

République populaire et démocratique d'Algérie
Ministère de l'enseignement supérieur et de la recherche
Université Ziane Achour de Djelfa
Faculté des Sciences de la Nature et de la Vie
Département de Biologie

**Polycopié de la matière de génie génétique incluant
Les cours, les travaux dirigés et les travaux pratiques**

Niveau : Troisième année licence

Spécialité : Biologie moléculaire

M. Fathi Berrabah

Maitre de conférences classe B en biologie

Responsable du module

Année universitaire 2018/2019

Sommaires

Partie I : Cours

Chapitre I : Généralités en biologie moléculaire	1
1. La cellule, unité de base du vivant	1
2. L'ADN base de l'information génétique	3
3. La transcription et la traduction	5
4. Promoteur de gènes et régulation cis-trans	8
5. Les autres niveaux de régulation d'expression de gènes	10
6. La régulation par petits ARNs	11
6. Notions d'espèces modèles	13
7. Introduction au génie génétique	14
Chapitre II : Les outils enzymatique du génie génétique	16
1. Les enzymes de restriction	16
1.1. Le phénomène de restriction	16
1.2. Nomenclature des enzymes de restriction	16
1.2.1. Enzymes de type I	16
1.2.1. Enzymes de type II	17
1.2.2. Enzymes de type III	17
1.3. Les type de coupures induites par les enzymes de restrictions	17
2. Les autres enzymes d'usage courant en biologie moléculaire	17
2.1. Les polymérases	17
2.2. Les ligases	18
2.3. Les exonucléases	18
Chapitre III : L'hybridation moléculaire	20
1. Rappel sur le principe de la réaction d'hybridation	20
1.1. Notion de température de fusion de l'ADN	20
1.2. Facteurs influençant la température de fusion	20
2. L'hybridation en phase liquide	20
2.1. Principe	20
2.2. Applications de l'hybridation moléculaire en phase liquide	20
3. L'hybridation sur support solide	21
3.1. Principe	21
3.2. Les supports utilisés pour immobiliser les acides nucléiques	22

Sommaires

2. L'hybridation <i>in situ</i>	22
Chapitre VI : clonage moléculaire	24
1. Le principe du clonage	24
2. Les banques d'ADN	24
2.1. Les banques d'ADN génomique	24
2.1.1. Etablissement de la banque d'ADN	24
2.2. Les banques d'ADNc	26
2.2.1. Le passage de l'ARN à l'ADN	26
2.2.2. Le choix du vecteur	26
Chapitre V : vecteurs de clonage	28
1. Généralités sur les vecteurs	28
1.1. Concept de vecteur	28
1.2. Propriétés que doit posséder un vecteur	28
1.3. Principes généraux d'utilisation d'un vecteur	28
2. Les plasmides	29
2.1. L'utilisation d'un plasmide	29
2.2. Processus globale de clonage par plasmide	30
2.2. Préparation des plasmides pour le clonage	31
2.2.1. Clonage par restriction ligation	31
2.2.1. Crible de clone	33
2.2.3. Le clonage par recombinaison	35
2.3. Les différents types de plasmides	37
2.3.1 Les plasmide de première génération	37
2.3.2. Les plasmides de seconde génération	37
2.3.3. Les plasmides de troisième génération	37
3. Les phages	37
3.1. Utilisation des phages	37
3.2. Préparation d'un phage	38
3.3. Les différents phages utilisés en biologie moléculaire	38
3.3.1. Les phage de première génération. : Le phage λ	38
3.3.2. Les phages de seconde génération	39
4. Les autres types de vecteur	39

Sommaires

4.1. Les cosmides	39
4.2. Les vecteurs «navettes»	39
4.3. Les vecteurs viraux eucaryotes	39
Chapitre VI : sondes moléculaires	41
1. Le concept de sonde	41
2. Les agents de marquages	41
2.1 Les isotopes radioactifs	41
2.2. Marquage non radioactif	41
3. Quelques stratégies de marquage	41
3.1. La « Nick translation »	41
3.2. La « Random printing »	42
3.3. Le marquage des sondes synthétiques (Oligo-nucléotides de synthèse)	42
3.4. Les sondes ARN (ribosondes)	42
Chapitre VIII : la transformation génétique	44
1. Principe de la transformation	44
2. Mécanisme naturel de transfère d'ADN	44
3. Les techniques de transformation	45
3.1. Transformation par choc physique	45
3.1.1. Transformation par choc thermique	45
3.1.2. Transformation par choc électrique	46
3.2. Transformation chimique	47
3.3. Biolistique	48
3.4. Microinjection	49
3.5. Agrotransformation	49
Chapitre IX : Mutation et mutagénèse	54
1. Evolution des génomes	54
2. Les mutations à la base de la diversité génétique	54
3. Les différents types de mutation	55
4. Les sources naturelles de mutagénèse	56
5. L'approche de génétique directe et génétique reverse en recherche	56
6. La transmission de mutation	57
7. La mutagénèse aléatoire	58

Sommaires

7.1. Mutagenèse par agents chimique	58
7.2. Mutagénèse par agents physique	59
7.4. Mutagénèse insertionnelle à l'ADNT	59
8. Le crible de mutants	59
9. Identification de mutations	61
9.1. Identification de mutations ponctuelles	61
9.2. Identification de mutation causée par une insertion	61
10. Validation de gènes candidat	62
11. Mutagénèse dirigée	62
12. Mutagénèse en système bactérien & fongique	62
13. Invalidation de gènes par approche RNA silencing	63
14. Invalidation de gènes par approche ribozyme	64

Partie II : Travaux dirigés

TD1. Evolution expérimentale	67
TD2. Caractérisation d'un mutant	73
TD3. Clonage moléculaire	79
TD4. Génomique	86
TD5. Transcriptome et analyses fonctionnelle	93

Partie III : Travaux pratiques

Travail pratique numéro 1	102
Travail pratique numéro 2	113
Travail pratique numéro 3	127
Travail pratique numéro 4	137

Liste des Abréviations

ADN : Acide Désoxyribonucléique
ARN : Acide Ribonucléique
ARNm : Acide Ribonucléique message
ARNt : Acide Ribonucléique transfert
ARNr : Acide Ribonucléique ribosomique
miRNA : micro Acide Ribonucléique
siRNA : petit Acide Ribonucléique interférents
RE: Réticulum Endoplasmique
AG: Appareil de Golgi
C1: Carbone 1
C3: Carbone 2
C5: Carbone 5
ORF: Open Reading Frame ; Séquence Ouverte de lecture
bHLH: basic Helix loop Helix
ADNase: Enzyme qui dégrade l'ADN
ARNase: Enzyme qui dégrade l'ARN
PCR: Polymerase Chain Reaction
Taq: ADN polymerase
FISH: Fluorescence *In Situ* Hybridization
ADNg: ADN génomique
ADNc: ADN complémentaire
YAC: Yeast Artificial Chromosome
BAC: Bacterial Artificial Chromosome
LacZ: gène qui code la β -galactosidase
OGM: Organisme Génétiquement Modifié
PEG: Polyéthylène Glycol
Ti: Tumor induced
Vir: Virulence
ADNT: ADN *d'A tumefascience*
KO: Knuck Out
KI: Knuck In
KD: Knuck Down

Liste des Abréviations

EMS: Ethyle Méthyl Sulfonate

WT: Wilde Type

RecA: Recombase A

Liste des figures & tableaux

Figure 1. Ultrastructure d'une bactérie, d'une cellule animale et végétale	3
Figure 2. Structure de l'ADN	5
Figure 3. Étapes clés de la transcription des gènes	6
Figure 4. Les différentes phases de la traduction des ARNm	8
Figure 5. Éléments cis-trans dans la régulation d'expression des gènes	10
Figure 6. Les différents niveaux de régulation de l'expression génique	11
Figure 7. Les étapes clés du RNA silencing	13
Figure 8. Étape d'amplification de l'ADN par PCR	21
Figure 9. Principe du clonage moléculaire	24
Figure 10. Les étapes relatives de la création des banques	25
Figure 11. Prototype d'un vecteur plasmidique	30
Figure 12. Les différentes étapes du clonage plasmidique	31
Figure 13. Le clonage par restriction-ligation	32
Figure 14. Génération de site de restriction par technique PCR	33
Figure 15. Principe de crible bleu-blanc	34
Figure 16. Principe du clonage par recombinaison	36
Figure 17. Les étapes de clonage dans un phage	38
Figure 18. Principe de la transformation	44
Figure 19. La transformation par choc thermique	46
Figure 20. Principe de l'électroporation	47
Figure 21. Transformation chimique de levure par lithium acétate	48
Figure 22. La transformation par biolistique	49
Figure 23. Schémas du prototype du plasmide Ti	50
Figure 24. Mécanisme d'infection par ADNT d' <i>A. tumefaciens</i>	52
Figure 25. La génétique directe et reverse en recherche	57
Figure 26. Principales étapes de recherche en génétique directe	60
Figure 27. Protéines RecA dans la recombinaison homologue	63
Figure 28. Principe du RNA silencing	64
Figure 29. Cycle d'action du ribozyme	65
Tableau 1. Principaux vecteurs et leurs caractéristiques	29

Partie I : Cours

Chapitre I : Généralités en biologie moléculaire

1. La cellule, unité de base du vivant

La cellule constitue l'unité de base des organismes vivants, on distingue deux grands types cellulaires, les procaryotes dépourvus d'un noyau qui est remplacé par un espace nucléoplasmique et les cellules eucaryotiques qui présentent un noyau dans lequel est inclus le matériel génétique (figure 1).

Une cellule est composée d'une membrane plasmique qui délimite son espace interne ou intracellulaire du milieu externe ou extracellulaire. La membrane plasmique est une bicouche lipidique contenant des protéines assurant les fonctions biologiques essentielles au maintien de l'homéostasie cellulaire. Chez les bactéries et les cellules végétales, on note la présence d'une paroi entourant la membrane plasmique, elle intervient dans la protection de la cellule (figure 1).

Le cytoplasme est la matrix intracellulaire dans laquelle le reste des composés de la cellule sont maintenus, il comprend le cytosol contenant l'ensemble des ions et molécules nécessaires aux échanges et aux modifications chimiques cellulaires. Le cytosquelette quant à lui est responsable du maintien structurel de la cellule et il constitue un des réseaux de transport intracellulaire.

Au sein du cytoplasme on distingue la présence de plusieurs organites cellulaires, chacun présente une membrane qui sépare son espace interne du reste du cytoplasme (figure 1). Le noyau est l'organite qui contient la base de l'information génétique qui est sous forme d'Acide Désoxyribo-Nucléique (ADN) et de protéines associées. L'ADN contient l'ensemble des gènes responsables des caractères d'un individu, L'expression des gènes (ou expression génique) aboutie par le processus de transcription à la production d'Acide Ribo Nucléique (ARN) qui peuvent dans le cas des ARN messager (ou ARNm) coder des protéines essentielles à la formation et au maintien de la cellule et de l'organisme plus largement. D'autres ARN interviennent dans la biosynthèse de protéines, tels que les ARNs ribosomiaux (ARNr) et les ARNs de transport (ARNt).

Le réticulum endoplasmique (RE) est un organite au contact du noyau (figure 1), il contient les ARNr qui traduisent l'information contenue dans les ARNm en protéines libérées dans la lumière du réticulum. L'appareil de Golgi (AG) est responsable de la maturation des protéines

Partie I : Cours
Chapitre I : Généralités en biologie moléculaire

et des modifications post traductionnelles, il fait partie de la voie par laquelle les protéines transitent avant leurs sécrétions dans le milieu extracellulaire, les protéines destinées au cytoplasme sont généralement traduites directement dans le cytosol.

La mitochondrie est l'organite responsable de la production d'énergie par la catalyse de composés polycarbonés. Enfin le chloroplaste retrouvé chez les organismes photosynthétiques est responsable de la transformation de l'énergie lumineuse en énergie chimique (figure 1).

Partie I : Cours
Chapitre I : Généralités en biologie moléculaire

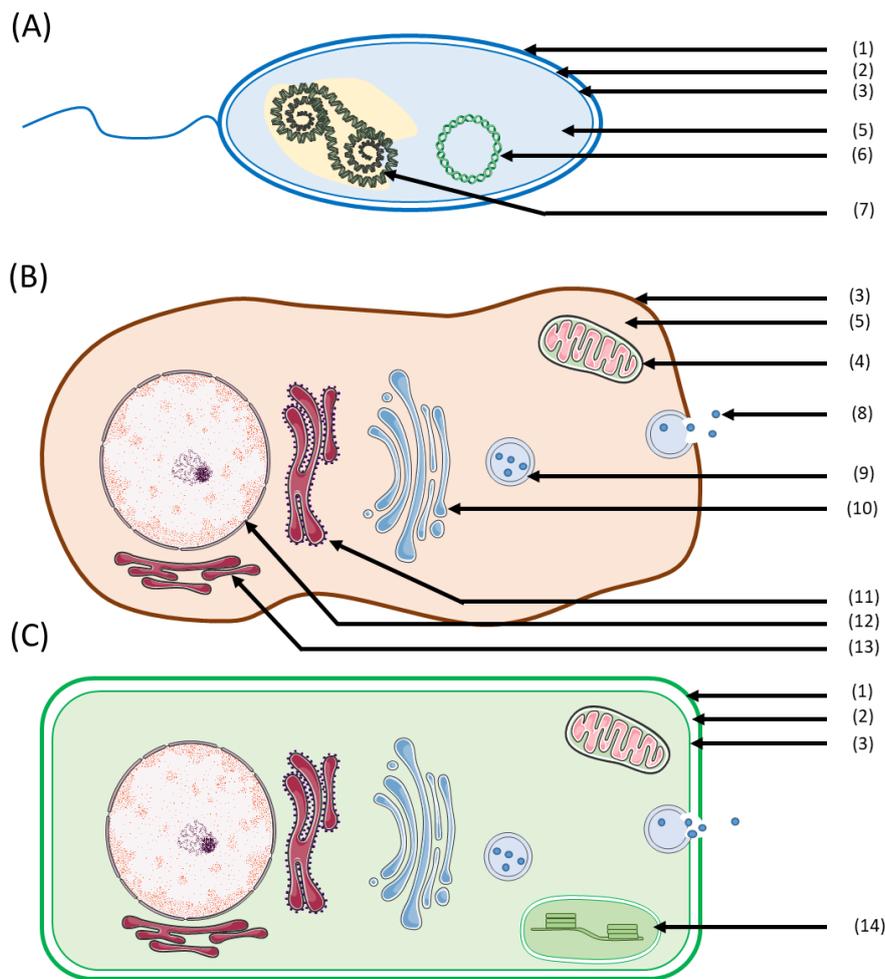


Figure 1. Ultrastructure d'une bactérie, d'une cellule animale et végétale. Le schéma représentant les composés principaux retrouvés au sein d'une bactérie (A), une cellule animale (B) ou d'une cellule végétale (C). (1) : paroi cellulaire (qui est un peptidoglycane chez les bactéries et de nature pectocellulosique dans les cellules végétales), (2) ; espace periplasmique, (3) ; membrane plasmique, (4) ; mitochondrie, (5) ; cytoplasme, (6) ; plasmide, (7) ; espace nucléoplasmique contenant le génome bactérien, (8) ; protéines sécrétées, 9 ; vésicule de sécrétion, 10 ; Appareil de golgi, 11 ; réticulum endoplasmique rugueux, 12 ; noyau, 13 ; réticulum endoplasmique lisse, 14 ; chloroplaste.

2. L'ADN base de l'information génétique

L'ADN est un enchainement de désoxyribonucléotides qui sont des composés chimiques comportant un désoxyribose lié à un phosphate en position C5 et une base azotée en position C1. On distingue deux grands types de bases azotées, (i) les bases puriques comportant deux

Partie I : Cours

Chapitre I : Généralités en biologie moléculaire

noyaux aromatiques, elle comprend la guanine et l'adénine, (ii) les bases pyrimidiques comportant un noyau aromatique, ce sont la cytosine et la thymine, cette dernière est remplacée par l'uracile au sein des ARNs qui à l'inverse de l'ADN sont un enchainement de ribonuécotides. Deux nucléotides successifs sont liés par le phosphate entre le carbone C5 et C3, l'orientation de la séquence d'ADN est par convention du carbone 5 au 3 (5'->3').

Chaque séquence d'ADN ou brin, s'hybride de manière antiparallèle à un second brin complémentaire produisant une hélice bi-caténaire. L'association entre les brins est permise grâce à des liaisons hydrogènes établies entre les bases azotées de telle sorte qu'une base pyrimidique se lie à une base purique ; ainsi l'adénine établie deux ponts hydrogènes avec la thymine et la guanine réalise trois ponts hydrogènes avec la cytosine (figure 2).

Dû au phosphate, l'hélice bi-caténaire est chargée négativement et s'enroule sur des protéines basiques chargées positivement nommées histones. L'association ADN-histones forme les nucléosomes qui se polymérisent et s'associent à des protéines de structures pour former la chromatine, une forme condensée de l'ADN.

La succession des nucléotides d'une molécule d'ADN détermine sa séquence, en fonction de cette dernière une information peut être portée et décodée par les machineries cellulaires telles que la transcription.

Partie I : Cours

Chapitre I : Généralités en biologie moléculaire

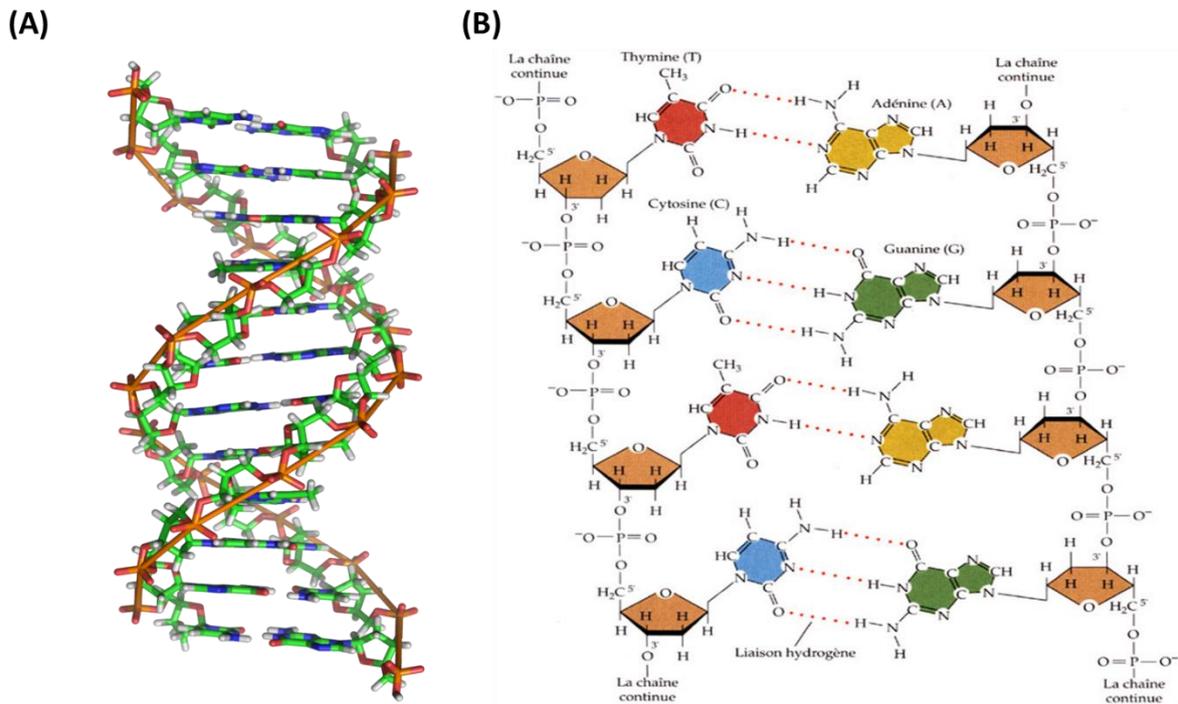


Figure 2. Structure de l'ADN. (A) structure tridimensionnelle d'une molécule d'ADN bicaténaire. (B) Composition chimique de l'ADN ; chaque brin est composé d'un enchaînement de désoxyribonucléotides qui sont une association entre un désoxyribose liée à un phosphate en C5 et une base azotée en C1. Les désoxyribonucléotides d'un brin sont liés entre eux par des phosphates établissant une liaison entre le carbone C5 et C3 de deux désoxyriboses successifs. Deux brins s'hybrident de manière complémentaire par l'établissement de deux ponts hydrogènes entre l'adénine (jaune) et la thymine (rouge) ou trois ponts hydrogènes entre la guanine (vert) et la cytosine (bleu).

3. La transcription et la traduction

Le génome d'un organisme correspond à l'ensemble des séquences d'ADN dont certaines sont des gènes, l'expression de ces derniers contrôle l'ensemble des événements requis pour l'accomplissement du cycle de vie d'un organisme.

Le phénomène d'expression résulte de la transcription aboutissant à la production d'ARNm à partir de la séquence ouverte de lecture (ORF : Open Reading Frams) du gène et de la traduction qui aboutit à l'assemblage d'une séquence protéique spécifique suite à la lecture de l'ARNm, les protéines produites interviennent ainsi dans l'ensemble des processus biochimiques.

Partie I : Cours

Chapitre I : Généralités en biologie moléculaire

La transcription débute par la fixation sur le promoteur de facteurs de transcriptions qui recrutent l'ARN polymérase. En synergie avec les facteurs et cofacteurs de transcription l'ARN polymérase procède à la synthèse d'une molécule d'ARNm à partir de la séquence de l'ADN du gène, l'ARNm synthétisé est complémentaire à l'ADN (figure 3). Chez les procaryotes, l'ARN messager produit est directement traduit, alors que chez les eucaryotes il subit un processus de maturation ou d'épissage lors de sa migration du noyau vers le cytoplasme, au cours de laquelle les introns (séquences non-codantes) sont retirés pour aboutir à la synthèse d'un ARNm mature contenant uniquement les exons (séquences codantes ou CDS).

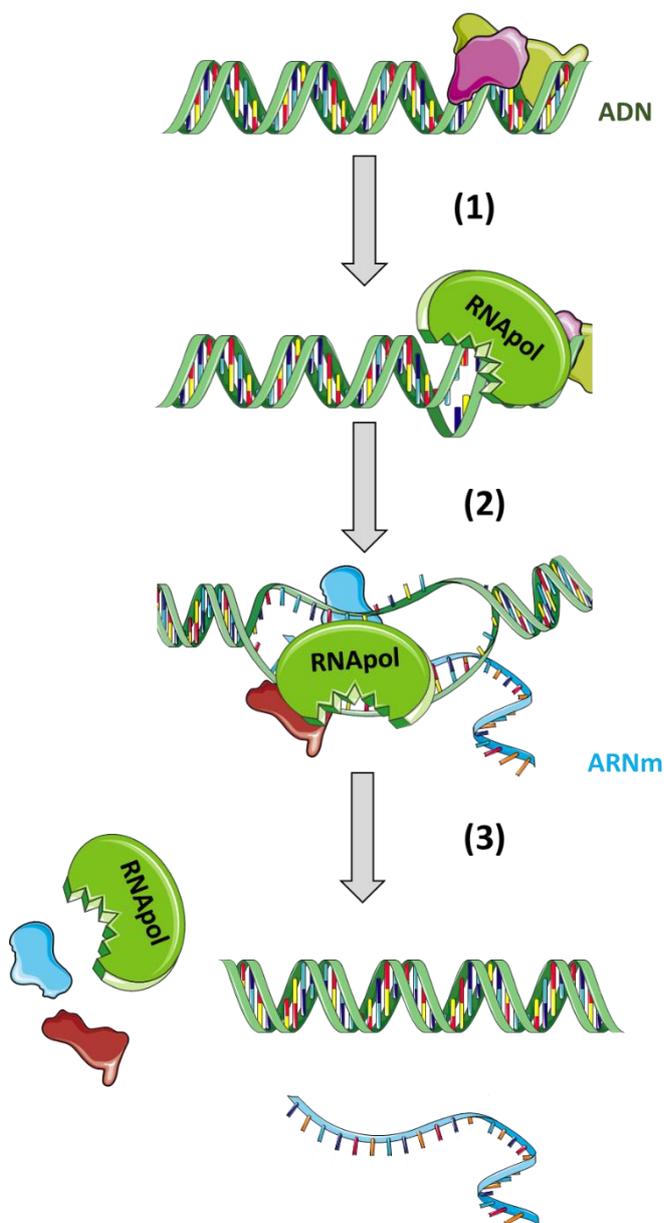


Figure 3. Étapes clés de la transcription des gènes. Après la reconnaissance des éléments cis par les facteurs de transcriptions ou éléments trans, le recrutement de l'ADN polymérase à lieu (1). Suite à cela l'ouverture de la molécule d'ADN à lieu et l'ARN polymérase à avec l'aide des cofacteurs de transcription procède à la biosynthèse d'une molécule d'ARNm complémentaire à l'ADN (2). Une fois la transcription achevée, l'ARN polymérase et les cofacteurs de transcription se décrochent et libèrent un ARNm (3).

Partie I : Cours

Chapitre I : Généralités en biologie moléculaire

Dans le cytosol, les ARNm sont reconnus par les facteurs de traduction, les ARN ribosomiaux et leurs protéines associées. L'ensemble procède au décodage de la séquence de l'ARNm via le processus de traduction (figure 4).

Le ribosome (ARNr et protéine ribosomique) se fixe sur le premier triplet (codon) de nucléotide qui est l'ATG, ce dernier est reconnu par l'ARN de transfert complémentaire au niveau du site A (amino acide) du ribosome, l'ARNt porte l'acide aminé méthionine (ARNt-met). Suite à la fixation du première ARNt, le complexe ribosomique se décale, l'ARNt-met est transloqué vers le site P (peptide) du ribosome, libérant ainsi le site A qui reçoit un autre ARNt possédant l'anti codon complémentaire au second codon de la séquence. Il est important de noter que chaque codon reconnaît un anti codon d'un ARNt complémentaire, ce dernier porte un acide aminé spécifique.

Une liaison peptidique est établie entre les acides aminés des deux ARNt (en position P et A), la liaison établie entre l'ARNt et la méthionine est rompue, puis l'ARNt est rejeté hors du ribosome via le site E (expulsion) au moment du second décalage. Ce processus est réitéré jusqu'à la lecture entière de la séquence d'ARNm (figure 4).

Partie I : Cours

Chapitre I : Généralités en biologie moléculaire

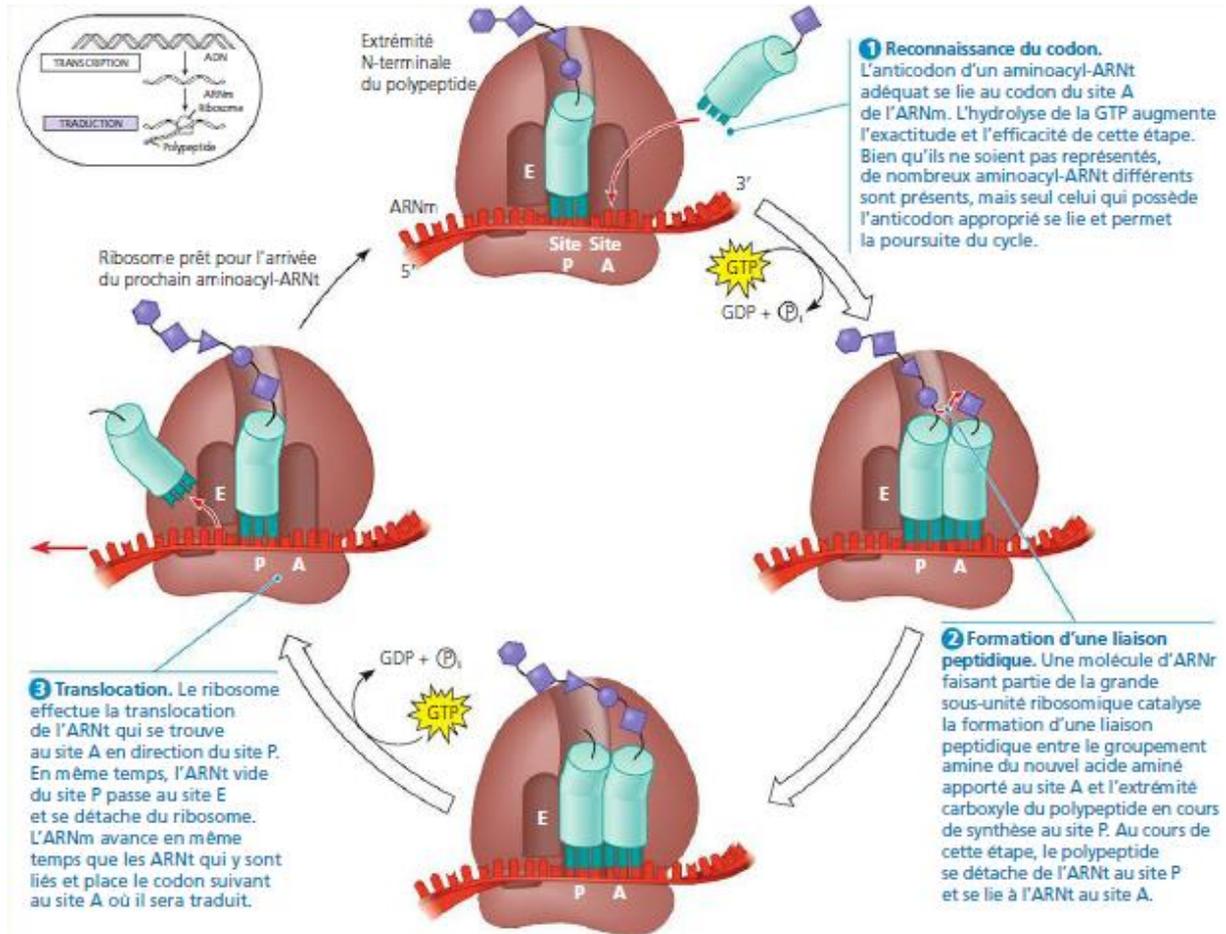


Figure 4. Les différentes phases de la traduction des ARNm. Quatre phases cycliques peuvent être différenciées : (1) la phase de reconnaissance du codon pendant laquelle après la fixation du ribosome, l'ARNt complémentaire (anti-codon) au triplet de nucléotide de l'ARNm s'hybride au niveau du site ribosomique A. (2) Suite à cela une liaison entre l'acide aminé porté par l'ARNt du site A et le reste du peptide porté par un ARNt situé dans le site P à lieu. (3) après la formation de la liaison, la chaîne peptidique est transférée sur l'ARNt du site A et le second ARNt est transféré vers le site E (4) dans lequel il est exclu du ribosome.

