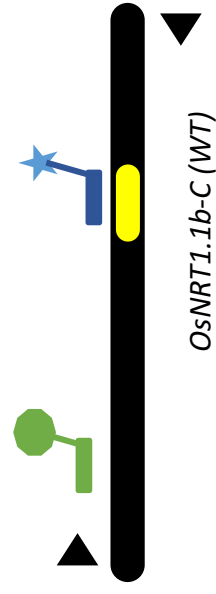
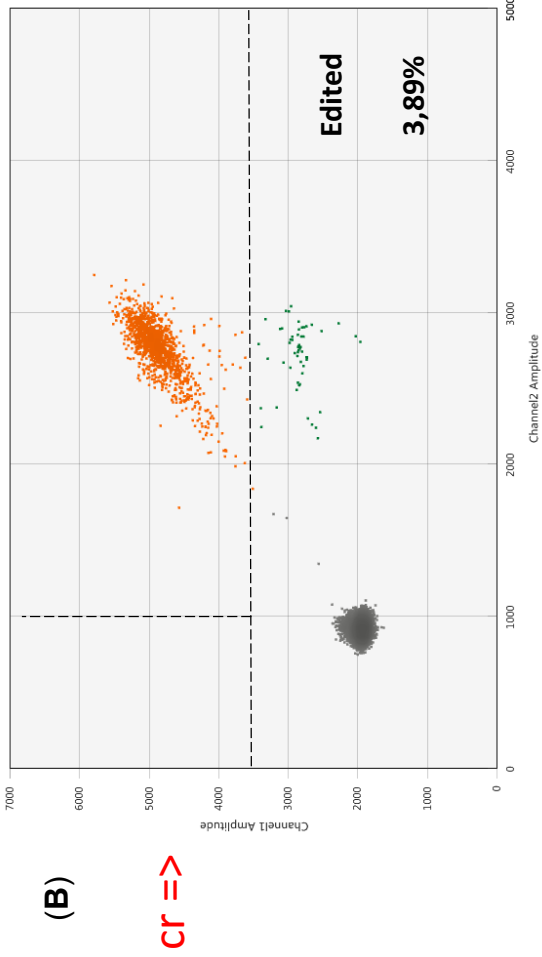


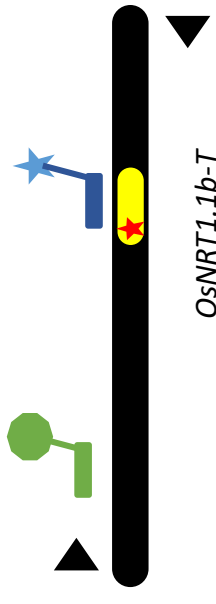
Figure 3



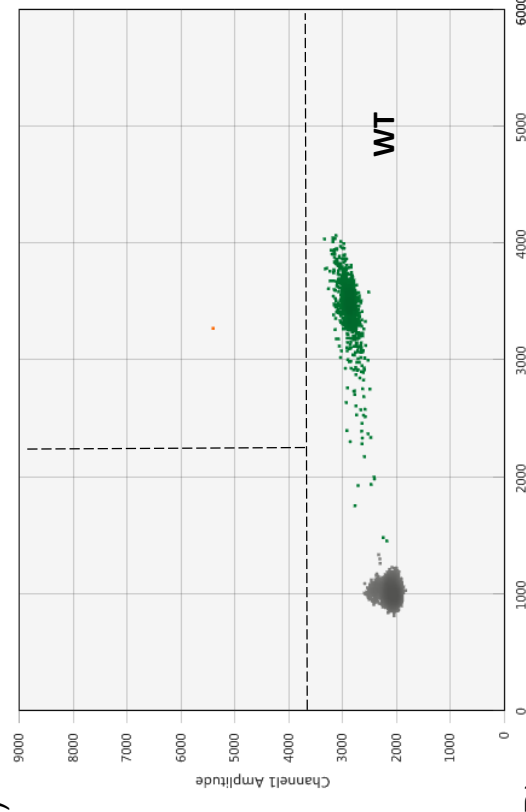
WT =>



+ Cr =>



(C)



(D)

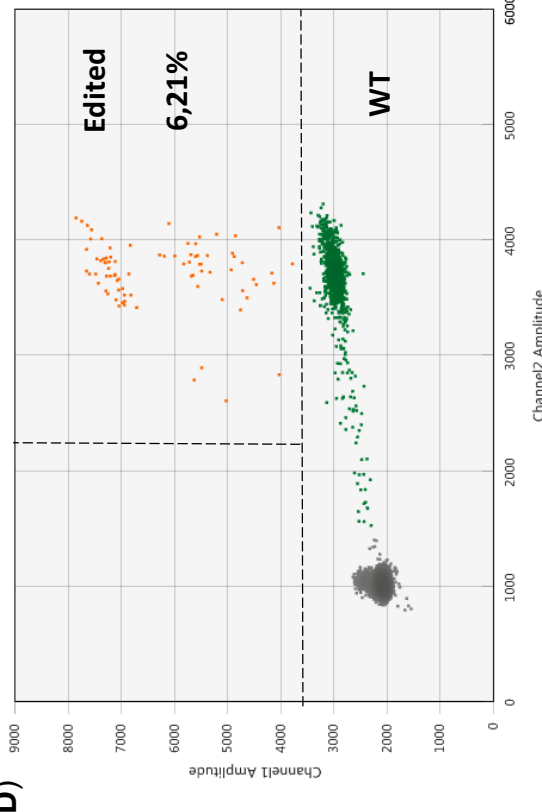


Figure 4

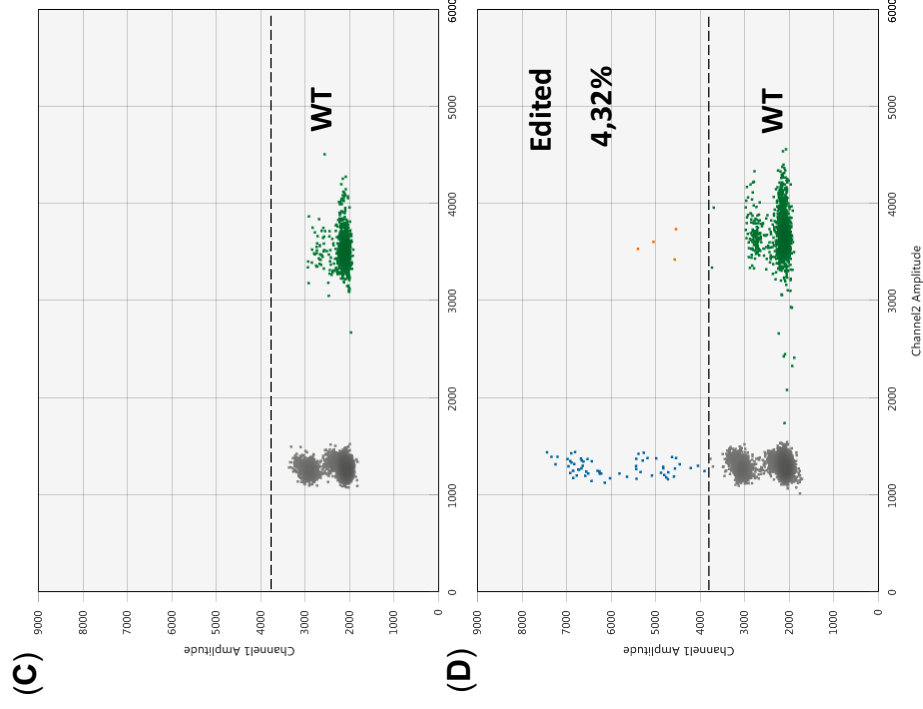
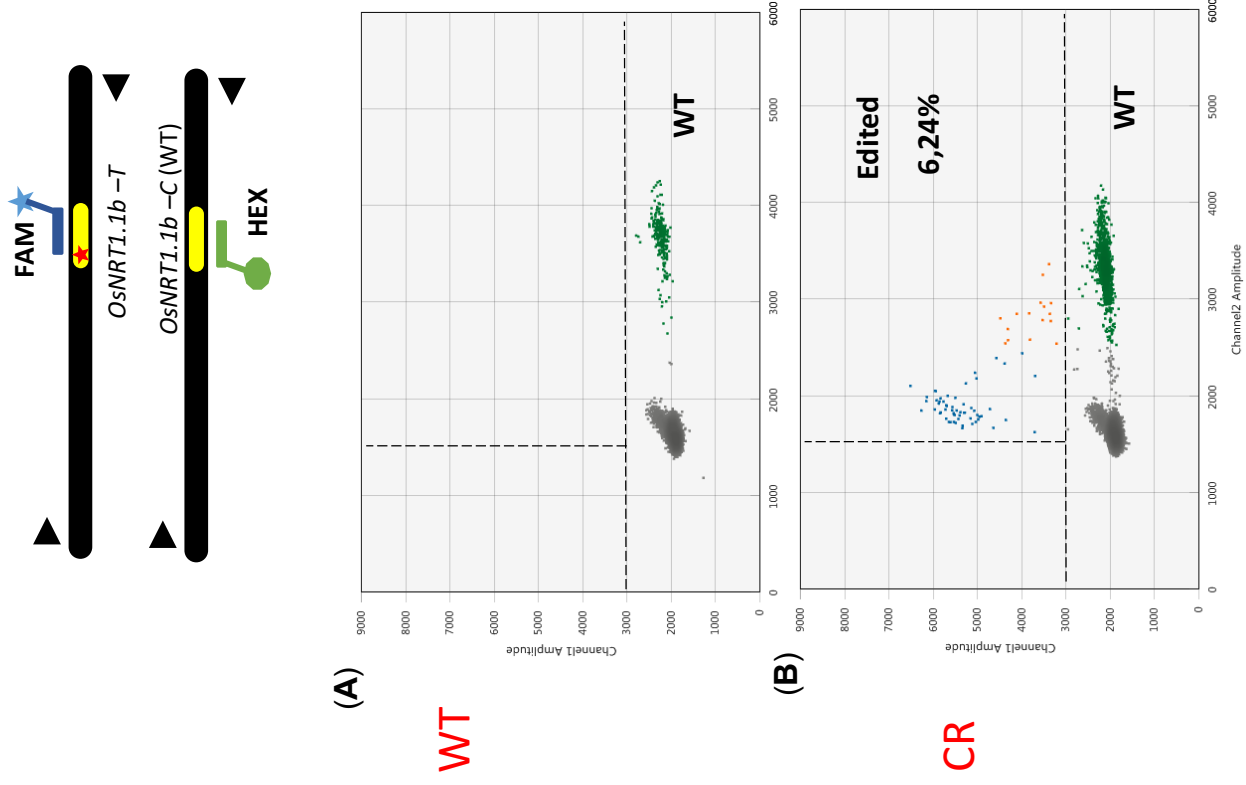


Figure 5

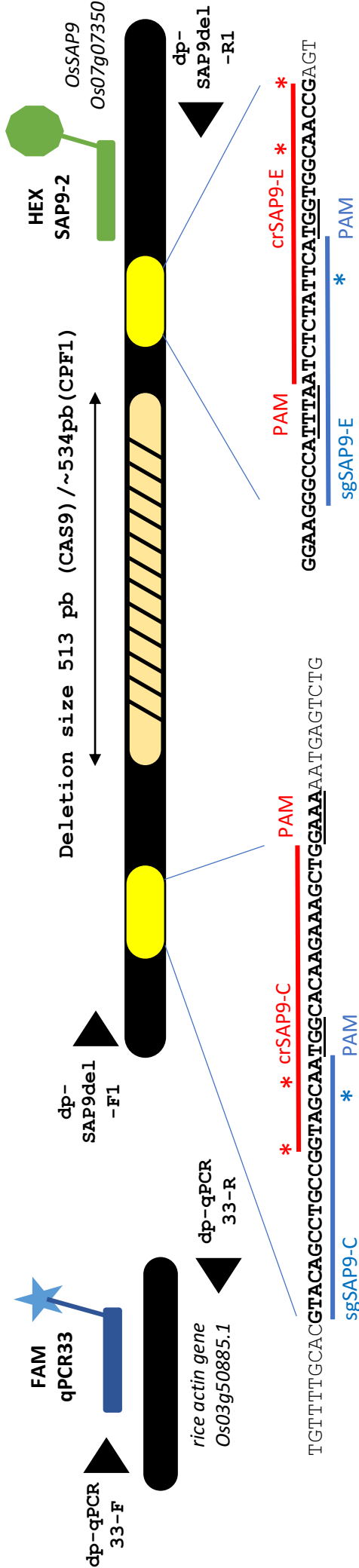
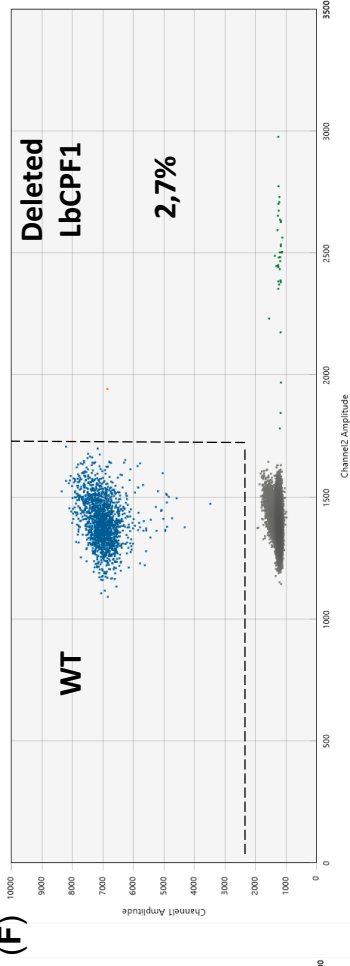
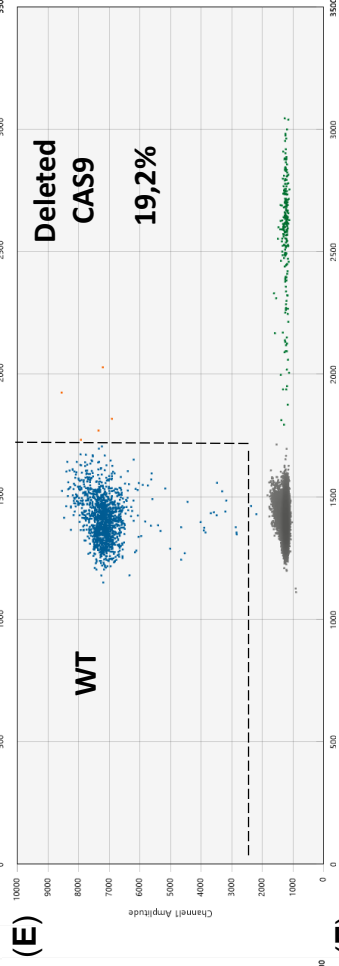
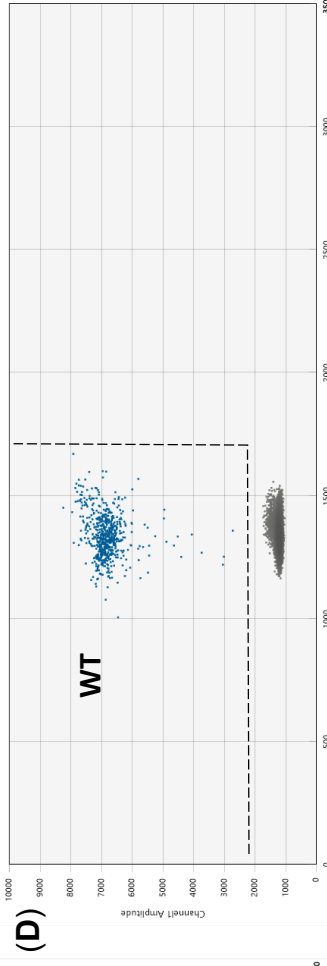
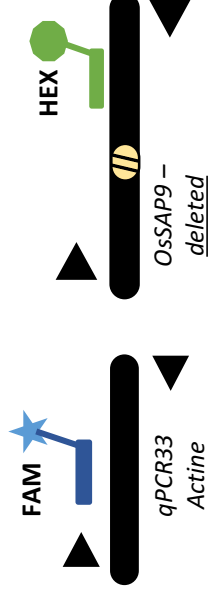
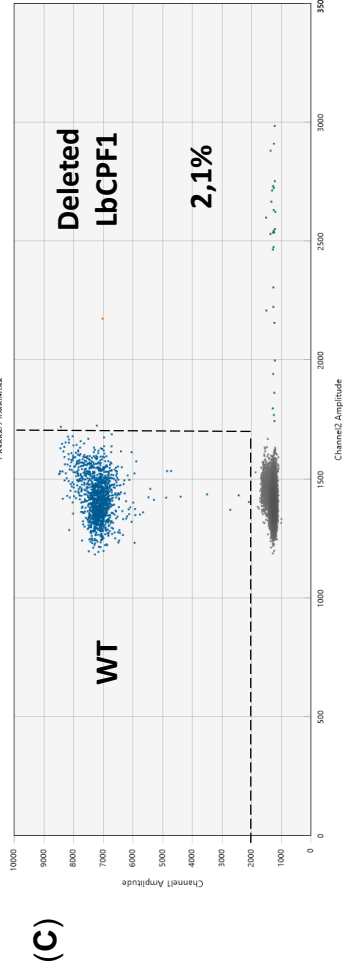
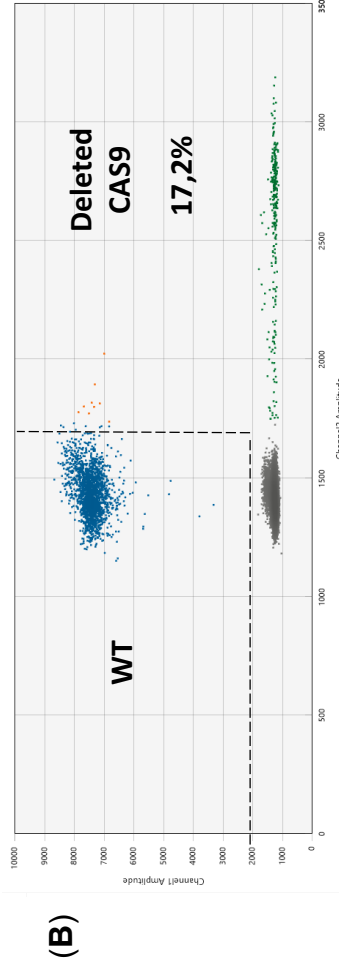
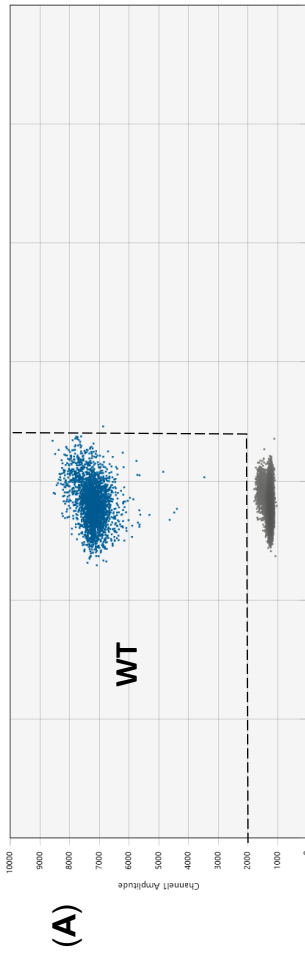
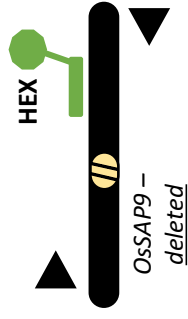
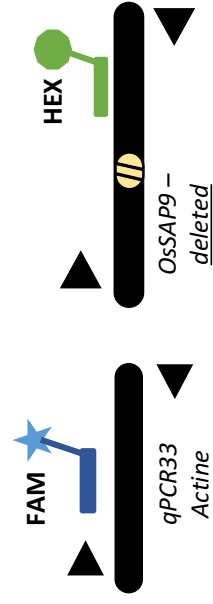


Figure 6



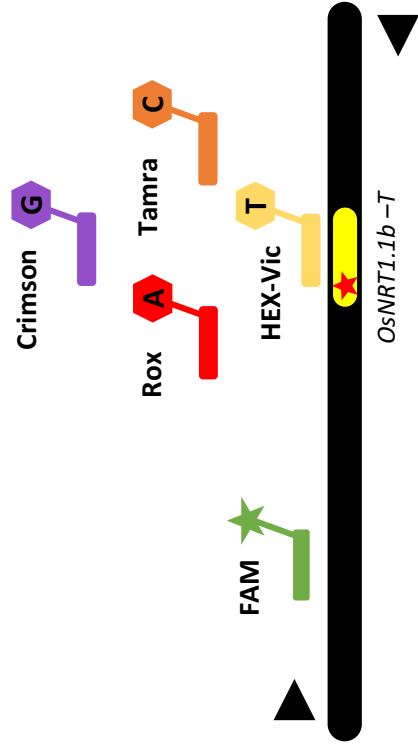
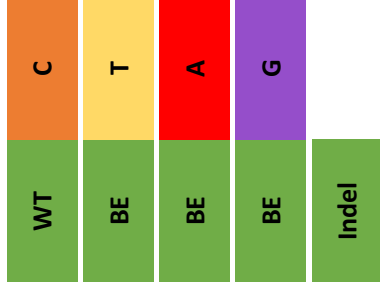
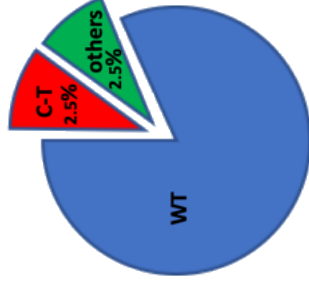
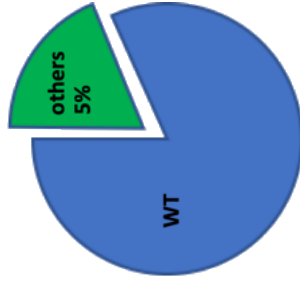
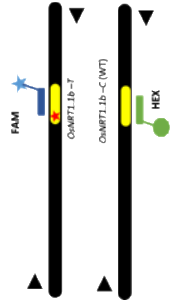
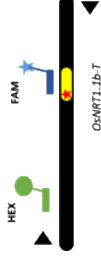
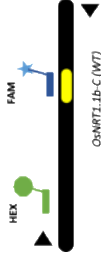


Figure 8

Dropoff **DropOn** **DropOn** **DropOn**



5%

0%

0%

0%

5%

2.5%

2.6%

2.5%

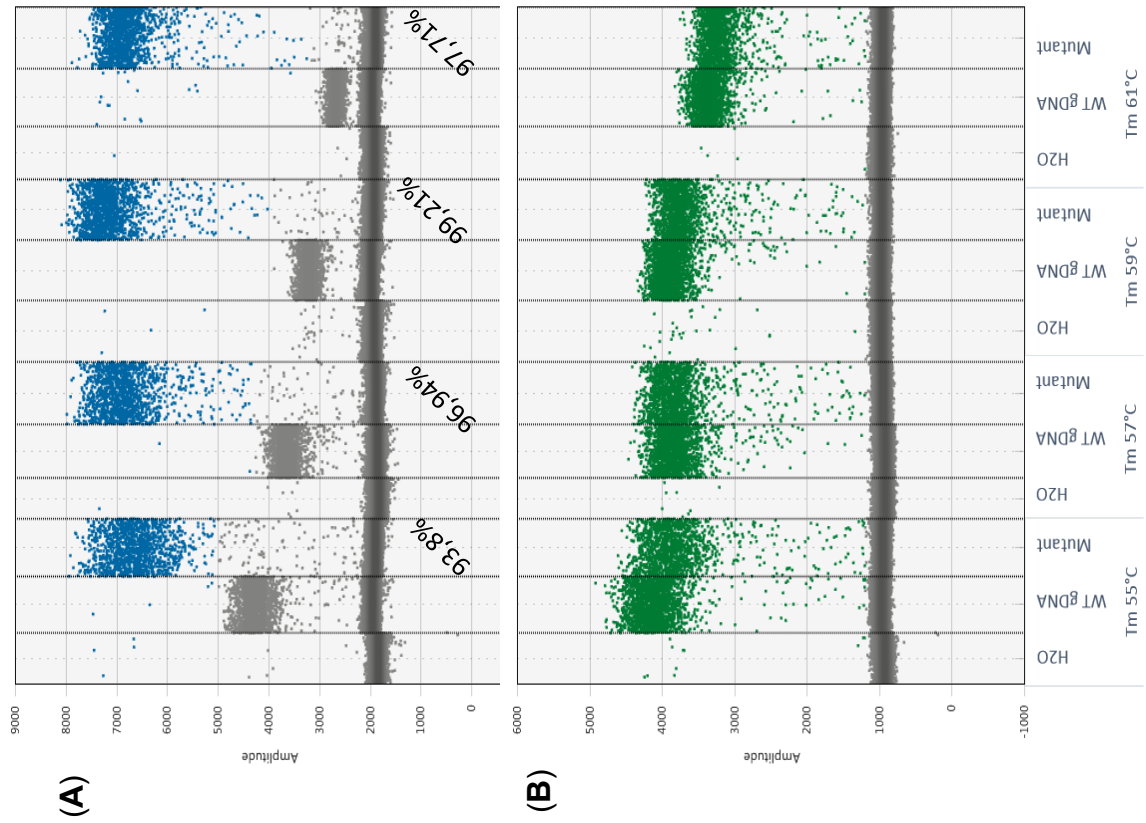
5%

5%

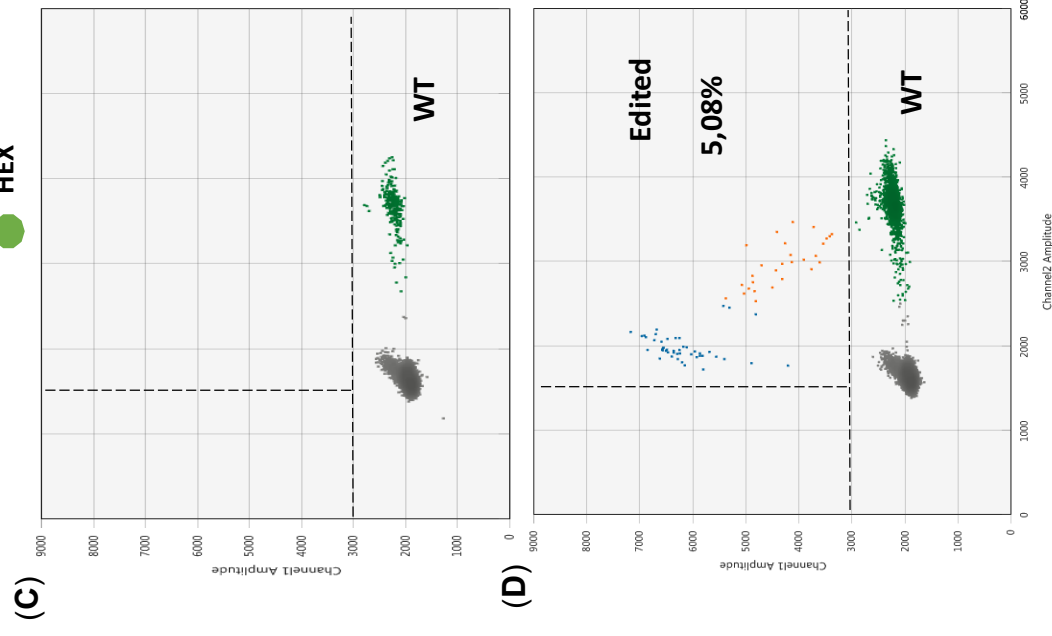
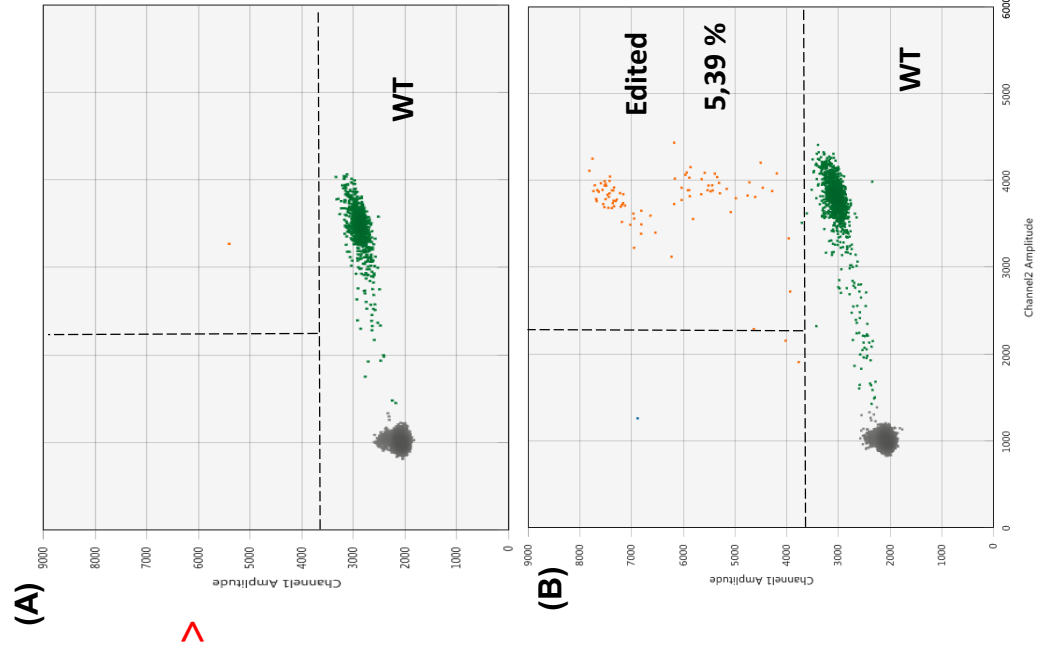
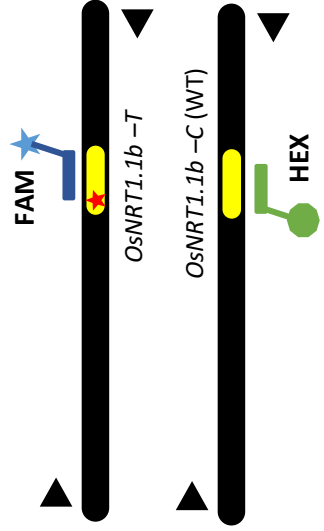
5%

5%

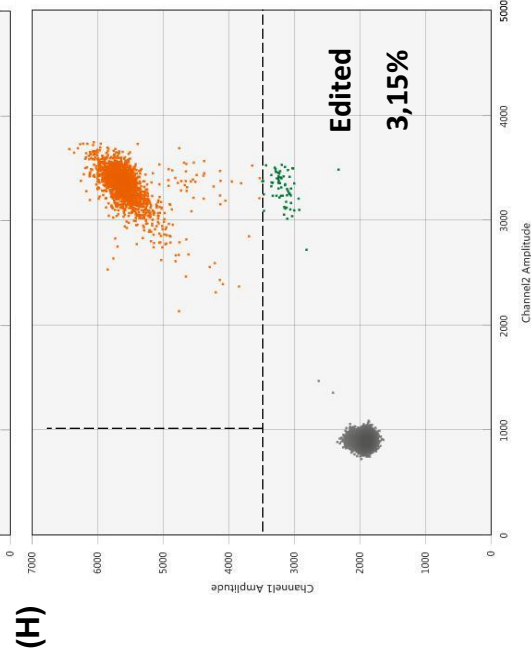
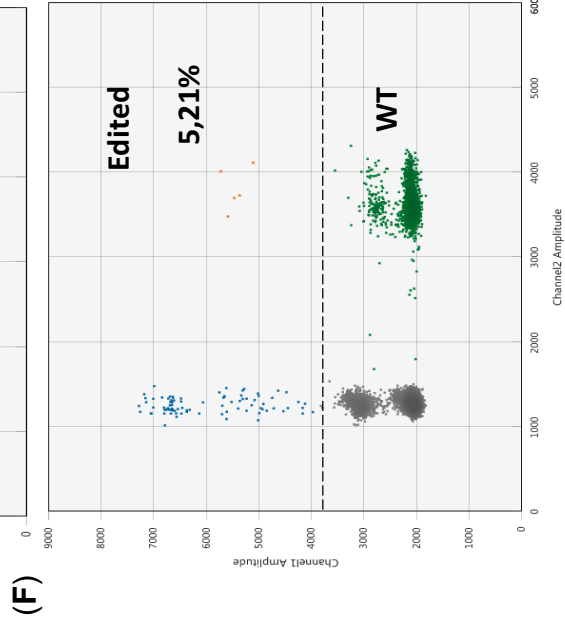
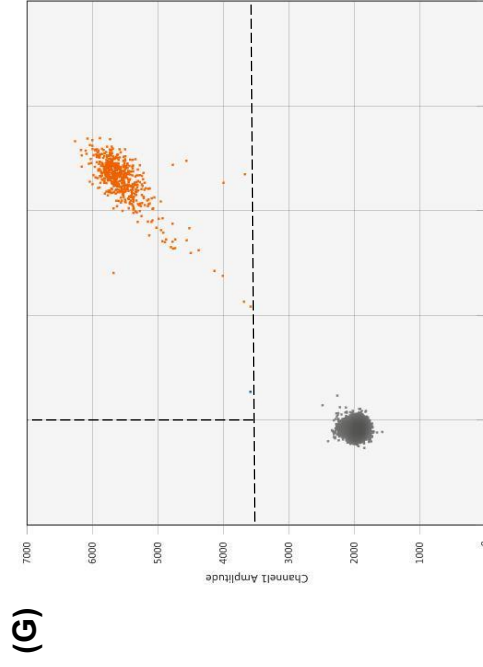
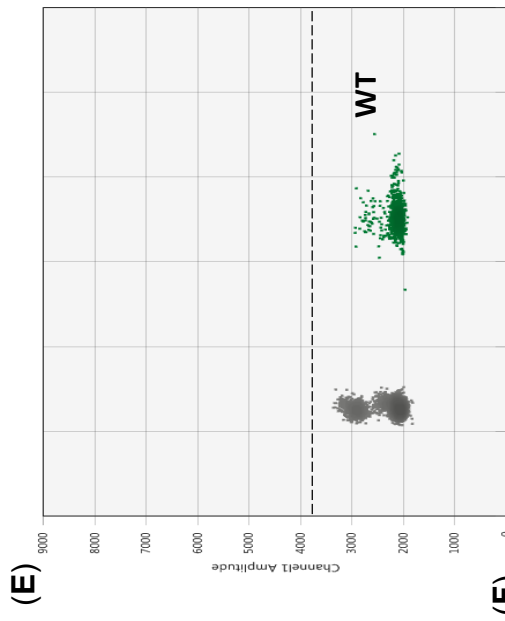
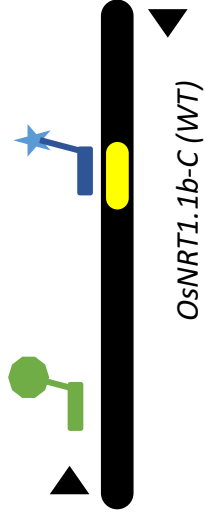
Figure 9

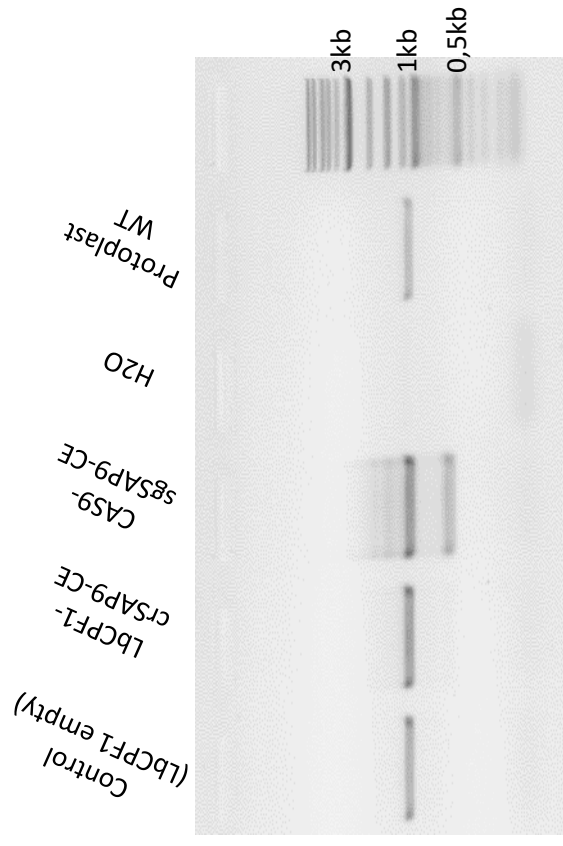


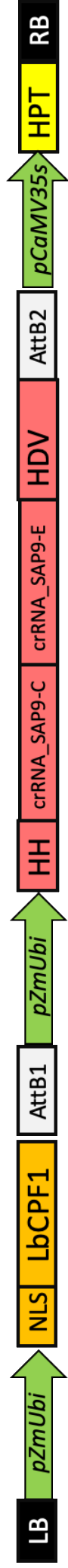
Supplemental Figure 1



⇒







Supplemental Figure 4

Kitaake genotype	Mutation	Description	Number of plants
WT	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 1 : CGGCGACGGCGAGCAAGTGG Allele 2 : CGGCGACGGCGAGCAAGTGG	WT	12
HO	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 1 : CGGCGATGGCGAGCAAGTGG Allele 2 : CGGCGATGGCGAGCAAGTGG	Homozygous Base Editing C in T	10
HT	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 1 : CGGCGATGGCGAGCAAGTGG Allele 2 : CGGCGACGGCGAGCAAGTGG	Heterozygous Base Editing C in T	12
HT	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 1 : CGGCGATGGCGAGCAAGTGG Allele 2 : CGGCGAGGGCGAGCAAGTGG	Base Edition heterozygous G and T	2
HT	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 1 : CGGCGAGGGCGAGCAAGTGG Allele 2 : CGGCGAGGGCGAGCAAGTGG	Heterozygous Base Editing C and G	1
HT	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 1 : CGGCG.AGG Allele 2 : CGGCGAGGGCGAGCAAGTGG	Heterozygous deletion - 15	1
HT	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 1 : GACGG.TGT Allele 2 : CGGCGAGGGCGAGCAAGTGG	Heterozygous deletion - 15	1
HT	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 1 : CGGCGAGGGCGAGCAAGTGG Allele 2 : CGGCGAGGGCGAGCAAGTGG	Heterozygous deletion - 16	1
Indéterminé	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 2 : CGGC.	No determinated	3
CHHD genotype	Mutation	Précision	Nombre de plantes
WT	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 1 : CGGCGACGGCGAGCAAGTGG Allele 2 : CGGCGACGGCGAGCAAGTGG	WT	2
HO	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 1 : CGGCGATGGCGAGCAAGTGG Allele 2 : CGGCGATGGCGAGCAAGTGG	Homozygous Base Editing C in T	1
HT	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 1 : CGGCGATGGCGAGCAAGTGG Allele 2 : CGGCGAGGGCGAGCAAGTGG	Heterozygous Base Editing C in T	3
HT	Wt : GGCGGAGTCGCGGGCAGGGCCGAGCAAGTGGAGGCT Allele 1 : CGGCGAGGGCGAGCAAGTGG Allele 2 : CGGCGAGGGCGAGCAAGTGG	Heterozygous Base Editing C and G	1

Tissue	Position	Gene	Primer name	Oligo sequence	Amplicon size	Probes name	Probes sequence	Tm probe	Fluorophore	
All	LOC_0503g50885	<i>qPCR33</i>	dp-qPCR 33-F	TGGACAAAGITTTCAACCGGC	153	HEX-qPCR33	CTACCTGTGGTTCACACTGTGA	61	HEX	
			dp-qPCR 33-F	GTACCCTCATCAGGCATC		FAM-qPCR33	CTACCTGTGGTTCACACTGTGA			62,9
	LOC_0510g40600	<i>O5NRT1.1b</i>	dp-F1 NRT1.1	GTCGGTTGGACCGATAAGA	243	HEX NRT1.1b-WT	CTCCACTTGCTCGCGTTCG	61	HEX	
			dp-R1 NRT1.1	ACCACCTGCTCACCTCT		HEX NRT1.1b-control	ACCGTTAGAACCAACCAAGGTAATAA			HEX
						FAM NRT1.1b-T	CTCCACTTGCTCGCCATCGC			FAM
	LOC_0507g07350	<i>O5SAP9</i>	dp-SAP9del-F1	CACCCATTTGCAGCCTGTTT	759/253*	HEX-SAP9-2	CCAGTAGGTTCTTCAGGTCT	61,2	HEX	
		dp-SAP9del-R1	GCCAGAGAAATGAAACCATG							

Table 1: primers and probes for ddPCR

Primer Name	Primer sequence	Purpose	Amplicon size	Spacer	Construct	Species
sgNRT.1.1-F	GGCAGGCGGACGGCAGCAAGTGG	Cloning sgRNA BE		CGGCGACGGCGAGCAAGTGG	nCAS9	<i>Oryza sativa</i>
sgNRT.1.1-R	AAACCACCTTGTCTGGCGTGC	Cloning sgRNA BE				
NRT-F3	ATCGGCCGCTAGGGATATT	Sequencing	533	NA	NA	<i>Oryza sativa</i>
NRT-R3	GTAGAAAGGGACGGTGAGGA	Sequencing		NA	NA	<i>Oryza sativa</i>
sgSAP9-C-F	GTGT GTACAGCCTGCCGGTAGCAA	Cloning sgRNA (in Bsal)		GTACAGCCTGCCGGTAGCAA	Multiplex deletion CAS9	<i>Oryza sativa</i>
sgSAP9-C-R	AAACT TGCTACCGGCAGGCTGTAC	Cloning sgRNA				
sgSAP9-E-F	TGTT GGCCATTTAAATCTCTATTCA	Cloning sgRNA (in BtgZI)		GGCCATTTAAATCTCTATTCA	Multiplex deletion CAS9	<i>Oryza sativa</i>
sgSAP9-E-R	AAACT GAAATAGAGATTTAAATGGCC	Cloning sgRNA				
crSAP9-C-F		Cloning crRNA		CAGCTTTCTTGCCCATTCGCTAC	Multiplex deletion LbCpf1	<i>Oryza sativa</i>
crSAP9-C-R		Cloning crRNA				
crSAP9-E-F	AGAT CAGCTTTCTTGCCATTGCTACTAATTTCTACTAA GTGTAGATATCTCTATTTCATGGTGGCAACCG	Cloning crRNA		ATCTCTATTTCATGGTGGCAACCG	Multiplex deletion LbCpf1	<i>Oryza sativa</i>
crSAP9-E-R	GGCC CGGTTGCCACCAGTAATAGAGATATCTACACTTAG TAGAAATTAGTAGCAATGGCACAGAAAGCTG	Cloning crRNA				
p53-F1	CCTACTGCGCAATGCTGATA	Sequencing SAP9 deletion	1116/595*	NA	NA	<i>Oryza sativa</i>
p53-R1	GGCTGCAGTAGCAAGGTTTC	Sequencing SAP9 deletion		NA	NA	<i>Oryza sativa</i>
BtgZI primer	cgttaggggagacaggttt	Sequencing pENTRY multiplex clone (1st spacer)				
Bsal primer	atcatgaaccaacggcctgg	Sequencing pENTRY multiplex clone (2d spacer)				
UbiRac primer	TGCAATGCTCAITATCTCTAGAG	Sequencing pZmUbi-OsCas9-HPT and pUbi_LbCpf1-destvect4.0 after LR reaction				

Table 2: Primers used for cloning, mutation detection and sequencing

Gene	Construct	Cultivar	Edition
<i>OsNRT1.1</i>	nCAS9, Base editing C->T	<i>Oryza sativa japonica</i> cv Chhomrong dhan	57% (n=7)
<i>OsNRT1.1</i>	nCAS9, Base editing C->T	<i>Oryza sativa japonica</i> cv Kitaake	55,8% (n=43)
<i>OssAP9</i>	CAS9 deletion	<i>Oryza sativa japonica</i> Nipponbare	20,8% (n=24); 16,8% (n=107)*
<i>OssAP9</i>	LbCPF1 deletion	<i>Oryza sativa japonica</i> Nipponbare	0% (n=40); 1,03% (n=97)*

Table 3: Mutations identified in planta for *OsNRT1.1* and *OssAP9* genes

Probes colors	Dyes	Name	Sequences			
			Primer F	Primer R	Probe	
Green	Fam	REF_NRT_Control	GTCGGTTTGGACCGATAAGA	ACCACCTGCTTCACCTCCT	ACCGTTAGAACCACCAAGGTAATAA	
Yellow	Hex-Vic	NRT1.1-T	GTCGGTTTGGACCGATAAGA	ACCACCTGCTTCACCTCCT	CTCCACTTGGCTCGCCATCGC	
Orange	Tamra	NRT1.1-C	GTCGGTTTGGACCGATAAGA	ACCACCTGCTTCACCTCCT	CTCCACTTGGCTCGCCGTCG	
Red	Rox	NRT1.1-A	GTCGGTTTGGACCGATAAGA	ACCACCTGCTTCACCTCCT	CTCCACTTGGCTCGCCTCGCG	
Crimson	Cy5	NRT1.1-G	GTCGGTTTGGACCGATAAGA	ACCACCTGCTTCACCTCCT	CTCCACTTGGCTCGCCCTCG	

Table 4

En conclusion nous avons réussi à détecter 5 % d'édition en moyenne sur plusieurs répliquas de protoplastes. De plus, nous avons pu déterminer précisément l'efficacité de délétion par deux endonucléases différentes par une autre stratégie de coamplification applicable pour de grandes délétions.

Le DropOff est le système le plus performant pour déterminer un taux de mutation global, induit par une construction CRISPR/CAS9. Le DropOn permet d'estimer le taux de mutation spécifique *ie* attendu, par exemple pour tester l'efficacité d'un construit d'édition à la base près. Alternativement, la stratégie de compétition permet elle aussi de déterminer le taux d'édition spécifique ou global, bien qu'il existe un biais. En revanche, la stratégie par co-amplification est beaucoup plus délicate à interpréter et nécessite d'utiliser les valeurs obtenues sur le WT pour déterminer, de manière indirecte les taux d'édition. Hormis dans des cas spécifiques, comme pour déterminer le taux de délétion entre la CPF1 et la CAS9 ((Herbert et al., in prep) **Figure 7**), il est plus prudent, lorsque cela est possible, d'utiliser les stratégies DropOff et DropOn pour estimer le taux de mutation global, ou spécifique.

Dans le cas de l'édition à la base près, où l'on souhaite estimer à la fois ce taux de mutation spécifique, c'est-à-dire le taux de mutation attendue (par exemple la transversion C-T pour *NRT1.1b*), et le taux de mutation non attendue (par exemple toutes les autres mutations non attendues comme les insertions, délétions, ou d'autres substitutions), de base il faut utiliser le DropOn, pour mesurer ce taux de mutation spécifique, et le DropOff pour estimer le taux global de mutation. En théorie, si le taux global de mutation est supérieur au taux de mutation spécifique, alors le construit induit majoritairement des mutations non attendues. Malheureusement, avec la technologie BioRad qui utilise deux fluorochromes, il n'est pas possible d'estimer simultanément ces deux valeurs. Or, dans le cas de l'édition à la base près de *NRT1.1b* chez le riz, nous avons vu (**Annexe 2**) que le construit en transformation stable, induisait environ 15 % de mutations non attendues, principalement, des petites délétions.

En utilisant les cinq fluorochromes avec la technologie ddPCR QIAcuity, nous allons essayer d'estimer l'ensemble des mutations monobases en protoplastes (Herbert et al., in prep ; **Figure 8**). Cette technologie représente un progrès considérable. Elle permettra de prédire à la fois la sensibilité et la spécificité de construits CRISPR, comme les construits d'édition à la base près en particulier pour des espèces comme la vigne, où il est critique d'estimer ces effets avant la transformation stable en raison du temps et de la quantité de travail que représente la génération de plantes transgéniques éditées.

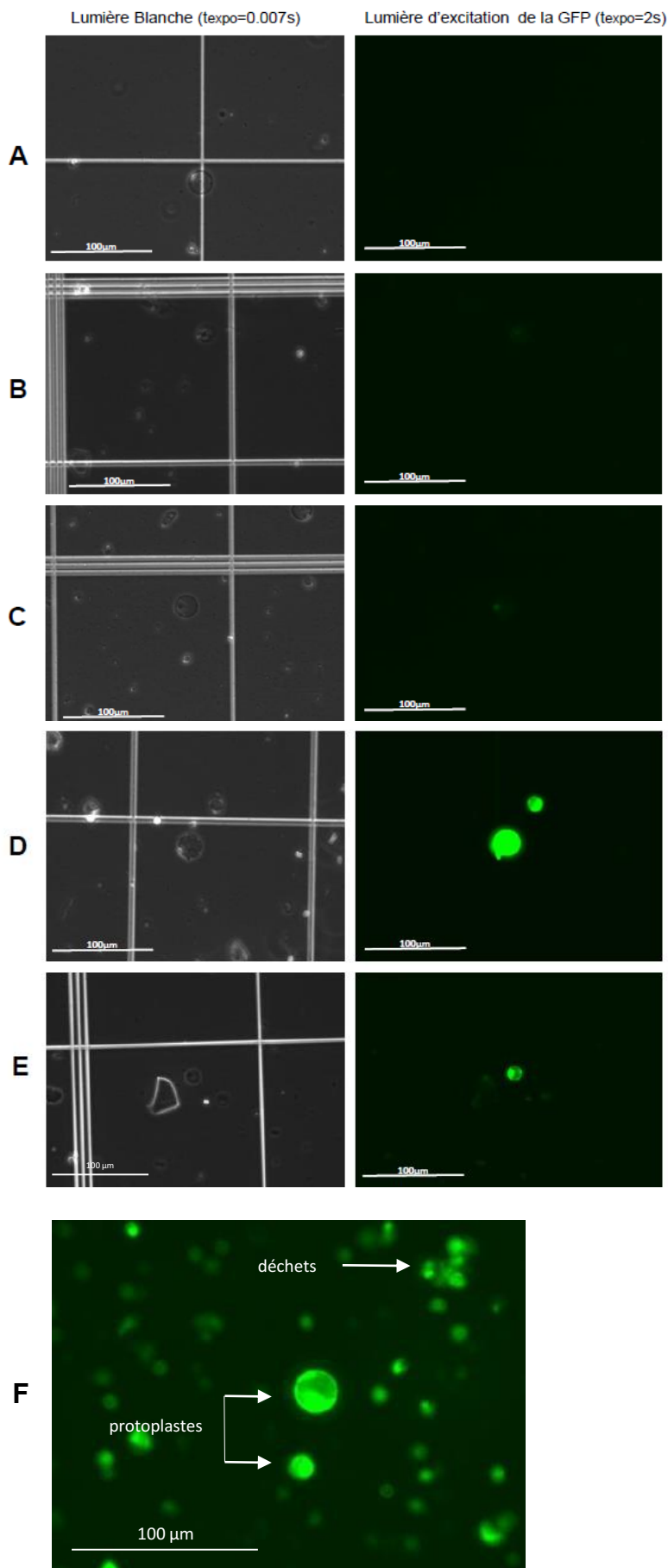


Figure 24 : Transformation de protoplastes de riz avec la eGFP

Le test de transformation est visualisé 44h après transformation au microscope à fluorescence avec un temps d'exposition de 0,007s en filtre LB (lumière blanche), et 2s en filtre L5 (excitation 494 nm, réception 512nm).

Les témoins négatifs sans ADN ni PEG (**A**), sans ADN avec PEG (**B**), et avec ADN (pHKN29) sans PEG (**C**) ne présentent pas de fluorescence.

Le témoin positif avec PEG et le pHKN29 montre une très forte fluorescence (**D**) .

Les protoplastes transformés avec la protéine eGFP et du PEG présentent une fluorescence moins intense mais bien visible (**E**) .

Deux protoplastes transformés avec la protéine eGFP. La GFP s'accumule également autour des déchets. La photo a été prise 14 h après transformation (temps d'exposition = 5 s (**F**)).

III] Transformation transitoire à l'aide de RiboNucléoProtéines (RNP) chez le riz : résultats préliminaires

Dans le cadre des optimisations possibles de la technologie CRISPR, nous avons également initié des travaux préliminaires visant à démontrer qu'il était possible d'utiliser la protéine CAS9 complexée au guide *in-vitro* (RiboNucleicProtein complexe) pour induire des mutations en protoplastes chez le riz. Cette technique permet de s'affranchir de la transformation stable de construits T-DNA, et présente donc plusieurs avantages : 1. Cela évite d'avoir à éliminer le transgène par ségrégation, notamment pour des espèces pérennes à multiplication végétative comme la vigne où la ségrégation du transgène est impossible sans passer par des systèmes de Heatshock-transposons FRT (Dalla Costa et al., 2018). 2. Elle n'introduit pas d'ADN, ce qui réduit le risque d'introduire des fragments du vecteur binaire qui se produisent parfois lors de transformations utilisant *Agrobacterium tumefaciens* (Buck et al., 2000). 3. Enfin, dans le cadre réglementaire de certains pays, cette technique n'est pas considérée comme de la transgénèse ce qui facilite l'exploitation des plantes éditées, pour l'amélioration variétale (Dalla Costa et al., 2019 ; Sprink et al., 2016). Elle nécessite par contre de pouvoir cribler les événements d'édition et de régénérer des plantes fertiles à partir de protoplastes, technique qui est maîtrisée chez très peu d'espèces. Il s'agit de résultats préliminaires, mais encourageants qu'il conviendrait de développer plus avant dans la suite de mon travail de thèse.

Dans un premier temps, pour vérifier qu'il était possible de transformer des protoplastes avec des protéines, nous avons utilisé de la protéine pure améliorée eGFP (enhanced GFP, CliniScience ref. 4999-100, resuspendue à 10 ug/uL dans du PBS 1x) (excitation à 480 nm) comme contrôle de transformation en comparant la fluorescence obtenue avec celle du plasmide, en suivant le protocole classique de transformation utilisant du PEG. Afin de nous débarrasser de la protéine GFP qui pourrait rester dans le milieu extérieur aux protoplastes, nous avons traité le culot contenant les protoplastes transformés avec de la trypsine à 0.025 % pendant 5 min avant d'être centrifugé et resuspendu dans du W5 comme le protocole de (Bès et al., 2021). Dans l'esprit, cette étape est comparable au traitement DNase utilisé dans le protocole de transformation utilisant de l'ADN.

Dans un premier temps, dans la condition utilisant du PEG, les protéines ont eu tendance à s'agréger autour des protoplastes et des débris cellulaires. Le PEG, hydrophobe, a tendance à agréger les protéines et l'ADN autour des membranes plasmiques et des débris (**Figure 24**). Il est par exemple possible de voir des protoplastes 14h après transformation entourés de eGFP, puis 44 h après transformation la eGFP semble devenir uniquement intracellulaire. L'explication la plus simple est que l'eGFP s'accumule d'abord sur les membranes, sous l'effet du PEG, puis pénètre par endocytose dans les protoplastes (**Figure 24-E et F**). Nous avons donc démontré simplement qu'il était possible de transformer directement des protéines, en utilisant le protocole de transformation PEG riz. Ce premier résultat encourageant ne nous permettait pas de conclure directement que le complexe RNP allait pénétrer. En effet, la eGFP a une taille de 27 kDa alors que le complexe RNP a une taille proche de 160 kDa ce qui aurait probablement un effet sur l'efficacité de pénétration.

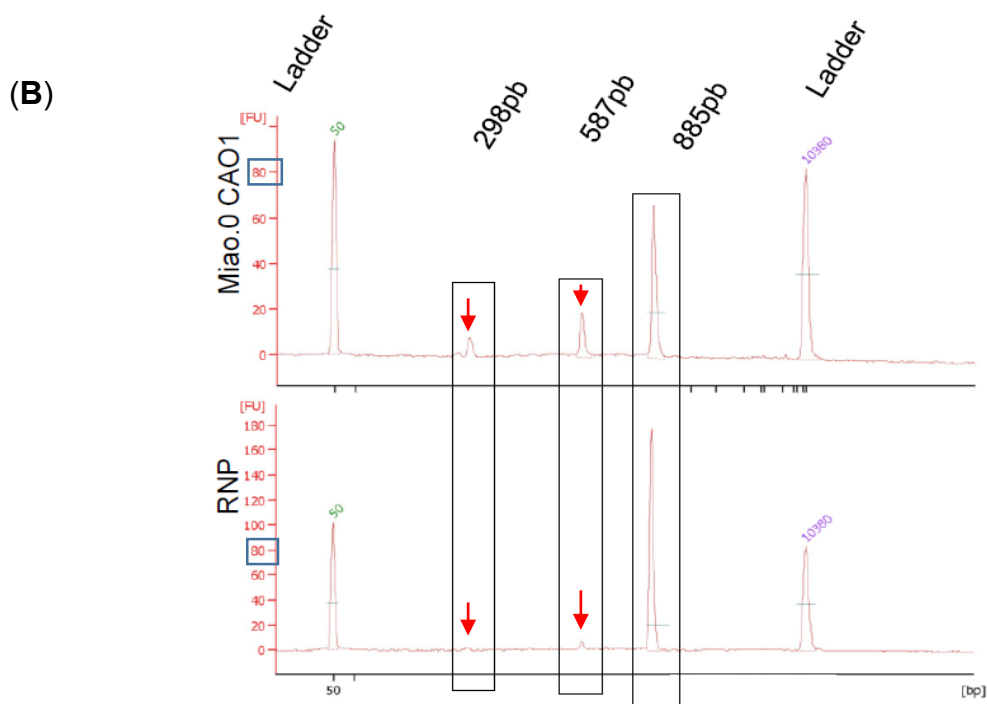
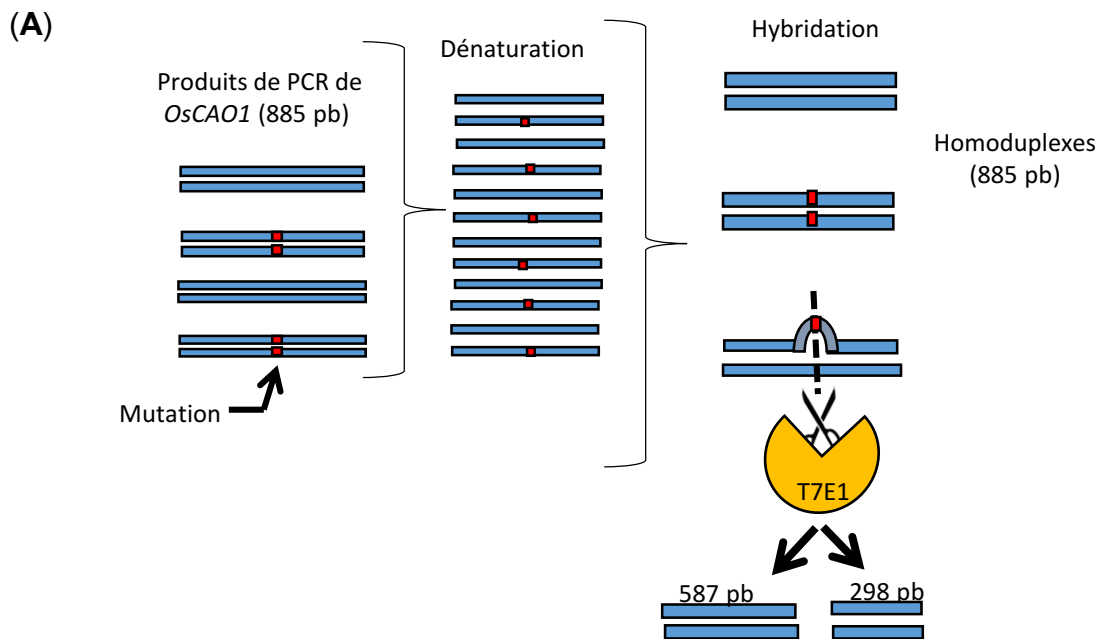


Figure 25 : Résultats du test T7E1 sur les protoplastes transformés par Miao.0 CAO1 et par RNP CAS9-sgRNACAO1.

Principe de la digestion des hétéroduplexes PCR reformés par la T7E1. Les substrats de l'endonucléase T7E1 sont des ADN doubles brins déformés ayant subi un brutal désappariement-réappariement. La T7E1 va couper tous les hétéroduplexes contenant des mésappariements. En mélangeant des produits de la PCR issus d'un mélange de gènes sauvages et de gène édité, les hétérodimères vont former des boucles qui seront clivés par la T7E1. **(A)**

Résultats obtenus par le test T7E1 sur Bioanalyseur (électrophorégramme).

Les protoplastes transformés par le plasmide présentent les bandes aux tailles attendues en cas de mutations au locus *CAO1* à 587 pb et 298 pb. Les protoplastes transformés par RNP présentent également les bandes. Les pics sont beaucoup plus faibles suggérant une moindre efficacité que le plasmide probablement dû à une faible quantité de complexes RNP transformés. Les bandes à 885 pb représentent les homoduplexes des produits PCR reformés **(B)**.

Nous avons ciblé le gène *OsCAOI*, pour lequel nous avons des sgRNA ayant démontrés leur efficacité en protoplastes et *in-planta*. Nous sommes donc repartis d'une séquence sgRNA fonctionnelle pour produire les sgRNA ciblant le gène *OsCAOI in-vitro*. Nous avons utilisé le kit Clontech Guide-it sgRNA In Vitro Transcription and Screening System (ref 631438, 631439, 631440) en suivant les recommandations du fournisseur, en clonant le sgRNA sous contrôle du promoteur T7 et avons utilisé de la CAS9 commerciale (ThermoFisher ref 15403923 GeneArt Platinum Cas9 Nuclease).

Lors de la transformation par RNP, nous avons rajouté une étape de « pré-chargement » de la CAS9 : 20 µg de CAS9 et 5 µg de sgRNA sont incubés pendant 10 min à 25 °C avant la mise en contact avec les protoplastes. Puis nous avons vérifié que le complexe incubé était capable *in-vitro*, de digérer le produit PCR contenant la cible *OsCAOI* (**Annexe 3**). Les résultats de cette transformation après le test T7E1 ont été analysés sept fois au bioanalyseur (Agilent2100). Les produits de PCR purifiés issus des protoplastes transformés par RNP présentent une bande intense à 885 pb et deux bandes beaucoup moins intenses aux tailles de 587 pb et 298 pb (**Figure 25**). Ces bandes sont exactement aux tailles attendues si des mutations apparaissent au locus *OsCAOI* (Bès et al., 2021). Il semble donc que, malgré la faible quantité de sgRNA par rapport à la CAS9, le complexe RNP puisse induire des mutations au locus *CAOI* en protoplastes. Quantitativement, au vu du ratio d'intensité entre les bandes obtenues avec le RNP et le plasmide, il semble néanmoins que l'efficacité de coupure soit très inférieure avec la transformation RNP par rapport à la transformation du plasmide Miao.0CAO1, probablement dû aux concentrations en CAS9 et en sgRNA (Bès et al., 2021 ; Miao et al., 2013).

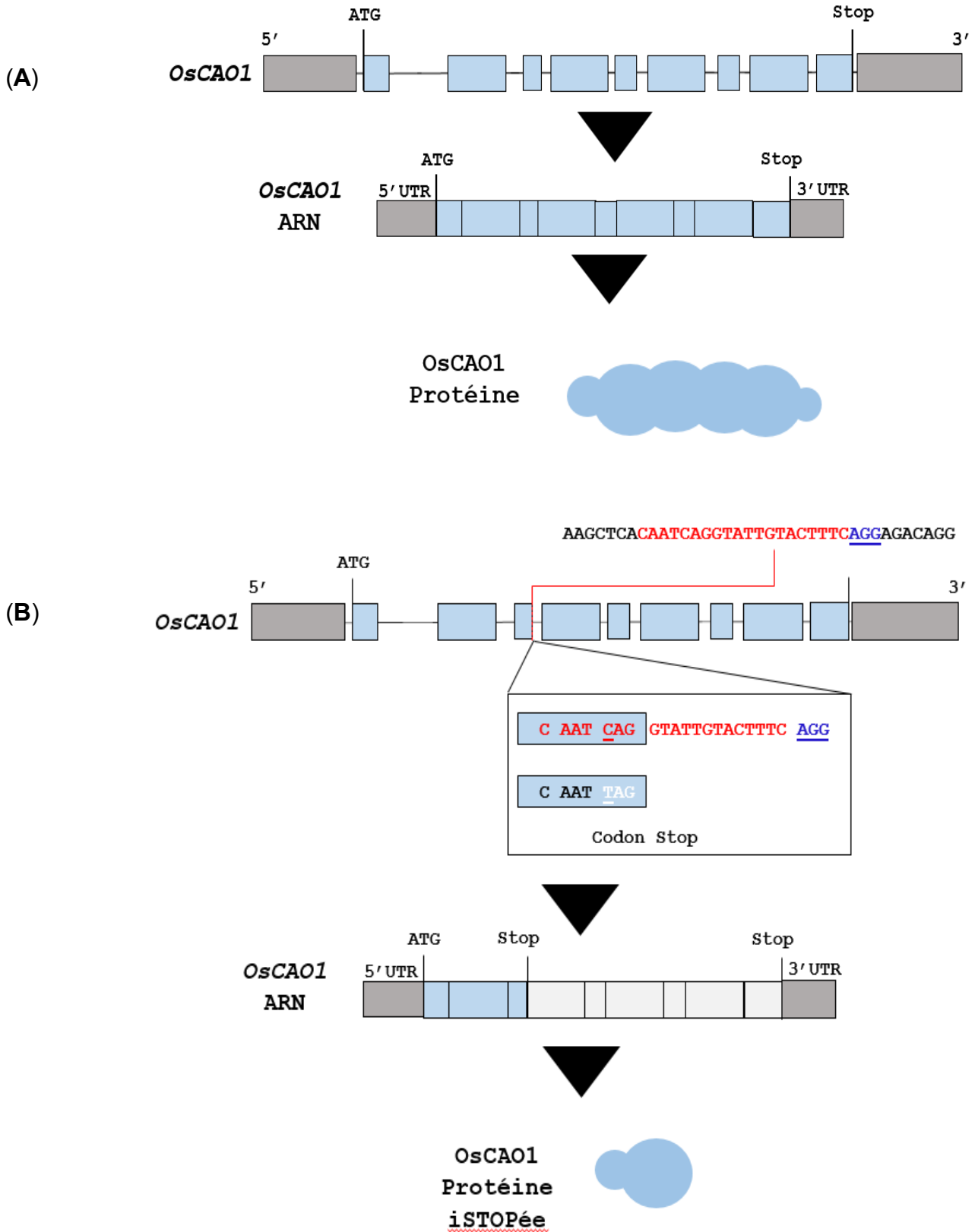
Nous n'avons pas eu le temps de répéter ce résultat préliminaire intéressant. Il est sans doute possible d'optimiser l'utilisation de RNP chez le riz, par exemple en augmentant la concentration du complexe ou encore en augmentant le temps de contact avec le PEG. En effet, la seule condition RNP qui a pu être testée en protoplastes a consisté à utiliser 20 µg de CAS9 pour 5 µg de sgRNA, alors que dans une publication utilisant une approche similaire (Woo et al., 2015), les auteurs utilisent au moins quatre fois plus de sgRNA pour induire des mutations détectables. Néanmoins, ces résultats démontrent que l'utilisation de RNP pour induire des mutations ciblées est parfaitement possible, en modifiant à minima, le protocole utilisé pour transformer de l'ADN en protoplastes et ouvre de nouvelles perspectives. Les RNP peuvent être utilisés par bombardement de cals, ou en protoplastes à la condition de pouvoir les régénérer.

Chapitre 2 : Application de l'édition des génomes au riz et à la vigne

Dans le chapitre précédent, nous avons démontré en système protoplaste, l'efficacité de nos constructions CRISPR/CAS9 pour générer des décalages du cadre de lecture (frameshifts) et pour induire des modifications à la base près. Dans ce second chapitre, dans un premier temps, nous avons testé l'efficacité de ces différentes constructions *in-planta* en comparant l'efficacité de la CAS9 et de la CPF1 et également démontré qu'il est possible de générer des mutants KO prédictibles à l'aide d'un construit « édition à la base près », stratégie qui a été nommée iSTOP en cellules animales (Billon et al., 2017). Dans une deuxième partie, nous allons aborder plus en détail les résultats obtenus chez le riz Kitaaké pour l'édition à la base près du gène *OsNRT1.1b* et les mutations KO sur les gènes *OsBT* et *OsNRT1.1b*. Nous analyserons différents paramètres agrophysiologiques pour l'ensemble de ces mutants sur deux conditions de culture enrichies en azote afin de tester leur impact sur la NUE. Puis nous mesurerons l'effet des mutations de *OsNRT1.1b* et *OsBT* sur l'influx de ^{15}N et leurs expressions. Les mêmes mutations que pour Kitaaké ont été obtenues sur Chhomrong Dhan (CHHD), une variété népalaise très cultivée sur les hauts plateaux malgaches. Les résultats des expérimentations au champ en Colombie sur les lignées CHHD éditées à la base près pour *OsNRT1.1b*, KO pour *Osbt* et *Osnrt1.1b*, et leurs effets sur la NUE, viendront clore cette deuxième partie sur le riz. Pour finir ce chapitre, nous analyserons les résultats préliminaires obtenus chez la vigne via la transformation transitoire de divers plasmides CRISPR en protoplastes en appliquant la détection par ddPCR. Nous traiterons ensuite de la transformation stable des plasmides CRISPR ciblant le gène *VvGAI*. Nous testerons pour finir l'application sur une alternative d'amélioration en utilisant des plasmides CRISPR « optimisés » pour la vigne.

I] Les technologies CRISPR : ou comment générer une série allélique à un locus pour faciliter l'analyse fonctionnelle et l'utilisation des construits d'édition comme NBT (New Breeding Technologies)

Les technologies d'édition du génome permettent d'obtenir des mutations ciblées. Elles dépendent à la fois de la nucléase, et du construit utilisé. Théoriquement, il est possible de générer une série allélique de mutations distinctes qui facilitent, par exemple, l'analyse de la fonction d'un gène. Nous avons choisi, dans l'étude présentée dans l'article « Herbert et al. 2020 Beyond Seek and Destroy: how to Generate Allelic Series Using Genome Editing Tools, Rice 13 :5 » de tester cette possibilité, en comparant les allèles produits par la CAS9, la CPF1 au locus *OsCAO1*. Le locus *OsCAO1* est un gène « proof of concept ». Lorsqu'il est muté, on observe un phénotype de plante jaune, avec un retard de croissance, mais qui n'est pas stérile ni létal. Enfin, nous avons également testé la possibilité d'induire des pertes de fonction ciblées, en utilisant « l'édition à la base près » avec le système base editor BE3 pour créer un codon STOP en convertissant spécifiquement un C d'un acide aminé CAG, en T, pour créer un TAG,




directement dans le gène (**Figure 26**). L'intérêt de ces différents construits pour générer une série allélique sera également discuté. Le but étant de proposer le meilleur outil pour la stratégie de recherche envisagée.

SHORT COMMUNICATION

Open Access

Beyond Seek and Destroy: how to Generate Allelic Series Using Genome Editing Tools



Leo Herbert^{1,2†}, Anne-Cécile Meunier^{1,2†}, Martine Bes^{1,2}, Aurore Vernet^{1,2}, Murielle Portefaix^{1,2}, Franz Durandet^{1,2}, Remy Michel^{1,2}, Christian Chaine^{1,2}, Patrice This^{1,2}, Emmanuel Guiderdoni^{1,2} and Christophe Périn^{1,2*} 

Abstract

Genome editing tools have greatly facilitated the functional analysis of genes of interest by targeted mutagenesis. Many usable genome editing tools, including different site-specific nucleases and editor databases that allow single-nucleotide polymorphisms (SNPs) to be introduced at a given site, are now available. These tools can be used to generate high allelic diversity at a given locus to facilitate gene function studies, including examining the role of a specific protein domain or a single amino acid. We compared the effects, efficiencies and mutation types generated by our LbCPF1, SpCAS9 and base editor (BECAS9) constructs for the *O5CAO1* gene. SpCAS9 and LbCPF1 have similar efficiencies in generating mutations but differ in the types of mutations induced, with the majority of changes being single-nucleotide insertions and short deletions for SpCAS9 and LbCPF1, respectively. The proportions of heterozygotes also differed, representing a majority in our LbCPF1, while with SpCAS9, we obtained a large number of biallelic mutants. Finally, we demonstrated that it is possible to specifically introduce stop codons using the BECAS9 with an acceptable efficiency of approximately 20%. Based on these results, a rational choice among these three alternatives may be made depending on the type of mutation that one wishes to introduce, the three systems being complementary. SpCAS9 remains the best choice to generate KO mutations in primary transformants, while if the desired gene mutation interferes with regeneration or viability, the use of our LbCPF1 construction will be preferred, because it produces mainly heterozygotes. LbCPF1 has been described in other studies as being as effective as SpCAS9 in generating homozygous and biallelic mutations. It will remain to be clarified in the future, whether the different LbCPF1 constructions have different efficiencies and determine the origin of these differences. Finally, if one wishes to specifically introduce stop codons, BECAS9 is a viable and efficient alternative, although it has a lower efficiency than SpCAS9 and LbCPF1 for creating KO mutations.

Introduction

The demonstration in 2012 and 2013 of the use of the site-specific nuclease CAS9 in eukaryotic systems to precisely mutagenize a DNA region was the first milestone in a revolution in functional biology (Cong et al. 2013; Jinek et al. 2012; Mali et al. 2013). Now commonly used in functional analysis, this technology has considerable potential for molecular breeding based on the impact of variant single nucleotides on traits of agronomic interest

(Es et al. 2019). It is now possible to evaluate and reproduce the effect of a molecular polymorphism in plants and to test its effects on one or more agronomic traits (see, for instance, (Rodriguez-Leal et al. 2017; Zhou et al. 2019; Zsogon et al. 2018)).

CAS9 is an endonuclease guided by an RNA and is currently the most widely used site-specific nuclease. The first-generation 'base editor' was later developed to introduce specific base changes in a target sequence. This was first accomplished by creating chimeric proteins merging a CAS9 nickase (nCAS9), in which either the HNH or the RuvC-like catalytic site has been inactivated, with a cytidine deaminase (Komor et al. 2016; Lu and Zhu 2017). These base editors allow targeted

* Correspondence: perin@cirad.fr

[†]Leo Herbert and Anne-Cécile Meunier are co-first authors and contributed equally to this work

¹CIRAD, UMR-AGAP, F-34398 Montpellier, France

²Université de Montpellier, Cirad, Inra, Montpellier SupAgro, F-34000 Montpellier, France

mutations to be induced without the integration of foreign DNA and without double-strand breakage. nCAS9 can always attach itself to a specific region and cut the nontargeted strand. The cytosine of the ssDNA is then converted to a uracil by the cytosine deaminase, and then, the uracil is replaced by a thymine during the cell cycle or by repair (Komor et al. 2016). Uracil DNA glycosylase (UDG) catalyzes the removal of U from DNA in cells and initiates base-excision repair (BER), including reversion of the U:G pair to a C:G pair. In second generation BECAS9, adding UGI (Uracil Glycosylase Inhibitor) via fusion increases the efficiency of BE by 3-fold. This technology can be used to introduce a given SNP but also to introduce a stop codon at a specific position, a technology called iSTOP (Billon et al. 2017).

CRISPR-Cpf1 (CAS12a) is a new site-specific nuclease that differs significantly from SpCAS9 (Zetsche et al. 2015). The protospacer used is T rich, i.e., 'TTTN', compared to that of SpCAS9, which is GC rich, i.e., 'NGG', allowing it to target regions rich in AT nucleotides, and it creates nucleotide overhangs while SpCAS9 creates blunt double-strand breaks. The mutations induced after repair by these two site-specific nucleases are different, potentially generating distinct alleles.

This set of gene editing tools enables the generation of a greater diversity of mutations that will facilitate the mutational and functional dissection of a given gene. Molecular diversity is a key element in generating phenotypic diversity, which can allow access to the function of a gene or be used in selection. For example, CRISPR/CAS9 technology has been used to generate a series of alleles in the promoter of a gene involved in fruit size in tomato (Rodriguez-Leal et al. 2017), redomesticate wild crop relatives (Lemmon et al. 2018) and improve agronomic traits in rice (Zhou et al. 2019).

The development and comparison of genome editing systems frequently uses marker genes to phenotypically identify visually introduced mutations. Phytoene desaturase (PDS), therefore inactivation, leads to an Albino phenotype (Qin et al. 2007). These two systems have been used frequently to test the effectiveness of the CRISPR/CAS9 system or HR (Homologous Repair) replacement rates (see (Charrier et al. 2019; Wilson et al. 2019) for two recent examples). Nevertheless, there are several limitations to their use. The mutation in the PDS gene is lethal, making it difficult to analyze molecular lesions and, above all, to transmit these lesions to subsequent generations. When the *CAOI* gene is mutated, the plants are shorter, with a pale-yellow phenotype, but remain fertile (Miao et al. 2013). The mutation is semi-dominant and fertile and therefore makes it possible to follow the mutations over several generations. The *CAOI* gene therefore seems to be a good model for

estimating the type of mutations and effectiveness of different GE systems and testing the transmission of mutations over generations. We propose, and illustrate in this paper, the value of using the *CAOI* gene as a marker gene to test the efficacy and type of mutations generated by GE and propose also that it can be used in the future as a complementary marker gene to the PDS in plants.

Few examples are available of comparing the efficacy and value of using SpCAS9, LbCPF1 and BECAS9 (see (Lee et al. 2019) for an example in maize) to create an allelic series for a given gene in a single experimental system. To illustrate the characteristics and complementarity of these three tools, we have chosen to mutagenize the *OsCAOI* gene (*Chlorophyll A Oxygenase 1*) (Miao et al. 2013) in rice. We also used a strategy called iSTOP to offer an alternative to the knockout (KO) of a single gene using BECAS9 (Billon et al. 2017).

Results

Constructs and targets

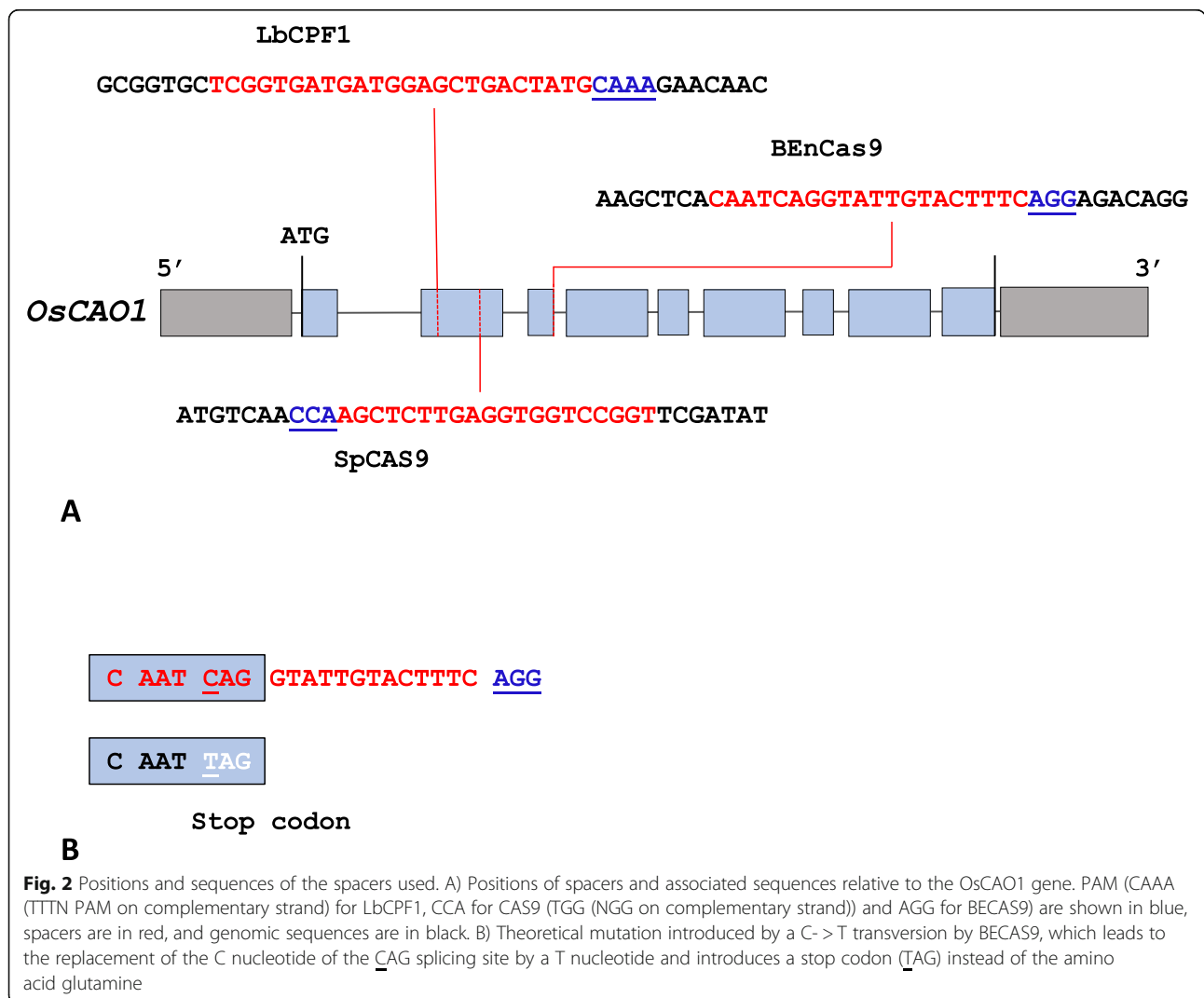
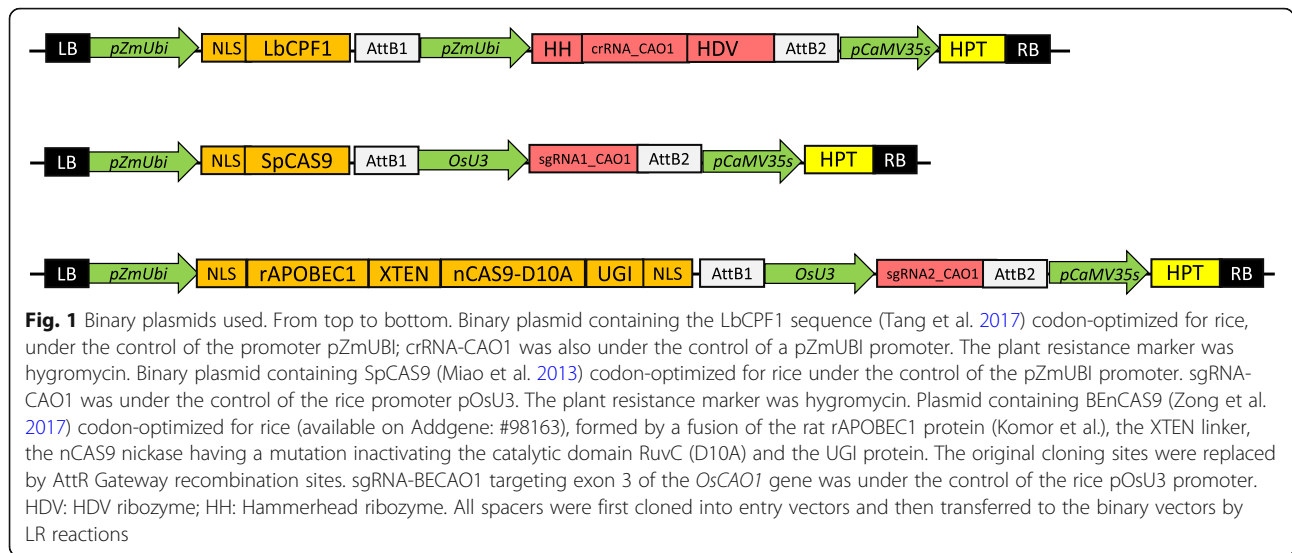
Three binary constructs were prepared (Fig. 1). The first was designed for LbCPF1 expression, the second for SpCAS9 expression, and the last to express a nickase CAS9 (mutation D10A) fused to a rAPOBEC1 protein and a UGI protein (Tang et al. 2017). All coding sequences were placed under the control of the same promoter, pZmUBI, and flanked with one or two NLS sequences to ensure nuclear transport (Fig. 1). Finally, the selectable marker gene *HPT*, which provided rice cells with tolerance to the antibiotic hygromycin, was included in all the T-DNA constructs to facilitate comparisons between the three systems in plants.

Three different spacers were designed for SpCAS9, BECAS9 and LbCPF1 targeting *OsCAOI*. *OsCAOI* exon 2 was targeted by a spacer for SpCAS9 (PAM TGG) and by a crRNA for LbCPF1 (PAM TTTG) (Fig. 2a). Exon 3 was targeted by a spacer (PAM AGG) for BECAS9 (Fig. 2a) to introduce a C->T transversion at the CAG splicing site that converted it into a TAG stop codon. This strategy, called iSTOP, was designed to introduce a stop codon at a specific position while maintaining the reading frame (Fig. 2b).

Mutations generated by SpCAS9, LbCPF1 and BECAS9

All three constructs were transformed into the Kitaake genotype, and 35, 25 and 44 T0 plants were generated, respectively; the mutations of these T0 lines were analyzed by PCR sequencing (Additional file 1: Table S1).

The efficiencies of the three systems, calculated as the ratio of the number of mutations observed to the number of plants regenerated and analyzed, were 94%, 72% and 36%. The SpCAS9 system was the most effective, followed by LbCPF1. In comparison, BECAS9 generated only a 36% mutation rate.



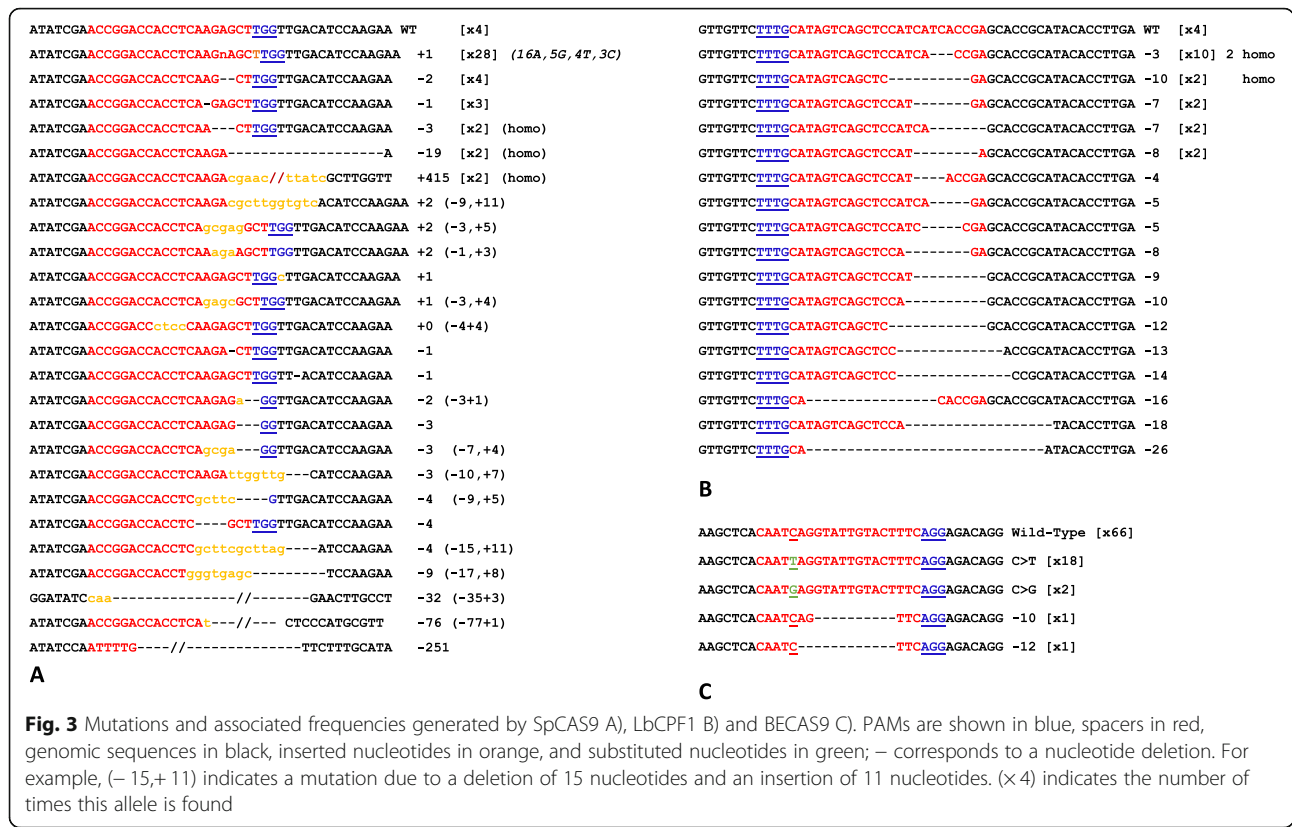


Fig. 3 Mutations and associated frequencies generated by SpCAS9 A), LbCPF1 B) and BECAS9 C). PAMs are shown in blue, spacers in red, genomic sequences in black, inserted nucleotides in orange, and substituted nucleotides in green; – corresponds to a nucleotide deletion. For example, (– 15,+ 11) indicates a mutation due to a deletion of 15 nucleotides and an insertion of 11 nucleotides. (x 4) indicates the number of times this allele is found

The mutations were aligned with the wild-type sequences to compare the types of mutation introduced by each construct (Fig. 3). SpCAS9 mainly generated single-nucleotide insertions 3 nucleotides upstream of the PAM with a preference for the insertion of the nucleotide A (16, 5, 4 and 3 with nucleotides A, G, T and C, respectively) (Fig. 3a). The PAM was also deleted in 36% of the mutants analyzed. For LbCPF1, the situation was completely different. We did not observe any insertions but only deletions of variable size, ranging from 3 to 26 nucleotides, downstream of the PAM (100% of the mutants analyzed). Moreover, in all plants analyzed, the PAM was always present, unlike in SpCAS9 (Fig. 3b). Finally, BECAS9 generated, as expected, mostly C->T transversions located 16 nucleotides upstream of the PAM (82% of the mutants analyzed). More surprisingly, we also identified C->G transversions (10% of the mutants analyzed) and deletions of – 10 and – 12 nucleotides (20%) located 3 nucleotides upstream of the PAM (Fig. 3c).

Comparison of the three systems for generation of homozygous, heterozygous and biallelic mutations

SpCAS9 was the nuclease that generated the most homozygous mutations (37%) compared to LbCPF1 and BECAS9 (12% and 14%) (Table 1). LbCPF1 and SpCAS9 produced similar proportions of biallelic mutants (46% and 40%). The three constructs induced comparable rates of heterozygosity, between 11 and 18%, and the inefficiency of each of the constructs, measured by the number of wild-type plants produced, was on the order of 6% for SpCAS9, 28% for LbCPF1 and 63% for BECAS9 (Table 1).

If we first analyze the ability to restore the open reading frame, significant differences appear between SpCAS9 and LbCPF1 (Table 2). LbCPF1 generated 4 times more in frame deletion than SpCAS9. SpCAS9 mutations are mostly (90%, Table 2) out frame mutations. LbCPF1 produced only simple deletions without insertions (100%, Table 3) while the vast majority of SpCAS9 mutations are

Table 1 Homozygous, heterozygous, biallelic and wild-type plants produced using SpCAS9, LbCPF1 and BE_nCAS9D10A expressed as % of analyzed regenerated plants

	Homozygous	Biallelic	Heterozygous	Wild type
SpCAS9 (n = 35)	37.1% (13)	45.7% (16)	11.4% (4)	5.7% (2)
LbCPF1 (n = 25)	12% (3)	40% (10)	16% (4)	28% (7)
BE_nCAS9D10A (n = 44)	13.6% (6)	0% (0)	22.7% (10)	63.6% (28)

Table 2 Deletions out of frame (Out-frame), deletions that maintain reading frame (In-Frame) generated using SpCAS9, LbCPF1 and BE_nCAS9D10A expressed as % of total mutant alleles identified. A total of 60, 30 and 22 mutants alleles were identified for SpCAS9, LbCPF1 and BE_nCAS9D10A respectively (see also Fig. 3)

	InFrame	OutFrame
SpCAS9	10% (6)	90% (56)
LbCPF1	43% (13)	57% (17)
BE_CAS9	95% (21)	5% (1)

insertions (52%) and complex insertion/deletions (21%). Although BECAS9 generated mainly C- > T transversions, we also found small deletions in 2% of the mutants analyzed (Table 3).

SpCAS9 generated mainly KO mutations via frame-shift leading to the appearance of a premature stop codon (PSC), and we observed the expected phenotype in most of the mutant plants (Fig. 4). In this example, when 35 mutants were compared to the control, 33 plants had mutations, and 29 had the expected pale-yellow phenotype (Fig. 4a). The two plants with phenotypes similar to WT harbored nucleotide deletions in multiples of 3 that maintained the reading frame of *OscAO1*. Similarly, half of the LbCPF1 plants did not have a visible phenotype, because 13 of the 30 alleles contained triplet nucleotide deletions. BECAS9 generated stop codons, and a comparison of heterozygous plants for C- > T transversions allowed us to highlight a difference between the phenotype of the heterozygous and the homozygous T0 plants (Fig. 4b). Both the heterozygous and homozygous mutations yielded the yellow leaf blade phenotype, but the plants harboring homozygous mutations exhibited a more severe phenotype with a shorter habit.

Discussion

Mutations produced by cellular repairs induced by double-strand breaks generated by SpCAS9 are not random. The vast majority of the mutations introduced correspond to insertions of a single base (see, for instance (Lee et al. 2019), in maize) and, to a lesser extent,

Table 3 Deletions, insertions, deletions plus insertions (Ind/del) and substitutions generated using SpCAS9, LbCPF1 and BE_nCAS9D10A expressed as % of total mutant alleles identified. A total of 60, 30 and 22 mutants alleles were identified for SpCAS9, LbCPF1 and BE_nCAS9D10A respectively (see also Fig. 3)

	Deletions	Insertions	Ind/del	Substitutions
SpCAS9	27% (16)	52% (31)	21% (13)	0% (0)
LbCPF1	100% (30)	0% (0)	0% (0)	0% (0)
BE_CAS9	9% (2)	0% (0)	0% (0)	91% (20)

to short deletions generated by cut site microhomologies (Allen et al. 2018). Moreover, these insertions of a single base are not random because of the creation of 5 nt overhangs by SpCAS9, which are filled by polymerases and then religated (Lemos et al. 2018) so CAS9 induced in majority duplications of the fourth nucleotide in 5' of the PAM. Consequently, in our case, as predicted, the most frequent mutation observed corresponds to a duplication of the 4th nucleotide (A) upstream of the PAM (Lemos et al. 2018); this case was found in 16 of 68 cases, although we also detected single-nucleotide insertions of C, T, and G at lower frequencies (see Fig. 3a). It is therefore possible to predict to some extent the type of mutations introduced using a given guide for SpCAS9, although the full diversity of the mutations generated by SpCAS9 is sometimes difficult to predict and may include deletions, deletions and insertions, and insertions of other single nucleotides. SpCAS9 also frequently generates homozygous and biallelic mutations, which have a very high probability of generating KO mutations in the primary transformants, providing a rapid way to assess the function of a gene directly after transformation. However, this feature can be problematic for genes involved with regeneration, viability or fertility. GC composition affects the effectiveness of SpCAS9. It appears that the GC percentage of the area targeted by the sgRNA is correlated with the effectiveness of the SpCAS9 (Ma et al. 2015; Raitskin et al. 2019). Considering that most genes have higher GC contents in their 5' regions this suggest the SpCAS9 is more efficient to target 5' end genes.

LbCPF1 mainly generates heterozygous deletions. The cuts made by LbCPF1 have nucleotide overhangs, while the SpCAS9 cuts are blunt double-strand breaks. The overhangs are repaired more frequently by resection of the extremities and by MMEJ, inducing small deletions, which are a common result found when using LbCPF1 (Kim et al. 2016; Lee et al. 2019; Li et al. 2018; Tang et al. 2017; Xu et al. 2017). These small deletions sometimes lead to a significant frequency of in-frame deletions, allowing one or more nucleotides to be eliminated while maintaining the reading frame. This feature could theoretically allow us to study the effects of one or more amino acid deletions on protein function and represents a comparative advantage of LbCPF1 over SpCAS9. Generating mainly small deletions proved also an advantage for abolishing the recognition of promoter sequences by TALE effectors of bacterial blight thereby conferring partial or complete resistance to different strains (Oliva et al. 2019).

The high number of heterozygotes generated by LbCPF1, previously reported in the literature (Li et al. 2018; Xu et al. 2017), makes it possible to study the function of genes for which KO is lethal or interferes

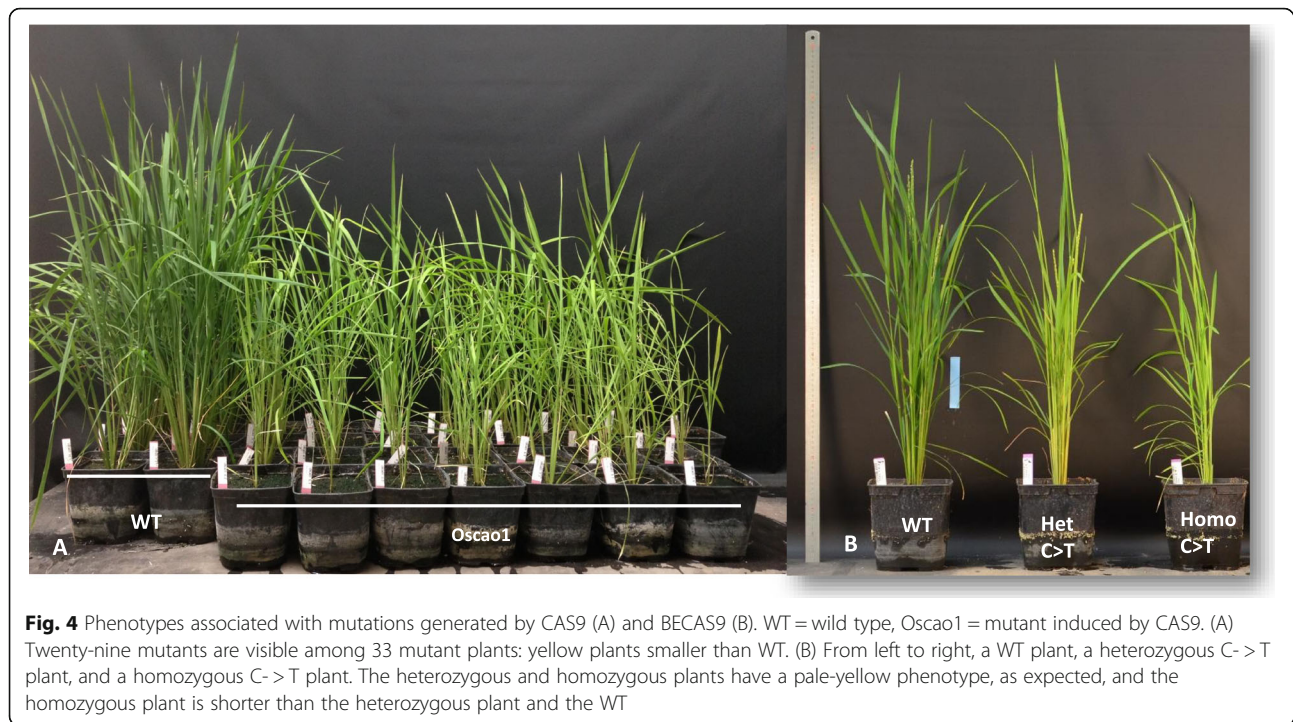


Fig. 4 Phenotypes associated with mutations generated by CAS9 (A) and BECAS9 (B). WT = wild type, Oscao1 = mutant induced by CAS9. (A) Twenty-nine mutants are visible among 33 mutant plants: yellow plants smaller than WT. (B) From left to right, a WT plant, a heterozygous C->T plant, and a homozygous C->T plant. The heterozygous and homozygous plants have a pale-yellow phenotype, as expected, and the homozygous plant is shorter than the heterozygous plant and the WT

with regeneration, while SpCAS9 mainly generates homozygous or biallelic mutants. Some LbCPF1 constructs appear to be as effective as SpCAS9 in rice and more effective than the construct we used (Tang et al. 2017). The reason is not clear and may depend on the guide used. This question would merit further comparative analysis of the different types of LbCPF1 constructs used in different publications.

BECAS9 has proven its effectiveness as part of the iSTOP strategy and is now a mature technology despite its limitations, particularly the C->T conversion rate, which is variable depending on the targeted region. For example, we obtained 55% conversion of the *OsNRT1.1B-C* in *OsNRT1.1B-T* alleles (data not shown), compared to a 22% conversion rate in *OsCAOI* (our results), and the average rates obtained by (Zong et al. 2017) were also approximately 43%. We also obtained significant C->G conversion rates of approximately 10% of the mutant lines, as found by (Lu and Zhu 2017), suggesting that the system is not perfect and generates a small number of C->G conversions that can also be used to create new alleles. Finally, as in (Lu and Zhu 2017), we found a low frequency of small deletions (10%) probably due to nickase CAS9 (D10A) activity, which nicks the unedited strand. Cytosine base editors, BECAS9 in our paper, induce genome-wide mutations in rice (Jin et al. 2019) and in mouse embryos (Zuo et al. 2019), even with a high-fidelity version of the cytosine base editor (Jin et al. 2019), and the off-targets cannot be predicted in silico. This is not the case for the

adenosine base editor (Jin et al. 2019; Zuo et al. 2019), and overall, these results suggest that using the adenine base editor (Yan et al. 2018) instead of cytosine base editor in the future may be more efficient, or alternatively, that the cytosine base editor should be optimized. Moreover, the inherent constraint of BECAS9, i.e., the C->T transversion, is mostly restricted to the -4 to -8 interval upstream of the PAM, limiting the accessible targets of the BECAS9 and the iSTOP strategy. Using a new BECAS9 with a new PAM and adenosine deaminase should help in the future to increase the usefulness of this approach (Hua et al. 2019).

Conclusion

The revolution in genome editing has provided the recent possibility to generate allelic series in any gene in plants, combined with an expansion of PAM specificity that will facilitate subtle and powerful functional analysis of any gene and any protein domain. For instance, allelic series were used to analyze the function of *OsIAA23* when simple KO lines did not have any visible phenotype (Jiang et al. 2019). Removing 13 amino acids, however, induced a severe phenotype. In this example, mosaic transcripts of *OsIAA23* were produced in frameshift mutants that did not represent “true” KOs (Jiang et al. 2019). A base editor could have been efficiently used to generate predictable KOs with premature termination of transcription. Alternatively, LbCPF1 could have been employed to generate longer deletion mutations. SpCAS9, LbCPF1 and BECAS9 have complementary characteristics; for instance, LbCPF1

can be used to delete one or several amino acids in frame, while BECAS9 can be used to introduce a frameshift mutation, to modify the phosphorylation status of a single amino acid, or to introduce a stop codon or any single amino acid change. With the development of CRISPR-CAS9 tools recognizing different PAMs (Hua et al. 2019), the range of nucleotides accessible to BECAS9 will also expand, giving rise to a large set of genome editing tools for deciphering gene functions in detail. Although off-target activity of SpCAS9 has been described in most eukaryotes and is highly variable depending on the sgRNA designed (O'Geen et al. 2015; Tsai and Joung 2016), it seems that off-target mutations are rare in plants and can almost always be predicted in silico (Li et al. 2019; Tang et al. 2018; Young et al. 2019), even if there is still debate regarding the level of off-target activity in plants (Zhang et al. 2018). Far fewer data are available concerning the off-target activity of LbCPF1, but the situation seems similar to that of SpCAS9 (Tang et al. 2018). Several complementary strategies have been developed, including the construction of high-fidelity SpCAS9, specialized deep learning software to help users reduce the probability of designing sgRNAs with high off-target potential, and technologies to predict or identify at a genome-wide level the frequencies of off-targets (O'Geen et al. 2015; Tsai and Joung 2016). Last, we think the *CAO1* gene is a marker gene complementary to the PDS gene for future study to compare GE efficiency in plants. i) it is not lethal, ii) the phenotype including heterozygotes is visible very early, iii) it is well conserved and can therefore be used in many plant species iv) It is also theoretically possible to estimate the efficacy of HDR with this gene, by restoring, for example, the framework for reading mutations in plants mutated in *CAO1*. We therefore propose its use as a marker gene to evaluate in the future the effectiveness and comparison of current and future publishing technologies.

Recently, a promising new technology derived from CRISPR/CAS9, prime-editing, has been developed. The system consists of a CAS9 nickase merged with a reverse transcriptase that is programmed by a prime editing guide (Anzalone et al. 2019). This system has the potential to introduce a very large range of variations, including deletions, insertions, base modifications and modifications combining all three at once. This system represents the next frontier of genome editing in plants and rice to be implemented.

Material and methods

CRISPR design

The targets in *OsCAO1* (LOC_Os10g41780) for LbCPF1, BECAS9 and SpCAS9 were defined with the CRISPOR-Tefor software (<http://crispor.tefor.net/>) by minimizing the number of possible off-targets (Additional file 1: Table S1). The crRNA for BECAS9 was designed to introduce a stop

codon by transversion of a C to T in position -4 to -8 of the PAM, the most favorable position for the use of cytosine deaminase, after analysis of the complete gene sequence on OryGenesDB (<http://orygenesdb.cirad.fr/>) (Komor et al. 2016; Zong et al. 2017).

Cloning spacers in pentry vectors

Cloning was carried out in entry vectors by digestion ligation using *Bsa* I sites. Briefly, the forward and reverse primers (Additional file 1: Table S1) corresponding to each spacer having sites complementary to the input vectors were ordered (Promega, USA) and then annealed: primers (100 μ M, 2.8 μ L each in 50 μ L final volume) were denatured at 95 °C for 5 min and then allowed to stand at room temperature for 1 h. Each entry vector (2 μ g) was digested with *Bsa* I restriction enzyme (40 U) in a 50 μ L final volume. Six microliters of the annealed primers was added for the ligation step together with the plasmid DNA of each entry vector (100 ng, 2.5 μ L), 10x ligase buffer (2 μ L), and T4 ligase (400 UI) in a final volume of 20 μ L. The ligation reaction was performed at 16 °C overnight. Then, 4 μ L of each ligation product was transformed into 50 μ L of chemocompetent DH5 alpha bacteria by heat shock. Positive clones were identified by *Bsa* I restriction (Additional file 1: Table S1).

Cloning by LR in binary destination vectors

crRNA (LbCPF1) under the control of the ZmUbi promoter was flanked by the hammerhead (HH) and hepatitis delta virus (HDV) ribozyme RNAs for precise crRNA processing (Tang et al. 2017). The sgRNAs (CAS9 and BECAS9) were under the control of the OsU3 promoter. Coding cassettes containing the guide RNAs were transferred to the destination binary vectors pUbi_LbCPF1-destvect4.0, pZmUbi-OsCas9-HPT and pBE_nCAS9-HPT_AttR-ccdB and transformed into chemocompetent DH5 alpha bacteria by heat shock. After verification by sequencing, the plasmids were transferred by electroporation to *Agrobacterium tumefaciens* strain EHA105. A final check by restriction was carried out on plasmids extracted from *A. tumefaciens* before genetic transformation.

Agrobacterium-mediated rice transformation

Genetic transformation was carried out in the Kitaake variety according to the protocol published elsewhere using mature seed embryo-derived secondary calluses (Sallaud et al. 2004). The selection was performed on a hygromycin medium.

Analysis of mutations

A total of 35 SpCAS9, 25 LbCPF1 and 44 BE_nCAS9D10A plants were regenerated and analyzed by

PCR sequencing using primers to amplify the areas targeted by LbCPF1, SpCAS9 and BECAS9 (Additional file 1: Table S1). The sequences obtained after cleaning were aligned with the reference sequence for each targeted locus. Mutations were analyzed manually with reference to the wild-type sequence.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12284-020-0366-y>.

Additional file 1: Table S1. Primers used for cloning, mutation detection and sequencing

Acknowledgments

This work was supported by several grants, including the Investissement d'Avenir program of the French National Agency of Research for the Project GENIUS (ANR-11-BTBR-0001_GENIUS), the CGIAR research program on RICE (CRP RICE) and the Labex AGRO ANR 10-LABX-0001 project GENERICE. Leo Herbert is partially supported by a Montpellier SupAgro PhD fellowship.

Authors' Contributions

LH and ACM built the Base editor, CPF1 and CAS9 constructions. FD, ACM and LH analyzed the mutations generated. Genetic transformations and monitoring of greenhouse plants were done by AV, MP, RM and CC. CP, EG and PT planned the experiments and wrote a first version of the paper. All authors reviewed and participated in the corrections and final version of the manuscript.

Funding

Not applicable.

Availability of Data and Materials

All the datasets, including videos and photographs, are included in the article and are also available from the corresponding author upon reasonable request. All the transgenic lines are also available through MTA from the corresponding author.

Ethics Approval and Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Competing Interests

The authors declare that they have no competing interests.

Received: 25 September 2019 Accepted: 14 January 2020

Published online: 28 January 2020

References

- Allen F, Crepaldi L, Alsinet C, Strong AJ, Kleshchevnikov V, De Angeli P, Palenikova P, Khodak A, Kiselev V, Kosicki M et al (2018) Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat Biotechnol*. <https://doi.org/10.1038/nbt.4317>
- Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, Chen PJ, Wilson C, Newby GA, Raguram A et al (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576(7785): 149–157
- Billon P, Bryant EE, Joseph SA, Nambiar TS, Hayward SB, Rothstein R, Ciccio A (2017) CRISPR-Mediated Base editing enables efficient disruption of eukaryotic genes through induction of STOP codons. *Mol Cell* 67(1068–1079):e1064
- Charrier A, Vergne E, Dousset N, Richer A, Petiteau A, Chevreau E (2019) Efficient targeted mutagenesis in apple and first time edition of pear using the CRISPR-Cas9 system. *Front Plant Sci* 10:40
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA et al (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823
- Es I, Gavahian M, Marti-Quijal FJ, Lorenzo JM, Mousavi Khaneghah A, Tsatsanis C, Kampranis SC, Barba FJ (2019) The application of the CRISPR-Cas9 genome editing machinery in food and agricultural science: current status, future perspectives, and associated challenges. *Biotechnol Adv* 37:410–421
- Hua K, Tao X, Zhu JK (2019) Expanding the base editing scope in rice by using Cas9 variants. *Plant Biotechnol J* 17:499–504
- Jiang M, Hu H, Kai J, Traw MB, Yang S, Zhang X (2019) Different knockout genotypes of OslAA23 in rice using CRISPR/Cas9 generating different phenotypes. *Plant Mol Biol* 100:467–479
- Jin S, Zong Y, Gao Q, Zhu Z, Wang Y, Qin P, Liang C, Wang D, Qiu JL, Zhang F et al (2019) Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science* 364:292–295
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–821
- Kim D, Kim J, Hur JK, Been KW, Yoon SH, Kim JS (2016) Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat Biotechnol* 34: 863–868
- Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533:420–424
- Lee K, Zhang Y, Kleinstiver BP, Guo JA, Aryee MJ, Miller J, Malzahn A, Zarecor S, Lawrence-Dill CJ, Joung JK et al (2019) Activities and specificities of CRISPR/Cas9 and Cas12a nucleases for targeted mutagenesis in maize. *Plant Biotechnol J* 17:362–372
- Lemmon ZH, Reem NT, Dalrymple J, Soyk S, Swartwood KE, Rodriguez-Leal D, Van Eck J, Lippman ZB (2018) Rapid improvement of domestication traits in an orphan crop by genome editing. *Nat Plants* 4:766–770
- Lemos BR, Kaplan AC, Bae JE, Ferrazzoli AE, Kuo J, Anand RP, Waterman DP, Haber JE (2018) CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. *Proc Natl Acad Sci U S A* 115:E2040–E2047
- Li J, Manghwar H, Sun L, Wang P, Wang G, Sheng H, Zhang J, Liu H, Qin L, Rui H et al (2019) Whole genome sequencing reveals rare off-target mutations and considerable inherent genetic or/and somaclonal variations in CRISPR/Cas9-edited cotton plants. *Plant Biotechnol J* 17:858–868
- Li S, Zhang X, Wang W, Guo X, Wu Z, Du W, Zhao Y, Xia L (2018) Expanding the scope of CRISPR/Cpf1-mediated genome editing in Rice. *Mol Plant* 11:995–998
- Lu Y, Zhu JK (2017) Precise editing of a Target Base in the Rice genome using a modified CRISPR/Cas9 system. *Mol Plant* 10:523–525
- Ma X, Zhang Q, Zhu Q, Liu W, Chen Y, Qiu R, Wang B, Yang Z, Li H, Lin Y et al (2015) A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol Plant* 8:1274–1284
- Mali P, Esvelt KM, Church GM (2013) Cas9 as a versatile tool for engineering biology. *Nat Methods* 10:957–963
- Miao J, Guo D, Zhang J, Huang Q, Qin G, Zhang X, Wan J, Gu H, Qu LJ (2013) Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res* 23:1233–1236
- O'Geen H, Yu AS, Segal DJ (2015) How specific is CRISPR/Cas9 really? *Curr Opin Chem Biol* 29:72–78
- Oliva R, Ji C, Atienza-Grande G, Huguet-Tapia JC, Perez-Quintero A, Li T, Eom JS, Li C, Nguyen H, Liu B et al (2019) Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat Biotechnol* 37:1344–1350
- Qin G, Gu H, Ma L, Peng Y, Deng XW, Chen Z, Qu LJ (2007) Disruption of phytoene desaturase gene results in albino and dwarf phenotypes in *Arabidopsis* by impairing chlorophyll, carotenoid, and gibberellin biosynthesis. *Cell Res* 17:471–482
- Raitskin O, Schudoma C, West A, Patron NJ (2019) Comparison of efficiency and specificity of CRISPR-associated (Cas) nucleases in plants: an expanded toolkit for precision genome engineering. *PLoS One* 14:e0211598
- Rodriguez-Leal D, Lemmon ZH, Man J, Bartlett ME, Lippman ZB (2017) Engineering quantitative trait variation for crop improvement by genome editing. *Cell* 171(470–480):e478
- Sallaud C, Gay C, Larmande P, Bes M, Piffanelli P, Piegu B, Droc G, Regad F, Bourgeois E, Meynard D et al (2004) High throughput T-DNA insertion mutagenesis in rice: a first step towards in silico reverse genetics. *Plant J* 39: 450–464
- Tang X, Liu G, Zhou J, Ren Q, You Q, Tian L, Xin X, Zhong Z, Liu B, Zheng X et al (2018) A large-scale whole-genome sequencing analysis reveals highly

- specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biol* 19:84
- Tang X, Lowder LG, Zhang T, Malzahn AA, Zheng X, Voytas DF, Zhong Z, Chen Y, Ren Q, Li Q et al (2017) A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nat Plants* 3:17018
- Tsai SQ, Joung JK (2016) Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat Rev Genet* 17:300–312
- Wilson FM, Harrison K, Armitage AD, Simkin AJ, Harrison RJ (2019) CRISPR/Cas9-mediated mutagenesis of phytoene desaturase in diploid and octoploid strawberry. *Plant Methods* 15:45
- Xu R, Qin R, Li H, Li D, Li L, Wei P, Yang J (2017) Generation of targeted mutant rice using a CRISPR-Cpf1 system. *Plant Biotechnol J* 15:713–717
- Yan F, Kuang Y, Ren B, Wang J, Zhang D, Lin H, Yang B, Zhou X, Zhou H (2018) Highly efficient A to G. C Base editing by Cas9n-guided tRNA adenosine Deaminase in Rice. *Mol Plant* 11:631–634
- Young J, Zastrow-Hayes G, Deschamps S, Svitashv S, Zaremba M, Acharya A, Paulraj S, Peterson-Burch B, Schwartz C, Djukanovic V et al (2019) CRISPR-Cas9 editing in maize: systematic evaluation of off-target activity and its relevance in crop improvement. *Sci Rep* 9:6729
- Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A et al (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163:759–771
- Zhang Q, Xing HL, Wang ZP, Zhang HY, Yang F, Wang XC, Chen QJ (2018) Potential high-frequency off-target mutagenesis induced by CRISPR/Cas9 in *Arabidopsis* and its prevention. *Plant Mol Biol* 96:445–456
- Zhou J, Xin X, He Y, Chen H, Li Q, Tang X, Zhong Z, Deng K, Zheng X, Akher SA et al (2019) Multiplex QTL editing of grain-related genes improves yield in elite rice varieties. *Plant Cell Rep* 38:475–485
- Zong Y, Wang Y, Li C, Zhang R, Chen K, Ran Y, Qiu JL, Wang D, Gao C (2017) Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nat Biotechnol* 35:438–440
- Zsogon A, Cermak T, Naves ER, Notini MM, Edel KH, Weinl S, Freschi L, Voytas DF, Kudla J, Peres LEP (2018) De novo domestication of wild tomato using genome editing. *Nat Biotechnol*. <https://doi.org/10.1038/nbt.4272>
- Zuo E, Sun Y, Wei W, Yuan T, Ying W, Sun H, Yuan L, Steinmetz LM, Li Y, Yang H (2019) Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* 364:289–292

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

En conclusion, le choix de l'endonucléase est un des paramètres important à prendre en compte en amont, pour analyser la fonction d'un gène par génome editing. Brièvement, la CAS9 reste la plus versatile dans les différents types de mutations (délétions, insertions) et la plus efficace, tandis que la LbCPF1 fait plutôt de courtes délétions ne changeant pas le cadre de lecture pour 43 % d'entre elles, et la BECAS9 fait très peu de délétions et comme attendu génère très majoritairement des substitutions, mais est la moins efficace (37 % seulement de plantes éditées). Les mutations sont pour la majorité bialléliques avec la CAS9 et la LbCPF1 (46 et 40 % respectivement), auquel s'ajoute un faible taux de mutations homozygotes (12 à 37 %). Pour finir, nous avons aussi montré que pour le riz *OsCAOI* était un excellent gène marqueur grâce à son phénotype rapidement visible et non létal, plus intéressant que la PDS, fréquemment utilisée chez les dicotylédones et les monocotylédones. L'approche que nous proposons permet de générer un jeu de mutations qui se manifestent pour la plupart par un phénotype mutant (plantes jaune pâle et de croissance limitée par rapport au témoin) et l'analyse moléculaire permet de préciser le type d'événements induit par chaque construit.

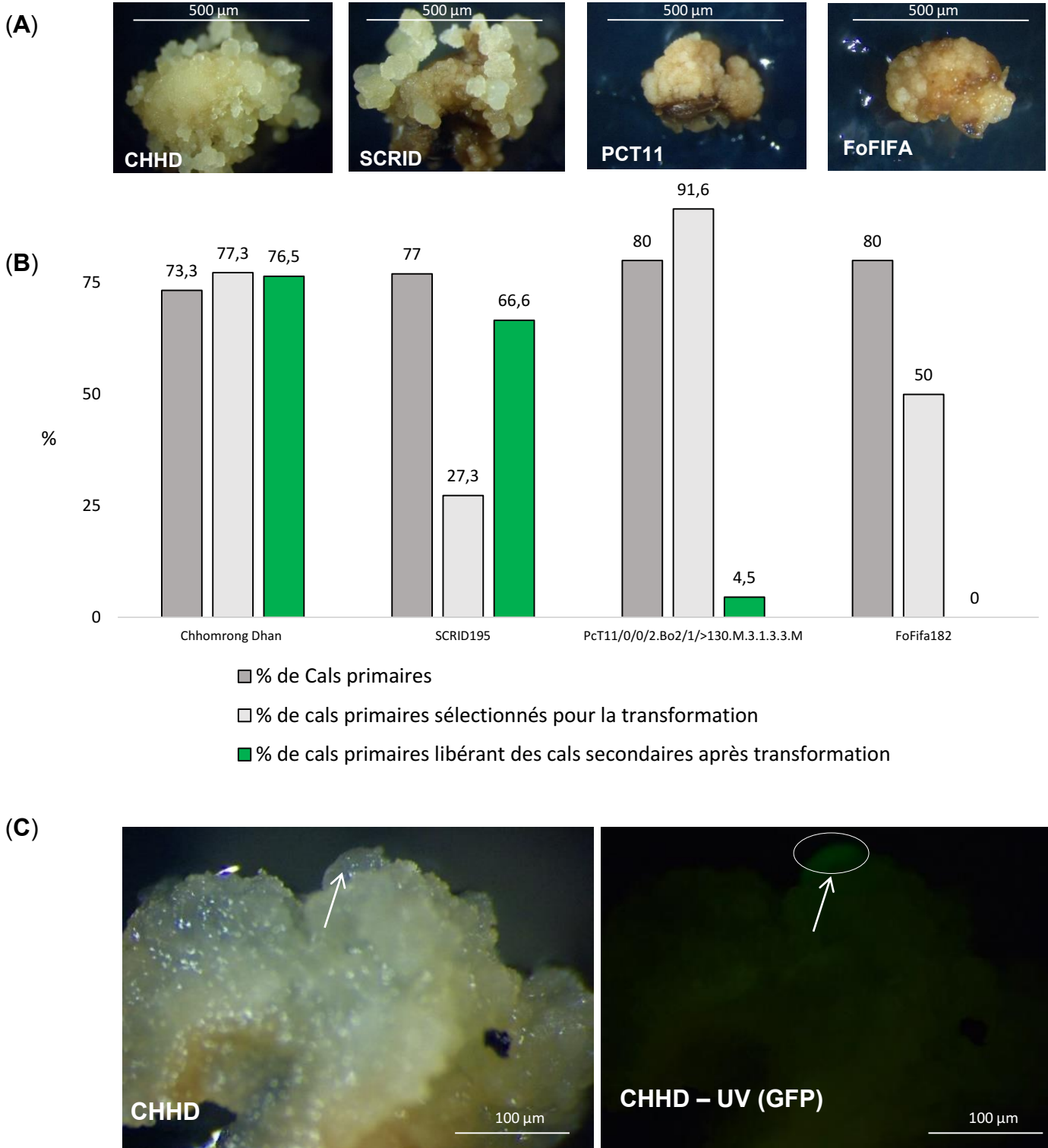


Figure 27 : Production de cals primaires et secondaires pour la transformation des lignées du programme Generice

Parmi les quatre géotypes sélectionnés CHHD (Chhomrong Dhan) et SCRID ont la capacité à produire des structures secondaires en forme de « nodules » blancs (A).

Les taux de production de cals primaires et secondaires sont estimés avant transformation. Globalement CHHD présente un meilleur taux de production de cals secondaires (B).

CHHD est le plus réactif à la transformation par *Agrobacterium* avec un plasmide GFP (C).

III] Amélioration des traits agronomiques de la NUE chez le riz à l'aide des technologies CRISPR : vers leur utilisation comme NBT

Afin d'étudier l'impact agronomique de l'édition à la base près du gène *OsNRT1.1b* et de la mutation KO sur le gène *OsBT* par CRISPR, nous avons analysé différents paramètres agrophysiologiques sur deux conditions de cultures enrichies en azote en chambre de culture, en serre de confinement, et au champ en Colombie.

Deux cultivars de la sous-espèce de riz *japonica* ont été sélectionnés pour les transformations stables et pour produire l'ensemble des mutants :

- Le premier est *Oryza sativa japonica* Kitaaké. C'est un cultivar simple à cultiver en serre, non sensible à la photopériode et aisément transformable au laboratoire par *Agrobacterium*. Ce sera notre plante de « référence » pour les expérimentations en serre, d'expression ARN, ou d'influx.
- Le second cultivar, Chhomrong dhan (CHHD), a été sélectionné dans le cadre du projet Agropolis « Generice ». C'est le cultivar de référence à Madagascar, exploité comme riz pluvial d'altitude (au-delà de 1200 mètres d'altitude) et utilisé en sélection, source majeure des variétés améliorées développées pour cet usage.

Nous avons sélectionné CHHD pour les raisons qui vont suivre. Quatre cultivars candidats pour la culture d'altitude avaient initialement été sélectionnés dans le projet (PcT11/0/0/2.Bo2/1/>130.M.3.1.3.3.M, FoFifa182, SCRID195 et CHHD) (**Figure 27-A**). CHHD a été très réactif *in-vitro* pour l'étape d'embryogenèse somatique. Il en a été de même, mais dans une moindre mesure pour la variété SCRID qui a fait beaucoup d'organogénèse racinaire en présence d'auxine (2,4-D à 2,5 mg/L), mais moins de cals primaires ont réagi à l'auxine pour produire des cals secondaires après transformation (**Figure 27-B et C**) (Meynard et al 2017 non publié).

Les variétés FoFIFA et PCT11 sont très sensibles aux *Agrobactéries* (même à une DO très faible, autour de 0.01). De plus il y a eu très peu, voire pas du tout de production de cals secondaires aptes à la transformation pour ces deux variétés. Les cals primaires sont brûlés après les trois jours de co-culture avec *Agrobacterium* à 25 °C (**Figure 27-A**). Enfin, nous avons reséquéncé les cibles dans ces différents génotypes, et avons écarté la variété PCT11 qui présentait déjà l'allèle *indica* au locus *NRT1.1b* (**Annexe 4**).

Nous avons sélectionné le cultivar Kitaaké pour les expérimentations en serre et en hydroponie car il est très simple à cultiver et à transformer. Pour ce faire, huit génotypes dans le fond Kitaaké ont été sélectionnés : 6 mutants issus de cals différents, et deux témoins : le témoin Kitaaké initial non transformé (WT_A) et une lignée issue de transformation avec un plasmide témoin (WT_B). Sept de ces huit génotypes sont donc issus de transformation *in-vitro* par *Agrobacterium*. Concernant les lignées éditées, 2 lignées ont été obtenues par KO du gène *Osbt* (frameshift), 2 lignées par KO du gène *Osnrt1.1b* (frameshift), et 2 par édition à la base près (C>T) du gène *OsNRT1.1b-T*. Le détail précis des mutations sur les gènes est présenté à la **Figure 28**.

Lignée	GÉNOTYPE - Code	Mutation	Précision
WT non transgénique	WT_A	Wt : CATCCGTACTAGCA CAAGATGTGCGAGGGTGTGC GGGCCCGCGACGCCGCCG Allèle 1 : CAAGATGTGCGAGGGTGTGC Allèle 2 : CAAGATGTGCGAGGGTGTGC	n/a , plante mère multipliée en serre
WT T237 4.1/4	WT_B	Wt : CATCCGTACTAGCA CAAGATGTGCGAGGGTGTGC GGGCCCGCGACGCCGCCG Allèle 1 : CAAGATGTGCGAGGGTGTGC Allèle 2 : CAAGATGTGCGAGGGTGTGC	n/a contrôle transformé
KO bt T236 1.1/9	KO_bt_C	Wt : CATCCGTACTAGCA CAAGATGTGCGAGGGTGTGC GGGCCCGCGACGCCGCCG Allèle 1 : CAAGATGTGCGAGGGT GTGC Allèle 2 : CAAGATGTGCGAGGGT GTGC	+1 G homozygote
KO bt T236 11.1/8	KO_bt_D	Wt : CATCCGTACTAGCA CAAGATGTGCGAGGGTGTGC GGGCCCGCGACGCCGCCG Allèle 1 : CAAGATGTGCGAGGGT GTGC Allèle 2 : CAAGATGTGCGAGGGT GTGC	+1 G homozygote
KO nrt T238 2.1/24	KO_nrt_E	Wt : cr1 : ACCATGCGCGATCATCGAC cr2 : CGCCAGATGACCACCTTCG (rev) cr 1 : ACCATGCGCGAT C GACAGG (-3) cr 2 : CGCC AGATGACCACCTTCG (-1)	2 CRISPR, - 3 et - 1 nucléotides homozygote
KO nrt T238 12.1/29	KO_nrt_F	Wt : cr1 : ACCATGCGCGATCATCGAC cr2 : CGCCAGATGACCACCTTCG (rev) cr 1 : ACCATGCGCGATCAT C AGAC (+1) cr 2 : ACGCG ATGACCACCTTCG (-3)	2 CRISPR, + 1 A et - 3 nucléotides homozygote
BE NRT T234 4.1/11	BE_NRT1.1_G	Wt : GCGGAGTCG CGGCGA CGGCGAGCAAGTGGAGGCT Allèle 1 : GCGGAG T GGCGAGCAAGTGG Allèle 2 : GCGGAG T GGCGAGCAAGTGG	Base Edition homozygote C en T
BE NRT T234 9.1/22	BE_NRT1.1_H	Wt : GCGGAGTCG CGGCGA CGGCGAGCAAGTGGAGGCT Allèle 1 : GCGGAG T GGCGAGCAAGTGG Allèle 2 : GCGGAG T GGCGAGCAAGTGG	Base Edition homozygote C en T

Figure 28 : Présentation des 8 génotypes sur fond Kitaaké utilisés dans l'étude sur la NUE: Code et génotypes
La séquence du protospacer WT est présentée en rouge. Pour les deux plantes sauvages le protospacer de leurs allèles non mutés est présenté en bleu.
Les mutations sont présentées sur les deux allèles, en rouge pour délétions et insertions, ou en vert pour les base éditées.
Cas particulier, cr1 et cr2 correspondent aux deux CRISPR du plasmide multiplexe.

Lignée	GENOTYPE-code	Mutation	Précision
n/a (WT CHHD)	WT_A1	Wt : CGTACTAGCA CAAGATGTGCGAGGGTGTGC GGGCC Allèle 1 : CAAGATGTGCGAGGGTGTGC Allèle 2 : CAAGATGTGCGAGGGTGTGC	n/a , plante mère multipliée en serre
T183 8.1/3A	WT_B	Wt : CGTACTAGCA CAAGATGTGCGAGGGTGTGC GGGCC Allèle 1 : CAAGATGTGCGAGGGTGTGC Allèle 2 : CAAGATGTGCGAGGGTGTGC	n/a contrôle transformé
KO bt T179 1.3/6F	KO_bt_C	Wt : CGTACTAGCA CAAGATGTGCGAGGGTGTGC GGGCC Allèle 1 : CAAGATGTGCGAGGGT TTGC Allèle 2 : CAAGATGTGCGAGGGT TTGC	+1 T homozygote
KO bt T179 13.1/11A	KO_bt_D	Wt : CGTACTAGCA CAAGATGTGCGAGGGTGTGC GGGCC Allèle 1 : CAAGATGTGCGAGGGT ATGC Allèle 2 : CAAGATGTGCGAGGGT ATGC	+1 A homozygote
KO nrt T182 7.1b/6D	KO_nrt_E	Wt : GCGGAGTCG CGGCGACGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCGACGGCGAGCAAG GTGG Allèle 2 : CGGCGACGGCGAGCAAG GTGG	+1 G homozygote
KO nrt T182 9.1/1C	KO_nrt_F	Wt : GCGGAGTCG CGGCGACGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCGACGGCGAGCAAG TGG Allèle 2 : CGGCGACGGCGAGCAAG TGG	+1 T homozygote
BE NRT T184 1.1a/2C	BE_NRT1.1_T1	Wt : GCGGAGTCG CGGCGA CGGCGAGCAAGTGGAGGCT Allèle 1 : CGGCGATGGCGAGCAAGTGG Allèle 2 : CGGCGATGGCGAGCAAGTGG	Base Edition homozygote C en T
BE NRT T184 1.1b/8B	BE_NRT1.1_T2	Wt : GCGGAGTCG CGGCGA CGGCGAGCAAGTGGAGGCT Allèle 1 : CGGCGATGGCGAGCAAGTGG Allèle 2 : CGGCGATGGCGAGCAAGTGG	Base Edition homozygote C en T
BE NRT T184 1.1a/1C	BE_NRT1.1_G1	Wt : GCGGAGTCG CGGCGA CGGCGAGCAAGTGGAGGCT Allèle 1 : CGGCGAGGGCGAGCAAGTGG Allèle 2 : CGGCGAGGGCGAGCAAGTGG	Base Edition homozygote C en G
BE NRT T184 1.1a/1E	BE_NRT1.1_G2	Wt : GCGGAGTCG CGGCGA CGGCGAGCAAGTGGAGGCT Allèle 1 : CGGCGAGGGCGAGCAAGTGG Allèle 2 : CGGCGAGGGCGAGCAAGTGG	Base Edition homozygote C en G

Figure 29 : Présentation des génotypes Chhomrong dhan sélectionnés: Code et génotypes
La séquence du protospacer WT est présentée en rouge. Pour les deux plantes sauvages le protospacer de leurs allèles non mutés est présenté en bleu.
Les mutations sont présentées sur les deux allèles, en rouge pour les insertions, ou en vert pour les base éditées.

Les géotypes Kitaaké utilisés ne contiennent plus le T-DNA. Les lignées T1 ont été préalablement semées puis les descendances portant les mutations à l'état homozygote mais ne contenant plus le T-DNA, ont été sélectionnées. Les analyses agronomiques et physiologiques ont été faites sur des plantes T2.

La variété CHHD est plus longue à cultiver en serre de confinement que la précédente, et arrive à maturité seulement après 6 mois de culture. Dix géotypes de la variété CHHD ont été sélectionnés pour être analysés au CIAT (Colombie) où la maturité arrive en 3 mois : huit mutants édités et deux témoins. Il s'agit du témoin CHHD initial non transformé et une lignée issue de transformation avec un plasmide témoin (**Figure 29**). Deux des lignées éditées ont été obtenues par KO du gène *Osbt* (frameshift), 2 lignées par KO du gène *Osnrt1.1b* (frameshift), et 4 par édition à la base près (C>T) du gène *OsNRT1.1b-T*. Deux des éditions à la base près correspondent bien à une mutation C en T, mais les deux autres correspondent à une mutation C en G par erreur du Base-Editing. Au niveau de la protéine au lieu de convertir la Thréonine (C) en Méthionine (T), il l'a convertie en Arginine (G). Cet acide aminé est chargé positivement, sans sulfure, et polaire comme la Thréonine. Nous verrons si cela a un effet physiologique sur la fonction du transecteur *OsNRT1.1b*.

1) Kitaaké : ARN, influx ¹⁵N, et Physiologie

a) ARN et influx ¹⁵N

Les cultures ont été réalisées en hydroponie sur de fortes concentrations d'azote, à 5 mM de KNO₃ ou de NH₄NO₃. Le nitrate induit l'expression de *OsNRT1.1b* (Chen and Ma, 2015) et nous pourrions ainsi comparer l'efficacité de prélèvement sur une solution où seul le NO₃⁻ est disponible, à une solution où le NH₄⁺ et le NO₃⁻ seront disponibles. Nous nous attendons à faire intervenir les différents acteurs de l'ammonium (AMTs) dans cette dernière (Xuan et al., 2017).

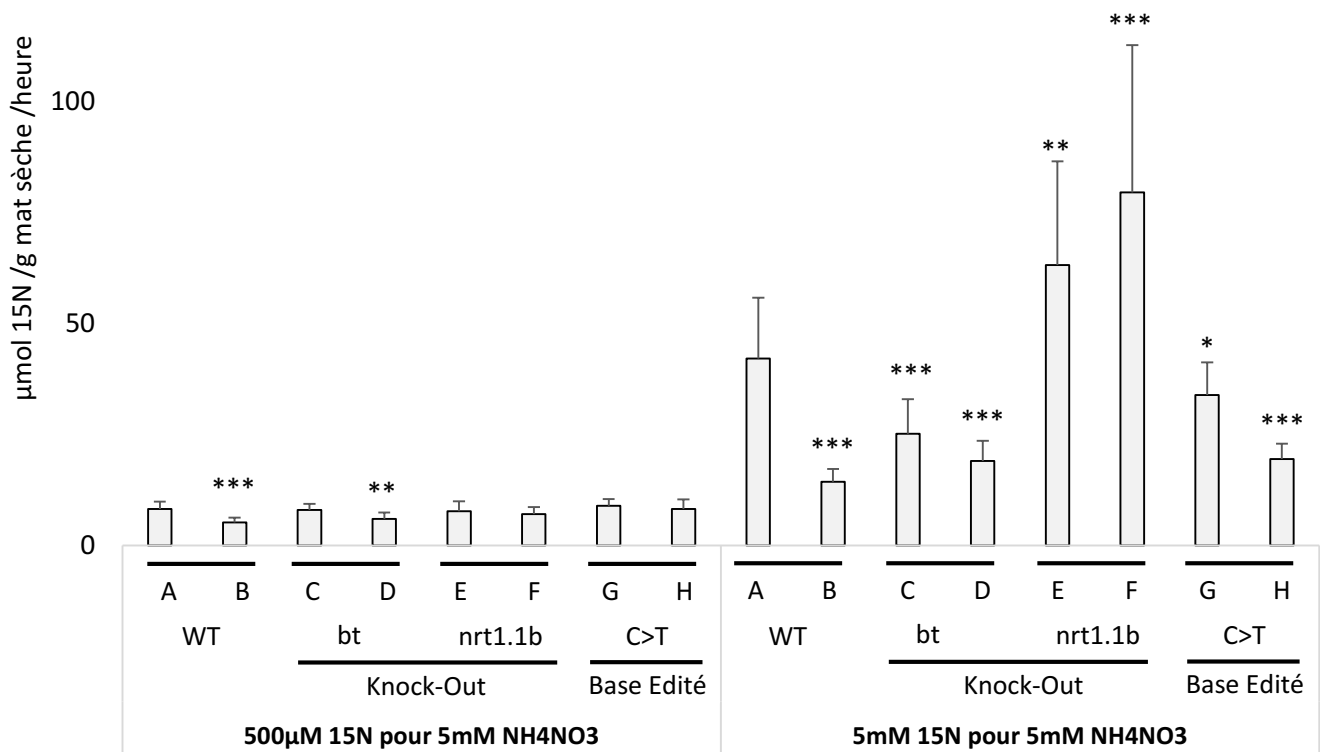
Après 1 mois de culture, les ARN totaux racinaires sont analysés sur 3 plantes par géotype et par condition.

En parallèle un influx est fait à 0.5 et 5 mM en utilisant un pulse d'azote lourd ¹⁵N-KNO₃ sur 12 racines par géotype et par condition pendant 30 jours. Pour rappel, les racines des 2 conditions de culture sont mises en contact 5 min avec une solution contenant soit 0.5 mM soit 5 mM de ¹⁵N-KNO₃. Cela permettra de voir la vitesse de prélèvement en fonction de la quantité d'azote disponible suivant le milieu par une analyse au spectromètre de masse à isotopes lourds.

Nous nous sommes basés sur la bibliographie pour définir cette expérimentation : Araus et collaborateurs (Araus et al., 2016) avait en effet également effectué un influx après culture sur NH₄NO₃. Hu et collaborateurs (Hu et al., 2015) avaient réalisé un test de pulse à l'aide de 5 mM de KNO₃ pendant 24h avec des cultures de seulement 10 jours et démontré l'intérêt de l'utilisation de ¹⁵N-KNO₃ versus ¹⁵N-NH₄Cl, dans la mesure où l'allèle de *indica* permet une meilleure accumulation de ¹⁵N-NO₃ par rapport à l'allèle *japonica*, mais ne montre pas de différence d'accumulation avec du ¹⁵N-NH₄Cl.

Les résultats préliminaires de cette expérimentation montrent que :

(A)



(B)

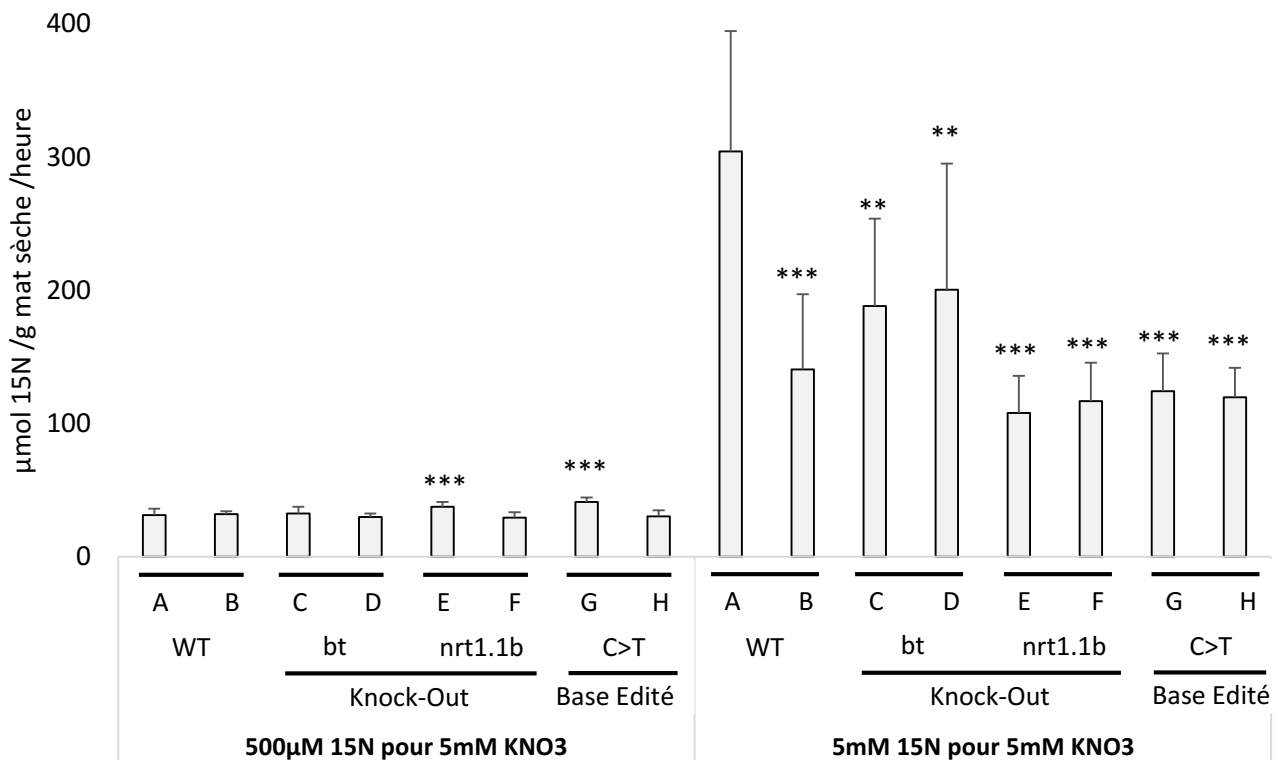


Figure 30 : Résultats d'influx pour les différentes lignées éditées

Influx des différentes lignées à $0,5 \text{ mM}$ et 5 mM de $\text{N}^{15}\text{KNO}_3$ cultivées sur 5 mM de NH_4NO_3 (A).