4. Promoteur de gènes et régulation Cis-trans

Un gène est une partie de l'ADN comprenant une information impliquée dans le fonctionnement d'un organisme ou d'une cellule. Les gènes sont placés sous un contrôle stricte afin de synchroniser et d'harmoniser leurs effets.

Partie I : Cours

Chapitre I : Généralités en biologie moléculaire

Le promoteur est la séquence d'ADN responsable du contrôle de l'expression d'un gène, il est généralement situé en amont de ce dernier et il exerce son effet via ces interactions avec les protéines requises pour l'expression des gènes tels que les facteurs de transcriptions.

Certains gènes sont exprimés de manière constitutive au cours du cycle de vie, ces gènes sont communément appelés « gènes de ménages » ou « house keeping gens » telle que le gène codant la flagelline bactérienne. Les autres gènes sont soumis à une régulation spatio-temporelle, elle est due à l'activation spécifique de facteurs de transcriptions qui interviennent dans le recrutement de la machinerie transcriptionnelle.

Un promoteur contient généralement deux types de séquences, celles de structure qui assure le maintien structurelle du promoteur et les séquences Cis liant les facteurs de transcription ou éléments Trans. La reconnaissance entre le facteur de transcription et le promoteur a lieu via l'interaction entre des motifs protéiques du facteur de transcription et les séquences Cis. Plusieurs motifs existent telle que le « leucines zipper riche » qui est riche en résidus leucines ou les « bHLH » qui forment une structure en hélice-boucle-hélice (figure 5).

La reconnaissance entre les protéines et l'ADN résulte de la capacité de ces dernières à établir des interactions chimiques thermodynamiquement favorables, ceci dépend dans le cas d'une régulation Cis-Trans de la séquence d'ADN et du type de motifs protéique, ainsi certains motifs protéiques reconnaissent de manière spécifique certaines séquences Cis (figure 5).

Partie I : Cours
Chapitre I : Généralités en biologie moléculaire

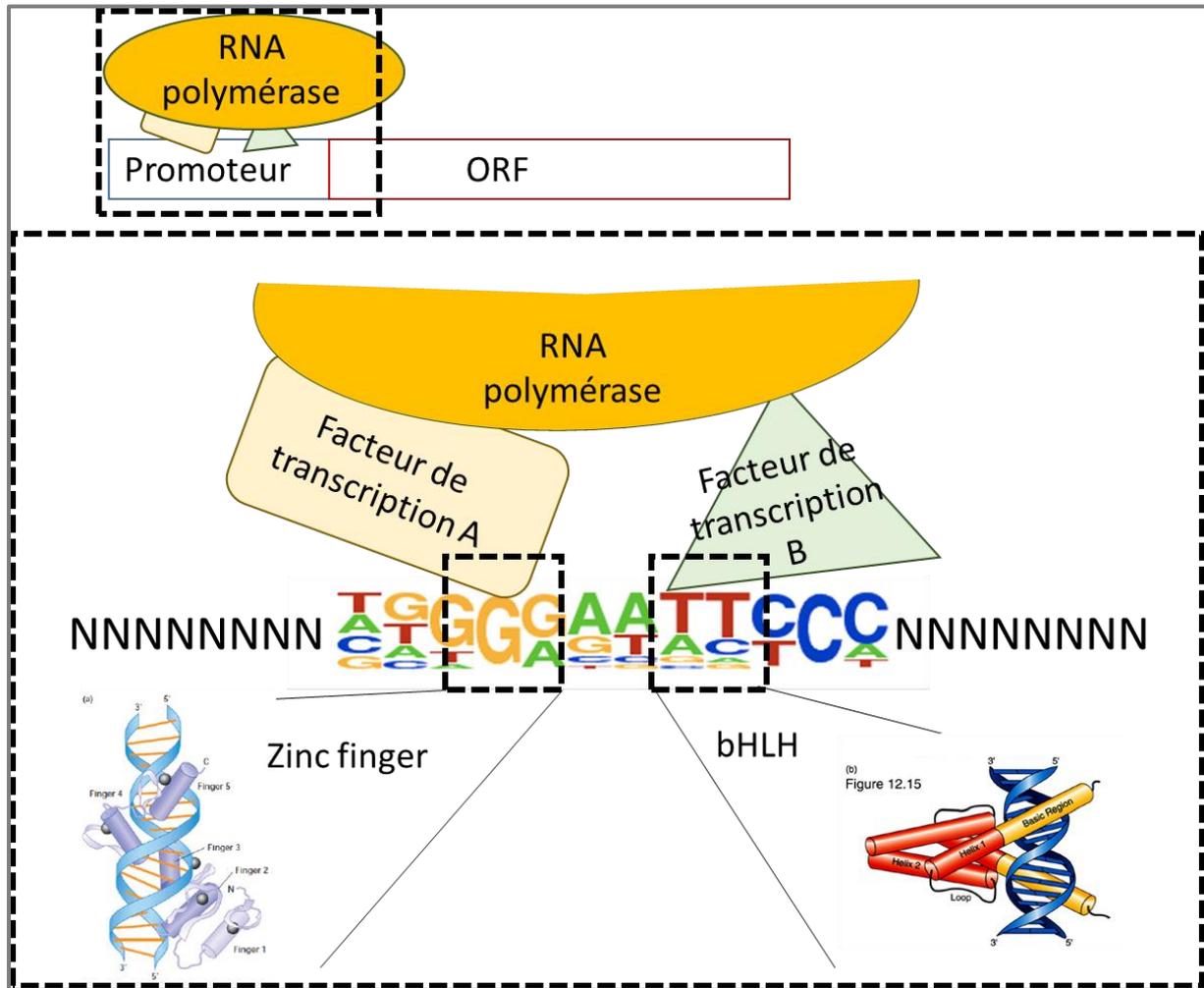


Figure 5. Éléments cis-trans dans la régulation d'expression des gènes. Les facteurs de transcriptions reconnaissent des motifs d'ADN ou éléments Cis au sein du promoteur via des motifs protéique. Après laissent les facteurs de transcription recrute l'ARN polymérase (RNA polymérase).

5. Les autres niveaux de régulation d'expression de gènes

Au cours de leurs évolutions les organismes ont développés plusieurs niveaux de régulation des gènes afin d'assurer un contrôle stricte de l'expression. Un gène échappant à la régulation peut représenter un danger pour une cellule et l'organisme à plus grande échelle.

A chaque étape de la synthèse d'une protéine et une fois qu'elle devient fonctionnelle la cellule peut intervenir pour arrêter le processus de production. Ainsi la biosynthèse et l'activation du facteur de transcription est le première élément de contrôle, en plus de la régulation de l'accessibilité du promoteur aux facteurs.

Partie I : Cours

Chapitre I : Généralités en biologie moléculaire

Lors de la transcription, de la maturation des ARNm et de la traduction, la cellule peut arrêter à tous moment la synthèse de la protéine, une fois que cette dernière est produite les modifications post traductionnelles peuvent avoir lieu afin de contrôler la durée de vie et l'activité de la protéine. De plus, si une protéine à la capacité de former des complexes protéiques, l'assemblage du complexe peut être régulé (figure 6).

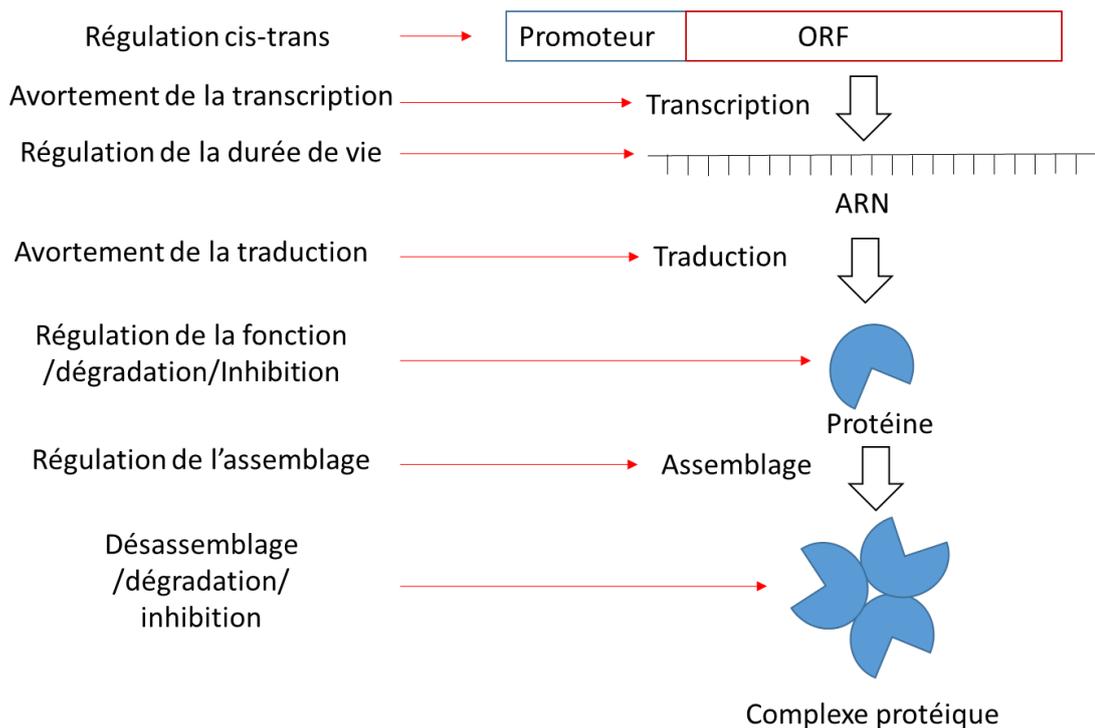


Figure 6. Les différents niveaux de régulation de l'expression génique. La cellule peut réguler l'expression des gènes et de leur effet à toute les étapes du processus : au niveau du promoteur par l'interaction entre les éléments cis et trans, elle peut avorter la transcription et la traduction. Réguler la durée de vie des ARNm et des protéines. Réguler l'activité des protéines par les modifications post traductionnelles et contrôler la formation des complexes protéiques.

6. La régulation par petits ARNs

Certains gènes produisent après transcription des ARNs de petite taille, destinés à intervenir dans la régulation d'expression de gènes cibles qui leurs sont au moins partiellement

Partie I : Cours

Chapitre I : Généralités en biologie moléculaire

homologue. Les petits ARN ou microARN (ou miRNA) sont produit par la machinerie transcriptionnelle et présentent une séquence conservée avec le gène qu'ils ciblent.

Les miRNA sont produit à partir de précurseurs qui sont des ARNs formant une structure tige boucle, en épingle ou double brin. Ces précurseurs sont reconnus par un complexe de protéines appelé DICER qui les clives pour libérer les miRNA. Ces derniers sont pris en charge par des protéines ARGONAUT et RISC, le complexe miRNA/ARGONAUT/RISC traque dans la cellule les ARNm complémentaires, qui correspondent aux gènes régulés par le miRNA (figure 7).

Dans le cas où le complexe reconnaît un ARNm, il procède au recrutement du DICER qui clive le transcrit et libère des petits ARNs anti sens nommés siRNA (small interference RNA) qui sont à leur tour recrutés par ARGONAUT et RISC afin de reconnaître les transcrits cibles restants dans la cellule. Ce processus est désigné par l'extinction d'expression de gènes ou « silencing » (figure 6).

Partie I : Cours
Chapitre I : Généralités en biologie moléculaire

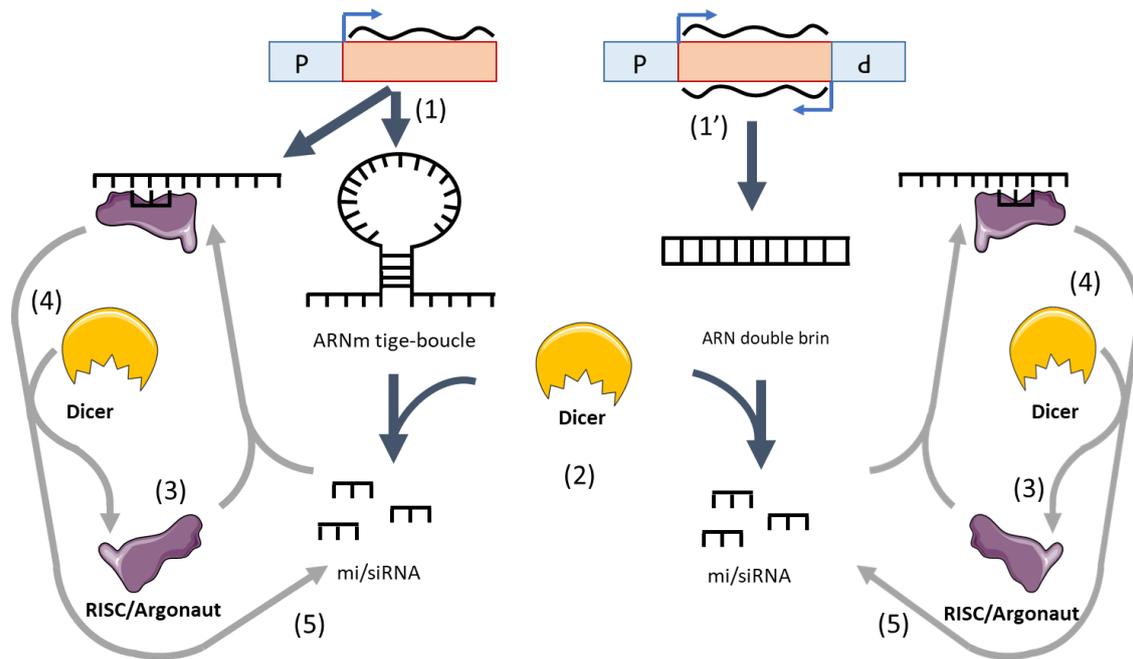


Figure 7. Les étapes clés du RNA silencing. L'ARN précurseur de miRNA est transcrit et présentant une structure modifiée. Deux processus peuvent aboutir à la formation d'un ARN de structure modifiée (1) après transcription les ARN peuvent présenter une structure tige boucle liée à la présence de séquences complémentaires ou (1') une double transcription donne un ARN double brin. Les ARN précurseur sont reconnus par le DICER qui les clives et libères des petits ARNs appelé miRNA (2). En association avec les protéines RISC/Argonaut les miRNA reconnaissent les transcrit des gènes cibles (3) et en collaboration avec le DICER, il les dégrade en petit ARN interférents appelés siRNA (3), de même manière que les miRNA, les siRNAs procèdes à la dégradation du reste des transcrits. La réaction s'arrête après la dégradation de l'ensemble des ARNm.

6. Notions d'espèces modèles

Dans le but de synchroniser les efforts et d'avancer de manière plus rapide dans la découverte des processus biologiques, la communauté scientifique a décidée de choisir certaines espèces sur lesquelles les efforts de recherches sont synchronisés, ces espèces sont nommées modèle car les mécanismes moléculaires mis en évidence dans ces dernières peuvent être utilisés de base de réflexion pour d'autre espèces proches ou éloignées.

Partie I : Cours

Chapitre I : Généralités en biologie moléculaire

Le choix d'une espèce modèle repose sur des critères biologiques et techniques, ainsi le cycle de vie est pris en considération, une espèce ayant un cycle court est plus intéressante qu'une espèce ayant un cycle long.

La complexité réduite d'un génome et une faible taille sont aussi des critères importants en plus de la possibilité d'appliquer des approches expérimentales en biologie fonctionnelle sur l'organisme en question. En fonction du règne plusieurs espèces ont été choisies, par exemple *Arabidopsis thaliana*, *Medicago truncatula* ou *Lotus japonicus* font partie des modèles végétaux, alors que *Escherichia coli* est un modèle bactérien.

7. Introduction au génie génétique

Le génie génétique est une science qui s'attache à l'étude du vivant ainsi que son amélioration, il comprend l'ensemble des techniques et approches utilisés dans la modification et l'étude des gènes. L'approche employée dans ce cours consistera à présenter les techniques les plus utilisées en recherche et génie génétique intégrées dans leurs contextes biologiques et conceptuels.

Partie I : Cours

Chapitre I : Généralités en biologie moléculaire

Références bibliographiques

- (1) Alberts, Bruce et al. 2008. *Molecular Biology of THE CELL Fifth Edition*. Edition, F. Garland Science.
- (2) Ameziane, Nadjma, Marc Bogard, and Jerome Lamoril. 2006. *Principes de Biologie Moléculaire En Biologie Clinique*. First. Elsevier.
- (3) Bartlett, John M S, and David Stirling. 226 *PCR Protocols*. Second. Human Press Inc.
- (4) Callen, Jean-claude. 2005. *Biologie Cellulaire, Des Molécules Aux Organismes*. Second. SCIENCES SUP.
- (5) Callen, Jean-Claude. 2009. *Biologie Cellulaire En 30 Fiches*. DUNOD.
- (6) David P. Clark, Nanette J. Pazdernik. 2012. *Biotchnology*. First. AP Cell.
- (7) Geoffrey M, Cooper, and Hausman Robert E. 2006. *The Cell A Molecular Approach*. Fouth. ASM Press.
- (8) Hart, Tony, and Paul Shears. 1999. *ATLAS DE POCHE DE MICROBIOLOGIE*. First. Flammarion.
- (9) KIERSZENBAUM, Abraham L. 2002. *Histologie et Biologie Cellulaire*. First. de boeck.
- (10) Kumar, Surinder. 2016. *Essentials of Microbiology*. First. The Health Sciences Publisher.
- (11) Magdeldin, Sameh. 2012. *GEL ELECTROPHORESIS Principles and Basics*. First. InTech.
- (12) Pevsner, Jonathan. 2015. *Bioinformatics and Functional Genomics*. third. Wiley Blackwell.
- (13) Pevzner, Pavel, and Ron Shamir. 2011. *Bioinformatics for Biologists*. First. Cambridge University Press.
- (14) P.S, VERMA, and V, K, AGRAWAL. 2005. *Multicolour Edition CELL BIOLOGY , GENETICS , MOLECULAR BIOLOGY*. First. S. CHAND & COMPANY LTD :
- (15) Serre, Jean-louis. 2006. *Génétique*. third. DUNOD.
- (16) Susan J, Karcher. 1995. *Molecular Biology A Project Approach*. First. Academic Press.
- (17) Swynghedauw, Bernard, and Jean-Sébastien Silvestre. 2008. *BIOLOGIE ET GÉNÉTIQUE MOLÉCULAIRES*. Third. DUNOD.
- (18) Tagu, Denis. 1999. *Principes Des Techniques de Biologie Moléculaire*. First:. INRA édition.
- (19) Tang, Yi-Wei et al. 2015. *Molecular Medical Microbiology*. Second. Elsevier.

Partie I : Cours

Chapitre II : Les outils enzymatique du génie génétique

1. Les enzymes de restriction

Les enzymes de restrictions sont des protéines ayant une activité enzymatique caractérisée par une capacité de couper l'ADN double brin, de ce fait elles font partie de la grande famille des endonucléases.

Les enzymes de restriction ont été identifiées au début des années 50 chez les bactéries. Ces enzymes permettent de dégrader l'ADN étranger pénétrant au sein des cellules bactériennes incluant ceux des phages.

1.1. Le phénomène de restriction

Le phénomène de restriction est une réaction chimique qui aboutit au clivage d'un ADN bicatenaire de manière séquence spécifique. Cette réaction est médiée par les enzymes de restrictions.

Il existe dans la nature un nombre important d'enzymes isolées à partir des bactéries, chaque enzyme reconnaît une séquence spécifique de 4 à 8 paires de bases qui forment un palindrome, c'est-à-dire ayant la même séquence sur les deux brins, mais en direction opposée.

La coupure de l'ADN peut aboutir à la libération de bouts francs ou de bouts cohésives, ces derniers à l'inverse des bouts francs sont complémentaire et peuvent se ré-hybrider (reformer un double brin).

1.2. Nomenclature des enzymes de restriction

Dans la nature on peut identifier trois grandes classes d'enzyme de restriction, les enzymes de classe/type I, II et III.

1.2.1. Enzymes de type I

Les enzymes de type I sont des hétéro multimères dont la caractéristique principale réside dans le fait que la séquence reconnue et celle qui subit la réaction de clivage sont séparées. Ainsi l'enzyme après reconnaissance clive de manière non-spécifique une séquence d'ADN situé de 100 à 1000 pb en aval du site de reconnaissance.

Partie I : Cours

Chapitre II : Les outils enzymatique du génie génétique

1.2.1. Enzymes de type II

Les enzymes de type II sont des monomères ou homo multimères. A l'inverse des enzymes de type I, celle de type II réalisent la reconnaissance et le clivage sur la même séquence, de ce fait le clivage devient site spécifique. Ce type d'enzyme est celui le plus employé en génie génétique, notamment pour le clonage moléculaire.

1.2.2. Enzymes de type III

De manière similaire aux enzymes de classe I, le type III se présente sous forme d'un complexe hétéro multimérique dont la séquence de reconnaissance et celle de clivage sont séparées physiquement. A l'inverse de la classe I, les enzymes de classe II clivent une séquence moins éloignée de la séquence de reconnaissance (25 à 27 pb en aval de la séquence de reconnaissance).

1.3. Types de coupures induites par les enzymes de restrictions

Les enzymes de restriction peuvent générer deux types de coupures au sein des séquences, la première permet de libérer des bouts francs qui ne peuvent s'hybrider entre eux, le second libère des bouts à extrémités cohésives qui peuvent s'hybrider par complémentarité de séquence. Le deuxième type de coupure est privilégié en génie génétique car il permet de réaliser le clonage moléculaire par technique de restriction-ligation.

2. Les autres enzymes d'usage courant en biologie moléculaire

En génie génétique il existe en plus des enzymes de restrictions utilisées traditionnellement pour le clonage moléculaire, d'autres familles d'enzymes utilisées dans différentes expérimentations.

2.1. Les polymérases

Les ADN polymérases sont des enzymes ayant la capacité de néo-synthétiser de l'ADN à partir d'une matrice d'ADN simple brin ou d'un ARN pour les reverse transcriptases. La taq polymérase est l'une des ADN polymérase les plus utilisée au monde, sa spécificité réside dans sa capacité à résister à de forte température. Elle est utilisée dans la PCR (Polymerase Chain Reaction) qui est une technique permettant d'amplifier de manière importante et spécifique des fragments d'ADN. La reverse transcriptase est utilisée pour la synthèse d'ADN complémentaire (ou ADNc) obtenus à partir d'une matrice d'ARNm.

Partie I : Cours
Chapitre II : Les outils enzymatique du génie génétique

2.2. Les ligases

Les ligases sont des enzymes qui ont l'aptitude de créer une liaison phosphate entre deux bases azotées successives. La réaction utilise de l'ATP qui est hydrolysé pour fournir l'énergie ainsi que le phosphate servant à engendrer la liaison.

Les ligases sont notamment utilisées en clonage moléculaire, par exemple afin de souder deux bouts cohésifs générés sous l'action d'une enzyme de restriction.

2.3. Les exonucléases

Les exonucléase sont des enzymes qui dégradent des acides nucléiques à partir de leurs extrémités (exo). Cette famille regroupe les ADNase qui dégradent l'ADN et les ARNase qui dégradent l'ARN. Ces enzymes sont employé en biologie moléculaire afin de nettoyer des surfaces de manipulation ou d'éliminer de l'ADN ou de l'ARN d'échantillons.

Partie I : Cours
Chapitre II : Les outils enzymatique du génie génétique

Références bibliographiques

- (1) Burrell, Michael M. 1993. *Enzymes of Molecular Biology*. First Edit. Humana Press.
- (2) Dalbey, Ross, Carla Koehler, and Fuyuhiko Tamanoi. 2007. *The Enzymes*. First Edit. Academic Press.
- (3) Rittié, Laure, and Bernard Perbal. 2008. "Enzymes Used in Molecular Biology: A Useful Guide." *J Cell Commun Signal* 2: 25–45.
- (4) Wilson, K. 2010. "Enzymes." In *Principles and Techniques of Biochemistry and Molecular Biology*, eds. Keith Wilson and JohnEditors Walker. Cambridge University Press, 581–624.

Partie I : Cours

Chapitre III : L'hybridation moléculaire

1. Rappel sur le principe de la réaction d'hybridation

L'hybridation est une réaction qui a lieu entre deux molécules d'acides nucléiques complémentaires, elle consiste à l'établissement de ponts hydrogènes entre les deux molécules, permettant la formation d'une molécule bicaténaire.

1.1. Notion de température de fusion de l'ADN

La température de fusion est définie comme la température à laquelle la moitié d'une molécule d'ADN bicaténaire se sépare. Cette température varie d'une molécule à une autre, ainsi que d'une condition à une autre. Elle est utilisée afin d'apprécier la température à laquelle deux molécules d'ADN/ARN s'hybrident.

1.2. Facteurs influençant la température de fusion

Plusieurs paramètres peuvent influencer la température de fusion notamment la nature du milieu chimique au sein duquel a lieu la réaction et le nombre de mésappariements entre les bases.

Le nombre de pont hydrogène au sein d'une molécule influence la température de fusion ainsi plus le nombre de liaisons est élevé et plus la température de fusion doit être importante. Deux paramètres principaux influencent le nombre de ponts hydrogènes, le premier est la longueur de la molécule et le second est la composition relative en Guanine/Cytosine.

2. L'hybridation en phase liquide

2.1. Principe

L'hybridation en phase liquide consiste à réaliser la réaction d'hybridation au sein d'un milieu contenant un tampon et du formamide. Sous l'agitation thermique les deux molécules se rapprochent l'une sur l'autre à une température inférieure à celle de la température de fusion.

2.2. Applications de l'hybridation moléculaire en phase liquide

Il existe un nombre important d'applications liées à l'hybridation moléculaire en phase liquide, l'une des plus importantes est l'amplification d'ADN par PCR (Polymerase Chain Reaction). Cette technique s'appuie sur l'utilisation d'un couple d'oligonucléotides de tailles variant entre 16 à 32 nucléotides qui s'hybrident de manière spécifique de part et d'autre de la séquence d'ADN à amplifier.

Partie I : Cours

Chapitre III : L'hybridation moléculaire

La réaction a lieu dans un tampon en présence d'oligonucléotides triphosphates, de $MgCl_2^+$ et nécessite l'action d'une ADN polymérase thermorésistante telle que la Taq. Après dénaturation de l'ADN à $94^\circ C$, une hybridation des amorces est réalisée en abaissant la température de 54 à $64^\circ C$, enfin à $72^\circ C$ l'enzyme se fixe sur l'amorce hybridée et procède à la néosynthèse du fragment d'ADN en intégrant les nucléotides complémentaires à la matrice. La procédure est répétée plusieurs fois (ou plusieurs cycles) afin d'augmenter la quantité du fragment ciblé (figure 8).