Influx des différentes lignées à $0,5 \text{ mM}$ et 5 mM de $\text{N}^{15}\text{KNO}_3$ cultivées sur 5 mM de KNO_3 (B).

Les données d'influx sont présentées en $\mu\text{mol } \text{N}^{15} / \text{g mat sèche} / \text{heure}$. Les codes correspondent au tableau en Figure 28. Les étoiles représentent les résultats des T tests démontrant une différence par rapport au témoin WTA (* < 0,05 ; ** < 0,01 ; *** < 0,0005)

Les deux témoins ne se comportent pas de la même manière sur de fortes concentrations en azote. Le WT non transformé (WT A (**Figure 28**)) présente un influx 2.5 à 3 fois plus élevé sur fortes concentrations que le contrôle, 'transformé', ie ayant subi la phase de culture *in-vitro* (**Figure 30-A et B**). Ce résultat surprenant est difficile à expliquer : 1. Soit le WT 'transformé', possède une ou plusieurs mutations qui affectent le transport d'azote, dans ce cas, il faudrait refaire l'expérimentation en comparant d'autres WT transformés avec d'autres vecteurs afin de vérifier si c'est un effet stochastique pour ce contrôle. 2. Soit la transformation elle-même a un effet sur le transport. La transformation génétique, en particulier affecte aléatoirement des locus en modifiant épigénétiquement d'autres locus (Kaeppeler et al., 2000). Encore une fois, il faudrait tester d'autres contrôles de transformation pour vérifier cette hypothèse. Dans la suite des analyses de cette première approche, nous nous référerons au contrôle non transformé (WT A) pour comparer l'effet des mutations sur le transport. Cette plante ne présente aucun biais de passage en transformation.

Le contrôle (WT A) a un influx quatre fois supérieur sur 5 mM KNO₃ par rapport au milieu 5 mM NH₄NO₃ avec 0.5 mM ¹⁵N. Même observation avec 5 mM ¹⁵N, qui lui est sept fois supérieur. Chez *A. thaliana*, l'ion NH₄⁺ est connu pour inhiber les transporteurs d'azote à haute affinité (HATS), et nos résultats suggèrent que c'est probablement le cas aussi chez le riz. Pour rappel OsNRT1.1b peut se transformer en HATS si il est phosphorylé lorsque le nitrate est en faible concentration (Chen and Ma, 2015; Hu et al., 2015). C'est un résultat intéressant car il semblerait bien qu'ici aussi le NH₄⁺ en forte quantité réprime les transporteurs à haute affinité, compliquant l'interprétation d'un effet du NO₃⁻ seul (Xuan et al., 2017). Par ailleurs, en conditions agronomiques au champ notamment, la source d'azote est souvent mixte (urée CH₄NO₂ complété avec du NH₄ et du NO₃ organique). Il conviendrait donc d'essayer de séparer les effets de ces deux sources d'azote pour une meilleure compréhension des mécanismes liés à l'absorption d'azote chez le riz, en testant séparément l'effet du NH₄⁺ et du NO₃⁻.

Sur NH₄NO₃, les deux KO *nrt1.1b* présentent des influx supérieurs (63 et 79 μmol ¹⁵N) par rapport au WT (42 μmol ¹⁵N) et aux autres mutations pour 5 mM de ¹⁵N. À contrario, ces deux KO *nrt1.1b* présentent des influx plus faibles (107 et 116 μmol ¹⁵N) sur KNO₃ par rapport au contrôle (304 μmol ¹⁵N). C'est un résultat étonnant qui semble suggérer que OsNRT1.1b a probablement une autre fonction, comme celle de réprimer d'autres transporteurs d'azote comme *OsNRT1.1a* par exemple (Wang et al., 2020), ou aussi que *NRT1.1b* a une fonction de « sensing » en collaboration avec les transporteurs à ammonium ou d'autres transporteurs comme *OsNRT1.1a*. Cette fonction ne serait pas visible en présence de KNO₃, mais uniquement en présence de NH₄NO₃ suggérant une interaction complexe entre NH₄⁺ et *NRT1.1b*.

Les deux KO bt sont globalement inférieurs (25 et 19 μmol ¹⁵N) au contrôle WT_A (42 μmol ¹⁵N) sous 5 mM NH₄NO₃ et sous 5 mM de KNO₃ (188 et 200 μmol ¹⁵N contre 304 chez WT_A) (**Figure 30-A et B**).

Les deux lignées éditées « à la base près » ne présentent jamais de meilleurs influx dans aucune des conditions mesurées (**Figure 30-A et B**). Seul le mutant Be NRT1.1b-T (échantillon G) présente un profil d'influx à 0.5 mM ¹⁵N légèrement plus important sur KNO₃ uniquement (41 μmol ¹⁵N). Auparavant il avait été montré que les lignées *japonica* disposant du gène *OsNRT1.1b_indica* sous promoteur 35S présentaient une meilleure accumulation du ¹⁵N en 24h (Hu et al., 2015).

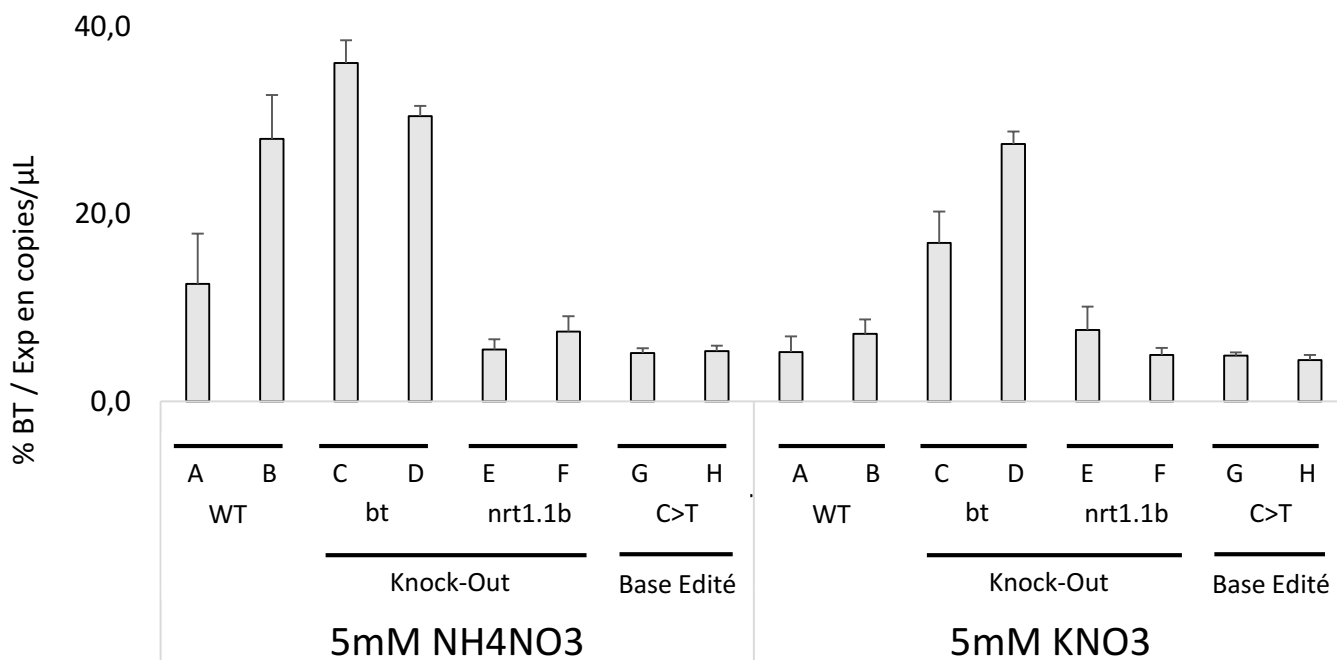


Figure 31 : Etude du niveau d'expression du gène *OsBT* dans les différentes lignées éditées dans les deux conditions de culture à 5 mM NH_4NO_3 et 5 mM KNO_3

L'expression de *OsBT* est normalisée par rapport à celle du gène de ménage *Exp* (EXP) sur. Les données sont représentées en % de copies/μL (n=3).

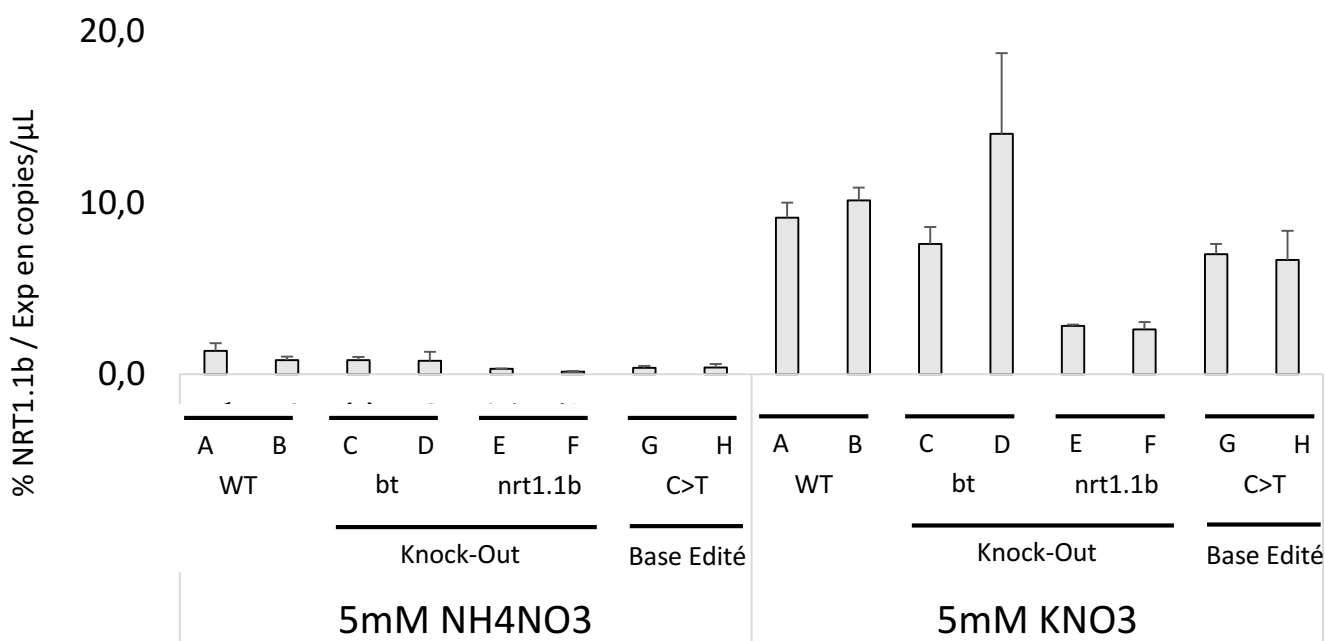


Figure 32 : Etude du niveau d'expression du gène *OsNRT1.1b* dans les différentes lignées éditées dans les deux conditions de culture à 5 mM NH_4NO_3 et 5 mM KNO_3

L'expression de *OsNRT1.1b* est normalisée par rapport à celle du gène de ménage *Exp* (EXP). Les données sont représentées en % de copies/μL (n=3).



Figure 33 : Alignement des acides aminés de la protéine *BT* WT et de la protéine mutée par KO CRISPR

L'alignement des séquences en acide aminés sur l'exon 1, ou les exons 1 et 2 après épissage, présente bien un décalage du cadre de lecture dû à l'ajout d'un nucléotide G dans les deux lignées CRISPR. La protéine ne semble plus fonctionnelle comme chez le WT.

Afin d'affiner nos résultats, nous avons mesuré les niveaux d'expression des gènes *OsNRT1.1b* et *OsBT* par rapport au gène (*EXP*), un contrôle d'expression constitutif, sur les deux milieux (KNO_3 et NH_4NO_3) pour tous les mutants analysés (**Figure 31 et 32**). Dans ces résultats préliminaires obtenus par ddPCR Evagreen, on peut observer que :

Le niveau d'expression du gène *OsBT* n'est pas modifié chez les mutants *NRT1.1b* par rapport aux WT, quel que soit le milieu, ce qui suggère que *OsNRT1.1b* ne régule pas l'expression du gène *OsBT*. En revanche, son niveau d'expression est plus fort chez les deux mutants KO bt pour les deux conditions, de trois fois plus en NH_4NO_3 à cinq fois plus en KNO_3 (**Figure 31**). Nous ne savons pas si cette surexpression est une conséquence des mutations « frameshifts » et si la protéine OsBT est encore fonctionnelle. L'alignement de la séquence en acide aminé par rapport au WT (**Figure 33**) suggère que le transcrite code pour une protéine tronquée, mais il reste possible que la protéine BT soit encore partiellement fonctionnelle. Une autre hypothèse peut être proposée, BT est rétro-régulé, c'est-à-dire que le niveau d'expression du gène *OsBT* est corrélé à la concentration de protéine BT fonctionnelle.

Le gène *OsNRT1.1b* est 10 fois moins exprimé sur NH_4NO_3 que sur KNO_3 , suggérant que le NH_4^+ inhibe son expression chez le riz comme celle d'autres transporteurs comme ceux de la famille des NRT2 (Xuan et al., 2017). Les mutants *Osnr1.1b* présentent un niveau d'expression plus faible que dans les lignées contrôles. Il semble n'y avoir en revanche aucun impact de l'édition « à la base près » sur son niveau d'expression, confirmant que la transversion C-T n'affecte pas le niveau d'expression du gène. C'est ce qui avait été observé entre les allèles *OsNRT1.1b* d'*indica* et de *japonica* (Chen and Ma, 2015) (**Figure 32**).

Difficile de conclure sur cette partie. En prenant en compte les résultats préliminaires dans leur globalité, ils semblent montrer que :

Pour *OsBT* : Dans les deux KO_bt (C et D), quelle que soit la condition le gène est plus exprimé malgré le fait qu'il soit muté et il qu'il y ait un influx plus faible. Nous pourrions nous demander si le niveau d'influx est inversement proportionnel à l'expression de BT. Cette hypothèse est mise à mal car dans les autres génotypes qui ont des niveaux d'expression similaires de *OsBT* entre eux, l'influx est inférieur pour les mutants édités à la base près (G et H) dans les deux conditions, et supérieur sur NH_4NO_3 pour les mutants KO_nrt (E et F). Il semblerait donc ne pas y avoir de corrélation entre l'expression de *OsBT* et l'influx.

Pour *OsNRT1.1b* : Il y a 9 fois plus d'expression sur KNO_3 , et 7 fois plus d'influx sur KNO_3 que sur NH_4NO_3 chez le contrôle (WT A). Nous pourrions nous demander si cette augmentation d'influx pourrait être liée à l'expression du gène. Sur KNO_3 cette hypothèse pourrait être possible car les mutants KO_nrt (E et F) présentent une réduction d'expression et d'influx. Sauf que ces KO_nrt (E et F) présentent une augmentation d'influx de $^{15}\text{N-KNO}_3$ sur NH_4NO_3 . Difficile de conclure sur une corrélation entre l'expression de *OsNRT1.1b* et l'influx sauf pour dire que cette corrélation n'est pas directe.

Pour aller plus loin dans notre réflexion, il faudrait réaliser un Western blot pour vérifier l'absence effective de protéines fonctionnelles car les niveaux des transcrits ne sont pas nuls ni chez *OsBT* ni chez *OsNRT1.1b*. De plus, il faudrait également réaliser un test d'influx « lent » en 24h pour éventuellement retrouver et confirmer une meilleure acquisition et accumulation du nitrate par les lignées éditées à la base près dans les mêmes conditions que Hu et al., (2015).

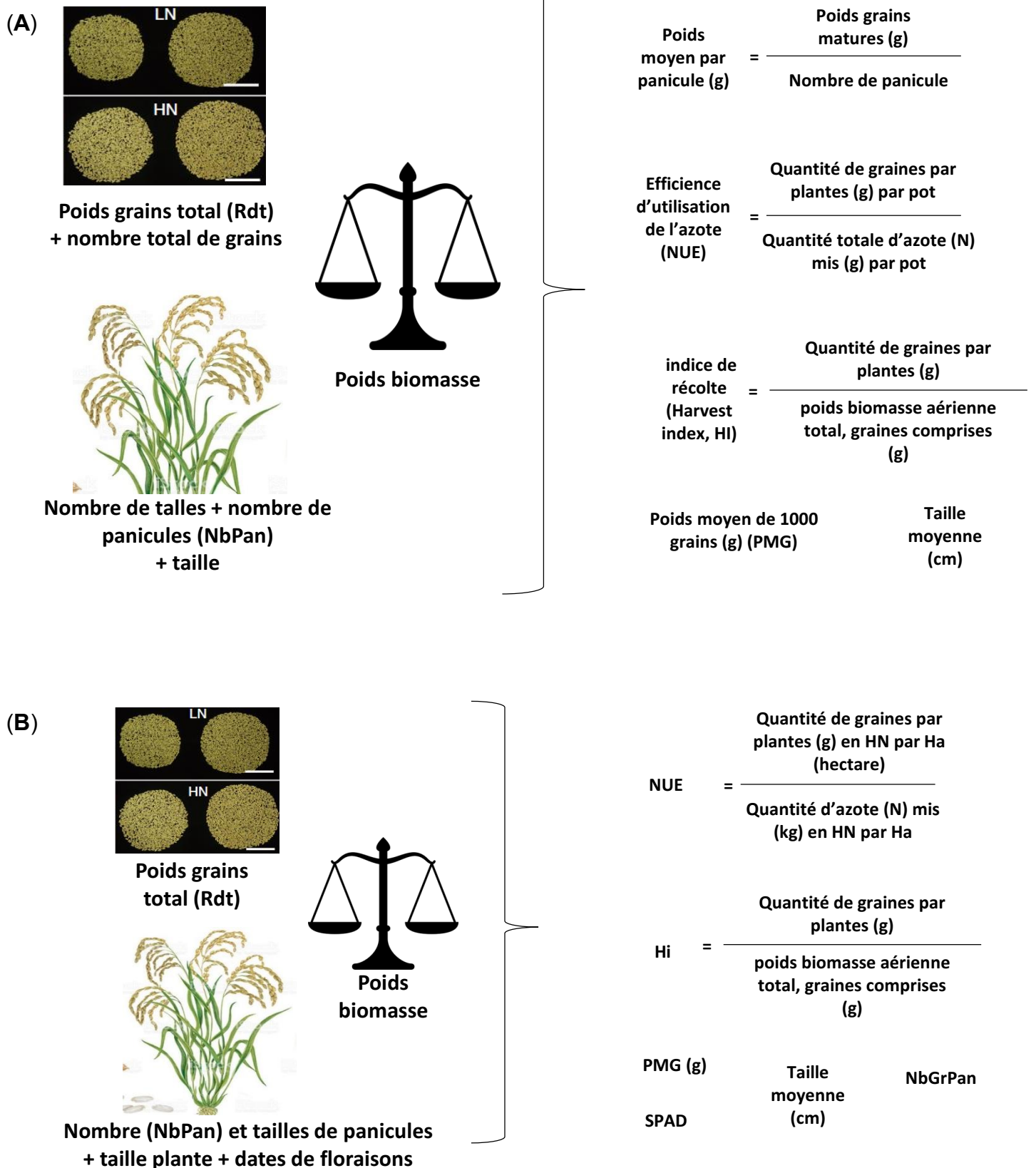


Figure 34 : Principaux paramètres mesurés sur Kitaaké en serre et Chhomrong dhan en Colombie

Les Kitaaké cultivés en semi hydroponie ont été phénotypés pour leur taille, leur biomasse, leur nombre de thalles et panicules, leur rendement, et leur nombre de grains. Il est ensuite possible de calculer le poids moyen de grains matures produits par panicule, la NUE, l'indice de récolte, et le poids moyen de 1000 grains. Ici n=15 plantes par condition et par génotype pour chaque expérimentation (A).

Les CHHD cultivés au champ seront phénotypés sur leurs taille, leur biomasse, le SPAD, leur nombre et taille de panicules, leur rendement, et le poids de 1000 grains. Il est ensuite possible de calculer la NUE, et l'indice de récolte (HI). Ici n=30 plantes par condition et par génotype (B).

Pour finir, les résultats obtenus pour les mutants KO *nrt1.1b* ont été différents des résultats attendus (*c.-à-d.* des influx plus faibles car la protéine est mutée). Cet effet peut-être dû à un fonctionnement encore partiel de la protéine OsNRT1.1b ou à un effet régulateur d'autres acteurs comme *OsNRT1.1a* (Wang et al., 2020).

b) Physiologie

Pour chaque expérimentation, 120 plantes de Kitaaké ont été cultivées sur faible (LN) ou fort (HN) nitrate à raison de 15 plantes analysées par géotypes en HN et en LN. Elles sont maintenues sur des tablar en serre de confinement dans 100 L de milieu semi-hydroponique Yoshida et Araus (**Annexe 5**). Seule la concentration en KNO₃ entre les milieux HN et LN, varie d'un facteur dix (0.5 mM LN vs 5 mM HN). Le milieu est changé toutes les semaines, pendant 12 semaines.

Nous ne présenterons dans cette partie que les résultats de NUE (utilisation de l'azote suivant le milieu par pot), d'Indice de Récolte (HI, Harvest Index) qui est souvent proche de 0.5 ce qui correspond à un poids en graines similaire au poids de la biomasse aérienne, et le rendement en graine par plante. Le reste des paramètres morphologiques mesurés (**Figure 34-A**) est présenté en annexe. Les deux contrôles, celui ayant subi la transformation et le contrôle non transformé, ne sont jamais significativement différents, et ce sur les deux expérimentations.

Deux répétitions ont été réalisées ; la première a été lancée en juin, et la seconde en août. Les deux conditions de culture sont différentes pour la durée du jour. L'analyse de variance nous a permis de montrer qu'il existe un effet majeur du traitement et du géotype sur l'expérimentation (**Annexe 6**). Malheureusement nous avons dû retirer l'expérimentation sur LN d'août, car il y a eu un problème sur tout un tablar (60 plantes sur 120) : les plantes de tous les géotypes étaient très en retard par rapport au second tablar LN et l'expérimentation de juin (**Annexe 7**). Les plantes ayant été randomisées, certains géotypes se trouvaient en bien plus grand nombre sur le tablar défaillant, jusqu'à 12 individus sur 15. Nous ne savons pas quelle est l'origine du problème, qui n'est visible que sur un seul tablar.

i) Plantes cultivées en HN (5 mM KNO₃)

En juin la NUE du WT_A est à 2.12 g_{grains}/g_{KNO₃} et tous les mutants KO et BE sur *OsNRT1.1b* présentent une décroissance de la NUE significative comprise entre 0.4 à 0.6 g_{grains}/g_{KNO₃} (soit 16 à 22 %). Chez les KO bt, seul le KO_bt_C a une NUE significativement inférieure au WT à 1.76 g_{grain}/g_{KNO₃} (**Figure 35**).

Les indices de récolte pour juin sont globalement tous identiques autour de 0.5 (**Figure 35**). Seul le HI du BE_NRT1.1_H est significativement inférieur avec une valeur de 0.48 (P < 0.05). En effet, sa biomasse moyenne est inférieure au WT_A (18.82 g contre 20.5 g) (**Annexe 8**) tout comme son rendement (8.9 g contre 10.7 g) (**Figure 35**) ce qui a influencé l'indice de récolte.

Les rendements en graines en juin sont globalement 10 fois supérieurs en HN. Tous les mutants sont significativement inférieurs au contrôle, avec une diminution de rendement comprise de 21 à 27 %, sauf pour le KO_bt_D (**Figure 35**).

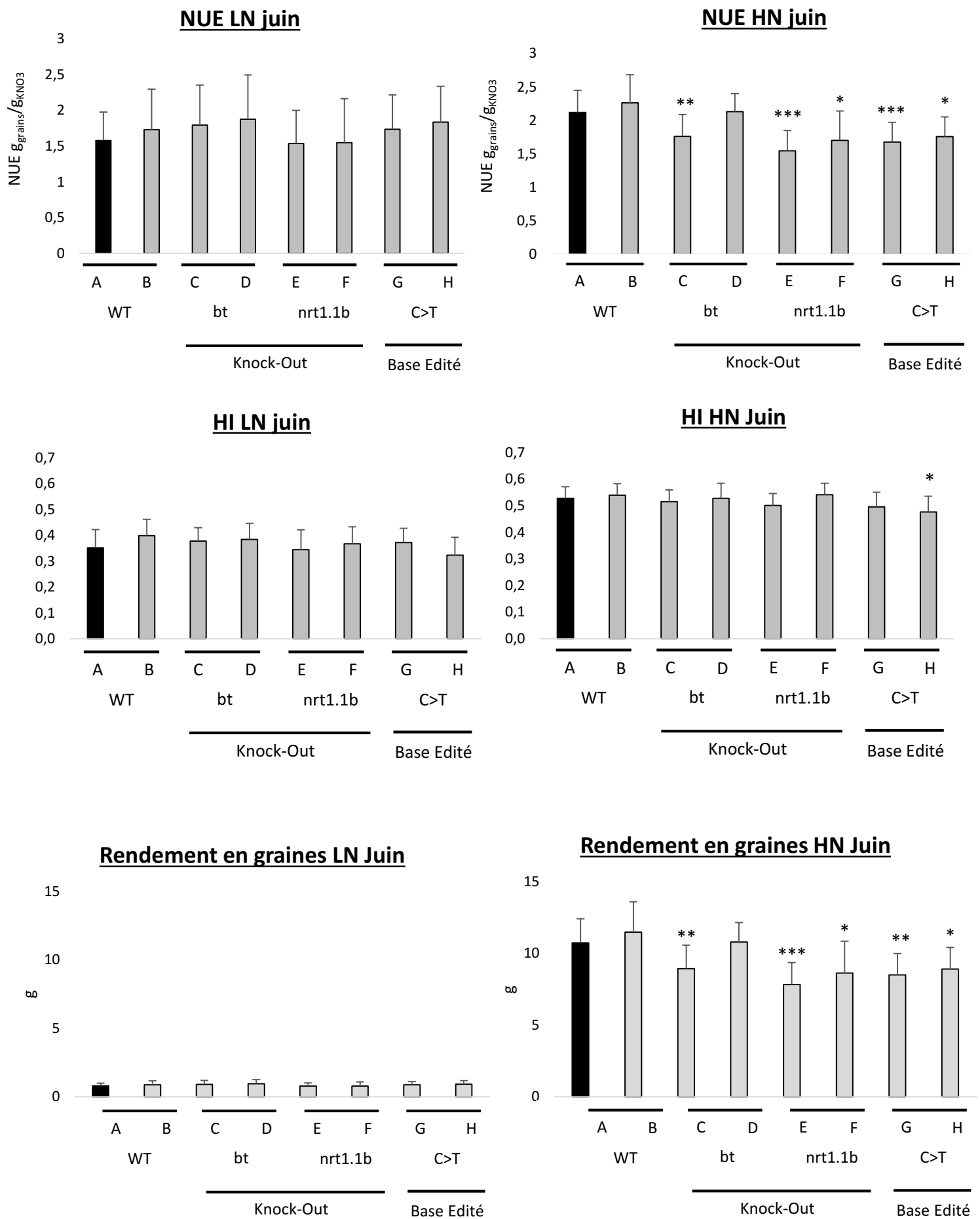


Figure 35 : Mesures de la NUE, du HI, et du Rendement en graines et chez Kitaaké réalisés en Juin en serre

Données de HI, Rendement, et NUE réalisées dans les deux conditions LN et HN en Juin.

Un T test est réalisé sur n=15 individus par génotype par rapport au WT_A (* < 0,05 ; ** < 0,01 ; *** < 0,0005)

En août, la NUE du WT_A est à 1.20 $\text{g}_{\text{rain}}/\text{g}_{\text{KNO}_3}$, soit une différence de 0.92 $\text{g}_{\text{rain}}/\text{g}_{\text{KNO}_3}$ par rapport à juin. Cette décroissance entre juin et août s'applique à tous les géotypes. Les géotypes BT_C, KO nrt (E et F) et BE NRT1.1b_H sont significativement inférieurs aux témoins WT (18 à 48 %) (**Annexe 30**).

Les indices de récolte pour août sont autour de 0.5 avec quelques variations. Le second contrôle WT_B (0.45 au lieu de 0.49 chez le WT_A), le KO_BT_C, le KO_nrt1_E, et le BE_NRT1.1_H ont significativement baissé par rapport à WT_A dans la mesure où leur rendement en grains a significativement baissé (**Annexe 30**).

Le rendement HN en août a diminué de 43 % pour le contrôle, 6.09 g contre 10.73 g en juin. Les mutants KO_nrt sont significativement inférieurs au contrôle, avec une diminution de rendement comprise de 33 à 48 %. Idem pour le KO_bt_C avec moins 28 % de rendement (**Annexe 30**).

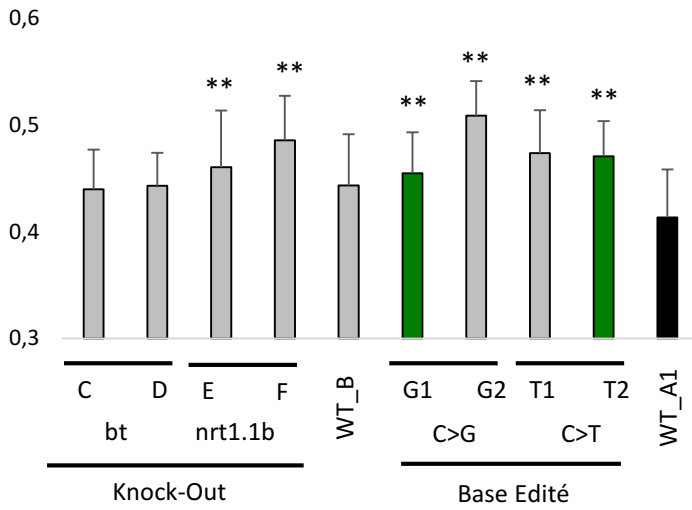
ii) Plantes cultivées en LN (0.5 mM KNO₃)

Suite au retrait de l'expérimentation d'août, nous n'allons observer que les résultats de juin (**Figure 35**). Dans un premier temps, les valeurs de NUE ne sont pas significativement différentes les unes des autres. De même le HI ne varie pas non plus significativement. Preuve que l'expérimentation a fonctionné pour l'apport de nitrate, le HI en LN est inférieur en moyenne de 0.2 unité pour tout le monde par rapport au HN (0.35 pour WT1 LN contre 0.53 en HN, effet traitement hautement significatif, **Annexe 6**). En LN, les rendements en graines sont tous autour de 0.8 et 0.9 g et ne sont pas significativement différents.

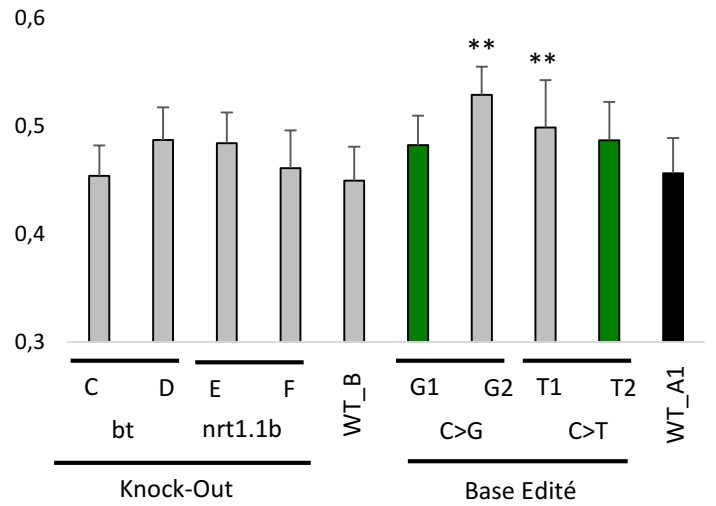
Pour conclure sur ces géotypes de Kitaaké cultivés sur KNO₃, il semblerait ressortir que dans des conditions de culture en serre de confinement sur cette source d'azote, il n'y a pas d'effet des mutations en condition de faible nitrate. Cependant les différents caractères observés diminuent en forte concentration de nitrate par rapport au contrôle dans les différents mutants. L'effet sur l'indice de récolte est moins net car il y a aussi une baisse de la biomasse aérienne. Il semblerait que les effets de la mutation Base Edité (-T) ne permettent pas une meilleure NUE sur KNO₃ et se rapprocheraient de la mutation KO du gène *OsNRT1.1b* en la réduisant de près de 20 %.

Plusieurs hypothèses peuvent être proposées : Le KNO₃ ne doit pas agir seul sur *OsNRT1.1b* ou *OsBT*, il faut ajouter du NH₄NO₃ pour voir leurs fonctions (Araus et al., 2016 ; Hu et al., 2015), mais cela implique l'activation des transporteurs d'ammonium. La rhizosphère pourrait aussi impacter sur l'apport d'azote, choses difficiles à estimer dans des conditions « confinées » (Wang et al., 2020 ; Zhang et al., 2019). Globalement, il est difficile d'expliquer que pour les lignées BE, leur NUE et leurs rendements soient inférieurs aux plantes contrôles. Les essais au champ (cf chapitre suivant), suggèrent que la source d'azote ainsi que les conditions de culture utilisées n'étaient probablement pas optimales.

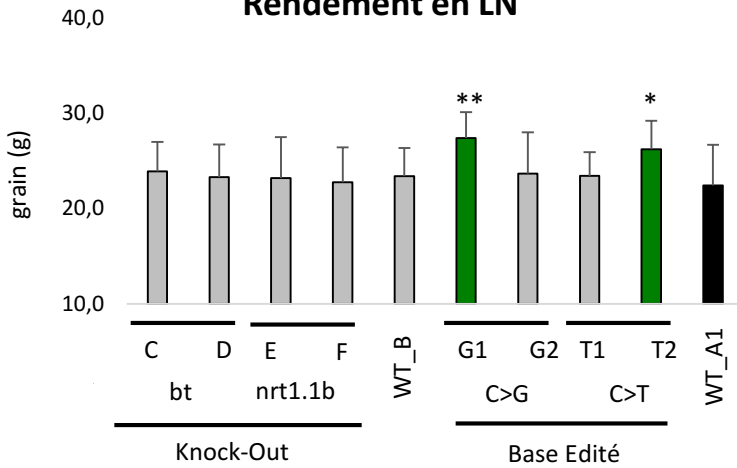
HI en LN



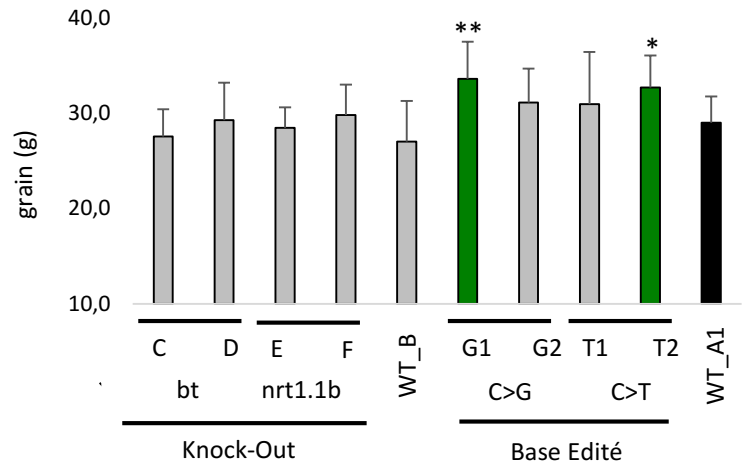
HI en HN



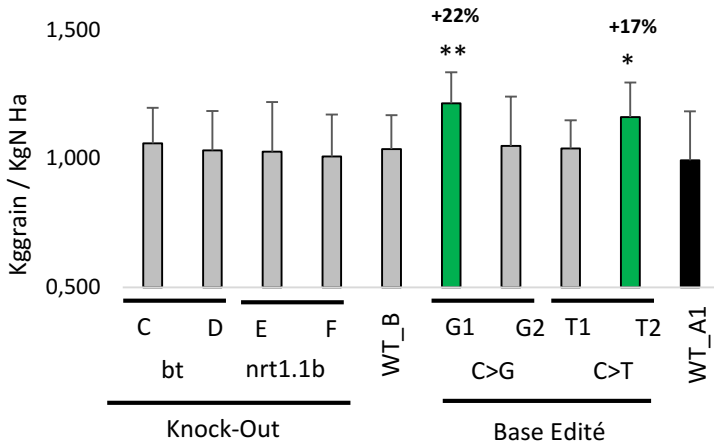
Rendement en LN



Rendement en HN



NUE évaluée en LN



NUE évaluée en HN

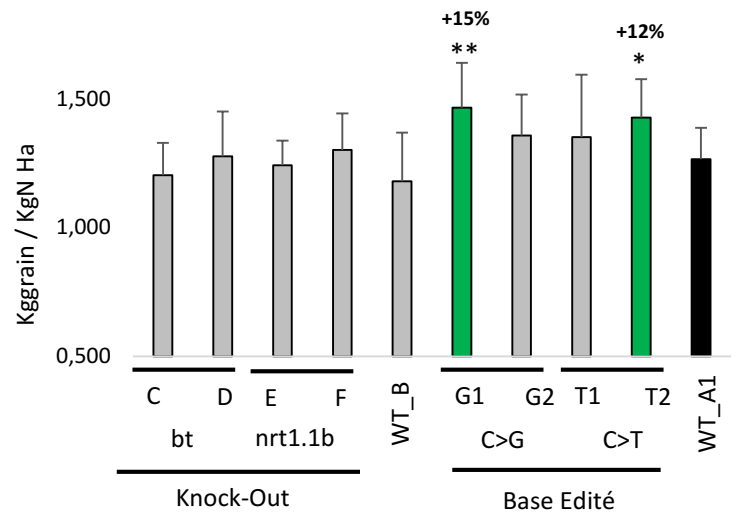


Figure 36 : Mesures du HI, du rendement, et de la NUE des lignées Chhomrong Dhan au champ en Colombie
 Représentation de l'HI, du rendement moyen par plante, et de la NUE par hectare des cultures des lignées CHHD en Colombie en LN et HN d'urée. Les barres vertes représentent les génotypes édité à la base près dans *OsNRT1.1b* qui sont positivement sortis du lot dans les différentes expérimentations.
 Significativité calculée avec des T-Test et test de Dunnett par rapport au WT A1 pur (n=30 plantes par condition par génotype ; * < 0,05 ; ** < 0,01 ; *** < 0,0005).

2) Chhomrong dhan (CHHD) : Physiologie au champ en Colombie

Le génotype CHHD est très long à cultiver en serre de confinement, et arrive à maturité seulement après 6 mois de culture. Dix génotypes de la variété CHHD (**Figure 29**) ont été sélectionnés pour être envoyés au CIAT Colombie où la maturité arrive en 3 mois. Nous avons rajouté un mutant supplémentaire, une des lignées testées dispose d'une mutation C en G par erreur du Base-Editing. En effet au lieu de convertir la Thréonine (C) en Méthionine (T), il l'a convertie en Arginine (G). Cet acide aminé est chargé, large comme la Méthionine, sans sulfure, et polaire comme la Thréonine. Nous verrons si cette arginine a un effet physiologique sur le transcepteur OsNRT1.1b. Pour rappel, ils ont été cultivés au champ sur 0 kg et 90 kg d'urée, à raison de 30 plantes analysées par génotype et par condition. Les détails sont disponibles dans le chapitre « Matériel et méthodes » pour la composition du sol (présenté en **Figure 60**).

En Colombie, en plus du nombre de panicules, du rendement, du poids de 1000 grains, de la taille, et du poids total de la matière sèche, des analyses de SPAD (teneur en chlorophylle), de temps de floraison, et de tailles de panicules ont été rajoutées (**Figure 34-B**). Comme pour les lignées Kitaakés, nous nous concentrerons sur le rendement, l'indice de récolte (HI), et la NUE. Le reste des résultats sera présenté en annexe sous forme de graphes. Nous pourrions nous y référer pour expliquer le résultat sur certains génotypes. Le WT A1 (Chhomrong Dhan non transformé) servira de comparaison pour les tests de Student ou de multiple comparaison (Dunnnett). Les deux témoins WT_A1 et WT_B ne sont pas significativement différents pour HI, rendement en grain et NUE. Le témoin de référence WT_A1 apparaît en noir sur les graphiques.

L'indice de récolte (HI) est le premier élément qui va permettre de contrôler si l'expérimentation a fonctionné ou non entre les deux milieux.

En LN, on observe que le WT non transformé présente un HI de 0.41 (**Figure 36**). Cela correspond à une biomasse sèche de la partie aérienne (hors grains) plus importante que le rendement en grains. Tous les mutants NRT1.1b sont significativement plus élevés : les mutants KO nrt E et F présentent un HI de 0.46 et 0.48, et les lignées éditées à la base près G et T (1 et 2) présentent des HI de 0.45 à 0.51 (**Figure 36**).

En HN, le WT non transformé présente un HI de 0.456. Seuls le G2 et le T1 sont significativement différents avec des valeurs de 0.53 et 0.5 (**Figure 36**). Cette significativité pour G2 et T1 est due au fait que les plantes ont une biomasse moins importante que le WT non transformé (**Annexe 9**) et un rendement légèrement supérieur aux WT bien que non significatif. Les KO bt C et D ne présentent pas de HI significativement différente en LN ni en HN. On peut en déduire que dans ces conditions de forte concentration en azote, la plante CHHD arrive à mieux utiliser l'azote lorsqu'il est encore plus présent pour l'allouer aux graines.

Le rendement en graines fait ressortir les deux génotypes édités à la base près G1 et T2 en LN et HN (**Figure 36**). En LN les génotypes Bases Edités G1, avec 27.41 g et T2 avec 26.21 g de rendement sont significativement supérieurs au WT non transformé avec 22.4 g. Sur HN, mêmes observations avec les Bases Edités G1, avec 33.6 g et T2 avec 32.72 g de rendement significativement supérieurs par rapport au WT non transformé avec 29 g.

Le calcul de la NUE par hectare (Ha) va dans le même sens (**Figure 36**). Attention cependant sur ces graphiques, l'écart-type est un pseudo-écart-type calculé manuellement en :

LN

HN

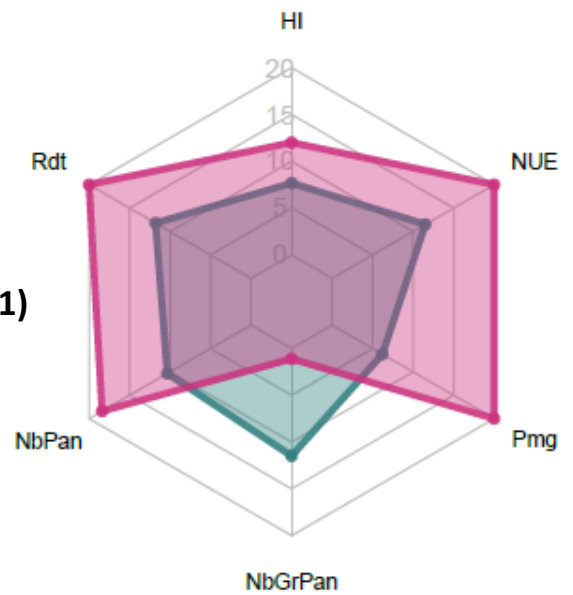
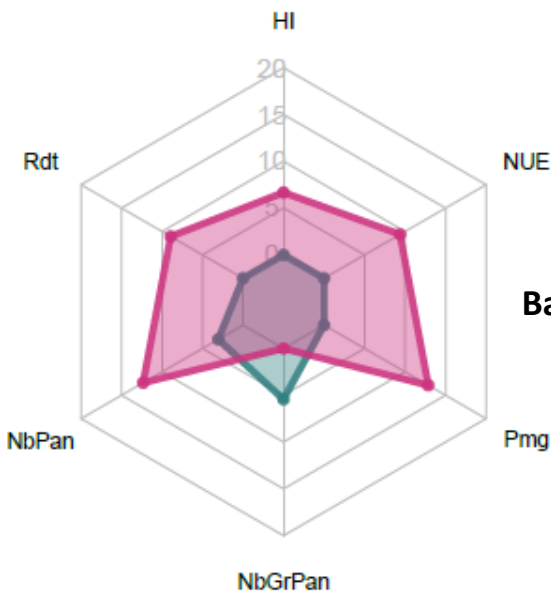
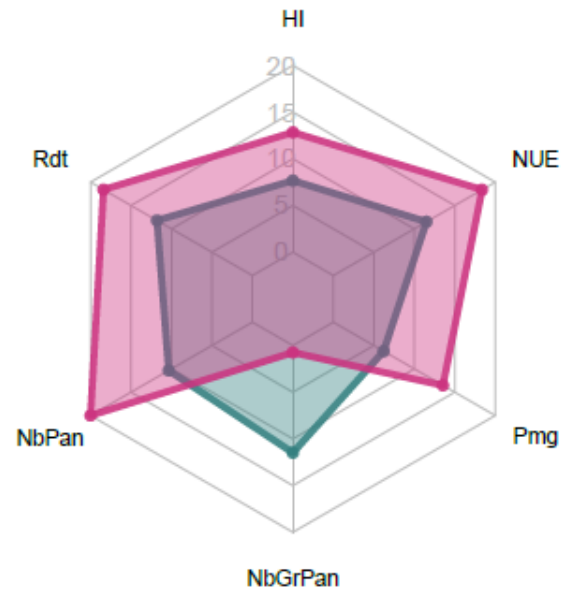
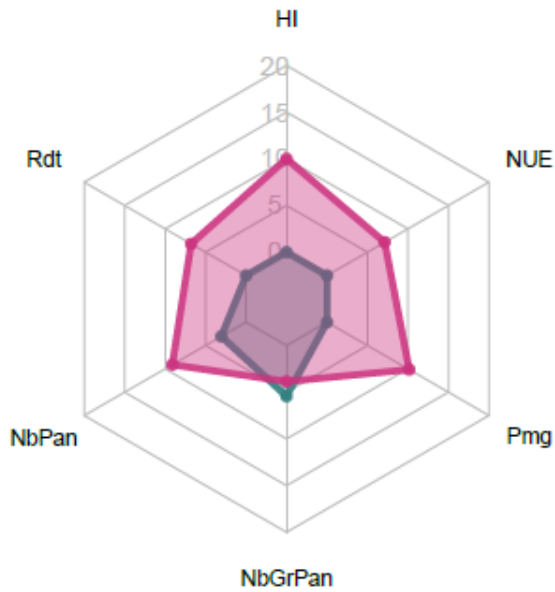


Figure 37 : Radar des lignées Base Edités de CHHD

Les Radar représentent les résultats globaux en LN (droite) et HN (gauche) :

- du rendement (Rdt)
- de l'indice de récolte (HI)
- de la NUE au champ (NUE)
- du poids de 1000 grains (Pmg)
- du nombre de grains moyen par panicule (NbGrPan)
- du nombre de panicules (NbPan).

Les plantes éditées à la base près sont en rose ; et la plante contrôle sauvage WT_A1 sert de référence en bleu à titre de comparaison.

soustrayant les 2 écarts-types (HN-LN) moyens divisés par l'apport d'engrais (90 kg ou 0 kg + quantité naturellement présente dans le sol dans 1 Ha). En LN, les Bases Edités G1, avec une production de 1.214 kg grains/kg urée Ha ; et T2 avec 1.161 kg grains/kg urée Ha, sont significativement supérieures au WT non transformé (0.993 kg grains/kg uréeHa). À noter que les Bases Edités G2 et le T1 présentent également une augmentation avec des valeurs de 1.049 et 1.038 kg grains/kg urée ; mais ne sont pas significativement différents du témoin, confirmant ce que nous avons dit pour le rendement.

En HN, à nouveau les Bases Edités G1, avec une production de 1.465 kg grains/kg urée Ha et T2 avec 1.427 kg grains/kg urée Ha, sont significativement supérieures au WT non transformé (1.265 kg grains/kg uréeHa). Les G2 et les T1 présentent également une augmentation à 1.357 et 1.351 kg grains/kg urée, mais ne sont pas significativement différentes du témoin (**Figure 36**).

Pour être encore plus précis, si nous considérons que le WT non transformé présente un rendement de 100 %, les Bases Edités G2 et T1 présentent une NUE plus élevée de 5.7 % et 4.6 % en LN, et de 7.4 % et 6.8 % en HN, mais non significatifs à chaque fois. Les plantes soeurs G1, T2 présentent quant à elles une NUE supérieure en LN de 22.3 %, et 17 % et de 15.8 % et 12.8 % en HN respectivement.

Si nous observons les autres caractères, nous pouvons constater que les deux génotypes édités à la base près G1, et T2 présentent significativement des panicules plus grandes et plus nombreuses (**Annexe 10**), mais qui portent en moyenne légèrement moins de grains par panicules que le témoin (entre 8 et 9 grains de moins que le WT) (**Annexe 11**). Cependant cet écart est rattrapé avec un poids de 1000 grains significativement supérieur au WT (**Annexe 11**). En moyenne, nous avons estimé le nombre de grains par plante en regardant le poids de 1000 grains moyen et le rendement moyen. Pour finir, les deux génotypes édités à la base près G1 et T2 présentent plus de grains par plante en LN et HN que le WT (**Annexe 11**). Les taux en chlorophylle donnés par le SPAD sont significativement supérieurs aussi pour les lignées G1 et T2 (**Annexe 12**). Le SPAD est un proxy de la teneur en azote dans les feuilles, suggérant que le taux d'azote végétatif est lui aussi, probablement plus élevé.

Nous pouvons représenter une grande partie de ces résultats des deux lignées éditées à la base près G1 et T2 sous la forme d'un Radar global en les comparant au contrôle non transformé (**Figure 37**). Pour la plupart des caractères (excepté le nombre de grains par panicule), les lignées éditées ont une meilleure performance que la lignée témoin. Leurs homologues G2 et T1 ont une même tendance (**Annexe 13**).

En conclusion globale sur cette partie de tests en Colombie, les génotypes édités à la base près en Chhomrong dhan dans *OsNRT1.1b* avec l'allèle de la variété *indica* (BE-T, Méthionine) ou le nouvel allèle (BE-G, Arginine) ont une NUE significativement supérieure, de plus de 12.8 % et 15.8 % au champ. La mutation KO bt n'apporte pas de meilleure NUE, et les mutations KO nrt1.1b n'ont pas présenté de diminution significative de la NUE en comparaison, ce qui suggère qu'il y aurait potentiellement une compensation par d'autres acteurs intervenant dans la régulation de l'absorption de nitrate chez CHHD, ou le riz en général (Wang et al., 2020).

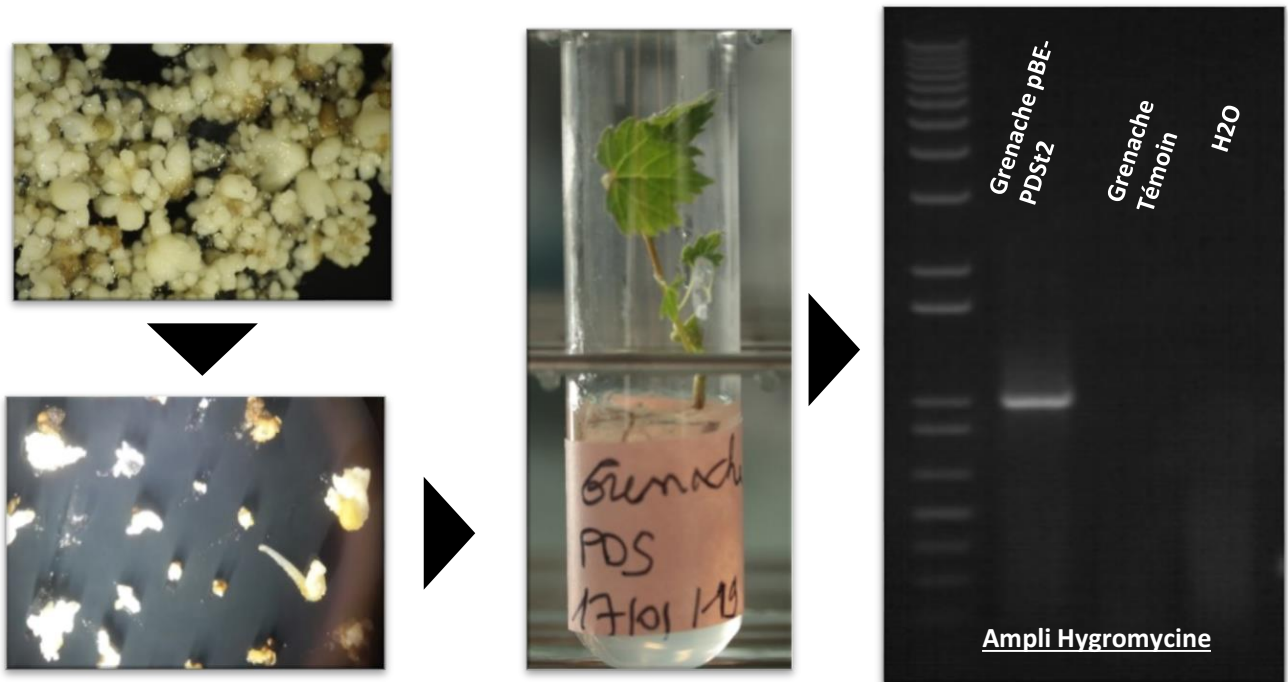


Figure 38 : Transformation et régénération chez la vigne (2018)

En suivant le protocole de Bouquet et al. (2006), à partir de cultures d'anthers, il a été possible de transformer stablement par *Agrobacterium* et régénérer des plants de vigne. La présence du T-DNA inséré est visible en amplifiant le gène de résistance à l'hygromycine

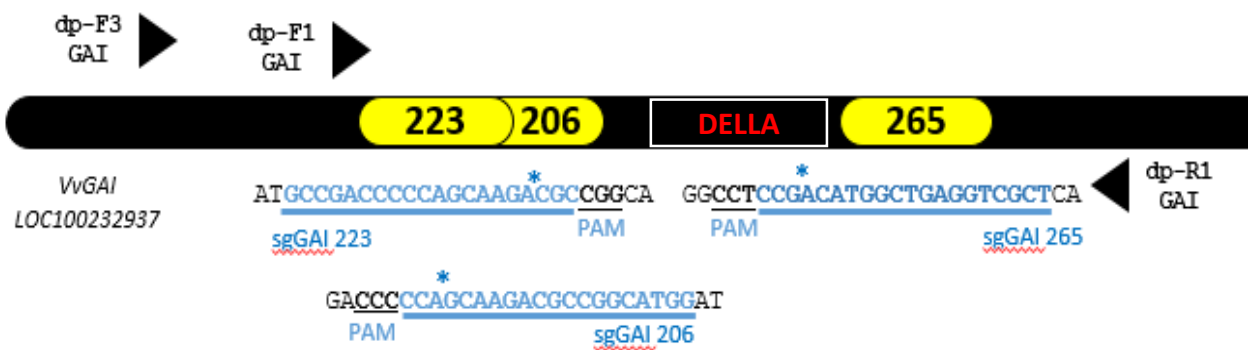


Figure 39 : Position des sites CRISPR simplex ou multiplexes sur VvGAI (vigne)

Schéma du gène *VvGAI* avec le positionnement des protospacers 223/206/265 pour les différentes expérimentations CRISPR. Leur séquence est présentée en bleu en 5'-3', le PAM est souligné. Les différentes amorces pour la ddPCR (dpF3/dpF1/dpR1) sont représentées par des flèches noires. Les étoiles bleues correspondent au site de coupe de la CAS9.

III] Amélioration de la Vigne via CRISPR : transformations stables et analyse des mutants potentiels pour le nanisme

1) Premier test avec les plasmides monocotylédones

Le développement du protocole et des expérimentations pour la transformation stable de la vigne a débuté en 2018 en parallèle du développement de la transformation transitoire pour gagner du temps. Nous avons finalisé les plasmides monocotylédones 2G9_GAI193 (32 pb en amont du domaine DELLA) (CAS9 riz et promoteur non codant du sgRNA U3 riz) et pBE_nCAS9_PDS-t2 (nCAS9 céréales, et promoteur non codant du sgRNA U3 riz). Le premier plasmide ciblait le gène *VvGAI*, et le second devait éditer à la base près le gène *VvPDS* pour voir si l'édition à la base près fonctionnait et avait un quelconque impact sur la couleur de la vigne. Le protocole utilisé a été celui de Bouquet (2006) sur du Grenache et du Portan. Au bout d'un an, la régénération a donné une centaine de plantes pour les deux cultivars (**Figure 38**). Une première analyse ADN s'est faite en amplifiant le gène de résistance à l'hygromycine des plasmides, présent uniquement chez les plantes transformées. Ensuite, chez ces plantes transformées, les gènes *VvPDS* et *VvGAI* ont été amplifiés puis séquencés. Malheureusement sur une centaine de plantes régénérées, aucune plante transformée ne présentait de mutations au locus respectif. Cet échec nous a encouragés à utiliser des plasmides CRISPR dicotylédones (pDE_CAS9 Arabidopsis), mais surtout à réaliser des tests d'efficacité par transformation transitoire avant transformation stable.

2) Second test : amélioration des plasmides et tests préalables par ddPCR

a) Constructions CRISPR Vigne

Comme première cible, nous sommes restés sur le gène *VvGAI* qui est à la fois très conservé entre génotypes et a priori facile à visualiser après régénération. Une seule cible a été construite dans un premier temps ; un CRISPR ciblant le gène conservé *VvGAI* : plasmide dénommé pDE_CAS9_cr265GAI (KO après le domaine DELLA) (**Annexes 1-B et C, 14, 15**).

Un construit multiplex, comprenant deux spacers (20 pb spécifiques du crRNA) a été réalisé pour déléter (retirer) entièrement le domaine DELLA. Il s'agit des plasmides CRISPR multiplex pDE_CAS9_cr265-223GAI (délétion de 40 pb) et pDE_CAS9_cr265-206GAI (délétion de 60 pb) (**Annexes 1-B et D, 14, 15**). Ces deux construits ont en commun le spacer cr265, utilisé également pour le construit simplex, mais ont deux spacers distincts pour la délétion (cr223 et cr206).

Les amorces et sondes utilisées pour le gène *VvGAI* sont présentées dans la **Figure 39**.

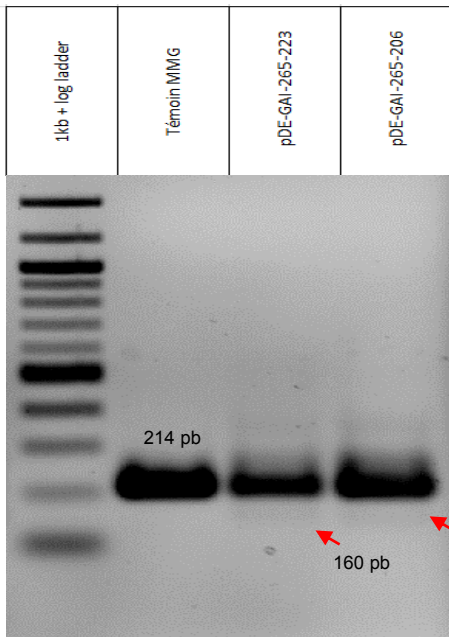


Figure 40 : Résultat du test délétion des CRISPR vigne

PCR emboîtée sur les protoplastes transformés avec les multiplexes CRISPR. Les flèches rouges représentent la taille attendue délété autour de 160 pb très peu visible (214 pb si WT).

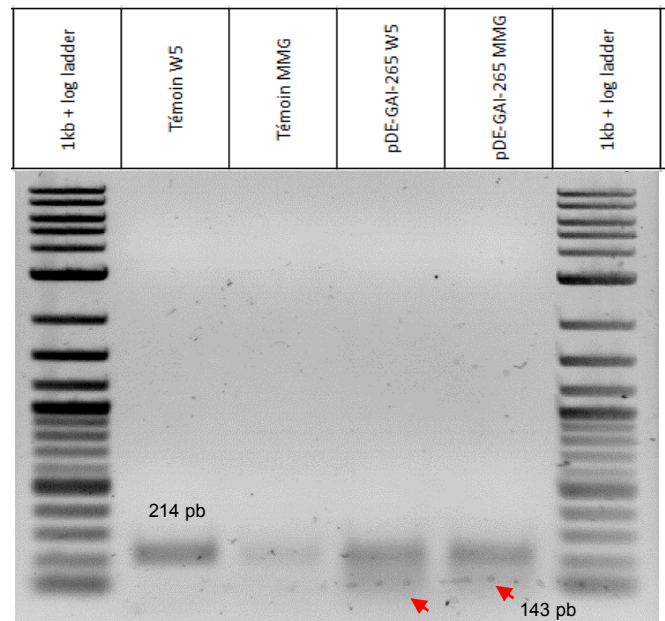


Figure 41 : Résultat de la T7E1 sur le CRISPR simplex vigne

Idem PCR emboîtée sur *VvGAI*, puis test T7E1 sur les protoplastes transformés avec le plasmide simplex dicotylédone avec 2h d'incubation dans le W5 ou dans le MMG avant transformation. Les flèches rouges représentent la taille attendue coupée par la T7E1 autour de 143 pb très peu visible (214 pb si WT, 143+71 pb si coupée).

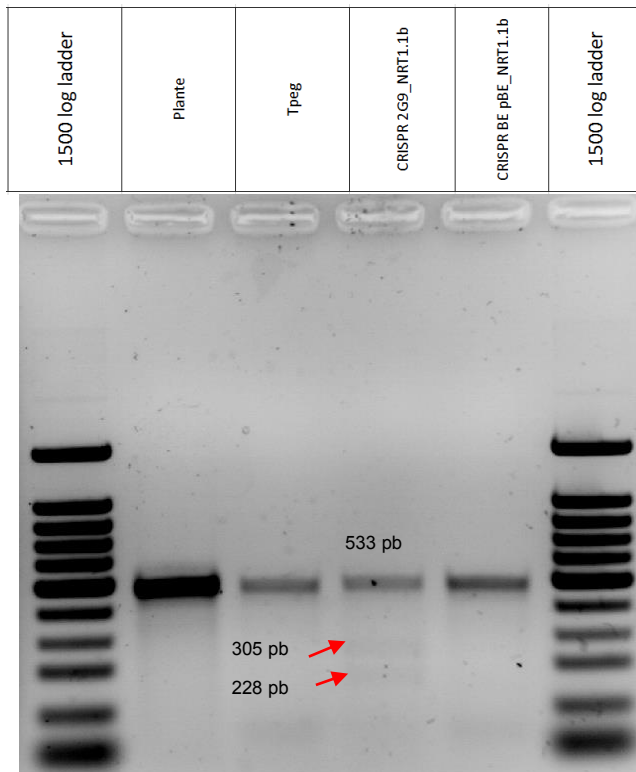


Figure 42 : Résultat de la T7E1 sur CRISPR simplex ou Base Editing en riz

PCR emboîtée sur *OsNRT1.1b*, puis test T7E1 sur les protoplastes transformés avec le plasmide simplex KO ou BE contre *OsNRT1.1b*.

Les flèches rouges représentent la taille attendue coupée par la T7E1 autour de 305 pb et 228pb peu visibles (533 pb si WT).

La détection des tailles attendues par le test T7E1 n'est visible que chez le CRISPR KO.

b) Détection des mutations en protoplastes par la méthode PCR ou T7EI

Pour commencer nous avons suivi un protocole similaire à celui décrit chez le riz dans (Bès et al., 2021). La détection des mutations se fait 48 h après transformation des protoplastes. Un traitement à la DNase de 30 minutes est réalisé sur les protoplastes transformés afin d'éliminer l'ADN provenant des protoplastes morts durant l'étape de purification. Une fois les protoplastes culotés, le surnageant est éliminé, et une extraction d'ADN est effectuée.

Une PCR de la zone théoriquement mutagénisée est réalisée. Deux PCR successives « emboîtées » sont nécessaires en raison de la faible quantité d'ADN extrait des protoplastes transformés. Si une délétion est induite par les multiplexes CRISPR, une PCR emboîtée est réalisée et la délétion est visible directement sur un gel à 2 % d'agarose comme une bande supplémentaire de plus faible poids moléculaire (WT à 200 pb, délétés entre 140 et 160 pb) (**Figure 40**). Pour les constructions CRISPR simplex, les produits PCR ont été purifiés, puis dénaturés pour produire des ADN simple brin. Ils sont ensuite rapidement réhybridés pour former des des homoduplexes et surtout des hétéroduplexes : ADN double brin WT/mutés. Ces hétéroduplexes sont reconnus et clivés spécifiquement par une endonucléase T7EI. Le mélange est déposé sur un gel d'agarose (Bès et al., 2021). Deux bandes de poids moléculaire plus faibles, dont la taille correspond au clivage au site muté sont alors visibles si le construit est fonctionnel. Cette méthode est illustrée dans la partie « Matériel et méthodes » (**Figure 59**). Cette méthode est qualitative, elle permet de savoir si le plasmide CRISPR est fonctionnel, et ce en moins de 10 jours avec une sensibilité autour de 20 %, c'est-à-dire qu'il faut un taux d'au moins 20 % de mutations pour que ce soit détectable.

Les deux gels présentés (**Figure 40 et 41**) ont été obtenus avec l'expérimentation qui a donné près de 72 % de taux de transformation, pour maximiser nos chances de détection. Pour les construits multiplexes, nous sommes à la limite du seuil de détection (**Figure 40**). L'absence de résultats à l'aide de la T7EI ne permet pas de conclure qu'un construit ne fonctionne pas et surtout cette technologie ne permet pas vraiment de comparer quantitativement l'efficacité de plusieurs construits.

Par exemple, sur le riz, la détection des substitutions du plasmide Base-Editor sur *OsNRT1.1b* (pBE_nCAS9-NRT1.1) n'a pas été possible en utilisant le test T7EI. Cela peut être dû à une efficacité trop faible d'édition ou alternativement la difficulté de couper au niveau d'un site unique. Néanmoins le crRNA NRT1.1b, identique à celui utilisé au BE, fonctionne bien puisque qu'il induit bien des mutations détectables par T7EI lorsqu'il est cloné dans un plasmide CRISPR CAS9 riz (2G9_NRT1.1b) (**Figure 42**). Alternativement, il se pourrait qu'il soit impossible de détecter beaucoup d'événements d'édition à la base près en protoplastes. Des étapes spécifiques de réparation, liées peut-être à la phase du cycle cellulaire, sont en jeu dans cette technologie, et les protoplastes dans nos conditions ne se divisent pas ou très peu d'entre eux le font.

En conclusion, même s'il est possible de détecter par PCR ou en utilisant la T7EI la fonctionnalité de certains de nos construits, nous avons cherché à utiliser une technologie alternative et pour comparer l'efficacité de nos constructions en utilisant des protoplastes de vigne.

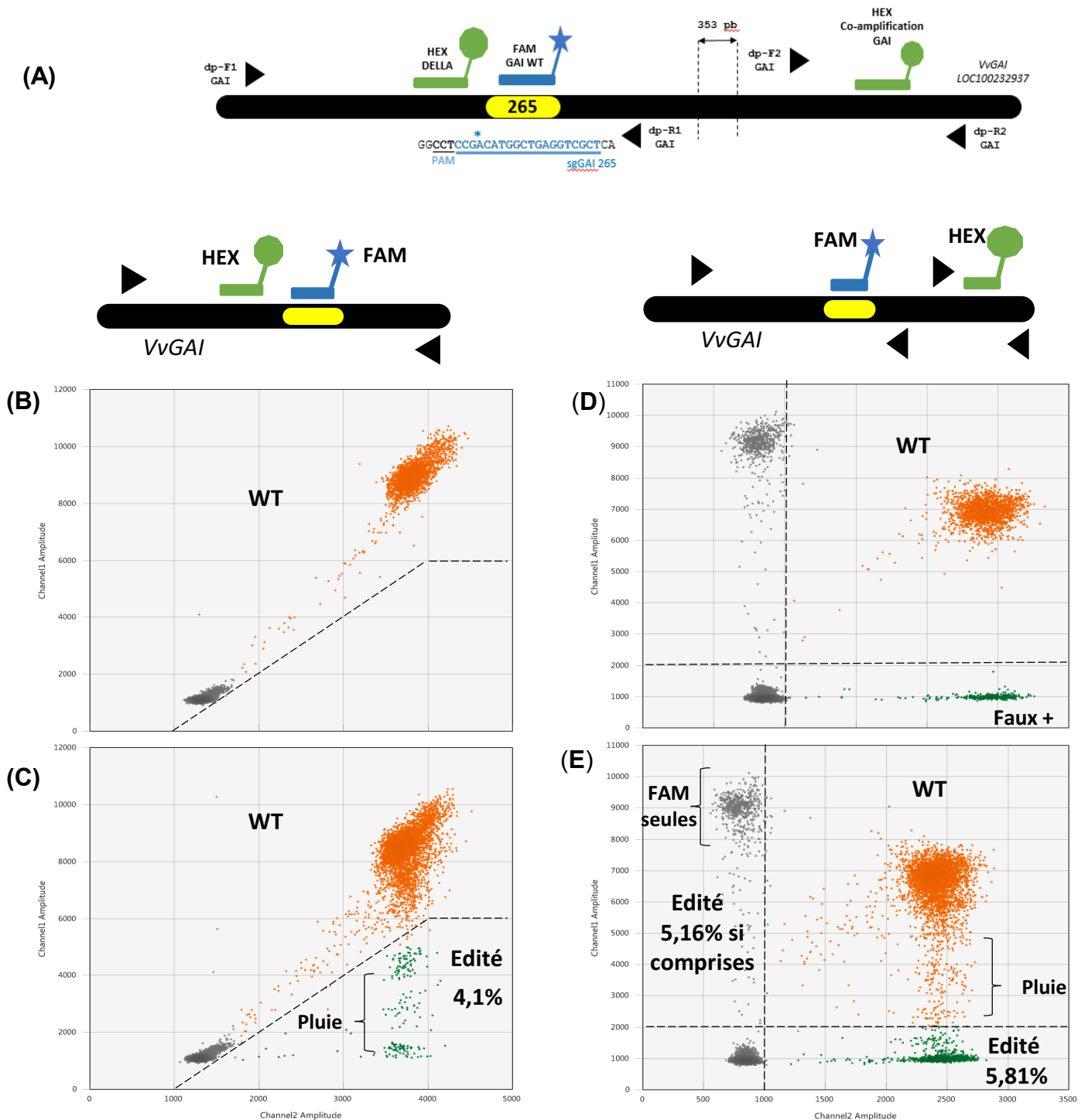


Figure 43 : ddPCR pour le CRISPR simplex sur *VvGAI*, par DropOff et Coamplification

Schema du gene *VvGAI* avec positionnement du protospacer 265 (jaune) pour le CRISPR simplex. La séquence est présentée en bleue en 5'-3', le PAM est souligné. L'étoile bleue correspond au site de coupure de la CAS9. Les différentes amorces pour la ddPCR (dpF1/dpR1 et dpF2/dpR2) sont représentés par les flèches noires. La sonde FAM est en bleu et se fixe sur le protospacer WT uniquement. Les deux sondes HEX sont en vert et servent de références interne (A).