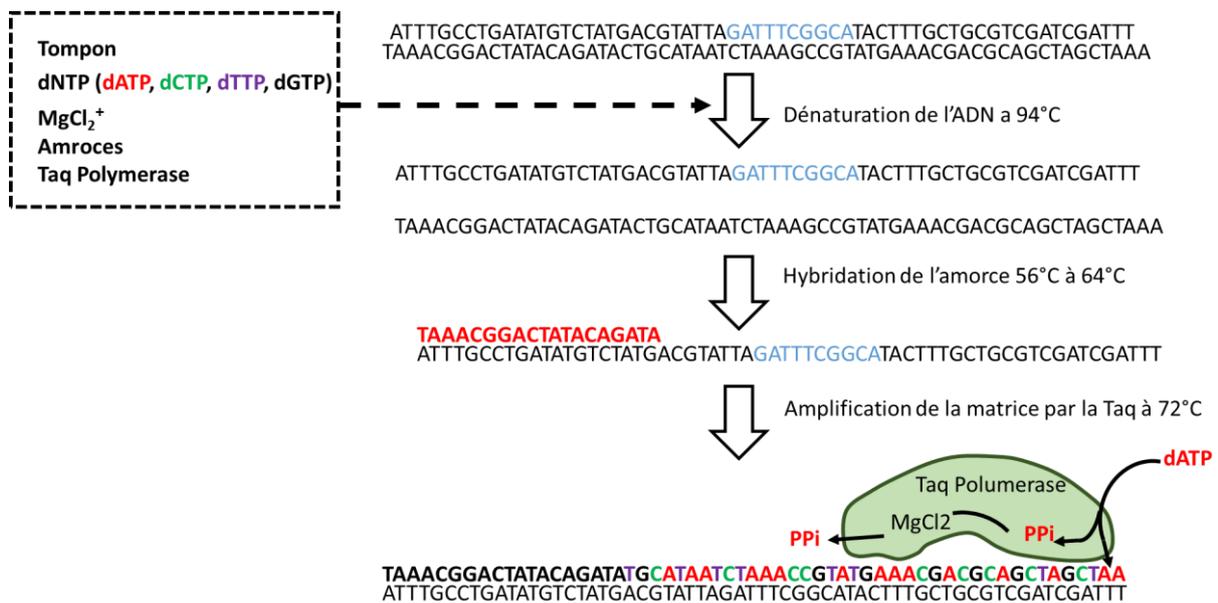


Figure 8. Étape d'amplification de l'ADN par PCR. Dans un tampon contenant des dNTP, du $MgCl_2^+$, des amorces et la taq, l'ADN est ajouté. Une dénaturation à $94^\circ C$ est réalisée, puis les amorces s'hybrident à une température variant de 56 à $64^\circ C$ en fonction de leurs compositions en G/C et de leurs longueurs. A $72^\circ C$ la Taq procède à l'intégration des nucléotides complémentaires à la séquence d'ADN. Chaque nucléotide triphosphate est hydrolysé et le pyrophosphate généré (PPI) est évacué par le $MgCl_2^+$.

3. L'hybridation sur support solide

3.1. Principe

L'hybridation des fragments d'ADNs cibles a lieu sur un support solide sur lequel est placé l'ADN appât complémentaire, l'ADN cible est apportés sous forme liquide. Cette approche permet un lavage des ADNs qui ne s'hybrident pas.

Partie I : Cours

Chapitre III : L'hybridation moléculaire

3.2. Les supports utilisés pour immobiliser les acides nucléiques

Plusieurs types de supports peuvent être utilisés, la nitrocellulose à 80°C permet la création de liaisons irréversibles avec les acides nucléiques à de hautes concentrations ioniques. Les membranes synthétiques telles que le nylon peuvent être aussi utilisées. Ce type de support après activation au rayon ultraviolet crée des associations plus stables avec les acides nucléiques que celles établies avec la nitrocellulose. Pour certaine technique telle que l'analyse d'expression de gènes par puce à ADN, une lame de verre est utilisée comme support d'immobilisation.

4. L'hybridation *in situ*

C'est une technique qui est employée pour la réalisation d'hybridation d'acides nucléiques sur des tissus fixés. Elle consiste en l'utilisation de sondes d'ADN marquées qui sont appliquées sur des coupes d'organe ou des cellules, les sondes ciblent les ARNm de gènes dont l'expression doit être localisée au sein de l'organe. Elles peuvent aussi cibler des séquences d'ADN dans le cas du FISH (Fluorescence In Situ Hybridation) afin de localiser un gène au sein du génome. Après hybridation, une étape de lavage à lieu dans le but de retirer le surplus de sondes, puis une révélation est réalisée afin de mettre en évidence les zones au sein desquelles les sondes se sont hybridées.

Le marquage des sondes peut être réalisé à chaud dans le cas d'un marquage avec un isotope radioactif tel que le phosphate radio marqué (P32) ou un marquage à froid qui consiste à l'ajout d'une molécule non-radioactive sur la sonde.

Partie I : Cours
Chapitre III : L'hybridation moléculaire

Références bibliographiques

- (1) Anderson, M. L. M. 1999. *Nucleic Acid Hybridization*. First Edit. CAB Direct.
- (2) Bartlett, John M S, and David Stirling. 226 *PCR Protocols*. Second. Human Press Inc.
- (3) Innis, Michael A, David H Gelfand, and John C Sininsky. 1999. *PCR Protocol for Functional Genomics*. First Edit. Academic Press.
- (4) Liehr, Thomas. 2017. *Fluorescence In Situ Hybridization (FISH)*. Springer, Berlin, Heidelberg.
<https://doi.org/10.1007/978-3-662-52959-1>.
- (5) Nuovo, Gerard J. 1994. *PCR In Situ Hybridization*. Humana Press.
<https://doi.org/10.1385/0-89603-280-9:223>.

Partie I : Cours

Chapitre VI : clonage moléculaire

1. Le principe du clonage

Le clonage moléculaire consiste à introduire une séquence d'ADN appelée insert dans un vecteur. Le vecteur recombiné (contenant l'insert) est introduit dans un virus ou un microorganisme (bactérie, levure). Le type de vecteur varie en fonction de l'organisme receveur. L'organisme est transformé par la construction moléculaire, après introduction, une phase de sélection est réalisée afin d'isoler les microorganismes transformés (figure 9).

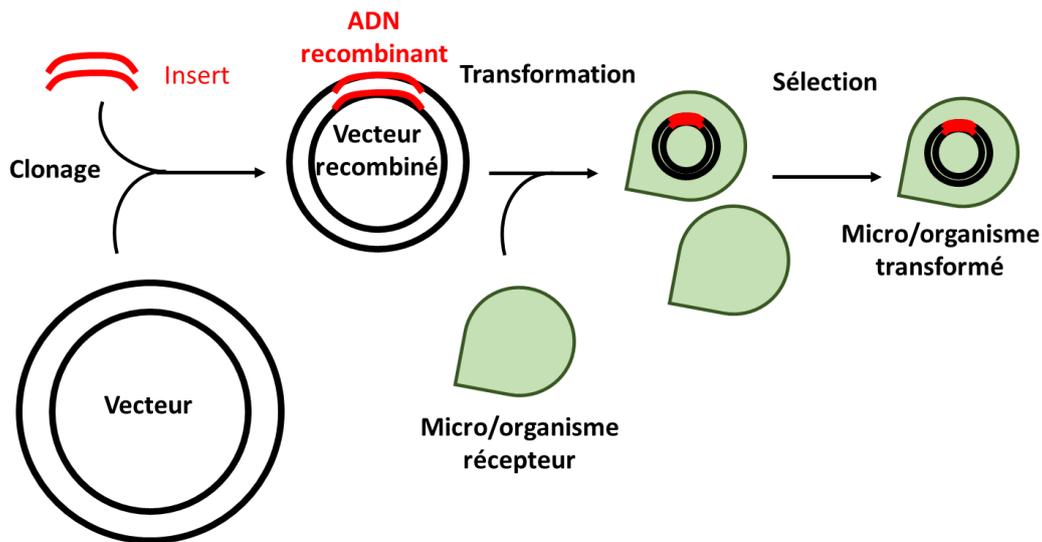


Figure 9. Principe du clonage moléculaire. L'insert (ADN à cloner) est introduit au sein d'un vecteur. Après intégration le vecteur est dit recombiné et il est introduit par transformation dans un microorganismes/virus récepteur. Les microorganismes ayant acquis les vecteurs recombinés sont sélectionnés.

2. Les banques d'ADN

2.1. Les banques d'ADN génomique

Un des progrès considérable en génétique a été la mise en place d'approche permettant d'obtenir la séquence des génomes. L'une d'entre elle consiste à stocker les séquences de génome entier, ce sont les banques d'ADN, qui ont contribuées grandement dans l'essor de la discipline.

2.1.1. Etablissement de la banque d'ADN

La création d'une banque d'ADN a pour objectif de stocker et de multiplier un génome entier pour un organisme donnée et cela à terme afin de séquencer le génome.

Partie I : Cours

Chapitre VI : clonage moléculaire

La création d'une banque passe par différentes étapes, après extraction de l'ADN génomique, une sonication ou une digestion enzymatique est réalisée afin de le fragmenter. Les fragments purifiés sont clonés dans des vecteurs dédiés de telle sorte que chaque fragment se retrouve inséré dans un vecteur. Suite à cela les vecteurs recombinés sont introduits dans le microorganisme récepteur. Après sélection des clones (microorganismes) transformés ces derniers sont multipliés et stockés. A ce stade la banque d'ADN est nommée banque non-ordonnée car les clones possédant des séquences d'ADN successives au sein du génome ne sont pas connus (figure 10).

Plusieurs techniques peuvent être utilisées afin d'ordonner les clones tel que l'hybridation de sonde, la digestion partielle ou par séquençage des extrémités de chaque insert.

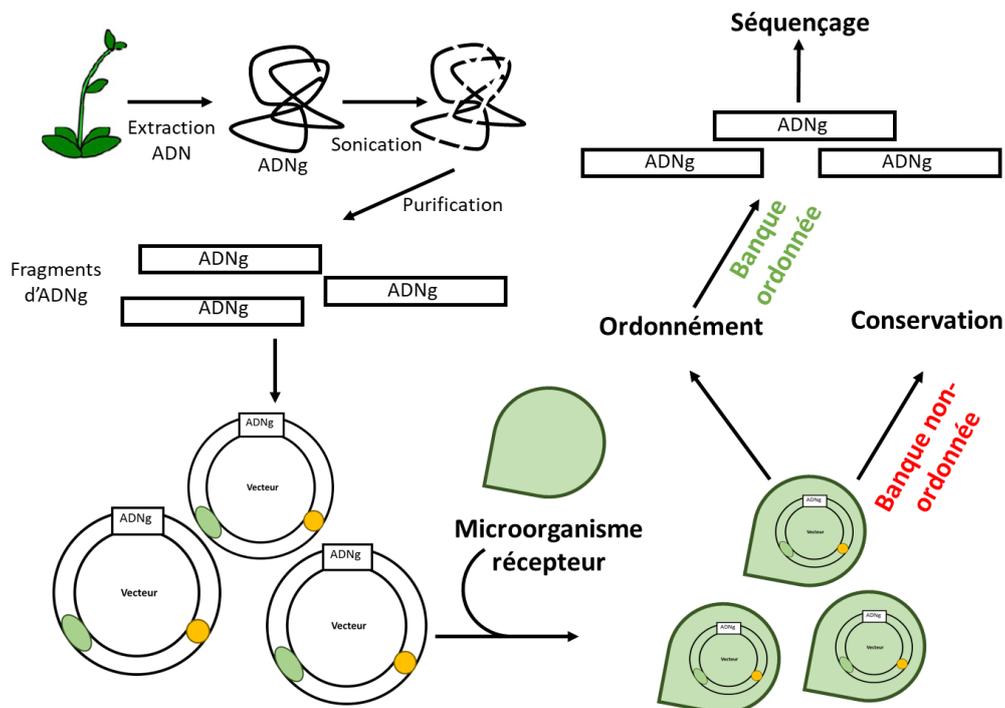


Figure 10. Les étapes relatives de la création des banques. L'ADN de l'organisme étudié est extrait puis fragmenter et les fragments sont purifiés. Ces derniers sont clonés dans des vecteurs qui sont introduits dans un microorganisme. Les individus transformés sont sélectionnés puis conservés ce qui permet l'obtention d'une banque d'ADN non-ordonnée. L'ordonnement des clones consiste à connaître ceux qui possèdent les fragments d'ADNs successifs dans le génome.

Partie I : Cours

Chapitre VI : clonage moléculaire

2.2. Les banques d'ADNc

Les banques d'ADNc sont des banques produites à partir d'ADN complémentaire issus de la reverse transcription des ARNm. Elle vise à stocker au sein d'une banque l'ensemble des transcrits produit dans un organisme ou un tissu dans une condition donnée et reflète les partie du génome qui sont exprimés.

Après conversion des ARNm en ADN complémentaires, ces derniers sont clonés dans des vecteurs et introduits au sein de microorganismes qui sont multipliés et conservés.

2.2.1. Le passage de l'ARN à l'ADN

La transformation des ARNm en ADN complémentaire passe par l'action d'une reverse transcriptase qui est une ADN polymérase synthétisant de l'ADN à partir de l'ARN. En fonction du type d'ARNm à reverse transcrire eucaryotique ou procaryotique, des oligonucléotides polythymine (pour les eucaryotes) ou des hexamères aléatoires (pour les procaryotes) sont respectivement utilisés.

Après hybridation des oligonucléotides la reverse transcriptase se fixe sur le duplex ARN-ADN et procède à la biosynthèse d'ADN à partir de l'ARNm. Par la suite l'ARN est éliminé sous l'action d'une ARNase ou de la soude. L'ADNc simple brin est ensuite transformés en double brin sous l'action d'une ADN polymérase. Des sites de restrictions peuvent être introduits aux extrémités des ADNc par la ligation d'adaptateurs (séquence d'ADN possédant des sites de restrictions).

2.2.2. Le choix du vecteur

Lors de la création d'une banque d'ADN (génomique ou d'ADNc) le choix du vecteur est crucial, pour la création d'une banque d'ADNg les vecteurs supportant de grands fragments d'ADNs sont privilégiés tel que le chromosome artificiel de levure.

Des vecteurs d'expressions caractérisées par la présence d'un promoteur en amont du site de clonage sont utilisés pour la création de banque d'ADNc, cela permet entre autre de produire des protéines à partir des ADNc.

Partie I : Cours
Chapitre VI : clonage moléculaire

Références bibliographiques

- (1) Brown, T. A. 2016. *Gene Cloning and DNA Analysis: An Introduction, 7th Edition*. 7th Edition. ed. Wiley-Blackwell.
- (2) Chen, Bing-Yuan, and Harry W Janes. 2002. *PCR Cloning Protocols*. Second Edition. Humana Press.
- (3) Green, Michael R., and Joseph Sambrook. 2012. *Molecular Cloning: A Laboratory Manual*. Fourth Edition. ed. U.S Cold Spring Harbor Laboratory Press.
- (4) Primrose, Sandy B., and Richard Twyman. 2006. *Principles of Gene Manipulation and Genomics*. First Edition. ed. Wiley-Blackwell.
- (5) Sambrook, Joseph F., and David Russell. 2006. *The Condensed Protocols from Molecular Cloning: A Laboratory Manual*. First Edition. ed. U.S Cold Spring Harbor Laboratory Press.
- (6) Tagu, Denis, and C. Moussard. 2006. *Techniques for Molecular Biology*. First Edition. ed. CRC Press.

Partie I : Cours

Chapitre V : vecteurs de clonage

1. Généralités sur les vecteurs

1.1. Concept de vecteur

Les vecteurs sont des molécules d'ADN extrachromosomiques qui sont utilisées afin de stocker et de multiplier un fragment d'ADN. L'ADN à manipuler est intégré au sein du vecteur qui est lui-même incorporé au sein d'un microorganisme compatible.

Il existe plusieurs types de vecteurs, on peut citer les vecteurs de clonage qui sont utilisés pour le stockage et la multiplication de fragments d'ADN et les vecteurs d'expression qui permettent d'exprimer un gène au sein du microorganisme.

1.2. Propriétés que doit posséder un vecteur

Il existe une diversité de vecteurs associée à une diversité de microorganismes, mais l'ensemble des vecteurs doivent présenter les caractéristiques suivantes :

- i) Une séquence lui permettant de se multiplier au sein du microorganisme.
- ii) Un site de multicolonnage au sein duquel l'insert va être intégrée.
- iii) Un gène de sélection doit être présent afin de sélectionner les individus ayant acquis le vecteur.

1.3. Principes généraux d'utilisation d'un vecteur

Les principaux vecteurs de clonages sont les plasmides bactériens, les vecteurs phagiques (qui sont des virus infectant les bactéries) ainsi que les chromosomes artificiels de levure et de bactérie, YAC et BAC respectivement.

Le choix du vecteur dépend du type d'approche choisie lors de l'étude, un vecteur plasmidique sera préférentiellement utilisé pour le séquençage de petits fragments d'ADN (jusqu'à 2 kilo bases), alors que des vecteurs phagiques, YAC (Yeast Artificial Chromosome) ou BAC (Bacterial Artificial Chromosome) sont utilisés pour réaliser des banques d'ADN génomiques (voir tableau 1).

Partie I : Cours

Chapitre V : vecteurs de clonage

Tableau 1. Principaux vecteurs et leurs caractéristiques

Vecteur	Taille de l'insert	Exemple	Caractéristiques
Plasmide	<20 kb	puc19	Se réplique indépendamment du chromosome, plusieurs copies peuvent être contenues dans une cellule, division cellulaire rapide
Bactériophage	9-25 kb	λ1059	Empaqueté dans un phage modifié afin de générer des doubles ou simples brin d'ADN au sein de l'hôte
Cosmide	30-47 kb	pJC720	Peut-être empaqueté dans un phage et répliqué dans une bactérie comme un plasmide
BAC	75-300 kb	pBAC108L	F plasmide modifié, très stable dans une cellule
YAC	100-1,000 kb	pYAC	Replicable dans <i>S. cerevisiae</i>

2. Les plasmides

2.1. L'utilisation d'un plasmide

Le plasmide est un matériel génétique extrachromosomique sous forme d'une molécule d'ADN circulaire de 3 à 10 kb, capable de se répliquer dans une bactérie. Naturellement il contient des gènes impliqués dans l'adaptation des bactéries à l'environnement telle que les gènes de pathogénicité.

En génie génétique, le plasmide a été modifié pour devenir un plasmide générique, ce dernier est utilisé comme vecteur d'inserts dont la taille est inférieure à 20kb.

Un vecteur plasmidique générique contient une origine de réplication qui permet à l'ADN de se répliquer. Un site de multicolonnage au sein duquel l'insert est introduit et un gène de résistance à un antibiotique qui permet de sélectionner les bactéries ayant intégrées le vecteur (figure 11).

Partie I : Cours

Chapitre V : vecteurs de clonage

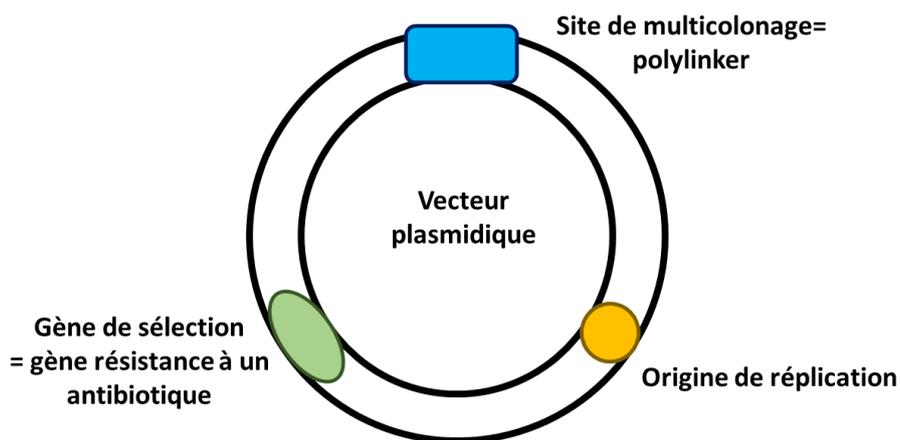


Figure 11. Prototype d'un vecteur plasmidique. Un plasmide doit posséder au moins un site de multicolonnage pour l'introduction de l'insert, une origine de réplication pour pouvoir se multiplier dans la bactérie et un gène de résistance à un antibiotique afin de sélectionner les bactéries ayant acquis le vecteur après la transformation.

2.2. Processus globale de clonage par plasmide

Généralement la procédure suivie au cours d'un clonage en système bactérien est la suivante : après insertion de l'insert, le plasmide est introduit dans une cellule bactérienne receveuse par technique de transformation (figure 12).

Les bactéries transformées sont sélectionnées sur un milieu solide comportant un antibiotique pour lequel le plasmide comporte un gène de résistance. Les clones résistants sont isolés et cultivés en milieu liquide en présence de l'antibiotique afin de maintenir la pression de sélection sur les clones.

Après une période d'incubation à température optimale pour la multiplication bactérienne, une suspension cellulaire est obtenue, une partie est utilisée pour la réalisation d'un aliquote stocké à de basse température (à -80°C ou -20°C) afin de sauvegarde le clone.

Les copies d'ADN plasmidiques restantes peuvent être extrait par différents procédés, le plus utilisé est l'extraction sur clone ou mini prep, au cours de laquelle les bactéries sont lysées et le plasmide est extrait, puis purifié.

Partie I : Cours

Chapitre V : vecteurs de clonage

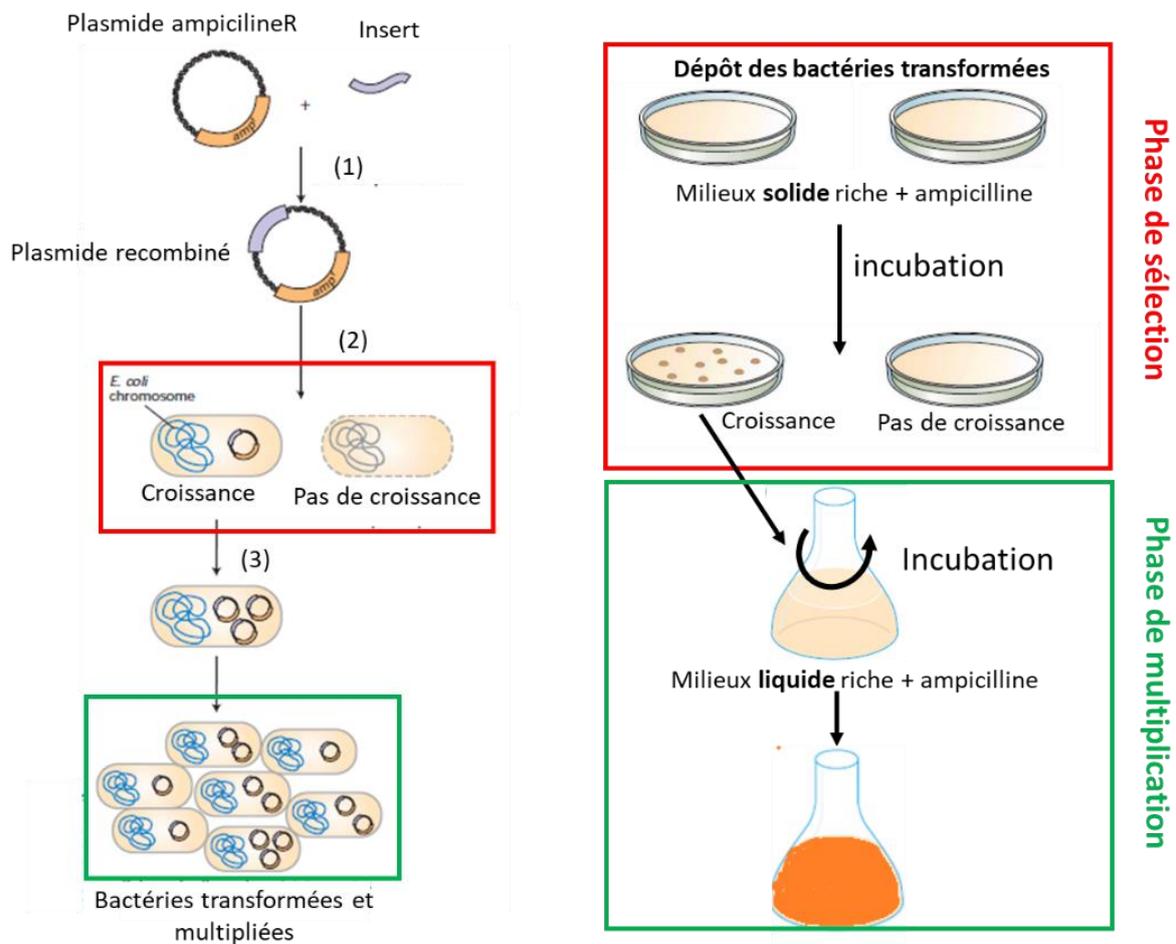


Figure 12. Les différentes étapes du clonage plasmidique. Après transformation des bactéries par le plasmide recombiné, une phase de sélection à lieu sur un milieu solide contenant un antibiotique dont la résistance est porté sur le plasmide (exemple : ampicilline). Les clones obtenus sont multipliés dans un milieu liquide contenant l'antibiotique afin de maintenir le plasmide dans la cellule hôte.

2.2. Préparation des plasmides pour le clonage

La préparation du plasmide consistera essentiellement à introduire l'insert au sein du site de multicolonnage, pour cela plusieurs approches sont disponibles.

2.2.1. Clonage par restriction ligation

L'étape d'introduction de l'insert au sein du vecteur est sensible dans le processus de clonage, plusieurs approches existent actuellement, certaines sont standardisées et commercialisées de telle sorte que le processus d'insertion devient simple et peu couteux.

Partie I : Cours

Chapitre V : vecteurs de clonage

Parmi les techniques classiques et les plus couramment utilisées le clonage par restriction et ligation. Il consiste à réaliser une double digestion de l'insert et du plasmide avec deux enzymes de restrictions libérant des bouts cohésifs (figure 13). L'insert et le plasmide digérés sont mis ensemble, par complémentarité les deux molécules s'hybrident et reformes le plasmide recombiné, enfin une ligase est utilisée afin de souder l'insert et le plasmide.

Cette méthode permet d'introduire l'insert de manière orienté (clonage directionnelle) et sans que le vecteur se re-circularise, ce qui peut conduire à la production d'un vecteur vide.

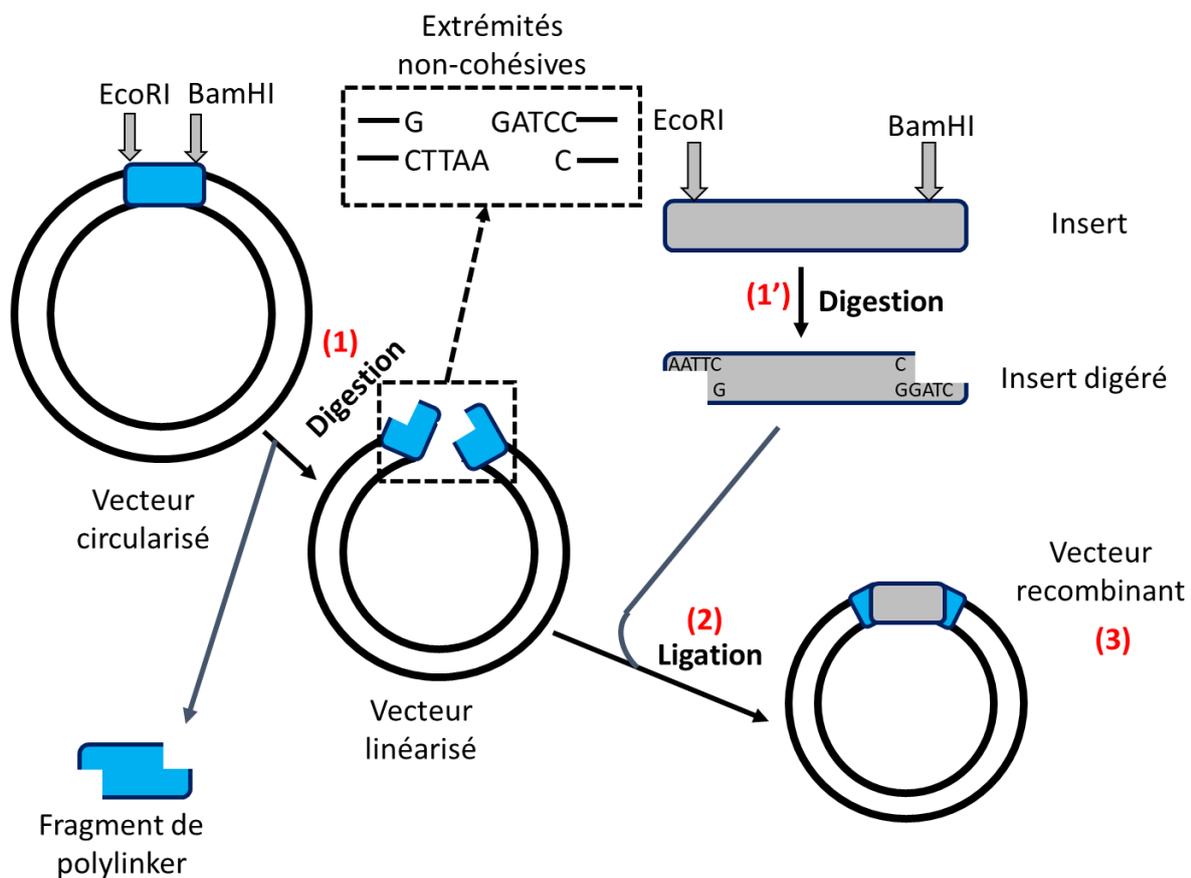


Figure 13. Le clonage par restriction-ligation. Le plasmide et l'insert sont digérés par deux enzymes de restrictions coupant au site de multicolonnage du vecteur et de part et d'autre de l'insert. Au cours de la ligation l'insert se fixe par complémentarité au plasmide et une ligase soude les deux molécules.

L'un des problèmes critiques rencontré lors d'un clonage est la création de site de restriction au niveau de l'insert. En effet les enzymes choisies doivent avoir la capacité de couper les

Partie I : Cours

Chapitre V : vecteurs de clonage

extrémités de l'insert sans altérer la séquence à cloner et sans produire un décalage de sa séquence qui pourrait conduire à la production d'une protéine aberrante.