L'expérimentation de DropOFF présente un seul nuage composé de gouttelettes doubles fluorescentes en orange pour les WT (B), et d'un nuage supplémentaire de gouttelettes vertes lors de l'ajout du plasmide CRISPR suite à la non fixation des sondes FAM au niveau du protospacer (C). Un taux de 4,1% d'édition est estimé par superposition 2D des deux représentations avec le calcul $[\text{HEX} / ((\text{FAM et HEX}) + \text{HEX})]$.

L'expérimentation de Co-amplification présente trois nuages composés de gouttelettes doubles fluorescentes (WT), vertes (faux positifs seuls) et grises (ou bleues si comptées) chez les protoplastes WT (D). Lors de l'ajout du plasmide CRISPR ces trois nuages réapparaissent mais en augmentant pour les vertes et faisant apparaître une pluie (E). Ici le biais d'amplification indique entre 5,16% et 5,81% d'ADN édités avec le calcul : $[\text{HEX} / ((\text{FAM et HEX}) + \text{HEX}) \text{ chez la condition CRISPR}] - [\text{HEX} / ((\text{FAM et HEX}) + \text{HEX}) \text{ chez le contrôle}]$.

c) Détection en ddPCR pour évaluer l'efficacité de mutation globale du construit cr265

La détection des CRISPR vignes a été réalisée par DropOFF et Coamplification. Nous avons utilisé des protoplastes transformés, avec un taux autour de 70 %, estimation obtenue à l'aide du plasmide contrôle GFP positif (pHKN29). Les amorces et sondes utilisées pour le gène *VvGAI* sont présentées dans la **Figure 43-A**.

i) DropOff pour évaluer l'efficacité de mutation globale du construit cr265

Pour rappel une seule amplification est faite sur le gène *VvGAI*. La sonde HEX sert de contrôle de la présence d'ADN et la sonde FAM reconnaît le protospacer WT (**Figure 43-A**). Dans les protoplastes non transformés (contrôle), les deux sondes se fixent bien sur la zone amplifiée. Elles laissent apparaître un cluster bien défini de gouttelettes oranges visibles sur le graphique 2D (**Figure 43-B**). En revanche, les protoplastes transformés avec la construction utilisant le cr265 (plasmide pDE_CAS9_265, avec AtCAS9 + AtU6-26 : sgRNAcr265) présentent deux clusters de gouttelettes : un orange représentant les ADN WT, et un vert représentant uniquement la fixation de la sonde HEX. Ces gouttelettes vertes représentent les ADN avec le protospacer muté. Au total, il y a un taux d'édition de 4.1 % dans ces protoplastes de vigne au locus *VvGAI* (**Figure 43-C**).

Un point important a été de savoir comment dissocier dans les échantillons possédant de la « pluie », entre les gouttelettes oranges (ADN WT) et vertes (ADN édité). Cette limite d'acceptation « WT » ou « Edité » a été obtenue en superposant les échantillons contrôles et ceux avec le plasmide CRISPR. Le seuil est donc directement lié aux échantillons contrôles. On peut observer une « pluie » de gouttelettes qui se détachent du nuage orange (gouttelettes positives [FAM et HEX]) (**Figure 43-C**). Cette pluie est interprétée comme différentes mutations (insertions, délétions...) qui affectent plus ou moins fortement la fixation de la sonde FAM en DropOff. Cette pluie n'apparaît que lors de l'ajout du plasmide pDE_CAS9_cr265GAI, et pas avec d'autres plasmides tel que le contrôle GFP+ (pHKN29) (**Annexe 16**) confirmant que cette « pluie » correspond probablement à des mutations variables au site ciblé. Plusieurs autres possibilités sont envisageables : des gouttelettes ayant incorporé deux ADNs WT et muté (Findlay et al., 2016), des gouttelettes « endommagées » ou avec des problèmes d'amplification (Witte et al., 2016). Pour certaines publications elles sont considérées comme « WT » (Berman et al., 2015) et mettent en évidence un problème d'amplification (Collier et al., 2017). Dans le cas du DropOff après plusieurs expérimentations nous avons observé que cette « pluie » apparaît uniquement lors de l'ajout du plasmide CRISPR (**Annexe 16** et **Figure 43-C**). Nous avons considéré de ce fait qu'elle représentait les mutants une fois le seuil défini par le contrôle. Afin de maximiser la fixation des sondes pour éviter des artefacts dans la « pluie », un gradient de 57 °C à 61 °C a été testé sur du Merlot (plante WT), et les meilleurs résultats (pas d'autohydrolyse, pas de pluie chez le WT) ont été obtenus entre 60 °C et 61 °C.

Pour tenter d'éliminer cette pluie chez les échantillons édités, il aurait été tenté d'augmenter encore plus le T_m jusqu'à 64-65 °C. Cependant, cela aurait eu pour effet de baisser l'intensité de fluorescence de la sonde FAM et de rapprocher les gouttelettes positives des négatives (manuel de référence BioRad, partie « PCR Optimization using thermal gradients »).

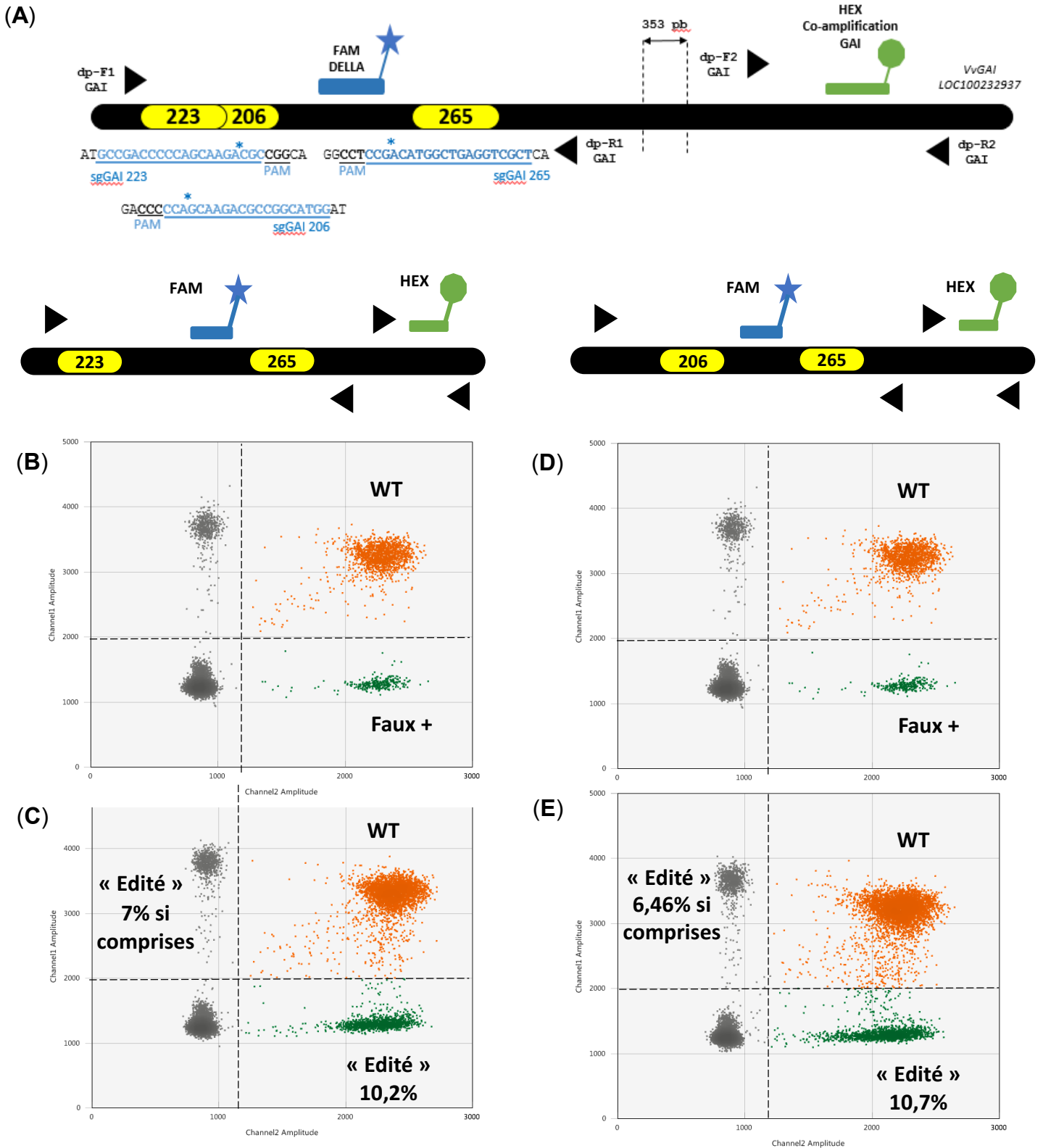


Figure 44 : Design de la ddPCR et du CRISPR Multiplex sur *VvGAI*

Schema du gène *VvGAI* (rappels) avec positionnement des protospacers 265/223/206 (jaune) pour les CRISPR. Leur séquence est présentée en bleu en 5'-3', le PAM est souligné. Les étoiles bleues correspondent aux sites de coupure de la CAS9. Les différentes amorces pour la ddPCR (dpF1/dpR1 et dpF2/dpR2) sont représentées par les flèches noires. La sonde FAM est en bleu et se fixe sur le domaine DELLA. Les sondes Hex sont en vert et servent de références **(A)**.

Les expérimentations de Co-amplifications présentent trois nuages composés de gouttelettes doubles fluorescentes (WT), vertes (faux positifs seuls) et grises (ou bleues si comptées) chez les protoplastes WT **(B)** et **(D)**. Lors de l'ajout du plasmide CRISPR ces trois nuages réapparaissent mais en augmentant pour les vertes et faisant apparaître une pluie **(C)** et **(E)**. Ici le biais d'amplification indique entre 7% à 10,2% (265-223) et entre 6,46% à 10,7% (265-206) d'ADN potentiellement édité avec le calcul :

$$[\text{HEX} / ((\text{FAM et HEX}) + \text{HEX})] \text{ chez la condition CRISPR} - [\text{HEX} / ((\text{FAM et HEX}) + \text{HEX})] \text{ chez le contrôle}.$$

Nous avons donc choisi de maximiser la différence d'amplitude de fluorescence entre les gouttelettes positives et négatives en sélectionnant le meilleur rapport Tm/séparation à 60 °C (selon la méthode préconisée par le fabricant).

Néanmoins, ce résultat démontre que le construit cr265 est capable d'induire des mutations au locus *VvGAI* avec un taux significatif.

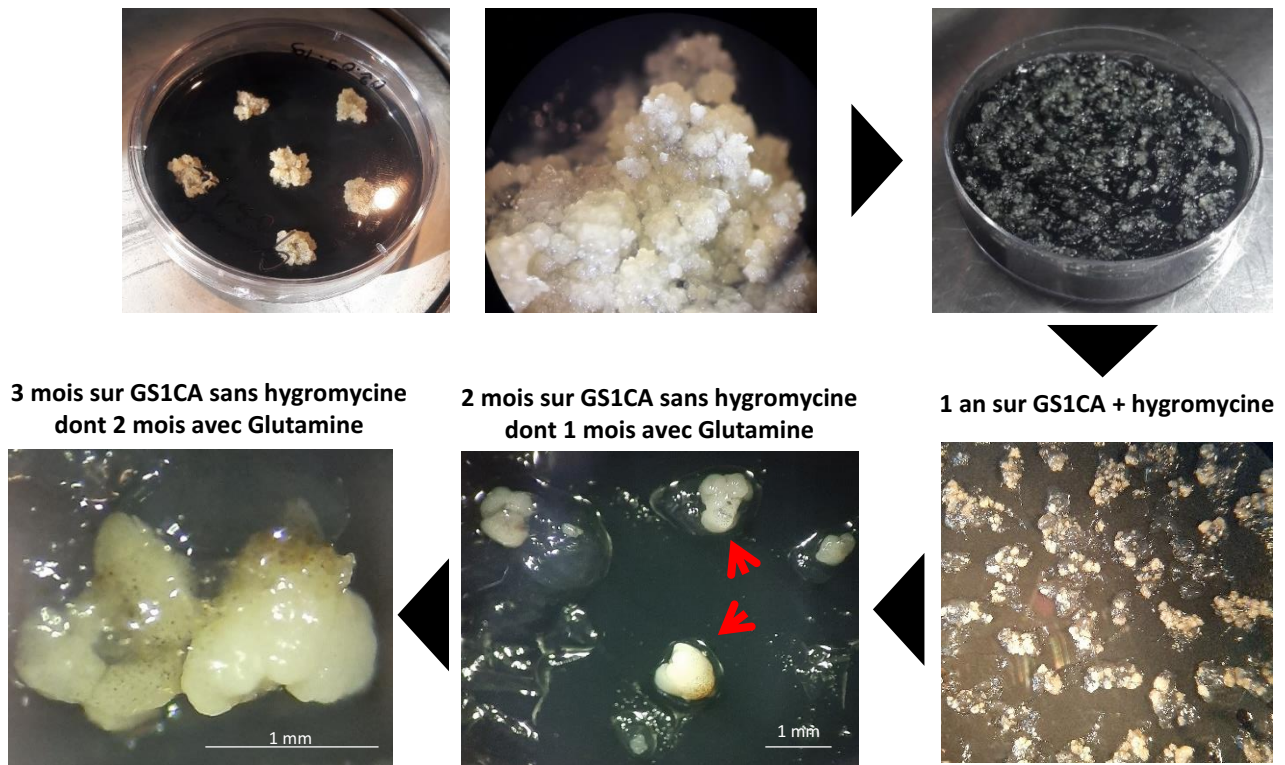
ii) Le cas de la Coamplification pour une mutation inconnue ou une délétion

La Coamplification pour la détection d'une mutation inconnue est basée sur la bonne répartition des ADN dans les gouttelettes. Les deux amplifications doivent idéalement se retrouver dans la même gouttelette. Ici les deux amplifications sont sur le même gène, à 353 pb d'écart (**Figure 43-A**), alors que pour le riz, il s'agissait de deux amplicons sur deux chromosomes distincts. Nous allons voir que le biais d'amplification (une amplification plus facile que l'autre) peut poser des problèmes dans l'interprétation.

Trois clusters apparaissent dans les deux représentations : un cluster orange majoritaire représentant les gouttelettes avec les 2 amplifications (WT) ; un vert censé représenter les mutants car c'est la seule amplification de la référence ; et un amas gris (bleu si pris en compte dans le calcul comme WT au locus ciblé) représentant l'amplification autour de la zone mutée (**Figure 43-D et E**). Le biais d'amplification ou de répartition est dû au fait que les amas gris et verts sont toujours présents dans l'échantillon WT (**Figure 43-D**). Il faut alors utiliser le ratio de l'échantillon WT pour en déduire le ratio après édition avec le plasmide CRISPR. C'est-à-dire appliquer la formule : $\text{HEX}/[(\text{HEX et FAM})+\text{HEX}]$ chez l'échantillon avec le CRISPR, moins $\text{HEX}/[(\text{HEX et FAM})+\text{HEX}]$ chez le contrôle. Par ce calcul, nous obtenons une estimation de 5.81 % d'édition par le CRISPR simplex pDE_CAS9_cr265GAI (**Figure 43-E**). Or nous avons omis complètement les gouttelettes FAM seules, qui à la base sont censées représenter le témoin d'ADN présent dans la gouttelette. Si nous les réintégrons, nous obtenons une estimation de 5.16 % de mutants (**Figure 43-E**).

Pour le test des deux multiplexes définis pour déléter le domaine DELLA (pDE_CAS9_cr265-223GAI et pDE_CAS9_cr265-206GAI) il a fallu effectuer une coamplification avec la sonde FAM non pas sur le site du cr265, mais au niveau du domaine DELLA (**Figure 44-A**). Nous avons appliqué à nouveau le principe de la coamplification présenté ci-dessus pour le calcul des mutants en les comparant aux WT (**Figure 43-B et D**). A cause des gouttelettes grises (FAM seules), deux possibilités à nouveau. Si les FAM seules sont éliminées du calcul, les protoplastes présentent 10.2 % et 10.7 % de délétion (**Figure 44-C et E**). Si toutes les FAM sont incluses dans le calcul, le taux de délétion passe à 7 % et 6.46 % (**Figure 44-C et E**).

Les résultats permettent néanmoins de conclure que les CRISPR semblent efficaces et induisent une délétion pour des taux compris entre 5 et 10%. Cependant, si nous comparons au DropOff qui a un taux inférieur (4 %), nous pourrions imaginer que nous sommes plus près de 5 % de délétion. Malheureusement, il aurait fallu disposer d'un vrai contrôle Dropoff pour mesurer ce taux de délétion et le comparer avec les stratégies de co-amplification. Malheureusement, une sonde HEX de référence n'était pas positionnable sur l'amplicon produit



3 mois sur GS1CA sans hygromycine
dont 2 mois avec Glutamine

2 mois sur GS1CA sans hygromycine
dont 1 mois avec Glutamine

1 an sur GS1CA + hygromycine

Figure 45 : Transformation et régénération chez la vigne (2019) : stade embryonnaire

Des transformations stables par Agrobactérium de plasmides CRISPR dicotylédones (pDE_GAI265 / 265-223 / 265-206) ont été faites à partir de cultures d'anthers de Grenache à partir du protocole de Bouquet et al, (2006).

Les embryons ne se sont pas développés après 8 mois sur GS1Ca+hygromycine. Après avoir retiré l'hygromycine et ajouté de la Glutamine, les embryons ont commencé à se différencier en prenant la forme « cœur » après 2 mois.

pour l'hybridation avec en suivant les consignes imposées par BioRad. Néanmoins, les résultats obtenus sur le riz suggèrent que les taux estimés par co-amplification se rapprochent des taux estimés par Dropoff. Nous pouvons donc estimer les taux de délétion en protoplastes, dans une fourchette comprise entre 5 et 10%.

3) transformation stable de la vigne avec les construits validés

Trois génotypes avaient été retenus : Grenache, Portan, et Merlot. Seul le Grenache à produit des structures secondaires à partir de ses cals primaires. L'hygromycine comprise dans le milieu GS1Ca n'a pas été retirée après 2 à 3 mois comme initialement prévu par le protocole, mais après 8 mois (une des conséquences du confinement). Nous voulions la laisser 4 mois pour éviter de régénérer des non transformés. En conséquence cela a eu en partie pour effet de bloquer les cellules sous forme de cals non différenciés (**Figure 45**, photo en bas à droite).

Pour faciliter leur différenciation en embryons, deux stratégies ont été mises en place après avoir complètement retiré l'hygromycine des milieux. Nous avons ajouté de la glutamine dans le GS1Ca et placé la moitié des embryons à la lumière (indirecte) à 26 °C.

Au bout d'un mois, les « cals embryogènes » ont gonflé et commencé à blanchir. Et au bout de deux mois, la différenciation a débuté. Des embryons en forme de « cœur » de moins d'un millimètre ont commencé à se développer, soit isolés sous des amas de cellules encore non différenciées, soit en grappe avant de se détacher. Seuls les embryons de Grenache arrivent à se différencier (**Figure 45**). Une hypothèse serait que les embryons de Portan et Merlot n'étaient pas au bon stade (**Annexe 17**), ce qui est très probable car ils n'ont pas régénéré comme le Grenache. Une autre hypothèse serait que tous les embryons actuels seraient bien mutés KO dans *Vvgai* ce qui affecterait la dédifférenciation embryonnaire et ralentirait la croissance. Cette dernière hypothèse est néanmoins peu probable car toutes les mutations par CRISPR ne sont pas des KO, il pourrait y avoir une substitution, ou une délétion d'un seul codon, qui n'affecte pas le cadre de lecture de la protéine.

A ce stade de la thèse, il n'a pas été possible d'analyser l'ADN des embryons (taille < 2 mm) et confirmer s'ils sont transformés et mutés. Ils sont encore en cours de régénération pour plusieurs mois, et nous espérons pouvoir présenter des résultats de plantules avec différents types de mutations pour l'oral.

4) Optimisation des construits CRISPR/CAS9 chez la vigne

Pour la vigne, nous avons également choisi de développer de nouvelles constructions pour lesquelles les plasmides ont été optimisés pour cette espèce via le promoteur non codant contrôlant l'expression du sgRNA ou l'optimisation des codons de la CAS9.

La plupart des constructions utilisées sont dérivées de plasmides utilisés chez *A. thaliana* et sont donc potentiellement optimisables pour une utilisation avec la vigne. En effet, chaque acide aminé peut être codé par plusieurs codons, phénomène appelé « biais d'usage des codons ».

	U3 Riz (code plasmide 217)	AtU6-26 Arabidopsis (code plasmide pEN_Chimera-C1.1)	VvU6 Vitis (code plasmide pEN_C1.1-VvU6-1.0)
CAS9 Riz (code plasmide 2G9)	2G9-217-265gai A	2G9-pEN-265gai B	2G9-VvU6-265gai C
CAS9 Arabidopsis (code plasmide pDE-Cas9_Hpt)	pDE-217-265gai D	pDE-265gai E	pDE-VvU6-265gai F
CAS9 Vigne (code plasmide pDVitis_VvCas9_Hpt)	pDVitis-217-265gai G	pDVitis-pEN-265gai H	pDVitis-VvU6-265gai i

Figure 46 : « Matrice » des combinaisons des différentes CAS et promoteurs non codants

On teste ici l'effet des promoteurs non codants OsU3/AtU6/VvU6, en combinaison avec les protéines optimisées OsCAS9/AtCAS9/VvCAS9. Le CRISPR utilisé est le GAI_cr265 qui avait montré une édition avec le plasmide pDE_GAI265 (AtU6 + AtCAS9).

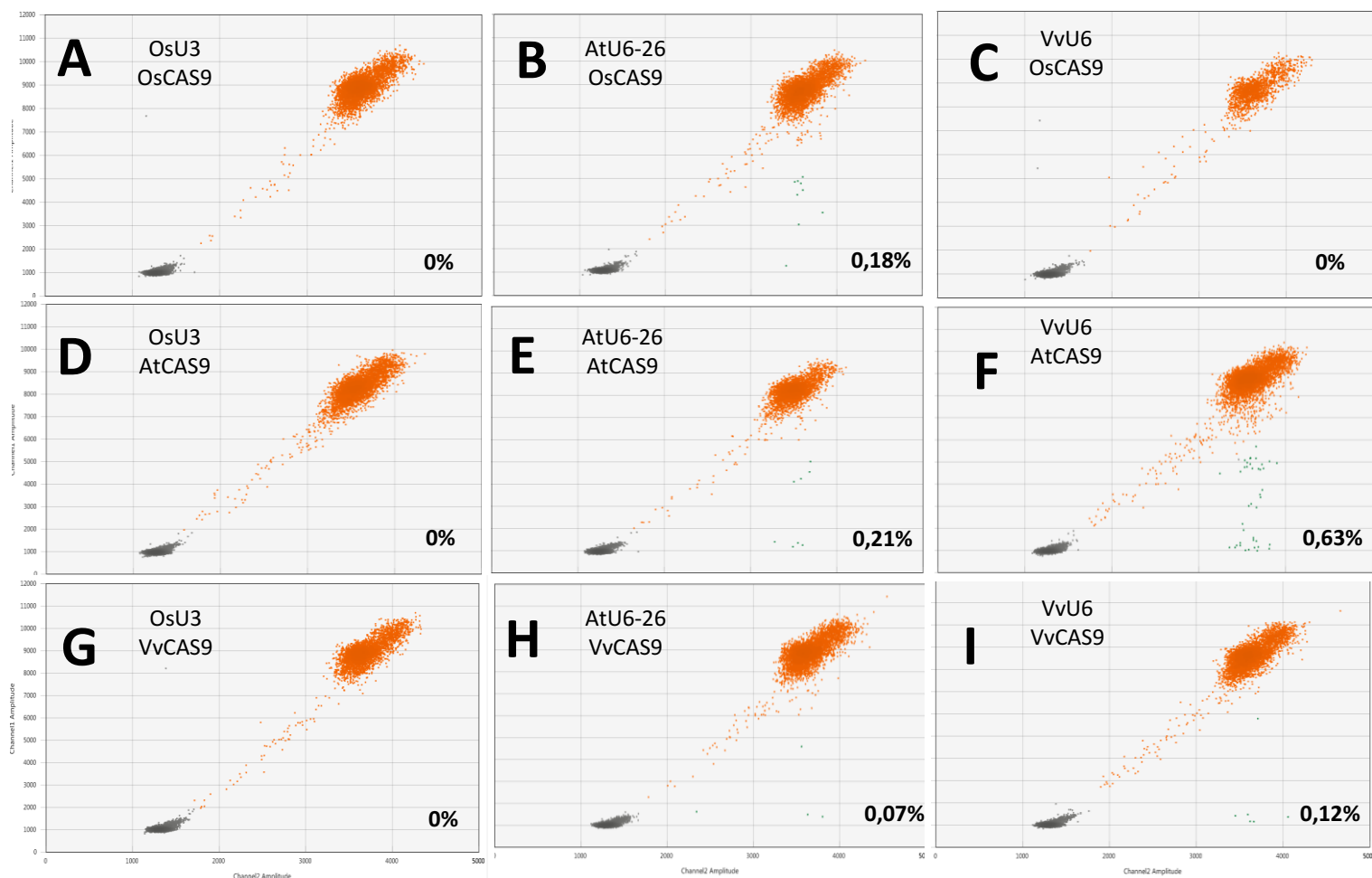


Figure 47 : ddPCR des combinaisons plasmidiques – Résultats préliminaires

Résultat des expérimentations DropOFF avec les différents plasmides décrits à la Figure 48. Les plasmides A, C, D, et G présentent 0 % d'édition. B et E près de 0,2 % ; H et I présentent près de 0,1 %. La combinaison VvU6 avec AtCAS9 (F) semble la plus efficace avec 0,63 % d'édition.

Il est donc possible d'optimiser les séquences codantes et ainsi optimiser la synthèse de la protéine ou l'expression du sgRNA et augmenter l'efficacité des CRISPR. A l'aide de la technologie ddPCR, qui permet de comparer quantitativement différents construits, il semblait intéressant de tester l'effet combiné de différents promoteurs riz, vigne, Arabidopsis et différentes CAS9 optimisées riz, vigne ou Arabidopsis.

Le spacer crGAI-265 a été utilisé car nous avons démontré qu'il était fonctionnel dans le plasmide dicotylédone (pDE-GAI-265, CAS9 Arabidopsis et pAtU6-26 avant le sgRNA) (**Annexe 1-B et C**). Il a donc été à nouveau cloné dans deux nouveaux plasmides d'entrées.

Concernant les nouveaux vecteurs d'entrée, nous avons utilisé le vecteur d'entrée riz (promoteur non codant pOsU3 (**Annexe 1-E**)), et nous avons développé un nouveau vecteur d'entrée à partir de pEN-C1.1 (**Annexe 1-C**). En se basant sur sa séquence et sa structure fournie par Jakab et coll. (Jakab et al., 2006), le pVvU6 a été substitué jusqu'à la TATA box au pAtU6-26 dans pEN-C1.1. Seules les 30 bases après la TATA box ont été conservées jusqu'au sgRNA afin de respecter l'initiation de la transcription (**Annexe 18**). Le plasmide final obtenu sera nommé pEN-C1.1_VvU6-1.0 (**Annexe 1-F**).

Concernant les vecteurs binaires, nous avons produit un nouveau vecteur à partir de pDE-CAS9-Hpt. La VitisCAS9 a été entièrement optimisée à partir de la table de Kazusa (<https://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=29760>) (**Annexe 19-A**) qui décrit les codons les plus utilisés chez la vigne dans les séquences codantes. La codon optimisation a été réalisée manuellement sur près de 4140 pb, à partir de la séquence de la CAS9 du pDE-CAS9-Hpt et la CAS9 vigne a été substituée à la CAS9 codon optimisée du vecteur pDE-CAS9-Hpt. Il a fallu aussi faire attention à un maximum de 4 T consécutifs, et du coup changer quelques Serines sous forme AGT au lieu de TCT, et des Leucines CTT au lieu de TTG, qui étaient pourtant théoriquement plus efficaces (**Annexe 19-B et C**). Le plasmide final obtenu a été renommé pDVitis-VvCAS9-Hpt (**Annexe 1-G**).

A partir de ces vecteurs, nous avons généré les vecteurs binaires finaux portant soit le promoteur non codant Riz OsU3, soit AtU6-26, soit VvU6, associé à chaque CAS9 disponible (riz, Arabidopsis, ou vigne) (**Figure 46**), soit 9 construits binaires différents, ciblant *VvGAI* avec le même sgRNA portant le spacer cr256. L'objectif final était de comparer les taux d'édition de chacun des construits par ddPCR en utilisant la technologie DropOff.

Les résultats préliminaires sont issus d'une seule expérimentation en protoplastes vigne, obtenus avec un taux de transformation estimé à 24 % à l'aide de trois contrôles GFP (pHKN x2, et pDE_GFP). Les résultats présentés sont préliminaires et il sera nécessaire de répéter si possible en augmentant le taux de transformation situé ici, plutôt dans la fourchette très basse. Le taux d'édition obtenu avec le vecteur utilisé précédemment (AtCAS9 et U6 Arabidopsis) n'est que de 0.21 % d'édition. Dans de précédentes transformations, il était 20x plus performant (**Figure 43-C**), ce qui semble confirmer le problème de transformation. Les premiers résultats permettent néanmoins, avec beaucoup de prudence de dégager de premières tendances, mais encore une fois, nécessiteront de répéter cette expérimentation, plusieurs fois pour conclure (**Figure 47**).

L'utilisation du U3 de riz, quelle que soit la CAS9 utilisée, n'a pas permis de détecter d'édition

en protoplastes de vigne (**Figure 47-A, D, et G**), alors que les promoteurs U6 de vigne et d'*Arabidopsis* ont permis, à l'exception de la combinaison VvU6/CAS9 riz, d'obtenir des taux d'édition oscillant entre 0.07 % et 0.63 % (**Figure 47-B, E, F, H, I**). Ce résultat suggère que le U3 de riz est beaucoup moins voire non fonctionnel chez la vigne, comparé au U6 de vigne et d'*A. thaliana*. Les taux d'édition obtenus avec les promoteurs U6 de vigne et d'*Arabidopsis* semblent équivalents, en tout cas, il n'est pas possible de les départager au vu des faibles taux d'édition.

La CAS9 d'*Arabidopsis* est supérieure à la CAS9 codon optimisée de vigne ou de riz, car elle a permis d'obtenir les meilleurs taux d'édition, quels que soient les promoteurs utilisés. Encore une fois, les taux sont proches, ces résultats ne sont pas définitifs. La CAS9 codon optimisée de riz n'a permis d'obtenir des éditions que combinée avec le promoteur U6 d'*Arabidopsis*.

Enfin, la meilleure combinaison semble être l'association du promoteur U6 de vigne avec la CAS9 codon optimisée d'*A. thaliana* (pDE-VvU6-265gai). Malgré les faibles taux d'édition, la sensibilité de la ddPCR a néanmoins permis d'obtenir une première série de comparaisons entre tous les vecteurs binaires (**Figure 47-F**).

Des transformations seront conduites en protoplastes, en 2021 à l'issue de ma thèse afin d'identifier la combinaison la plus performante, promoteur U6 et CAS9. Ces plasmides vont également être transformés stablement en vigne par *Agrobacterium* pour comparer leur efficacité *in-planta* courant 2021. En combinant les résultats qui seront obtenus en transformation stable et en transformation transitoire, il sera à la fois possible de déterminer si la transformation transitoire est un bon prédicteur de l'efficacité d'édition en plante, mais également d'identifier la meilleure combinaison promoteur/CAS9 pour de futures expérimentations utilisant la technologie CRISPR/CAS9 chez la vigne.

Discussion générale

Dans le cadre des deux chapitres de résultats de la thèse, nous avons tout d'abord abordé les développements méthodologiques qui ont été indispensables au bon déroulement de mon travail, comme le développement d'un système de prototypage et de la ddPCR pour tester nos construits. Nous avons ensuite abordé les résultats des travaux spécifiques sur le riz pour la génération de séries alléliques et pour l'édition à la base près des gènes d'intérêt agronomiques *OsNRT1.1b* et *OsBT*, puis les travaux spécifiques sur la vigne, en choisissant le gène *VvGAI* comme modèle. Nous allons successivement discuter certains points en lien avec les développements de la thèse, avant une discussion plus générale.

II Réflexions sur la mise au point des différentes techniques

1) Les protoplastes comme système de prototypage de construits CRISPR : avantages et limitations

Un facteur limitant dans l'utilisation de la technologie CRISPR est la possibilité de générer des plantes transgéniques. Chez la vigne, pour laquelle il faut au moins une année pour produire des plantes transformées, il était donc critique d'évaluer en amont l'efficacité des construits CRISPR utilisés. Dans la première partie de la thèse, nous nous sommes attardés sur les différents systèmes de mise en place et de tests de nos expérimentations afin de créer un système de prototypage en protoplastes et évaluer la meilleure stratégie pour rapidement tester et comparer nos constructions avant transformation génétique.

L'utilisation de protoplastes de riz ou de vigne a permis de tester en grand nombre des plasmides CRISPR pour différentes cibles au cours d'une seule expérimentation. L'expérience sur les protoplastes de riz a servi de base pour le protocole vigne et ses améliorations, notamment côté transformation. Les expériences menées sur le riz démontrent leurs très bonnes reproductibilités (Bès et al., 2021). Côté vigne, après avoir obtenu de très bons résultats en 2019 (**Figure 22** et **Figure 43**), nous avons eu de très faibles taux de transformation en 2020. L'utilisation de cultures cellulaires plus âgées, et potentiellement plus instables, en est probablement la raison principale. Malheureusement, les expérimentations menées en 2020 n'ont pas permis de conclure sur les différentes optimisations testées (codon optimisation de la CAS9, promoteurs U6 de vigne...). Malgré tout, cela n'a pas empêché de valider la fonctionnalité de plusieurs plasmides, avec des taux d'édition très faibles, grâce à la ddPCR (**Figure 47**). Il conviendrait peut-être de trouver une solution alternative pour la production de protoplastes. D'autres tissus sources pourraient être utilisés, comme les feuilles à l'aide de protocoles dérivés (Cantu and Walker, 2019; Dalla Costa et al., 2019; Malnoy et al., 2016). Nous avons essayé d'adapter le protocole du riz sur des feuilles de vitro plants de vigne en rajoutant une enzyme de digestion (pectolyase). Trop peu de protoplastes avaient été obtenus et les taux de transformation n'étaient que de 10 %. Pour estimer les taux de transformation en vigne, nous avons utilisé un construit comprenant une GFP, de taille et de structure assez différente de nos construits CAS9 (Kumagai and Kouchi, 2003). Il est censé représenter un taux moyen de transformation dans un échantillon contrôle pour toute l'expérimentation. Cependant, il ne

permet pas de mesurer le taux de transformation réel pour chacune des constructions dans chaque échantillon. De plus, l'efficacité réelle d'édition est bien sûr dépendante du taux de transformation dans chaque échantillon. Une solution consisterait à utiliser le sgRNA_GAI265 (sous AtU6-26 qui fonctionne bien) dans le plasmide pDE_atCAS9_GFP. Ce vecteur, construit pendant la thèse permettra d'estimer simultanément dans le même échantillon le taux de transformation (GFP +) et le taux d'édition (par CRISPR) pour en estimer l'efficacité réelle.

2) La ddPCR pour quantifier le taux de mutations global et spécifique comparé au séquençage NGS : combiner le meilleur des deux mondes

L'utilisation de la ddPCR pour évaluer les taux d'édition a permis d'écartier certaines constructions pour la transformation stable. De plus, malgré de très faibles taux d'édition obtenus pour la vigne, nous avons pu déterminer quantitativement le taux de mutations pour chaque construction (Herbert et al., in prep ; **Figures 4 et 7**). Le plasmide base editor utilisé pour le riz, n'avait jamais été testé dans l'équipe avant ce travail. Nous avons pu montrer que le taux d'édition était autour de 5 % en protoplastes, et nous a permis d'utiliser cette construction en transformation stable du riz. Nous savions que nous pouvions l'utiliser pour d'autres cibles (iSTOP) grâce aux sorties des premières plantes éditées à la base près sur *OsNRT1.1b* (Herbert et al., 2020). Pour la vigne nous avons éliminé plusieurs constructions ciblant *VvGAI* grâce à la ddPCR (**Figure 47 et Annexe 20**) et évalué l'efficacité de certains plasmides, malgré les très faibles taux (pDE initial vers 5 % en transitoire dans nos conditions (**Figure 43**)). En effet, l'utilisation de la T7EI n'aurait pas permis de conclure, sa sensibilité étant au moins 100 fois inférieure à la ddPCR.

La technologie DropOff permet d'estimer le taux de mutation global, *ie* le nombre de mutants générés quelle que soit la nature de ces mutations, et le DropOn le taux de mutations spécifiques, par exemple pour le « Base Editor ». En combinant ces deux technologies, il est théoriquement possible de comparer ces deux taux et d'estimer si une construction pour l'édition à la base près est efficace, générant les mutations attendues avec un taux élevé. Malheureusement, il n'est pas possible d'estimer simultanément au sein d'un même échantillon le taux de mutation global et spécifique. Une solution alternative, théoriquement très prometteuse, permet d'obtenir ces taux en utilisant plus de deux, au moins trois, mais théoriquement jusqu'à cinq fluorochromes en même temps avec la technologie QIAcuity. Nous sommes en train de la tester sur les échantillons présentés dans l'article sur la ddPCR.

Les taux et le type de mutations ont fréquemment été estimés dans le passé par séquençage NGS d'amplicons obtenus après transformation de protoplastes (Wang et al., 2016a). Si quelques laboratoires ont déjà utilisé la ddPCR chez les plantes (Collier et al., 2017), cette technique est maintenant utilisée en routine dans les systèmes animaux (Findlay et al., 2016 ; Garcia et al., 2017). L'avantage du NGS est d'obtenir exhaustivement les mutations générées par une construction CRISPR/CAS9 et de permettre leur quantification. La ddPCR de son côté présente de nombreux avantages en termes de coûts et de labeur. Le séquençage NGS peut

néanmoins être utilisé pour vérifier le type de mutations induites par une construction, en complémentarité avec la ddPCR. Il est en effet possible de combiner ddPCR et séquençage NGS, puisqu'il est possible d'extraire les amplicons des gouttelettes générées par la lecture de droplet (BioRad) si nécessaire. Il est donc envisageable d'imaginer une stratégie en deux temps. L'utilisation de la ddPCR pour évaluer l'efficacité en protoplastes d'un grand nombre de construits, puis pour les plus efficaces, en extraire l'ADN et obtenir de manière exhaustive l'ensemble des mutations générées pour avoir non seulement une idée quantitative, mais aussi qualitative des mutations générées et donc la spécificité d'une construction CRISPR.

3) Quelle fiabilité des résultats obtenus en transformation transitoire pour prédire l'édition et la spécificité de construits CRISPR/CAS9 en transformation stable ?

Par ailleurs, il fallait s'assurer que notre stratégie utilisant des transformations transitoires *in-vitro* permettait effectivement de prédire l'édition *in-planta*. Nous avons donc utilisé le riz, pour vérifier que ces prédictions *in-vitro* étaient bien corrélées aux efficacités *in-planta*. En effet, dans le fond génétique Kitaaké, la production de plantes transgéniques est bien maîtrisée et ne prend que quelques mois ce qui nous a permis de corréler nos efficacités d'édition du génome en protoplastes et en plantes (Herbert et al., in prep). Chez le riz, les taux obtenus *in-planta* sont systématiquement supérieurs à ceux obtenus en transitoire avec le « Base Editor » (5 % détectés en protoplastes contre 50 % de plantes régénérées éditées). L'explication la plus vraisemblable est que lors de la transformation stable en riz, qui s'opère très tôt (Sallaud et al., 2003), il y a une expression forte et permanente de la CAS9 ce qui augmente la probabilité que les cellules transformées et sélectionnées sur antibiotique soient éditées.

Nous avons pu *in fine* reproduire *in-planta* les mutations d'intérêt agronomique : pour *OsNRT1.1b* une mutation par substitution *OsNRT1.1b* C>T à l'aide du « Base Editor », similaire à celle présentée par (Hu et al., 2015) et une mutation par décalage du cadre de lecture dans *OsBT* (mutation perte de fonction pour reproduire les résultats obtenus par (Araus et al., 2016)). Ces deux mutations ont été introduites dans deux fonds génétiques *japonica* distincts : Kitaaké, génotype modèle utilisé fréquemment au laboratoire et au niveau international (Tyagi and Mohanty, 2000), et Chhomrong Dhan, cultivé à Madagascar, mais jamais transformé ou édité par CRISPR, deux premières.

Chez la vigne, pour maximiser nos chances nous avons utilisé deux stratégies : introduire un frameshift dans le gène *VvGAI* par CRISPR/CAS9, ou en déléant son domaine DELLA en espérant conserver le cadre de lecture (délétion multiple de 3). Il est bon de rappeler que pour la vigne, le gène *VvGAI* avec le motif muté DELHA n'est pas KO, et il semblerait que l'introduction du gène *VvGAI* avec de grands domaines délétés en *Arabidopsis* ait le même effet (Zhong and Yang, 2012). Pour les plasmides optimisés, les expérimentations doivent être validées, suite aux problèmes de transformation transitoire, mais quelques premiers résultats encourageants montrent que les plasmides disposant de l'AtCAS9 ou VvCAS9, et AtU6 ou VvU6 sont intéressants. Ils devraient être utilisés pour des transformations stables en collaboration avec l'UMR EGFV de INRAe de Bordeaux.

A

crRNA (+PAM en rouge)	Séquence OFF target semblables (2 OFF)	Position des OFF targets p/r au crRNA	Code d'amplification
CGGCGACGGCGAGCAAGTGGAGG	CGTCGACGGCGAGCAAGTGGAGG	.. * *	1
CGGCGACGGCGAGCAAGTGGAGG	CGGCGACGGCGAGCAAGCGGCGG * ..*	2
CGGCGACGGCGAGCAAGTGGAGG	CGACGACGGCGAGCCAGTGGTGA	.. * *	3
CGGCGACGGCGAGCAAGTGGAGG	CGGCGAAGGCGAGCCAGTGGTGA **	4
CGGCGACGGCGAGCAAGTGGAGG	CGGCGACGGTGAGGAAGTGGCGG * ..*	5
CGGCGACGGCGAGCAAGTGGAGG	CGGCGACGGCGAGGACGTGGAGG * ..*	6
CGGCGACGGCGAGCAAGTGGAGG	CGGCGACGGCGAGGAGGTGGTGG * ..*	7
CGGCGACGGCGAGCAAGTGGAGG	CGGCGACGGCGAGCATCTGGTGG ** ..	8
CGGCGACGGCGAGCAAGTGGAGG	CGGCGACGGCGAGCATCTGGTGG ** ..	9
CGGCGACGGCGAGCAAGTGGAGG	CGGCGACGGCGAGGGAGTGGTGG ** ..	10
CGGCGACGGCGAGCAAGTGGAGG	CGGCGACGGCGAGCATCTGGTGG ** ..	11
CGGCGACGGCGAGCAAGTGGAGG	CGGCGACGGCGAGCTCGTGGTGG ** ..	12

B

OFF 1	crRNA	CGTCGACGGCGAGCAGGTGGAGG
	WT Kita	TGTGGACGGCGAGCAGGTGGAGG
	Kita BE_NRT1.1_G (T)	TGTGGACGGCGAGCAGGTGGAGG
	Kita BE_NRT1.1_H (T)	TGTGGACGGCGAGCAGGTGGAGG
	CHHD BE_NRT1.1_T2 (T)	TGTGGACGGCGAGCAGGTGGAGG
CHHD BE_NRT1.1_T1 (T)	TGTGGACGGCGAGCAGGTGGAGG	
OFF 3	crRNA	CGACGACGGCGA-GCC-AGTGGTGA
	WT Kita	CGCCGTCTCGCCGCCGACTGGTGT
	Kita BE_NRT1.1_G (T)	CGCCGTCTCGCCGCCGACTGGTGT
	Kita BE_NRT1.1_H (T)	CGCCGTCTCGCCGCCGACTGGTGT
	CHHD BE_NRT1.1_T2 (T)	CGCCGTCTCGCCGCCGACTGGTGT
CHHD BE_NRT1.1_T1 (T)	CGCCGTCTCGCCGCCGACTGGTGT	
OFF 6	crRNA	CGGCGACGGCGAGGACGTGGAGG
	WT Kita	CACCGCCGCCGCAACCACGGCGG
	Kita BE_NRT1.1_G (T)	CACCGCCGCCGCAACCACGGCGG
	Kita BE_NRT1.1_H (T)	CACCGCCGCCGCAACCACGGCGG
	CHHD BE_NRT1.1_T2 (T)	CACCGCCGCCGCAACCACGGCGG
OFF 8	crRNA	CGGCGACGGCGAGCATCTGGTGG
	WT Kita	CGCCCTCGCCGTGCCAAGGGGA
	Kita BE_NRT1.1_G (T)	CGCCCTCGCCGTGCCAAGGTGA
	Kita BE_NRT1.1_H (T)	CGCCCTCGCCGTGCCAAGGTGA
	CHHD BE_NRT1.1_T2 (T)	AGCCATCGGCGTGCCAAGATGA
CHHD BE_NRT1.1_T1 (T)	CGCCATCGGCGTGCCAAGGTGA	
OFF 10	crRNA	CGGCGACGGCGAGGGAGTGGTGG
	WT Kita	CGGCGACGGCGAGGGAGTGGTGG
	Kita BE_NRT1.1_G (T)	CGGCGACGGCGAGGGAGTGGT-G
	Kita BE_NRT1.1_H (T)	CGGCGACGGCGAGGGAGTGGTGG
	CHHD BE_NRT1.1_T2 (T)	CGGCGACGGCGAGGGAGTGGT-G
CHHD BE_NRT1.1_T1 (T)	CGGCGACTGCGAGTTAGTG-TTG	

Figure 48 : Analyse préliminaires des Off-Targets du crNRT1.1b

Tableaux des résultats préliminaires des séquençages autour des OFF targets du crNRT1.1b à 2 mésappariements.

Présentation des séquences et positions des mésappariements (en vert, ou *) vis-à-vis du crNRT1.1b (A).

Présentation des résultats des séquençages sur cinq sites pour quatre plantes éditées à la base près pour *NRT1.1b* (B).

Attention, il n'y a pas eu d'amplification du WT_A1 CHHD et la qualité du couple 8 et 10 n'a pas permis d'avoir un séquençage très propre (exemple avec le CHHD BE_NRT1.1_T1 OFF 10)

L'idée est d'améliorer l'efficacité des construits *in-planta*, pour augmenter le taux de mutation avec la même efficacité de transformation afin de réduire le nombre de plantes à régénérer et à analyser. L'idée même de s'affranchir des plasmides en passant par des RNP (Malnoy et al., 2016 ; Woo et al., 2015) semble intéressante, à la condition que les protoplastes soient régénérables, ou bien de passer par de la transformation par bombardement des complexes RNP à l'aide de billes d'or.

4) Comment améliorer la sécurité des technologies CRISPR/CAS9 ? mutations aléatoires, off-Targets et in-targets

a) Off-Targets

Le sgRNA peut cibler avec une certaine tolérance, prédictible en partie *in silico* (Li et al., 2019), d'autres endroits du génome. Pour avoir une estimation de ce taux de off-Target uniquement au niveau des protospacers similaires pour le « Base Editor », nous avons recherché la présence de mutations au niveau de ces sites prédits, qui possédaient le moins de mésappariements (2) pour les 20 pb du crNRT1.1b. Nous avons pu en tester 8 sur les 12 mésappariements prédits de 2 bases dans le génome du riz, sur près de 1200 off-Targets à plus de 2 mésappariements. Cinq d'entre eux ont pu être séquencés sur les plantes Kitaaké et CHHD WT et édités à la base près (**Figure 48-A**). Aucun OFF-Target C>T en 7e position (**Figure 48-B**) n'a été mis en évidence. Les résultats sont rassurants, mais ne permettent pas d'exclure totalement qu'il n'y ait aucun site « Off-Target » induit par le « Base Editor ». Seul un reséquençage complet du génome des plantes éditées pourrait le confirmer. Indépendamment du sgRNA, il a été démontré que la cytosine déaminase du « Base Editor » avait une tendance à modifier aléatoirement dans le génome d'autres cytosines (Jin et al., 2019). Ce problème, qui n'était pas connu lorsque le projet a été lancé, conduit à suggérer d'utiliser préférentiellement des Adénines en « Base Editor » (ABE) alternativement aux cytosines. En effet, si les Cytosines « Base Editor » (CBE) génèrent bien des mutations aléatoires dans des régions transcrites du génome à cause de la C-déaminase et/ou l'UGI, non prédictibles par les logiciels comme Cas_OFFfinder, cela ne semble pas être le cas des ABE (Jin et al., 2019). Dans le futur, il sera donc nécessaire d'utiliser des ABE pour introduire des mutations à la base près chez le riz et plus globalement chez les plantes pour minimiser le risque d'introduire des mutations non voulues qui ne peuvent-être, totalement éliminées.

Enfin, la transformation génétique passe par une phase de culture *in-vitro* qui est également connue pour introduire un certain nombre de mutations génétiques, ou épigénétiques dont la fréquence est corrélée à la longueur de la phase de culture *in-vitro* (Kaeppeler et al., 2000). Plus globalement, ces mutations dues à la CIV peuvent être gênantes s'il a des conséquences sur la performance agronomique des plantes produites. En comparant plusieurs plantes issues d'un même cycle de transformation, et en les comparant aussi à un contrôle, il est possible d'éliminer celles qui présentent des effets délétères et de ne conserver que celles qui sont les plus conformes. Enfin, la culture *in-vitro*, a été utilisée depuis plus de 50 ans en amélioration génétique (haploïdes doublés, fusion de protoplastes, sauvetage d'embryons,

amélioration génétique (haploïdes doublés, fusion de protoplastes, sauvetage d'embryons mutagenèse aléatoire... (Sathishkumar et al., 2019)) et ces mutations non voulues, délétères peuvent être éliminées par évaluation phénotypique et rétrocroisements. Mais dans la perspective de la nouvelle réglementation européenne sur la définition de plantes OGM (jugement Cour Européenne de Justice), l'existence de mutations hors site freinera inévitablement l'acceptabilité de la méthodologie CRISPR.

b) In-Targets

Sur le protospacer, il peut aussi y avoir certaines modifications non désirées, en effet nous avons remarqué que par moment le système « Base Editor » pouvait faire des erreurs et convertir la C en G ou effectuer une délétion. Nous avons en effet obtenu 3 plantes Kitaaké et 1 plante CHHD avec une erreur de conversion, et 3 plantes Kitaaké délétées (**Annexe 2**). Ces taux restent très faibles (Ren et al., 2018) et il est facile d'éliminer les plantes ainsi mutées par séquençage. A noter que nous sommes en train de tester les protoplastes avec la technologie QIAcuity afin d'avoir aussi accès à ces taux en transitoire (Herbert et al., in prep ; **Figure 8**).

En conclusion sur cette partie, évaluer les construits CRISPR en protoplastes, et combiner la détection de leur efficacité par ddPCR est une nécessité pour décider des construits à passer en transformation stable, en particulier pour la vigne et pour les espèces récalcitrantes à la transformation. Enfin, afin de vérifier les potentielles mutations Off-Targets des outils prédictifs peuvent être utilisés. Certaines approches comme le CIRCLE-seq (Li et al., 2019 ; Tsai et al., 2017) pourraient être utilisées pour cibler et analyser les off-Targets potentiels, et comparer la spécificité des construits les plus efficaces. Malgré tout pour vérifier toutes les mutations il reste la possibilité de reséquencer le génome entier, mais les coûts sont élevés et les mutations off-Target induites par une construction CRISPR devront être distinguées de celles induites par la phase de culture *in-vitro*, multipliant le nombre de plantes dont le génome serait à reséquencer et à comparer.

III) Réflexions sur les technologies dérivées de CRISPR et l'utilisation sur des gènes d'intérêt agronomique

1) La technologie CRISPR/CAS9 et ses dérivés : un outil puissant et sans équivalent pour générer rapidement et simplement des séries alléliques et faciliter l'analyse fonctionnelle et la sélection

Nous avons pu montrer qu'il était possible d'orienter et de rationaliser les stratégies d'édition, en choisissant judicieusement l'endonucléase (CPF1 vs CAS9) ou un outil dérivé de la CAS9 (le « Base Editor ») pour faciliter l'analyse fonctionnelle en générant, dans un même fond génétique, une série allélique (Herbert et al., 2020). Cela peut se faire simplement en générant des knock-out par frameshift (CAS9), ou bien en délétant un codon (CPF1), ou encore en affectant la traduction de l'ARNm (BE iSTOP). Sans parler des évolutions à venir ou déjà en place comme l'utilisation d'adénine déaminases (Jin et al., 2019), ou bien les CAS9_cis régulatrices (Li et al., 2020a) par CRISPR interférence ou CRISPR activation. Nous avons démontré la faisabilité de cette approche sur un gène modèle, qui nous a permis d'obtenir une gamme de phénotypes rapidement observables et non létaux. Ce modèle facilitera la validation future de nouveaux outils dérivés de la technologie CRISPR/CAS9. Générer des séries alléliques permet de faciliter l'analyse fonctionnelle d'un ou plusieurs gènes et ouvre la voie à l'utilisation potentielle de la technologie CRISPR/CAS9 pour la sélection, pour des caractères simples et pour lesquels la nature moléculaire des allèles favorables est connue. Ces séries alléliques pour un ou plusieurs gènes facilitent déjà la création de nouveaux allèles quantitatifs (Rodríguez-Leal et al., 2017), la néodomestication, la redomestication (Pourkheirandish et al., 2020), ou la validation de QTLs dans différents fonds génétiques (Xu et al., 2016 ; Zhou et al., 2018).

2) Bien que la transformation génétique soit bien maîtrisée chez le riz, améliorer et rendre les protocoles de transformation les plus génériques possibles reste une nécessité chez la plupart des espèces végétales pour lesquelles peu de génotypes sont transformables avec une efficacité minimale

Cette réussite a néanmoins nécessité l'optimisation de la transformation chez CHHD. Il s'agit, encore une fois, du principal facteur limitant l'utilisation de ces technologies chez les plantes. Une grande partie des génotypes de riz sont accessibles à la transformation génétique. Les cultivars de la sous-espèce *japonica* peuvent être transformés par callogénèse de scutellum de grains matures (Sallaud et al., 2003, 2004), bien qu'il soit nécessaire, comme dans le cas de CHHD, d'ajuster le protocole au cas par cas. Pour la sous-espèce *indica*, il est nécessaire d'utiliser des embryons immatures (Hiei and Komari, 2006), ce qui est laborieux et rend indispensable de maintenir des plantes donneuses en serre au bon stade de floraison.

Pour la vigne, bien que des protocoles publiés existent, la transformation génétique reste longue et laborieuse. Dans le cadre de la thèse, lors des premiers tests, nous avons obtenu des plantes

transgéniques, nous n'avons pas encore réussi à générer des plantes éditées. Le fait d'avoir utilisé pour les premières expérimentations un construit CRISPR/CAS9 provenant du riz, et dont nous avons montré à posteriori, par ddPCR, qu'il avait une efficacité de zéro (**Annexe 20**), est sans doute responsable de cet échec, bien que nous ayons eu des plantes transformées. Dans quelques semaines, nous saurons probablement si nous avons obtenu des plantes éditées par CRISPR/CAS9 pour le gène *VvGAI* en utilisant des construits dicotylédones. Par ailleurs, d'autres équipes ont déjà obtenu des plantes éditées chez la vigne (Malnoy et al., 2016 ; Nakajima et al., 2017), ce qui nous permet de rester relativement optimistes pour la suite.

Au-delà du riz, la facilité de transformation reste la principale difficulté pour tirer le maximum des technologies d'édition. Cela nécessite de développer des recherches autour de l'optimisation des méthodes de transformation pour un nombre croissant d'espèces et de cultivars, en particulier pour des cultivars élites. Il a ainsi été démontré que la surexpression des facteurs de transcription GRF (GROWTH-REGULATING FACTORS) rendait les cellules et les tissus plus compétents pour la régénération pour de nombreuses espèces (Kong et al., 2020). Enfin, résultats encourageants, les taux d'édition comparés entre système transitoire et transformation stable sont systématiquement en faveur de la transformation stable. Par exemple, le base editor n'avait que 5 % d'efficacité en protoplaste, contre 55 % en stable (Herbert et al., 2020) chez Kitaaké. Ce résultat démontre qu'en transformation stable, au moins chez le riz, même si on espère des résultats similaires chez la vigne, les taux d'efficacité de construits CRISPR seront systématiquement plus élevés.

3) Générer et tester de nouveaux allèles pour l'efficacité d'utilisation de l'azote chez le riz

a) Exploration de la fonction de *OsNRT1.1b* et de *OsBT* dans le transport d'azote en plantules de riz.

OsNRT1.1b est un transporteur de nitrate à faible affinité. Il peut devenir un transporteur à forte affinité une fois phosphorylé (Plett et al., 2010 ; Wang et al., 2020). *OsBT* quant à lui est un régulateur négatif des gènes d'absorption du nitrate à forte affinité comme NRT2.1 et NRT2.4, transporteurs fortement actifs lors d'une carence en nitrate (Araus et al., 2016).

Nous avons voulu voir par complémentarité si BT régulait aussi NRT1.1b chez le riz en conditions de faible azote. Cela nous aurait permis d'avoir deux gènes interagissant entre eux pour améliorer la NUE. C'est pour cela que nous avons fait des expérimentations d'influx et d'expression avec des mutants perte de fonction KO *Osbt* et *Osnrt1.1b*, ainsi que des mutants édités à la base près dans *OsNRT1.1b*, théoriquement capables d'augmenter la NUE (Hu et al., 2015 ; Zhang et al., 2019).

Pour cela, nous avons testé deux milieux, un enrichi en KNO_3 , car le NO_3^- est un inducteur des transporteurs de nitrate en général, et un milieu enrichi en NH_4NO_3 , car le NH_4^+ est un répresseur des transporteurs de nitrate à forte affinité de la famille des NRT2 (Xuan et al., 2017).

Cela nous permettait de ne pas avoir l'impact de ces transporteurs à forte affinité dans cette condition enrichie en ammonium. L'impact du potassium (K) n'est pas pris en compte ici. D'après la bibliographie en comparant KNO₃ et KCl, il s'avère que l'acteur principal de la croissance et du phénotype des plantes est le NO₃⁻ (Blevins et al., 1974 ; Ruffel et al., 2011). Malheureusement nous n'avons pas assez de grains pour certains génotypes pour envisager un contrôle KCL en parallèle.

Les résultats des lignées Kitaakés ont été plus difficiles à analyser. En hydroponie, ils ont montré que le KNO₃ semble plus facilement assimilable, quel que soit le milieu. Difficile de conclure pour la suite car les deux WT (non transformé A, transformé B) réagissent différemment. On pourra noter que sur fort nitrate d'ammonium, les KO *nrt1.1b* n'arrivent pas à réguler l'influx et ont une tendance à l'augmenter, tandis que sur fort KNO₃, les différents mutants NRT1.1b sont moins efficaces que le sauvage non transformé.

Les analyses d'expression sont difficiles à interpréter, au vu du faible échantillonnage par génotype. Il en ressort tout de même que *OsNRT1.1b* est 10 fois plus exprimé sur nitrate de potassium que sur nitrate d'ammonium. *OsBT* est toujours exprimé, mais tend à être surexprimé quand il est muté, un résultat difficile à interpréter.

b) Impact des mutations BT KO, NRT1.1b KO, et BE sur la NUE en serre dans un fond Kitaaké : analyse des raisons de l'échec du phénotypage en serre et recommandations pour des expérimentations futures

En serre, en apportant comme source d'azote uniquement du KNO₃, les résultats obtenus sont à l'opposé de ce qui était supposé. Nous nous attendions à avoir une augmentation de la NUE comme décrite par (Araus et al., 2016 ; Hu et al., 2015) en ne mettant que le NO₃⁻ à disposition pour faciliter son absorption. Or nous avons observé une décroissance de la NUE dans les mutants *NRT1.1b* (KO et BE) en forte concentration de nitrate de potassium. De plus il y a eu un problème de répétabilité entre les deux expérimentations à cause de la différence des conditions environnementales (notamment de la lumière) malgré l'utilisation du génotype Kitaaké. Les résultats en fort nitrate y étaient similaires, mais moins forts dans la deuxième répétition biologique. Pour finir les KO *bt* ont été peu concluants car il y a deux comportements opposés entre les deux génotypes utilisés possédant des allèles KO. Un point important pour terminer est l'absence de la lignée NIL *japonica* avec le NRT1.1b-*indica* de (Hu et al., 2015) et la lignée KO *bt* de (Araus et al., 2016) dans les expérimentations comme contrôles positifs. Nous n'avons tout simplement pas pu les obtenir. Les fonds génétiques utilisés par Hu et Araus étaient différents (Nipponbare et Donjing) et nous n'aurions de toute façon pas pu les comparer à nos génotypes mutés, mais ils auraient pu servir de contrôle.

Ces résultats, malgré les précautions prises, ont été très décevants. Ils n'ont pas permis de conclure sur l'effet des mutations sur la NUE et démontrent que le phénotypage en serre reste délicat et bien sûr soumis aux fluctuations saisonnières qui en impactant la lumière, rendant illusoire une bonne répétabilité des expérimentations. Une solution aurait pu être

d'utiliser des chambres climatiques, mais on se confrontait à plusieurs problèmes. Il n'était pas possible d'analyser 240 plantes en même temps en chambre climatique. De plus, nous ne pouvions utiliser que des serres S2, ce qui nous aurait empêché de conduire ces analyses en chambre de culture, sauf en se concentrant sur quelques génotypes ou avec un effectif par génotype beaucoup plus faible et donc une puissance expérimentale plus faible. Or l'effet sur la NUE attendue n'est pas gigantesque, de l'ordre de la dizaine voire de la vingtaine de pour cent.

Dans le futur, en raison des limitations décrites, il serait plus judicieux d'analyser moins de génotypes, en conditions contrôlées en chambres climatiques. Nous avons envisagé de faire le type d'analyses réalisées sur Kitaaké pour les lignées éditées CHHD, mais les conditions de culture en serre ne sont absolument pas adaptées à la culture de cette variété d'altitude. En utilisant des chambres climatiques, nous aurions pu mener aussi des analyses pour ce cultivar, avec, toutes les limitations décrites ci-dessus. La nouvelle serre construite au CIRAD (serre Abiophen), qui dispose de trois chambres climatiques, pourrait permettre de réaliser ce type d'expériences, mais elle n'a pas encore l'agrément pour la culture de plantes transgéniques. Enfin, le choix de la source d'azote sera également à réfléchir. Si le choix d'un système semi-hydroponique semblait judicieux au départ, car permettant de contrôler plus finement la source d'azote, notre choix initial fut finalement très éloigné des conditions agronomiques où la source d'azote est très différente. Ce choix explique peut-être également les différences obtenues lors des essais de NUE au champ (cf. ci-dessous), bien que d'autres hypothèses (cultivar CHHD vs Kitaaké, et conditions de champ vs conditions en serre) soient également plausibles.

Pour des analyses plus poussées de l'effet du type d'azote, on pourrait donc envisager de tester KNO_3 et NH_4NO_3 en chambre de culture. Les chambres de cultures de la serre Abiophen sont autorégulées et programmables sur la quantité de CO_2 , l'hygrométrie, la température. En utilisant Kitaaké et Chhomrong Dhan WT non transformés, les Ko nrt1.1b, les BE NRT1.1b-T et -G, et comme sources d'azote du KNO_3 , du NH_4NO_3 , et de l'urée, cela permettrait de comparer à la fois l'effet des mutants NRT1.1b pour la NUE, pour deux génotypes et trois différents milieux azotés. Cette expérimentation, qui n'est pas réalisable au champ, mais complémentaire, permettrait de trancher sur l'effet de la source d'azote et orienterait sur l'impact et la fonction du gène *NRT1.1b* dans la NUE, dans une perspective plus fondamentale, sur la fonction de *OsNRT1.1b* dans la nutrition azotée chez le riz.

c) Impact des mutations NRT1.1b KO, BE et BT KO sur la NUE au champ dans un fond CHHD : les lignées bases éditées, possédant un allèle *indica* C>T ou C>G semblent avoir un meilleur rendement et une meilleure NUE.

Le choix du cultivar Chhomrong Dhan pour le projet GeneRice a d'abord été un choix logique puisque cette variété a réagi positivement à la transformation et est cultivée à Madagascar. Rapidement, il nous a cependant grandement compliqué la tâche. CHHD à un cycle de culture en serre long (plus de 5 mois) et nécessite des conditions de culture assez particulières, car il s'agit d'une variété népalaise qui pousse habituellement en altitude et qui

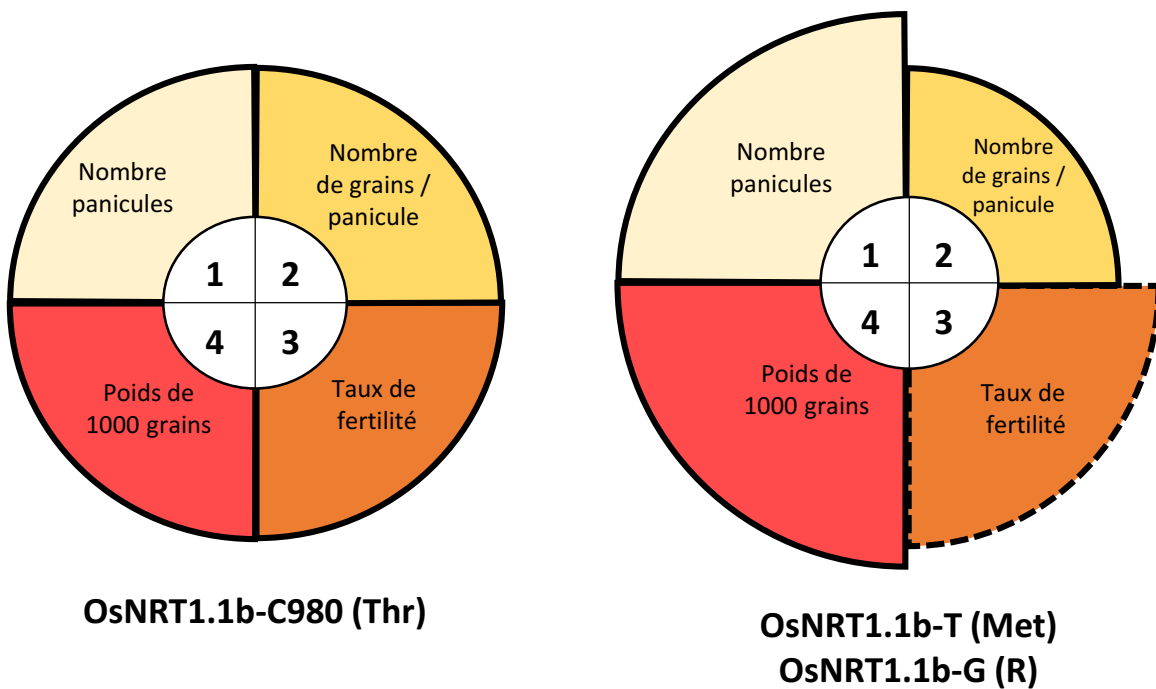


Figure 49 : Roue du rendement entre WT CHHD et les Base-Edités CHHD en Colombie

Schema descriptif représentant les différents paramètres du rendement global d'une plante. L'étape la plus précoce est le nombre de panicules produites, et la plus tardive le remplissage des grains via leur poids pour 1000.

Le WT CHHD cultivé en Colombie est présenté à gauche, tandis que les deux lignées éditées à la base près sont à droite avec une augmentation importante des étapes 1 et 4 et plus faible de l'étape 3; et une diminution de l'étape 2.

est photosensible (donc de jours courts), conditions assez éloignées de nos conditions de culture en serre. Enfin, même si nous avons réussi à le transformer, son taux de régénération est bien inférieur, d'un facteur 10, à celui de Kitaaké. Nous avons néanmoins réussi à éditer Chhomrong Dhan à la base près pour *OsNRT1.1b* à l'aide du « Base Editor ». En changeant la Thréonine 327 en Méthionine (BE-T) pour transformer l'allèle *japonica* en un allèle *indica* plus performant, ou introduire un nouvel allèle Arginine (BE-G), l'efficacité d'utilisation de l'azote, et le rendement, ont augmenté en moyenne entre les deux conditions de 14.5 % pour *OsNRT1.1bT* à 18.5 % pour *OsNRT1.1bG*. Les deux lignées alternatives (C>T et C>G), présentent-elles aussi une amélioration de la NUE l'ordre de 7 %, mais malheureusement non significatif. Cela peut s'expliquer par une puissance insuffisante (nombre de plantes analysées par génotype) pour cet essai au champ. Cela n'en demeure pas moins un résultat encourageant. Il semblerait bien que la substitution de la Thréonine soit responsable de cet effet. Néanmoins il ne faut pas négliger qu'il peut toujours y avoir eu un effet épigénétique, chromatinien, ou d'insertion du transgène dans une région codant pour un caractère non observé, qui aurait pu influencer sur le phénotype et expliquer ces différences entre allèles. Il faudra refaire un essai au champ, en n'utilisant que les allèles édités à la base près, avec un nombre plus élevé de plantes par génotype. Il est important aussi, de rappeler qu'il s'agit d'une évaluation de la NUE. Si nous avions voulu connaître la NUE totale exacte, nous aurions dû faire un dosage complet du taux d'azote dans la partie aérienne et dans les grains (Araus et al., 2016).