Plusieurs approches peuvent être utilisées afin d'éviter le problème liée au choix du site de restriction, l'une des plus courante consiste à amplifier de manière spécifique via la technique de PCR l'insert avec des amorces présentant des extras bouts à l'extrémité 5' qui lors de l'amplification génère des sites de restriction adéquats. A la fin de la réaction le fragment d'ADN amplifié présent l'insert entouré par les sites de restriction (figure 14).

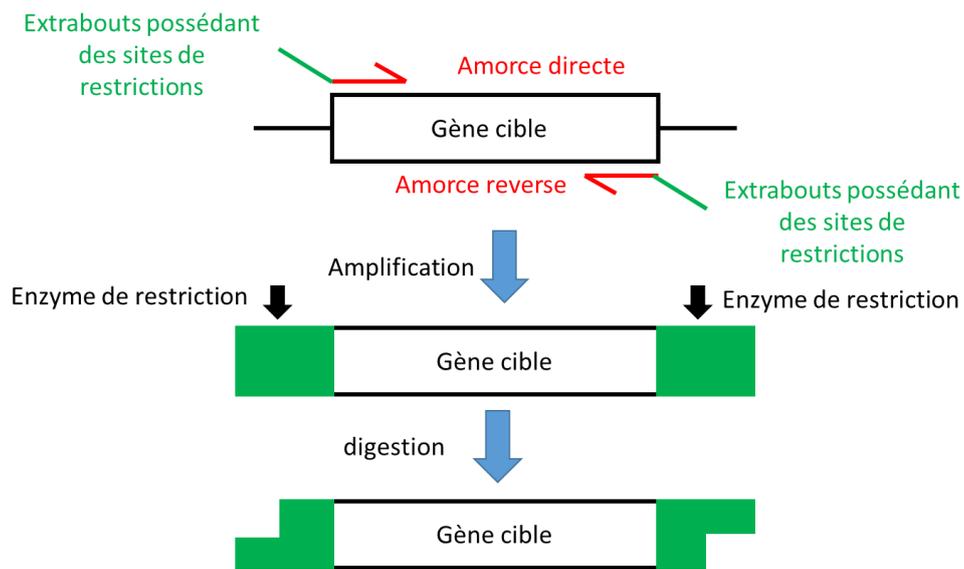


Figure 14. Génération de site de restriction par technique PCR. Une PCR est réalisée avec des amorces contenant des extras bouts afin d'amplifier l'insert. Les amorces s'hybrident aux extrémités de la séquence et lors de l'amplification elles génèrent un fragment d'ADN contenant l'insert bordé par les sites de restriction.

2.2.1. Crible de clone

Dans certains cas et pour des raisons diverses, les clones issus de la transformation par le vecteur recombiné contiennent un vecteur dépourvu d'insert, on parle dans ce cas de clones négatifs ou vides.

Le crible de clones représente un ensemble de techniques utilisées afin de sélectionner les clones comportant le vecteur contenant l'insert et d'éviter ainsi les clones comportant un vecteur vide.

Partie I : Cours

Chapitre V : vecteurs de clonage

Le crible de clones bleu-blanc est une technique qui se base sur la présence ou l'absence de l'activité β -galactosidase (où β -gal). Cette enzyme catalyse dans la nature la dégradation du lactose en glucose et galactose. Quand le lactose est remplacé par du X-gal, un analogue incolore, il est transformé par l'enzyme en 5-hydroxy-4-chloro-3-indole qui au contact de l'oxygène se transforme en 5,5'-Dibromo-4,4'-Dichloro indigo dont la couleur est bleu.

Le vecteur utilisé pour réaliser un crible bleu-blanc contient au sein du site de multiclonnage le gène codant la β -gal. Au cours du clonage l'insert est introduit au niveau du gène, après transformation les clones positifs contenant l'insert n'expriment pas l'enzyme, alors que les clones négatifs produisent cette protéine.

La discrimination entre les clones positifs et négatifs s'effectue sur milieux solide contenant du X-Gal, les clones positifs présentent une couleur blanche et les clones négatifs présentent une couleur bleu (figure 15).

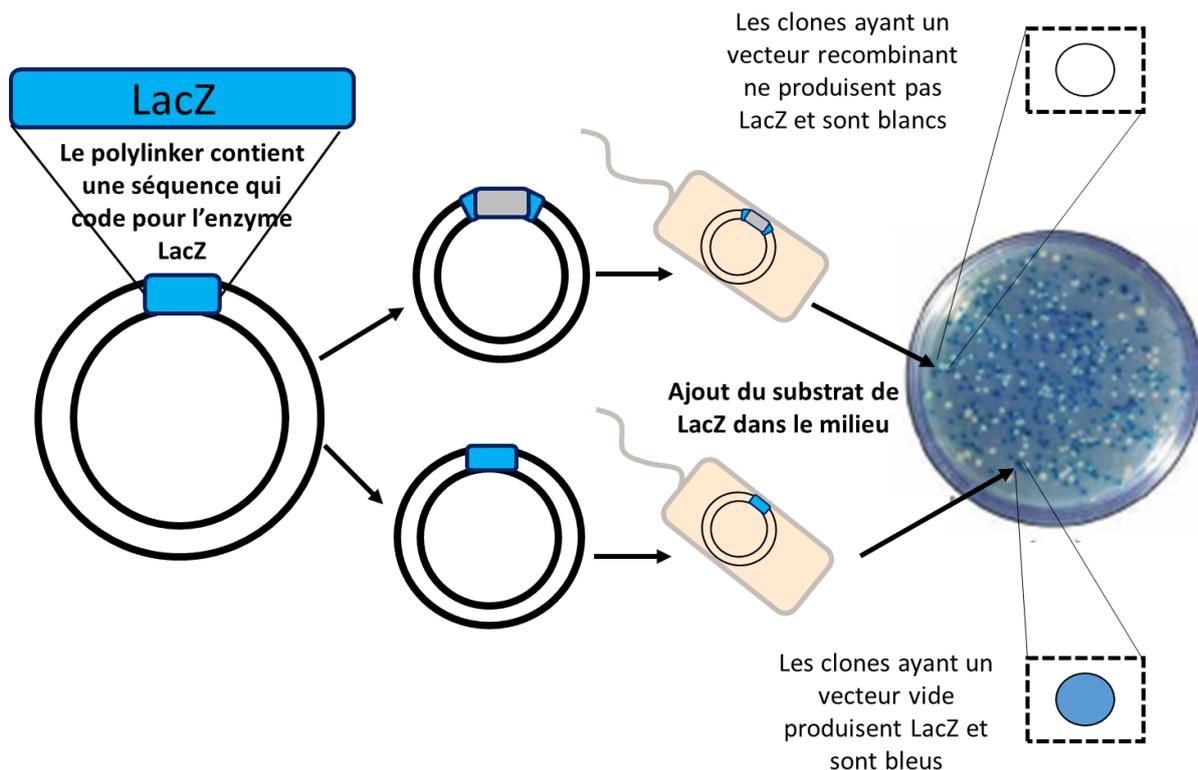


Figure 15. Principe de crible bleu-blanc. Le vecteur contient le gène LacZ codant la β -galactosidase au niveau du site de multiclonnage. En fonction de l'introduction ou pas de l'insert au sein du site de multiclonnage, les bactéries transformées produisent une couleur blanche ou bleu respectivement au contact du X-gal.

Partie I : Cours
Chapitre V : vecteurs de clonage

2.2.3. Le clonage par recombinaison

Le clonage par recombinaison appelé aussi système GET WAY consiste introduire par recombinaison homologue l'insert au sein du site de multicolonnage sous l'action d'une recombinase. La recombinase procède lors du clonage à l'introduction de l'insert par recombinaison homologue au niveau du site de multicolonnage qui contient une cassette toxique (gène toxique), lors de la recombinaison la cassette toxique est retirée. La sélection des clones positive est basé sur l'absence de la cassette toxique, les clones négatifs sont ainsi incapables de survivre (figure 16).

Partie I : Cours
Chapitre V : vecteurs de clonage

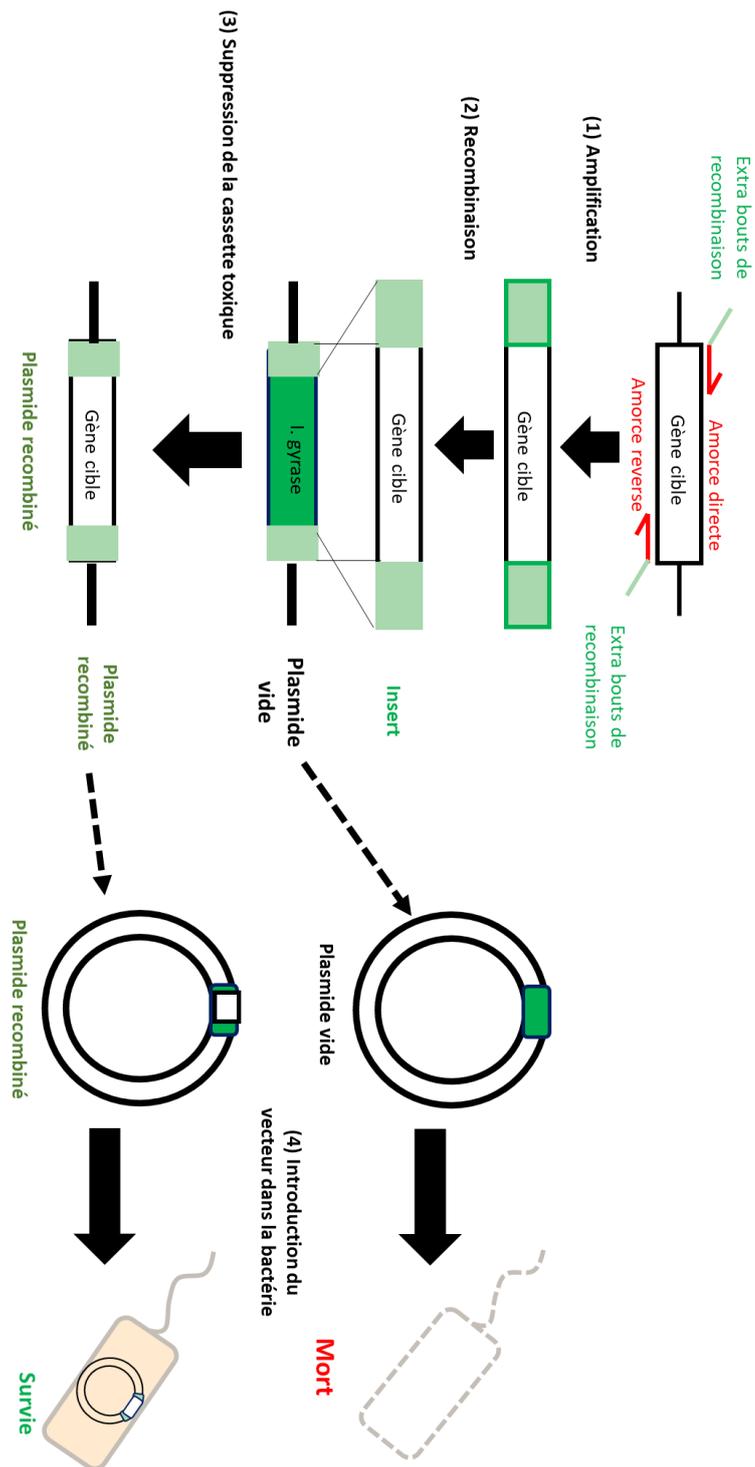


Figure 16. Principe du clonage par recombinaison. Le vecteur Get WAY contient deux séquences de recombinaison entourant une cassette toxique au niveau du site de multicolonnage. L'insert est introduit par recombinaison homologue, après transformation seulement les clones positifs (contenant l'insert) survivent.

Partie I : Cours

Chapitre V : vecteurs de clonage

2.3. Les différents types de plasmides

2.3.1 Les plasmide de première génération

Ce sont les premiers plasmides utilisés, ils sont à l'état naturel et non-modifiés tel que les plasmides ColE1, RSF2124 et pSC 101.

2.3.2. Les plasmides de seconde génération

Ce sont des plasmides dérivés de ceux de premier génération, mais ayant subis des modifications afin de les améliorer, notamment par : i) l'ajout de gène de résistance à différents antibiotiques, ceci afin d'améliorer le processus de sélection des clones. ii) Par l'ajout d'un site de multicolonnage possédant des sites de restrictions pour plusieurs enzymes dans le but de faciliter l'intégration de l'insert.

2.3.3. Les plasmides de troisième génération

Ce sont des plasmides de deuxième générations améliorés pour éviter le passage par une phase de sous clonage, l'insert peut directement être intégré dans le plasmide d'intérêt.

3. Les phages

3.1. Utilisation des phages

Les phages (ou bactériophages) sont des virus capables d'infecter les bactéries et s'insèrent au sein de leur chromosome afin de se multiplier, en fonction du cycle infectieux on distingue deux type de phages ;

i) Les phages lysogènes, qui après pénétration dans la cellule entre dans une phase de latence. Au cours de cette période le phage ne se réplique qu'au cours de la réplication du chromosome. Sous certaines conditions telles qu'en cas de stress infligé à la bactérie, le phage peut enclencher une phase de multiplication conduisant à la lyse cellulaire.

ii) Les phages lytiques quant à eux se répliquent de manière importante dès la pénétration aboutissant à la destruction de la bactérie.

Certains phages utilisent l'ADN comme source de l'information génétique, dans ce cas la molécule dite ADN phagique correspond aux gènes impliqués dans les différentes phases du cycle.

Partie I : Cours

Chapitre V : vecteurs de clonage

3.2. Préparation d'un phage

La préparation du phage comprend le clonage qui consiste à introduire l'insert au sein de l'ADN d'un phage lytique. L'ADN recombiné est introduit dans un phage recréé artificiellement. Ce dernier est multiplié au sein d'une bactérie receveuse, puis stocké à de basse température (-20°C ou -80°C) (figure 17).

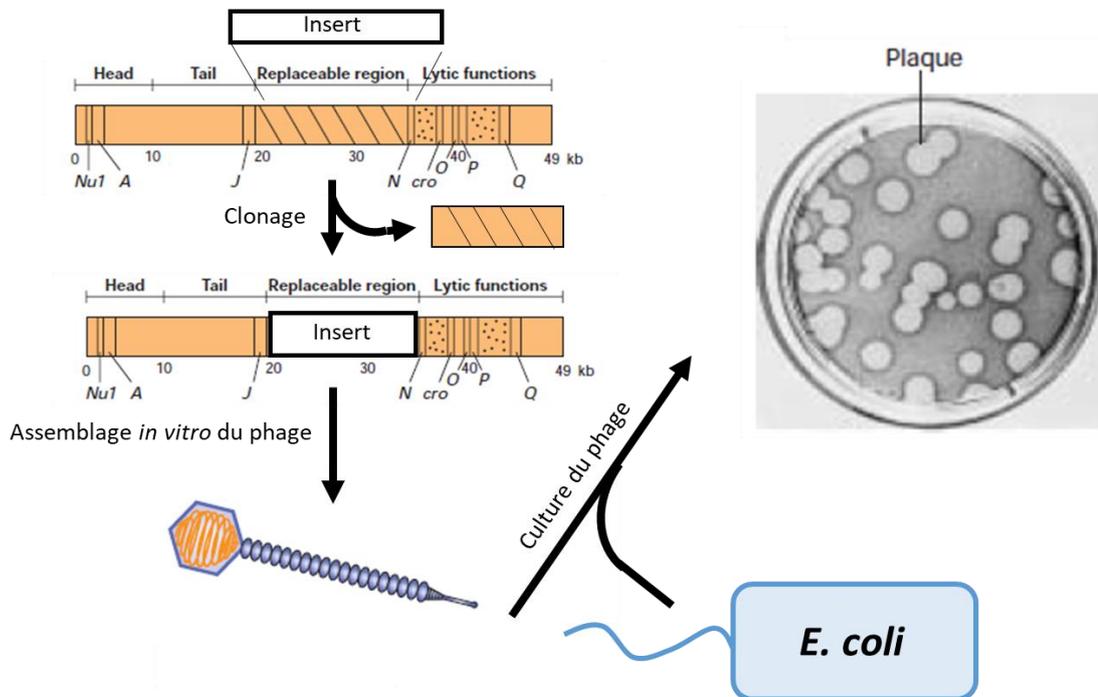


Figure 17. Les étapes de clonage dans un phage. L'insert est introduit au niveau du génome du phage qui est artificiellement reconstruit. Le virus est multiplié par application sur un tapis de bactérie receveuses. Après incubation des plaques apparaissent au sein du tapis, elles correspondent aux zones au sein desquelles le phage s'est multiplié et les bactéries ont été lysées.

3.3. Les différents phages utilisés en biologie moléculaire

3.3.1. Les phage de première génération. : Le phage λ

Le phage λ est un bacériophage capable d'infecter la bactérie *Escherichia coli*. De cycle tempéré il peut réaliser en fonction des conditions un cycle lytique aboutissant à la production de milliers de copies du virus et à la mort de la bactérie, ou un cycle lysogène au cours duquel il intègre le génome bactérien sans provoquer la mort.

Partie I : Cours

Chapitre V : vecteurs de clonage

Possédant un ADN double brin de 48502 pb de forme linéaire, il se circularise au sein de la bactérie. Le génome du phage est composé des gènes nécessaires à son maintien et à sa réplication au sein d'une cellule hôte. Les séquences d'ADN à insérer de taille allant de 40 à 50 kb peuvent être introduites entre les séquences Cos. Le phage λ est largement utilisé en génie génétique notamment pour la création de banque d'ADN.

3.3.2. Les phages de seconde génération

Les phages de seconde génération sont des virus dérivés de phage naturel tel que le phage λ . Ils sont modifiés afin d'améliorer les propriétés du phage, par exemple dans le cas du phage Embl 3 & 4 des sites de restriction sont introduits entre les séquences Cos afin de permettre un clonage directionnel de l'insert.

4. Les autres types de vecteur

4.1. Les cosmides

Un cosmide est un vecteur contenant en plus du site de multicolonnage, des gènes lui permettant d'être multiplié dans une bactérie (origine de réplication, gène de sélection), à cela s'ajoute les gènes Cos permettant l'encapsidation dans un phage λ .

Le cosmide est encapsidé dans le phage, puis après infection de la bactérie par le virus, le vecteur se multiplie tel un plasmide.

4.2. Les vecteurs «navettes»

Les vecteurs navettes sont des vecteurs possédant les gènes nécessaires pouvant leur permettre d'être utilisés dans deux microorganismes différents. Le cosmide est un bon exemple de vecteur navette.

4.3. Les vecteurs viraux eucaryotes

Les vecteurs viraux eucaryotes sont largement utilisés en génie génétique, ils permettent notamment d'exprimer des protéines au sein de cellules eucaryotes. Par exemple le baculovirus se multiplie au sein des cellules d'insecte, lors de son cycle infectieux, une de ces protéines, la polyhidrine est produite au sein de la cellule infectée pour protéger le virus contre la dégradation. En génie génétique le gène de la polyhidrine est remplacé avec celui codant la protéine d'intérêt.

Partie I : Cours
Chapitre V : vecteurs de clonage

Références bibliographiques

- 1- Casali, Nicola, and Andrew Preston. 2003. *E. Coli Plasmid Vectors*. First Edit. Humana Press.
- 2- Haase, Santiago, Leticia Ferrelli, Matías Luis Pidre, and Víctor Romanowski. 2013. "Genetic Engineering of Baculoviruses." In *Current Issues in Molecular Virology - Viral Genetics and Biotechnological Applications*, ed. Victor Romanowski. IntechOpen.
- 3- Lodish, Harvey et al. 2000. *Molecular Cell Biology*. 4th editio. Connected, Media.
- 4- Rapley, Ralph, and John M. Walker. 1998. *Molecular Biomethods Handbook*. First Edit. Humana Press.
- 5- Rodriguez, Raymond L., and David T. Denhardt. 1988. *Vectors A Survey of Molecular Cloning Vectors and Their Uses*. First Edit. ed. Elsevier. <https://doi.org/10.1016/C2013-0-01066-3>.

Partie I : Cours

Chapitre VI : sondes moléculaires

1. Le concept de sonde

Les sondes sont des oligonucléotides synthétisés chimiquement et marqués. De taille variable en fonction de leur application, ils sont employés dans de nombreuses techniques telles que l'hybridation *in situ*, le northern ou southern blot.

2. Les agents de marquages

On peut distinguer deux grands types de marquages, celui dit chaud obtenu par l'ajout d'un isotope radioactif et le marquage froid obtenu par l'ajout d'une molécule non-radioactive telle qu'un fluorochrome.

2.1 Les isotopes radioactifs

L'isotope le plus largement utilisé pour le marquage chaud est le phosphate radio marqué ou P32. Il est intégré au sein de l'oligonucléotide lors de la phase biosynthèse et sa révélation est réalisée par autoradiographie.

2.2. Marquage non-radioactif

Consiste à l'ajout d'une molécule non-radioactive telle que la biotine ou un fluorochrome. L'ajout peut se faire lors de la biosynthèse (par utilisation de nucléotides marquées) ou après l'étape de biosynthèse.

3. Quelques stratégies de marquage

3.1. La « Nick translation »

Lors de ce processus, l'ADN à marquer est traité avec une désoxyribonucléase (DNAse I) qui ouvre des brèches au sein du double brin. Suite à cela une ADN polymérase I synthétise un nouveau brin à partir de l'extrémité 3' OH, alors que l'activité 5' → 3' exonuléasique de l'enzyme dégrade la partie 5' phosphate de la même brèche.

L'enzyme procède à la néo synthèse du fragment au site où a été générée la brèche. Lors de la biosynthèse des nucléotides radioactifs ou fluorescents peuvent être ajoutés afin de marquer l'ADN synthétisé.

Partie I : Cours

Chapitre VI : sondes moléculaires

3.2. La « Random printing »

L'ADN à marquer est dénaturé thermiquement afin de séparer le double brin, suite à cela des amorces aléatoires (qui correspondent à toutes les associations possibles de nucléotides) sont ajoutés et vont s'hybrider de manière aléatoire. Les amorces permettent à l'ADN polymérase I de se fixer sur l'ADN et de biosynthétiser la séquence complémentaire en intégrant des nucléotides marqués.

3.3. Le marquage des sondes synthétiques (Oligo-nucléotides de synthèse)

L'ADN est synthétisé par réaction chimique avec des nucléotides marqués, le premier nucléotide de la séquence est fixé sur un support, puis les nucléotides à intégrer sont ajoutés progressivement de manière unique (un seul nucléotide à chaque cycle d'intégration) et dans l'ordre déterminé jusqu'à obtenir la séquence entière. La réaction de ligation entre les nucléotides est réalisée chimiquement.

3.4. Les sondes ARN (ribosondes)

Ce sont des sondes ARN pouvant être radio marquées utilisées dans la cartographie des ARNs ainsi que l'hybridation *in situ*. Ce type de sonde permet un meilleur rendement d'hybridation car les associations ARN/ADN, ARN/ARN sont moins stables que les associations ADN/ADN.

Partie I : Cours
Chapitre VI : sondes moléculaires

Références bibliographiques

- 1- Johnson, Iain, and Michelle T.Z. Spence. 2010. *Molecular Probes Handbook, A Guide to Fluorescent Probes and Labeling Technologies*. First Edit. Invitrogen.
- 2- Metfies, Katja, and Linda Medlin. 2005. "Ribosomal RNA Probes and Microarrays: Their Potential Use in Assessing Microbial Biodiversity." In *Methods in Enzymology*, Elsevier, 258–78. [https://doi.org/10.1016/S0076-6879\(05\)95016-7](https://doi.org/10.1016/S0076-6879(05)95016-7).
- 3- Osada, Hiroyuki. 2000. *Bioprobes: Biochemical Tools for Investigating Cell Function*. First Edit. Springer.
- 4- Symons, Robert H. 1989. *Nucleic Acid Probes*. CRC Press.
- 5- Vizzini, Priya, Lucilla Iacumin, Giuseppe Comi, and Marisa Manzano. 2017. "Development and Application of DNA Molecular Probes." *AIMS Bioengineering* 4(1): 113–32.

Partie I : Cours

Chapitre VIII : la transformation génétique

1. Principe de la transformation

La transformation en méthodologie moléculaire consiste à introduire une molécule d'ADN exogène ou endogène modifiée au sein d'une cellule eucaryote, d'une cellule procaryote ou même d'un tissu ou d'un organisme (figure 18).

Dans le cas d'un microorganisme ou d'un organisme la transformation aboutie à la création d'un organisme génétiquement modifié (OGM).

Il existe plusieurs approches afin de réaliser une transformation, la technique employée dépend du type de cellule transformée.

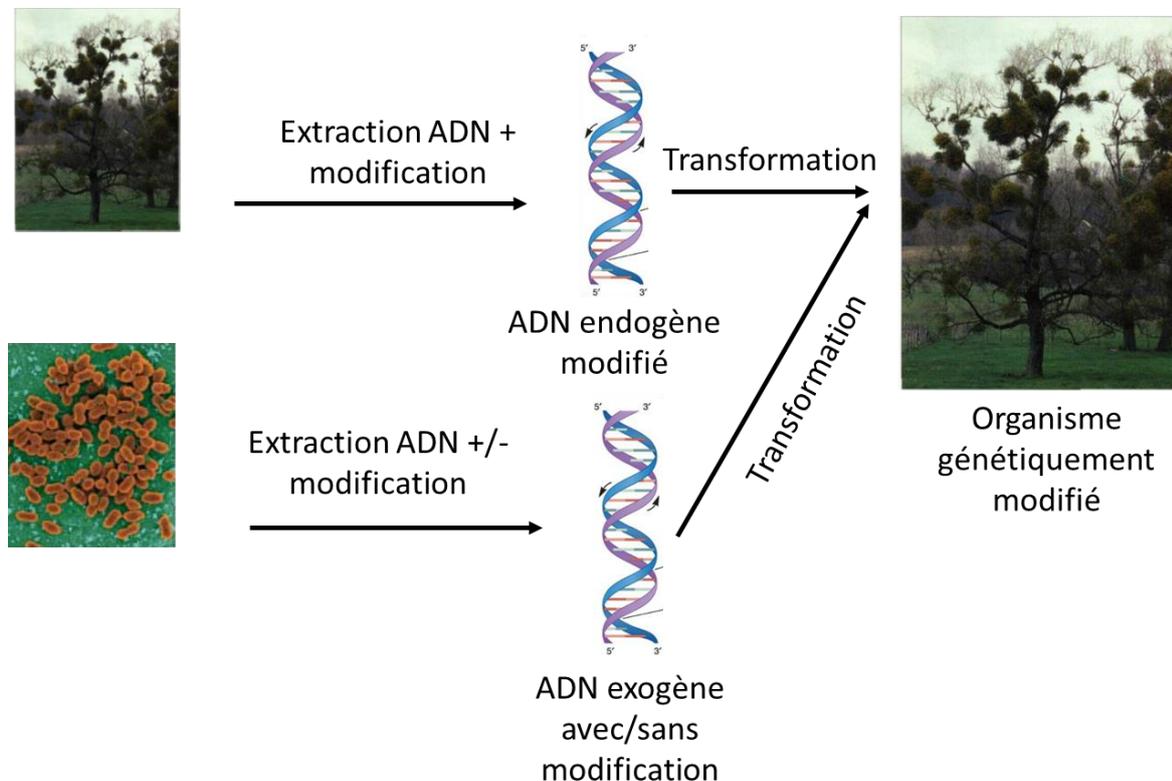


Figure 18. Principe de la transformation. L'ADN issu d'un organisme étranger ou de l'organisme étudié est modifié puis introduit par technique de transformation. Cela abouti à la création d'un OGM.

2. Mécanisme naturel de transfère d'ADN

Dans la nature le phénomène de transformation peut avoir lieu, par exemple chez les bactéries, ceci abouti à une acquisition d'une molécule d'ADN pouvant contenir de nouveaux

Partie I : Cours

Chapitre VIII : la transformation génétique

gènes pour l'organisme transformé. Ce mécanisme biologique a sans doute participé aux phénomènes évolutifs et micro évolutif qui ont permis l'émergence de nouvelles formes de vie sur terre.

3. Les techniques de transformation

Il existe plusieurs techniques permettant la transfère d'un ADN dans une cellule ou un organisme.

3.1. Transformation par choc physique

La transformation par choque physique est couramment utilisée en recherche afin de transformer des cellules procaryotes, telle qu'au cours de l'introduction d'un plasmide au sein d'une bactérie. Le choc thermique et le choc électrique sont deux types de transformation par choc physique.

3.1.1. Transformation par choc thermique

La transformation par choc thermique peut être utilisée sur des bactéries thermocompétantes (qui sont capables de supporter les variations de températures exercées). Les bactéries sont incubées avec l'ADN en présence de CaCl_2 quelque minute à faible température (4°C), puis transférées rapidement à une température élevée (42°C) ce qui aboutit à la formation de nano pores par le CaCl_2 dans l'enveloppe bactérienne. L'ADN travers la cellule via les pores formés (figure 19).

Partie I : Cours

Chapitre VIII : la transformation génétique

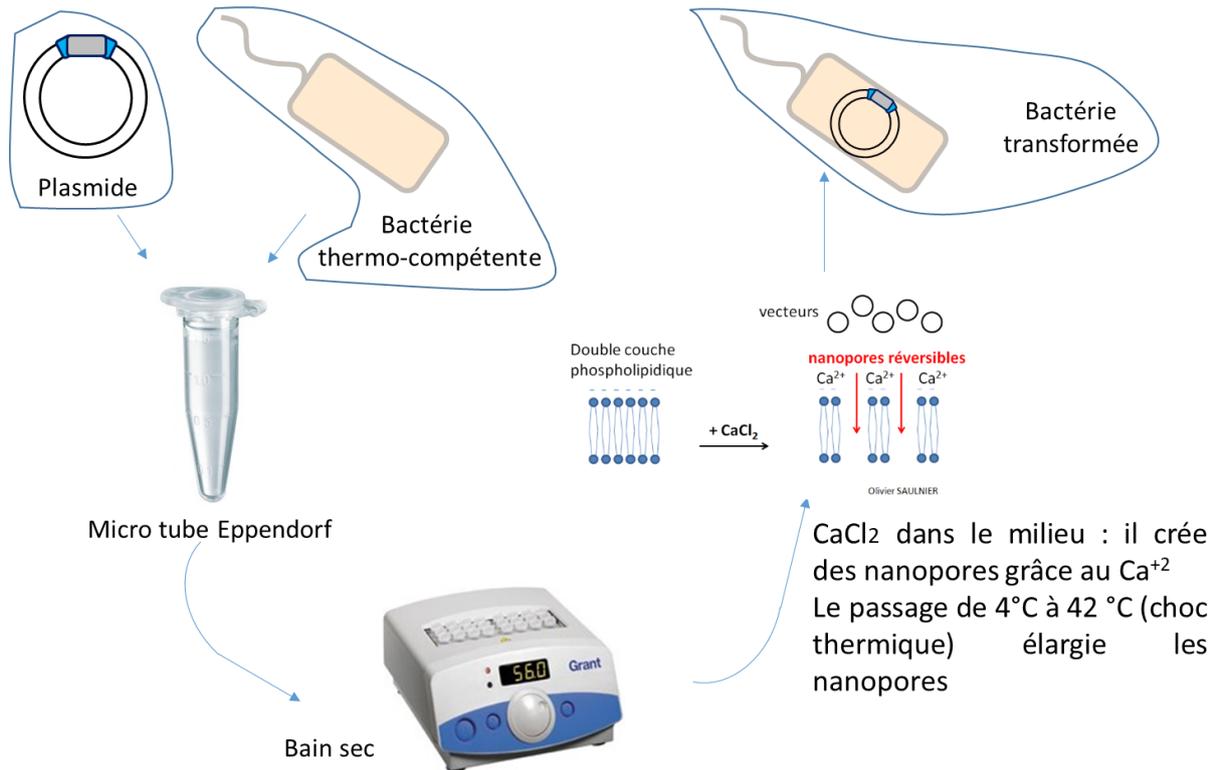


Figure 19. La transformation par choc thermique. Le plasmide et les bactéries sont soumis à un choc thermique en présence de CaCl₂. Cela provoque des pores au niveau des membranes par lesquels le plasmide pénètre dans la cellule.

3.1.2. Transformation par choc électrique

La transformation par choc électrique ou électroporation consiste à mettre les cellules électrocompétantes (incluses dans un milieu dépourvue de sels) avec l'ADN dans une cuve d'électroporation. Suite à cela la cuve est placée dans un générateur électrique (électroportateur) dans lequel la suspension est soumise à un choc électrique qui provoque la formation de pores au sein de l'enveloppe bactérienne par lesquels passe l'ADN (figure 20).

Partie I : Cours

Chapitre VIII : la transformation génétique

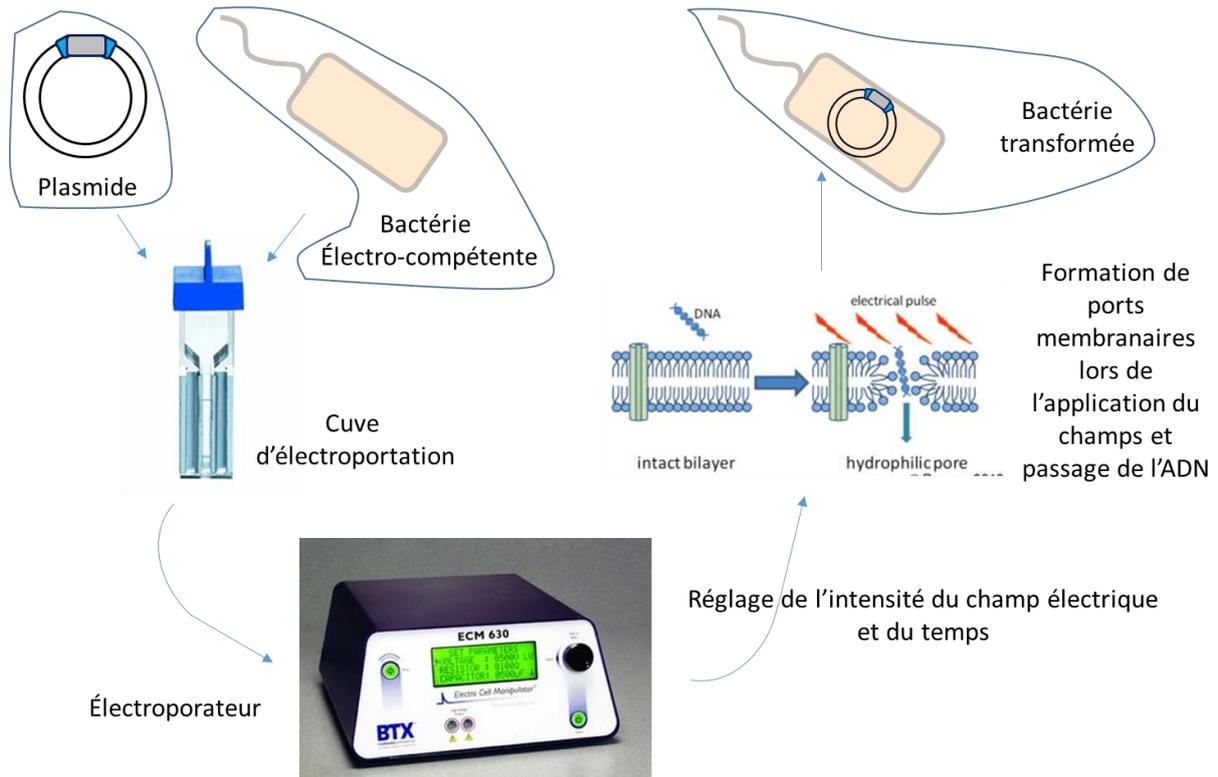


Figure 20. Principe de l'électroporation. Les cellules et l'ADN à transférer sont soumis à un choc électrique qui provoque la formation de pores par lesquels l'ADN traverse.

3.2. Transformation chimique

La transformation chimique consiste à transférer de l'ADN dans une cellule par l'ajout d'agents chimique. La transformation des levures par lithium acétate est un exemple de transformation chimique.

Elle consiste à incuber des levures en présence de lithium acétate, de Polyéthylène glycol (PEG) ainsi que d'un ADN entraîneur (exemple ADN de sperme de saumon) ,en plus de l'ADN à transférer.

La suspension est incubée à 30°C, puis transférée à 42°C et enfin 30°C. Le lithium acétate sous l'effet du changement de températures provoque la formation de pores au sein de l'enveloppe, le PEG permet de lier l'ADN aux parois de la cellule et de générer un choque osmotique permettant la pénétration de l'ADN par l'ADN entraîneur (figure 21).

Partie I : Cours

Chapitre VIII : la transformation génétique

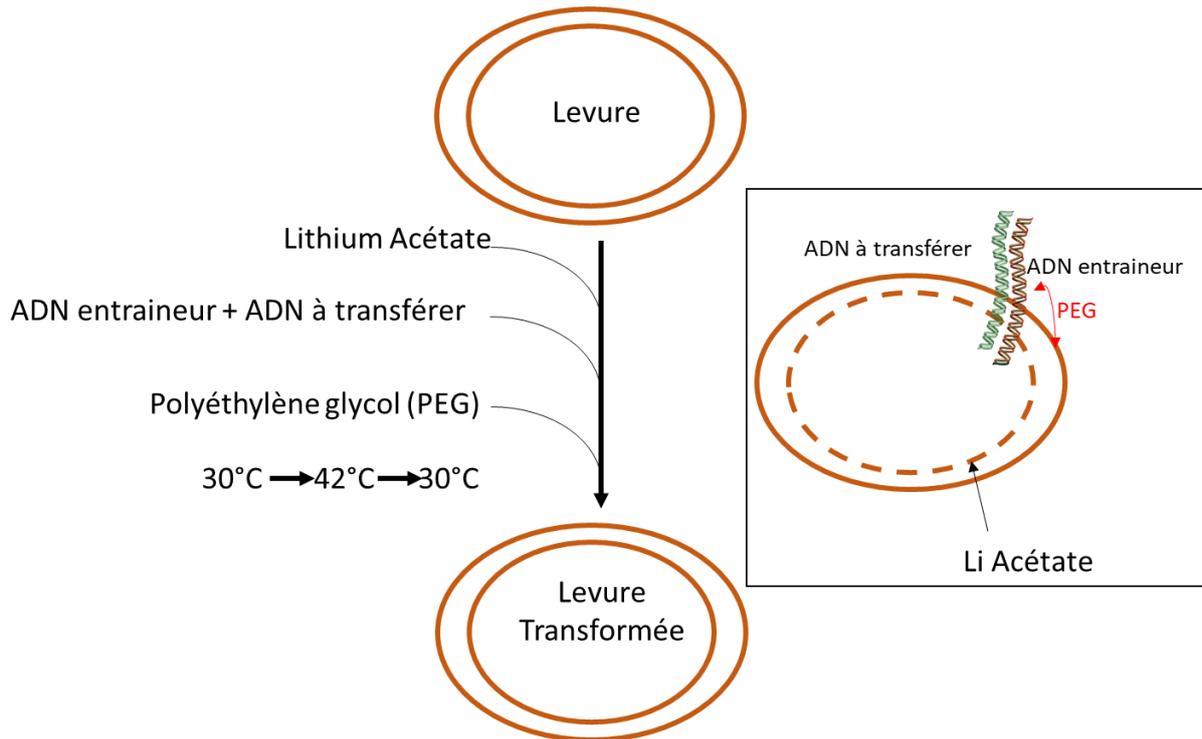


Figure 21. Transformation chimique de levure par lithium acétate. Les levures sont soumises à un choc thermique de 42°C en présence de lithium acétate qui perméabilise la membrane plasmique, du PEG qui fixe l'ADN à la paroi et de l'ADN entraîneur qui entraîne l'ADN dans la cellule.

3.3. Biolistique

La technique de transformation par la technologie de biolistique est très utilisée dans le domaine végétal. Elle consiste à l'introduction de fragment d'ADN dans des cellules de la plante par la propulsion de microbilles. Les acides nucléiques (transgène) sont fixés sur les microbilles qui peuvent être en tungstène ou en or. Le complexe ADN-billes est propulsé grâce à un canon à particule sur des tissus de plantes.

Certaines microbilles se logent dans les noyaux où l'ADN peut s'intégrer au sein du génome. Les cellules transformées sont sélectionnées via la résistance à un herbicide porté par une partie du transgène. Par régénération *in vitro*, des plantes OGMs régénérées à partir des cellules transformées sont produites (figure 22).

Partie I : Cours

Chapitre VIII : la transformation génétique

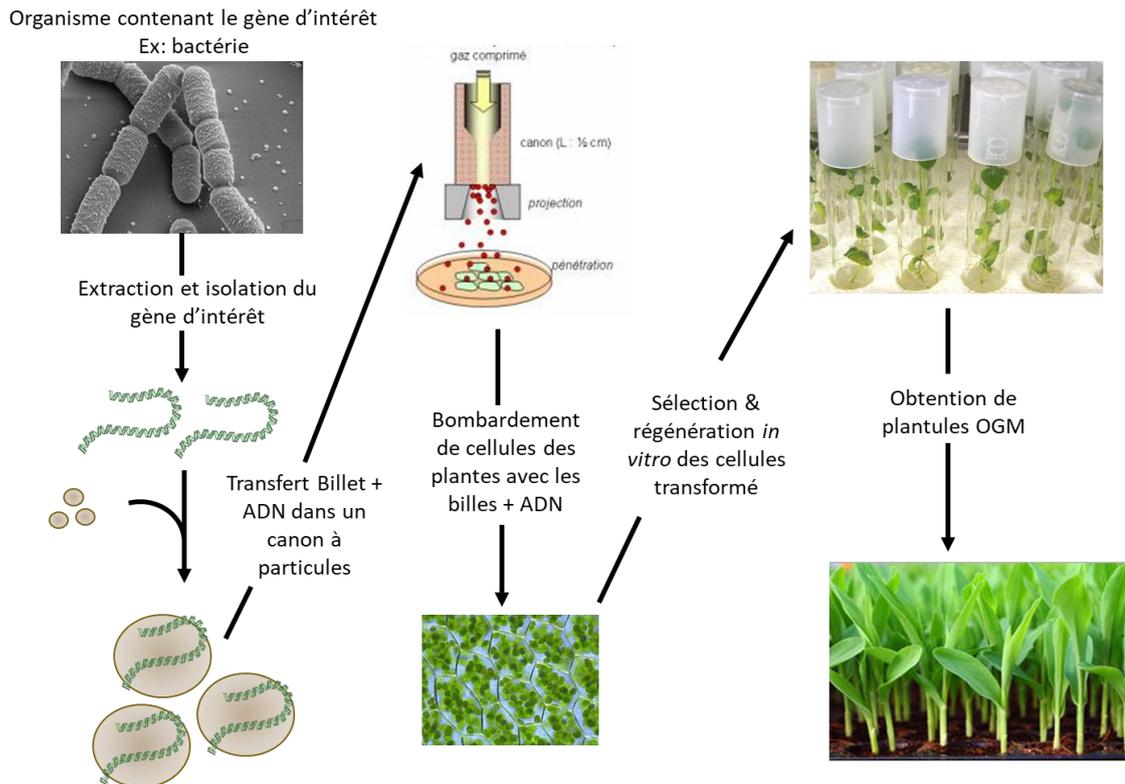


Figure 22. La transformation par biolistique. Les gènes issus d'un organisme telle qu'une bactérie sont extraits puis fixés sur des microbilles, qui sont propulsées par un canon à particule sur des cellules. Une sélection des cellules ayant intégrées l'ADN est sélectionnée puis des plantes sont régénérées à partir de ces cellules.

3.4. Microinjection

La microinjection consiste à introduire de l'ADN directement dans une cellule par injection, elle est utilisée dans la transformation de cellules animales. Ces dernières sont fixées par un micromanipulateur et un micro-injecteur injecte l'ADN au sein de la cellule. La transformation de cellules sexuelles est privilégiée car elle permet d'exprimer le transgène chez l'animal entier.

3.5. Agrotransformation

La bactérie phytopathogène, *Agrobacterium tumefaciens* est l'agent responsable de la tumeur du collet chez de nombreuses plantes. Elle vit à l'état épiphyte sur les végétaux, dans le cas d'une blessure, l'activation des gènes de virulence Vir a lieu suite à la perception du signal de dommage.

Partie I : Cours

Chapitre VIII : la transformation génétique

La virulence de cette bactérie est déterminée par la présence d'un plasmide portant les gènes de virulences, le plasmide Ti (Tumor induce). Ce dernier présente deux régions, la première est nommée ADNT, elle est située entre deux séquences flanquantes nommées bordures droite et gauche. Elle comporte les gènes oncogènes (de biosynthèse d'auxine et de cytokinine) et le gène de biosynthèse des opines qui sont des composés polycarbonnés utilisés pour la croissance de la bactérie. L'ADNT est transférée au cours du processus d'infection de la bactérie vers la cellule hôte. La deuxième région du plasmide comporte les gènes de virulences ou gènes Vir responsable de la production et du transfert de l'ADNT ainsi que du gène de dégradation des opines (figure 23).

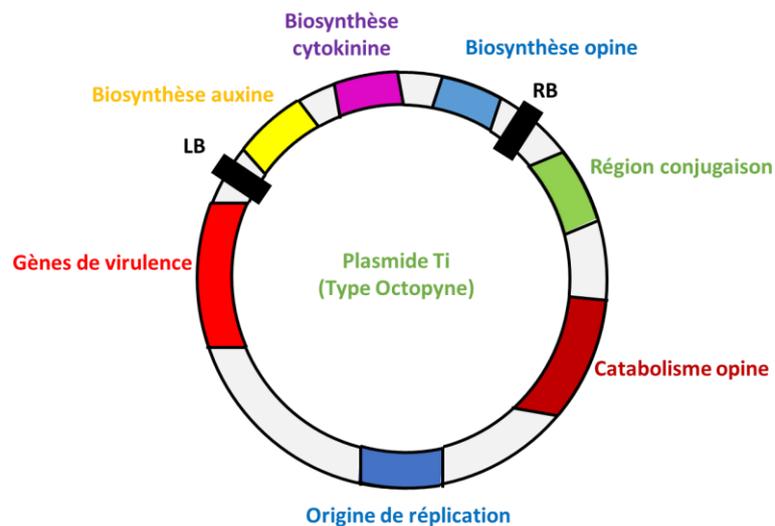


Figure 23. Schémas du prototype du plasmide Ti. Le plasmide Ti comporte une origine de répllication, un gène de catabolisme d'opine, une région contenant les gènes de virulence et un ADNT situé entre les bordures droite et gauche. Cette partie du plasmide contient des de biosynthèse d'auxines, de cytokinine et d'opine.

Lors du processus infectieux, les protéines Vir codées par le plasmide de virulence Ti active la répllication de l'ADNT qui contient des gènes oncogènes (de biosynthèse d'auxine et de cytokinine) ainsi qu'un gène de biosynthèse d'opine (molécule poly-carbonées utilisée comme source d'énergie).

L'ADNT répliqué est ensuite transféré au sein du noyau des cellules de la plante toujours sous l'action des protéines Vir, au sein du génome de la plante, l'ADNT est intégré par recombinaison hétérologue et les gènes sont induits ce qui conduit à une stimulation de la division des cellules

Partie I : Cours
Chapitre VIII : la transformation génétique

de l'hôte et à la formation d'une tumeur. Cette dernière libère des opines utilisés pour la croissance d'*A. tumefaciens* (figure 24).

L'agro transformation consiste à utiliser *A. tumefaciens* comme agent de transformation. Un plasmide de virulence Ti atténué (ou désarmé) est employé, au sein duquel l'ADNT a été remplacés par le transgène. Le plasmide de virulence atténué est ensuite transféré au sein de la bactérie.

La transformation des cellules végétale est réalisée par l'incubation de cellules de plante en présence de la bactérie ou le traitement d'un organe telle que les fleurs. Les cellules ou individus transformés sont sélectionnés par l'expression d'un gène de résistance à un herbicide codé par le transgène.

Partie I : Cours
Chapitre VIII : la transformation génétique

Cellule Hôte

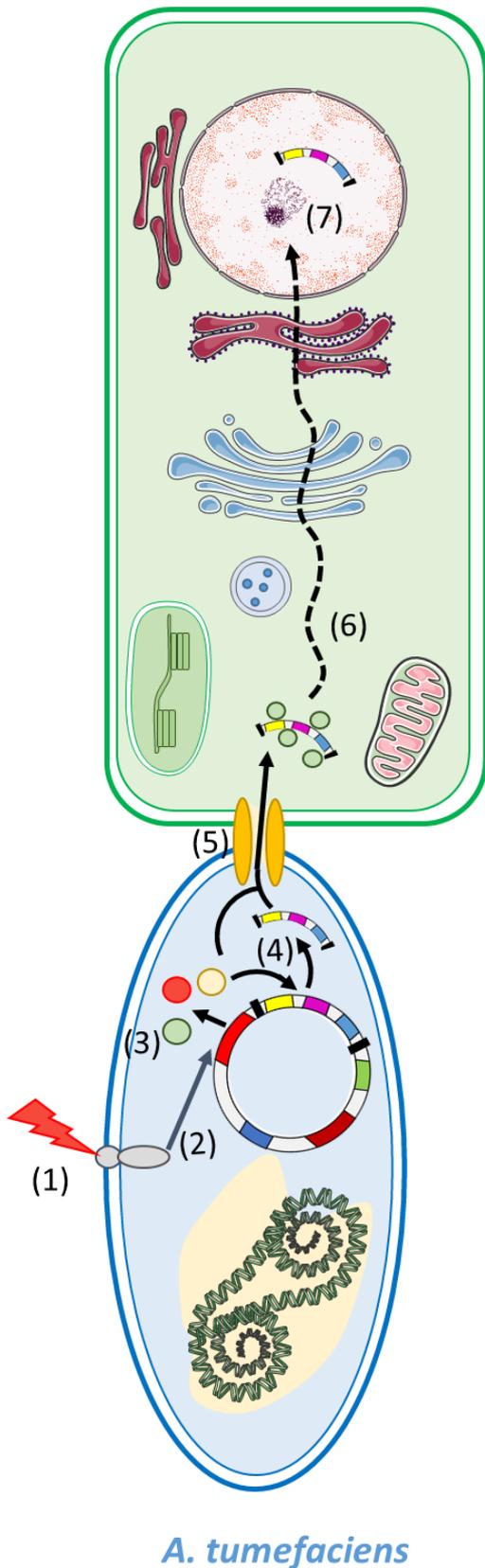


Figure 24. Mécanisme d'infection par ADNT d'*A. tumefaciens*. Après la perception de signaux extracellulaire (1) *A. tumefaciens* exprime les gènes de virulence ou *Vir* (2). Les protéines *Vir* (3) réplique l'ADNT (4) et le transfert vers la cellule végétale (5), puis vers le noyau (6) au sein duquel les gènes contenus dans l'ADNT sont exprimés.

Partie I : Cours
Chapitre VIII : la transformation génétique

Références bibliographiques

- 1- Eynard, Nathalie, and Justin Teissié. 2000. *Electrotransformation*. Springer Berlin Heidelberg.
- 2- Gietz, R. Daniel, and Robin A. Woods. 2018. "Genetic Transformation of Yeast." *BIOTECHNIQUES* 30(4). <https://doi.org/10.2144/01304rv02>.
- 3- Johnston, S. A. 1990. "Biolistic Transformation: Microbes to Mice." *Nature* 346(6286): 776–77.
- 4- Wang, Kan. 2015. *Agrobacterium Protocols*. ed. Kan Wang. Springer US. <https://doi.org/10.1007/978-1-4939-1658-0>.

Partie I : Cours

Chapitre IX : Mutation et mutagenèse

1. Evolution des génomes

La théorie de l'évolution ou théorie Darwinienne postule que l'ensemble des êtres vivants présente une origine commune. L'ADN est un bien commun que partage l'ensemble des êtres vivants sur terre. Le génome peut être de taille et de complexité différentes en fonction de l'organisme, ceci dépend en partie du milieu dans lequel se développe ce dernier, par exemple les plantes présentes un taille et une complexité de génome plus importante que celui humain, ceci est dû au fait qu'elles sont soumises à des stress environnementaux plus importants.

Au cours de l'évolution, différentes forces ont modifiés les génomes, l'endoréduplication (duplication multiple d'une partie ou de la totalité d'un génome) et la duplication de gènes ont permis d'augmenter leur taille, le processus de mutagenèse et de recombinaison ont permis d'augmenter leur complexité.

Dans le cas de la mutagenèse, en fonction de l'emplacement de la mutation (exemple ADN codant ou non-codant), du type de mutation (voir ci-dessous) l'effet peut être positif, négatif ou neutre. La sélection naturelle telle qu'énoncée par Charles Darwin permet de trier les mutations et les modifications de manière plus générale de telle sorte qu'une modification positive sera rapidement fixée, alors qu'un changement négatif sera rejeté. Les modifications neutres quant à elles sont maintenues ou perdues et participent à la complexification des séquences génomiques.

2. Les mutations à la base de la diversité génétique

Le phénotype est défini comme l'ensemble des modifications observables chez à un organisme. Il résulte de l'expression du génotype défini par l'ensemble des gènes. La diversité des phénotypes découle de la diversité des allèles retrouvés au sein d'une population, ainsi que de leur expression.

L'apparition de nouveaux allèles est en partie liée à la modification d'allèles ancestraux ou préexistants notamment par l'ajout de mutations. Ces dernières peuvent affecter l'expression du gène en touchant le promoteur ou affecter la séquence de la protéine en touchant le cadre ouvert de lecture.

Partie I : Cours

Chapitre IX : Mutation et mutagenèse

Les effets d'une mutation à l'échelle d'un organisme dépendant du niveau et de la localisation de l'expression du gène, de son importance et dans le cas d'une mutation négative, de la présence de copie du gène pouvant compenser l'effet de la mutation, on parlera de redondance fonctionnelle qui est observée dans le cas des familles multigéniques.

La taille de la population portant l'allèle, ainsi que la pression de sélection détermine le maintien d'une mutation ; une mutation bénéfique sera maintenue car conférant un avantage et sera largement dispersée dans une population. De plus l'incidence du nombre d'individu sur l'apparition de nouvelles modifications de l'allèle sera d'autant plus grande que la taille effective de la population est importante.

3. Les différents types de mutation

Plusieurs types de mutation conduisant à différents effets sur les gènes peuvent être observés. Une mutation de transition provoque la substitution entre des bases de même type (pyrimidique ou purique), alors qu'une mutation de transversion provoque la substitution d'une base pyrimidique par une base purique et inversement. Une mutation peut en fonction de sa position augmenter ou réduire l'effet d'un gène ; effet knock in et knock down respectivement, annuler l'effet du gène ; knock out (K.O).

Une mutation affectant le promoteur peut réduire ou annuler son activité, si les modifications affectent les éléments cis, inversement certaines mutations modifient le profil d'expression ou augmentent l'activité du promoteur en mimant des motifs cis. Une mutation affectant le cadre ouvert de lecture peut ne pas avoir d'effet si la mutation touche une séquence non-codante (exemple intron) ou si la mutation est non-sens (ne modifie pas l'acide aminé).

Si une mutation a lieu dans la séquence codante du gène, elle peut avoir un effet si elle affecte de manière drastique la séquence de la protéine, par exemple dans le cas de mutation intégrant un codant stop au milieu de la séquence aboutissant à une protéine tronquée, ou une mutation ajoutant ou réduisant 1 à 2 nucléotides provoquant un décalage du cadre de lecture et donc à une modification drastique de la séquence de la protéine.

De manière plus subtile, une mutation peut changer un acide aminé clé dans l'activité de la protéine conduisant à une modification d'activité, par exemple une mutation provoquant l'ajout d'un acide aminé chargé négativement (exemple acide aspartique) au niveau d'une

Partie I : Cours

Chapitre IX : Mutation et mutagénèse

région de la protéine dont la phosphorylation est responsable de son activation provoque la formation d'une protéine constitutivement active.

4. Les sources naturelles de mutagénèse

Plusieurs événements peuvent naturellement produire des mutations dans les génomes, des agents chimiques ou physiques présents dans l'environnement peuvent provoquer des mutations. C'est le cas des rayons ultraviolets qui induisent la formation de dimère de thymine entre deux bases adjacentes, provoquant ainsi leur dissociation des adénines correspondantes. Le mécanisme de réparation de l'ADN responsable de changer les nucléotides peut dans certain cas faire des erreurs et introduire par exemple des nucléotides G ou C.

D'autre part, des mécanismes biologiques tel que la recombinaison entre deux molécules d'ADN qui entraîne un échange de fragments entre les deux brins est une source de mutations spontanées. Au moment de la réplication de l'ADN, l'ADN polymérase peut introduire des erreurs lors du processus en changeant le nucléotide, particulièrement si l'enzyme est dépourvue d'activité d'autocorrection « proof reding ». Elle peut aussi provoquer la réduction ou l'ajout de nucléotides par glissement, ceci est particulièrement observé au niveau des séquences répétées ou riche en A-T.

5. L'approche de génétique directe et génétique reverse en recherche

En recherche moderne deux approches sont employées pour identifier des gènes impliqués dans un processus donné, la première est une approche sans a priori, elle consiste à produire des mutations spontanées dans le génome de manière artificiel et aléatoire. Après la mutagénèse un crible est réalisés afin d'identifier les mutants ayant un phénotype intéressant. Par la suite le gène responsable du phénotype est identifié, cette approche est nommée « génétique directe ».

La deuxième approche appelé génétique reverse consiste à introduire des mutations de manière ciblé dans des gènes potentiellement impliqués dans le processus étudié via la mutagénèse dirigée. Le phénotype du mutant est observée afin de valider ou d'invalider le rôle du gène (figure 25).

Partie I : Cours

Chapitre IX : Mutation et mutagénèse

La génétique reverse est apparue avec la découverte des séquences des séquences génomiques et donc aux gènes pour un organisme donné.