Les deux allèles édités à la base près (C-T et C-G) ont une meilleure NUE lorsqu'on les cultive sur une forte quantité d'urée, en plus de la quantité d'azote naturellement présente dans le sol au champ. Ils ont aussi été capables de produire plus, avec un rendement plus élevé en graines et en biomasse que le témoin. Pour finir, leur index de récolte (HI) est plus important à biomasse « égale » que celui du témoin WT lorsque l'apport d'urée est supprimé, et que la seule source disponible est celle du sol (**Figure 37**). Globalement, il semblerait donc que les conversions de ces acides aminés aient eu une influence positive sur le rendement en général.

Un point important au champ et qui est plus complexe à estimer en serre de confinement est l'impact de la rhizosphère. Il a en effet été démontré que chez *indica* la quantité de bactéries sur les racines étaient supérieure de 29 % à celle présente chez *japonica*, et que le responsable de ce recrutement n'était autre que *OsNRT1.1b* (Zhang et al., 2019). Cela pourrait donc expliquer les diverses augmentations agronomiques observées au champ avec les allèles édités à la base près. Des expérimentations de coculture en présence de bactéries bénéfiques devraient être menées pour explorer cette hypothèse.

La représentation du rendement peut se faire via une roue où 4 leviers sont modifiables au cours de la croissance des plantes (**Figure 49**). La première variable fixée correspond au nombre de panicules ; le second à la taille et au nombre de grains par panicule ; le troisième au taux de fertilité obtenue à floraison qui donnera le nombre de grains ; et le quatrième au remplissage pour avoir le poids de 1000 grains. Si nous considérons que le contrôle a pour représentation cette roue à « 100 % » (**Figure 49**), alors les lignées éditées à la base près G et T cultivées sur urée ont été influencées positivement durant l'étape 1 (effet précoce), 3 (un peu moins), et 4 (effet tardif) ; mais négativement à l'étape 2 (**Figure 49**). Il semblerait donc que les plantes BE-T ou BE-G pallient la diminution légère du nombre de grains par panicule en

produisant plus de panicules. De plus, elles ont un meilleur remplissage des grains. Il semble donc bien que *OsNRT1.Ib* joue un rôle dans le rendement chez Chhomrong Dhan. Un dernier indice nous avait déjà laissé à penser que ces lignées éditées à la base près étaient plus performantes, une analyse SPAD à floraison avait donné des résultats plus élevés chez les lignées éditées à la base près (**Annexe 13**) en condition de faible et fort azote, confirmant, que le taux de chlorophylle chez lignées éditées à la base près était plus élevé.

4) Recommandations pour le futur : combiner conditions contrôlées, conditions de champ en utilisant une source unique d'azote proche des conditions agronomiques

Dans la mesure du possible, il faudra conduire de nouvelles expérimentations au champ, pour obtenir une évaluation réelle de l'impact des mutations sur la NUE, et surtout répéter les résultats obtenus pour les lignées éditées à la base près, qui sont les plus prometteurs. L'impact de ces mutations peut être estimé entre 7-15 % d'effet positif sur la NUE et le rendement. Plus globalement, c'est un effet considérable obtenu de plus chez une variété considérée comme rustique, qui a une croissance déjà bonne sur des milieux pauvres en azote lorsqu'elle est cultivée sur les hauts plateaux malgaches.

5) Vers l'utilisation de la technologie CRISPR/CAS9 chez la vigne : outils de prototypage, constructions CAS9 optimisées et production de plantes éditées pour le gène *VvGAI*

En optimisant la production et la transformation de protoplastes, nous avons réussi à évaluer l'efficacité de plasmides CRISPR pour la vigne et obtenir une première estimation des taux d'édition. En parallèle, nous avons obtenu des premiers résultats qui orienteront le choix de construits optimisés pour l'édition de gènes d'intérêt agronomiques chez la vigne. Nous avons également réussi à transformer et régénérer des embryons de vigne, malheureusement non éditées. Nous avons tenté de gagner du temps en utilisant un construit fonctionnel en riz, car les embryons de vigne sont longs à régénérer. Ce premier résultat nous a permis de confirmer la fonctionnalité du protocole de transformation et nous a rassurés sur sa faisabilité. Ce résultat, combiné avec les premiers résultats obtenus de comparaisons de constructions CRISPR/CAS9 optimisées vigne (**Figure 43 et Figure 47**) a confirmé l'intérêt de tester en amont la fonctionnalité des construits CRISPR/CAS9 en transformation transitoire en utilisant la ddPCR, validant l'ensemble de la chaîne de prototypage. En effet, nos résultats préliminaires indiquent qu'en transformation transitoire, ce plasmide riz, n'est pas fonctionnel (**Annexe 20**), permettant aussi de confirmer que l'approche que nous avons suivie (constructions de nouveaux vecteurs vigne) était une nécessité. Même si préliminairement le vecteur comportant le promoteur *VvU6* et la *CAS9* d'*Arabidopsis* semble le plus efficace dans ces travaux préliminaires (**Figure 47**), il faudra répéter ce résultat pour confirmer le classement des combinaisons testées. En parallèle, nous allons lancer des expériences de transformation stable avec ces constructions.

Après plus d'un an de régénération et l'utilisation de glutamine pour stimuler la régénération, les premiers embryons commencent à apparaître, au moment de la rédaction de ce manuscrit, retardant la production de plants de vigne édités. Plusieurs raisons à ce retard. Le premier est le stade des cals utilisés (**Annexe 17**). Peut-être étaient-ils trop jeunes au moment où nous les avons transformés. La comparaison entre les cals embryogènes de Chardonnay et de Portan et ceux obtenus à partir du Grenache est parlante. Seuls les cals issus du cultivar Grenache ont survécu. Le temps passé en sélection sur l'hygromycine, plusieurs mois au lieu d'un seul, a probablement bloqué la formation d'embryons au stade protos-cals (**Figure 45**, photo en bas à droite) et inhibé leur différenciation. Alternativement, on peut aussi faire les hypothèses que la plupart des cals seraient non transformés et donc bloqués sur hygromycine, ou que la mutation KO de *VvGAI* ralentirait extrêmement le développement embryonnaire. Cette dernière est peu probable car certaines mutations n'affectent pas le cadre de lecture, et tous les embryons mis en culture ne sont probablement pas mutés.

L'ajout de la glutamine a débloqué cette phase de stase et permis l'apparition de proto-embryons. Nous pourrions faire des tests supplémentaires en complétant ces milieux avec d'autres acides aminés comme de la proline. Elle est par exemple indispensable pour des espèces difficiles à régénérer comme *Miscanthus* (Yoo et al., 2018). Dans le futur, le choix du génotype sera critique (Gribaudo et al., 2017). En effet, et c'est un résultat très classique, certains génotypes réagissent très peu à la transformation et à la régénération. Le système RNP est une alternative intéressante à la transformation de cals embryogènes (Malnoy et al., 2016), à la condition de pouvoir régénérer des protoplastes, mais également si le taux d'édition est important, pour éviter d'avoir un cribler un grand nombre de plantes transformées pour identifier la mutation recherchée.

Dans quelques semaines ou mois, nous saurons si nous avons des plantes éditées chez la vigne, ce qui nous permettrait de conclure sur la fonctionnalité complète du système mis en place chez la vigne pour tester et produire des plantes éditées pour des gènes d'intérêt agronomiques.

III] Discussion globale

La question principale, initiale, de la thèse était de répondre à la question :

« Est-il possible d'utiliser la technologie CRISPR pour intégrer des caractères d'intérêt agronomiques ? »

La réponse est « oui », chez le riz, probablement chez la vigne, et potentiellement chez toutes les espèces accessibles à la transformation stable. Ces résultats ne sont pas complètement nouveaux. En effet, chez le riz, introduire des mutations à effet agronomique a été réalisé pour le rendement comme vu au préalable (Xu et al., 2016). Les outils CRISPR permettent d'aller plus loin, par exemple en générant de la variabilité allélique à effet quantitative dans un seul fond génétique (Rodríguez-Leal et al., 2017), ou domestiquer *de novo* des plantes (Lemmon et al., 2018 ; Zsögön et al., 2018). Il semble donc possible de domestiquer des riz sauvages (Lacchini et al., 2020; Li et al., 2016). Avec l'avantage d'utiliser différents sgRNA, différentes CAS9 reconnaissant différents PAM, utiliser la CPF1, il serait possible de modifier en une seule génération un grand nombre de gènes d'intérêt. Donc en conclusion, clairement, et sans ambiguïté, la réponse à la question initiale de la thèse est « oui, il est possible d'utiliser la technologie CRISPR pour intégrer des caractères d'intérêt chez le riz et la vigne ».

Notre apport majeur, dans le cadre de cette thèse est d'avoir contribué à faciliter l'exploitation de l'outil CRISPR/CAS9 pour introduire des mutations à effet agronomique, notamment chez la vigne. Un apport important pour la vigne a aussi été de développer la méthodologie d'obtention de protoplastes. Le système rapide de test que nous avons choisi basé sur les protoplastes a ensuite donné de très bons résultats en peu de temps. Couplé à la ddPCR il permet d'analyser rapidement un grand nombre de constructions. Le transfert de ces technologies sur la vigne n'a cependant pas été simple. Nous avons rencontré de nombreux problèmes comme la répétabilité complexe de la transformation transitoire. L'utilisation de la ddPCR combinée à l'utilisation de protoplastes, en particulier avec la technologie QIAcuity dans le futur, permettra de tester à la fois l'efficacité et la spécificité de construits CRISPR/CAS9. Cela permettra d'optimiser en amont les construits à utiliser chez de nombreuses espèces comme la vigne pour lesquelles la régénération est délicate et laborieuse. Il est en effet désormais possible de tester 5 fluorochromes en même temps, et c'est ce que nous sommes en train de tester sur les protoplastes ayant été transformés avec le plasmide d'édition à la base près dans *OsNRT1.1b* afin de déterminer précisément les taux de mutations spécifiques et non spécifiques à un locus donné.

Trois gènes d'intérêts agronomiques, *OsNRT1.1b*, *OsBT*, *VvGAI* ont été ciblés dans la thèse. Si inactiver ces trois par introduction de frameshift, ne posait pas de problème technique, l'introduction de mutations à la base près, en utilisant le « Base Editor » (nCAS9BE-sgRNA), au début de ce projet, de plus dans une variété exploitée au sud et jamais modifiée avant (CHHD), représentait un vrai défi qui a été relevé malgré les difficultés. La technologie multiplexe permettrait de cibler plusieurs locus simultanément, ce qui augmente l'intérêt de cette technologie si on souhaite l'utiliser comme outil complémentaire de la sélection classique. De plus cette méthode a un intérêt décuplé pour la vigne où les cycles de transformation sont

longs et laborieux (**Figure 38 et 45**). D'autres technologies sont disponibles et seront disponibles dans le futur. Il est possible d'« éteindre », ou d'« allumer » un gène sans même affecter sa séquence par CRISPR interférence ou CRISPR activation en ciblant directement le promoteur d'un gène d'intérêt (Li et al., 2020a). Le remplacement allélique par HDR (Meynard et al., 2020; Svitashv et al., 2016), un des objectifs initiaux de cette thèse, n'est toujours pas maîtrisé, en tout cas pas en routine avec des taux réalistes, bien que des travaux soient encore en cours pour en améliorer l'efficacité.

Chez la vigne il faudra maîtriser la reproductivité de la transformation transitoire en protoplastes, et essayer de l'augmenter comme chez le riz. Les résultats obtenus sont néanmoins, déjà très encourageants et ont permis l'utilisation des NBT pour tester un caractère d'intérêt. La transformation stable et la régénération seront aussi des points à retravailler malgré la réussite d'un premier essai.

Doit-on et peut-on utiliser la technologie CRISPR/CAS9 en sélection ?

Une dernière question, la plus importante sans doute, est : « Doit-on utiliser la technologie CRISPR/CAS9 comme outil de sélection ? ». Question à subdiviser en deux sous-questions :

- Est-il rentable, en terme financier par rapport à la sélection classique d'utiliser CRISPR/CAS9 ?
- Au vu des conséquences juridiques et de propriété intellectuelle que va entraîner l'utilisation de cette technologie, est-il éthique et acceptable par la société d'utiliser cette technologie en sélection ?

Pour la première question, la rentabilité en termes de rapport coût/bénéfice dépend d'un troisième facteur, réglementaire et juridique. Si on se positionne dans un cadre légal européen, les plantes éditées sont considérées comme des OGM, décision de la Cour Européenne de Justice (2018), et soumise au long, complexe, et coûteux processus réglementaire d'autorisation des OGM (Sprink et al., 2016). Il faudrait, en sus des coûts de production des plantes transgéniques (plusieurs milliers d'euros), démontrer qu'il n'y a pas d'autres mutations introduites (plusieurs milliers d'euros pour le reséquençage du génome des plantes à inscrire), conduire des essais pour démontrer l'innocuité sanitaire des grains produits (plusieurs dizaines de milliers d'euros), réaliser des essais au champ dans des conditions de confinement (plusieurs dizaines de milliers d'euros). Cette liste, loin d'être exhaustive, démontre que le coût serait extrêmement élevé (plusieurs millions d'euros au final) et il faudrait de longues années (au moins une dizaine) pour qu'elle puisse être peut-être un jour effectivement cultivée, annihilant l'intérêt de la technologie pour gagner du temps sur l'introduction par sélection d'un allèle favorable. Dans ce cadre réglementaire, la réponse est clairement « non ». Si on se positionne dans un cadre juridique comme celui des USA, de la Colombie ou même de la Chine, la situation est complètement différente. Si les plantes éditées ne contiennent plus de T-DNA, elles ne sont pas soumises à la réglementation OGM. Dans ce cadre, alors la réponse à la question initiale est « oui », d'autant plus s'il est possible de cibler plusieurs locus en même temps.

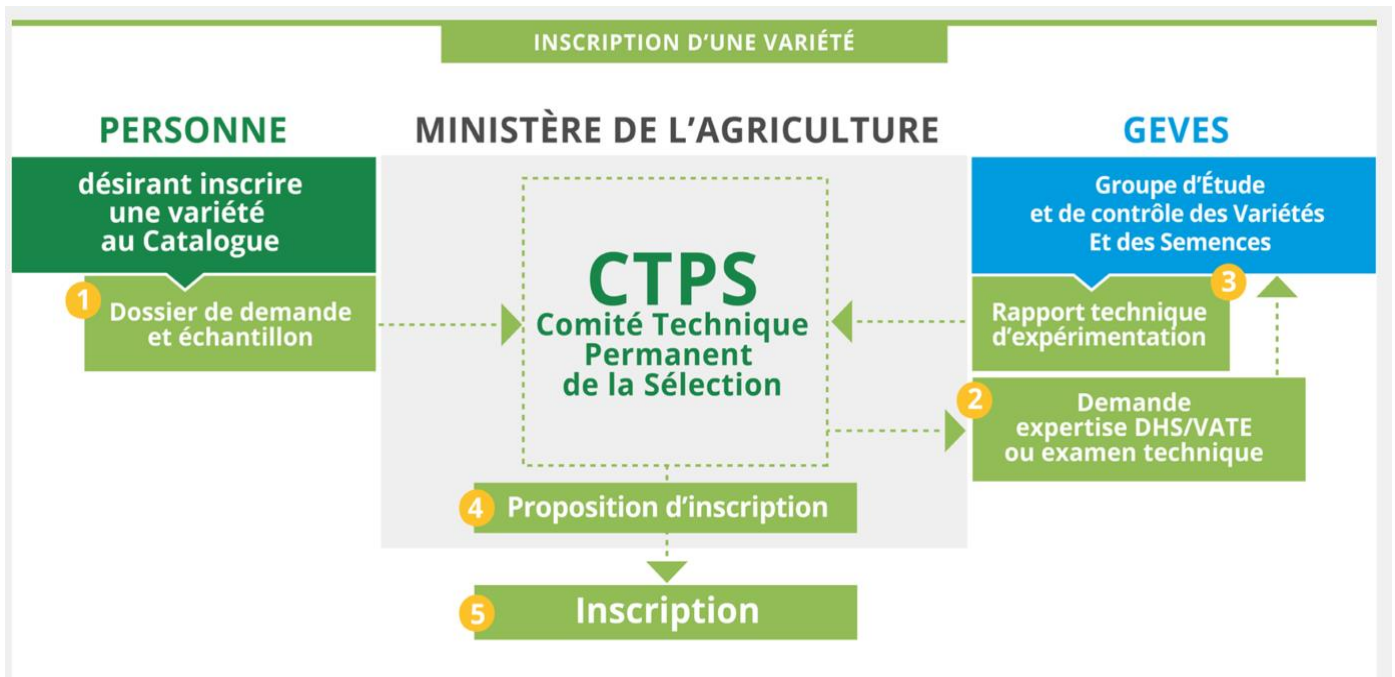


Figure 50 : Diagramme d'inscription d'une nouvelle variété

Afin qu'une nouvelle variété puisse être utilisée, un dossier doit être monté et présenté devant un recours d'experts qui validera ou non son inscription au catalogue.

Pour la deuxième question, « au vu des conséquences juridiques et de propriété intellectuelle que vont entraîner l'utilisation de cette technologie, est-il éthique et acceptable pour la société d'utiliser cette technologie en sélection ? », la réponse est plus nuancée, mais défavorable tout de même. Les OGM font l'objet d'un rejet sociétal, pour des raisons qui, pour certaines sont parfaitement légitimes et recevables, et dépendent de l'interlocuteur.

Ainsi, une variété, pour être inscrite au Catalogue Officiel français des espèces et variétés de plantes cultivées, et être commercialisable doit passer les tests DHS (elle doit être Distincte, Homogène et Stable) et les tests VATE, prouvant ses Valeurs Agronomique, technologique et environnementale. (<https://www.gnis-pedagogie.org/inscription-varietes-catalogue-officiel/#quels-sont-les-tests-pour-linscription>).

Le système de protection des variétés quant à lui repose en Europe sur le COV (Certificat d'obtention végétale). Le COV protège spécifiquement les nouvelles variétés végétales, dès lors qu'elles sont distinctes, homogènes et stables (DHS). Les variétés protégées par le COV peuvent être utilisées par des sélectionneurs pour développer leurs propres variétés. Ce système permet à tous de profiter collectivement du progrès génétique global pour une espèce, y compris pour développer des variétés au sud. Théoriquement, il est possible aussi à des agriculteurs de multiplier les semences pour leur propre utilisation, même si cette clause est inutilisable dans le cas de variétés hybrides, qui dominent le marché des semences.

Si une plante est éditée, elle peut aussi relever du système des brevets. Ce dernier ne peut porter que sur des plantes — et non pas sur la variété — ou sur la technique qui a permis d'en caractériser certaines propriétés. Il n'est alors plus question de réutilisation, ou de multiplication de semences, car il s'agirait du piratage. Les conséquences seraient potentiellement graves, pour la sélection et pour les agriculteurs, en particulier du sud qui n'auront pas les moyens de payer les royalties demandées pour une variété brevetée et seront obligés de se fournir chez un semencier, à chaque nouveau semis. En terme plus direct se poserait rapidement la question de l'appropriation du vivant par les obtenteurs des variétés (**Figure 50**).

Enfin, se posera également, dans le cadre des brevets et de l'introduction d'allèles favorables, leur origine, et le « benefit-sharing » (<https://www.cbd.int/abs/infokit/revised/web/all-files-en.pdf>). Le « benefit-sharing » se réfère à la manière dont des ressources génétiques (allèles favorables dans ce cas) ont été obtenues et du partage des bénéfices qui auront été produits. La plupart, si ce n'est tous, des allèles favorables obtenus ont été identifiés par clonage de QTL, génétique d'association, à partir de ressources génétiques d'origines diverses, très souvent d'ailleurs, provenant de pays ou de régions du monde situées au sud, ou exploitant des connaissances traditionnelles. Si des variétés produites en exploitant ces ressources sont commercialisées, alors un partage de ces bénéfices doit se faire avec les pays, les groupes qui ont identifié, mis en évidence ou exploité ces ressources génétiques (**Figure 51**). Il s'agit de questions très complexes avec des intérêts financiers à la clef potentiellement très importants et des intérêts opposés par exemple entre des compagnies semencières et les détenteurs de ces ressources génétiques.

Pour finir, ces technologies nécessitent des laboratoires de biologie moléculaire et cellulaire de pointe et renforcent les inégalités entre le sud et le nord pour leur appropriation. Dans le cadre

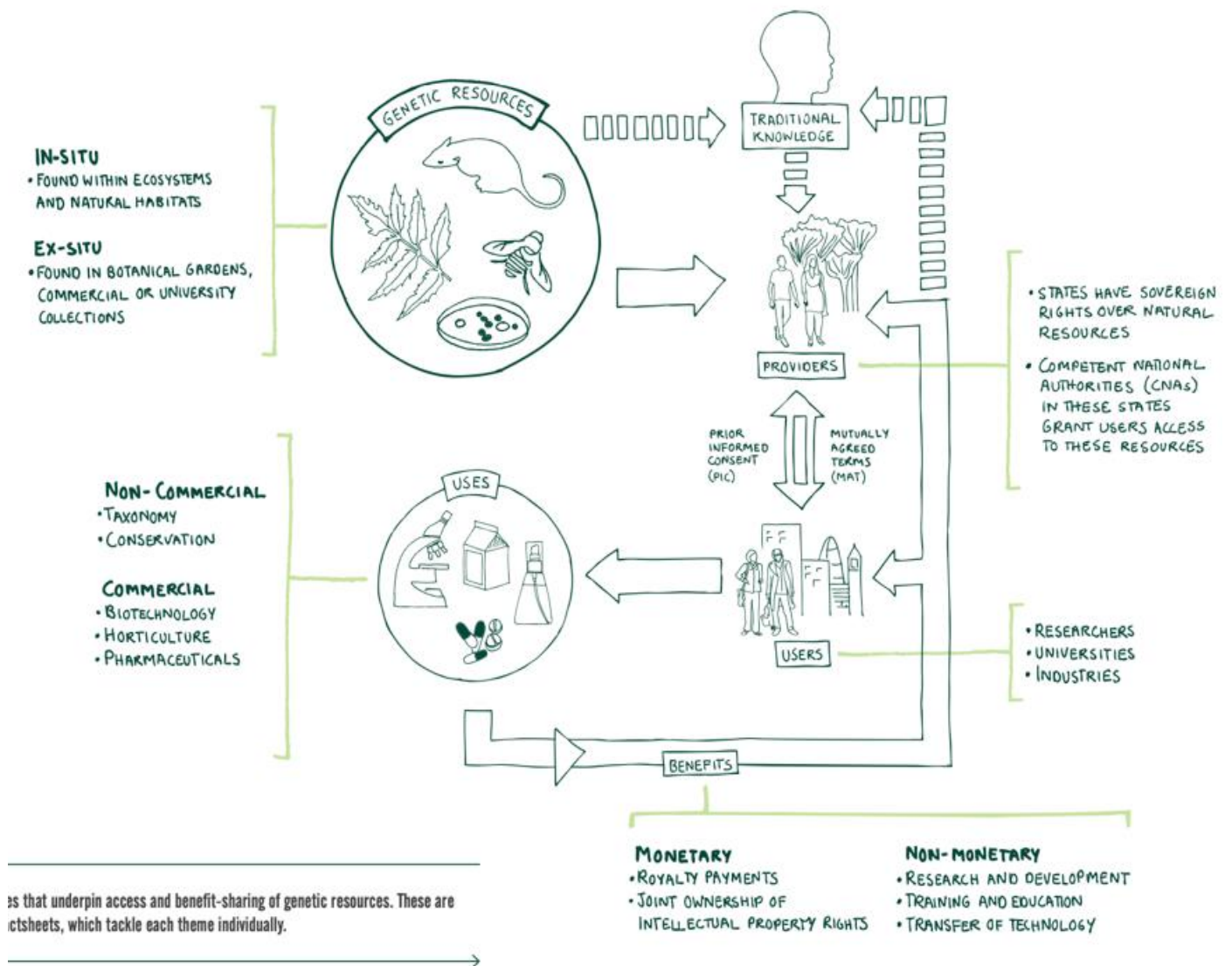


Figure 51 : Diagramme théorique des avantages du « Benefit-Sharing »

Le benefit-sharing se réfère à la manière dont des nouvelles ressources génétiques (allèles favorables dans ce cas) ont été obtenues et du partage des bénéfices qui auront été générés entre les utilisateurs et les fournisseurs.

Tiré de <https://www.cbd.int/abs/infokit/revised/web/all-files-en.pdf>.

de cette thèse, qui s'appuyait sur un projet Agropolis GENERICE, ce travail a permis d'initier des discussions avec le laboratoire (Responsable Professeur Tendro) qui s'occupe de la surveillance et l'autorisation — ou non — de cultures de plants OGM à Madagascar. L'idée a été de fournir des éléments pour permettre l'identification de plants de CHHD édités et d'éclairer sur une future mise en place de conditions permettant à Madagascar d'exercer sa pleine souveraineté sur l'utilisation et le déploiement — ou pas — de plantes éditées sur son sol.

Les technologies d'édition par CRISPR sont donc des outils formidables pour étudier la fonction de gènes d'intérêt agronomique chez les plantes cultivées, et elles sont aussi importantes pour la validation de polymorphismes d'intérêt. Avant ce travail, par exemple, nous n'avions aucune idée si l'allèle *indica* introduit dans CHHD pourrait avoir un effet favorable sur la NUE pour ce cultivar. Nous avons couplé cela à l'utilisation de nouvelles techniques pour accélérer le temps de travail et d'études avec succès. La transposition dans une espèce à cycle long ne s'est pas faite sans problèmes, mais au final l'intérêt du développement de toutes ces technologies et stratégies est encore plus évident pour les plantes pérennes (forestières, fruitières ou à multiplication végétative).

Matériel et Méthodes

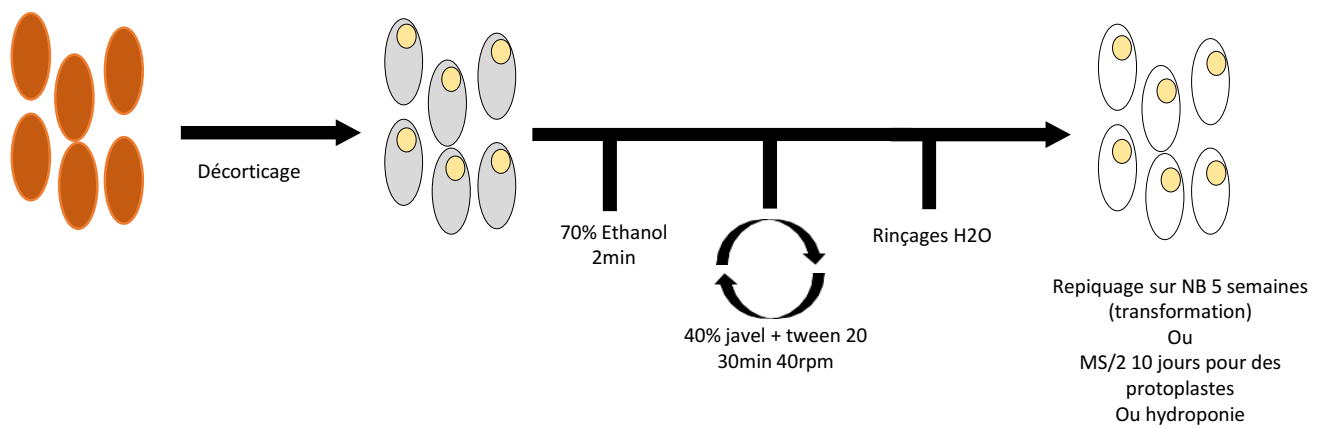


Figure 52 : Stérilisation des grains de riz

Les grains sont écosés, passés à l'éthanol, puis incubés sous agitation 30min avec de la javel. Une fois rincés à l'eau, ils peuvent être utilisés pour les transformations stables, pour la production de protoplastes, ou pour la culture en hydroponie.

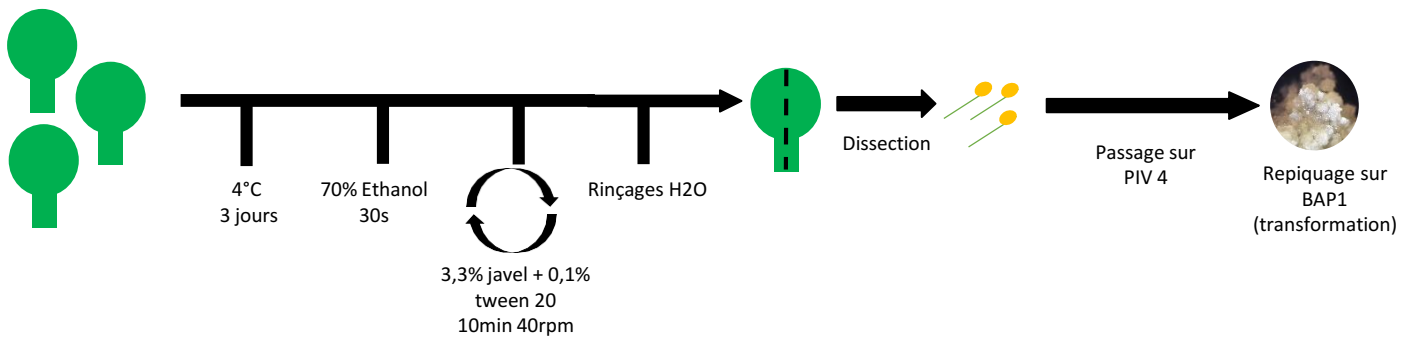


Figure 53 : Stérilisation des anthères de vigne

Les bourgeons floraux sont récoltés et mis 4°C trois jours, avant d'être stérilisés à l'éthanol et à la javel. Une fois rincés ils sont disséqués et les anthères sont prélevées afin d'être mises sur un milieu PIV 4 semaines pour la production de cals, Les cals sont repiqués sur un milieu BAP1 pour la transformation.

I] Matériel végétal et culture

Riz : la variété Kitaaké, *Oryza sativa* L. ssp. *japonica* à un cycle court, est insensible à la photopériode, se transforme très facilement et son génome a été entièrement séquencé (International Rice Genome Sequencing, 2005). Elle a donc été choisie comme modèle pour générer rapidement des mutants par édition du génome. Nous avons également utilisé en parallèle, dans le cadre du projet GENERICE (Agropolis Fondation), une lignée d'origine népalaise cultivée sur les hauts plateaux malgaches, *Oryza sativa* L. ssp. *japonica* Chhomrong Dhan (CHHD, Raboin et al., 2010). Elle a un cycle long, est plus délicate à transformer et à manipuler, mais représente une variété d'application agronomique dans ce travail de thèse. Les mutants CRISPR dans la NUE (Nitrogen Use Efficiency) ont donc été produits en parallèle chez Kitaaké et Chhomrong Dhan. Les mutants iSTOP sur *OsCAOI* ont été produits uniquement chez Kitaaké. La variété Nipponbare, *Oryza sativa* L. ssp. *japonica* a été utilisée pour produire des protoplastes à partir de jeunes plantules cultivées *in-vitro*, pour les tests d'efficacité de ddPCR.

Vigne : *Vitis vinifera* cv Merlot et Cabernet Sauvignon ont été utilisées en cultures cellulaires pour leur bonne réactivité et leur production de cellules embryogènes. *Vitis vinifera* cv Portan, Chardonnay et Grenache ont été utilisés pour produire des cals embryogènes et proviennent aussi des cultures et collections de l'INRAE.

1) Stérilisation des semences et des explants

Riz : Les grains sont désinfectés avant leur mise en culture en hydroponie, en culture *in-vitro* pour la production de protoplastes, ou avant l'étape d'induction pour la transformation stable de grains matures. Sous hotte stérile, ils sont décortiqués pour retirer la glume et la glumelle, puis ils sont mis en contact 2 min avec de l'éthanol 70 % sous agitation. Ils sont ensuite incubés 30 minutes à 40 rpm sous agitateur rotatif (argos rotoflex r2001) dans une solution de 50 mL d'eau de Javel 40 % (v : v) (9.6 % chlore) à laquelle sont ajoutées 4 gouttes de tween 80 % (Sigma P4780). Les grains sont ensuite rincés en condition stérile 5 fois à l'aide de 30 mL d'eau distillée stérile puis mis en culture en hydroponie ou *in-vitro* (cf plus bas) (**Figure 52**).

Vigne : Pour la transformation stable de la vigne, des inflorescences immatures (petits bourgeons floraux de 2-3 mm) sont prélevées au champ ou en serre (INRAE, campus de

Supagro à Montpellier). Elles sont ensuite désinfectées 2 min avec de l'éthanol 70 % puis pendant 15 min dans une solution de 10 g/l NaDCC, 0.1 % tween 20. En début de désinfection, les bourgeons sont soumis rapidement deux fois 20 secs à un vide partiel afin de faciliter leur stérilisation. Puis, ils sont rincés quatre fois avec de l'eau distillée autoclavée. Les anthères et leurs filaments sont récupérés par dissection des bourgeons au scalpel et déposés par 30 dans une boîte de Petri (55 mm) sur un milieu d'induction de cals embryogènes contenant BAP (9 μ M) et 2-4 D (4.5 μ M) (**Annexe 21**) pendant 4 semaines. Les bourgeons sont repiqués sur le même milieu tous les mois (**Figure 53**). Les cals induits sont ensuite maintenus jusqu'à leur utilisation en transformation stable sur un milieu de maintien contenant BAP (1 μ M) et 2-4 D (5 μ M) (voir **Annexe 21**).

2) Conditions en chambre de culture

Riz : Pour la production de protoplastes, des plantules de riz sont cultivées à l'obscurité en chambre de culture (27.2 °C, 25 % hygrométrie) dans des boîtes de pétri 20x20 (Corning 431301) sur un milieu de 200 mL de MS/2 (Duchefa, M0221.0050), avec 2.15 g de MS low salt et 51.5 mg de vitamines MS pour 1 L, solidifié avec 4.5 g de phytigel (CAS 71010-52-1) en condition stérile (cf. protocole de désinfection). Les boîtes sont ensuite incubées verticalement à l'obscurité à 27 °C pendant 8 à 11 jours.

Pour les expérimentations de physiologie, les riz sont placés en bassine de 5 L de milieu hydroponie (**Annexe 22**) avec 12h jour, 12h nuit, chambre LED 29 °C jour et 25 °C nuit avec 70 % d'humidité pendant 4 semaines.

Vigne : Les cellules sont cultivées à 26 °C sous agitation (120 rpm) en milieu liquide MS/2 avec 1 μ M BAP et 1 μ M NAA à pH 5.7. Elles sont ensuite repiquées toutes les 2 semaines pour garder une bonne réactivité de multiplication. La chambre de culture présente une température constante de 26 °C, et une photopériode de 14h de jour. Les cellules en culture sont utilisées pour la production de protoplastes.

3) Serre de confinement

Riz : Les plantules ou les graines sont mises à germer dans des pots de 3 L. Les compositions des terreaux utilisés sont différentes pour les deux cultivars utilisés :

Pour Kitaaké terreau 2 (Tref2 : 28 % tourbe blonde fine, 30 % tourbe blonde irlandaise moyenne, 18 % fibres coco, 10 % perlite n° 2, 9 % sable volcanique, 5 % argile, pH=4,5, Ec=0,2) plus l'engrais Basacote plus 6M (Compo Expert) 16-8-12 NPK (16 % azote, 8 % P₂O₅, 12 % K₂O, et 2 % MgO + 5 % soufre) à 2 g/L.

Chhomrong Dhan (CHHD) se cultive de deux manières différentes :

– Avec le terreau 9 (70 % tourbe blonde, 30 % tourbe noire de sphaignes) complété de 1L de pouzzolane 7-15 mm/L (pour 14L de terreau), et 2.5 g/L d'engrais Basacote plus 6M (Compo Expert) 16-8-12 NPK (16 % azote, 8 % P₂O₅, 12 % K₂O, et 2 % MgO + 5 % soufre). Ce mélange était utilisé historiquement au CIRAD pour les multiplications traditionnelles. Initialement nous avons voulu conserver ce terreau afin d'éviter toute carence.

– Avec le terreau 2 (Tref2) plus l'engrais Basacote high K (Compo Expert) 13-5-18 NPK (13 % azote, 5 % P₂O₅, 18 % K₂O, et 2 % MgO + oligo-éléments) à 1.5 g/L + 1.5 g/L de Sifor organic fertilizer. Lors des multiplications, nous nous sommes aperçus que les plantes ne présentent pas de carences apparentes. C'est le mélange utilisé maintenant avec CHHD.

Les plantes en pot sont ensuite transférées en serre de confinement, posées sur des tablars avec en permanence une lame d'eau de 3 cm. Les conditions en serre sont : 28 °C le jour et 24 °C nuit. L'hygrométrie est à 60 %, l'éclairage d'appoint est à 700 μmol/m²/s en plus de la lumière naturelle.

Environ 40 jours après floraison, les plantes sont mises à sécher pendant une semaine et les panicules matures sont récoltées et mises à sécher individuellement dans des enveloppes à 42 °C 48 h dans une étuve. Elles sont ensuite stockées en chambre froide à 4 °C.

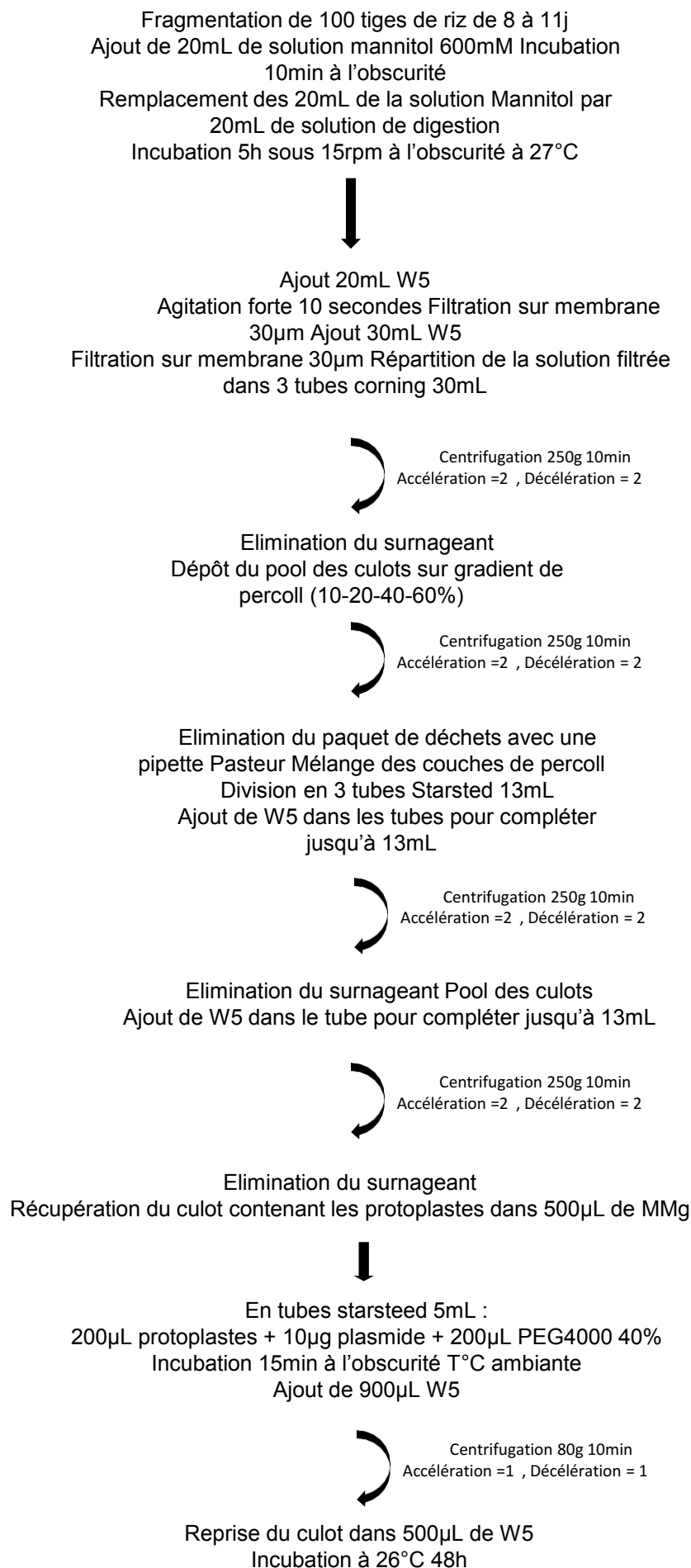


Figure 54 : Workflow de la production et transformation de protoplastes de riz

Description rapide du protocole de production de protoplastes de plantules de riz développé au sein de l'équipe DAR.

II) Méthodes de transformations

1) Transformation transitoire du riz

La production et transformation de protoplastes de riz a été décrite et optimisée dans l'article (Bès et al., 2021) (**Figure 54**) (**Annexe 23**). Brièvement, une centaine de jeunes pousses de riz cultivées en conditions stériles pendant 8 à 11 jours sont hachées puis mises en contact 10 min avec 20 mL d'une solution de Mannitol (600 mM mannitol, 2 mM MES, pH=5.7). Après avoir retiré la solution de mannitol, le macérat est digéré dans un volume de 20 mL comprenant de la Macérozyme et de la Cellulase pendant 5 h sous agitation à 15 rpm (0.75% Macerozyme R10 Onozuka, 1.5% Cellulase RS Onozuka, 0.1% BSA, 7.5 mM CaCl₂ 2H₂O, 600 mM Mannitol, 10 mM MES, pH=5.7).

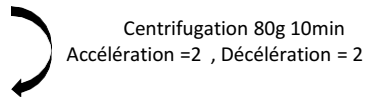
Un volume de 50 mL de solution de lavage W5 (154 mM NaCl, 125 mM CaCl₂ 2H₂O, 5 mM KCl, 2 mM MES, pH=5.7) est ajouté à la solution de feuilles digérées. Deux filtrations successives de cette solution sont réalisées à travers une toile de nylon avec des mailles de 0.30 µm. Les protoplastes sont centrifugés à 250 g 10 min avec une faible accélération/frein dans des tubes corning 30 mL (Accélération/Décélération = 2/2). Ces conditions de centrifugation seront utilisées pour toutes les purifications de protoplastes. Après avoir éliminé le surnageant, 20 mL de W5 sont ajoutés, et une nouvelle centrifugation est effectuée. Le surnageant est éliminé et le volume restant contenant les protoplastes culotés est déposé sur un gradient de percoll réalisé à l'aide de solutions percoll-W5 60 %, 40 %, 20 %, 10 %, à raison de 2 mL par couche en tube Starsted 13 mL. Une nouvelle centrifugation est effectuée afin de regrouper et d'éliminer les déchets entre les couches 20 % et 40 %. Après avoir complété le tube avec la solution de lavage W5, une nouvelle centrifugation permet de rincer les protoplastes et éliminer le percoll. Le culot est récupéré et centrifugé une dernière fois avec 13 mL de W5. Les protoplastes culotés et purifiés sont récupérés et mis en contact une heure avec une solution de MMG (400 mM mannitol, 15 mM MgCl₂ 6H₂O, 4 mM MES, 5 mM CaCl₂ 2H₂O, pH=5,7) pour être transformés. La concentration de protoplastes par mL est estimée à l'aide d'une cellule de Fuchs et Rosenthal. Les protoplastes sont alors dilués dans le MMG pour obtenir une concentration entre 1x10⁶ et 4x10⁶ protoplastes vivants par mL.

La transformation se fait avec 200 µL de PEG à 40 % (40 % PEG4000 Fluka, 200 mM mannitol, 100 mM CaCl₂ 2H₂O, pH=5,7) auxquels est ajouté un volume égal de protoplastes et

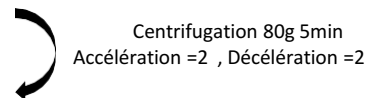
25mL de culture cellulaire
Ajout de 25mL de de digestion enzymatique
Incubation 14h sous 25rpm à l'obscurité à 25°C



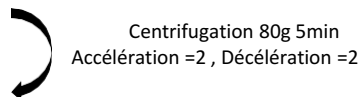
Répartition de 5x10mL en tubes 12mL
Ajout 1mL W5 délicatement dessus (2 phases
visibles)



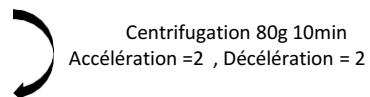
Récupération des coussins de protoplastes à
l'interface des 2 couches en tubes 12mL.
Ajout de 10mL de W5.



Elimination du surnageant et pool des culots
Ajout de 10mL de W5.



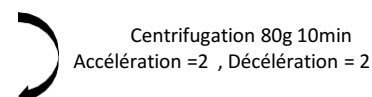
Reprise du culot dans 2mL MMg
A conserver pendant 2h à 4°C



Elimination du surnageant
Récupération du culot contenant les
protoplastes dans du MMg pour une
concentration de 4.5×10^6 protoplastes par mL
maximum



En tubes starsteed 5mL :
200µL protoplastes + 20µg plasmide + 200µL PEG6000 40%
Incubation 1h à l'obscurité à 27°C
Ajout de 900µL W5
Incubation 10min à l'obscurité T°C ambiante



Reprise du culot dans 450µL de W5
Incubation à 25°C 48h

Figure 55 : Workflow de la production et transformation de protoplastes de vigne

Description rapide du protocole de production de protoplastes de plantules de vigne développé à l'INRA par Mme Lauvergeat, réajusté par Mme Laucou et Mr Herbert. D'après la base de Hichri et al., (2010).

10 µg d'ADN plasmidique. Le plasmide pHKN29 est utilisé comme contrôle de transformation (**Annexe 1A**). pHKN29 est un plasmide binaire dérivé du plasmide pCAMBIA1300, qui contient un promoteur 35S suivi de la séquence de la sGFP (Hirotaka Kumagai and Hiroshi Kouchi, 2003). Le PEG, l'ADN plasmidique et les protoplastes sont mis 15 min en contact dans le noir, puis 900 µL de la solution W5 sont ajoutés. Les protoplastes sont alors culotés par centrifugation. Le surnageant est éliminé et les protoplastes sont resuspendus dans 500 µL de W5.

Pour les protoplastes de riz, les taux de transformation pour pHKN29 oscillent entre 60 % et 90 % (80 % en moyenne) au bout de 42 h suivant la période et le génotype mesurés en estimant le nombre de protoplastes GFP sur le nombre total de protoplastes vivants à l'aide d'une cellule de Fuchs et Rosenthal. La détection des mutations se fait 48 h après la transformation. Un volume de 50 µL de tampon 10x RQ1 et 50 µL de DNase RQ1 (Promega) est ajouté aux 500 µL de protoplastes transformés. L'ensemble est incubé pendant 30 min à 26 °C pour éliminer l'ADN provenant de protoplastes morts durant l'étape de purification. Un volume de 50 µL de solution Stop RQ1 est ajouté, puis les protoplastes sont culotés à 500 g pendant 3 min. Une fois le surnageant éliminé, une extraction d'ADN est effectuée sur le culot de protoplastes à l'aide du kit DNeasy Plant Minikit (Qiagen, 69104), à partir de l'étape d'ajout du buffer de lyse à 65 °C. Le volume final d'extraction se fait dans 150 µL de tampon AE.

2) Transformation transitoire de la vigne

La production et la transformation de protoplastes de vigne ont été décrites par Mme Virginie Lauvergeat (INRAe Bordeaux) à partir d'un protocole de Hichri et al., (2010) et optimisé avec l'aide de Mme Valérie Laucou (DAAV INRAe) (**Figure 55**) (**Annexe 24**). Cinquante mL de culture cellulaire de Merlot sont centrifugés 5 min à 3100 rpm. Les cellules culotées sont mises en contact avec 50 mL d'une solution enzymatique (0.6% Macerozyme R10 Onozuka, 1.2% Cellulase RS Onozuka, 0.4 M Saccharose, 0,319 % Macroéléments Gamborg B5 medium, pH=5.7) pendant 14h à 25 rpm à 26 °C.

Les protoplastes sont répartis dans 5 à 6 tubes de 12 mL. Puis 1 mL de solution W5 (154 mM NaCl, 125 mM CaCl₂·2H₂O, 5 mM KCl, 5 mM glucose, pH=5.7) sont délicatement déposés à la surface pour effectuer une séparation des phases par centrifugation à 80 g pendant 10 min (Accélération/Décélération = 1/1). Ces conditions seront conservées pour toutes les

centrifugations de protoplastes. Un coussin de protoplastes apparaît ensuite à l'interface des deux phases. Les coussins formés dans chaque tube sont prélevés dans de nouveaux tubes de 12 mL. Dix millilitres de W5 sont ajoutés et mélangés, afin de refaire un nouveau lavage par centrifugation pendant 5 min. Les surnageants sont éliminés, les culots assemblés, et 10 mL de W5 sont ajoutés pour réaliser un deuxième lavage par centrifugation pendant 5 min.

Pour préparer la transformation, les protoplastes sont repris dans 2 mL de MMG riz (400 mM mannitol, 15 mM MgCl₂ 6H₂O, 4 mM MES, 5 mM CaCl₂ 2H₂O, pH=5.7) et mis à 4 °C pendant 2 h. Ils sont ensuite resuspendus dans une solution de MMG à une concentration de 4.5x10⁶ protoplastes par mL. La transformation est adaptée du protocole riz et se fait avec 200 µL de PEG à 40 % (40% PEG6000 Fluka, 0.4 M mannitol, 100 mM Ca(NO₃)₂, pH=5.7) à volume égal de protoplastes, mais avec 20 µg de plasmide. En tant que témoin, 10 µg d'un plasmide fluorescent 35S:GFP (pHKN29) est aussi utilisé. Le PEG et les protoplastes restent 60 min en contact dans le noir à 25 °C avant de recevoir 900 µL de W5. Après 10 min d'incubation à l'obscurité à température ambiante, ils sont culotés par centrifugation pendant 5 min. Le surnageant est éliminé et les protoplastes sont resuspendus dans 500 µL de W5, puis les tubes sont incubés à l'horizontale à 26 °C.

Le taux de transformation de ce plasmide oscille entre 20 et 70 % au bout de 42 h suivant la période. Une influence négative de la température en hiver a été observée par simple comptage sur cellule de Fuchs et Rosenthal, avec en moyenne 25 % de transformation, alors qu'en été cette moyenne est autour de 50 %. La détection des mutations se fait 48h après la transformation. Les 500 µL de protoplastes sont mis en contact avec 50 µL de tampon 10x RQ1 et 50 µL de DNase RQ1 (Promega) pendant 30 min à 26 °C pour éliminer l'ADN provenant de protoplastes morts durant l'étape de purification. 50 µL de Stop RQ1 sont ajoutés, puis les protoplastes sont culotés à 100 g pendant 3 min. Une fois le surnageant éliminé, une extraction d'ADN est effectuée à l'aide du kit DNeasy Plant Minikit (Qiagen, 69104), à partir de l'étape d'ajout du buffer de lyse à 65 °C et un prétraitement avec 1 % de PVP-40 et la RNase du kit (8 µL) pendant 15 min. Le volume final d'extraction se fait avec dans 150 µL de tampon AE.

3) Transformation stable du riz

La transformation stable du riz est réalisée à partir du protocole de (Sallaud et al., 2003) par l'équipe transformation, et décrite en **Figure 56** et les solutions sont décrites en **Annexe 25**.

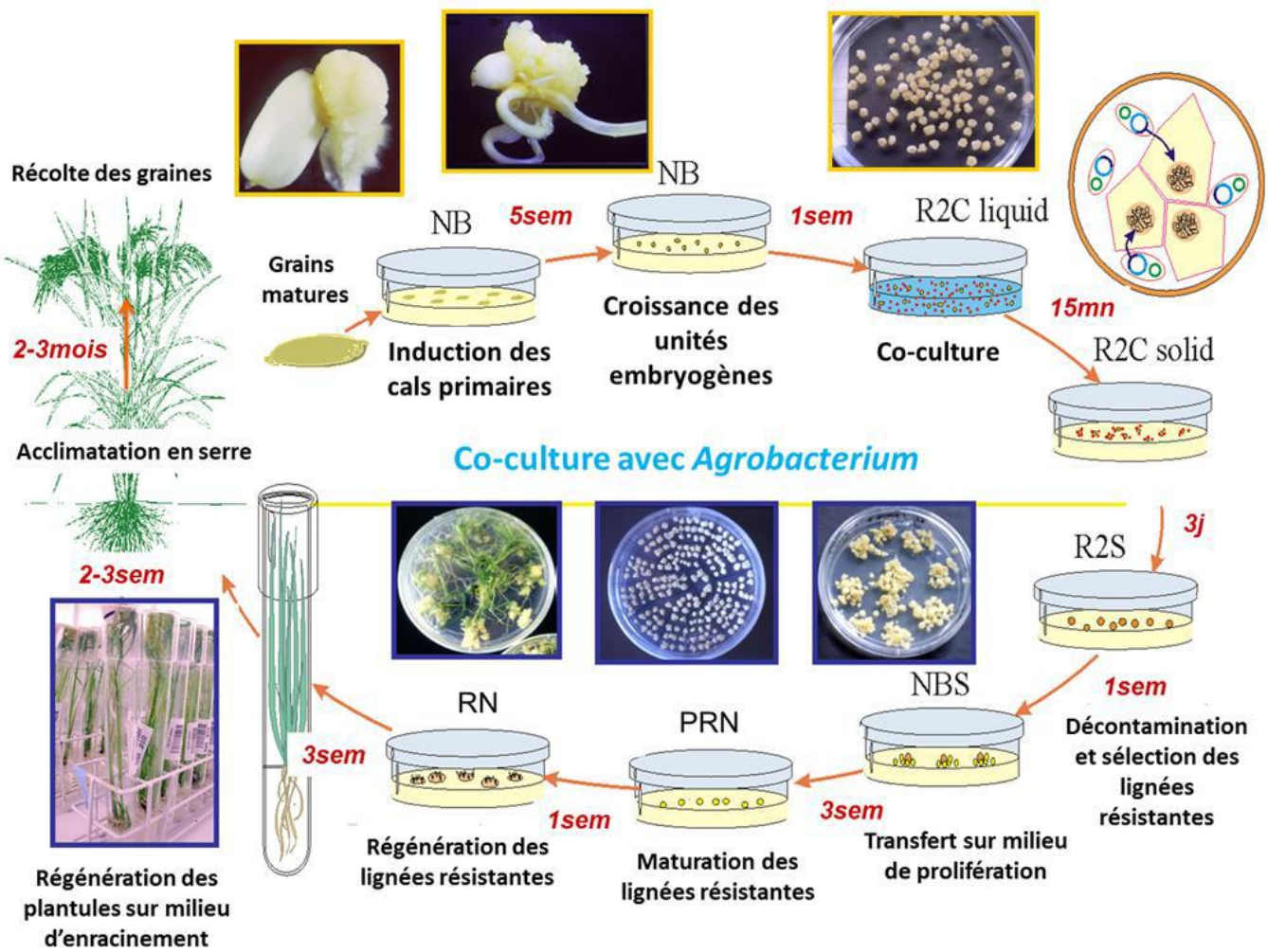


Figure 56 : Transformation stable du riz par Agrobacterium

Photos des étapes de transformation et régénération du riz. Réadaptée de la publication de Sallaud et al. 2003.

Pour induire la formation de cals, une dizaine de grains décortiqués et désinfectés (cf. conditions de cultures stériles ci-dessus) sont placés sur le flanc sur du milieu NB en boîte de Petri pendant 5 semaines en chambre de culture (obscurité, 28 °C, hygrométrie 60 %). Un cal primaire se forme alors à partir du scutellum et va libérer des cals secondaires. Ces cals secondaires portent de très nombreux embryons. Les cals secondaires embryogènes, ronds jaunes et compacts sont sélectionnés et mis par 50 sur du milieu NB en boîtes de Petri pendant 7 à 10 jours (obscurité, 28 °C, hygrométrie 60 %). Trente à cinquante unités embryogènes de 3 à 5 mm sont ensuite utilisées par construit pour la transformation par *Agrobacterium tumefaciens*.

Les agrobactéries doivent être virulentes au moment de la mise en co-culture. Les colonies de souches EHA105 ont préalablement reçu les vecteurs à transformer (T-DNA avec CRISPR et gène de résistance antibiotique en plante par exemple). Elles sont cultivées dans 10 mL de milieu LB — Rifampicine (25 mg/L) — Spectinomycine (50 mg/L, antibiotique de sélection à adapter suivant le vecteur) en tube 50 mL, à 28 °C sous agitation à 120 rpm toute la nuit. Deux boîtes de Petri avec 20 mL de milieu AB solide, sontensemencées avec 200 µL de culture liquide pour chaque construction, puis mises à incuber 3 jours à 28 °C. À l'aide d'une spatule, le tapis bactérien est récolté et les bactéries sont resuspendues par vortex dans 30 mL d'un milieu R2-CL. La densité optique (DO), mesurée à 600 nm est ensuite ajustée à 0.01 en ajoutant le volume de R2-CL nécessaire.

La transformation est alors réalisée en plongeant les 30 unités embryogènes dans les 30 mL de R2-CL contenant les agrobactéries pendant 15 min dans un tube de 50 mL. Les cals sont ensuite séchés successivement sur 6 feuilles de papier Wattman pour retirer l'excédent d'agrobactéries. Dix cals séchés sont ensuite déposés par boîte de Petri contenant 20 mL de milieu R2-CS solide, et incubés à 25 °C pendant 3 jours à l'obscurité. Ils sont ensuite transférés sur milieu R2S contenant comme antibiotique, l'hygromycine et un bactériostatique, la céfotaxime (200 µg/mL), pour contrôler la propagation des agrobactéries, pendant une semaine à 28 °C à l'obscurité. Les cals avec une couverture visqueuse, due à un surdéveloppement d'agrobactéries ou nécrosés sont éliminés.

Les cals sont alors transférés sur milieu NB-S pendant 10 jours (28 °C, obscurité) pour la dernière étape de sélection. Les boîtes sont aérées 15 minutes à l'issue de ces 10 jours afin de favoriser la croissance des amas globuleux qui représentent les zones résistantes en prolifération. Les boîtes sont remises en culture une semaine de plus dans les mêmes conditions, puis les nodules qui ont bourgeonné sont étalés autour du cal co-cultivé.

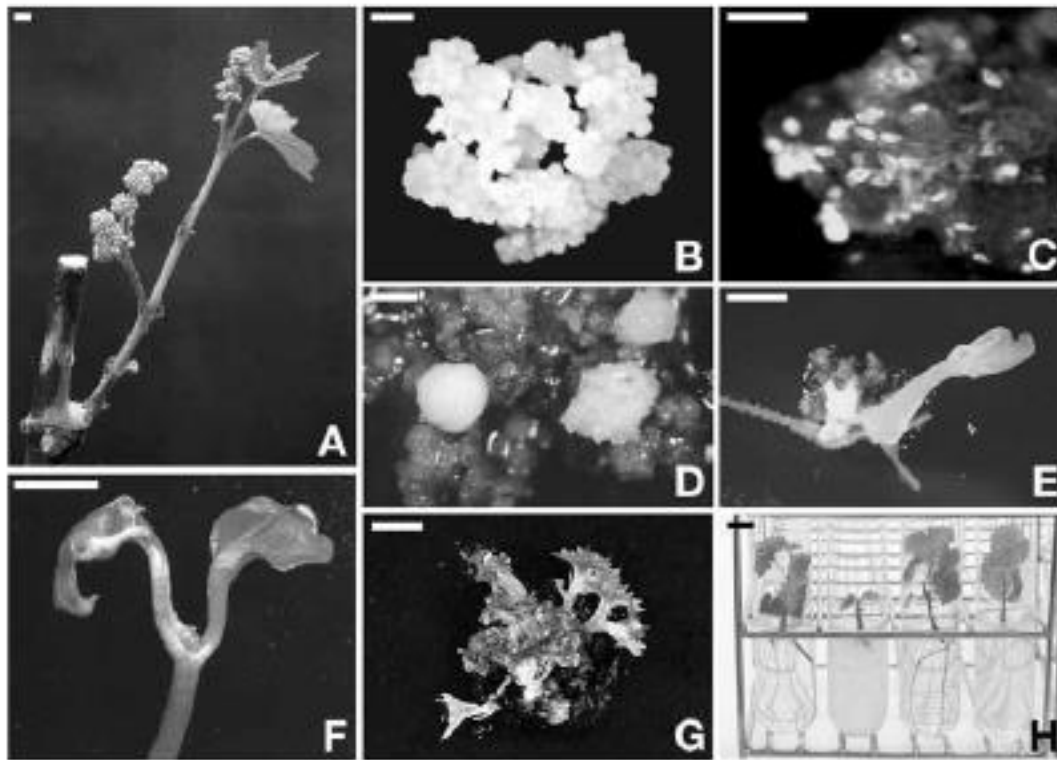


Fig. 1. Successive stages of the transformation procedure in grapevine. (A) Inflorescence with immature flower buds produced on cutting and ready for anther culture; bar = 5 mm). (B) Embryogenic calli grown on GSICA medium prior to co-cultivation; bar = 1 mm. (C) Bright clusters of *gfp*-expressing cells observed 5 wk after co-cultivation; bar = 50 µm. (D) Clusters of kanamycin resistant embryogenic cells visually selected among dead tissue two months after co-cultivation; bar = 1 mm. (E) Germination of a kanamycin resistant embryo on MG1 medium; bar = 5 mm. (F) Well-developed embryo before trimming of cotyledons and roots; bar = 5 mm. (G) Growth and axillary branching of the embryo apical meristem on BFe2 medium with 50 µg/mL kanamycin; bar = 5 mm. (H) Rooting of transformed plantlets on GNBC medium with 50 µg/mL kanamycin. *Note:* Second plant from the left is an untransformed plantlet.

Figure 57 : Transformation stable de la vigne par *Agrobacterium* – Tiré de Bouquet et al., 2006
 Photos des étapes de transformation et régénération de la vigne. Planche photo issue de la publication de Bouquet et al. 2006.

Les cals résistants sont transférés sur le milieu de prérégénération en boîte de Petri sur milieu PRN avec hygromycine (50 µg/mL) pendant 7 jours (28 °C, obscurité). Les embryons somatiques se développent et les cals sont alors transférés, sept par boîte de Petri sur milieu de régénération RN pendant 2 jours à 28 °C, à l'obscurité. Les boîtes RN sont ensuite placées en chambre de culture à la lumière pour stimuler la germination des embryons somatiques (110-130 mM/mPAR; 500 µEm⁻²s⁻¹) pendant 3 semaines (28 °C). Les plantules sont alors transférées en tubes à essai avec 13 mL de milieu d'enracinement MS (50 g/L saccharose, 2.6% phytigel, pH=5.8). Après deux semaines sur milieu d'enracinement, elles sont ensuite acclimatées à la serre en les laissant une nuit, tube ouvert avant leur transfert en pot.

4) Transformation stable de la vigne

La transformation stable de la vigne se base sur l'adaptation du protocole de Bouquet et al., (2006), et décrite **Figure 57**. Pour préparer les expérimentations de transformation les milieux solides sont répartis par 10 mL en boîte de pétri de 55 mm de diamètre, et ceux avec du GS1Ca sont toujours complétés avec du charbon actif. L'ensemble des solutions sont décrites en **Annexe 21**.

Des cals issus de cultures d'anthères de Grenache, Portan ou Chardonnay sont obtenus après 4 semaines de culture sur le milieu d'induction de cals embryogènes, puis maintenus jusqu'à la transformation sur un milieu de maintien de cal embryogènes. Ils sont transférés sur un milieu d'induction d'embryons contenant 1 µM BAP et 10 µM NOA (GS1Ca + Charbon actif) (6 demi-cals par boîte) à 28 °C à l'obscurité pour la formation des pro-embryons pendant quatre semaines.

Les agrobactéries contenant les constructions à transférer sont mises en culture dans du LB-Spectinomycine (50 mg/L) (résistance fournie par le plasmide) — Rifampicine (25 mg/L) comme pour le riz, à 26 °C sous 120 rpm sur la nuit. Elles sont centrifugées 10 min à 3000 g et resuspendues dans des erlenmeyers avec 100 mL de milieu liquide Induction Medium Agrobactéries puis incubées 2 h à 100 rpm à 26 °C. Elles sont finalement centrifugées une dernière fois 10 min à 3000 g et resuspendues dans du milieu LCM à une DO de 0.4 mesurée à 550 nm (soit entre 30 et 50 mL de LCM par construit).

La coculture est réalisée avec 15 mL de la suspension bactérienne en milieu LCM et 10 à 35 cals en tubes Falcon de 50 mL. Les tubes sont agités vigoureusement à la main pendant 2 à 3 s puis

les cals sont incubés 10 min à 25 °C (temps extrêmement important, ne pas excéder 12 min de coculture). Les cals et fragments de cals sont récupérés par filtration à l'aide d'une cupule de 100 µm (Falcon cell strainer 100 µm), séchés ensuite sur du papier Wattman pour retirer l'excédent de liquide. Pour finir, les cals sont récupérés dans la cupule à l'aide d'un scalpel puis déposés sur boîtes de Petri contenant du CM-medium (GSC1Ca + 100 µM Acétosyringone). Les boîtes sont mises à sécher 5 min sous hotte puis laissées 3 jours à 28 °C à l'obscurité.

Les proembryons non contaminés par les agrobactéries sont rincés avec 20 mL de milieu LCM additionné d'Augmentin (concentration finale à 200 mg/L) en tube Falcon de 50 mL, et la gélose est séparée des proembryons lors du rinçage. Les proembryons sont récupérés dans une cupule de 100 µm séchée brièvement avec du papier Wattman. Chaque proembryon est prélevé avec un scalpel et déposé sur boîtes de Petri contenant du GS1Ca additionné d'Augmentin (200 mg/L) et de Céfotaxime (200 mg/L) pour 3 semaines de culture dans le noir à 28 °C.

Les minuscules proembryons sont transférés sur boîtes de Petri avec du GS1Ca contenant de l'Augmentin (200 mg/L), Céfotaxime (200 mg/L) et Hygromycine (25 mg/L) et repiqués tous les mois, au moins 3-5 fois sur le même milieu à 28 °C à l'obscurité. Puis les repiquages se font sur du milieu GS1Ca sans antibiotique avec ajout de Glutamine à 2 mM (292.3 mg pour 1L). Au fur et à mesure de leur développement, une part significative des proembryons vont blanchir et former de petits embryons munis d'une « pseudo » racine et de deux « pseudo » cotylédons. Ils vont alors être transférés par dix en boîte magenta avec un milieu RIM (Bouquet et al., 2006) lorsqu'ils atteignent ½ centimètre ils sont placés en chambre de culture à 25 °C à une intensité lumineuse de 15 µ Em⁻²s⁻¹. Ils sont individualisés en tube à essais sur un milieu MS/2 contenant 6.5 g de MicroAgar (Duchefa, M1002) lorsqu'ils atteignent deux centimètres et qu'ils possèdent des feuilles bien individualisées.

Plasmide	Enzyme	Séquence flanquante (rouge) à rajouter aux 20pb
pENTRY simplex 217 (un seul sgRNA, promoteur U3 riz)	BsaI	(5'-3') F : GGCA ATGC(...)TTTT (3'-5') R : AAAC AAAA(...)GCAT
pENTRY multiplex 289 (deux sgRNA, promoteur U3 riz)	BtgZI	(5'-3') F : TGGT ATGC(...)TTTT (3'-5') R : AAAC AAAA(...)GCAT
	BsaI	(5'-3') F : GTGT ATGC(...)TTTT (3'-5') R : AAAC AAAA(...)GCAT
pENTRY pEn.C1-1 et pEn-Chimera simplex dicotylédones (un seul sgRNA, promoteur AtU6-26).	BbsI	(5'-3') F : ATTG ATGC(...)TTTT (3'-5') R : AAAC AAAA(...)GCAT
pEN-Manticora multiplex dicotylédones (deux sgRNA, deux promoteurs AtU6-26).	BsmB1 et Bbs1 idem	(5'-3') F : ATTG ATGC(...)TTTT (3'-5') R : AAAC AAAA(...)GCAT

Echange de la cassette
AttL-sgRNA-AttL du
pENTRY
avec la cassette
AttR-ccdB-
chloramphénicol-AttR du
pDESTY

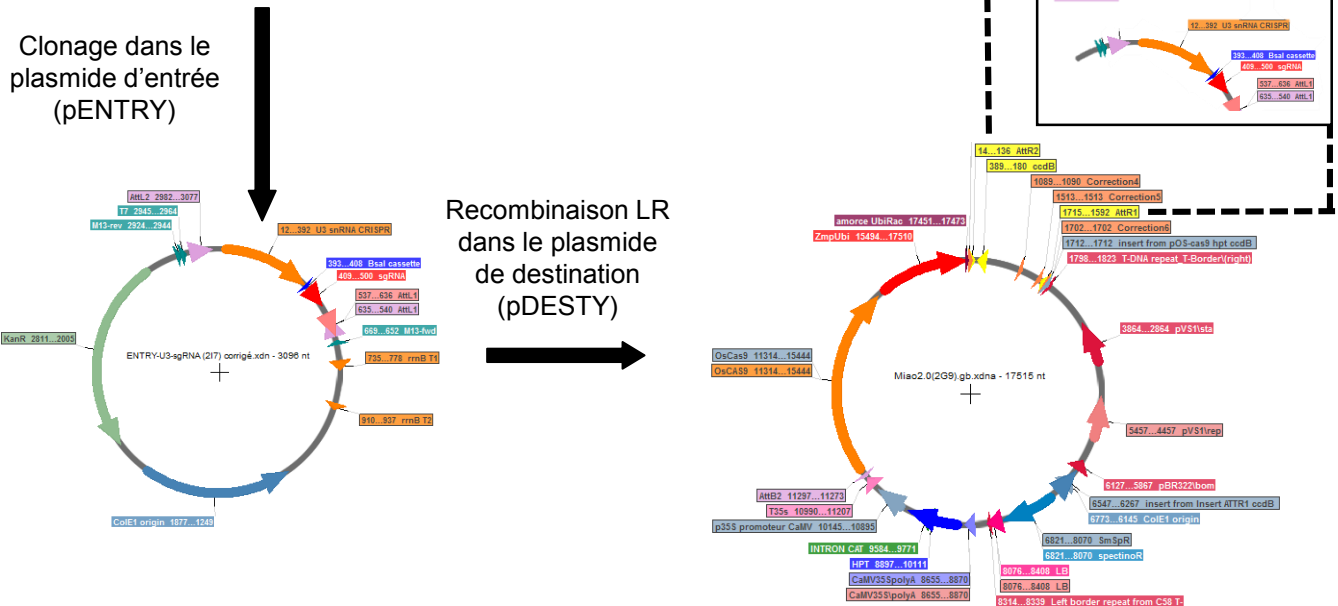


Figure 58 : Principe du clonage des CRISPR et recombinaisons

Les crRNA de 20pb sont dessinés et complétés pour être clonés dans les plasmides pENTRY au début du sgRNA. Ces plasmides sont ensuite recombinaisonnés via des sites AttR-AttL dans les plasmides de destination avec notamment la séquence de la CAS9. Il s'agit de la stratégie utilisée au laboratoire.