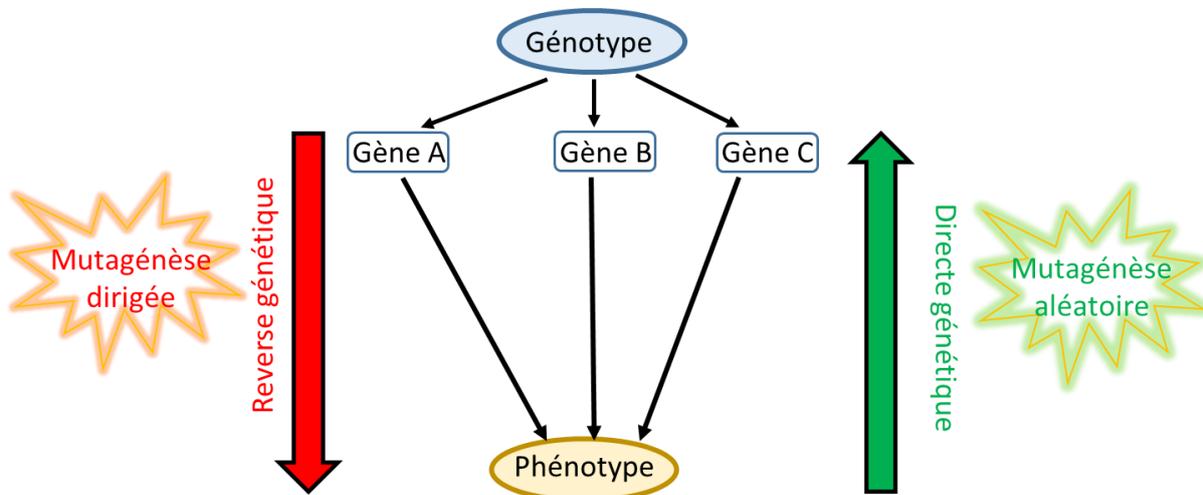


Figure 25. La génétique directe et reverse en recherche. La génétique directe repose sur la création de mutations aléatoires, puis l'isolation de mutant ayant un phénotype en adéquation avec le processus étudié, le gène responsable est par la suite identifié. La génétique reverse repose sur la connaissance du génome, les gènes d'intérêts sont mutés de manière ciblée puis le phénotype du mutant est analysé.

6. La transmission de mutation

La transmission d'une mutation est un élément important à prendre en considération lors de la réalisation de la mutagénèse. On peut identifier deux modes de transmission de mutations, la première est observée dans les mutations somatiques qui affectent des cellules différenciées et qui est limitée aux cellules dérivant de cette dernière. L'autre type de mutation est les mutations sexuelles qui affectent les cellules germinales, elles sont donc transmises à l'ensemble des cellules d'un individu.

En recherche, quand aucun a priori n'est émis pour l'étude d'un gène ou d'une fonction la mutagénèse est réalisée sur des cellules sexuelles afin de détecter tous les effets potentiels de la modification sur un l'organisme, par ailleurs une mutagénèse sur certaines cellules peut-être envisagé afin de confirmer des hypothèses fonctionnelles.

Partie I : Cours

Chapitre IX : Mutation et mutagénèse

7. La mutagénèse aléatoire

La mutagénèse aléatoire est utilisée en génétique directe afin d'identifier des gènes impliqués dans un processus biologique donné. Les mutations produites affectent de manière aléatoire les séquences d'un génome, elles doivent être saturantes afin de toucher un nombre maximum de gènes, par exemple dans le cas de la mutagénèse à l'agent chimique EMS (Ethyl Methyl Sulfonate) un nombre moyen de 5×10^2 à 5×10^4 modifications a lieu par gène pour un génome donnée.

Plusieurs agents mutagènes peuvent être utilisés afin de produire des mutations spontanées dans un génome. Le choix de la méthode employée dépend principalement de l'organisme étudié. Par exemple la mutagénèse à l'ADN T qui consiste à introduire de l'ADN modifié du pathogène *A. tumefaciens* dans des cellules de plante ne peut être appliquée que sur certaines espèces hôtes du pathogène.

7.1. Mutagenèse par agents chimique

La mutagénèse chimique consiste à traiter des cellules grâce à des agents chimiques qui provoquent la formation de mutations. Trois catégories de mutagènes chimiques sont identifiées.

La première provoque des modifications de nucléotides en miment et remplacent des bases azotées, c'est le cas par exemple du 5-Bromouracil qui est un analogue de la thymine et qui s'apparie à la guanine.

La deuxième catégorie de mutagène chimique provoque la modification de bases azotées, dans le cas de l'EMS, entre autre une modification de la guanine en O6-méthyle guanine a lieu, cette base s'apparie à une thymine provoquant ainsi une transition de pair de nucléotide de G-C en A-T.

La troisième catégorie d'agents chimiques est représentée par les intercalants de l'ADN, tel que le bromure d'etidium utilisé en biologie moléculaire pour la révélation de l'ADN. Les intercalant de l'ADN s'introduisent entre les pairs de bases et provoque des distorsions au niveau de la séquence aboutissant à la formation de mutation.

Partie I : Cours

Chapitre IX : Mutation et mutagénèse

7.2. Mutagénèse par agents physique

L'irradiation aux UV de cellules introduit des mutations, ces dernières peuvent être corrigées dans certain cas par le système enzymatique photolyase/glycylase qui remplace les nucléotides endommagés, par ailleurs si le nombre de mutations est trop important le système est saturé et un nombre conséquent de mutation sont introduites.

Le bombardement aux neutrons est un autre type de mutagénèse physique, elle consiste dans la projection de neutron par un canon à particule sur une molécule d'ADN. Cela provoque des cassures d'ADN, l'introduction de mutation à lieu lors de la réparation des brins cassés.

7.4. Mutagénèse insertionnelle à l'ADNT

La mutagénèse insertionnelle à l'ADNT consiste à utiliser une version modifiée du plasmide Ti dont les gènes de l'ADNT ont été remplacés par un ou des gènes de sélection (gène de résistance aux antibiotiques et/ou de résistance à un herbicide). Le plasmide Ti modifié est re-introduit dans la bactérie et le processus d'infection est recrée artificiellement de telle sorte que l'ADNT modifié intègre le génome de l'hôte.

Au cours du processus plusieurs milliers d'ADNT sont insérés dans le génome assurant ainsi une bonne saturation de ce dernier en mutation.

8. Le crible de mutants

Suite au processus de mutagénèse aléatoire, des centaines de mutants sont obtenus, chaqu'un avec plusieurs dizaines de milliers de mutations. Certains mutants présentent des mutations sans intérêts par rapport au processus biologique étudié, alors que d'autre présente des mutations intéressantes amplifiant ou bloquant le phénomène.

Le crible consiste à sélectionner les mutants portants des mutations intéressantes (affectant de manière significative des gènes impliqués dans le processus). Deux grands types de cribles peuvent être utilisés :

i) Le premier est le crible phénotypique, largement employé en recherche, il se base sur l'identification de mutants comportant des variations phénotypiques en adéquation avec le processus étudié dans les conditions expérimentales pendant lesquelles le processus a lieu.

Partie I : Cours

Chapitre IX : Mutation et mutagénèse

Par exemple pour l'étude de la motilité bactérienne, des mutants ayant un mouvement modifié ou altéré sont recherchés.

ii) Le deuxième type est le crible moléculaire, il est utilisé quand les variations phénotypiques sont absente ou trop subtiles pour être détectées. Dans ce cas une analyse d'expression de gènes ciblent préalablement identifiés comme activés par le processus peut être réalisée chez des mutants, afin d'identifié des individus dans lequel le processus est invalidé ou altéré. La figure 26 résume les principales étapes expérimentales en génétique directe.

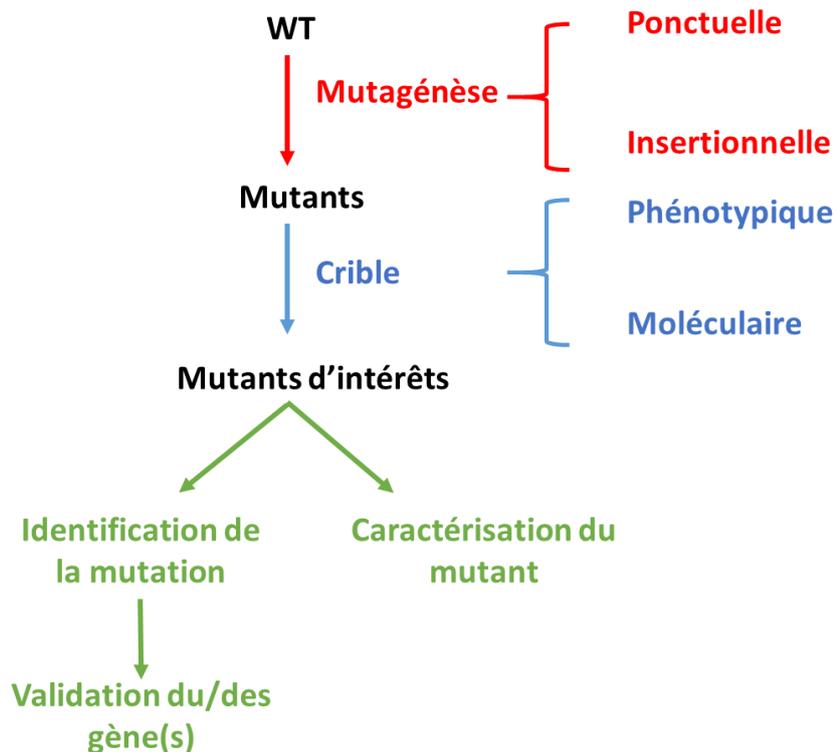


Figure 26. Principales étapes de recherche en génétique directe. A partir d'un individu sauvage (WT) une mutagénèse aléatoire est réalisée ce qui aboutit à la formation de mutants. Un crible est réalisé afin d'identifier les mutants comportant un phénotype intéressant pour l'étude. Ces derniers sont caractérisés afin de mieux comprendre le rôle du gène muté d'un part et d'autre part le gène responsable du phénotype mutant est recherché.

Partie I : Cours

Chapitre IX : Mutation et mutagenèse

9. Identification de mutations

Une fois les mutants isolés, la localisation de la mutation impliquée est indispensable afin d'identifier le gène affecté. Plusieurs approches peuvent être utilisées afin de retrouver une mutation, le choix de l'approche dépendant du type de mutation réalisée.

Les mutants doivent d'abord subir une série de croisement avec un fond sauvage (back cross) afin de purifier le fond génétique et d'éliminer les mutations parasites. Généralement le back cross est accompagné par le processus d'identification de la mutation.

9.1. Identification de mutations ponctuelles

Dans le cas de mutation ponctuelle telle que celle obtenues par mutagenèse EMS, une analyse de co-ségrégation de marqueurs moléculaires avec le phénotype peut être réalisée sur une lignée en ségrégation ; obtenue par back cross. La descendance obtenue suit dans la majorité des cas une distribution mendélienne avec présence ou absence du phénotype mutant.

Un génotypage par amplification spécifique de marqueurs est réalisé pour chaque plante, puis comparé à la lignée sauvage afin d'identifier les variabilités. A l'issue de l'analyse certains marqueurs sont non-variables entre les lignées, alors que d'autres sont variables, parmi eux certains sont systématiquement retrouvés avec le phénotype mutant et sont donc en co-ségrégation avec ce dernier.

Un séquençage (analyse de la séquence ADN) des régions comportant les marqueurs co-ségrégants est réalisée et comparée à la lignée sauvage afin d'identifier les mutations.

Les nouvelles technologies de séquençage permettent actuellement à moindre coût de séquencer un génome entier et d'identifier des variations d'un nucléotide ou SNP (Single nucleotide permutation), une autre stratégie consiste à ré-séquencer le génome de mutant ayant un fond génétique pure et de le comparer avec le sauvage.

9.2. Identification de mutation causée par une insertion

L'identification causée par une mutagenèse intentionnelle est relativement plus simple que l'identification d'une mutation ponctuelle, plusieurs approches existent ; l'une d'entre elle consiste à identifier les séquences flanquantes l'ADN insérées au cours de la mutagenèse.

Partie I : Cours

Chapitre IX : Mutation et mutagénèse

Après purification du fond génétique, l'ADN du mutant est extrait, fragmenté et cloné dans des plasmides, puis des bactéries sont transformées.

L'insert responsable de la mutation doit contenir un gène de résistance à un antibiotique, ainsi les bactéries transformées sont sélectionnées sur un milieu comportant l'antibiotique.

Le plasmide des bactéries résistance est extrait et l'insert séquencé. Il comporte deux types de séquences, celles de l'élément mutagène ainsi que des fragments du gène muté flanquant l'élément d'insertion. Le gène est retrouvé via la recherche des séquences identifiées dans le génome.

10. Validation de gènes candidat

Suite au processus d'identification des mutations, un ou plusieurs gènes potentiellement responsables du phénotype mutant sont identifiés, il faut par la suite valider ces gènes.

Deux grandes approches sont employées, la première consiste à identifier d'autres mutants des gènes candidats obtenus indépendamment, la présence d'au moins deux allèles différents du gène muté présentant le même phénotype confirme le rôle du gène.

La deuxième approche nommée teste de complémentation fonctionnelle, consiste à introduire une version non-mutée du gène dans le mutant, la restauration du phénotype sauvage valide l'implication du gène.

11. Mutagénèse dirigée

Utilisé en génétique reverse, la mutagénèse dirigée consiste à inactiver de manière ciblée un gène préalablement identifié comme potentiellement impliqué dans le processus étudié. Une connaissance préalable du génome est requise, la sélection du gène peut s'effectuer par différents critères ; profile d'expression, fonction putative de la protéine, recherches bibliographiques...etc.

12. Mutagénèse en système bactérien & fongique

Les bactéries et certains organismes fongiques réalisent naturellement de la recombinaison homologe qui consiste en un échange de fragments d'ADN entre deux séquences similaires.

Partie I : Cours

Chapitre IX : Mutation et mutagénèse

Par exemple chez les bactéries, la protéine RecA est impliquée dans la recombinaison homologue (figure 27).

L'approche employée dans ce type d'organisme consiste à introduire une version modifiée du gène cible, cette version est non-fonctionnelle et comporte un gène de sélection (résistance à un antibiotique) flanqué de régions homologues du gène.

La construction est introduite dans l'organisme qui procède au remplacement par recombinaison homologue du gène cible. Les microorganismes ayant intégrés la construction sont sélectionnés via le gène de résistance. Suite à cela la région génomique contenant le gène cible est re-séquencée afin de confirmer son remplacement par la version tronquée.

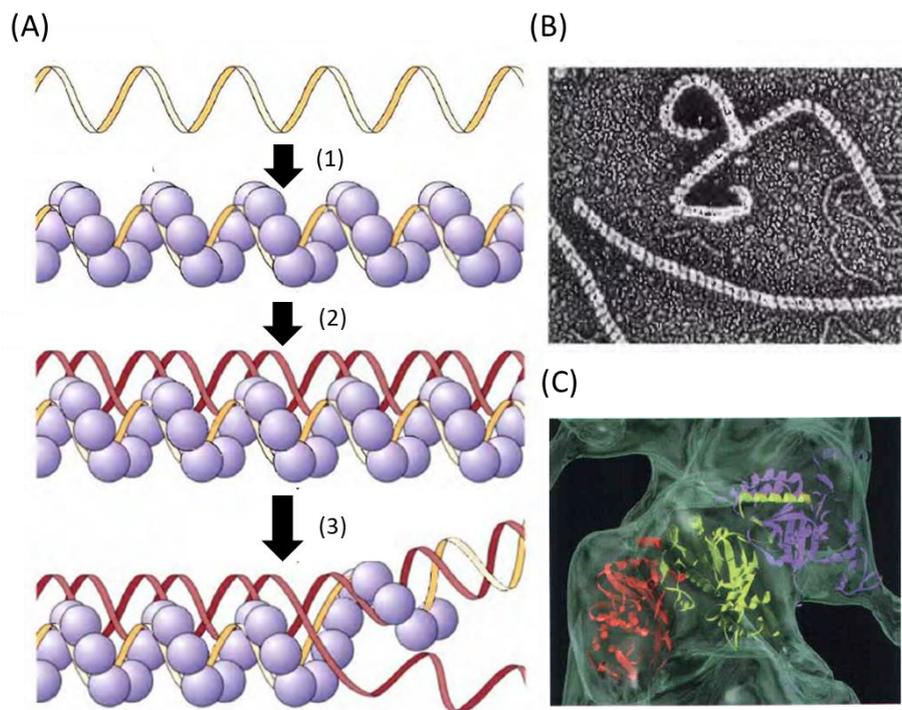


Figure 27. Protéines RecA dans la recombinaison homologue. Les protéines RecA forment un complexe au tour de l'ADN et procède à son intégration au niveau du génome. A) mécanisme de fonctionnement de la protéine RecA, B) observation en microscopie électronique des protéines RecA, C) structure 3D des protéines RecA.

13. Invalidation de gènes par approche RNA silencing

Cette approche exploite un mécanisme naturellement présent chez les eucaryotes qui est l'extinction d'expression de gènes ou RNA silencing (voir chapitre I). Lors de cette approche,

Partie I : Cours

Chapitre IX : Mutation et mutagénèse

une version modifiée du gène cible est construite de telle sorte que les ARNm produits présente une structure Tige-boucle. Ceci est obtenue généralement par ajout de plusieurs résidu adénine et thymine des deux côtés de la séquence du gène (voir figure 28).

La construction est exprimée dans l'organisme, les ARNm aberrant sont reconnus par la machinerie de silencing et une production de petit ARN interférents à lieu aboutissant à l'extinction d'expression du gène cible (figure 28).

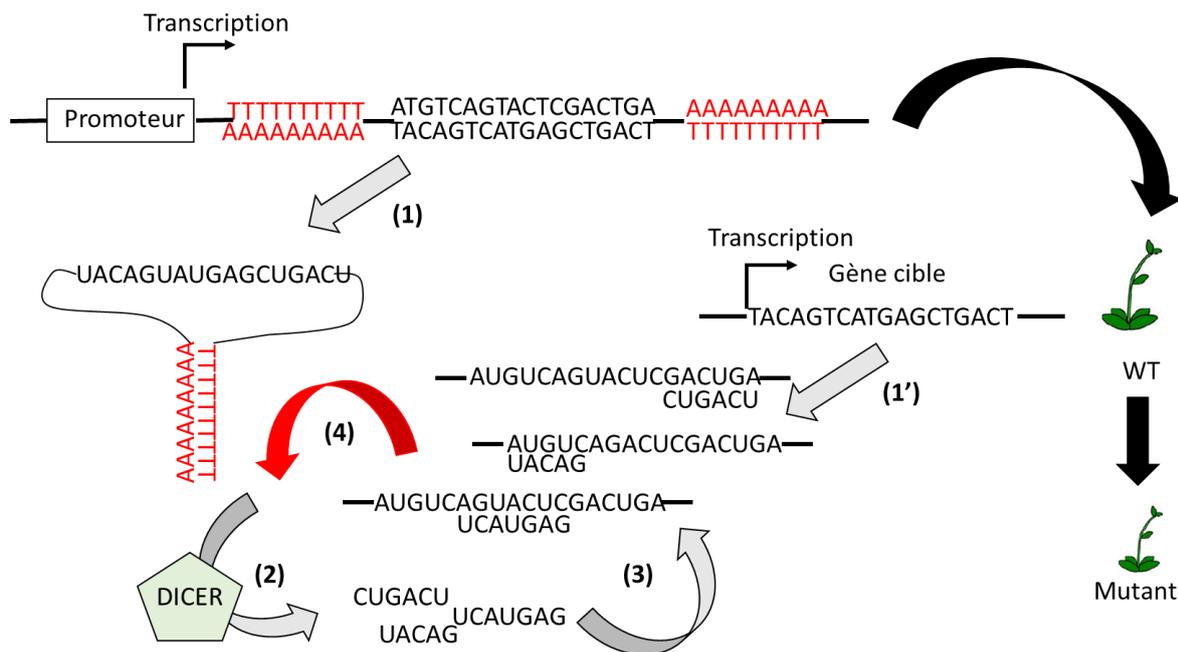


Figure 28. Principe du RNA silencing. Dans le but d'inactiver un gène, une version modifiée de ce dernier formant une structure tig-boucle est exprimée au sein de l'organisme (1). Les ARNm aberrants sont reconnus par la machinerie de silencing et des petits ARNs interférants sont formés (2) ce qui provoque la dégradation des ARNm endogènes (1', 4) après la reconnaissance. Le phénotype du mutant est analysé afin de vérifier l'implication du gène ciblé dans le processus étudié.

14. Invalidation de gènes par approche ribozyme

Le ribozyme est une structure d'ARN ayant une activité nucléase. Ceci lui est conféré par la formation d'une structure spéciale liée au repliement et à la stabilisation de sa séquence nucléique.

Partie I : Cours
Chapitre IX : Mutation et mutagénèse

Après reconnaissance d'un ARN, le ribozyme se lie, clive et libère les fragments d'ARN clivé. Un ribozyme peut effectuer plusieurs cycles de dégradation d'ARN cible. La reconnaissance de l'ARN cible par le ribozyme s'effectue via des régions complémentaires entre les deux ARNs.

L'approche d'inactivation est réalisée par l'expression d'un ribozyme modifié contenant une séquence complémentaire de l'ARNm du gène cible. Le ribozyme dégrade de manière spécifique les ARNm synthétisé (figure 29).

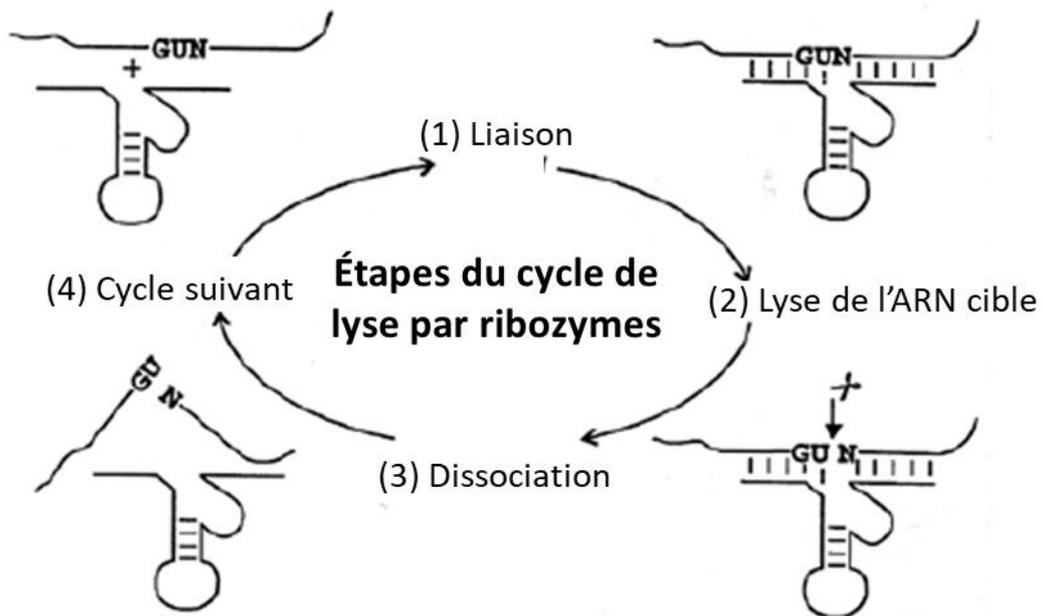


Figure 29. Cycle d'action du ribozym. Le ribozyme reconnaît l'ARN cible (1). Une modification structurel est provoquée par le ribozyme ce qui provoque la cassure de l'ARN cible (2). Puis le ribozyme se dissocie (3) pour réitérer le cycle sur un autre ARN (4).

Partie I : Cours
Chapitre IX : Mutation et mutagénèse

Références bibliographiques

- 1- Brahmachari, Goutam. 2017. *Biotechnology of Microbial Enzymes Production, Biocatalysis and Industrial Applications*. First Edit. ed. Goutam Brahmachari. Academic Press. <https://doi.org/10.1016/C2015-0-00048-X>.
- 2- Braman, Jeff. 2010. *In Vitro Mutagenesis Protocols*. Third Edit. ed. Jeff Braman. Humana Press.
- 3- Carmichael, Gordon. 2005. *RNA Silencing Methods and Protocols*. First Edit. ed. Gordon Carmichael. Humana Press.
- 4- Fedor, Martha J. 2000. "Structure and Function of the Hairpin Ribozyme."
- 5- Geoffrey M, Cooper, and Hausman Robert E. 2006. *The Cell A Molecular Approach*. Fouth. ASM Press.
- 6- Lodish, Harvey et al. 2000. *Molecular Cell Biology*. 4th editio. Connected, Media.
- 7- Meksem, Khalid, and Günter Kahl. 2010. *The Handbook of Plant Mutation Screening: Mining of Natural and Induced Alleles*. First Edit. eds. Khalid Meksem and Günter Kahl. Wiley Blackwell.
- 8- Sikora, Per et al. 2011. "Mutagenesis as a Tool in Plant Genetics, Functional Genomics, and Breeding." *International Journal of Plant Genomics* 2011: 13. <http://dx.doi.org/10.1155/2011/314829>.
- 9- Turner, Philip C. 1997. *Ribozyme Protocols*. First edit. ed. Philip C. Turner. Humana Press. <https://doi.org/10.1385/0896033899>.
- 10- Xu, Jian-zhong, and Wei-guo Zhang. 2016. "Strategies Used for Genetically Modifying Bacterial Genome: Ite-Directed Mutagenesis, Gene Inactivation, and Gene over-Expression." *J Zhejiang Univ Sci B* 17(2): 83–99. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4757579/>.

Partie II : Travaux dirigés

TD1. Evolution expérimentale

Les organismes eucaryotiques pluricellulaires peuvent interagir avec différents types de microorganismes, certains peuvent être pathogènes ou d'autres symbiotiques. Beaucoup de questions restent sans réponses quant aux forces évolutives qui sont à la base de l'émergence de ces différents types de microorganismes.

L'équipe du Dr. Jacques Batut et du Dr. Catherine Masson-Boivin au LIPM (Laboratoire des Interactions Plantes-Microorganismes) à Toulouse est pionnière dans la recherche en évolution expérimentale sur le pathogène *Ralstonia solanacearum*. Ce TD propose de retracer une partie de leurs travaux scientifiques.

Question 1. Ci-dessous un extrait de l'abstract issu de l'étude, veuillez déterminer le contexte général de l'étude ainsi que les objectifs fixés et la méthode employée ?

Réponse 1. Les auteurs travaillent sur l'évolution expérimentale (en laboratoire) de la bactérie pathogène *R. solanacearum* et de sa transformation en bactérie symbiotique. Les chercheurs veulent établir un lien entre deux phénomènes opposés, la pathogénicité et la symbiose. La méthode consiste à générer une bactérie chimère en introduisant le plasmide symbiotique de *C. taiwanensis* dans la bactérie *R. solanacearum*. Les chimères sont soumises à plusieurs cycles de culture *in planta* et *ex planta* (en milieu liquide) afin de reproduire le processus évolutif naturel.

Nitrogen-fixing symbionts of legumes have appeared after the emergence of legumes on earth, approximately 70 to 130 million years ago. Since then, symbiotic proficiency has spread to distant genera of α - and β -proteobacteria, via horizontal transfer of essential symbiotic genes and subsequent recipient genome remodeling under plant selection pressure. To tentatively replay rhizobium evolution in laboratory conditions, we previously transferred the symbiotic plasmid of the Mimosa symbiont *Cupriavidus taiwanensis* in the plant pathogen *Ralstonia solanacearum*, and selected spontaneous nodulating variants of the chimeric *Ralstonia* sp. using *Mimosa pudica* as a trap. Here, we pursued the evolution experiment by submitting two of the rhizobial drafts to serial *ex planta*–*in planta* (*M. pudica*) passages that may mimic alternating of saprophytic and symbiotic lives of rhizobia

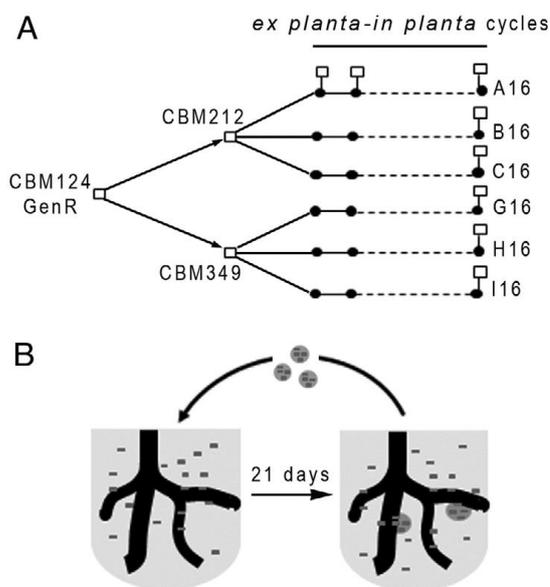
Partie II : Travaux dirigés

TD1. Evolution expérimentale

Question 2. *M. pudica* est l'hôte spécifique de *C. taiwanensis*, pourquoi les auteurs décident de tester la souche d'*R. solanacearum* évoluée avec *M. pudica* et non avec une autre espèce de légumineuse ?

Réponse 2. Premièrement car *M. pudica* est un hôte commun entre *R. solanacearum* et *C. taiwanensis*. Puis par rapport à la proximité génétique qui existe entre les deux bactéries ce qui sous-entend la présence de mécanisme moléculaire potentiellement conservés.

Après avoir transféré le méga plasmide symbiotique de *C. taiwanensis* à *R. solanacearum* les auteurs obtiennent deux souches chimères (CBM212 et CBM349), ils décident par la suite de soumettre ces souches à plusieurs cycles de culture *in planta* et *ex planta* selon la figure 1.



Question 3. Veuillez expliquer la figure 1 ?