III] Analyses moléculaires : ADN, ARN, plasmides, influx

Les dosages se font au Nanoquant 200 pro Tecan pour l'ADN, ou en puces RNA (RNA screen tape Agilent, 5067) ou au TapeStation (4200 Agilent Technologies) pour l'ARN.

1) Clonages en plasmides CRISPR

Les clonages en plasmides CRISPR des crRNA se font en suivant le protocole décrit par Bès et al., (2021). Pour le riz et la vigne, seuls les vecteurs d'entrée et de destination diffèrent. Les différentes étapes sont brièvement présentées en **Figure 58**.

Les séquences CDS des locus ciblés sont récupérées à partir de (NCBI <https://www.ncbi.nlm.nih.gov/> , OrygenesDB <https://orygenesdb.cirad.fr/>) pour le riz, ou (NCBI, Genoscope http://www.genoscope.cns.fr/cgi-bin/blast_server/projet_ML/blast.pl) pour la vigne. La sélection des cibles ou protospacers du CRISPR se fait à l'aide du logiciel CRISPR Tefor (<http://crispor.tefor.net/>). Un crRNA de 20 nucléotides est choisi à partir de deux critères : un nombre de coupures hors site (ou off-Targets) le plus faible possible et la plus grande dissimilarité dans la région en 3' du PAM (NGG) (le « seed ») de ses séquences hors site. Des oligonucléotides forward (F) et reverse (R) correspondant aux spacers auxquels sont ajoutés les demi-sites de clonage pour l'insertion dans les vecteurs d'entrée ; on ajoute 4 nucléotides en 5' sur le crRNA-F, et 4 autres nucléotides en 5' sur le crRNA-R pour le cloner dans le vecteur d'entrée (**Annexe 15**). Ces nucléotides dépendent du pENTRY utilisé et donc de l'enzyme de coupure utilisée. Ils sont décrits en **Annexe 14** :

- BsaI pour le pENTRY simplex 2I7 (un seul sgRNA, promoteur U3 riz, résistance Kanamycine (50 µg/mL)) (**Annexe 1E**).
- BtgZI et BsaI pour le pENTRY multiplex 2B9 (deux sgRNA, promoteur U3 riz, résistance Kanamycine (50 µg/mL)) (**Annexe 1H**).
- BbsI pour le pENTRY pEn.C1-1 et pEn-Chimera simplex dicotylédones (un seul sgRNA, promoteur AtU6-26, résistance Ampicilline (100 µg/mL)) (**Annexe 1C**).
- BsmB1 et BbsI pour le pEN-Manticora multiplex dicotylédone (deux sgRNA, deux promoteurs AtU6-26, résistance Ampicilline (100 µg/mL)) (**Annexe 1D**).

Les oligos F et R, 2 µg chacun, sont ajoutés dans un volume de 50 µL (qsp H₂O) et dénaturés 5 min à 95 °C. Ils sont hybridés en laissant le bloc refroidir pendant 1h.

La digestion des plasmides d'entrée a lieu à raison de 10 µL de plasmide à 100 ng/µL, 1 µL d'enzyme, 2 µL de tampon (Cutsmart), 7 µL H₂O, pendant 1 h 30 à 37 °C. Cas particulier du 2B9 (multiplex riz) : 3 h 30 à 60 °C pour BtgZ1. Les vecteurs sont déphosphorylés avec de l'Antartic Phosphatase (NEB, M0289L) en ajoutant aux 20 µL digérés 2.5 µL de tampon AP et 1 µL d'Antartic Phosphatase. La déphosphorylation a lieu pendant 1 h 30 à 37 °C suivie d'une étape d'inactivation de 5 min à 65 °C. Huit µL d'oligonucléotides F/R hybridés sont phosphorylés avec 1 µL de la T4-PNK (NEB, M0201S), 1.5 µL T4 buffer, 1 µL ATP 20 mM, 3.5 µL H₂O, pendant 1 h à 37 °C.

La ligation se fait sur la nuit à température ambiante en utilisant 2.5 µL de plasmide linéarisé, 6 µL d'oligos hybridés, 2 µL de tampon T4-ligase, 1 µL de T4-ligase, et 8 µL d'H₂O. Dans une dernière étape, 1 µL d'enzyme de digestion est alors ajouté et le mélange est incubé 1h à la température optimale de l'enzyme (**Annexe 15**). Toutes les transformations et multiplications de bactéries *E. coli* utilisent le protocole suivant : Deux µL du produit de ligation sont mélangés à 25 µL de bactéries chimiocompétentes DH5-alpha pendant 30 min à 4 °C. Les bactéries sont transformées par choc thermique à 37 °C pendant 1min30, puis laissées 2 min sur glace. Cinq cents µL de LB liquide sont ajoutés et les bactéries sont incubées 1 h 30 à 37 °C sous agitation à 225 rpm. Cent µL sont déposés sur une boîte de pétri avec un milieu solide LB + antibiotique correspondant au plasmide, et incubé une nuit à 37 °C. Huit colonies sont sélectionnées et inoculées dans 2 mL de LB liquide + antibiotique pendant 12 h. Les plasmides sont extraits à l'aide du kit (Wizard, A1330 miniprep), dosés et vérifiés par restriction avec l'enzyme utilisée pour le clonage. Les clones positifs, qui ont intégré le spacer, ont perdu leurs sites de restriction (BsaI par exemple). Un séquençage est alors effectué en utilisant le primer M13-R (**Annexe 26**).

Le clonage de la séquence du sgRNA complet se fait par recombinaison LR (Invitrogen, 11791020) dans un plasmide binaire de destination compatible avec le système gateway. Le(s) sgRNA(s) et son/leurs promoteur(s) sont clonés dans :

- un plasmide binaire qui code pour la protéine CAS9 codon optimisé riz sous contrôle du promoteur ZmUbi de maïs (modifié de Miao et al., 2013) : 2G9 (monocotylédones), résistance Spectinomycine (50 µg/mL) (Bès et al., 2021; Herbert et al., 2020) (**Annexe 1i**).

Un plasmide binaire pBE-nCas9 qui permet de faire du base-editing C>T avec un promoteur ZmUbi de maïs contrôlant l'expression de la CAS9 céréale, pBE-nCAS9— attR modifié (monocotylédones) (Herbert et al., 2020), résistance Kanamycine (50 µg/mL) (**Annexe 1J**).

- Un plasmide binaire pour dicotylédones pDE_Cas9 avec une Cas9 codon optimisée pour *Arabidopsis thaliana* avec son promoteur PcUbi4-2 et terminateur pea3A du persil (pour pEn.C1.1 et pEn-Chimera) (dicotylédones) (Fauser et al., 2014; Herbert et al., 2020), résistance Spectinomycine (50 µg/mL) (**Annexe 1B**).

Un volume de 1.5 µL du plasmide binaire de destination à 150 ng/µL est ajouté à 1.5 µL du vecteur d'entrée pENTRY contenant le spacer cible, à 4 µL de TE (Tris-EDTA) pH=8, et à 1.5 µL de LR clonase (ThermoFisher 11791019). La réaction de clonage a lieu 3h à 25 °C. La réaction est stoppée par ajout de 1 µL de Protéinase K (Thermofisher, EO0491) et incubation à 37 °C pendant 10 min. Les plasmides sont transformés, multipliés et sélectionnés dans des bactéries *E. coli* DH5-alpha. Les plasmides sont extraits en utilisant le kit miniprep (Wizard, A1330). Ils sont vérifiés par restriction à l'aide de deux enzymes (voir les cartes pour la sélection) (**Annexe 1A à J**). Un séquençage est aussi effectué avec un primer spécifique (**Annexe 26**).

Après validation des vecteurs binaires contenant les sgRNA des cibles définies, les plasmides sont multipliés dans 200 mL de culture LB + antibiotique de résistance. Des stocks glycérols de chaque construction sont réalisés, en combinant une solution de 500 µL de bactéries et 500 µL d'une solution de glycérol à 80 % stérile. Ils sont alors stockés à -80 °C. Une maxiprep est réalisée en parallèle en utilisant le kit Qiagen HiSpeed 12662. L'ADN plasmidique est stocké à -20 °C jusqu'à son utilisation pour les expériences de transformation de protoplastes.

Pour les transformations stables en riz et vigne, des agrobactéries (*Agrobacterium tumefaciens*) EHA-105 sont transformées par électroporation avec 1 µL du plasmide binaire et 20 µL de bactéries électrocompétentes EHA-105. Les plasmides sont insérés par électroporation (kV=1,8 ; ms=5,7). Cinq cents µL de LB liquide sont ajoutés et les bactéries sont incubées 1h30 à 28 °C sous agitation à 225 rpm dans un tube de 13 mL. Cent µL sont déposés sur une boîte de pétri contenant un milieu solide LB + antibiotique et Rifampicine (25 µg/mL). Les boîtes sont incubées deux jours à 28 °C.

2) Extraction d'ADN végétal

Le kit DNeasy plant (Qiagen, 69104) est utilisé pour extraire l'ADN des plantules, pour de petits effectifs (<15) ou pour extraire l'ADN des protoplastes chez le riz ou la vigne.

Le protocole a été adapté pour la vigne : la lyse à 65 °C a lieu avec un traitement 1 % PvP (polyvinylpyrrolidone) et la RNase du kit (8 µL). Le temps de lyse est de 15 min pour l'extraction d'ADN de protoplastes, et d'une heure pour l'extraction d'ADN à partir de plantes matures.

Une extraction au MATAB est effectuée lorsque le nombre de plants de riz à analyser est supérieur à 15, par exemple lors du génotypage des plantes issues de transformation génétique. Cent milligrammes d'une jeune feuille, soit un morceau de 3 cm, sont immédiatement congelés dans l'azote liquide dans un rack comprenant des tubes eppendorf de 1.2 mL (colonne de 8, plaque 96, Qiagen) contenant une bille en acier (2 mm). Ils sont ensuite broyés au broyeur à bille (Genogrinder) (fréquence = 30, 30 s, deux fois). La poudre est resuspendue dans 900 µL de MATAB (Tris-HCL 100 mM, NaCl 1.5 mM, 20 mM EDTA, 2% MATAB, 0.5% Na₂SO₃, 1% PEG6000, pH = 8) à 72 °C pendant 2 h 30. Les tubes sont mélangés par inversion toutes les 15-20 min. Un volume de 900 µL de CIAA (chloroforme-alcool isoamylique, proportions : 24/1) est ajouté, puis vortexé. Le mélange est centrifugé 20 min à 6000 rpm à température ambiante. Les 250 µL de surnageant sont prélevés et transférés dans un nouveau tube et l'ADN est précipité en ajoutant 700 µL d'isopropanol. Après centrifugation pendant 20 min à 6000 g, le surnageant est éliminé et le culot est lavé 2 fois avec 1 mL d'éthanol à 70 % puis centrifugation à 6000 g pendant 10 min. Le culot d'ADN est séché brièvement puis resuspendu dans 100 µL d'H₂O milliQ.

3) Extraction d'ARN racinaire total et RT-PCR

Après 28 jours de croissance en hydroponie (**Annexe 22**), 100 à 150 mg de tissus racinaires de riz sont broyés soit au pilon et au mortier avec de l'azote liquide, soit en tube eppendorf de 2 mL après avoir été congelées avec deux billes en acier (2 mm) dans l'azote liquide, à l'aide du broyeur à bille (Genogrinder) (fréquence = 30, 30 s, deux fois).

Deux protocoles ont été mis en place pour l'extraction d'ARN racinaire de riz :

- Extraction d'ARN au RNAzol. Cent milligrammes d'échantillons broyés sont resuspendus dans 1 mL de RNAzol-RT (Sigma, R4533). Le mélange est alors vortexé puis complété avec 400 µL d'H₂O RNase free. Après incubation 15 min à température ambiante, les tubes sont centrifugés à 12000 g pendant 15 min. Le surnageant, d'environ 1 mL, est combiné avec 1 mL d'isopropanol en tube 2 mL

ependorf RNase Free. Les tubes sont mélangés par inversion puis incubés 10 min à température ambiante. Après centrifugation à 12000 g pendant 10 min, le surnageant est éliminé. Un volume de 400 µL d'éthanol à 75 % est ajouté et l'ensemble est centrifugé 3 min à 8000 g. Le surnageant est éliminé délicatement à la pipette, et le culot est séché sous la hotte chimique, tubes inversés et ouverts pendant 10 min. Le culot est repris dans 50 µL d'H₂O RNase free et vortexé 5 min pour resuspendre l'ARN. La conservation des ARN se fait à -80 °C. Leur dosage se fait sur puce ARN au TapeStation par dépôt de 1 µL sur Puce ARN.

- Au robot KingFisher Duo prime (ThermoScientific) à l'aide du kit NucleoMag RNA (Macherey-Nagel, 744350). Cent cinquante milligrammes d'échantillons broyés sont resuspendus dans du tampon de lyse (500 µL MR1 + 8,7 µL TCep) puis placés 10 min sur glace avant d'être centrifugés à 20000 g 5 min pour ne garder que le surnageant. Le reste des étapes suit le protocole du kit. Brièvement, l'ARN va être lié à des billes magnétiques. Ensemble ils vont être mis dans des bains consécutivement pour y être purifiés. Ce kit utilise une DNase dans un de ses solvants pour un des bains. Le volume final d'H₂O RNase free ajouté dépend de la quantité initiale : 100 µL pour 150 mg de matériel.

Les RT-PCR sont réalisées à l'aide du kit SuperScript III Reverse Transcriptase (Invitrogen, 18080044) à partir des ARN totaux en utilisant un oligo-dT 12-18 (Invitrogen, 58862), et entre 150-600 ng d'ARN total. La DNase du kit n'est pas ajoutée car elle est déjà présente lors de l'étape d'extraction des ARN au robot. La RNaseOUT est omise. Le volume final de la réaction est de 20 µL. La solution de travail est alors obtenue en diluant la réaction de RT pour atteindre une concentration de 3 ng/µL d'ADNc théorique pour le passage en ddPCR Evagreen (4-5 ng totaux maximum/test evagreen) (cf plus bas).

4) PCR et purifications de produits PCR

Le dessin des amorces PCR est réalisé à partir de Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) en choisissant un T_m optimal à 65 °C (entre 60 et 70 °C). La spécificité des amorces est vérifiée sur OrygenesDB (<https://orygenesdb.cirad.fr/>) pour le riz et le Genoscope (http://www.genoscope.cns.fr/cgi-bin/blast_server/projet_ML/blast.pl) pour la vigne par BlastN. La séquence des amorces

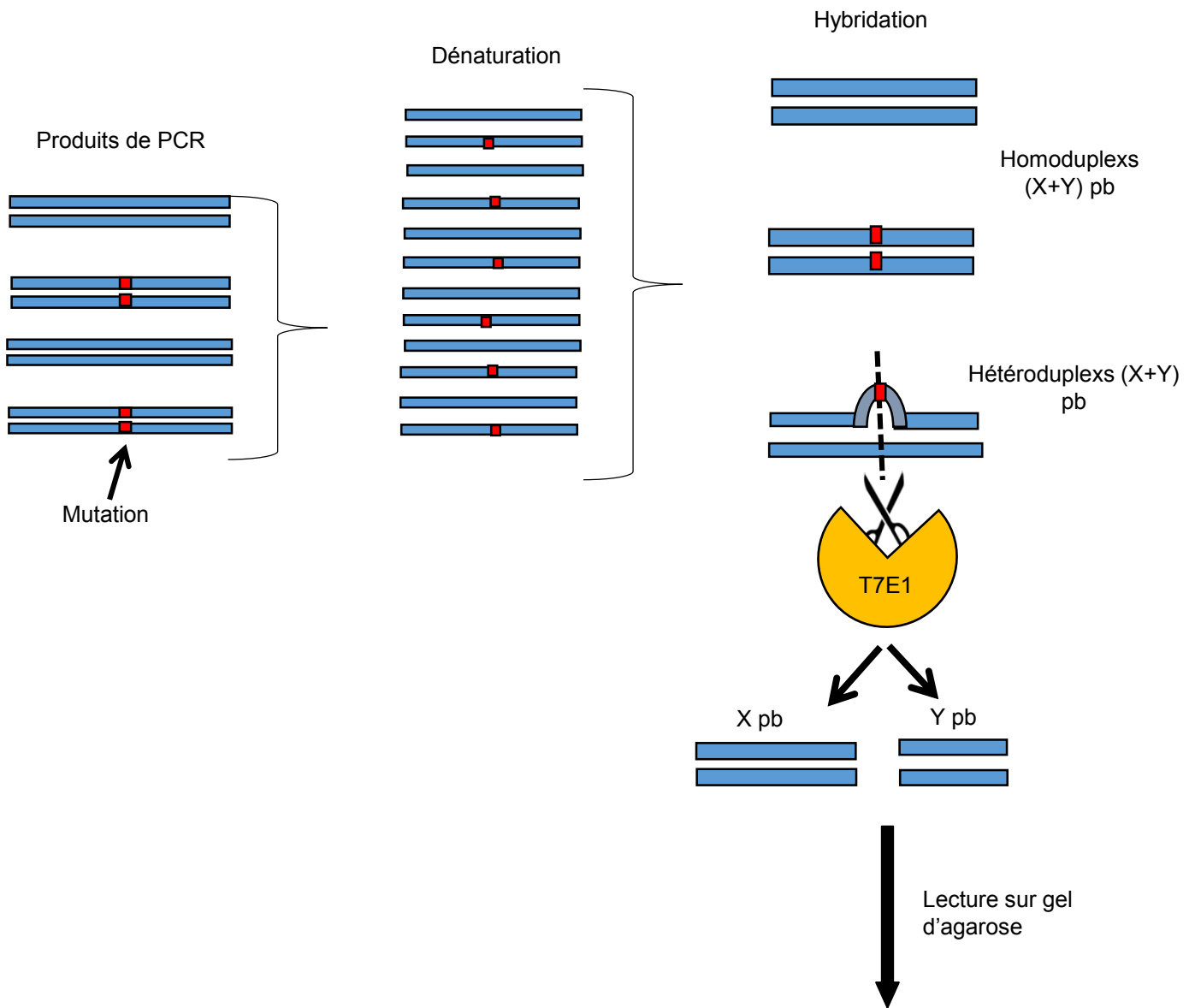
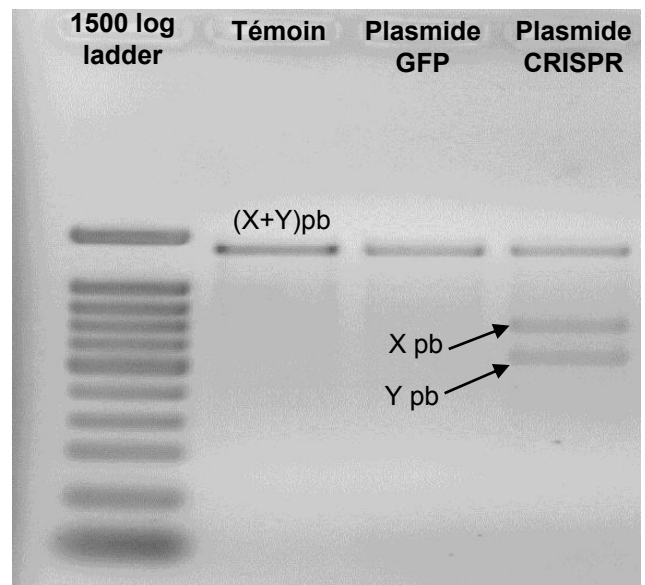


Figure 59 : Fonctionnement du test T7E1

Les produits PCR purifiés sont dénaturés et renaturés rapidement entre eux avant d'être lentement ramenés à température ambiante. L'endonucléase T7E1 va ensuite être ajoutée pour couper les hétéroduplexes.

Si les taux de mutations sont suffisants, une migration lente sur un gel d'agarose 2% permet de bien dissocier les amplifications WT (X+Y pb) des amplifications mutées (X pb et Y pb), Phot issue de l'article Bès et al, 2021.



utilisées est disponible en **Annexe 27**.

Un gradient PCR (T_m de 57 °C à 72 °C) est réalisé à l'aide du Thermocycleur (Mastercycler Gradient eppendorf) en utilisant deux Taq polymérases : la Diamond (Eurogentech, TAQ-1021-1000) pour génotypage à partir de quantités d'ADN supérieures à 50 ng, ou la KOD (Sigma, 71085) pour amplification à partir d'ADN de protoplastes ou quand la quantité d'ADN est limitante et inférieure à 5 ng. Les quantités de réactifs étaient aussi optimisées toujours sur de l'ADN génomique pour n'utiliser que 0.1 µL à 0.25 µL de Taq (**Annexe 28**).

Les purifications des produits PCR se font avec le kit PCR Purelink (Invitrogen, K310001). Brièvement, 40 µL de produits PCR sont mélangés avec 160 µL de tampon B2 et passés sur colonne de purification. L'élution se fait dans 50 µL d'H₂O MilliQ.

5) Tests T7E1

Le kit T7 endonuclease I de NEB (M0302L) a été utilisé pour détecter et couper les hétéroduplex entre brins d'ADN et détecter la présence de mutations induites par CRISPR-CAS9 dans les protoplastes. La technique est qualitative et permet de détecter les coupures, mais pas de quantifier l'efficacité (cf. le paragraphe ddPCR pour une mesure plus quantitative). Le gène ciblé par CRISPR est amplifié par PCR (**Annexe 29** et paragraphe précédent), puis purifié à l'aide du kit Purilink PCR purification dans 50 µL d'H₂O. L'étape de formation des hétéroduplex est réalisée à l'aide d'un réactif composé de 200 ng d'ADN des produits PCR purifiés (100 ng WT + 100 ng ADN mutant ; ou 200 ng d'ADN de protoplastes directement), 2 µL de tampon 10X NEBuffer 2 et complétés à 19 µL avec de l'eau nuclease-free. La réaction est réalisée à l'aide d'un thermocycleur : après une première dénaturation à 95 °C pendant 5 min, l'étape de réappariement rapide est réalisée grâce au passage de 95 °C à 85 °C en utilisant un gradient de -2 °C/sec puis plus progressivement de 85 °C à 25 °C à -0.1 °C/sec. Un volume de 1 µL d'endonuclease T7 est alors ajouté et l'ensemble est incubé 15 min à 37 °C. La réaction est stoppée à l'aide de 1.5 µL d'EDTA à 0.25 M. Un volume de 20 µL est déposé sur un gel d'agarose à 2 % pour visualiser la présence de trois bandes, les bandes homozygotes (non clivée par la T7) et deux bandes correspondants aux hétéroduplex (hétérozygotes) clivés par la T7 (**Figure 59**).

IV | ddPCR

Cette partie sur la ddPCR se base sur l'expérience acquise au cours de la thèse pour l'analyse, la conception des amorces, et des sondes, à l'aide du manuel de BioRad (BioRad) et en collaboration avec la startup IAGE France (<https://www.iage-france.com/>) présente au laboratoire. Se reporter également au **Chapitre 1** de la thèse et à l'article (Herbert et al., in prep).

1) Design TaqMan FAM/HEX

Les oligonucléotides, sondes (probes) sont toutes décrites en **Annexe 29**. La conception des amorces d'amplification forward et reverse et des sondes fluorescentes (FAM, HEX) suit les recommandations du fournisseur Biorad. Le dessin des amorces se fait avec le logiciel Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) avec les paramètres thermodynamiques et sels réglés en SantaLucia 1998. Les primers font 20 pb avec un taux de GC compris entre 50 et 60 %. Leur T_m doit être comprise entre 50 et 65 °C, en évitant les structures secondaires et les dimères, sans présence de plus de 3 G ou 3 C consécutifs, et quand cela est possible des G et des C en 3'. La taille de l'amplicon doit être comprise entre 60 et 200 pb avec une taille maximum de 250 pb. Enfin, les primers sont choisis de manière à être sur deux exons différents et consécutifs afin d'éliminer tout risque d'amplification d'ADN génomique bien que les étapes d'extraction d'ARN comprennent une étape de DNase.

Les sondes FAM ou HEX doivent faire entre 13 et 30 pb et avoir un taux de GC allant de 30 à 80 %, avec un T_m supérieur aux amorces de 3 à 10 °C (64 °C < 65 °C optimum < 70 °C) et ne pas commencer par un G. Ces sondes permettent de distinguer deux séquences qui diffèrent par un SNP (Single Nucléotide Polymorphism).

2) Réaction de ddPCR

La réaction de ddPCR est faite dans un volume final de 21 µL contenant 8 µL d'ADN de protoplastes (ou 4 ng d'ADN de plante), 11 µL de ddPCR Super Mix No UTP (2x Bio-Rad Cat #186-4021), 1 µL du mix [sonde FAM (250 nM) + primers F et R (900 nM chacun)], 1 µL

du mix [sonde HEX (250 nM) + primers F et R (900 nM chacun)] (commandés à Biorad), et qsp 21 μ L avec de l' H_2O miliQ RNase free si besoin. Les échantillons sont transférés dans une cartouche 8 puits spécifique pour générer les gouttelettes (droplet-generation cartridges, Biorad Cat#1864008 and #1863009). Un volume de 70 μ L d'huile (Droplet Generation Oil Bio-Rad cat #1863005) est ajouté dans le compartiment spécifié de la cartouche. La cartouche est mise dans l'appareil « QX200 droplet generator » pour effectuer l'émulsion par technologie microfluidique. Au final, 40 μ L de l'émulsion de chaque échantillon sont très délicatement transférés (pipette à un angle de 45° pour le prélèvement pour éviter l'éclatement des gouttelettes) dans une plaque PCR 96 (Bio Rad Cat#12001925). Elle est scellée avec une feuille en aluminium (Foil Heat Seal, Bio Rad Cat#1814040) avec l'appareil « PX1 plate sealer ». Les conditions PCR sont les suivantes : 10 minutes à 95 °C ; 40 cycles de 30 s à 95 °C et 1 minute entre 57 °C-61 °C (T_m optimisé suivant l'échantillon) ; et une dénaturation finale à 98 °C pendant 10 minutes. A noter qu'il est possible de faire varier le temps d'élongation de 1 à 3 min suivant la taille des amplifications. Les lectures se font ensuite sur l'appareil « QX200 droplet reader » puis avec le logiciel « Quanta soft software » (Analysis Pro) version 1.6.

3) Design Evagreen ARN

Les oligonucléotides utilisés pour les expérimentations avec l'Evagreen (un intercalant) sont identiques à ceux utilisés avec les sondes Taqman. Dans ce cas, la technologie utilise l'Evagreen comme intercalant fluorescent. En cas de multiplexage, par exemple avec un gène de référence, il faut néanmoins prévoir une différence de taille d'amplicon d'au moins 40 à 50 pb et ne pas dépasser une amplification de 200 pb (**Annexe 29**). Les primers sont choisis sur 2 exons successifs afin de n'amplifier que les cDNA correspondant à des ARN épissés afin d'éviter toute amplification à partir ADN génomique résiduel, bien qu'un traitement DNase est été réalisé avec le kit du robot KingFisher. La réaction de ddPCR est la même que celle décrite ci-dessus.

4) Tests d'optimisation

Les optimisations de ddPCR sont effectuées grâce à un gradient de température pour l'étape d'amplification PCR, comme pour une PCR classique, sur de l'ADN génomique (5 ng). Cette optimisation permet de faire disparaître au maximum la pluie (« rain ») lors de la lecture

de fluorescence pour mieux discriminer les gouttelettes entre elles (Herbert et al., in prep).

5) Extraction et purifications après amplification par ddPCR

Un protocole fourni par Biorad permet de récupérer les produits PCR après réaction de ddPCR. Les gouttelettes, après la réaction PCR, sont récupérées en tubes eppendorf 1.5 mL en minimisant la quantité d'huile récupérée. Vingt μL de tampon TE (pH=8) et 70 μL de chloroforme sont ajoutés et les tubes sont vortexés pendant 2 min. Cinq mélanges par pipetage aller-retour sont effectués pour faciliter la dissolution des gouttelettes. Les tubes sont alors centrifugés 10 min à 15500 g. De haut en bas du tube trois phases se forment après centrifugation : ADN, droplets résiduelles et brisées, chloroforme. Environ 25 μL de la phase supérieure transparente sont récupérés dans un tube de 1.5 mL. La purification des produits PCR est alors réalisée soit à l'aide du kit IPure V2 (Diagenode, C03010015).

VI Génotypage

Le génotypage des lignées transgéniques T-DNA CRISPR est identique pour le riz et la vigne. Les gènes de résistances à l'hygromycine ou le gène de la CAS9/nCAS9 sont amplifiés pour identifier les plantes transformées (T0) (**Annexe 27** et **28**) ou suivre la ségrégation du transgène (T1, T2). La présence de mutations induites par les construits CRISPR sont vérifiées directement par amplification PCR et séquençage des produits de PCR sur les plantes T0, T1 et T2.

1) Détermination du nombre de copies du T-DNA par qPCR

Les ADN génomiques extraits au MATAB sont dilués entre 0.5 et 5 ng/μL. Un quadruplat qPCR est réalisé par lignée pour la présence du gène de résistance à l'hygromycine (ou le gène de la CAS9) et le gène de référence (qPCR33). Un volume de 1 μL d'ADNg est ajouté à 0.3 μL d'amorces Forward (10 μM), 0.3 μL d'amorces Reverse (10 μM), 0.4 μL d'H₂O, et 3 μL de SYBRgreen QSP 6 μL. La réaction et la lecture en temps réel de la fluorescence sont réalisées à l'aide d'un LightCycler 480 (Roche) avec les conditions de PCR suivantes : 5 min à 95 °C ; 45 cycles avec (20 s 95 °C, 15 s 60 °C, 20 s 72 °C) ; puis 5 s à 95 °C. Puis une courbe de fusion est réalisée avec 1 min à 65 °C et une augmentation continue de 0,11 °C/s jusqu'à 95 °C pour s'assurer de l'amplification d'un seul ADN. Le nombre de copies du transgène est estimé après normalisation de la quantité d'ADN en utilisant le gène de référence et l'ADN d'une plante T0 contenant une copie du gène d'hygromycine comme comparaison. Brièvement, la qPCR de l'ADN de cette plante qui ne contient qu'une copie du T-DNA (vérifié par southern blot) sert de référence (on lui attribue la valeur 1) et le nombre de copies est estimé par rapport à cette valeur de référence.

2) Identification des mutations induites par CRISPR/CAS9

Les séquences cibles des CRISPR sont amplifiées par PCR (**Annexe 27**) et les produits de PCR séquencés. Les produits PCR (minimum 25 μL, sans tampon de charge) sont envoyés pour séquençage à la société Eurofins (<https://www.eurofinsgenomics.eu/en/home/>). Ils sont purifiés et séquencés à l'aide d'un des deux primers utilisés pour la PCR, avec la condition

optionnelle « riche en GC ». L'analyse des chromatogrammes de séquençage est réalisée manuellement à l'aide du logiciel FinchTV (<https://digitalworldbiology.com/FinchTV>). Les séquences sont ensuite comparées à la séquence de référence à l'aide du logiciel Multalin (<http://multalin.toulouse.inra.fr/multalin/>). En parallèle, l'analyse des chromatogrammes est faite avec le logiciel ICE Synthego (<https://ice.synthego.com/#/>) pour faciliter l'analyse des mutations complexes.

3) Test de résistance à l'hygromycine sur feuille

Afin de confirmer rapidement en serre la présence du T-DNA dans des transformants T0, T1 et T2, des tests de résistance sur feuille sont réalisés. Un volume de 4 µL d'une solution antibiotique (25 mg/mL hygromycine, 0.001 % triton-X100, 0.5 % gélatine) est déposé au centre d'une jeune feuille complètement émergée. Après 4 jours d'incubation, la plante est considérée comme homozygote pour la présence du T-DNA en l'absence totale de nécrose ou si une légère tache jaune, correspondant au dépôt, est présente. La plante est considérée comme hétérozygote (ou hémizygot) si une nécrose blanche, mais peu étendue est présente. La plante ne possède pas/plus de T-DNA si la nécrose s'étend le long des tissus vasculaires de la feuille.

A

Tratamiento 90 Kg de Nitrogeno				
Dosis en Gramos				
Fuente	Siembra	15 DDS	25 DDS	35 DDS
	22-may.-20	06-jun.-20	16-jun.-20	26-jun.-20
UREA (N 46%)		210	421	421
SAM (N 21%)	831	166	332	332
Fosfate Monopotasico (P 52%)	1066			
Fosfate Monopotasico (K 34%)	1071			
KCL			628	628
Sulfex Zn	231			
Sulfato de Fe	600			
Sulfato de Mg				
Agrimins	462			

C Soil Analysis – Pooled soil samples

Parameter	Unit	Values
pH	un	6.59
C Oxid	g/Kg	33.03
Organic matter	g/Kg	74.99
N-NO ₃	g/Kg	0.567
N-NH ₄	g/Kg	0.423
P-BrayII	mg/kg	35.78
Ca	cmol/kg	10.33
Mg	cmol/kg	4.41
K	cmol/kg	0.72
Na	cmol/kg	0.681
CIC	cmol/kg	36.683
Fe	mg/kg	2.107
Mn	mg/kg	18.659
Cu	mg/kg	0.157
Zn	mg/kg	6.868
B	mg/kg	0.757
S	mg/kg	158.656

B

Tratamiento 0 Kg de Nitrogeno				
Dosis en Gramos				
Fuente	Siembra	15 DDS	25 DDS	35 DDS
	22-may.-20	06-jun.-20	16-jun.-20	26-jun.-20
Fosfate Monopotasico (P 52%)	1066			
Fosfate Monopotasico (K 34%)	1071			
KCL			628	628
Sulfex Zn	231			
Sulfato de Fe	600			
Sulfato de magnesio	692	692		
Agrimins	462			

Sample	Deep (cm)	Sand	Clay	Silt
		%	%	%
1	10	62.39	19.22	18.38

Texture	Apparent density g/cc
Sandy loam	1.14

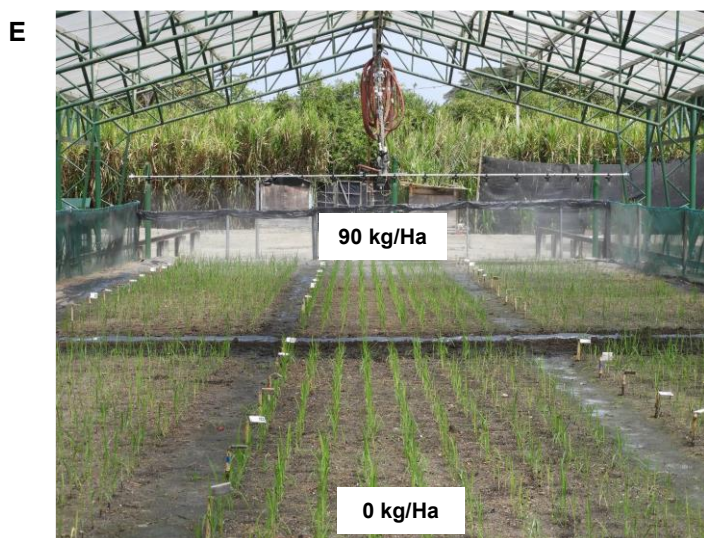
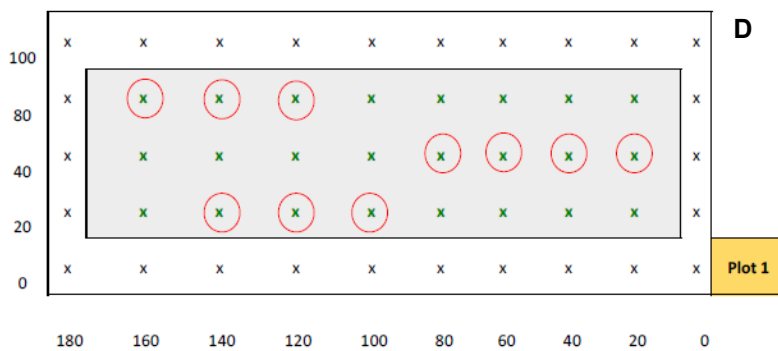


Figure 60 : Présentations des données de culture en Colombie (CIAT-champ)

Quantité de milieux par arrosage en HN (A) et LN (B) avec supplément d'urée (46-0-0 ; 46% N ; CH₄NO₂). Analyse de la composition du sol avant plantation (C). Disposition des plantes par « plot » (1,8m², 20cm entre plantes et 20cm entre rangée), avec retrait de l'effet de bord (surface blanche), et sélection aléatoire des individus (cercles rouges) (D). Photos des cultures en LN et HN au début et en fin de cycle (E).

VII Physiologie et NUE

Le phénotypage des plantes éditées pour la NUE a été effectué en utilisant deux stratégies : la première au champ en Colombie au CIAT pour le génotype Chhomrong Dhan et la seconde à Montpellier en serre de confinement sur le génotype Kitaaké en pots en semi-hydroponie avec une solution nutritive en faible et forte concentration en azote (Araus et al., 2016).

1) NUE en faible et forte concentration d'urée — Champ en Colombie (Chhomrong Dhan)

En Colombie les CHHD envoyés ont été multipliés sur place avant d'être semés pour les essais. Les essais ont eu lieu en champ sur 2 milieux, un à 90 kg/ha d'urée (46-0-0 ; 46 % N. ; CH₄NO₂) (HN) et l'autre à 0 (LN). Les quantités de milieux ajoutés par arrosage et les composés initiaux des sols sont présentés en **Figure 60**.

Les plantes sont disposées par « plot » de 1.8 m², 20 cm entre plantes et 20 cm entre rangées, avec le retrait de l'effet de bord pour la sélection « aléatoire » des individus (**Figure 60-D**). Les cultures en LN et HN au début et en fin de cycle sont présentées en **Figure 60-E**.

Il est important de préciser le calcul de l'évaluation la NUE par Hectare (Ha). Nous considérerons que les plantes ont une partie racinaire explorant jusqu'à 50 cm sous la surface du sol, et sont disposés tous les 20 cm. C'est pour ça qu'il s'agit **d'une évaluation**. Cela donne un volume de sol occupé de 0.02 m³ par plante (0.2*0.2*0.5). Le sol fait 1140 kg/m³ dans ces cultures au CIAT, soit 22.8 kg de sol par plante (1140*0.02). À 10 cm de profondeur, il a été mesuré 0.000990 kgN/kg sol (N-NO₃ et N-NH₄), ce qui revient à avoir 0.022572 kgN/plante (22.8*0.00099). Si la surface fait 1 ha (10 000 m²), alors cela revient à une densité de 250 000 plantes/Ha (10 000/(0.2*0.2) (surface par plante occupée)), donc 5643 kgN/Ha naturellement présent si rien n'est ajouté (250 000*0.022572). Il faut ensuite rajouter à cela 0 kg de N. (5643 kgN/Ha) ou 90 Kg de N. (5733 kgN/Ha) sous forme d'urée, puis calculer la NUE en divisant le rendement (en Kg par Ha) par cette valeur.

Pour avoir le rendement à l'hectare, il faut multiplier le rendement moyen par plante (kg) par la densité de plantes pour 1 hectare (250 000 ici, 10000/(0.2*0.2)).

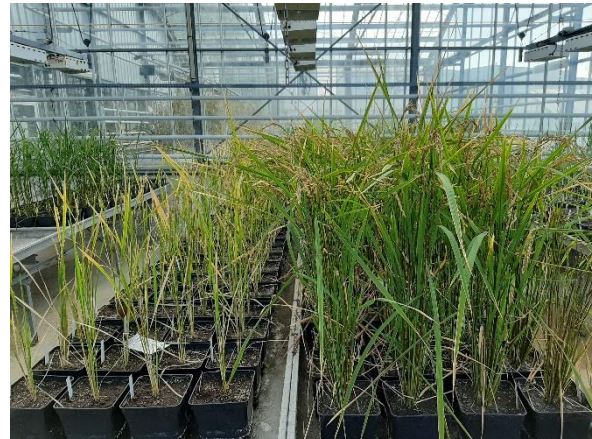
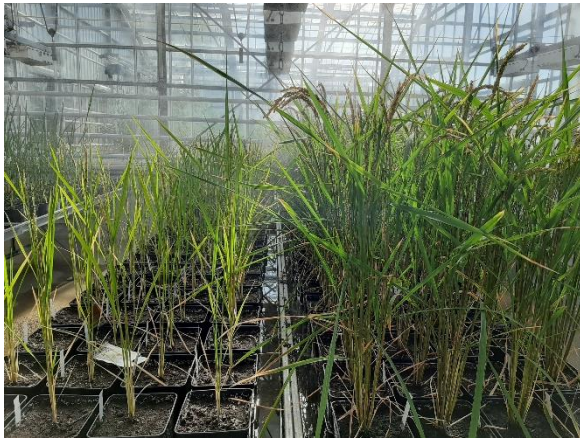
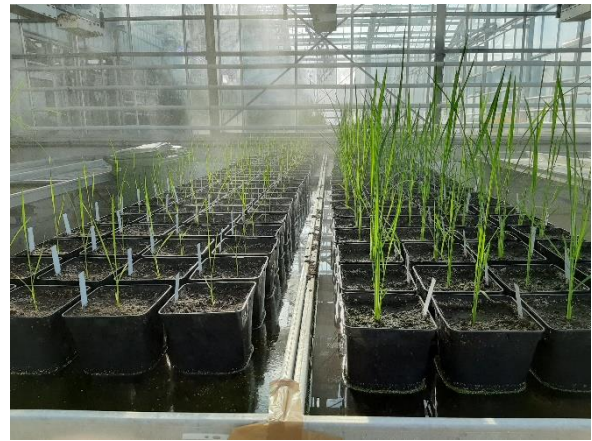


Figure 61 : Planches représentant les systèmes de culture LN et HN en semi hydroponie
Chaque photo représente un stage de culture de la germination à la plante mature prête à être prélevée sur 0,5 mM (gauche) ou 5 mM (droite) de KNO_3 . Les milieux nutritifs enrichis en KNO_3 sont renouvelés hebdomadairement pendant 12 semaines.

2) NUE en faible et forte concentration en KNO₃ — Serre de confinement (Kitaaké)

Une masse de 1500 g (+/-10 g) de terreau (Tref 2) est ajoutée par pot de 3 L. L'apport des nutriments se fait une fois par semaine à raison de 10 L de solution pour 12 plantes. La solution utilisée est similaire à celle utilisée pour l'hydroponie et préparée en grandes quantités en serre pour deux concentrations de KNO₃ différentes (100 L, Milieu de Yoshida modifié avec solution de microéléments (oligos) issus de Araus et al., 2016. (**Annexe 5**). Les pots sont arrosés à l'aide de la solution nutritive par-dessus afin de bien imbiber le terreau et éliminer les algues. De l'eau est ajoutée une fois par semaine si les pots ne sont plus complètement imbibés. Une graine de Kitaaké prégermée en boîte de Petri pendant 7 jours (pas de décorticage, pas de désinfection) est alors transférée dans chaque pot.

Le cycle de complet de Kitaaké, variété peu sensible à la photopériode, est d'environ 90 jours. Le cycle complet peut être présenté par la **Figure 61**. La date de floraison est notée. Et les paramètres analysés sont : la hauteur de la plante, le nombre de talles total (brin maître compris), le nombre de panicules final, le poids sec de la partie aérienne avec les panicules et grains (après mise en étuve à 42 °C pendant deux jours), le poids total des grains matures, et le nombre de grains pour avoir le poids moyen de 1000 grains.

La NUE est calculée par pot. En connaissant la quantité totale de KNO₃ mise (en g), il suffit de la diviser par 120 pots (« équitablement répartie »), puis diviser le rendement par cette valeur.

3) Influx 15N en racines et ARN

Les plantules sont cultivées pendant 28 jours (7 jours de croissance sur H₂O avant ajout du milieu nutritif) sur un milieu d'hydroponie (**Annexe 22**) en chambre de culture (12 h/12 h, LED, 29 °C, 70 % humidité, 25 °C nuit) sur 5 mM KNO₃ ou sur 5 mM NH₄NO₃ (**Figure 62**). Une solution de culture est préparée en remplaçant les concentrations de KNO₃ ou NH₄NO₃ par du 15N-KNO₃ (IN5547 Berry & Associates) à 0.5 ou 5 mM. Autour de 13h00, pic d'expression des transporteurs de nitrate, 3 racines de ces plantes sont prélevées, séchées sur du papier absorbant, mis dans l'azote liquide, et broyées (1 min fréquence broyage 30) pour effectuer une extraction d'ARN au robot KingFisher sur tout le matériel. Les autres racines des plantes sont placées 1 min dans une solution de lavage de CaSO₄ (180 mg pour 10L H₂O). Ensuite on les transfère dans la solution de ¹⁵N-KNO₃ pendant 5 min. Elles sont ensuite transférées 1 min



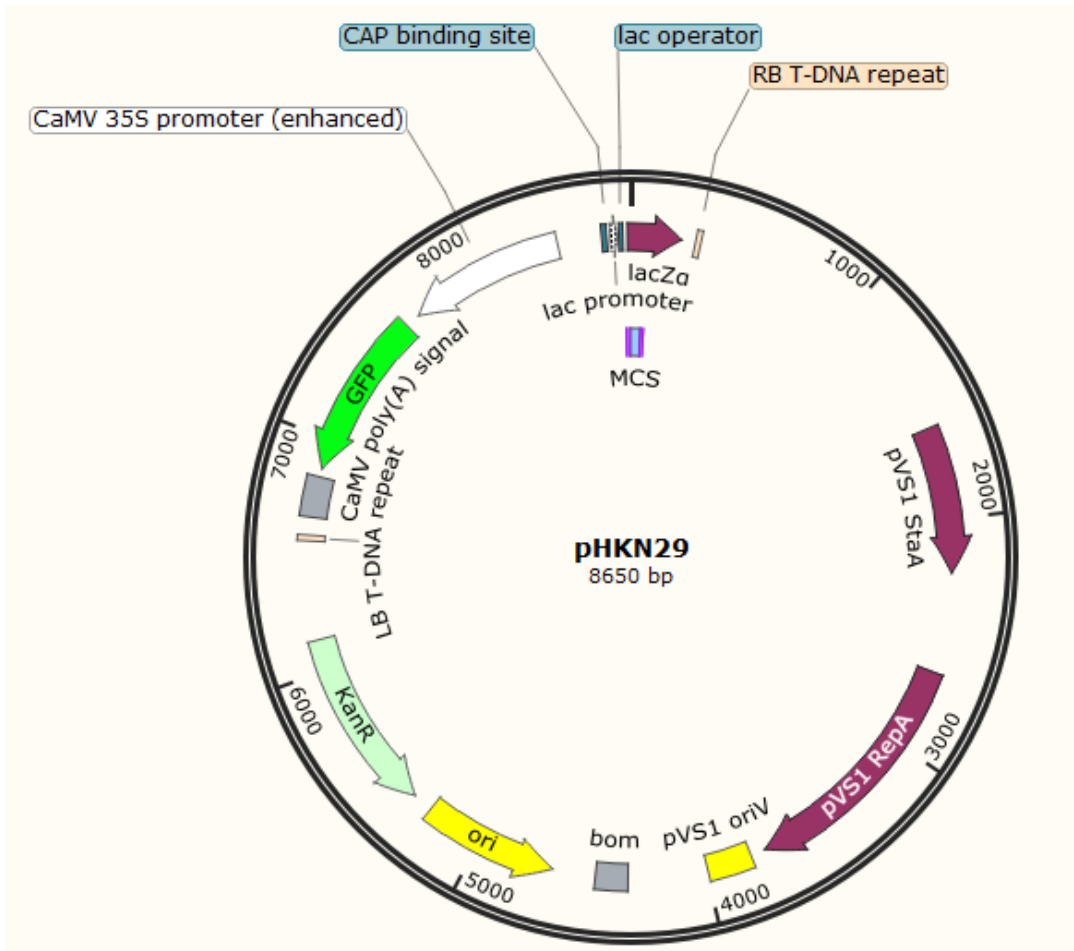
Figure 62 : Prélèvements racinaires en hydroponie pour l'expression et l'influx

Les différents géotypes des Kitaaké ont été cultivés un mois en hydroponie dans un milieu nutritif complété avec 5 mM NH_4NO_3 ou 5 mM KNO_3 .

Les prélèvements racinaires se feront sur 12 racines pour l'influx et 3 racines pour l'expression, pour chaque géotype.

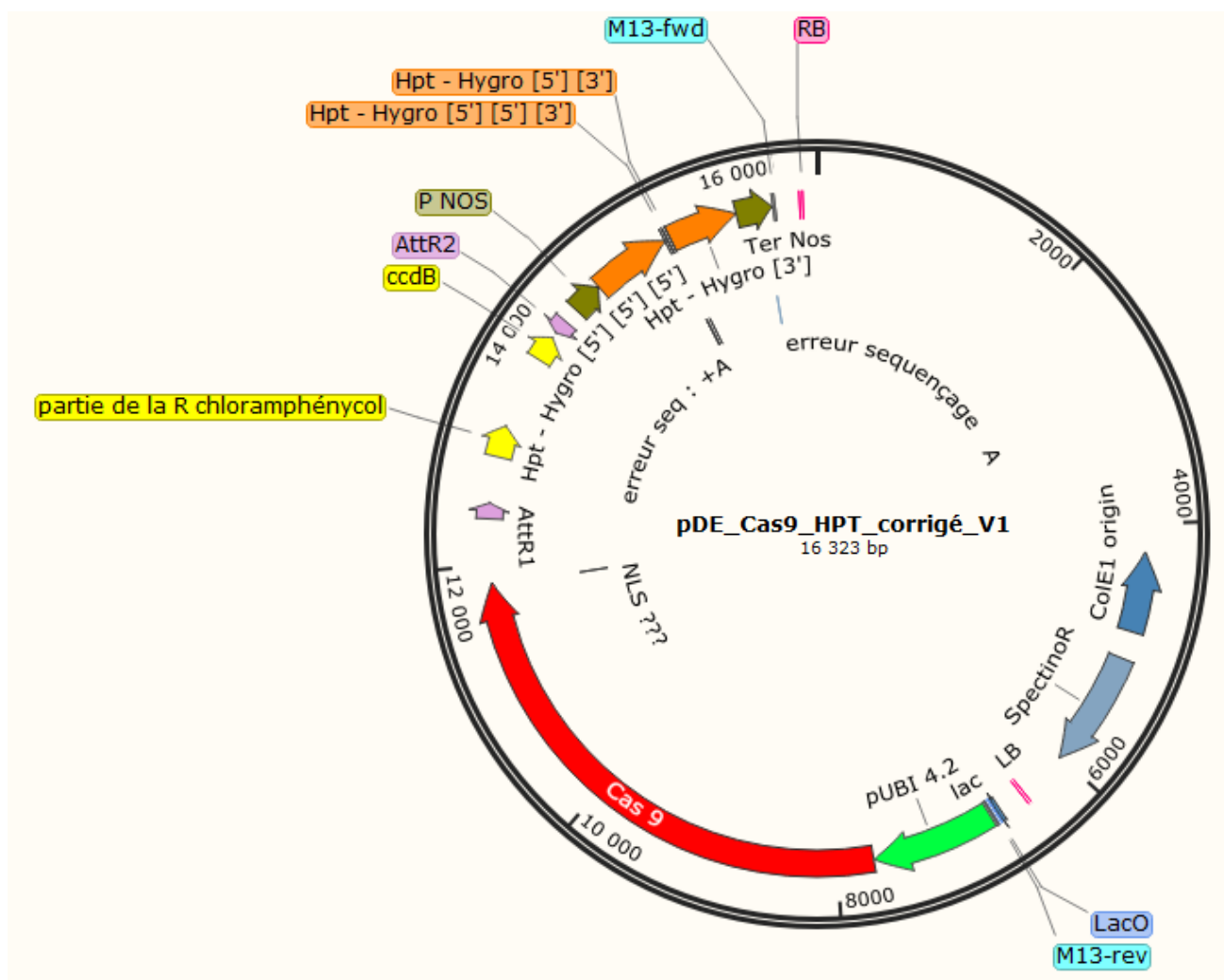
dans la solution de lavage de CaSO_4 pour retirer le ^{15}N résiduel. Les racines sont alors séchées brièvement sur du papier absorbant avant d'être mises dans des cupules en aluminium préalablement pesées une à une à la balance de précision. Elles sont séchées à l'étuve une semaine à $70\text{ }^\circ\text{C}$. Les racines dans les cupules sont ensuite pesées pour savoir leur poids sec exact avant d'être analysées au spectromètre de masse à isotopes lourds. Le taux de $^{15}\text{NO}_3^-$ sera donné en $\mu\text{mol/h/g}$ de MS (matière sèche) par racine.

Annexes



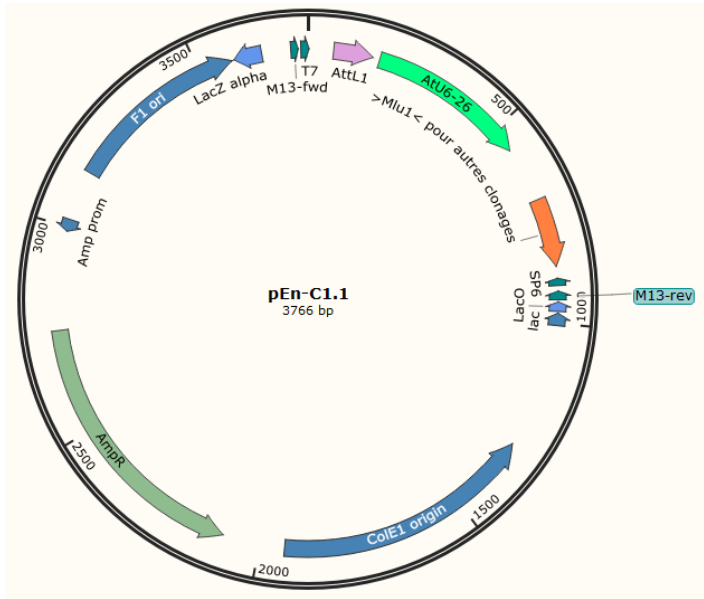
Nom	pHKN29 (témoin GFP)
Taille (pb)	8650
Résistance	Kanamycine
Enzymes de clonage	n/a
Enzymes de vérification	EcoRI + HpaI

Annexe 1A



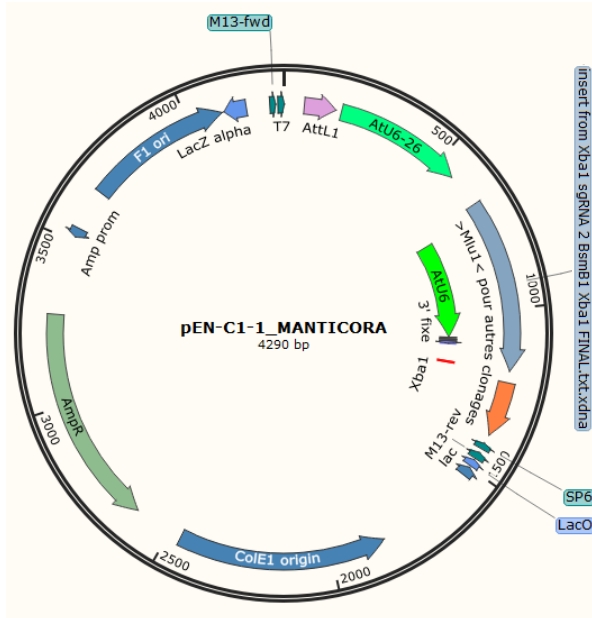
Nom	pDE_CAS9_Hpt
Taille (pb)	16323
Résistance	Spectinomycine + Chloramphénycol (si pas de LR) + Hygromycine en plante
Enzymes de clonage	LR nécessaire
Enzymes de vérification	Hind III + BspHI

Annexe 1B



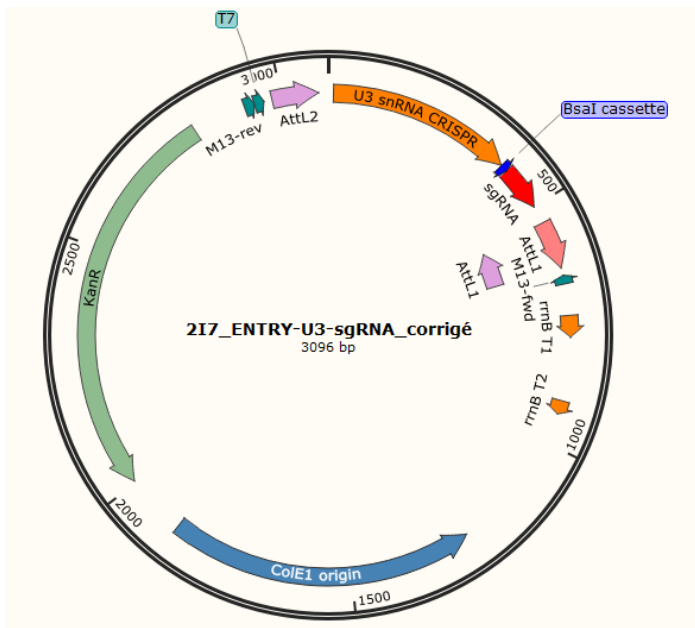
Nom	pEN-C1.1 (Chimera 1.1)
Taille (pb)	3766
Résistance	Ampicilline
Enzymes de clonage	BbsI
Enzymes de vérification	BbsI
Note importante	Plasmide identique au pEN-Chimera, mais dispose en plus de l'ajout d'un site de clonage multiple MLU1 (flèche orange)

Annexe 1C



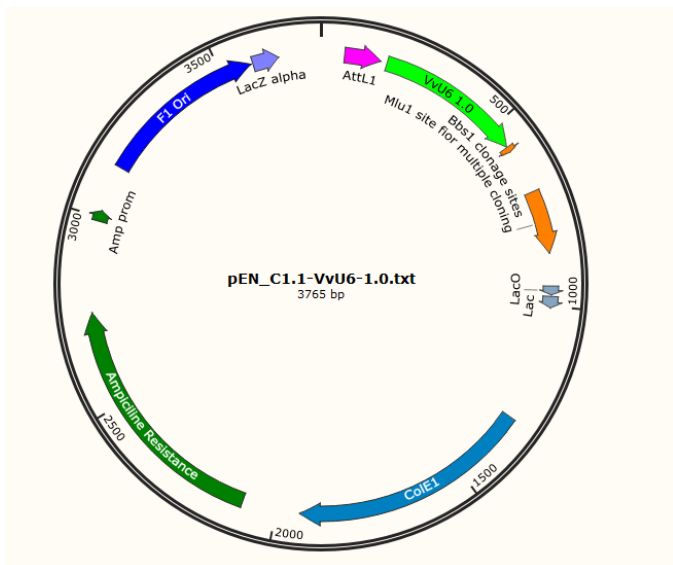
Nom	pEN-Manticora-C1.1
Taille (pb)	4290
Résistance	Ampicilline
Enzymes de clonage	BbsI + BsmBI
Enzymes de vérification	BbsI + EcoRI

Annexe 1D



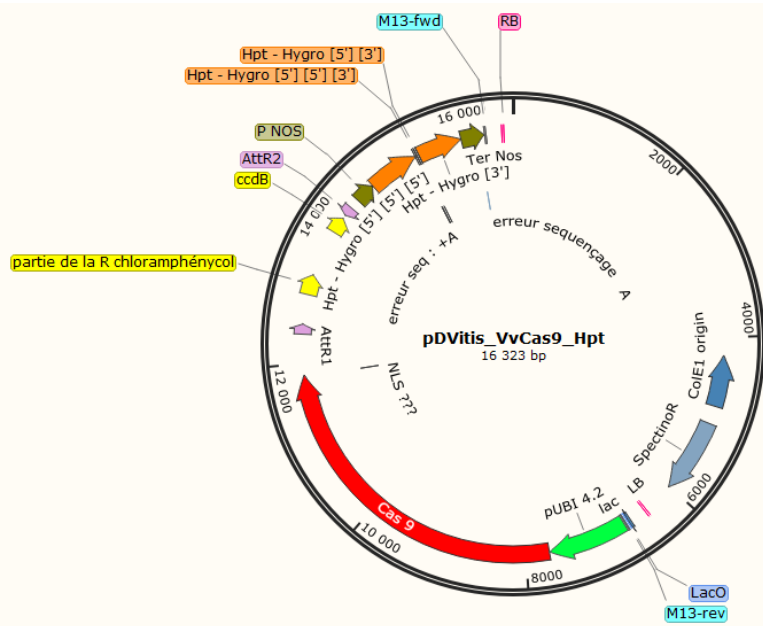
Nom	217
Taille (pb)	3096
Résistance	Kanamycine
Enzymes de clonage	BsaI
Enzymes de vérification	BsaI

Annexe 1E



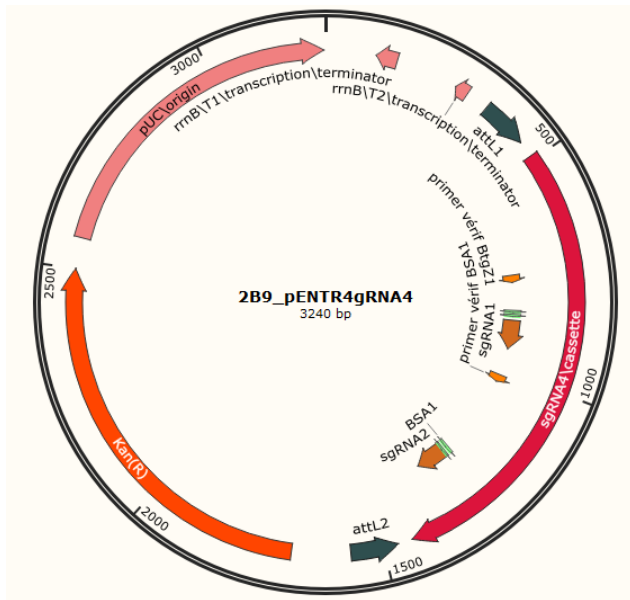
Nom	pEN-C1.1-VvU6-1.0
Taille (pb)	3765
Résistance	Ampicilline
Enzymes de clonage	Bbs1
Enzymes de vérification	BspHI + SpeI

Annexe 1F



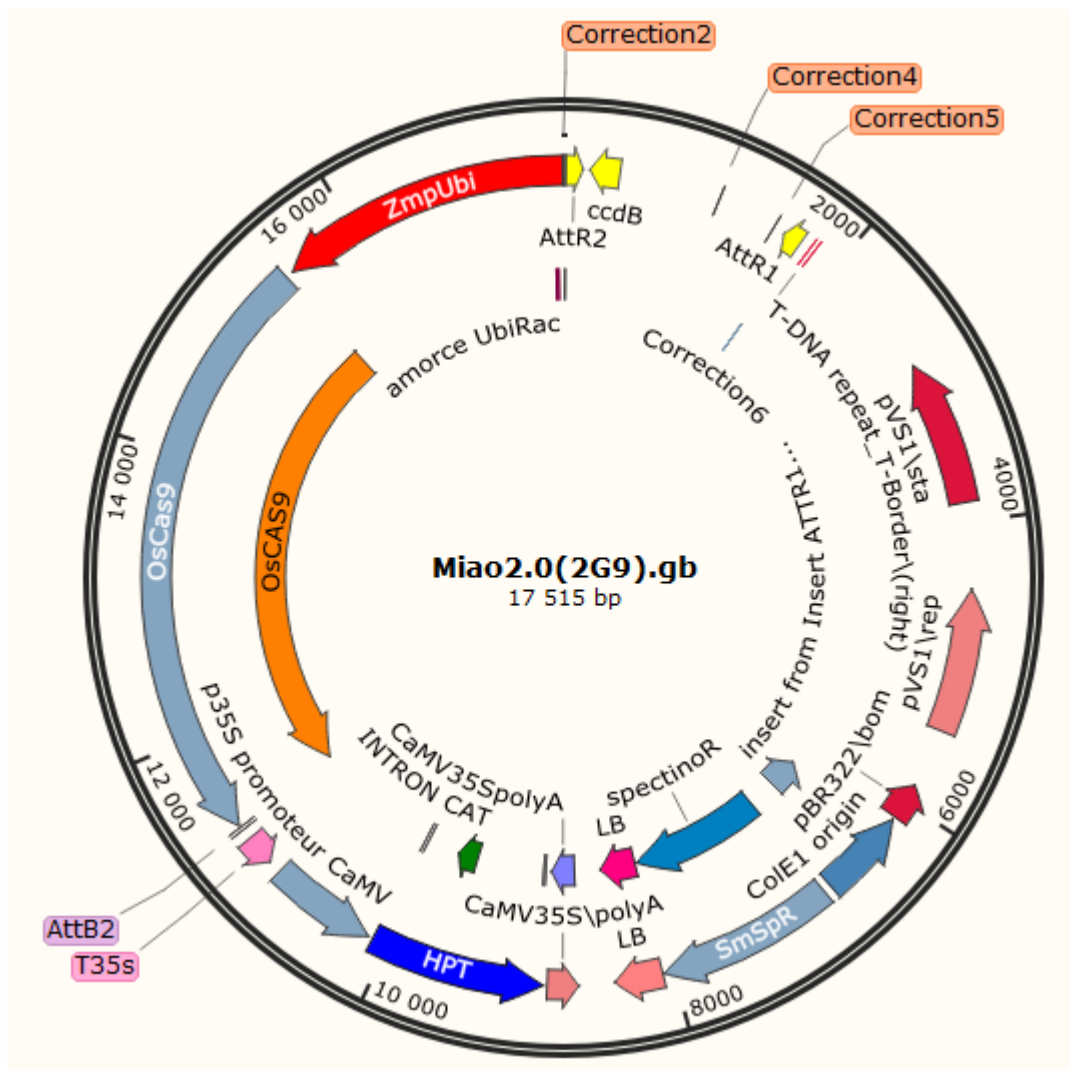
Nom	pDVitis_VvCAS9_Hpt
Taille (pb)	16323
Résistance	Spectinomycine + Chloramphénicol (si pas de LR) + Hygromycine en plante
Enzymes de clonage	LR nécessaire
Enzymes de vérification	BspHI + SpeI

Annexe 1G



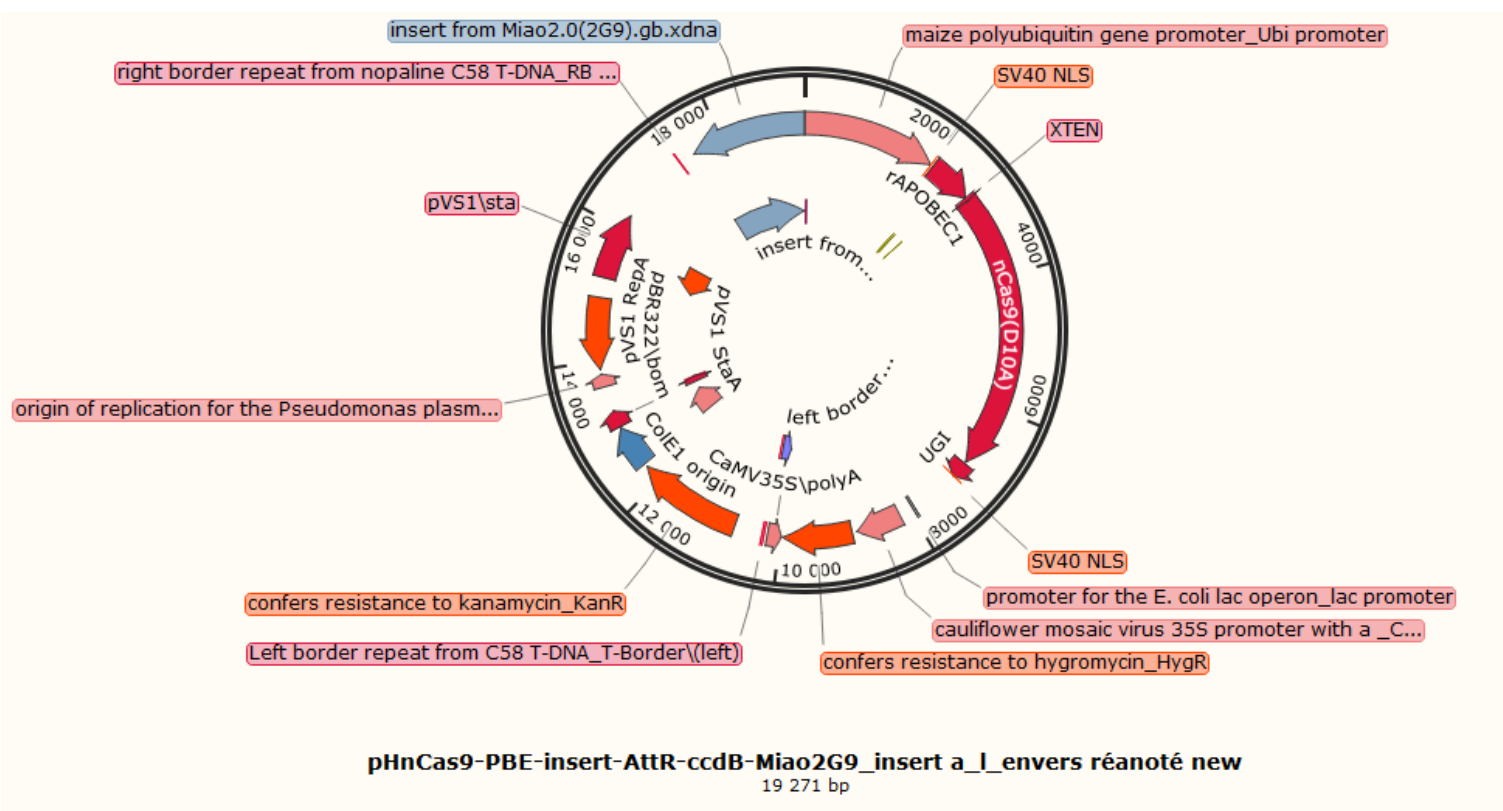
Nom	2B9
Taille (pb)	3240
Résistance	Kanamycine
Enzymes de clonage	BsaI + BtgZI
Enzymes de vérification	BsaI et/ou BtgZI

Annexe 1H



Nom	Miao 2.0 – 2G9
Taille (pb)	17515
Résistance	Spectinomycine + Chloramphénicol (si pas de LR) + Hygromycine en plante
Enzymes de clonage	LR nécessaire
Enzymes de vérification	SpeI + StuI

Annexe 1i

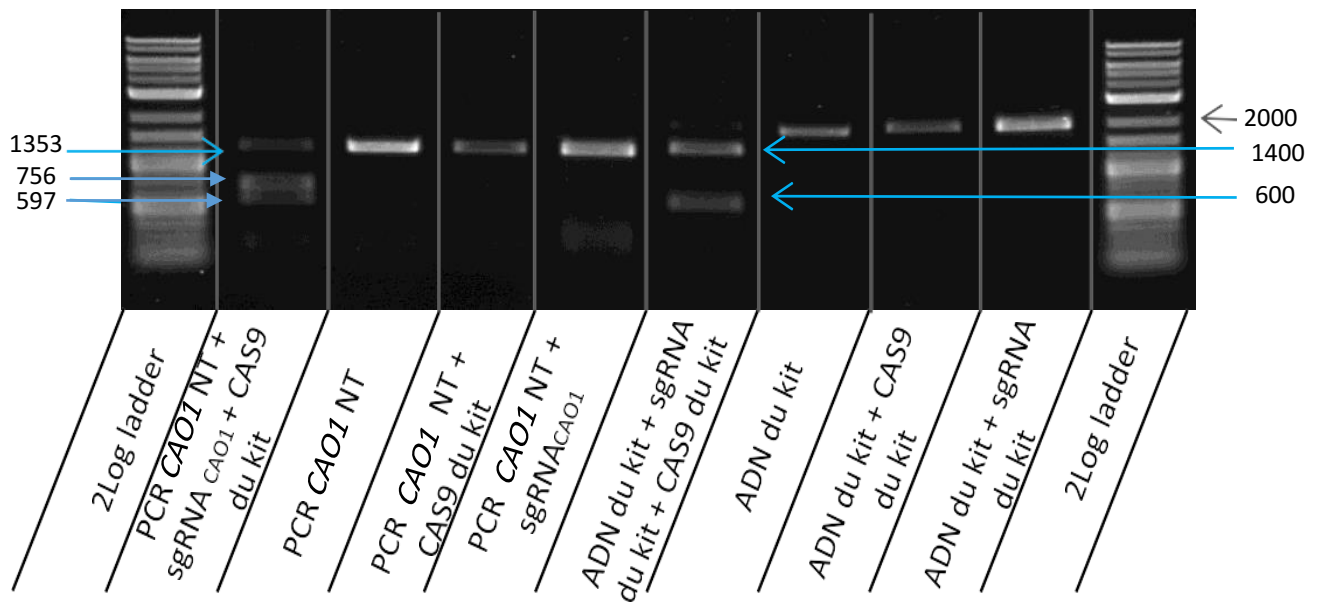


Nom	pBE-nCAS9-Attr-ccdb
Taille (pb)	19267
Résistance	Kanamycine + Chloramphénicol (si pas de LR) + Hygromycine en plante
Enzymes de clonage	LR nécessaire
Enzymes de vérification	KpnI et SpeI après LR

Annexe 1J

Génotype Kitaaké	Mutation	Précision	Nombre de plantes
WT	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCGACGGCGAGCAAGTGG Allèle 2 : CGGCGACGGCGAGCAAGTGG	WT	12
HO	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCGATGGCGAGCAAGTGG Allèle 2 : CGGCGATGGCGAGCAAGTGG	Base Edition homozygote C en T	10
HT	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCGATGGCGAGCAAGTGG Allèle 2 : CGGCGACGGCGAGCAAGTGG	Base Edition hétérozygote C en T	13
HT	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCGATGGCGAGCAAGTGG Allèle 2 : CGGCGAGGGCGAGCAAGTGG	Base Edition hétérozygote G et T	2
HT	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCGACGGCGAGCAAGTGG Allèle 2 : CGGCGAGGGCGAGCAAGTGG	Base Edition hétérozygote C en G	1
HT	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCG.AGG Allèle 2 : CGGCGACGGCGAGCAAGTGG	Délétion -15 hétérozygote	1
HT	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT Allèle 1 : GACGG.TGT Allèle 2 : CGGCGACGGCGAGCAAGTGG	Délétion -15 hétérozygote	1
HT	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCGACGGCGAGCAAGTGG Allèle 2 : CGGC.	Délétion -16 hétérozygote	1
Indéterminé	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT	Traces non déterminées (chimères possibles)	2
Génotype CHHD	Mutation	Précision	Nombre de plantes
WT	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCGACGGCGAGCAAGTGG Allèle 2 : CGGCGACGGCGAGCAAGTGG	WT	2
HO	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCGATGGCGAGCAAGTGG Allèle 2 : CGGCGATGGCGAGCAAGTGG	Base Edition homozygote C en T	1
HT	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCGATGGCGAGCAAGTGG Allèle 2 : CGGCGACGGCGAGCAAGTGG	Base Edition hétérozygote C en T	3
HT	Wt : GCGGAGTCGC CGGCGA CGGCGAGCAAGTGG AGGCT Allèle 1 : CGGCGACGGCGAGCAAGTGG Allèle 2 : CGGCGAGGGCGAGCAAGTGG	Base Edition hétérozygote C en G	1

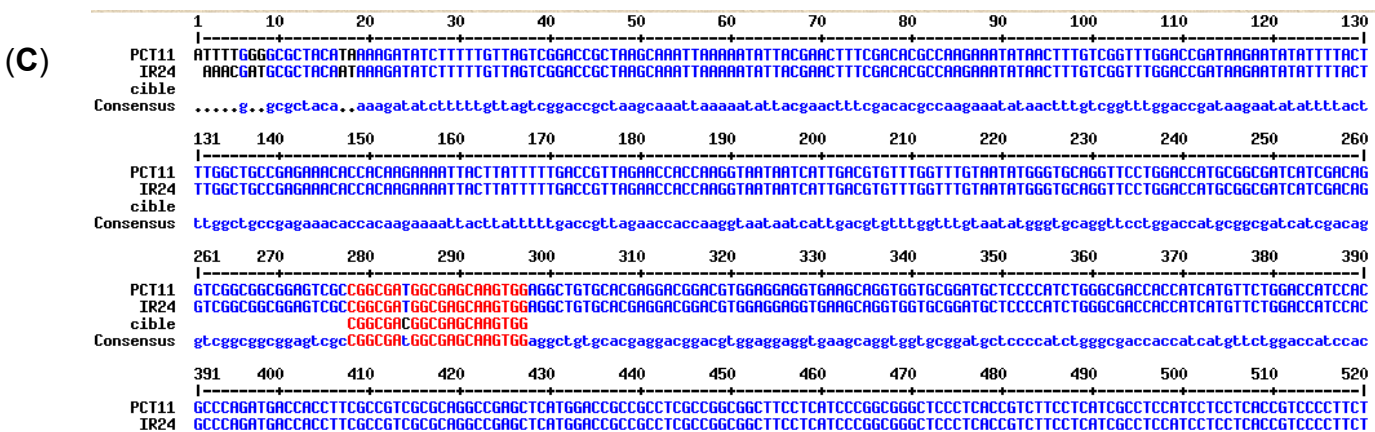
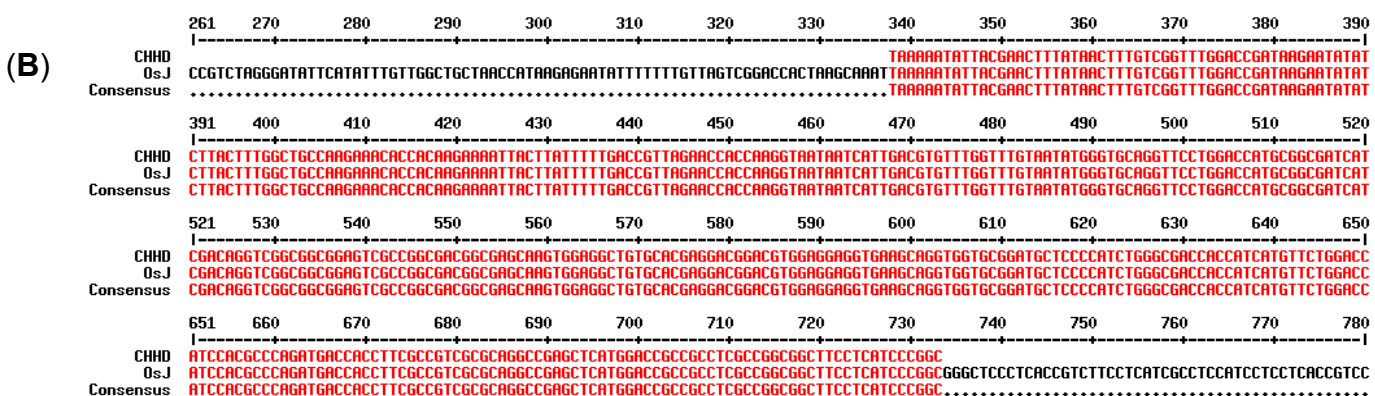
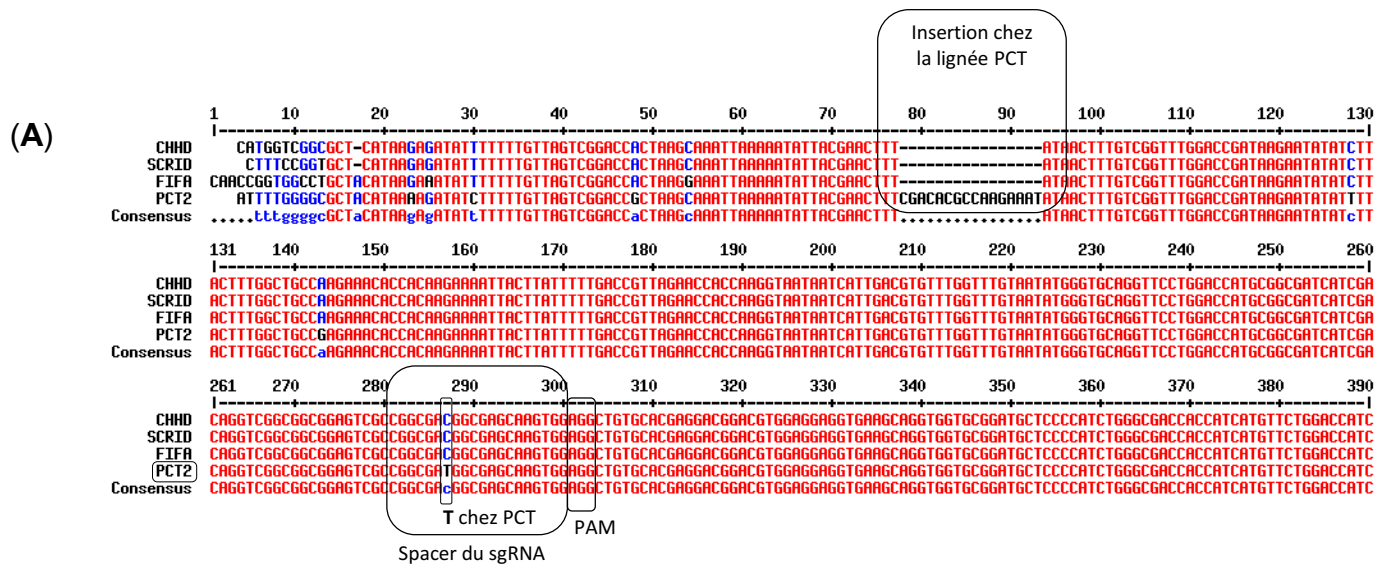
Annexe 2 : Tableau des séquences Base Edités sur *OsNRT1.1b* en Kitaaké et Chhomrong Dhan



Annexe 3 : Validation du sgRNA contre CAO1 par préchargement du sgRNA par la CAS9.

Pour vérifier si les sgRNA transcrits ciblent bien CAO1, 100 ng d'une amplification PCR de CAO1 chez Nipponbarre non Transformé (plante Wild Type, NT) de 1353 pb est mis en contact avec 20 ng de sgRNA et 500 ng de CAS9 préalablement incubés 10 min à 25°C.

S'il y a clivage, deux bandes à 756 pb et 597 pb sont observées pour CAO1 ainsi que deux bandes à 1400 pb et 600 pb pour l'ADN témoin du kit Clontech Guide-it sgRNA, ciblé par le sgRNA témoin du kit (fragment du chromosome 10 humain).



Annexe 4 : Séquences de *OsNRT1.1b* dans les lignées du projet GeneRice

Les comparaisons des séquences des 4 lignées sélectionnées (CHHD, SCRID, FoFIFA, et PCT11) autour de *OsNRT1.1b* sont présentées en (A). On peut y retrouver une insertion de 16 pb dans la lignée PCT11, et la présence déjà effective du C>T en position 980.

La comparaison de CHHD contre les bases de données de *japonica* est présentée en (B).

La comparaison de PCT11 contre les bases de données de *indica* est présentée en (C). La cible CRISPR a été rajoutée.

Solution	Sels	Poid	Solution-mère (g /L)	Concentration Solution Mère	Concentration finale	Volume	
		Moléculaire				Pour 1L	Pour 100L
Mg	MgSO ₄ · 7H ₂ O	246.49	197.2	800mM	1.5mM	1.2mL	120mL
Ca	CaCl ₂ 2H ₂ O	147	88,2	600mM	1.5mM	2.5mL	250mL
KH ₂ PO ₄	KH ₂ PO ₄	136.09	27.2	200mM	0.8mM	4mL	400mL
Fe-EDTA	FeSO ₄ ·7H ₂ O	278.00	5.56	20mM	60µM	3mL	300mL
	Na ₂ EDTA	372.24	7.44	20mM	60µM		
Nitrate de potassium	KNO ₃	101.11	35.4	350mM	0.5mM	1.43mL	143mL
					5mM	14.3mL	1430mL

	produit	Hydratation	Masse Mol (g/L)	quantité finale	en mMol	et donc en Mol	quantité (g) pour 10L	quantité (g) pour 100L	solution mère préparée	
<i>Oligo-éléments Viviana</i>	H ₃ BO ₃		61,83	50µM	0,05	0,00005	0,031	0,309		
	MnSO ₄	H ₂ O	169,02	50µM	0,05	0,00005	0,085	0,845		
	CoCl ₂	6H ₂ O	237,9	0,055µM	0,000055	0,00000055	0,00013	0,001	100mg	10mg pour 1mL donc 1mg dans 100µL
	Na ₂ MoO ₄	2H ₂ O	241,9	0,05µM	0,00005	0,00000005	0,00012	0,001	100mg	10mg pour 1mL donc 1mg dans 100µL
	ZnSO ₄	7H ₂ O	287,6	25µM	0,025	0,000025	0,072	0,719		
	KI		166	5µM	0,005	0,000005	0,008	0,083		
	CuSO ₄	5H ₂ O	249,7	0,08µM	0,00008	0,00000008	0,00020	0,002	200mg	20mg pour 1mL donc 2mg dans 100µL
pH = 5.5 au KOH										

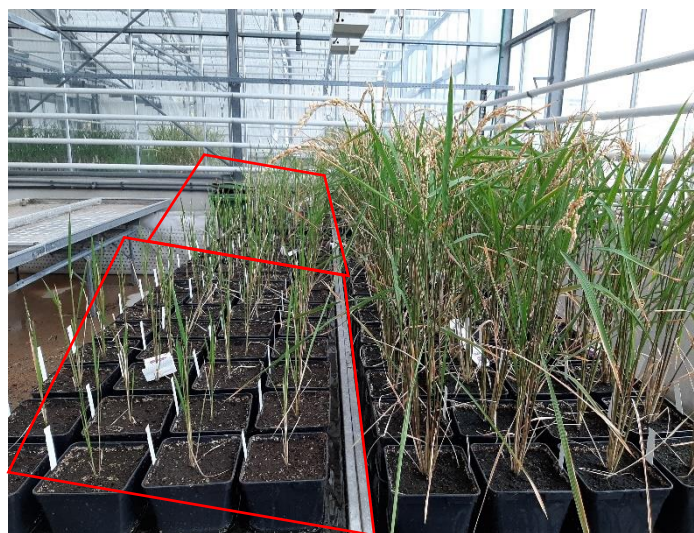
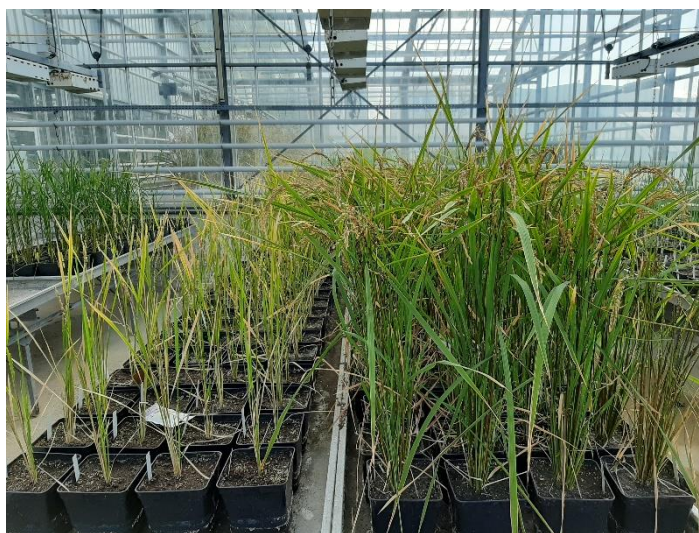
Annexe 5 : Milieux semi-hydroponie Riz Phénotypage (Medium Yoshida et Arous modifié)

	Variation observée sur les caractères	Traitement		Génotype		Intéraction génotype et traitement	
		Effet significatif	% d'explication de la variation	Effet significatif	% d'explication de la variation	Effet significatif	% d'explication de la variation
Juin	nbtalles	***	82,6	n/a	0,6	n/a	0,7
	biomasse_ssgrains	***	75,2	**	1,6	n/a	1,2
	rendement	***	85,7	***	1,8	***	1,7
	NUE_plante	*	1,9	**	6,3	n/a	3,8
	HI	***	49,1	**	3,5	n/a	0,6
	Hauteur	***	83,3	***	1,7	n/a	0,8
	nbgrains	***	86,9	***	1,6	***	1,5
	PoidsMilleGrains	***	6,4	***	18,9	***	9,0
	biomasse_totale	***	85,5	***	1,5	***	1,4
	nbpan	***	80,6	n/a	0,6	n/a	0,8
	grainparpan	***	81,3	***	2,6	*	1,1
Août - tablars LN pris en compte	nbtalles	***	74,4	***	4,0	n/a	1,1
	biomasse_ssgrains	***	75,1	***	4,9	**	2,3
	rendement	***	70,1	***	7,5	***	3,2
	NUE_plante	n/a	0,2	***	18,6	**	9,7
	HI	***	11,2	n/a	6,0	n/a	6,3
	Hauteur	***	81,1	***	5,4	*	1,3
	nbgrains	***	70,6	***	7,2	**	3,0
	PoidsMilleGrains	n/a	0,6	***	12,1	**	11,1
	biomasse_totale	***	74,7	***	6,2	***	2,8
	nbpan	***	74,3	***	3,4	n/a	1,3
	grainparpan	***	71,3	***	7,0	**	2,7

Annexe 6 : Anova des données Juin et Août des Kitaaké

Les données de LN et HN en Juin ont été comparées. Par analyse de variance on voit qu'entre 49 et 86 % des caractères observés sont dûs à l'effet du traitement, sauf pour la NUE ; 6 % de la NUE et 18 % du poids de 1000 grains sont dûs à un effet du génotype ; et que 9 % du poids de 1000 grains est dû à un effet d'interaction entre le génotype et le traitement.

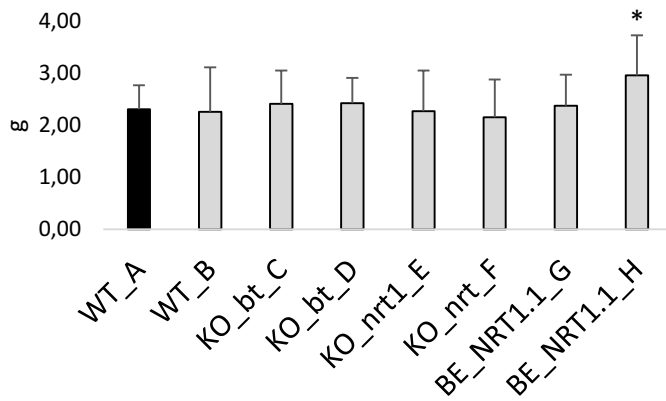
Toutes les données de LN et HN en Août ont été comparées. Nous avons cependant retiré les deux tablars LN.



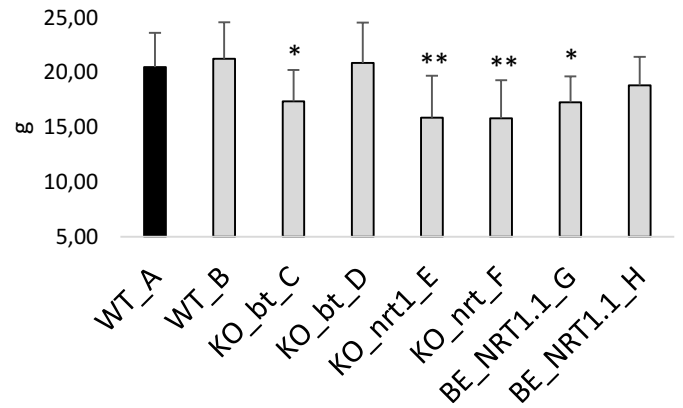
Annexe 7 : Comparaison des deux expérimentations Kitaaké en Juin et Août 2020 à floraison

Après 12 semaines de culture sur milieu LN (tablar de gauche) ou HN (tablar de droite), l'expérimentation de Juin présente des plantes vigoureuses (photo de gauche du 07/09/2020), tandis que celles d'Août présentent des plantes moins chargées (photo de droite du 26/10/2020). Sur cette dernière nous pouvons aussi voir une différence de taille entre les deux tablars (parallélogrammes rouges), le tablar le plus avancé présentant un fort retard de croissance.

Biomasse LN Juin



Biomasse HN Juin

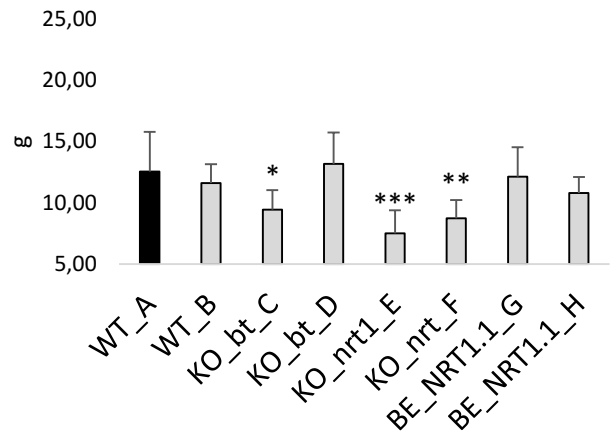


Annexe 8 : Comparaison des biomasses totales (panicules + grains) des lignées Kitaakés

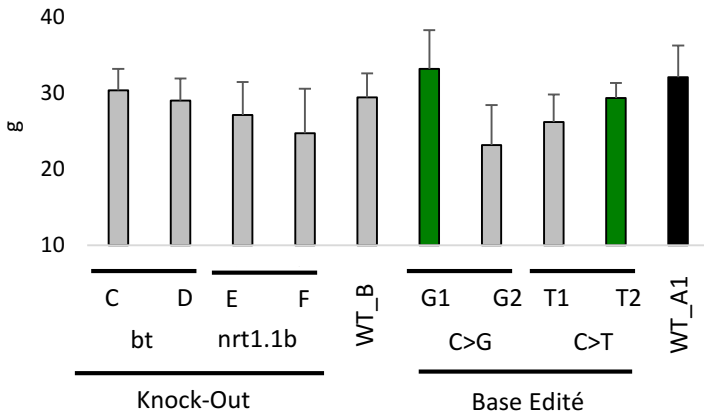
Les biomasses sur HN sont comparées entre LN et HN en Juin, et en HN en Août.

La biomasse en LN de juin est représenté sans pouvoir être comparée à celle d'Août (absente). Ici n=15 individus par génotype.

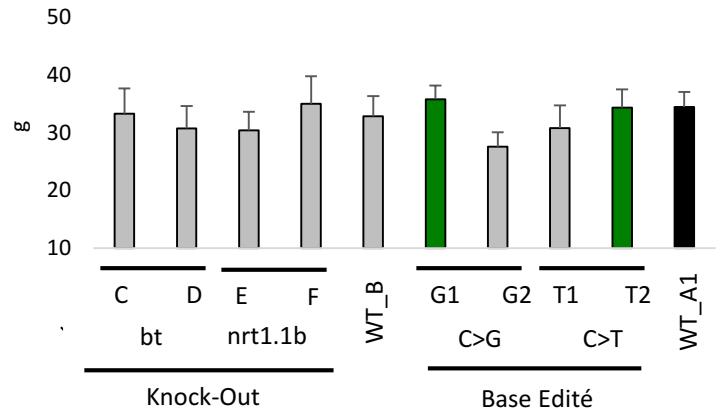
Biomasse HN Aout



Biomasse aérienne en LN



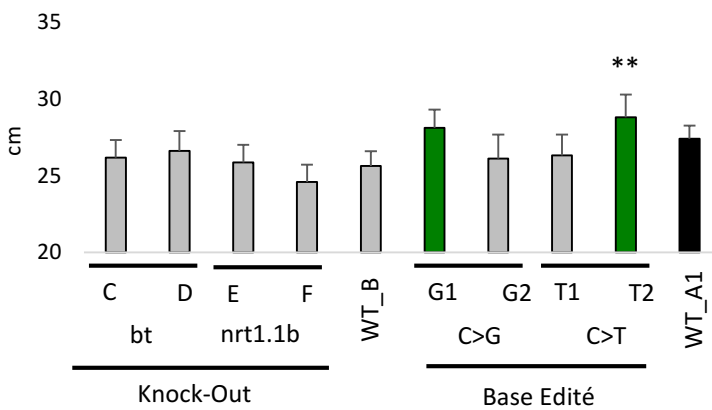
Biomasse aérienne en HN



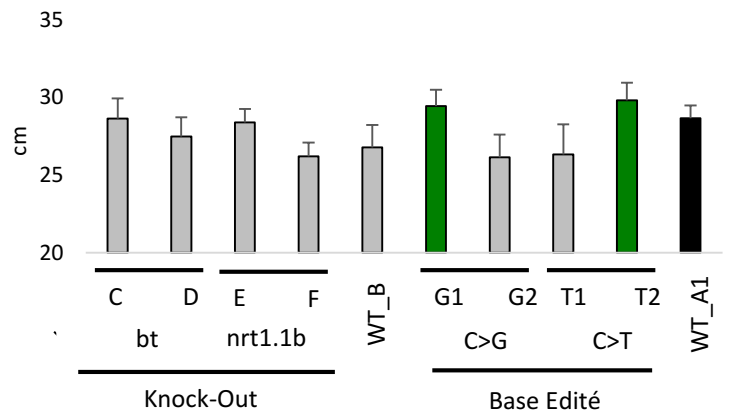
Annexe 9 : Représentation de la biomasse aérienne (en grammes) sans graine chez les lignées CHHD

Les barres vertes représentent les génotypes édité à la base près dans *OsNRT1.1b* qui sont positivement sortis du lot dans les différentes expérimentations.

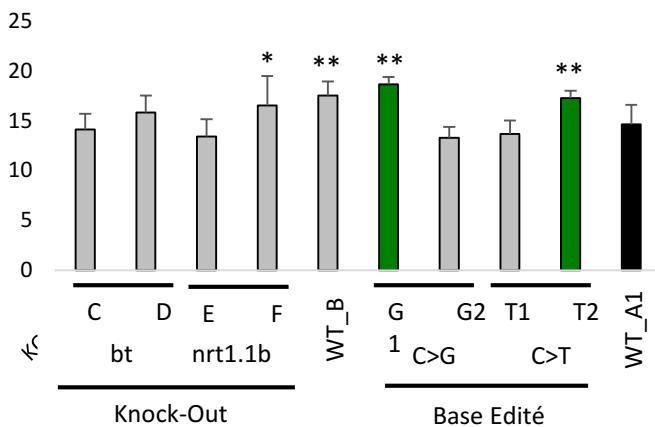
Taille des panicules en LN



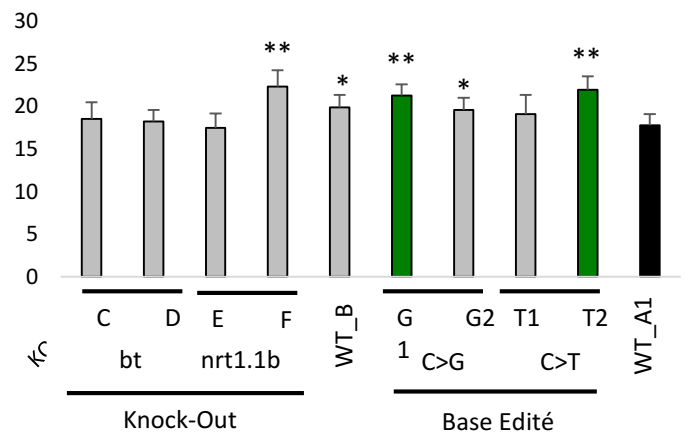
Taille des panicules en HN



Nombre panicules en LN



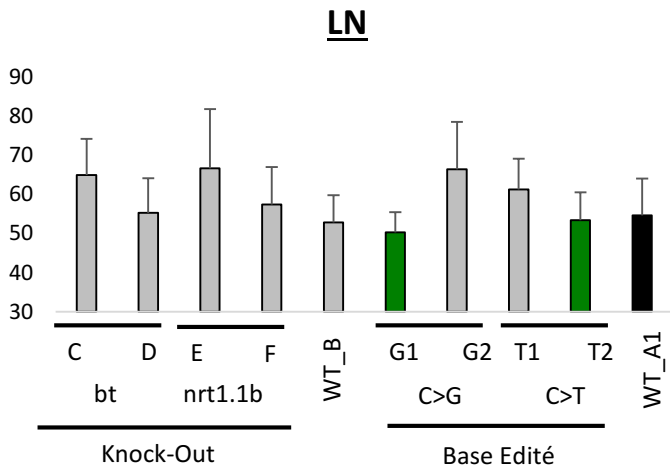
Nombre panicules en HN



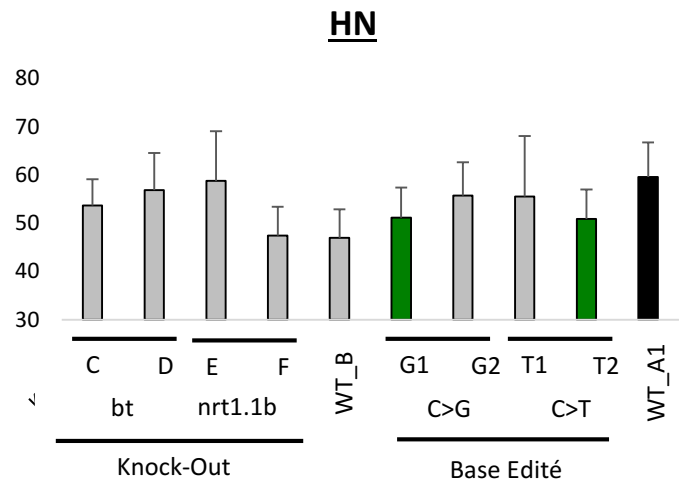
Annexe 10 : Représentation de la taille (cm) et du nombre de panicules chez les lignées CHHD

Les barres vertes représentent les génotypes édité à la base près dans *OsNRT1.1b* qui sont positivement sortis du lot dans les différentes expérimentations.

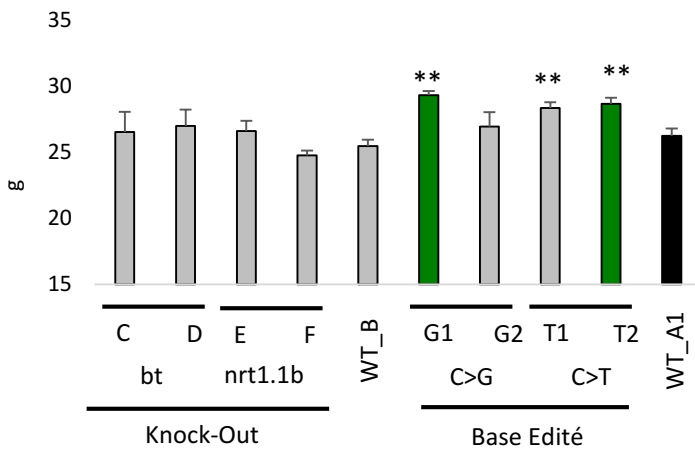
Nombre de grains par panicule en LN



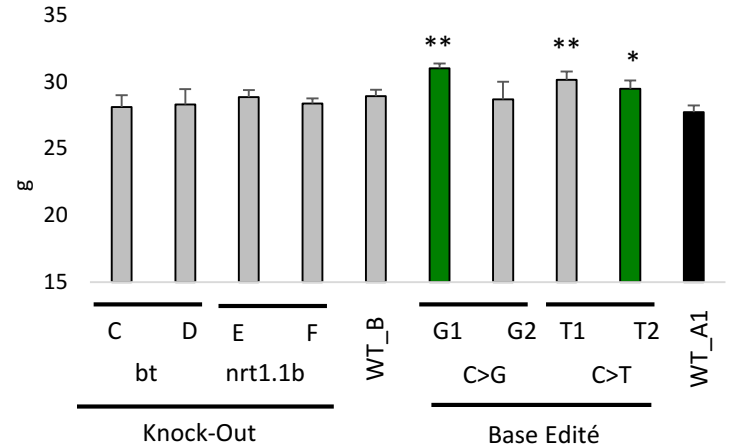
Nombre de grains par panicule en HN



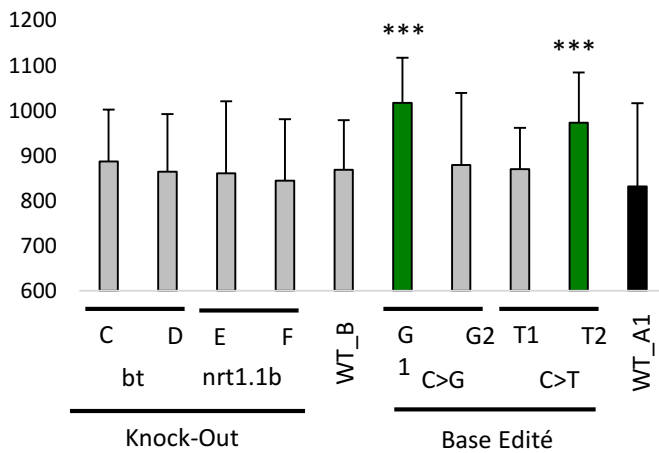
Poids de 1000 grains en LN



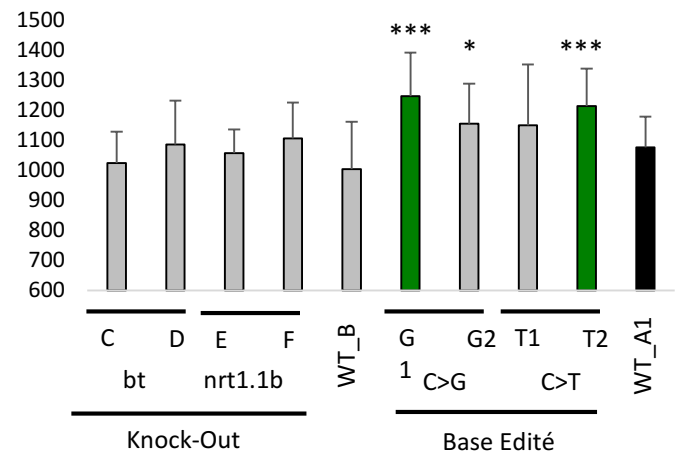
Poids de 1000 grains en HN



Nombre de grains par plante en LN



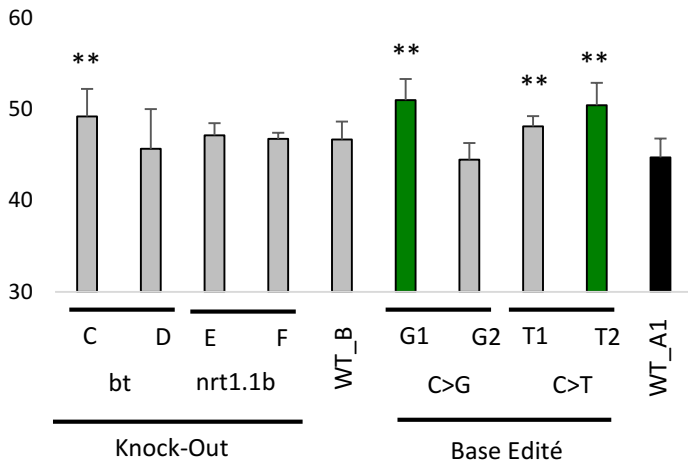
Nombre de grains par plante en HN



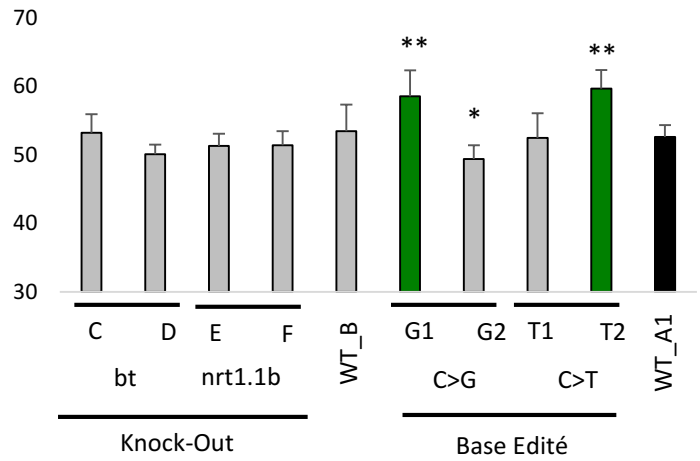
Annexe 11 : Présentation des résultats sur les grains des lignées CHHD (nombre moyen par plante et panicule, poids de 1000 grains en grammes)

Les barres vertes représentent les génotypes édités à la base près dans *OsNRT1.1b* qui sont positivement sortis du lot dans les différentes expérimentations.

SPAD en LN



SPAD en HN



Annexe 12 : SPAD des lignées CHHD pour la teneur en chlorophylle

Les barres vertes représentent les génotypes édités à la base près dans *OsNRT1.1b* qui sont positivement sortis du lot dans les différentes expérimentations.

Test statistique de Dunnett sur 20 individus par condition.



Annexe 13 : Radar des lignées Base Edités T1 et G2

Les Radars représentent les résultats globaux en LN (droite) et HN (gauche) du rendement (Rdt), de l'indice de récolte (HI), de la NUE, du poids de 1000 grains (Pmg), du nombre de grains moyen par panicule (NbGrPan), du nombre de panicules (NbPan).

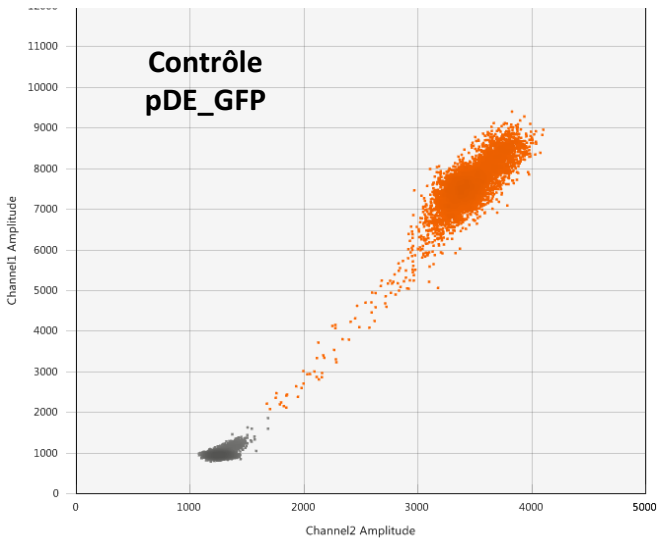
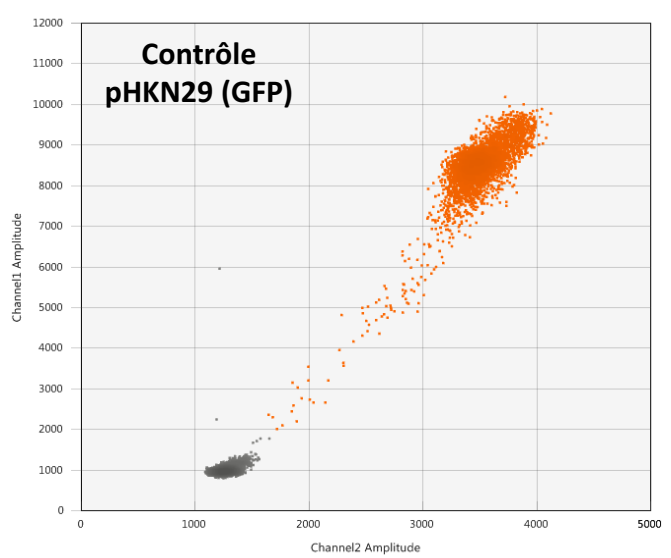
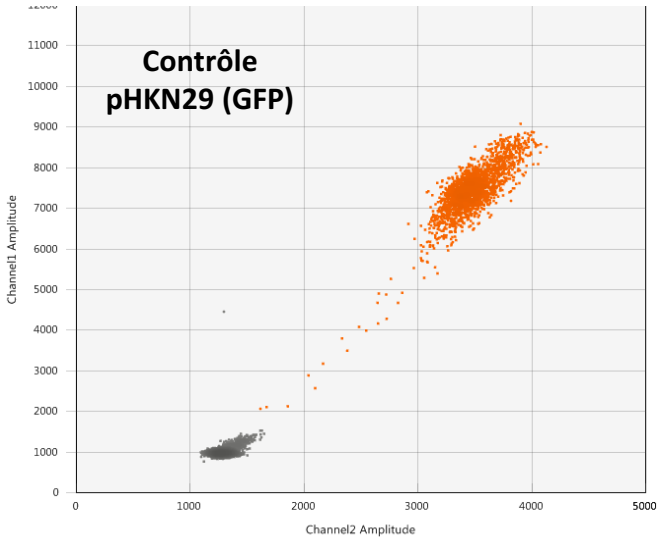
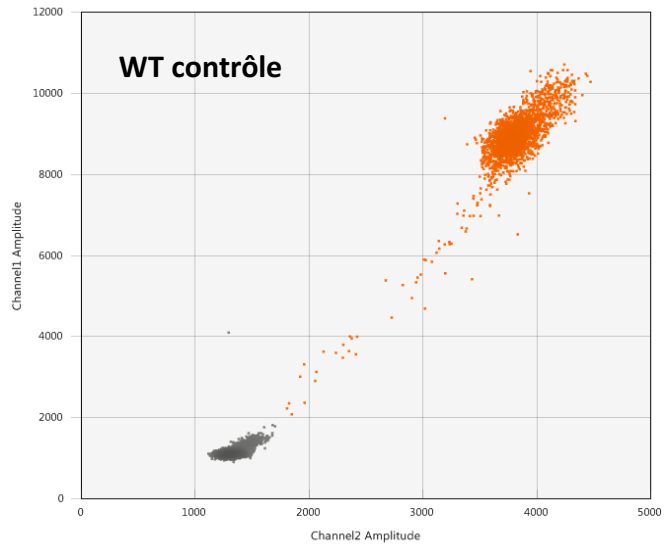
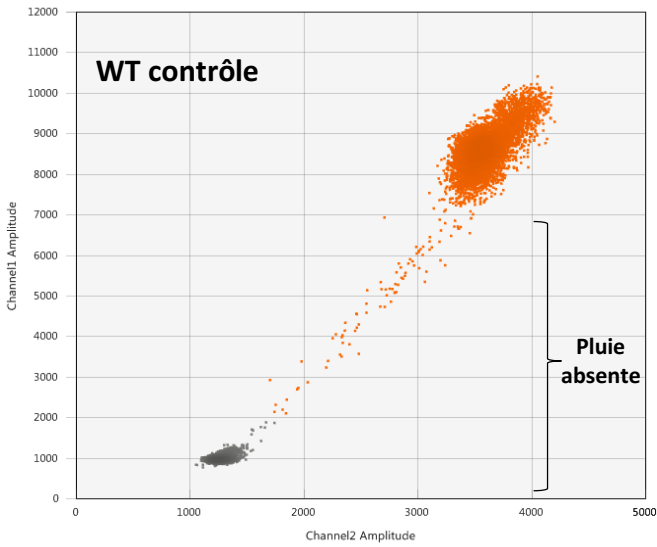
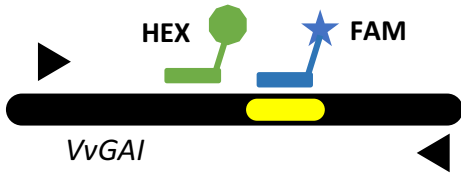
Les plantes éditées à la base près sont en rose et la plante contrôle sauvage WT_A1 sert de référence en bleu à titre de comparaison.

Primer	Séquence (sites de clonage en	CRISPR pour :	Fonctionnels ?
crGAI_265_Bbs1_F	ATTG AGCGACCTCAGCCATGTCGG	CRISPR KO vigne 265 (même que ultiplex) à cloner en Bbs1 sur pEN-Chimera ou pEN-C1,1	oui
crGAI_265_Bbs1_R	AAAC CCGACATGGCTGAGGTCGCT		
crGAI223-bsmb1-F	ATTG GCCGACCCCGAGCAAGACGC	crRNA 2 pour multiplex pEN-Manticora (dicoty) vigne contre GAI	oui
crGAI223-bsmb1-R	AAAC GCGTCTTGCTGGGGGTCGGC		
crGAI206-bsmb1-F	ATTG CCATGCCGGCGTCTTGCTGG	crRNA 2 pour multiplex pEN-Manticora (dicoty) vigne contre GAI	oui
crGAI206-bsmb1-R	AAAC CCAGCAAGACGCCGGCATGG		
VVGAI-CRISP-KO-R	GGCA ACATCTTACCCTTGCCGGTG	CRISPR Ko simplex contre GAI1 chez la vigne (193) en plasmide 217 riz	non
VVGAI-CRISP-KO-F	AAAC CACCGGCAAGGGTAAGATGT		
sgPDS-t2 F	GGCA TTTGTCTACTGCAAAATATT	CRISPR Ko simplex contre VvPDS-t2 (exon 2) et clonage en Base editor (pBE-nCas9) possible en plasmide 217 riz	non en Base editing
sgPDS-t2 R	AAAC AATATTTGCAAGTAGACAAA		
sgRNA_NRT1.1b-F	GGCA CGGCGACGGCGAGCAAGTGG	CRISPR contre NRT1.1 pour le base editing	oui
sgRNA_NRT1.1b-R	AAAC CCACTTGCTCGCCGTCGCCG		
crNRT1.1-Cas_1F	TGTT ACCATGCGGCGATCATCGAC	CRISPR contre NRT1.1 pour délétion	oui
crNRT1.1-Cas_1R	AAAC GTCGATGATCGCCGATGGT		
crNRT1.1-Cas_2F	GTGT CGAAGGTGGTCATCTGGGCG		
crNRT1.1-Cas_2R	AAAC CGCCCAGATGACCACCTTCG		
CrOsBTex1_LFs	GGCA CAAGATGTGCGAGGGTGTGC	CRISPR contre OsBT pour le KO (plantes utilisées)	oui
CrOsBTex1_LRs	AAAC GCACACCCTCGCACATCTTG		
CrOsBTex2_LFs	GGCA AGGGCCGTCGTCCTGATCCG	2eme CRISPR contre OsBT pour le KO (plantes non utilisées par la suite)	oui mais ne sera pas utilisé
CrOsBTex2_LRs	AAAC CGGATCAGGACGACGGCCCT		
pBE_crCAO1-bsa1-F	ggca CAATCAGGTATTGTACTTTC	Base editing de CAO1 pour témoin pBE par apparition codon stop sur séquence CAO1	oui
pBE_crCAO1-bsa1-R	aaac GAAAGTACAATACCTGATTG		

Annexe 14: Primers des CRISPR

Plasmide	Enzyme	Séquence flanquante (rouge) à rajouter aux 20pb
pENTRY simplex 217 (un seul sgRNA, promoteur U3 riz)	BsaI	(5'-3') F : GGCA ATGC(...)TTTT R: AAAC AAAA(...)GCAT
pENTRY multiplex 2B9 (deux sgRNA, promoteur U3 riz)	BtgZI	(5'-3') F : TGTT ATGC(...)TTTT R: AAAC AAAA(...)GCAT
	BsaI	(5'-3') F : GTGT ATGC(...)TTTT R: AAAC AAAA(...)GCAT
pENTRY pEn.C1-1 et pEn-Chimera simplex dicotylédones (un seul sgRNA, promoteur AtU6-26).	BbsI	(5'-3') F : ATTG ATGC(...)TTTT R: AAAC AAAA(...)GCAT
pEN-Manticora multiplex dicotylédones (deux sgRNA, deux promoteurs AtU6-26).	BsmB1 et Bbs1 idem	(5'-3') F : ATTG ATGC(...)TTTT R: AAAC AAAA(...)GCAT

Annexe 15: Plasmides et enzymes de clonage utilisés

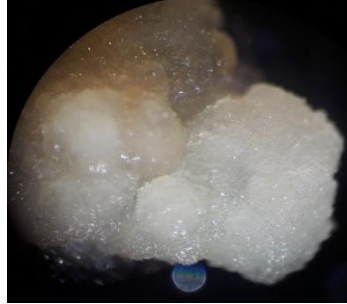


Annexe 16 : Comparaisons de DropOff sur des échantillons WT par ddPCR

Cinq contrôles sont représentés en DropOff pour vérifier la présence de la pluie. Aucune pluie ou droplet verte n'est détectée dans les contrôles WT ou de transformation à l'aide d'un plasmide GFP ou CRISPR-GFP vide.



Chardonnay



Portan



Grenache

Annexe 17 : Choix des cals embryogènes pour la transformation stable

Le choix des cals embryogènes au bon stade est crucial pour la transformation stable. Ici les cals de Chardonnay ont brunis et présentent une structure non embryogène à proscrire. Ceux du Portan semblent trop jeunes. Seuls ceux du Grenache semblent sortir du lot et présenter des structures secondaires pour la transformation stable.

(A)

```

1      10      20      30      40      50      60      70      80      90      100     110     120     130
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
AtU6-26 CTTTTTTCCTTCCTT-CGTTCATACAGTITTTTTTGTGTTATCAGCTTACATTTCTTGACCCGTAGCT-TTCGTTTCTCTCTTTTACCTTCCATTCGGAGTTTTGTATCTGTTCAT
VvU6    TTCCCTCTCAGCTCTCCCACTT-TTCTCATCAATTTTTTGGTTGCATTAATTTCCCAAAATTTGACAAATATGTTCTCCAAATTTGTGTAGTAAARAAACCAATAGAGTT---GGATGTT-TTCCTT
OsU3    AAGGGATCTTTAARCATACGAAACAGATCATTAAAGTT-CTCTGAGGCACTTAAAGTTATCAGGATGCATGGATCTTGGAGGATCAGATGTCCAGTCAGGGACCATAGCACAGAA---CAG
Consensus .....ct..g.TcttccaaCtT.cg.tCAta.a.tTTTt.tgTgctAt.a..t.a.Caaat.TTgaca.g.ag.t.ttc.t.TT.t..t.attAaat.tcCA.tcaGaGtt..tggatcttgtt.Cat

131     140     150     160     170     180     190     200     210     220     230     240     250     260
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
AtU6-26 AGTTTGTCCT-AGGATTAGAAATAGTTAGGCATCGAACCTTCAGGAAATTTGATTGAATAAACATCTTCATTTTARAGATATGAGATATCTTCAAAAGGCCCTGGGAATCTGAAGAGAGAGAGCGAG
VvU6    TGAATTGCTATAGATGAARATATAGTATACAAACAAATGAGAGAGTAAAGAGCAGACATATCCCAATCAATTAAGAAACACAGA---AAGTACCCCTCATTTCAGGTTGGCCACCATTA
OsU3    GCCTCTTACTGGTGTCTACCGCAATGCTGGAGAGCC---GGACACTTGGGTACGTTGGAAACCCAGTGTGAGAGATAGATTAAC---TGTAGGAGAAAGCATTCGTAGTGGCCCATGAGG
Consensus .g.TgTCTa.agGat.aaatg.aaatgca.caAaCc...gaaaa.Tga.taagtagaAaRtctcca.a.ttaAa.aA.gARgatAaac...aaaaggacct.gcAtC.G.a.gGaca.gcAg.

261     270     280     290     300     310     320     330     340     350     360     370     380     390
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
AtU6-26 CCCATTTATATGGAAAGAACATAGTATTTCTTATATAGGCCATTTAAGTTGAARACATCTTCARAGTCCCACATCGCTTAGATAGAAACAGAGCTGAGTT---TATATCAGCTAGAGTCGAGG
VvU6    AAGGCGGTGTCAGGTTGAGTGAARACGGCATCGTTTGGGCAAGCCGGGGTC-TTGACATCCCAATCCCA---TCCACATCGAACTTARAGCCAGATATGTTCCCTCATATATCAGCTAGAGTCGAGG
OsU3    CTTTCAGGACATGATTTGCAGTATGGCCGGCCCATTAC-GCAATTTGGACGACACAAAGACTAGTATTTAGTACCACCTCG---GCTATCCCATAGATCAAGCTGATTTAAAGAGTTGTGCAGATG
Consensus c...c.g..cagGta.aG.a.aAtGg.at.gcttat..c.GCc..tggcaggtgAaAc.ac.ttca.aagTCCACaTCG....g.TAgaacA.a.At.ct.agtT.aTaTtAcagctAaGaGtCGAaG

391     401
|-----+|
AtU6-26 TA--GTGATTG
VvU6    TA--GTGATTG
OsU3    ATCCGTG6CA
Consensus ta..GTGatg

```

(B)

```

1      10      20      30      40      50      60      70      80      90      100     110     120     130
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
AtU6-26 CTTTTTTCCTTCCTTCGTTTCATACAGTTTTTTTTTGTGTTATCAGCTTACATTTCTTGACCCGTAGCT-TTCGTTTCTCTCTTTTACCTTCCATTCGGAGTTTTGTATCTGTTCAT
VvU6    TTCCCTCTCAGCTCTCCCACTTTCCTCATCAATTTTTTGGTTGCATTAATTTCCCAAAATTTGACAAATATGTTCTCCAAATTTGTGTAGTAAARAAACCAATAGAGTT---GGATGTT-TTCCTT
Consensus .....CTcagTccccCaaCTTgcTCAtaAaGTTTTTggTTGcaTAaaagcTcaCaaaaTcTTGaaagaaGct.TTCcaaTTcTgcTgagTAAaaaaCCAaacaGAGTT...GgATGTT.TtCcaTa

131     140     150     160     170     180     190     200     210     220     230     240     250     260
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
AtU6-26 GTTTTGTCCT-AGGATTAGAAATAGTTAGGCATCGAACCTTCAGGAAATTTGATTGAATAAACATCTTCATTTTARAGATATGAGATATCTTCAAAAGGCCCTGGGAATCTGAAGAGAGAGAGCGAGC
VvU6    GATTGTCTATAGATGAARATATAGTATACAAACAAATGAGAGAGTAAAGAGCAGACATATCCCAATCAATTAAGAAACACAGA---AAGTACCCCTCATTTCAGGTTGGCCACCATTA
Consensus GattTGTcCa.aAaGATgAaAATaaaaagacAaCAaACaaacaAaaAagTARaGaAaCaAARCAATaTcCAcaaaTcAaAaaAaGAAaAaaAaa...AARAGaaCCCTcacaATCaGaaaGaaGAcacaCRgaa

261     270     280     290     300     310     320     330     340     350     360     370     380     390
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
AtU6-26 CCAATTTATATGGAAAGAACATAGTATTTCTTATATAGGCCATTTAAGTTGAARACATCTTCARAGTCCCACATCGCTTAGATAGAAACAGAGCTGAGTT---TATATCAGCTAGAGTCGAGG
VvU6    AAGGCGGTGTCAGGTTGAGTGAARACGGCATCGTTTGGGCAAGCCGGGGTC-TTGACATCCCAATCCCA---TCCACATCGAACTTARAGCCAGATATGTTCCCTCATATATCAGCTAGAGTCGAGG
Consensus acacggagaaGgGaaAGaaaARCaGcATcgcTTagagaaGCCcagggaa..TTGAaRaCaacaTTCaaA..TCCACATCGaaaagaTARGAaaRcaAaRgcTgacct..TATATCAGCTAGAGTCGAGG

391     400     410     420     430     440     450     460     470     480     490     499
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
AtU6-26 AGTGATTGGGGCTTCGAGAGACCTGTTTTAGAGCTAGAAATAGCAAGTTAAATAGGCTAGTCCGTTATCACTTGAAARAGTGGCACCAGTCGGTGCTTTTTTT
VvU6    AGTGATTGGGGCTTCGAGAGACCTGTTTTAGAGCTAGAAATAGCAAGTTAAATAGGCTAGTCCGTTATCACTTGAAARAGTGGCACCAGTCGGTGCTTTTTTT
Consensus AGTGATTGGGGCTTCGAGAGACCTGTTTTAGAGCTAGAAATAGCAAGTTAAATAGGCTAGTCCGTTATCACTTGAAARAGTGGCACCAGTCGGTGCTTTTTTT

```



Annexe 18 : sgRNA vigne

Comparaison de la séquence du promoteur du snRNA U6 de *Vitis vinifera* (VvU6) avec celle d'*Arabidopsis thaliana* (AtU6-26) et de riz (OsU3). Les séquences ont été récupérées au génoscope (A).

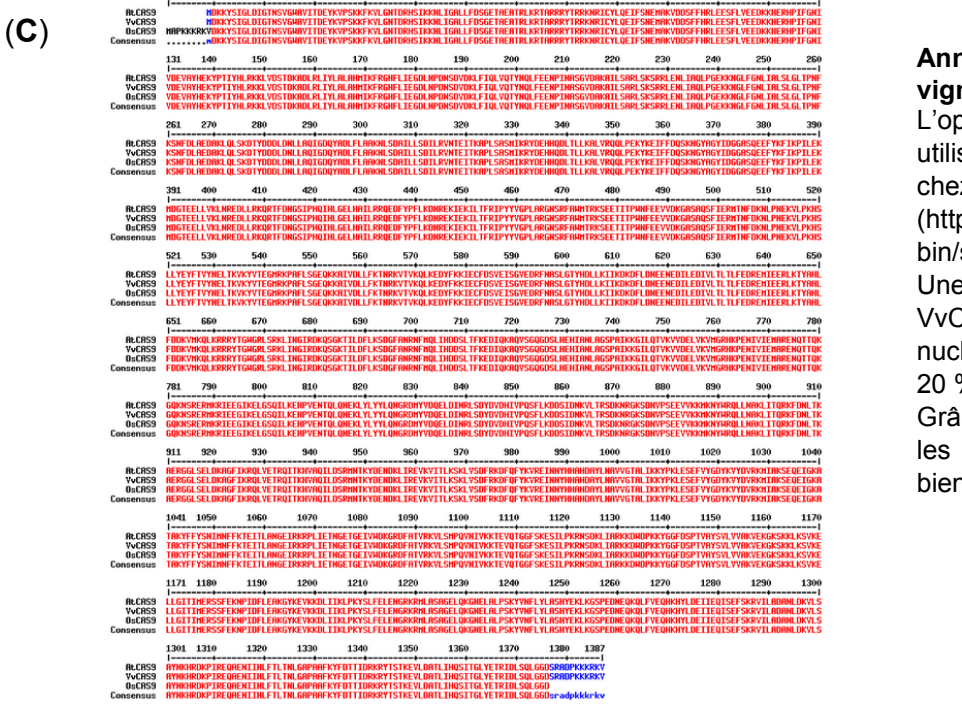
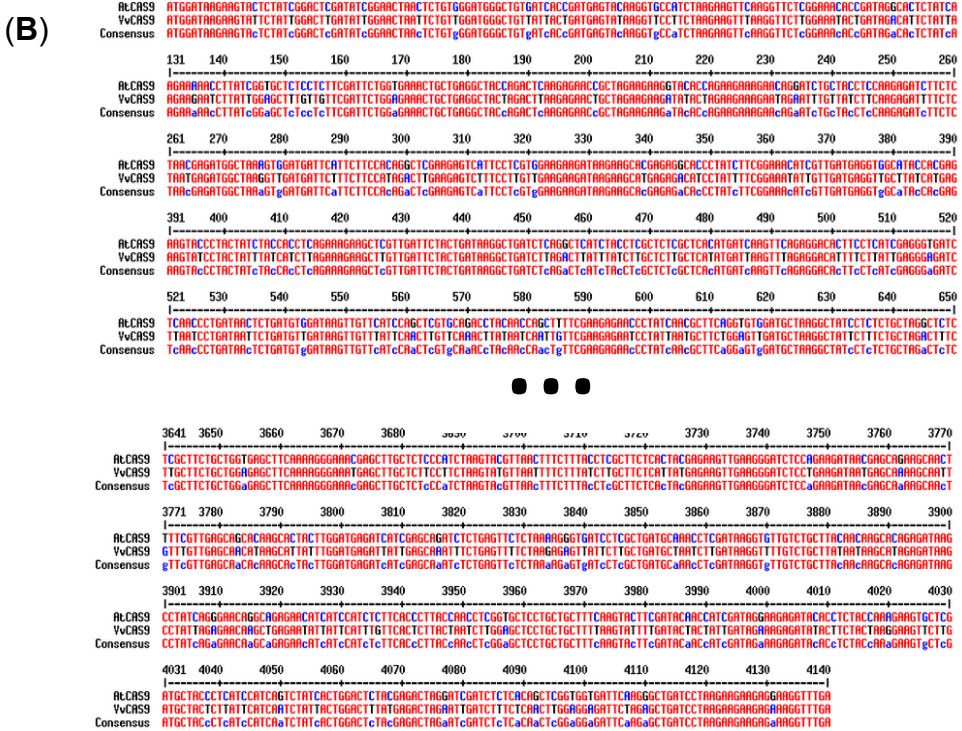
La construction du promoteur VvU6 pour le sgRNA est propre à la vigne jusqu'à la TATA box. Ensuite les 30 pb qui suivent jusqu'aux sites d'insertion de la cible CRISPR (20pb) ont été récupérées de celles du pENTRY_C1.1 (Chimera 1.1) (AtU6-26) pour éviter un décalage du cadre de lecture (B).

La construction finale est insérée dans pEN-C1.1 à la place de l'ATU6-26 (C).

(A) *Vitis vinifera* [gbpln]: 24140 CDS's (11557387 codons)

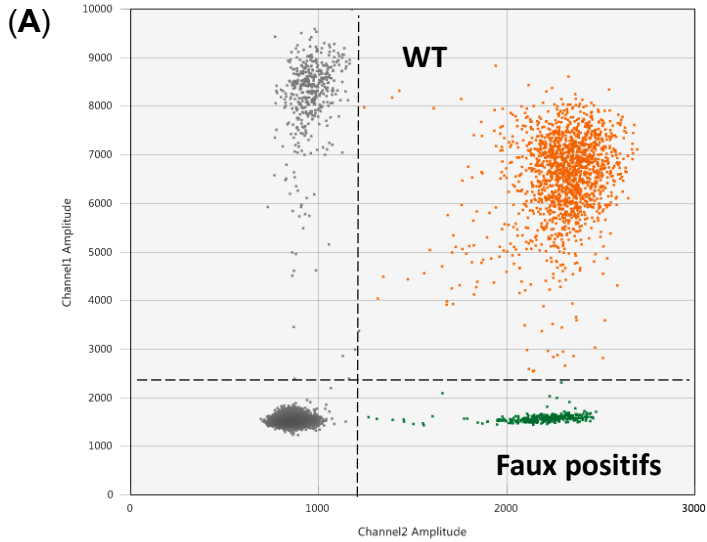
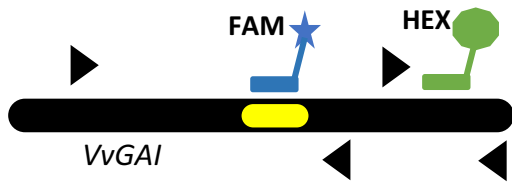
fields: [triplet] [frequency: per thousand] ([number])

UUU 25.2 (291490)	UUC 21.1 (244291)	UAU 17.9 (206846)	UGU 10.8 (125132)
UUC 18.6 (215288)	UCC 12.6 (145986)	UAC 10.5 (121432)	UGC 8.2 (94398)
UUA 13.5 (156031)	UCA 18.7 (215994)	UAA 0.6 (6828)	UGA 0.9 (10455)
UUG 25.1 (289607)	UCG 5.6 (64626)	UAG 0.6 (6815)	UGG 15.6 (180435)
CUU 22.1 (255515)	CCU 18.7 (216148)	CAU 17.9 (206688)	CGU 6.0 (69669)
CUC 13.5 (156457)	CCC 9.1 (105200)	CAC 9.1 (105740)	CGC 4.7 (53766)
CUA 11.2 (128918)	CCA 18.3 (211478)	CAA 22.2 (256139)	CGA 6.6 (76515)
CUG 12.2 (141100)	CCG 5.2 (60461)	CAG 15.9 (184009)	CGG 5.2 (59827)
AUU 26.1 (301597)	ACU 17.7 (204752)	AAU 26.1 (301492)	AGU 14.1 (162715)
AUC 15.7 (181341)	ACC 11.0 (127287)	AAC 14.6 (168187)	AGC 10.3 (118518)
AUA 13.7 (158431)	ACA 15.5 (179134)	AAA 29.2 (337782)	AGA 18.6 (214862)
AUG 26.1 (301968)	ACG 4.3 (49960)	AAG 33.6 (388634)	AGG 16.0 (185039)
GUU 23.3 (268788)	GCU 23.1 (267133)	GAU 34.9 (403518)	GGU 16.8 (194580)
GUC 10.9 (126346)	GCC 13.4 (155209)	GAC 16.1 (186142)	GGC 10.9 (126426)
GUA 10.4 (120284)	GCA 19.8 (229251)	GAA 32.4 (374017)	GGA 20.1 (231880)
GUG 19.6 (226898)	GCg 5.1 (58533)	GAG 32.4 (374052)	GGG 14.7 (169347)



Annexe 19 : Codon optimisation CAS9 vigne

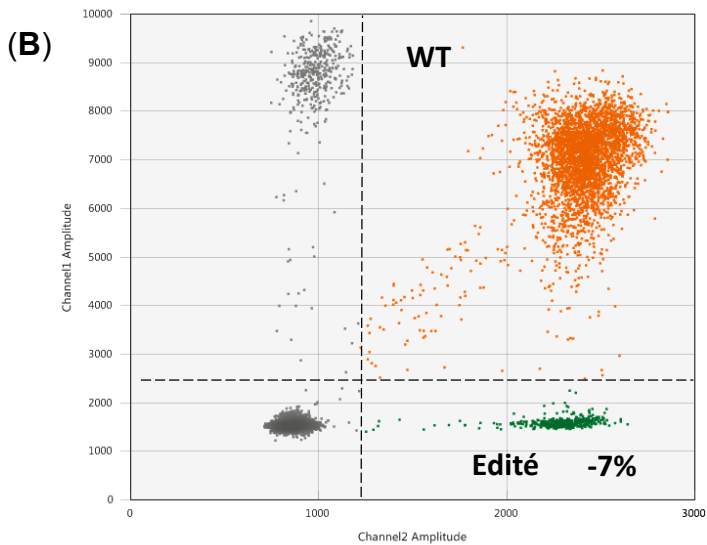
L'optimisation des codons a été réalisée en utilisant des acides aminés les plus fréquents chez la vigne décrits dans la Table de Kazusa (<https://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=29760>) (A). Une partie de la séquence complète de la VvCAS9 est présentée sous forme nucléotidique contre celle de AtCAS9. Environ 20 % des nucléotides ont été modifiés (B). Grâce au code dégénéré des acides aminés les séquences des 2 protéines CAS9 sont bien les mêmes (C).



Annexe 20 : Coamplification par ddPCR – cas d'un CRISPR non fonctionnel en Vigne

L'expérimentation de Coamplification présente trois nuages composés de droplet doubles fluorescents (WT), vertes (faux positifs seuls) et grises chez les protoplastes WT **(A)**. Lors de l'ajout du plasmide CRISPR ces trois nuages réapparaissent mais n'augmentent pas proportionnellement pour les vertes. Ici le biais d'amplification indique entre -7% d'ADN mutés, soit 0% d'édition **(B)**.

La confirmation se fera plus tard avec plus de 100 plantes régénérées mais aucune plante éditée malgré la présence du T-DNA.