Réponse 3. La figure 1 présente les résultats de l'évolution de souches de *Ralstonia* chimères (CBM124). Après transfère du plasmide symbiotique dans *R. solanacearum* elle est inoculée dans la plante *M. pudica*, deux souches filles sont isolées à partir de deux nodosités qui sont CBM212 et CBM349. Ces dernières sont ré-inoculées à partir d'une étape de culture liquide et les dérivés des souches sont isolés à partir de chaque nodosité formée, puis à leur tour mis en culture liquide et re-inoculée. Après 16 cycles, 6 souches sont isolées à partir de CBM212 et CBM349.

Question 4. Les auteurs procèdes à 16 cycles de cultures, dans quel but ?

Partie II : Travaux dirigés

TD1. Evolution expérimentale

Réponse 4. Les auteurs essayent de recréer le processus naturel d'évolution qui alterne des étapes d'interaction avec l'hôte et des étapes de vie dans le sol. A chaque cycle les chimères accumulent en théorie de nouvelles mutations qui augmentent leur fitness avec l'hôte.

Dans un second temps, le niveau d'infection des tissus de l'hôte par les clones chimères a été analysé, la figure 2 présente les résultats obtenus.

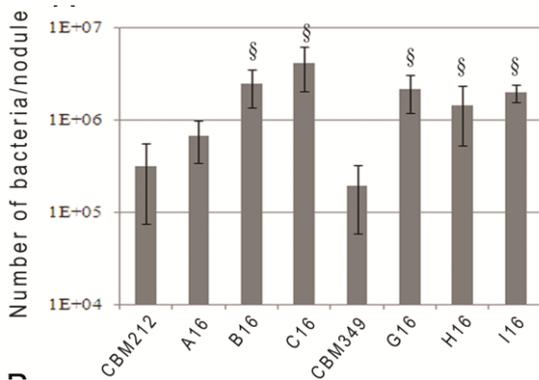


Figure 2. In planta fitness of evolved clones. Values are averages \pm standard deviation of the number of bacteria recovered per nodule from three independent experiments. In each experiment, nodules from 6 plants were harvested 21 days after inoculation (dpi). Symbol § indicates significant differences with the relative ancestor ($P < 0.05$; Kruskal-Wallis test)

Question 5. Discuter la méthodologie employée ?

Réponse 5. Chaque souche chimère est inoculée dans la plante, les nodosités formées sont récoltées, broyées et les débris sont nettoyés par centrifugation à basse vitesse. Le surnageant est étalé sur un milieu de culture après dilution et le nombre relatif de bactérie est évalué. Pour chaque souche, les nodosités de 6 plantes sont analysées, cela permet de réaliser des tests statistiques afin de vérifier la validité des observations.

Question 6. Pourquoi les auteurs décident-t'il de réaliser cette expérience ?

Réponse 6. Les auteurs réalisent cette expérience afin de voir les niveaux de survie au sein des nodosités formées, ceci permettra notamment de vérifier si l'ensemble des souches testées (chimères mères et descendantes) présentent le même niveau d'association avec l'hôte.

Question 7. Analyser la figure 2, quelle est votre conclusion ?

Réponse 7. La figure 2 présente le nombre de bactérie par nodule par souche testée. On observe que le nombre de bactérie n'est pas identique entre les différentes souches, les souches CBM212 et CBM349 présentent les valeurs les plus faibles, alors que leurs descendants présentent une plus grande survie avec C16 qui possède le meilleur niveau. Ces

Partie II : Travaux dirigés

TD1. Evolution expérimentale

résultats montrent que les souches chimères ne possèdent pas la même capacité d'adaptation et que l'évolution des souches mères a permis une amélioration de l'adaptation des chimères filles.

Une analyse live and dead straining est réalisée sur des tissus de l'hôte infectés par les bactéries, cette coloration permet de révéler les bactéries vivantes en vert et les bactéries mortes en rouge. La figure 3 présente les résultats obtenus

Question 8. Discutez la méthodologie employée ?

Réponse 8. Après inoculation des plantes avec les bactéries sauvages, les mutants non-fixateurs d'azote *nifH*, la souche chimère G16 et son ancêtre CBM349. Les nodosités sont récoltées, disséquées et colorées avec des colorants révélant les bactéries mortes en rouge et vivantes en vert.

Question 9. Qu'est-ce que la figure 3 apporte de plus par rapport à la figure 2 ?

Réponse 9. La figure 3 permet une observation *in vivo* et *in planta* du niveau de survie des bactéries en plus d'avoir une observation du niveau d'occupation des tissus.

Question 10. Analyser la figure 3, quelle est votre conclusion ?

Réponse 10. La figure 3 présente l'analyse des niveaux de survie d'une souche sauvage, non fixatrice, d'une chimère et de son ancêtre. On observe que le mutant non-fixateur présente un niveau de survie équivalent à celui de la souche sauvage ce qui montre que l'absence de fixation n'altère pas la survie. La chimère présente un niveau de survie important, mais réduit par rapport à la souche sauvage, alors que la souche ancêtre présente un niveau de survie très réduit. Ces résultats montrent que l'évolution expérimentale a permis d'améliorer la survie de la souche chimère, mais reste réduite par rapport à la condition de symbiote naturel.

Question 11. Discuter le choix des souches présentés par les auteurs ?

Réponse 11. La première souche présente le sauvage symbiotique et sert de référence, la souche *nifH* est une souche symbiotique mutée incapable de fixer l'azote, elle sert de référence négative car incapable d'achever l'interaction. La souche G16 et CBM349 sont les

Partie II : Travaux dirigés
TD1. Evolution expérimentale

tests, la comparaison des deux souches permet de connaître l'effet des cycles d'évolution expérimentale.

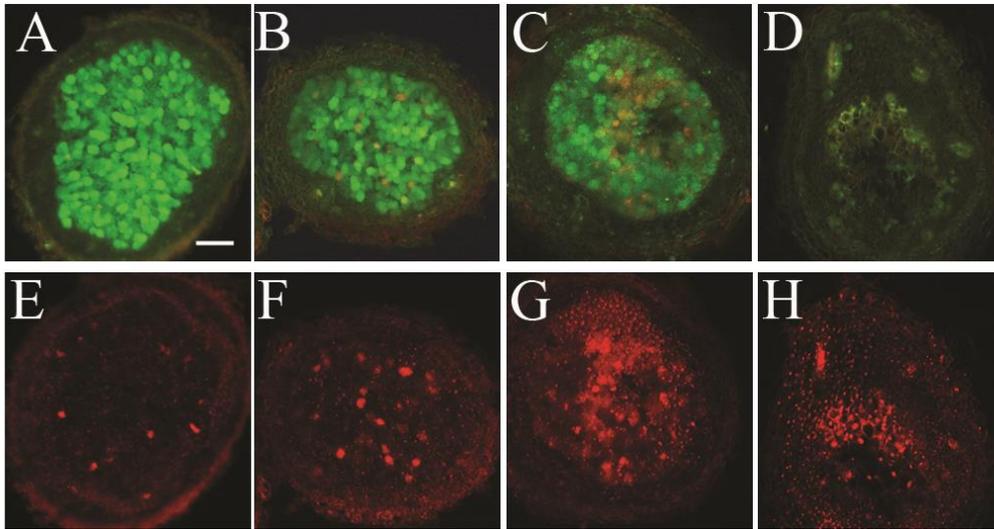


Figure 3. Persistence of endosymbiotic evolved clones. Live/dead staining of *Mimosa pudica* nodules formed by A and E, *Cupriavidus taiwanensis* LMG19424 and B and F, its *nifH* mutant CBM2432; C and G, *Ralstonia* G16, and D and H, CBM349. A through H, Fresh-sliced nodules harvested at 7 or 14 days post-nodule formation (dpnf) were stained with Syto9 (green signal, live bacteria) (A to D) and PI (red signal, dead bacteria) (E to H).

Les nécroses tissulaires ont été analysées par les auteurs chez l'hôte après inoculation avec des souches chimères adaptées. La figure 4 présente les résultats obtenus.

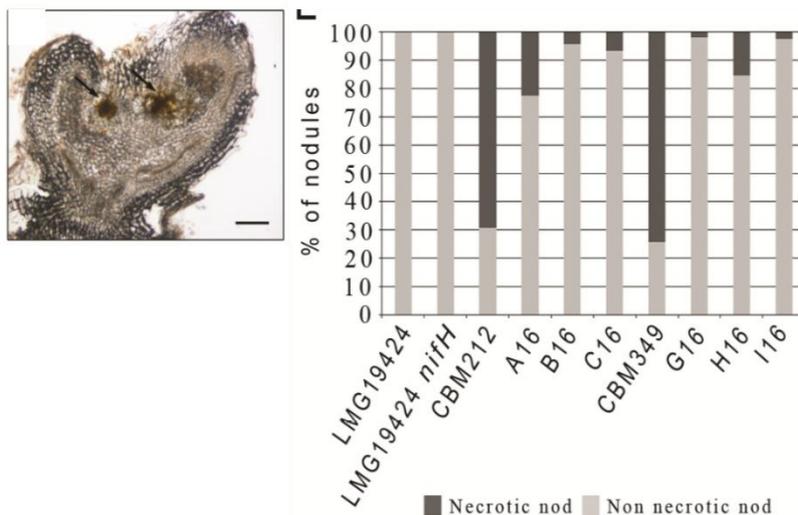


Figure 4. Necrotic area evaluation

Partie II : Travaux dirigés

TD1. Evolution expérimentale

Question 12. Discuter la méthodologie employée ?

Réponse 12. Après récolte des nodosités issues de plantes inoculées avec les différentes souches, des coupes histologiques sont réalisées et le nombre de nodosités nécrotiques et non-nécrotiques est évalué.

Question 13. Pourquoi les auteurs décident-t' ils de réaliser cette expérience ?

Réponse 13. Les auteurs souhaitent voir l'impacte des souches chimères sur la stimulation des réponses immunitaires chez l'hôte.

Question 14. Analyser la figure 4, quelle est votre conclusion ?

Réponse 14. La figure 4 présente le niveau de nécrose au sein de nodosités issus de plantes inoculées avec de bactéries symbiotique, mutant symbiotique non-fixateur, souche chimères mères et leurs descendants. On observe aucune nécrose dans le cas des symbiotiques (sauvage et mutant), alors que les nécroses sont observées chez les chimères avec une plus fort quantité chez les souches mères. Ces résultats indiquent que l'introduction du plasmide symbiotique ne suffit pas à établir une interaction complète, une évolution expérimentale doit être réalisée et elle a pour principale effet une réduction de la stimulation de la réponse immunitaire.

Question 15. Réalisez une synthèse des résultats obtenus ?

Réponse 15. Les étudiants doivent réaliser un travail de groupe afin de restituer les points clés de l'étude. Une note est attribuée en fonction de la qualité de la restitution et intègrent l'évaluation continue.

Partie II : Travaux dirigés

TD2. Caractérisation d'un mutant

Les légumineuses sont des plantes qui ont la capacité d'établir une interaction symbiotique avec des bactéries du sol fixatrices d'azote, les rhizobia. Au cours de cette interaction un nouvel organe racinaire est formé chez l'hôte, la nodosité fixatrice d'azote, au sein de laquelle les bactéries infectent massivement les nodosités. Dans le but de comprendre les mécanismes sous-jacent l'interaction entre la bactérie *Sinorhizobium meliloti* et son hôte *Medicago truncatula*, l'équipe du Pr. Michael K Udvardi a identifié un gène *RSD* impliqué dans le processus symbiotique, ce TD propose de retracer l'étude réalisée.

L'équipe a cherché à identifier un ou des gènes de *M. truncatula* impliqués dans le processus symbiotique par approche de génétique directe.

Question 1. Quel est une approche de génétique directe en recherche ?

Réponse 1. Une approche de génétique directe est basée sur l'identification de mutants obtenus par mutagenèse aléatoire ayant un phénotype en corrélation avec la problématique étudiée. Le(s) gène(s) responsable(s) du phénotype sont identifiés.

Question 2. Proposez une expérience utilisant la génétique directe afin d'identifier des gènes importants au cours de l'interaction symbiotique ?

Réponse 2. Réalisation de mutant à partir de la souche sauvage par mutagenèse aléatoire, puis criblage de mutant d'intérêt en utilisant les plantes WT ou mutées inoculées avec la bactérie WT en condition de carence en azote.

Question 3. Quels sont les phénotypes possibles pour les mutants obtenus ? Quelles sont les fonctions putatives des gènes associés ?

Réponse 3. Pas de variation de phénotype, le gène muté n'est pas responsable. Pas de symbiose, le gène muté est impliqué dans la mise en place de l'interaction. Augmentation de la symbiose, le gène muté est impliqué dans la répression de la symbiose.

L'équipe de recherche a identifié un mutant ayant le phénotype d'intérêt, la figure 1 vous présente le phénotype d'une plante WT et du mutant en condition de carence azotée et en présence de la souche de *S. meliloti*.

Question 4. Analysez la figure 2, quelle est votre conclusion ?

Partie II : Travaux dirigés

TD2. Caractérisation d'un mutant

Réponse 4. La figure présente le phénotype de la plante WT et du mutant en condition de carence en azote et inoculé par la rhizobia. Le mutant présente une réduction de la croissance et des signes de stress par rapport au WT. Donc le mutant est altéré dans sa capacité à établir une interaction symbiotique viable en milieu carencé.

Question 5. Es-que le phénotype du mutant est intéressant pour continuer l'étude ?

Réponse 5. Oui, car le phénotype est un dysfonctionnement symbiotique

Question 6. Proposez trois expériences afin de continuer l'étude ?

Réponse 6. Analyse de la fixation d'azote, analyse histologique des nodosités, identification du gène muté.

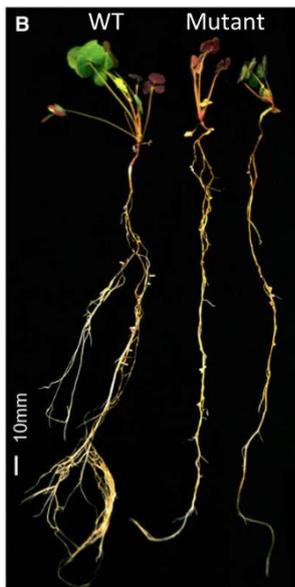


Figure 1. Phénotype d'une plante sauvage (WT) et du mutant en condition de carence azotée et inoculés avec *S. meliloti*

Dans le but d'identifier la cause ayant produit le phénotype du mutant, l'équipe a procédé à des caractérisations phénotypiques et physiologiques, ils ont d'abord observés le phénotype des nodosités WT et mutante (Figure 2a), puis analysé la capacité de fixation d'azote des nodosités (figure 2b), enfin une analyse de la survie intracellulaire des bactéries par live/dead staining est réalisée chez les deux génotypes (figure 2d).

Partie II : Travaux dirigés

TD2. Caractérisation d'un mutant

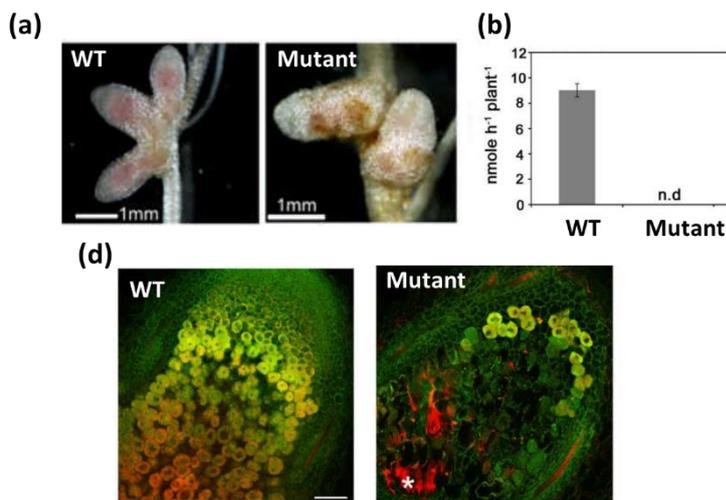


Figure 2. Analyse approfondie du phénotype du mutant, a : observation macroscopique des nodosités, b : analyse de la fixation d'azote, d : analyse de la survie intracellulaire

Question 7. Analysez la figure 2, quelles sont vos conclusions ?

Réponse 7. La figure 2 présente le phénotype de nodosité WT et mutant (a) le niveau de fixation d'azote chez les deux génotypes (b) ainsi que la survie des bactéries au sein des nodosités du WT et du mutant. On observe que les nodosités mutantes sont de couleur marron, la plante ne fixe pas l'azote et le nombre de bactérie vivante est largement réduit. Le mutant forme donc des nodosités modifiées avec une réduction du nombre de bactérie accompagné par une absence de fixation d'azote. Le dysfonctionnement du mutant ne réside pas dans la formation des nodosités, mais plutôt dans leur non-fonctionnalité.

Question 8. Vos observations corrèlent-elles avec celle de la figure 1 ? Commentez votre réponse.

Réponse 8. Oui, car elle confirme l'hypothèse du dysfonctionnement symbiotique,

L'équipe de recherche a voulu identifier le(s) gène(s) muté(s). En sachant que le mutant a été obtenu par mutagenèse aléatoire insertionnelle au transposon *tnt1*.

Question 9. Décrire le type de mutagenèse utilisée ?

Réponse 9. Elle consiste dans l'introduction d'un transposon au sein d'une cellule de plante par technique de transformation génétique, puis une régénération *in vitro* est réalisée afin de créer une nouvelle plante, au cours de la régénération le transposon se multiplie et se fixe de manière aléatoire dans le génome provoquant des milliers de mutations insertionnelles.

Partie II : Travaux dirigés
TD2. Caractérisation d'un mutant

Question 10. Quels autres types de mutagenèse pourrions-nous utiliser ? Comparez avec la mutagenèse au tnt1 ?

Mutagenèses	Points communs	Points de divergences
Ethyle méthyl sulfonate (EMS)	- Mutagenèse aléatoire - Pouvant provoquer un Kncuk out	- Modification d'une base azotée - Agent mutagène chimique - Mutation ponctuelle

Question 11. En sachant que l'élément tnt1 contient une caste de résistance à la kanamycine, veuillez présenter une méthode simple afin d'identifier le gène muté ?

Réponse 11. Le génome du mutant est extrait, puis fragmenté. Les fragments sont insérés dans des vecteurs et introduits dans des bactéries qui sont mises en culture sur un milieu contenant de la Kanamycine. Seules les bactéries possédant le fragment contenant le gène de résistance peuvent se développer. Puis l'insert est séquencé afin d'identifier les séquences bordant le transposon.

L'équipe de recherche a identifié un gène potentiellement responsable (*MtRSD ; Regulatory of Symbiosom Differentiation*), la figure 3 présente la séquence de la protéine.

Partie II : Travaux dirigés TD2. Caractérisation d'un mutant

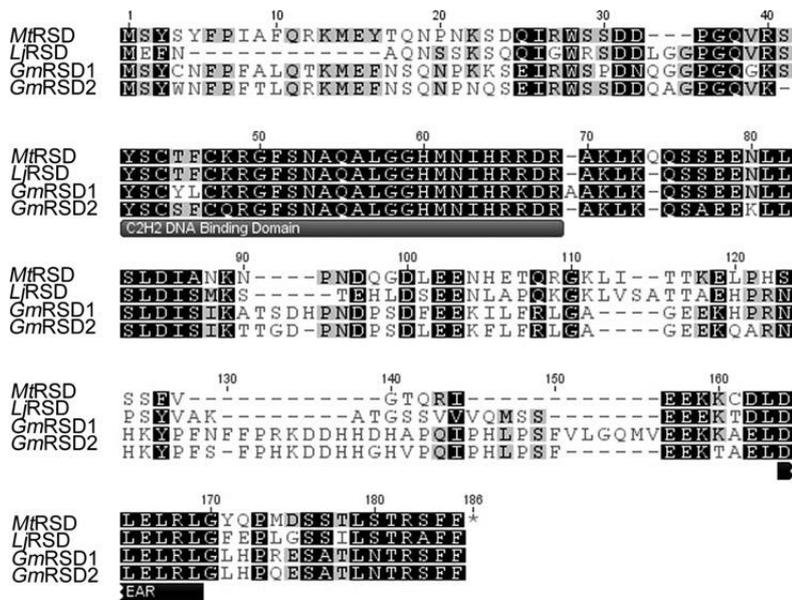


Figure 3. Séquence du gène RSD chez *M. truncatula*, *L. japonicus* et *G. max*.

Question 12. Pourquoi plusieurs espèces sont présentées ?

Réponse 12. Cela permet d'identifier les régions divergentes et conservées entre le gène étudié et des orthologues proche.

Question 13. Analysez la figure 3 ? En sache que le domaine C2H2 est impliqué dans la liaison des protéines à l'ADN, Proposez une hypothèse quant à la fonction putative du gène ? Discutez les limites de votre hypothèse ?

Réponse 13. On observe des régions divergentes et des séquences conservées, ces dernières sont C2H2 et EAR. Vu que le domaine C2H2 est impliqué dans la liaison à l'ADN ceci indique que la protéine est potentiellement un facteur de transcription. La protéine ne possède pas de séquence d'import dans le noyau (NLS) il faut donc confirmer expérimentalement cette hypothèse.

Question 14. Proposez une approche afin de valider le gène candidat ?

Réponse 14. Un test de complémentation fonctionnelle peut valider le candidat, le gène sous sa forme sauvage est introduit dans la plante mutante, s'il y'a une restauration du phénotype, ceci confirme l'implication du gène.

Partie II : Travaux dirigés
TD2. Caractérisation d'un mutant

Question 15. A la lumière de vos observations et de vos conclusions, proposez une hypothèse sur le mécanisme par lequel fonctionne *MtRSD* pour maintenir la symbiose effective

Réponse 15. Les étudiants doivent réaliser un travail de groupe afin de restituer les points clés de l'étude. Une note est attribuée en fonction de la qualité de la restitution et intègre l'évaluation continue

Partie II : Travaux dirigés

TD3 clonage moléculaire

Le clonage d'une séquence d'ADN est une étape souvent cruciale en recherche.

Ce Travail dirigé se propose d'aborder l'un des clonages classiquement utilisés, le clonage par digestion et ligation enzymatique.

Durant une étude ciblant la caractérisation du facteur de transcription de type C2H2, *RSD* de l'espèce *medicago truncatula* (voir TD : 2) les chercheurs ont voulu cloner le gène *MtRSD*.

Ce TD se propose de retracer l'étape de clonage du gène. Vous trouverez en pièces-jointe le schéma du vecteur que nous proposons d'utiliser (pUC19 ; figure 1) ainsi que de la séquence partielle du vecteur contenant la séquence du site de clonage multiple. Vous disposez aussi de la séquence du CDS (Coding DNA Sequence) de *MtRSD* (figure 2).

Question :

1- Proposer une approche/protocole afin de cloner le gène *MtRSD*.

Réponse 1. Des amorces qui possèdent des extrémités sont utilisées pour amplifier par PCR le gène *RSD* à partir de l'ADN de la plante WT *Medicago truncatula*. Une co-digestion des fragments PCR et du vecteur est réalisée. Après purification, le vecteur linéarisé ainsi que le fragment PCR digéré sont mis ensemble en présence d'une ligase ce qui provoque la recircularisation du vecteur sur l'insert. Le vecteur recombiné est transféré dans une bactérie qui est sélectionnée sur un milieu contenant l'antibiotique pour lequel le vecteur possède la résistance.

2- D'après vous, à partir de quels matériels génétiques nous pouvons démarrer ? Commenter l'avantage et les limites de chaque type de matériels.

Réponse 2. Méthode 1) L'ADN génomique permet d'obtenir le gène entier (exon + intron), mais ne permet pas l'expression *in vitro* du gène dans un système hétérologue (exemple une bactérie). Méthode 2) L'ADN complémentaire est généré après reverse transcription de l'ARNm codant de *RSD*, on peut exprimer la protéine *RSD* dans un système hétérologue. L'inconvénient c'est que l'on ne possède pas l'ensemble du gène et donc on ne peut pas vérifier s'il y a un épissage alternatif.

3- à quoi correspond le terme CDS ?

Partie II : Travaux dirigés TD3 clonage moléculaire

Réponse 3. CDS : coding direct sequence, cela correspond à la séquence codant la protéine (exon).

4- Le vecteur utilisé est pUC19, en se basant sur sa séquence quel est l'avantage de ce vecteur ? Et quels sont ces limites ?

Réponse 4. Le vecteur pUC19 présente un site de multicolonage au sein du gène lacZ, ce qui permet de réaliser un crible bleu-blanc. Ce plasmide ne possède pas de promoteur permettant la réalisation d'une expression *in vitro* du gène. Le gène de résistance correspond à celui de l'ampicilline, un antibiotique qui se dégrade dans le milieu, ce qui peut provoquer l'apparition de faux positifs.

5- En se basant sur la séquence du vecteur et celle du CDS, définir les enzymes de restrictions convenables pour le clonage de *MtRSD* dans pUC19. Commenter votre choix.

Réponse 5. Il faut prendre deux enzymes coupants du site de multi clonage, au niveau du gène RSD ; la première enzyme coupe avant le start de transcription du CDS et la seconde après le stop de transcription du CDS

6- Comment vérifier que la digestion enzymatique s'est bien passée ?

Réponse 6. On peut analyser l'activité lacZ en ajoutant le substrat de l'enzyme (X Gal) ou redigérer le vecteur recombinant avec une enzyme qui coupe une fois dans l'insert et une fois dans le vecteur. Si l'insert est présent deux bandes apparaissent après migration, mais s'il est absent seulement une seule bande sera présente.

7- Une fois le clonage terminé, décrivez le(s) critère(s) de sélection(s) appliqué(s).

Réponse 7. Résistance à l'ampicilline, l'expression de l'enzyme lacZ.

Les chercheurs souhaitent cloner le gène sous le contrôle d'un promoteur viral fort (CMV_35S), la séquence du promoteur vous est détaillée dans la figure 4.

8- En prenons en considération ces nouveaux éléments, veuillez répondre à la question 6 et 7.

Réponse 8. Trois enzymes doivent être appliquées, la première coupe au site de multi clonage et en amont du 35S, la seconde coupe en aval du 35S et avant le start de

Partie II : Travaux dirigés

TD3 clonage moléculaire

transcription du CDS et la dernière coupe au site de multicolonnage et en aval du stop du CDS. Ainsi on obtient une construction : 35S ::CDS.

Le critère de sélection utilisé est une enzyme capable de coupe une fois dans le vecteur, une fois dans le 35S et une fois dans le CDS.