Micro MS x100 solution mère	
MnSO4 H2O	3,2g
H3BO3	1,24g
ZnSO4 7H2O	1,72g
Na2MoO4 2H2O	0,05g
CuSO4 5H2O	0,005g
Ki	0,166g
CoCl2 6H2O	0,005g
H2O	qsp 2L
Stérilisation par filtration , 4°C	
Macro MS x10 solution mère	
KNO3	19g
NH4NO3	16,5g
CaCl2 2H2O	4,4g
MgSO4 7H2O	3,7g
KH2PO4	1,7g
H2O	qsp 1L
Stérilisation par filtration , 4°C	
Fer EDTA x200 solution mère	
FeSO4 7H2O	2,276g
Na2EDTA 2H2O	3,725g
H2O	qsp 0,5L
Faire bouillir l'eau puis rajouter les poudres	
Stérilisation par filtration , 4°C, obscurité	

Vitamines (dérivée Gamborg B5)		
	mg/L	µM
myo-inositol	100	554,94
Nicotinic acid	1	8,12
Pyridoxine HCl	1	4,86
Thiamine HCl	10	29,65

GS1Ca (avec Charbon)	
Macroéléments	50mL
Microéléments	10mL
Fer EDTA	5mL
Vitamines	1mL
saccharose	60g
NOA	1mL
BAP	1mL
IAA	20mL à 1 μ M
charbon	2,5g
bacterioagar	10g
H2O	qsp 1L
pH	5,7
Stérilisation par autoclave	
(après autoclave) Glutamine	2mM (292.3 mg pour 1L)

LCM medium	
Macroéléments	50mL
Microéléments	10mL
Fer EDTA	5mL
Vitamines	1mL
saccharose	60g
H2O	qsp 1L
pH	5,7
Stérilisation par autoclave	
RIM (milieu productur de racines)	
Macroéléments	25mL
Microéléments	5mL
Fer EDTA	2,5mL
Vitamines	0,5mL
saccharose	30g
IAA	5 μ M
bacterioagar	7g
H2O	qsp 1L
pH	6
Stérilisation par autoclave	

Induction medium Agrobactéries	
NH ₄ Cl	10g
MgSO ₄ 7H ₂ O	6,15g
KCl	1,5g
CaCl ₂	0,132g
FeSO ₄ 7H ₂ O	0,25g
Biotine	2,5µg (5µL à 500µg/mL)
NaH ₂ PO ₄	0,12g
MES	4,264g
glucose	2,5g
Acétylsyringone	100µM (500µL 100µM)
H ₂ O	qsp 0,5L
pH	3,9 environs naturellement
Stérilisation par filtration	

Induction Cal embryogène	
Macroéléments	50mL
Microéléments	5mL
Fer EDTA	2,5mL
Vitamines	1mL
saccharose	60g
BAP (8,9µM)	1mL à 5µmol/L + 4mL à 1µmol/L
2,4-D	900mL à 5µmol/L
Gelrite	3g
H ₂ O	qsp 1L
pH	5,8
Stérilisation par autoclave	

Maintien Cal embryogène	
Macroéléments	50mL
Microéléments	5mL
Fer EDTA	2,5mL
Vitamines	1mL
saccharose	30g
BAP (1µM)	1mL
2,4-D	5µmol
Gelrite	2,5g
H ₂ O	qsp 1L
pH	5,8
Stérilisation par autoclave	

Solution	Sels	Solution Mère	
MgK	KH ₂ PO ₄	136 g	10 L
	MgSO ₄	246 g	
Fer	NaFeEDTA	72 g	2L
Nitrate	KNO ₃	80,88 g	4L
	Ca(NO ₃) ₂	94,4 g	

Solution	Sels	Poid	Solution- mère (g /L)	Concentration	Concentration	Volume
		Moléculaire		Solution Mère	Final	Pour 10L
Oligo- elements	MnSO ₄ ·H ₂ O	169.02	1.690	10mM	20µM	20mL
	(NH ₄) ₆ Mo ₇ O ₂₄	1163	0.186	0.16mM	0.32µM	
	ZnSO ₄ ·7H ₂ O	287.54	0.201	0.7mM	1.4µM	
	CuSO ₄ ·5H ₂ O	249.68	0.199	0.8mM	1.6µM	
	H ₃ BO ₃	61.83	1.397	22.6mM	45.2µM	

Pour 10L de milieu hydroponie à :	1mM KNO ₃	0,5mM KNO ₃	5mM KNO ₃
MgSO ₄ /KH ₂ PO ₄	100mL		
solution Fer	10mL		
solution Oligos	20mL		
KNO ₃ /CA(NO ₃) ₂	25mL	12,5mL	125mL
pH=5,7			

		pour 200ml de solution
Mannitol	600mM	21,86 g
Mes	2mM	0.076 g

pH 5.7

SD (Solution Digestion)		pour 200ml de solution
CaCl ₂ , 2H ₂ O (MW 147)	7.5mM	222mg
Mannitol (MW 182.17)	600mM	21,86 g
MES buffer (MW 195,24)	10mM	390mg

pH 5.7

W5		pour 1000ml
NaCl (MW 58.4)	154mM	8.92g
CaCl ₂ , 2H ₂ O (MW 147)	125mM	18.36g
KCl (MW 74.55)	5mM	0,372g
MES buffer (MW 195,24)	2mM	0,388g

pH 5,7 (osmolarité = 634.83mosmol/L)

Solution enzymatique		pour 20ml de solution
Macerozyme R10 Onozuka(ref duchefa 00651803)	0,75%	150mg
Cellulase RS Onozuka (ref Duchefa 00578704)	1,5%	300mg
BSA 10%	0,1%	200ul solution 100X
QSP Solution de digestion (SD)		20ml

MMg		pour 200ml
Mannitol	400mM	14,56g
MgCl ₂ , 6H ₂ O (MW 203.3)	15mM	0,61g
MES buffer (MW 195,24)	4mM	0,160g
Cacl ₂ 2H ₂ O	5mM	111mg

RT pH 5,7 (osmolarité = 464.08mosmol/L)

PEG		pour 10ml
Mannitol	200mM	0,364g X
CaCl ₂	100mM	1mL de solution 1M
PEG4000 Fluka®	40%	4g X
Ajuster avec KOH à pH 5.7		

Annexe 23: Milieux protoplastes Riz – stérilisation par filtration

Milieu K4 (1L)
3,19 g de poudre de macroéléments Gamborg (Gamborg B5 medium including vitamins, Duchefa G0210),
136.92 g de saccharose (0,4 M final)
Ajouter 800 ml d'eau distillée
pH ajusté à 5,7 avec du KOH 0,1N
qsp 1 l d'eau distillée
Stériliser par autoclavage
Solution de digestion
Milieu K4 stérile + 1,2 % de cellulase R10 Onozuka (Duchefa, C8001) + 0,6 % macerozyme R10 Onozuka (Duchefa, M8002).
Stériliser par filtration 0.22 µm

Milieu W5 (1L)
154 mM NaCl : 8,99g
125 mM CaCl ₂ : 13,87g
5 mM KCl : 0,37g
5 mM Glucose : 0,9g
Ajouter 800 ml d'eau
pH ajusté à 5,7 avec du KOH 0,1N
qsp 1 l d'eau distillée
Autoclaver

Solution MMG (200ml)
0,4 M Mannitol : 14.56 g
4 mM MES : 1 g
15 mM MgCl ₂ : 0.61 g
5 mM Ca Cl ₂ : 111 mg
pH ajusté à 5,7 avec du KOH 0,1N
qsp 200 ml d'eau distillée
Stériliser par filtration

Solution PEG (10 ml)
40 % PEG 6000 : 20 g
0,4 M D-mannitol : 3,6 g
0,1 M Ca(NO ₃) ₂ 4H ₂ O : 1,18 g
qsp 50 ml d'eau distillée
Ajuster le pH à 5,7 – 6,0 max (attention pH très difficile à stabiliser), Stériliser par filtration

Composition des solutions mères pour la préparation du milieu de base type NB (N6/B5)					
Nom de la solution	PM g/L	Composés chimiques	Poids du composé chimique pour une solution stock d'un Litre (20 X)	Concentration finale du composé chimique par Litre (1 X)	Volume (20 X) à prendre pour 1 Litre de milieu (1 X final)
Macroéléments N6	101,1	KNO ₃	56,6 g	2830 mg	50 ml
	132,1	(NH ₄) ₂ SO ₄	9,26 g	463 mg	
	136,1	KH ₂ PO ₄	8,0 g	400 mg	
	147	CaCl ₂ .2H ₂ O	3,3 g	165 mg	
	246,5	MgSO ₄ .7H ₂ O	3,7g	185 mg	
Nom de la solution	PM g/L	Composés chimiques	Poids du composé chimique pour une solution stock d'un Litre 100 X	Concentration finale du composé chimique par Litre (1 X)	Volume (100 X) à prendre pour 1 Litre de milieu (1 X final)
Microélément B5	169	MnSO ₄ .H ₂ O	758 mg	10 mg	10 ml
	61,8	H ₃ BO ₃	300 mg	3 mg	
	287,5	ZnSO ₄ .7H ₂ O	200 mg	2 mg	
	166	KI	75 mg	0,75 mg	
	241,9	Na ₂ MoO ₄ .2H ₂ O	25 mg	0,25 mg	
	249,7	CuSO ₄ .5H ₂ O	2,5 mg (a)	0,025 mg	
	237,9	CoCl ₂ .6H ₂ O	2,5 mg (b)	0,025 mg	
Nom de la solution	PM g/L	Composés chimiques	Poids du composé chimique pour une solution stock d'un Litre 100 X	Concentration finale du composé chimique par Litre (1 X)	Volume (100 X) à prendre pour 1 Litre de milieu (1 X final)
Fer-EDTA N6/B5	372	Na ₂ EDTA	3,72 g	3,72 mg	10 ml
	278	FeSO ₄ .7H ₂ O	2,78 g	27,8 mg	
Nom de la solution	PM g/L	Composés chimiques	Poids du composé chimique pour une solution stock d'un Litre 100 X	Concentration finale du composé chimique par Litre (1 X)	Volume (100 X) à prendre pour 1 Litre de milieu (1 X final)
Vitamines B5	123,1	Acide Nicotinique	100 mg	1 mg	10 ml
	205,6	Pyridoxine-HCl	100 mg	1 mg	
	337,3	Thiamine-HCl	1000 mg	10 mg	

Composition du Milieu de culture d'induction NB (N6/B5)-2,4-D			
Nom du milieu de culture	Nom de la solution stock et du composant à prendre	Volume ou poids à prendre pour 500 ml	Volume ou poids à prendre pour 1 Litre
NB - induction	N6 Macro	25 ml	50 ml
	B5 Micro	5 ml	10 ml
	Fer-EDTA (N6/B5)	5 ml	10 ml
	B5 Vitamines	5 ml	10 ml
	L-Proline	250 mg	500 mg
	L-Glutamine	250 mg	500 mg
	Hydrolysate de caseine	150 mg	300 mg
	Myo-inositol	50 mg	100 mg
	Saccharose	15 g	30 g
	2,4 - D	1,25 mg	2,5 mg
	Phytigel	1,3 g	1,3 g
	pH = 5,8		

Composition du Milieu de culture NB Sélection (NB - S)			
Nom du milieu de culture	Nom de la solution stock et du composant à prendre	Volume ou poids à prendre pour 500 ml	Volume ou poids à prendre pour 1 Litre
NB - S (sélection II)	N6 Macro	25 ml	50 ml
	B5 Micro	5 ml	10 ml
	Fer-EDTA (N6/B5)	5 ml	10 ml
	B5 Vitamines	5 ml	10 ml
	L-Proline	250 mg	500 mg
	L-Glutamine	250 mg	500 mg
	Hydrolysate de caseine	150 mg	300 mg
	Myo-inositol	50 mg	100 mg
	Saccharose	15 g	30 g
	2,4 - D	1,25 mg	2,5 mg
	Céfotaxime	400 mg/ 0,5 ml	400 mg/1 ml
	Vancomycine	100 mg/ 0,5 ml	100 mg/1 ml
	si agent sélectif "Hygromycine"	50 mg/ 0,5 ml	50 mg/1ml
	si agent sélectif "Généticine"	300 mg/ 0,5 ml	300 mg/1 ml
	Phytigel	3,5 g	7 g
		pH = 5,8	

Composition des solutions mères pour la préparation du milieu de base type R2					
Nom de la solution	PM g/L	Composés chimiques	Poids du composé chimique pour une solution stock d'un Litre (10 X)	Concentration finale du composé chimique par Litre (1 X)	Volume (10 X) à prendre pour 1 Litre de milieu (1 X final)
Macroéléments R2-I	101,1	KNO ₃	40 g	4000 mg	100 ml
	132,1	(NH ₄) ₂ SO ₄	3,3 g	330 mg	
	138	NaH ₂ PO ₄ .H ₂ O	3,12 g	312 mg	
	246,5	MgSO ₄ .7H ₂ O	2,46g	246 mg	
Macroéléments R2-II	147	CaCl ₂ .2H ₂ O	1,46 g	146 mg	100 ml
Nom de la solution	PM g/L	Composés chimiques	Poids du composé chimique pour une solution stock de 100 ml 1000X	Concentration finale du composé chimique par Litre (1 X)	Volume (1000 X) à prendre pour 1 Litre de milieu (1 X final)
Microéléments R2	169	MnSO ₄ .H ₂ O	160 mg	1,6 mg	1 ml
	287,5	ZnSO ₄ .7H ₂ O	220 mg	2,2 mg	
	61,8	H ₃ BO ₃	283 mg	2,83 mg	
	249,7	CuSO ₄ .5H ₂ O	19,5 mg	0,2 mg	
	241,9	Na ₂ MoO ₄ .2H ₂ O	12,5 mg	0,125 mg	
Nom de la solution	PM g/L	Composés chimiques	Poids du composé chimique pour une solution stock d'un Litre (100 X)	Concentration finale du composé chimique par Litre (1 X)	Volume (100 X) à prendre pour 1 Litre de milieu (1 X final)
Fer-EDTA R2	372	Na ₂ -EDTA	0,177 g	1,8 mg	10 ml
	278	FeSO ₄ .7H ₂ O	1,25 g	12,5 mg	
Nom de la solution	PM g/L	Composés chimiques	Poids du composé chimique pour une solution stock d'un Litre (40 X)	Concentration finale du composé chimique par Litre (1 X)	Volume (40 X) à prendre pour 1 Litre de milieu (1 X final)
Vitamines R2	337,3	Thiamine HCL	40 mg	1,0 mg	25 ml

Composition de Milieu de coculture liquide R2 (R2-CL)			
Nom du milieu de culture	Nom de la solution stock et du composant à prendre	Volume ou poids à prendre pour 500 ml	Volume ou poids à prendre pour 1 Litre
R2 - CL (liquide)	R2 Macro I	50 ml	100 ml
	R2 Macro II	50 ml	100 ml
	R2 Micro	0,5 ml	1 ml
	R2 Fer-EDTA	5 ml	10 ml
	R2 Vitamines	12,5 ml	25 ml
	L-Glucose	5 g	10 g
	2,4 - D	1,25 mg	2,5 mg
	Acétosyringone	50 µM	100 µM
	pH = 5,2 stérilisation par filtration		

Composition de Milieu de coculture solide R2 (R2-CS)			
Nom du milieu de culture	Nom de la solution stock et du composant à prendre	Volume ou poids à prendre pour 500 ml	Volume ou poids à prendre pour 1 Litre
R2 - CS (solide)	R2 Macro I	50 ml	100 ml
	R2 Macro II	50 ml	100 ml
	R2 Micro	0,5 ml	1 ml
	R2 Fer-EDTA	5 ml	10 ml
	R2 Vitamines	12,5 ml	25 ml
	L-Glucose	5 g	10 g
	2,4 - D	1,25 mg	2,5 mg
	Acétosyringone	50 µM	100 µM
	Phytigel	3,5 g	7 g
pH = 5,2			

Composition de Milieu de culture type R2 Sélection (R2 - S)			
Nom du milieu de culture	Nom de la solution stock et du composant à prendre	Volume ou poids à prendre pour 500 ml	Volume ou poids à prendre pour 1 Litre
R2 - S (sélection I)	R2 Macro I	50 ml	100 ml
	R2 Macro II	50 ml	100 ml
	R2 Micro	0,5 ml	1 ml
	R2 Fer-EDTA	5 ml	10 ml
	R2 Vitamines	12,5 ml	25 ml
	Saccharose	15 g	30 g
	2,4 - D	1,25 mg	2,5 mg
	Céfotaxime	400 mg/ 0,5 ml	400 mg/1 ml
	Vancomycine	100 mg/ 0,5 ml	100 mg/1 ml
	si agent sélectif "Hygromycine"	50 mg/ 0,5 ml	50 mg/1 ml
	si agent sélectif "Généticine"	300 mg/ 0,5 ml	300 mg/1 ml
	Phytigel	3,5 g	7 g
	pH = 5,8		

Annexe 25: Milieux transformation riz (2)

Composition du Milieu de culture PRN-S (milieu de pré régénération)

Nom du milieu de culture	Nom de la solution stock et du composant à prendre	Volume ou poids à prendre pour 500 mL	Volume ou poids à prendre pour 1 Litre
PRN-S (Sélection III)	N6 Macro	25 ml	50 ml
	B5 Micro	5 ml	10 ml
	Fer-EDTA (N6/B5)	5 ml	10 ml
	B5 Vitamines	5 ml	10 ml
	L-Proline	250 mg	500 mg
	L-Glutamine	250 mg	500 mg
	Hydrolysat de caseine	150 mg	300 mg
	Myo-inositol	50 mg	100 mg
	Saccharose	15 g	30 mg
	ABA	2,5 mg	5 mg
	BAP	1 mg	2 mg
	ANA	0,5 mg	1 mg
	Céfotaxime	100 mg/0,125 ml	100 mg/0,25 ml
	Vancomycine	100 mg/0,5 ml	100 mg/1 ml
	si agent sélectif "Hygromycine"	50 mg/0,5 ml	50 mg/1ml
	si agent sélectif "Généticine"	300 mg/0,5 ml	300 mg/1 ml
Phytigel	3,5 g	7 g	
	pH = 5,8		

Composition du Milieu de culture RN (milieu de régénération)

Nom du milieu de culture	Nom de la solution stock et du composant à prendre	Volume ou poids à prendre pour 500 mL	Volume ou poids à prendre pour 1 Litre
RN (régénération)	N6 Macro	25 ml	50 ml
	B5 Micro	5 ml	10 ml
	Fer-EDTA (N6/B5)	5 ml	10 ml
	B5 Vitamines	5 ml	10 ml
	L-Proline	250 mg	500 mg
	L-Glutamine	250 mg	500 mg
	Hydrolysat de caseine	150 mg	300 mg
	Myo-inositol	50 mg	100 mg
	Saccharose	15 g	30 g
	BAP	1,5 mg	3 mg
	ANA	0,25 mg	0,5 mg
	Phytigel	1,75 à 2,25 g	3,5 à 4,5 g
		pH = 5,8	

Composition de Milieu de culture type P (milieu d'enracinement)

Nom du milieu de culture	Nom du composant stock à prendre	Poids à prendre pour 500 ml	Poids à prendre pour 1 Litre
Plant (enracinement)	Basal Salt Mixture (MS)	2,15 g	4,3 g
	Vitamine (MS)	51,5 mg	103 mg
	Saccharose	25 g	50 g
	Phytigel	1,3 g	2,6 g
		pH = 5,8	

Composition du Milieu de culture d'Agrobacterium AB

Tampon AB stock 20x	K2HPO4 60g/L
	NaH2PO4 20g/L
Autoclaver	
Sels AB stock 20x	NH4Cl 20g/L
	MgSO4.7H2O 6g/L
	KCl 3g/L
	CaCl2 0.20 g/L
	FeSO4.7H2O 50mg/L
Autoclaver	
Pour 1L de milieu AB, mélanger 5g de glucose, 15g de Difco bacto agar et 900 ml d'eau bidistillée et autoclaver. Ajouter ensuite 50ml de tampon AB autoclavé et 50ml de sels AB. Ajouter les antibiotiques appropriés à la souche et couler les boîtes de Pétri.	

Annexe 25: Milieux transformation riz (3)

Primer	Séquence	Utilisation pour séquençage de :
Verif BtgZI	CGTTGAGGGGAGACAGGTTT	vérification du clonage multiplex en 2B9 riz
UbiR-AC	TGCAATGCTCATTATCTCTAGAG	vérification de la LR en 2G9 riz ou pBE
M13F	TGTAACGACGCGCCAGT	séquençages divers (pENTRY ...)
M13R	CAGGAAACAGCTATGACCATG	séquençages divers (pENTRY dicot ...)
Manti-BsmB-LacO-R	GGAATTGTGAGCGGATAACAATT	sequençage 2eme crRNA pEN MANTICORA BsmB1
pDE_verif_LR_F	TGAGCTCCCTAGGCCTGTTA	Primer Forward de séquençages des plamides dicotylédons pDE-Hpt et Kana pour vérif LR

Annexe 26: Primers de séquençage

Primer	Séquence	Utilisation pour séquençage de :
OsNRT1.1big F3	ATCGGCCGTCTAGGGATATT	533pb autour du gène OsNRT1.1b
OsNRT1.1big R3	GTAGAAGGGGACGGTGAGGA	
OsBT_F1	GTAAGCGACGCGCGACGCC	Couples PCR F1-R1 (409 pb) et F2-R2 (771 pb) dans le gène OsBT et des sites de mutations des deux CRISPR (couple 1 pour CRISPR1 (plantes étudiées) et couple 2 pour CRISPR2, attention ne serviront pas)
OsBT_R1	CGGCGCTGCAGGATGCTCTCCA	
OsBT_F2	CTGGTTGGCGTCGGGCGTGT	
OsBT_R2	GCCGGGCTGCTTGGATTGGAC	
Cao1 F1	CGGGACGCTGATGACGTGC	Ampli large de OsCao1 (1353pb)
Cao1 R2	CGCTGTCTTTGCTGCCGA	Ampli interne de OsCao1 (855pb)
Cao1 F3	ATGACGTGCTTGGAGCAAGG	
Cao1 R3	CCCCAGTAACACAGGCTTA	
GAI-F1 U54	AATGTGGTCTCCAGCTGAT	PCR auour de VvGAI , 1177pb (design Valérie Laucou)
GAI-R1 L1212	TGGCTACCTCCGCATGG	
PDS-t2 F	ACCATAAAATGATTATGTAATGCAAA	Amorces autour de l'exon 2 du gène (307pb)
PDS-t2 R	TGTCTTATAATAACCAATAAGGGGAGA	
HYG F	CTGAACTACCGCGACGTCTG	ampli pour génotypage Hygro riz 2G9 (1200pb) /pBE (980pb)
HYG R	GGCGTCGGTTTCCACTATCG	
pBE-nCAS9-verif-F	TCGGGAACCTCATCGCTCTC	ampli pour génotypage dans nCAS9 du pBE riz (taille = 449pb ; tm=65°C)
pBE-nCAS9-verif-R	CTCCGTGCCGTCCATCTTCT	
oscas9F2	TGCCTGCGGAGGATAGCATGAAGCTC	ampli pour génotypage dans CAS9 2G9
oscas9R2	TACCACGAGAAGTACCCGACCATCT	

Annexe 27: Primers PCR

<u>osNRT1.1b</u> <u>F3 R3</u>		Taille	Taq	mix	Volume (µL)	<u>VvPDS t2 F1</u> <u>R1</u>		Taille	Taq	mix	Volume (µL)
		533	Diamond	10x	5			307	Diamond	10x	5
				dNTP	3					dNTP	3
°C	temps			MgCl2	2,5	°C	temps			MgCl2	2,5
94	5min			F	2,5	94	5min			F	2,5
94	45s	x35		R	2,5	94	45s	x35		R	2,5
62,5	45s		H2O	24,25	58,1	30s	H2O		24,25		
72	55s		Diamond	0,25	72	30s	Diamond		0,25		
72	5min		ADN	10	72	5min	ADN		10		
10	O/N					10	O/N				

<u>osBT F1 R1</u> <u>(17/01/18)</u>		Taille	Taq	mix	Volume (µL)	<u>VvGAI F1</u> <u>u54 R1</u> <u>I1212</u>		Taille	Taq	mix	Volume (µL)
		409	Diamond	10x	5			1158	Diamond	10x	5
				dNTP	3					dNTP	3
°C	temps			MgCl2	2,5	°C	temps			MgCl2	2,5
94	5min			F	2,5	94	5min			F	2,5
94	45s	x35		R	2,5	94	45s	x35		R	2,5
72	45s		H2O	24,25	60,3	45s	H2O		24,25		
72	30s		Diamond	0,25	72	1min20s	Diamond		0,25		
72	5min		ADN	10	72	5min	ADN		10		
10	O/N					10	O/N				

<u>osCAO1 F1</u> <u>R2</u> <u>(17/01/18)</u>		Taille	Taq	mix	Volume (µL)	<u>pBE-nCas9-</u> <u>Vérif F/R</u> <u>(28/01/19)</u>		Taille	Taq	mix	Volume (µL)
		1353	KOD	10x	5			449	Diamond	10x	2,5
				dNTP	5					dNTP	1,5
°C	temps			MgSO4	5	°C	temps			MgCl2	1,25
95	2min			F1	1,5	94	5min			F	1,25
95	20s	x35		R2	1,5	94	45s	x35		R	1,25
72	10s		H2O	21,5	66	30	H2O		16,625		
70	17s		KOD	0,5	72	40	Diamond		0,13		
70	1min		ADN	10	72	5min	ADN		1		
10	O/N					10	O/N				

<u>osCAO1 F3</u> <u>R3</u> <u>(17/01/18)</u>		Taille	Taq	mix	Volume (µL)	<u>Hygro F/R</u> <u>(28/01/19)</u>		Taille	Taq	mix	Volume (µL)
		885	KOD	10x	5			986	Diamond	10x	2,5
				dNTP	5					dNTP	1,5
°C	temps			MgSO4	5	°C	temps			MgCl2	1,25
95	2min			F3	1,5	94	5min			F	1,25
95	20s	x35		R3	1,5	94	45s	x35		R	1,25
72	10s		H2O	21,5	68	30	H2O		16,625		
70	17s		KOD	0,5	72	40	Diamond		0,13		
70	1min		ADN	10	72	5min	ADN		1		
10	O/N					10	O/N				

Annexe 28: Optimisations PCR

Nom	Séquence	Taille (pb)
dp-qPCR 33-F	TGGACAAAGTTTTCAACCGGC	153
dp-qPCR 33-R	GTACCCTCATCAGGCATC	
Probe HEX-qPCR33	CTACCTGTGGTTCTCACTTGTA	
Probe FAM-qPCR33	CTACCTGTGGTTCTCACTTGTA	
dp-F1 NRT1.1	GTCGGTTTGGACCGATAAGA	243
dp-R1 NRT1.1	ACCACCTGCTTCACCTCCT	
Probe FAM NRT1.1b-T	CTCCACTTGCTCGCCATCGC	
Probe HEX NRT1.1b-WT	CTCCACTTGCTCGCCGTCG	
Probe HEX NRT1.1b control	ACCGTTAGAACCACCAAGGTAATAA	
dp-F1 GAI	CACCGGCAAGGGTAAGATGTGG	169
dp-R1 GAI	AGAGAGGCCATCCTCCTGAGC	
Probe FAM GAI-WT	CTCCGACATGGCTGAGGTCGCTCAGAAG	
Probe HEX control	GAGCTTCTCGCTGTTTTGGGCTACAACG	
dp-F2 GAI	TTGGGGAGTCCACCGAGT	221
dp-R2 GAI	GACGAGCTAGACCTTCGGCG	
Probe HEX False-dropOFF	GCGGGAGCCATGCGGAAGGTAG	
Exp_F	ATGGGCAGAAGTCAAGATG	Amorces de ddPCR Evagreen sur Exp contrôle, 155pb
Exp_R	TTTGCACTTGGTCTCAGAGG	
ddPCR BT qF3	CTTCGTCCGCCTCCTCTA	Amorces de ddPCR Evagreen taille si épissage : 60 / 1,8kb sinon
ddPCR BT qR3	CGCTCTTCTCGTCAATCTCC	
ddPCR NRT1.1 qF3	CAAGCAGAGGATGCCACAC	Amorces de ddPCR Evagreen taille si épissage : 108 / 2kb sinon
ddPCR NRT1.1 qR3	CACAGCTCCACTTGCTC	

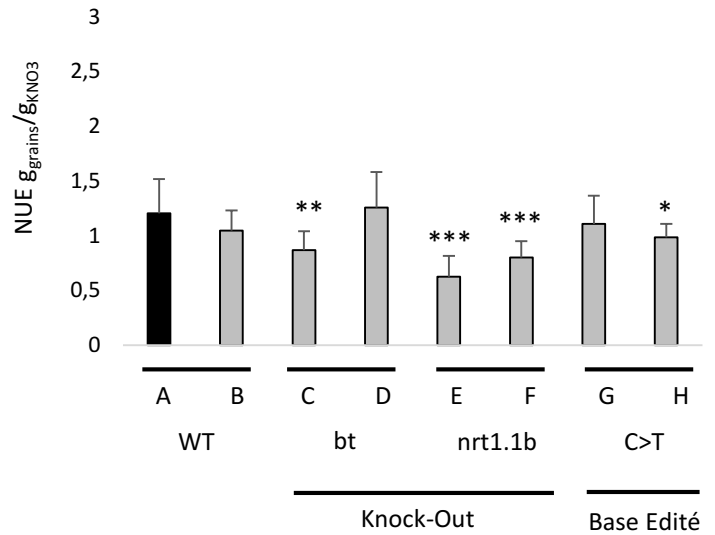
Annexe 29: Primers ddPCR

Annexe 30 : Mesures de la NUE, du HI, et du Rendement chez les lignées Kitaaké réalisées en serre en Août

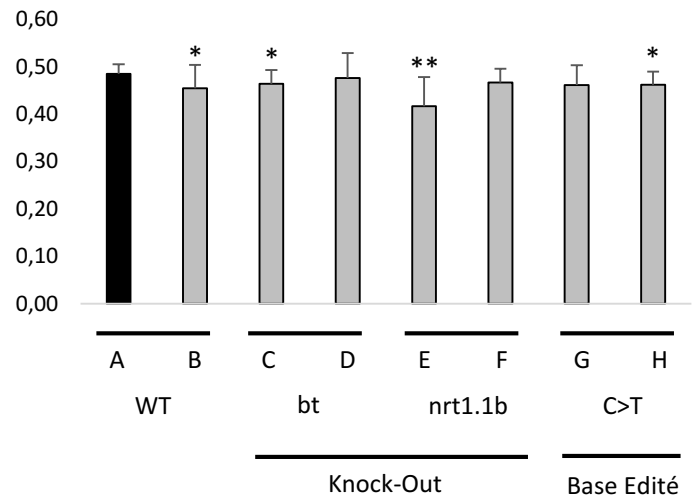
Les données de HI, Rendement, et NUE sont mesurées seulement en HN en Août.

Un T test est réalisé sur n=15 individus par génotype (* < 0,05 ; ** < 0,01 ; *** < 0,0005)

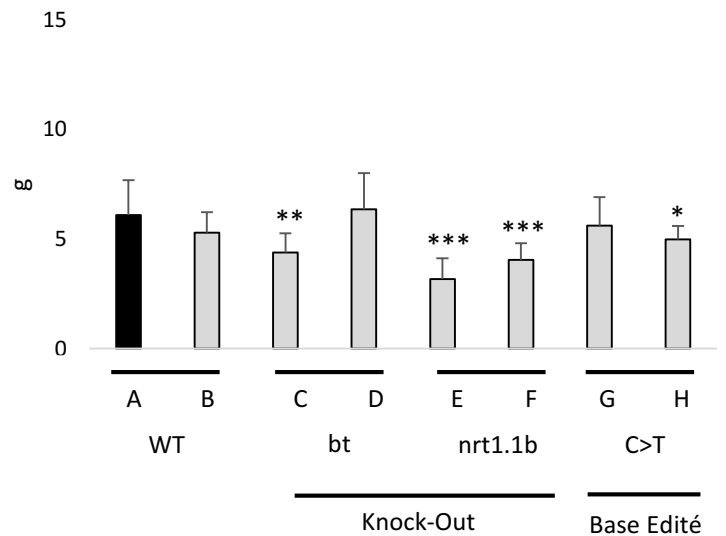
NUE HN aout



HI HN Aout



Rendement en graines HN Aout



Bibliographie

- Adam-Blondon, A.F.; Jaillon, O.; Vezzulli, S.; Zharkikh, A.; Troglio, M.; Velasco, R. (2011). Genome sequence initiatives. In: Genetics, genomics and breeding of grapes, editor(s) Adam-Blondon, A.F.; Martinez-Zapater, J.M.; Kole, C.). CRC Press.
- Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. *Nat Commun* 9, 1911.
- Akcakaya, P., Bobbin, M.L., Guo, J.A., Malagon-Lopez, J., Clement, K., Garcia, S.P., Fellows, M.D., Porritt, M.J., Firth, M.A., Carreras, A., et al. (2018). In vivo CRISPR editing with no detectable genome-wide off-target mutations. *Nature* 561, 416–419.
- Andersen, M.M., Landes, X., Xiang, W., Anyshchenko, A., Falhof, J., Østerberg, J.T., Olsen, L.I., Edenbrandt, A.K., Vedel, S.E., Thorsen, B.J., et al. (2015). Feasibility of new breeding techniques for organic farming. *Trends in Plant Science* 20, 426–434.
- Andersson, M., Turesson, H., Olsson, N., Fält, A.-S., Ohlsson, P., Gonzalez, M.N., Samuelsson, M., and Hofvander, P. (2018). Genome editing in potato via CRISPR-Cas9 ribonucleoprotein delivery. *Physiol Plantarum* 164, 378–384.
- Araus, V., Vidal, E.A., Puelma, T., Alamos, S., Mieulet, D., Guiderdoni, E., and Gutiérrez, R.A. (2016). Members of BTB gene family regulate negatively nitrate uptake and nitrogen use efficiency in *Arabidopsis thaliana* and *Oryza sativa*. *Plant Physiology* 171, 1523–1532.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* 315, 1709–1712.
- Bellegarde, F., Herbert, L., Séré, D., Caillieux, E., Boucherez, J., Fizames, C., Roudier, F., Gojon, A., and Martin, A. (2018). Polycomb Repressive Complex 2 attenuates the very high expression of the *Arabidopsis* gene NRT2.1. *Scientific Reports* 8.
- Berman, J.R., Cooper, S., Zhang, B., Karlin-Neumann, G., Litterst, C., Jouvenot, Y., Hefner, E., Miyaoka, Y., and Conklin, B.R. (2015). Ultra-Sensitive Quantification of Genome Editing Events Using Droplet Digital™ PCR. *Bulletin Biorad*, 6712.
- Bès, M., Herbert, L., Mounier, T., Meunier, A.-C., Durandet, F., Guiderdoni, E., and Périn, C. (2021). Efficient genome editing in rice protoplasts using CRISPR/CAS9 construct. In “Rice Genetic Engineering, Molecular Biology and Genome Editing” Protocol Series, 173-191.
- Billon, P., Bryant, E.E., Joseph, S.A., Nambiar, T.S., Hayward, S.B., Rothstein, R., and Ciccio, A. (2017). CRISPR-Mediated Base Editing Enables Efficient Disruption of Eukaryotic Genes through Induction of STOP Codons. *Molecular Cell* 67, 1068-1079.e4.
- BioRad Droplet Digital™ PCR Applications Guide.
- Blevins, D.G., Hiatt, A.J., and Lowe, R.H. (1974). The Influence of Nitrate and Chloride Uptake on Expressed Sap pH, Organic Acid Synthesis, and Potassium Accumulation in Higher Plants. *Plant Physiol.* 54, 82–87.
- Boch, J., Bonas, U., and Lahaye, T. (2014). TAL effectors - pathogen strategies and plant resistance engineering. *New Phytol* 204, 823–832.

- Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* *151*, 2551–2561.
- Bortesi, L., and Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances* *33*, 41–52.
- Boss, P.K., and Thomas, M.R. (2002). Association of dwarfism and floral induction with a grape ‘green revolution’ mutation. *Nature* *416*, 847–850.
- Bouquet, A., Torregrosa, L., Iocco, P., and Thomas, M.R. (2006). Grapevine (*Vitis vinifera* L.). *Methods in Molecular Biology*, Vol. 344: *Agrobacterium Protocols* 2, 273–285.
- Bowers, J., Boursiquot, J.M., This, P., Chu, K., Johansson, H., Meredith, C. (1999). Historical Genetics: The Parentage of Chardonnay, Gamay, and Other Wine Grapes of Northeastern France. *Science* *285*, 1562–1565.
- Bowers, J.E., and Meredith, C.P. (1997). The parentage of a classic wine grape, Cabernet Sauvignon. *Nat Genet* *16*, 84–87.
- Buck, S.D., Wilde, C.D., Van Montagu, M.V., and Depicker, A. (2000). T-DNA vector backbone sequences are frequently integrated into the genome of transgenic plants obtained by *Agrobacterium*-mediated transformation. *Molecular Breeding* *6*, 459–468.
- Canaguier, A., Grimplet, J., Di Gaspero, G., Scalabrin, S., Duchêne, E., Choisne, N., Mohellibi, N., Guichard, C., Rombauts, S., Le Clainche, I., et al. (2017). A new version of the grapevine reference genome assembly (12X.v2) and of its annotation (VCost.v3). *Genomics Data* *14*, 56–62.
- Cantu, D., and Walker, M.A. (2019). *The Grape Genome* (Cham: Springer International Publishing), 385p.
- Chaïb, J., Torregrosa, L., Mackenzie, D., Corena, P., Bouquet, A., and Thomas, M.R. (2010). The grape microvine - a model system for rapid forward and reverse genetics of grapevines: Grape microvines. *The Plant Journal* *62*, 1083-92
- Chatbanyong, R., and Torregrosa, L (2015) . A highly efficient embryo rescue protocol to recover a progeny from the microvine. *Vitis* *54*, 41-46.
- Chen, Z.C., and Ma, J.F. (2015). Improving Nitrogen Use Efficiency in Rice through Enhancing Root Nitrate Uptake Mediated by a Nitrate Transporter, NRT1.1B. *Journal of Genetics and Genomics* *42*, 463–465.
- Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S., et al. (2013). Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell* *155*, 1479–1491.
- Choi, J.Y., and Purugganan, M.D. (2018). Multiple Origin but Single Domestication Led to *Oryza sativa*. *G3* *8*, 797–803.

- Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*: Floral dip transformation of *Arabidopsis*. *The Plant Journal* *16*, 735–743.
- Collier, R., Dasgupta, K., Xing, Y.-P., Hernandez, B.T., Shao, M., Rohozinski, D., Kovak, E., Lin, J., de Oliveira, M.L.P., Stover, E., et al. (2017). Accurate measurement of transgene copy number in crop plants using droplet digital PCR. *The Plant Journal* *90*, 1014–1025.
- Collonnier, C., Guyon-Debast, A., Maclot, F., Mara, K., Charlot, F., and Nogu , F. (2017). Towards mastering CRISPR-induced gene knock-in in plants: Survey of key features and focus on the model *Physcomitrella patens*. *Methods* *121–122*, 103–117.
- Dalla Costa, Loredana Moffa, Stefano Piazza, and Mickael Malnoy (2018). “Clean” genome editing in grapevine (*Vitis* spp.). In: XII International Conference on grapevine breeding and genetics, Bordeaux, France, July 15-20, 2018: 213
- Dalla Costa, L., Malnoy, M., Lecourieux, D., Deluc, L., Ouaked- Lecourieux, F., Thomas, M.R., and Torregrosa, L.J.-M. (2019). The state-of-the-art of grapevine biotechnology and new breeding technologies (NBTS). *OENO One* *53*, 189–212.
- Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* *471*, 602–607.
- Demol Julien, Baudoin Jean Pierre, Louant Benoit Pascal, Mar chal R., Mergeai Guy, and Otoul Emile (2002). Am lioration des plantes: application aux principales esp ces cultiv es en r gions tropicales. (Les presses agronomiques de Gembloux), 585p.
- Deng, L., Wang, H., Sun, C., Li, Q., Jiang, H., Du, M., Li, C.-B., and Li, C. (2018). Efficient generation of pink-fruited tomatoes using CRISPR/Cas9 system. *Journal of Genetics and Genomics* *45*, 51–54.
- Deveau, H., Barrangou, R., Garneau, J.E., Labonte, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., and Moineau, S. (2008). Phage Response to CRISPR-Encoded Resistance in *Streptococcus thermophilus*. *Journal of Bacteriology* *190*, 1390–1400.
- Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* *34*, 184–191.
- Dor  and Varoquaux (2006). Histoire et am lioration de cinquante plantes cultiv es. (Institut national de la recherche agronomique, Paris). QUAE  ditions, pp. 840 (Institut national de la recherche agronomique, Paris).
- Doudna, J. (2017). CRISPR System: From Adaptive Immunity to Genome Editing. *Molecular Frontiers Journal* *01*, 76–91.
- Fausser, F., Schiml, S., and Puchta, H. (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *The Plant Journal* *79*, 348–359.

- Fayos, I., Meunier, A.C., Vernet, A., Sanz, S.N., Portefaix, M., Lartaud, M., Bastianelli, G., Périn, C., Nicolas, A., and Guiderdoni, E. (2020). Assessment of the roles of OsSPO11-2 and OsSPO11-4 in rice meiosis using CRISPR/Cas9 mutagenesis. *Journal of Experimental Botany* *71*, 7046-7058.
- Findlay, S.D., Vincent, K.M., Berman, J.R., and Postovit, L.-M. (2016). A Digital PCR-Based Method for Efficient and Highly Specific Screening of Genome Edited Cells. *PLOS ONE* *11*, e0153901.
- Galet (1988). *Cépages et vignobles de France*. Eds Galet P. (Montpellier), 553p.
- Garcia, J., Dusserre, E., Cheynet, V., Bringuier, P.P., Brengle-Pesce, K., Wozny, A.-S., Rodriguez-Lafrasse, C., Freyer, G., Brevet, M., Payen, L., et al. (2017). Evaluation of pre-analytical conditions and comparison of the performance of several digital PCR assays for the detection of major *EGFR* mutations in circulating DNA from non-small cell lung cancers: the CIRCAN_0 study. *Oncotarget* *8*, 87980–87996.
- Garneau, J.E., Dupuis, M.-È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* *468*, 67–71.
- Grantham, R., Gautier, C., and Gouy, M. (1980). Codon frequencies in 119 individual genes confirm consistent choices of degenerate bases according to genome type. *Nucl Acids Res* *8*, 1893–1912.
- Gribaudo, I., Gambino, G., Boccacci, P., Perrone, I., and Cuozzo, D. (2017). A multi-year study on the regenerative potential of several *Vitis* genotypes. *Acta Hort.* *1155*, 45–50.
- Harkess, A. (2019). Smashing Barriers in Biolistic Plant Transformation. *Plant Cell* *31*, 273–274.
- Herbert, Meunier, Périn, Bes, Pestel, Vernet, Chaîne, Portefaix, Michel, Meynard, et al. (in prep). Detection of CRISPR/ CAS9 mutations by Droplet Digital PCR: catch them all !
- Herbert, L., Meunier, A.-C., Bes, M., Vernet, A., Portefaix, M., Durandet, F., Michel, R., Chaîne, C., This, P., Guiderdoni, E., et al. (2020). Beyond Seek and Destroy: how to Generate Allelic Series Using Genome Editing Tools. *Rice* *13*, 5.
- Hess, G.T., Tycko, J., Yao, D., and Bassik, M.C. (2017). Methods and Applications of CRISPR-Mediated Base Editing in Eukaryotic Genomes. *Molecular Cell* *68*, 26–43.
- Hichri, I., Heppel, S.C., Pillet, J., Léon, C., Czemmél, S., Delrot, S., Lauvergeat, V., and Bogs, J. (2010). The Basic Helix-Loop-Helix Transcription Factor MYC1 Is Involved in the Regulation of the Flavonoid Biosynthesis Pathway in Grapevine. *Molecular Plant* *3*, 509–523.
- Hiei, Y., and Komari, T. (2006). Improved protocols for transformation of indica rice mediated by *Agrobacterium tumefaciens*. *Plant Cell Tiss Organ Cult* *85*, 271–283.
- Hirota Kumagai, and Hiroshi Kouchi (2003). Gene Silencing by Expression of Hairpin RNA in *Lotus japonicus* Roots and Root Nodules. *MPMI* Vol. 16, No. 8, Publication No. M-2003-0512-01R 663–668.

- Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* 157, 1262–1278.
- Hu, B., Wang, W., Ou, S., Tang, J., Li, H., Che, R., Zhang, Z., Chai, X., Wang, H., Wang, Y., et al. (2015). Variation in NRT1.1B contributes to nitrate-use divergence between rice subspecies. *Nature Genetics* 47, 834–838.
- Hua, K., Tao, X., and Zhu, J.-K. (2019). Expanding the base editing scope in rice by using Cas9 variants. *Plant Biotechnology Journal* 17, 499–504.
- Huglin and Schneider (1998). *Biologie et écologie de la vigne*. Tech et doc Lavoisier, p 370
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology* 169, 5429–5433.
- Ito, Y., Nishizawa-Yokoi, A., Endo, M., Mikami, M., and Toki, S. (2015). CRISPR/Cas9-mediated mutagenesis of the RIN locus that regulates tomato fruit ripening. *Biochemical and Biophysical Research Communications* 467, 76–82.
- Izawa, T., and Shimamoto, K. (1996). Becoming a model plant: The importance of rice to plant science. *Trends in Plant Science* 1, 95–99.
- Jacob, P., Avni, A., and Bendahmane, A. (2018). Translational Research: Exploring and Creating Genetic Diversity. *Trends in Plant Science* 23, 42–52.
- Jaillon, Aury, Noel, Policriti, and Clepet (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449, 463–467.
- Jakab, Dubresson, Bel, Hamiduzzaman, Mauch-Mani, and Neuhaus (2006). Reverse genetic methods in research on induced resistance of grapevine: development of a vector for shRNA production to induce gene silencing. *Induced Resistance in Plants against Insects and Diseases IOBC/Wprs Bull.* 29(8) 55–60.
- Jansen, Ruud., Embden, Jan.D.A. van, Gaastra, Wim., and Schouls, Leo.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 43, 1565–1575.
- Jin, S., Zong, Y., Gao, Q., Zhu, Z., Wang, Y., Qin, P., Liang, C., Wang, D., Qiu, J.-L., Zhang, F., et al. (2019). Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. 5.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337, 816–821.
- Joshi, B.K., and Gauchan, D. (2017). Released and promising crop varieties for mountain agriculture in Nepal (1959-2016) (Pokhra, Nepal: Biodiversity International).

- Kaeppler, S.M., Kaeppler, H.F., and Rhee, Y. (2000). Epigenetic aspects of somaclonal variation in plants. In *Plant Gene Silencing*, M.A. Matzke, and A.J.M. Matzke, eds. (Dordrecht: Springer Netherlands), pp. 59–68.
- Keller M (2010) *The science of grapevines: anatomy and physiology*, 1st edition. Elsevier Academic Press, Burlington
- Kiba, T., and Krapp, A. (2016). Plant Nitrogen Acquisition Under Low Availability: Regulation of Uptake and Root Architecture. *Plant and Cell Physiology* 57, 707–714.
- Kim, Y.-G., Shi, Y., Berg, J.M., and Chandrasegaran, S. (1997). Site-specific cleavage of DNA–RNA hybrids by zinc finger/FokI cleavage domain fusions. *Gene* 203, 43–49.
- Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424.
- Kong, J., Martin-Ortigosa, S., Finer, J., Orchard, N., Gunadi, A., Batts, L.A., Thakare, D., Rush, B., Schmitz, O., Stuiver, M., et al. (2020). Overexpression of the Transcription Factor GROWTH-REGULATING FACTOR5 Improves Transformation of Dicot and Monocot Species. *Front. Plant Sci.* 11, 572319.
- Kulcsár, P.I., Tálas, A., Huszár, K., Ligeti, Z., Tóth, E., Weinhardt, N., Fodor, E., and Welker, E. (2017). Crossing enhanced and high fidelity SpCas9 nucleases to optimize specificity and cleavage. *Genome Biol* 18, 190.
- Kumagai, H., and Kouchi, H. (2003). Gene Silencing by Expression of Hairpin RNA in *Lotus japonicus* Roots and Root Nodules. *MPMI* 16, 663–668.
- Lacchini, E., Kiegle, E., Castellani, M., Adam, H., Jouannic, S., Gregis, V., and Kater, M.M. (2020). CRISPR-mediated accelerated domestication of African rice landraces. *PLoS ONE* 15, e0229782.
- Lacombe (2012). Contribution à l'étude de l'histoire évolutive de la vigne cultivée (*Vitis vinifera* L.) par l'analyse de la diversité génétique neutre et de gènes d'intérêt, 328p.
- Le Gall, O., Torregrosa, L., Danglot, Y., Candresse, T., and Bouquet, A. (1994). Agrobacterium-mediated genetic transformation of grapevine somatic embryos and regeneration of transgenic plants expressing the coat protein of grapevine chrome mosaic nepovirus (GCMV). *Plant Science* 102, 161–170.
- Lemmon, Z.H., Reem, N.T., Dalrymple, J., Soyk, S., Swartwood, K.E., Rodriguez-Leal, D., Van Eck, J., and Lippman, Z.B. (2018). Rapid improvement of domestication traits in an orphan crop by genome editing. *Nature Plants* 4, 766–770.
- Li, J., Zhang, X., Sun, Y., Zhang, J., Du, W., Guo, X., Li, S., Zhao, Y., and Xia, L. (2018). Efficient allelic replacement in rice by gene editing: A case study of the *NRT1.1B* gene: Targeted elite allele replacement by CRISPR/Cas9 in rice. *Journal of Integrative Plant Biology*, 60, 536-540.

Li, J., Hong, S., Chen, W., Zuo, E., and Yang, H. (2019). Advances in detecting and reducing off-target effects generated by CRISPR-mediated genome editing. *Journal of Genetics and Genomics* 46, 513–521.

Li, J.-F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M., and Sheen, J. (2013). Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat Biotechnol* 31, 688–691.

Li, J.-Y., Wang, J., and Zeigler, R.S. (2014). The 3,000 rice genomes project: new opportunities and challenges for future rice research. *GigaSci* 3, 8.

Li, M., Li, X., Zhou, Z., Wu, P., Fang, M., Pan, X., Lin, Q., Luo, W., Wu, G., and Li, H. (2016). Reassessment of the Four Yield-related Genes *Gn1a*, *DEP1*, *GS3*, and *IPA1* in Rice Using a CRISPR/Cas9 System. *Front. Plant Sci.* 7, 377.

Li, Q., Sapkota, M., and van der Knaap, E. (2020a). Perspectives of CRISPR/Cas-mediated cis-engineering in horticulture: unlocking the neglected potential for crop improvement. *Hortic Res* 7, 36.

Li, Y., Zhu, J., Wu, H., Liu, C., Huang, C., Lan, J., Zhao, Y., and Xie, C. (2020b). Precise base editing of non-allelic acetolactate synthase genes confers sulfonylurea herbicide resistance in maize. *The Crop Journal* 8, 449–456.

Liu, C., Zhang, L., Liu, H., and Cheng, K. (2017). Delivery strategies of the CRISPR-Cas9 gene-editing system for therapeutic applications. *Journal of Controlled Release* 266, 17–26.

Lu, Y., and Zhu, J.-K. (2017). Precise Editing of a Target Base in the Rice Genome Using a Modified CRISPR/Cas9 System. *Molecular Plant* 10, 523–525.

Malnoy, M., Viola, R., Jung, M.-H., Koo, O.-J., Kim, S., Kim, J.-S., Velasco, R., and Nagamangala Kanchiswamy, C. (2016). DNA-Free Genetically Edited Grapevine and Apple Protoplast Using CRISPR/Cas9 Ribonucleoproteins. *Frontiers in Plant Science* 7, 1904.

Mandel, M., and Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology* 53, 159–162.

McGinn, M., Phippen, W.B., Chopra, R., Bansal, S., Jarvis, B.A., Phippen, M.E., Dorn, K.M., Esfahanian, M., Nazareus, T.J., Cahoon, E.B., et al. (2019). Molecular tools enabling pennycress (*Thlaspi arvense*) as a model plant and oilseed cash cover crop. *Plant Biotechnol J* 17, 776–788.

Meynard, Aurore Vernet, Meunier, A., Delphine Mieulet, Martine Bès, Murielle Portefaix, Jean-Christophe Breitler, Christophe Périn, and Emmanuel Guiderdoni (2020). Thirty Years of Genome Engineering in Rice: From Gene Addition to Gene Editing. In *Annual Plant Reviews Online* (2020), pp. 1–76.

Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., Zhang, X., Wan, J., Gu, H., and Qu, L.-J. (2013). Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Research* 23, 1233–1236.

- Mieulet, D., Aubert, G., Bres, C., Klein, A., Droc, G., Vieille, E., Rond-Coissieux, C., Sanchez, M., Dalmais, M., Mauxion, J.-P., et al. (2018). Unleashing meiotic crossovers in crops. *Nature Plants* 4, 1010–1016.
- Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *J Mol Evol* 60, 174–182.
- Mullins, M.G., Tang, F.C.A., and Facciotti, D. (1990). Agrobacterium-Mediated Genetic Transformation of Grapevines: Transgenic Plants of *Vitis rupestris* SCHEELE and Buds of *Vitis vinifera* L. *Nat Biotechnol* 8, 1041–1045.
- Muthayya, S., Sugimoto, J.D., Montgomery, S., and Maberly, G.F. (2014). An overview of global rice production, supply, trade, and consumption: Global rice production, consumption, and trade. *Ann. N.Y. Acad. Sci.* 1324, 7–14.
- Nacry, P., Bouguyon, E., and Gojon, A. (2013). Nitrogen acquisition by roots: physiological and developmental mechanisms ensuring plant adaptation to a fluctuating resource. *Plant and Soil* 370, 1–29.
- Nakajima, I., Ban, Y., Azuma, A., Onoue, N., Moriguchi, T., Yamamoto, T., Toki, S., and Endo, M. (2017). CRISPR/Cas9-mediated targeted mutagenesis in grape. *PLOS ONE* 12, e0177966.
- Nakamura, M., Gao, Y., Dominguez, A.A., and Qi, L.S. (2021). CRISPR technologies for precise epigenome editing. *Nat Cell Biol* 23, 11–22.
- Nekrasov, V., Wang, C., Win, J., Lanz, C., Weigel, D., and Kamoun, S. (2017). Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. *Sci Rep* 7, 482.
- Nishi, T., Yamada, Y., and Takahashi, E. (1968). Organ Redifferentiation and Plant Restoration in Rice Callus. *Nature* 219, 508–509.
- Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* 156, 935–949.
- Østerberg, J.T., Xiang, W., Olsen, L.I., Edenbrandt, A.K., Vedel, S.E., Christiansen, A., Landes, X., Andersen, M.M., Pagh, P., Sandøe, P., et al. (2017). Accelerating the Domestication of New Crops: Feasibility and Approaches. *Trends in Plant Science* 22, 373–384.
- Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A., and Liu, D.R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature Biotechnology* 31, 839–843.
- Péros, J.-P., Berger, G., Portemont, A., Boursiquot, J.-M., and Lacombe, T. (2011). Genetic variation and biogeography of the disjunct *Vitis* subg. *Vitis* (Vitaceae) *Journal of Biogeography* 38, 471–486.
- Piatek, A., Ali, Z., Baazim, H., Li, L., Abulfaraj, A., Al-Shareef, S., Aouida, M., and Mahfouz, M.M. (2015). RNA-guided transcriptional regulation *in planta* via synthetic dCas9-based transcription factors. *Plant Biotechnol J* 13, 578–589.

- Plett, D., Toubia, J., Garnett, T., Tester, M., Kaiser, B.N., and Baumann, U. (2010). Dichotomy in the NRT Gene Families of Dicots and Grass Species. *PLoS ONE* 5, e15289.
- Pourkheirandish, M., Golicz, A.A., Bhalla, P.L., and Singh, M.B. (2020). Global Role of Crop Genomics in the Face of Climate Change. *Front. Plant Sci.* 11, 922.
- Puchta, H. (2004). The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *Journal of Experimental Botany*, Vol 56, No. 409, pp. 1–14.
- Puchta, H., Dujon, B., and Hohn, B. (1993). Homologous recombination in plant cells is enhanced by *in vivo* induction of double strand breaks into DNA by a site-specific endonuclease. *Nucl Acids Res* 21, 5034–5040.
- Raboin, L.M., Ramanantsoanirina, A., Dzido, J.L., Andriantsimialona, D., Tharreau, D., Radanielina, T., and Ahmadi, N. (2010). Upland (aerobic) rice breeding for the harsh environment of the High Plateau of Madagascar. *Africa Rice Congress 2010*.
- Ran, F.A., Hsu, P.D., Lin, C.-Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., et al. (2013). Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell* 154, 1380–1389.
- Ren, B., Yan, F., Kuang, Y., Li, N., Zhang, D., Zhou, X., Lin, H., and Zhou, H. (2018). Improved Base Editor for Efficiently Inducing Genetic Variations in Rice with CRISPR/Cas9-Guided Hyperactive hAID Mutant. *Molecular Plant* 11, 623–626.
- Ren, C., Liu, X., Zhang, Z., Wang, Y., Duan, W., Li, S., and Liang, Z. (2016). CRISPR/Cas9-mediated efficient targeted mutagenesis in Chardonnay (*Vitis vinifera* L.). *Scientific Reports* 6.
- Ren, R., Gao, J., Lu, C., Wei, Y., Jin, J., Wong, S.-M., Zhu, G., and Yang, F. (2020). Highly Efficient Protoplast Isolation and Transient Expression System for Functional Characterization of Flowering Related Genes in Cymbidium Orchids. *IJMS* 21, 2264.
- Reustle, G., Harst, M., and Alleweldt, G. (1995). Plant regeneration of grapevine (*Vitis* sp.) protoplasts isolated from embryogenic tissue. *Plant Cell Reports* 15, 238–241.
- Rodríguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E., and Lippman, Z.B. (2017). Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. *Cell* 171, 470-480.e8.
- Ruffel, S., Krouk, G., Ristova, D., Shasha, D., Birnbaum, K.D., and Coruzzi, G.M. (2011). Nitrogen economics of root foraging: Transitive closure of the nitrate-cytokinin relay and distinct systemic signaling for N supply vs. demand. *Proceedings of the National Academy of Sciences* 108, 18524–18529.
- Saika, H., Mori, A., Endo, M., and Toki, S. (2019). Targeted deletion of rice retrotransposon Tos17 via CRISPR/Cas9. *Plant Cell Rep* 38, 455–458.
- Sallaud, C., Meynard, D., van Boxtel, J., Gay, C., Bès, M., Brizard, J.P., Larmande, P., Ortega, D., Raynal, M., Portefaix, M., et al. (2003). Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. *Theor Appl Genet* 106, 1396–1408.

- Sallaud, C., Gay, C., Larmande, P., Bès, M., Piffanelli, P., Piégu, B., Droc, G., Regad, F., Bourgeois, E., Meynard, D., et al. (2004). High throughput T-DNA insertion mutagenesis in rice: a first step towards *in silico* reverse genetics. *The Plant Journal* *39*, 450–464.
- Sánchez-León, S., Gil-Humanes, J., Ozuna, C.V., Giménez, M.J., Sousa, C., Voytas, D.F., and Barro, F. (2018). Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. *Plant Biotechnol J* *16*, 902–910.
- Sathishkumar, R., Kumar, S.R., Hema, J., and Baskar, V. (2019). Advances in Plant Transgenics: Methods and Applications, chapter "Application of genome engineering/editing tools in plants", Mohan, Shibao, and Silva, 143-165, and chapter "Genome Modification Approaches to Improve Performance, Quality, and Stress Tolerance of Important Mediterranean Fruit Species (*Olea europaea* L., *Vitis vinifera* L., and *Quercus suber* L.)", Cardoso, Figueiredo, Serrazina, Pires, Peixe, 273-312, (Singapore: Springer Singapore).
- Schindele, A., Dorn, A., and Puchta, H. (2020). CRISPR/Cas brings plant biology and breeding into the fast lane. *Current Opinion in Biotechnology* *61*, 7–14.
- Shimamoto, K., Terada, R., Izawa, T., and Fujimoto, H. (1989). Fertile transgenic rice plants regenerated from transformed protoplasts. *Nature* *338*, 274–276.
- Shukla, V.K., Doyon, Y., Miller, J.C., DeKolver, R.C., Moehle, E.A., Worden, S.E., Mitchell, J.C., Arnold, N.L., Gopalan, S., Meng, X., et al. (2009). Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature* *459*, 437–441.
- Smith, C.J., Castanon, O., Said, K., Volf, V., Khoshakhlagh, P., Hornick, A., Ferreira, R., Wu, C.-T., Güell, M., Garg, S., et al. (2019). Enabling large-scale genome editing by reducing DNA nicking (Synthetic Biology). [Nucleic Acids Res.](#) *48*, 5183–5195.
- Sprink, T., Eriksson, D., Schiemann, J., and Hartung, F. (2016). Regulatory hurdles for genome editing: process- vs. product-based approaches in different regulatory contexts. *Plant Cell Rep* *35*, 1493–1506.
- Srinivasan, C., and Mullins, M.G. (1980). High-frequency somatic embryo production from unfertilized ovules of grapes. *Scientia Horticulturae* *13*, 245–252.
- Steele, K.A., Gyawali, S., Joshi, K.D., Shrestha, P., Sthapit, B.R., and Witcombe, J.R. (2009). Has the introduction of modern rice varieties changed rice genetic diversity in a high-altitude region of Nepal? *Field Crops Research* *113*, 24–30.
- Sun, Y., Jiao, G., Liu, Z., Zhang, X., Li, J., Guo, X., Du, W., Du, J., Francis, F., Zhao, Y., et al. (2017). Generation of High-Amylose Rice through CRISPR/Cas9-Mediated Targeted Mutagenesis of Starch Branching Enzymes. *Front. Plant Sci.* *8*.
- Svitashev, S., Schwartz, C., Lenderts, B., Young, J.K., and Mark Cigan, A. (2016). Genome editing in maize directed by CRISPR–Cas9 ribonucleoprotein complexes. *Nature Communications* *7*, 13274.
- Tsai, S.Q., Nguyen, N.T., Malagon-Lopez, J., Topkar, V.V., Aryee, M.J., and Joung, J.K. (2017). CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR–Cas9 nuclease off-targets. *Nat Methods* *14*, 607–614.

- Tyagi, A.K., and Mohanty, A. (2000). Rice transformation for crop improvement and functional genomics. *Plant Science* 158, 1–18.
- Velasco, R., Zharkikh, A., Troglio, M., Cartwright, D.A., Cestaro, A., Pruss, D., Pindo, M., FitzGerald, L.M., Vezzulli, S., Reid, J., et al. (2007). A High Quality Draft Consensus Sequence of the Genome of a Heterozygous Grapevine Variety. *PLoS ONE* 2(12): e1326.
- Wang, H., Wang, W., Zhan, J., Huang, W., and Xu, H. (2015). An efficient PEG-mediated transient gene expression system in grape protoplasts and its application in subcellular localization studies of flavonoids biosynthesis enzymes. *Scientia Horticulturae* 191, 82–89.
- Wang, W., Akhunova, A., Chao, S., and Akhunov, E. (2016a). Optimizing multiplex CRISPR/Cas9-based genome editing for wheat (Genomics). *bioRxiv*, 10.1101/051342.
- Wang, W., Hu, B., Li, A., and Chu, C. (2020). NRT1.1s in plants: functions beyond nitrate transport. *Journal of Experimental Botany* 71, 4373–4379.
- Wang, X., Tu, M., Wang, D., Liu, J., Li, Y., Li, Z., Wang, Y., and Wang, X. (2018). CRISPR/Cas9-mediated efficient targeted mutagenesis in grape in the first generation. *Plant Biotechnology Journal* 16, 844–855.
- Wang, Y., Liu, X., Ren, C., Zhong, G.-Y., Yang, L., Li, S., and Liang, Z. (2016b). Identification of genomic sites for CRISPR/Cas9-based genome editing in the *Vitis vinifera* genome. *BMC Plant Biol* 16, 96.
- Wiley (2018). Annual Plant Reviews online.
- Witte, A.K., Mester, P., Fister, S., Witte, M., Schoder, D., and Rossmann, P. (2016). A Systematic Investigation of Parameters Influencing Droplet Rain in the *Listeria monocytogenes* prfA Assay - Reduction of Ambiguous Results in ddPCR. *PLOS ONE* 11, e0168179.
- Woo, J.W., Kim, J., Kwon, S.I., Corvalán, C., Cho, S.W., Kim, H., Kim, S.-G., Kim, S.-T., Choe, S., and Kim, J.-S. (2015). DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nature Biotechnology* 33, 1162–1164.
- Xu, R., Yang, Y., Qin, R., Li, H., Qiu, C., Li, L., Wei, P., and Yang, J. (2016). Rapid improvement of grain weight via highly efficient CRISPR/Cas9-mediated multiplex genome editing in rice. *Journal of Genetics and Genomics* 43, 529–532.
- Xuan, W., Beeckman, T., and Xu, G. (2017). Plant nitrogen nutrition: sensing and signaling. *Current Opinion in Plant Biology* 39, 57–65.
- Yamamoto, T., Iketani, H., Ieki, H., Nishizawa, Y., Notsuka, K., Hibi, T., Hayashi, T., and Matsuta, N. (2000). Transgenic grapevine plants expressing a rice chitinase with enhanced resistance to fungal pathogens. *Plant Cell Reports* 19, 639–646.
- Yan, F., Kuang, Y., Ren, B., Wang, J., Zhang, D., Lin, H., Yang, B., Zhou, X., and Zhou, H. (2018). Highly Efficient A·T to G·C Base Editing by Cas9n-Guided tRNA Adenosine Deaminase in Rice. *Molecular Plant* 11, 631–634.
- Yoo, J.H., Seong, E.S., Ghimire, B.K., Heo, K., Jin, X., Yamada, T., Clark, L.V., Sacks, E.J., and Yu, C.Y. (2018). Establishment of *Miscanthus sinensis* with decreased lignin biosynthesis

by *Agrobacterium*–mediated transformation using antisense COMT gene. *Plant Cell Tiss Organ Cult* 133, 359–369.

Zambryski, P., Joos, H., Genetello, C., Leemans, J., Van Montagu, M., and Schell, J. (1983). Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. *The EMBO Journal* 2, 2143–2150.

Zhang, Y., and Qi, Y. (2019). CRISPR enables directed evolution in plants. *Genome Biology* 20.

Zhang, J., Liu, Y.-X., Zhang, N., Hu, B., Jin, T., Xu, H., Qin, Y., Yan, P., Zhang, X., Guo, X., et al. (2019). NRT1.1B is associated with root microbiota composition and nitrogen use in field-grown rice. *Nature Biotechnology* 37, 676–684.

Zhang, Y., Su, J., Duan, S., Ao, Y., Dai, J., Liu, J., Wang, P., Li, Y., Liu, B., Feng, D., et al. (2011). A highly efficient rice green tissue protoplast system for transient gene expression and studying light/chloroplast-related processes. *Plant Methods*, 7, 30

Zhong, G.-Y., and Yang, Y. (2012). Characterization of grape Gibberellin Insensitive1 mutant alleles in transgenic *Arabidopsis*. *Transgenic Research* 21, 725–741.

Zhou, H., Liu, B., Weeks, D.P., Spalding, M.H., and Yang, B. (2014). Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Research* 42, 10903–10914.

Zhou, J., Xin, X., He, Y., Chen, H., Li, Q., Tang, X., Zhong, Z., Deng, K., Zheng, X., Akher, S.A., et al. (2018). Multiplex QTL editing of grain-related genes improves yield in elite rice varieties. *Plant Cell Reports* 38, 475–485

Zong, Y., Wang, Y., Li, C., Zhang, R., Chen, K., Ran, Y., Qiu, J.-L., Wang, D., and Gao, C. (2017). Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nature Biotechnology* 35, 438–440.

Zsögön, A., Čermák, T., Naves, E.R., Notini, M.M., Edel, K.H., Weinl, S., Freschi, L., Voytas, D.F., Kudla, J., and Peres, L.E.P. (2018). De novo domestication of wild tomato using genome editing. *Nat Biotechnol* 36, 1211–1216.

Résumé

L'édition du génome consiste à modifier dans le génome de façon très ciblée une ou plusieurs séquences d'ADN à un locus bien précis à l'aide de nucléases. La technologie CRISPR CAS9 (Clustered Regularly Interspaced Short Palindromic Repeats/Crispr Associated proteins) est basée sur l'induction de cassures doubles brins de l'ADN (Double Strand Break, DSB), ce qui entraîne l'apparition de mutations lors de la réparation de ces cassures. La majorité des mutations entraîne l'inactivation du gène cible par mutations de type insertion-délétion. Cependant d'autres types de mutations, plus spécifiques, à la base près, peuvent être générées par des technologies dites de « Base Editing ». Le développement de ces technologies faciliterait l'analyse de la fonction de l'impact de mutations dans des gènes d'intérêt agronomique chez le riz et la vigne.

Dans le cadre de cette thèse, dans un premier temps, nous avons développé des protocoles rapides de tests en protoplastes de riz et vigne pour faire de l'édition des génomes sur différentes cibles. De même grâce à l'utilisation de la technologie droplet digital PCR (ddPCR) nous avons pu déterminer les taux de mutations de différents construits et sélectionner les plus efficaces. Dans un second temps nous avons appliqué *in-planta* ces technologies pour étudier des caractères d'intérêts agronomiques. Nous avons démontré la possibilité de développer différentes versions alléliques par l'utilisation de plusieurs endonucléases, induisant différents phénotypes. Ensuite, nous avons reproduit un SNP naturel dans *OsNRT1.1b* par la technologie CRISPR « Base Editor », dans une lignée de riz cultivée à Madagascar, qui permettrait une meilleure utilisation de l'azote. Nous l'avons ensuite évaluée en Colombie pour tester son effet, démontrant que la mutation à la base près modifiant la Thréonine 327 en Méthionine (Thr327Met) permet une meilleure utilisation de l'azote au champ. Pour la vigne nous avons ciblé par la technologie CRISPR/CAS9 un gène de nanisme, *VvGAI*, afin d'espérer reproduire le modèle de microvigne. Des transformations stables ont été réalisées avec les construits correspondants, mais nous attendons la régénération de plantes afin de vérifier le phénotype.

Abstract

Genome editing consists in modifying in a very targeted way one or more DNA sequences at a very precise locus by means of nucleases. The CRISPR CAS9 technology (Clustered Regularly Interspaced Short Palindromic Repeats/Crispr Repeats/Crispr Asso proteins) is based on the induction of DNA double strand breaks (Double Strand Break, DSB) which leads to the appearance of mutations when the breaks are repaired. The majority of the mutations inactivate target genes by insertion-deletion mutations. The majority of mutations result in the inactivation of the target gene by insertion-deletion type mutations. However, other types of mutations, more specific, can be generated by so-called "Base Editing" technologies. The development of these technologies would facilitate the analysis of the function and impact of mutations in genes of agronomic interest in rice and grapevine.

In the framework of this thesis, we have first developed rapid protocols to test DNA construction in rice or grape protoplasts to perform genome editing on different targets. Using the new droplet digital PCR technology (ddPCR), we were able to determine the mutation rates of different constructs and select the best ones. Secondly, we applied these technologies *in-planta* to study traits of agronomic interest. We demonstrate the possibility to develop different allelic versions by using several endonucleases, inducing different phenotypes. Then, we reproduced a natural SNP in *OsNRT1.1b* using the CRISPR "Base Editor" technology in a rice variety cultivated in Madagascar. This mutation would theoretically allow a better Nitrogen Use efficiency. Then, we have tested it in Colombia to check its efficiency, demonstrating that single base substitution modifying Threonine 327 to Methionine (Thr327Met) allows better use of nitrogen in the field. For grapevine, using the CRISPR/CAS9 technology, we targeted a dwarfism gene, *VvGAI*, in order to hope to reproduce the microvine model. Stable transformation was performed with the corresponding constructs, but we are waiting for the plant regeneration in order to verify the phenotype.