En fonction des cas les possibilités sont :

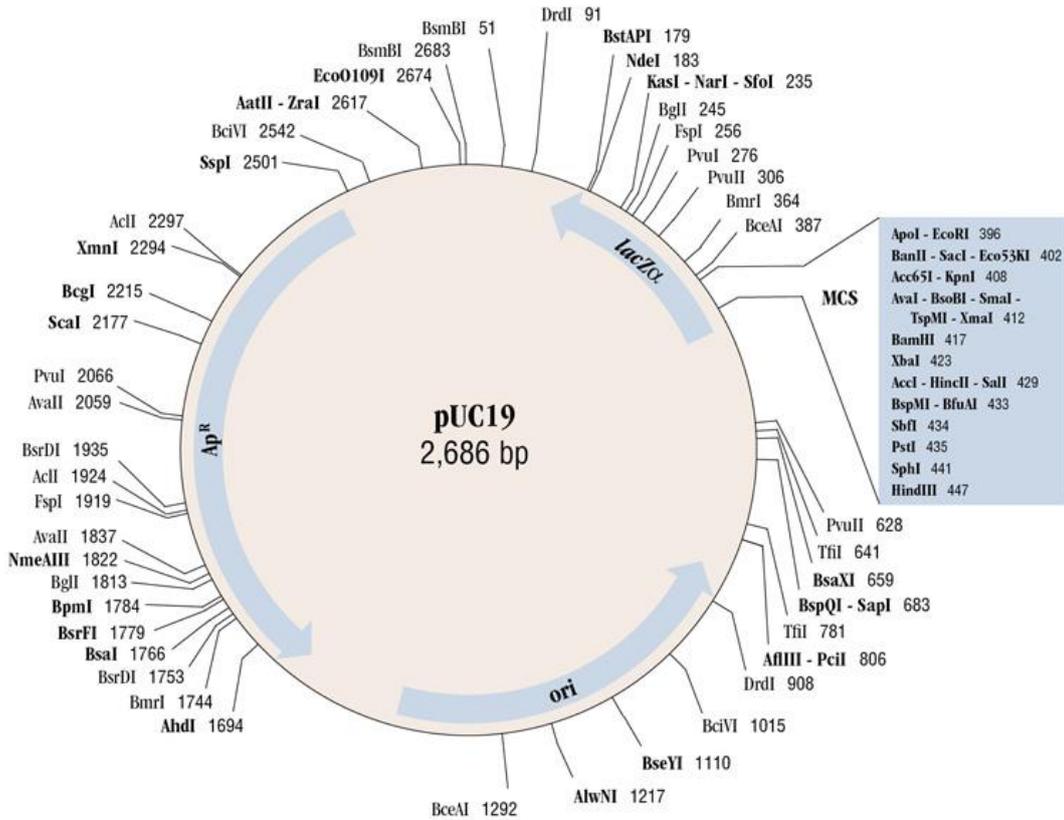
Clonage réussi : 3 bandes

Seulement 35S intégré : 2 bandes

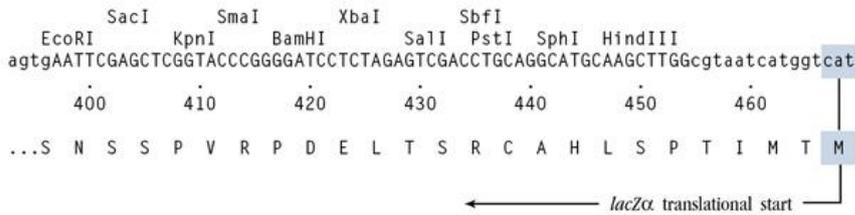
Seulement CDS intégré : 2 bandes

Clonage vide : 1 bande

Partie II : Travaux dirigés TD3 clonage moléculaire



pUC19 MCS



Partie II : Travaux dirigés
TD3 clonage moléculaire

>Cloning vector pUC19 partial sequence

LacZ

```
cgtaaggagaaaataccgcatcaggcgccattcgccattcaggctgcgcaactgttgg
gaagggcgatcgggtgcgggcctcttcgctattacgccagctggcgaaagggggatgt
gctgcaaggcgattaagttgggtaacgccagggttttccagtcacgacgttgtaaaa
cgacggccagtgaattcgagctcgggtaccggggatcctctagagtcgacctgcaggc
atgcaagcttggcgtaatcatggcatagctgtttcctgtgtgaaattgttatccgctca
caattccacacaacatacagagccggaagcataaagtgtaaagcctgggggtgcctaat
gagtgagctaactcacattaattgcgttgcgctcactgccgctttccagtcgggaaac
ctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgta
ttgggcgc
```

MCS

← LacZ start

Partie II : Travaux dirigés
TD3 clonage moléculaire

>CMV_35S

tgaattcgagctcgggtaccggggatcctctagagtcgacctgcaggcatgcaagcttcagaagaccaaagggc
aattgaaacctttcagaaaagggttaatttcgggaagccttcttgattcattgccatcgatctgtcactttatca
aaaggacagtagaaaaggaaggtggcgcctaccaatgccatcaatgcgataaaggaaaggctatcgttgaag
ataagcctgccgacagtgggtccagcagacagctcgggcccaccaccacgaaagaaaacgtggaaaaagaag
atgccgtttccactacgacaccttcaaagcaagtggattgatgttatatctcactgacgtaagggatgacgcac
aatcccactatccttcgcaagacccttctctatataagggaagttcatttcattggagaggacacgctgaaatca
ccagtctgaattcgagctcgggtaccggggatcctctagagtcgacctgcaggcatgcaagct

CAAT regulatory
sequence
TATA box

Partie II : Travaux dirigés
TD3 clonage moléculaire

>JX307863.1 Medicago truncatula RSD (RSD) mRNA, complete cds

Start

tgattcgagctcgggtacccggggatcctctagagtcgacctgcaggcatgcaagctATGTCTTATAGCTA
TTTCCCATTGCTTTCCAAAGAAAGATGGAATACACCCAAAACCCTAATAAATCTGAC
CAAATAAGATGGAGCTCTGACGATCCGGGTCAAGTTAGGTCTTATTCTTGACTTTTT
GCAAGAGAGGATTCTCTAATGCACAAGCATTAGGAGGCCACATGAATATCCATAGAA
GAGATAGAGCCAAGCTCAAACAACAATCTTCCGAAGAAAATTTGCTTTCTTTGGACAT
CGCAAACAAAAACCCTAATGATCAAGCAGATTCGGAAGAGAATCACGAAACTCAAAG
GGGTAAGTTAATTACTACAAAAGAACTTCCTCATGGTTCTTCATTTGTTGGAACTCAAA
GGATTGAAGAGAAGAAATGTGATTTAGACCTTGAGCTTAGGTTAGGTTATCAGCCAA
TGGATTCATCAACATTGAGCACTCGATCATTTTTCTGACCTTGAAAATTGGCAAAT
ATAGGTCACCTTTTTTTTTTTTCGTACCTACCTCTTCTTCCACTTCCATTTGAAAGTACAA
CTCGATTTTTTTTATATAAACATAAAGGTACATGTATTTGATCTATACCTTTATGCTTAT
GTAAAATACATGATGGAGTAAATAAATGGTGAAAATTTAATTTGGTGTGTTTATTAG
TATATCAAAAAATTTACATTTTCAATGGCATGTTTCCTACATTATTTAATAAATTTTTGA
TATCTACAATTATTGTGGGTAGTCAAATTTATTCAATAATGTAATTATATATTATTGAA
CGAAAGTGTGAGAGTTTTATATACTGGCGCTACATATATAAATTAATATCACGTTAG
TGCAATCGTGGGAACACTACCAAAAACATATATGCACTTTTATACATGTTAAAATTACT
TTTATTTTAGAACCGAAAAAAATTATTTGTATTTTTATACCAAAAAATATATGTTTTTAT
TTTTCTGCCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAtgaattcgagctcgggtaccc
ggggatcctctagagtcgacctgcaggcatgcaagct

Stop

Partie II : Travaux dirigés TD4 génomique

Les microARN sont des petits ARNs non codants impliqués dans la régulation de l'expression des gènes. L'étude des microARNs devient de plus en plus importante et un nombre croissant d'études sont publiées. Ce TD se propose de retracer l'une des études menée par l'équipe du Dr. Martins Crespi.

Question 1. Ci-dessous l'abstract complet de l'étude, veuillez déterminer les objectifs fixés et la méthode employée pour répondre aux objectifs ainsi que les principaux résultats ?

Posttranscriptional regulation of a variety of mRNAs by small 21- to 24-nucleotide RNAs, notably the microRNAs (miRNAs), is emerging as a novel developmental mechanism. In legumes like the model *Medicago truncatula*, roots are able to develop a de novo meristem through the symbiotic interaction with nitrogen-fixing rhizobia. We used deep sequencing of small RNAs from root apices and nodules of *M. truncatula* to identify 100 novel candidate miRNAs encoded by 265 hairpin precursors. New atypical precursor classes producing only specific 21- and 24-nucleotide small RNAs were found. Statistical analysis on sequencing reads abundance revealed specific miRNA isoforms in a same family showing contrasting expression patterns between nodules and root apices. The differentially expressed conserved and nonconserved miRNAs may target a large variety of mRNAs. In root nodules, which show diverse cell types ranging from a persistent meristem to a fully differentiated central region, we discovered miRNAs spatially enriched in nodule meristematic tissues, vascular bundles, and bacterial infection zones using in situ hybridization. Spatial regulation of miRNAs may determine specialization of regulatory RNA networks in plant differentiation processes, such as root nodule formation.

Réponses 1. Les petits ARN ou miRNA de 21 à 24 nucléotides régulent post traductionnellement l'expression des gènes. L'équipe a cherché à identifier les miRNA impliquées dans les régulations du méristème des nodosités et celle de l'apex des racines. 100 nouveaux candidats ont été identifiés et un nouveau précurseur atypique spécifique des 21 et 24 nucléotides petits ARN a été trouvé. Certains petits ARNs sont enrichis dans des parties spécifiques de la nodosité.

La figure 1 présente la méthodologie employée par l'équipe de recherche afin d'identifier les microRNAs.

Partie II : Travaux dirigés TD4 génomique

Question 2. Veuillez expliquer la figure 1 ?

Réponse 2. La figure présente les étapes expérimentales réalisées, les méristèmes de racines traité avec ou sans NaCl sont isolés, en plus des nodosités entières. Les ARNs extraits sont séquencés et le traitement bio-informatique est réalisé afin d'identifier les miRNAs produits.

Question 3. Quels conditions peuvent être ajouté à l'étude pour l'améliorée ?

Réponse 3. Une condition manque c'est l'utilisation de nodosités traitées au NaCl afin de vérifier si la physiologie des régulations impliquant des miRNAs est identique dans cette situation entre les méristèmes des racines et ceux des nodosités.

Question 4. Quels précaution doit prendre l'équipe au moment de la production des échantillons ?

Réponse 4. Les ARNs étant des molécules très sensibles, il faut donc travailler avec des équipements décontaminés d'ARNase. Les échantillons doivent être générés en même temps afin de réduire l'hétérogénéité des résultats.

Question 5. Quels sont les filtres bio-informatique que l'on doit appliquer afin d'identifier les miRNAs ?

Réponse 5. Le filtre doit exclure tous les ARNs ayant des tailles inférieures à 21 nucléotides et supérieur 24 à nucléotides afin de se focaliser uniquement sur cette famille.

Question 6. Que veut dire 21 et 24 nucléotides ?

Réponse 6. Le numéro indique le nombre de nucléotides, par ailleurs la séquence peut varier et la quantité relative d'un type (population possédant une séquence identique) de 21 ou 24 nucléotides peut également changer.

Partie II : Travaux dirigés TD4 génomique

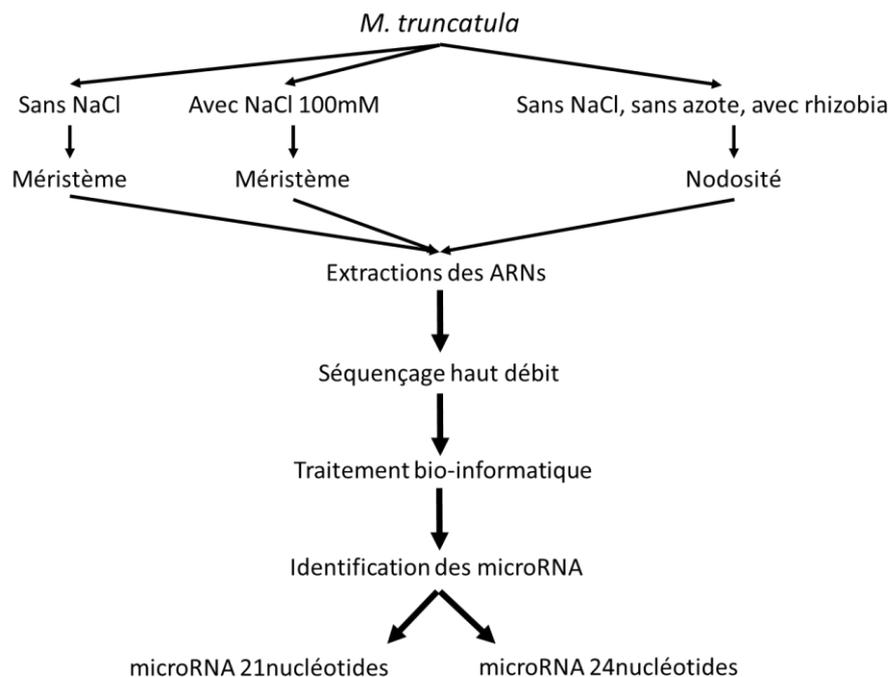


Figure 1. Méthodologie simplifiée d'identification des microRNAs de *M. truncatula*

L'équipe a évalué la densité et la diversité des microARNs ainsi que l'expression de gènes codant des enzymes qui produisent les microRNA (DCL1 & 2), la figure 2 présente les résultats obtenus.

Question 7. Comment les auteurs évaluent la diversité des miRNA ?

Réponse 7. Il compare les séquences obtenues après le séquençage haut débit entre les 21 ou 24 nucléotides.

Question 8. Analyser la figure 2? Quelle est votre conclusion par rapport à la diversité et l'abondance des miRNAs entre la nodosité et l'apex racinaire ?

Réponse 8. La figure 2 présente les résultats obtenus pour l'abondance et la diversité des petits ARNs de taille de 15 à 30 nt. On observe que les petits ARN de 24 nucléotides sont les plus abondants et les plus diversifiés dans la nodosité et l'apex racinaire, alors que ceux de 21 nucléotides sont abondants et diversifiés essentiellement dans l'apex racinaire. Donc les miRNA les plus abondants et les plus diversifiés sont ceux de 24 nucléotides pour les nodosités et 21/24 nucléotides pour l'apex racinaire.

Question 9. Présentez une méthode afin d'analyser l'expression de DCL 1 & 2 ?

Partie II : Travaux dirigés TD4 génomique

Réponse 9. Une méthode est la Reverse transcription PCR quantitative qui est dérivée de la PCR classique et qui permet de quantifier les ADNc d'un gène donnée en temps réel.

Question 10. A quoi correspond RBP1 ?

Réponse 10. RBP1 est un gène de ménage constitutivement exprimé, il sert de référence afin de normaliser les données d'expression pour les gènes d'intérêts.

Question 11. Analysez les résultats d'expression de DCL1 & 2 ? Quelle est votre conclusion ?

Réponse 11. DCL1 est exprimé dans les apex racinaires et dans les nodosités, alors que DCL3 n'est exprimé essentiellement que dans les nodosités. Le gène de référence RBP1 présente une expression stable, signe que les échantillons non pas subits de dégradation d'ARN, les variations observées sont réellement dû à l'expression des gènes et non à une dégradation des échantillons. DCL1 est donc exprimé dans les nodosités et l'apex racinaire, alors que DCL3 est essentiellement exprimé dans les nodosités.

Question 12. En prenons en considération les résultats de densité et de diversité des miRNA et l'expression des DCL, veuillez présenter une hypothèse pouvant expliquer les différences observées entre les nodosités et les apex racinaire ?

Réponse 12. Il est possible que la présence de DCL3 dans les nodosités permet la production de miRNA 24 nucléotides, alors que DCL1 permet la production des miRNA 21 nucléotides Il est possible que les deux enzymes sont concurrentes sur les même précurseur pour la production des petit ARNs. La réduction d'expression de DCL3 permet d'augmenter l'effet de DCL1 par réduction de compétition et donc d'augmenter la quantité des miRNA 21 nucléotides.

Partie II : Travaux dirigés TD4 génomique

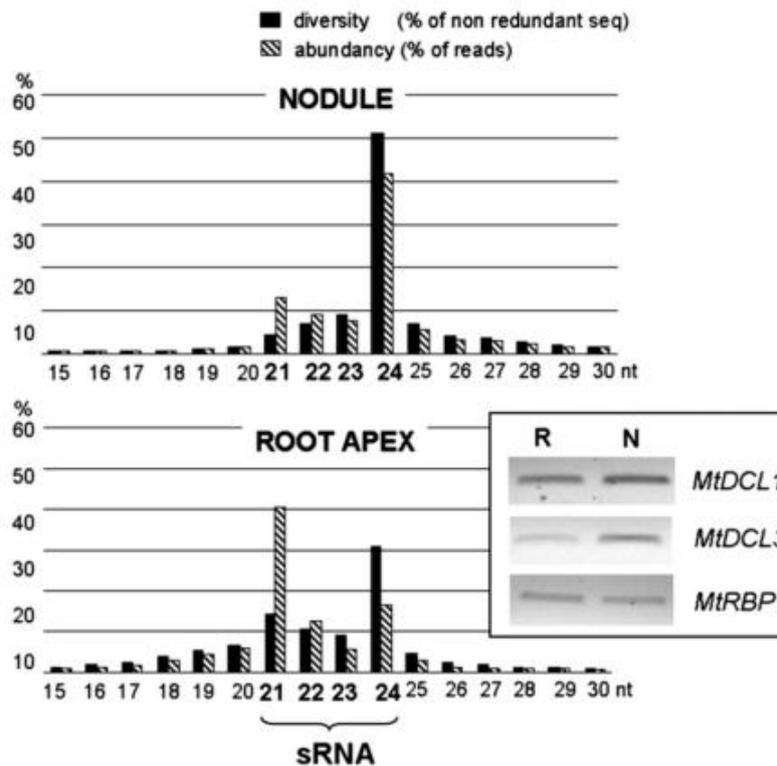


Figure 2. The diversity (percentage of total non-redundant sequences, black bars) and abundance (percentage of total reads, dashed bars) of 15 to 30 nucleotide RNAs are represented for nodule (top panel) and root apex (bottom panel) libraries. Inset shows expression profiles of the *M. truncatula* DCL1-like and DCL3-like genes in nodules and root apices. After reverse-transcription of DNase-treated total RNA from mature nodules (N) and root tips (R), PCR experiments were performed using primers specific for *M. truncatula* DCL1- and DCL3-like genes and for the Mt RBP1 gene as constitutive control.

La figure 3 présente la distribution des miRNA dans le génome de *M. truncatula*.

Question 13. Comment les auteurs réalisent la localisation des miRNA dans le génome de *M. truncatula* ?

Réponse 13. Une méthode afin d'identifier les miRNA dans le génome est de comparer leurs séquences avec celle du génome de la plante modèle. Cela permettra de révéler les régions codant les miRNA.

Question 14. Analysez la figure 3 ? Quelle est votre conclusion sur la localisation des miRNA dans le génome de *M. truncatula* ?

Partie II : Travaux dirigés TD4 génomique

Réponse 14. La figure 3 présente la localisation des petits ARN dans le génome de *M. truncatula*. On observe une hétérogénéité de la répartition des miRNA dans le génome. Les miRNA 21/22 nucléotides sont plus largement réparties que les 24 nucléotides. Ces résultats montrent que les miRNA sont répartis dans l'ensemble du génome et ne sont pas concentrés sur une région donnée.

Question 15. Présentez une synthèse des résultats obtenus ?

Réponse 15. Les étudiants doivent réaliser un travail de groupe afin de restituer les points clés de l'étude. Une note est attribuée en fonction de la qualité de la restitution et intègre l'évaluation continue

Partie II : Travaux dirigés
TD4 génomique

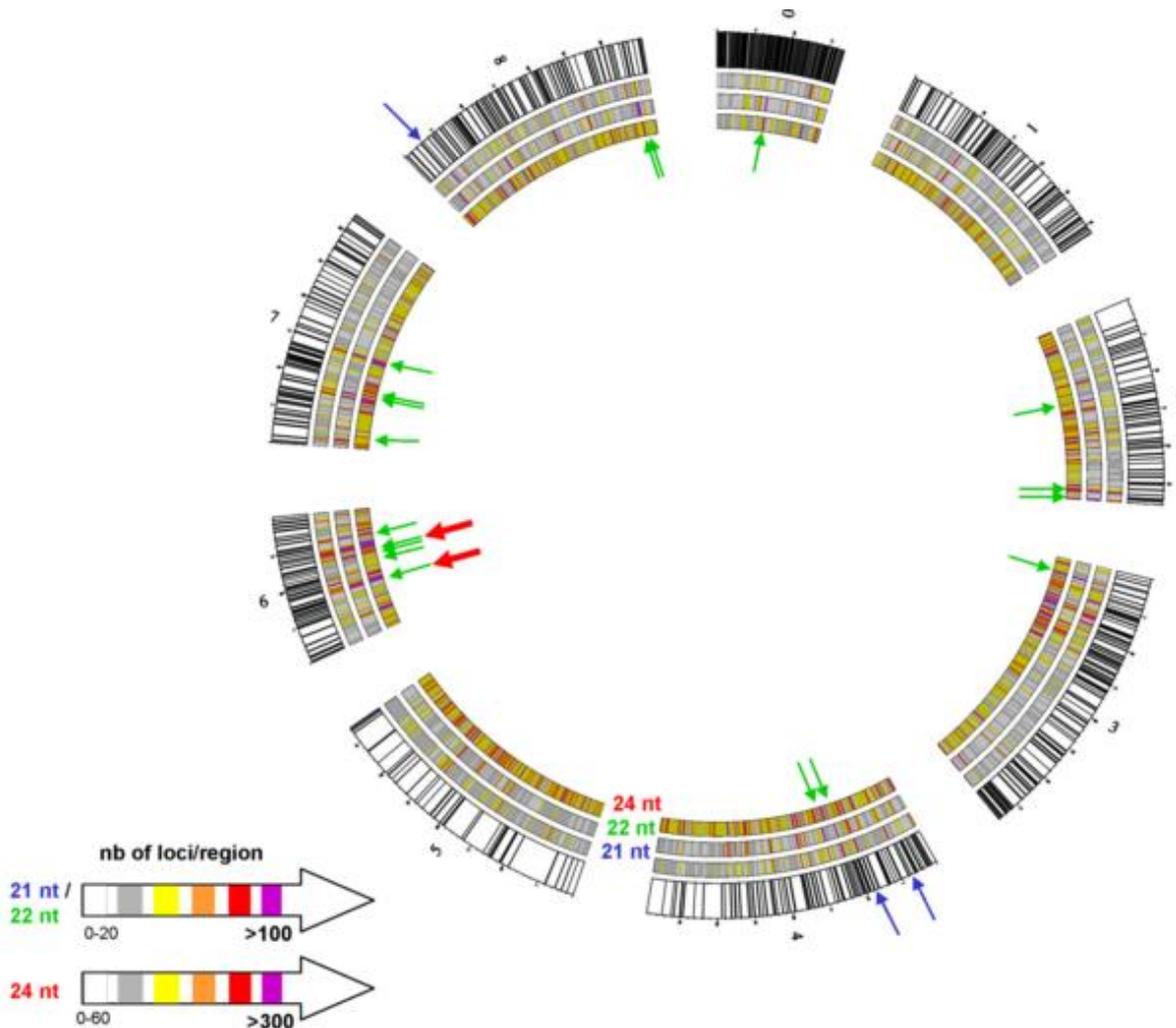


Figure 3. Non redundant sequences from 21-, 22-, and 24-nucleotide RNAs were independently mapped on the eight *M. truncatula* chromosomes (plus chromosome 0). Chromosomes are represented by the outermost circle with white bars corresponding to the 300-kb nonoverlapping genomic regions and black bars to the gaps (at least 10 kb) in the *M. truncatula* genome version 2. The number of sRNA loci per region is represented in successive circles (24 nucleotides being the innermost one) and visualized using a color scale from white (minimal density) to purple (maximal density). For 21- to 22-nucleotide RNAs, the density increment is based on 20 loci/region, whereas for the 24-nucleotide RNAs, which are much more diverse, the increment is based on 60 loci/region. Hot spots of 21 and 22 nucleotides are represented by blue and green arrows, respectively. In chromosome 6, two clusters of three consecutive regions show very high density of 24-nucleotide loci (red arrows). This representation has been generated with CIRCOS (<http://mkweb.bcgsc.ca/circos/>)

Partie II : Travaux dirigés

TD5 Transcriptome et analyses fonctionnelle

Les équipes du Dr. Peter Mergaert et du Pr. Eva Kondorosi sont parmi les premières à avoir décrit les mécanismes impliqués dans le phénomène de différenciation des rhizobia au sein des nodosités des légumineuses via l'effet de peptides antimicrobiens, les NCRs (Nodule Cystein Rich). Ce TD propose de retracer l'étude des NCRs.

Les auteurs ont identifiés 42 gènes induits au cours du développement nodulaire chez *M. truncatula*, les gènes identifiés ne présentent aucune homologie avec des séquences connues. Une analyse southern blot a été réalisée sur les gènes cibles. La figure 1 présente les résultats obtenus.

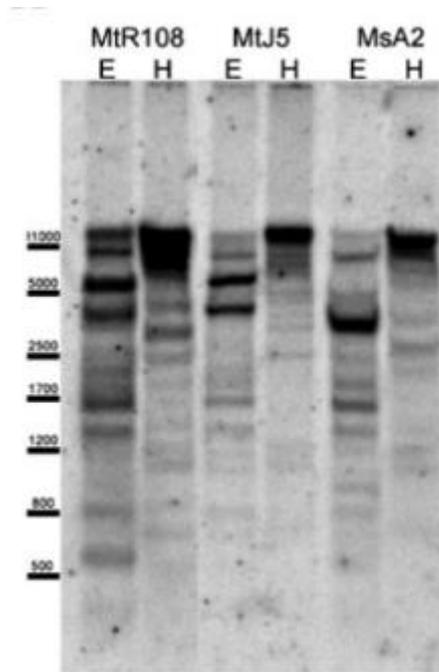


Figure 1. Southern hybridization of genomic DNA digested with EcoRI (E) or HindIII (H) from *M. truncatula* and alfalfa (*Medicago sativa*) with the NCR001 probe revealing multiple hybridizing bands. MtR108, *M. truncatula* line R108; MtJ5, *M. truncatula* line Jemalong J5; MsA2, alfalfa subsp. *varia* line A2. The HindIII digestion was partial

Question 1. Pour quel raison les auteurs décident-ils de réaliser un southern blot ?

Réponse 1. Les auteurs décident de réaliser cette expérience afin de déterminer le nombre relatif de copies des NCRs dans le génome de plusieurs écotypes de *M. truncatula*

Question 2. Décrire l'expérience réalisée ?

Réponse 2. L'ADN génomique des plantes est extrait, une digestion partielle est réalisée puis les acides nucléiques sont séparés sur gel d'agarose par électrophorèse. A l'issue de la réaction, les produits sont transférés sur une membrane de nitrocellulose sur laquelle les sites non-spécifiques sont bloqués. Puis une sonde marquée complémentaire de la

Partie II : Travaux dirigés TD5 Transcriptome et analyses fonctionnelle

séquence du gène *NCR001* est ajoutée. La sonde s'hybride sur les régions complémentaires, le marquage est par la suite révélé.

Question 3. Analysez la figure 1, quel est votre conclusion ?

Réponse 3. On observe plusieurs bandes dans le cas des différents écotypes dans les deux conditions de digestions. Ces résultats indiquent que le *NCR001* est présent chez les trois écotypes testés et qu'il existe en plusieurs copies dans les génomes des plantes.

L'analyse de la séquence de gènes a révélée la présence d'un grand nombre de résidus cystéines, les auteurs ont décidés de nommer les gènes en question NCRs (Nodule Cystein Rich). Le génome de *M. truncatula* présente 334 gènes codants des NCRs. Une analyse d'expression des gènes correspondant aux NCRs a été réalisée par puce à ADN. La figure 2 présente les résultats obtenus.

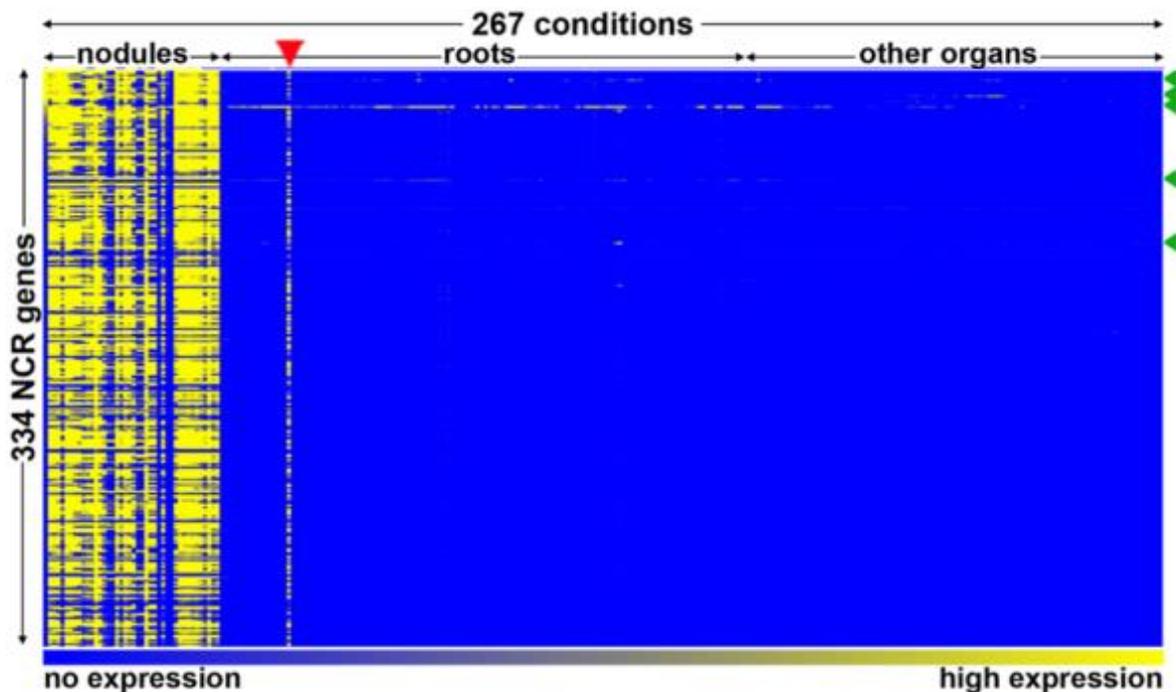


Figure 2. Heat map of NCR gene expression in the MtGEA compendium. The heat map shows the expression of 334 NCR genes (rows) in 267 experimental conditions (columns). Experiments are ordered as indicated above the columns. The color scale bar indicates the expression level from background level (blue) to maximum level (yellow). The red arrowhead indicates the mycorrhizal sample that is contaminated with nodules.

Partie II : Travaux dirigés

TD5 Transcriptome et analyses fonctionnelle

Question 4. Décrivez l'expérience permettant d'obtenir le résultat de la figure 2 ?

Réponse 4. Les ARNs pour différentes conditions sont extraits et reverse transcrits en ADNc. Ces derniers sont marqués par des sondes fluorescentes et hybridés sur une puce à ADN qui permet d'analyser l'expression de plusieurs milliers de gènes en parallèle. Les valeurs de fluorescence sont évaluées pour chaque gène et le niveau d'expression relative est déterminé.

Question 5. Analysé la figure 2, Quel est votre conclusion ?

Réponse 5. La figure 2 présent l'expression relative de la famille de gènes NCRs dans différentes conditions. On observe que les gènes sont spécifiquement exprimés dans les nodosités ce qui indique que ces gènes sont potentiellement impliqués dans le fonctionnement de cette organe.

Question 6. Les auteurs ont par la suite décidés d'étudier le rôle des NCRs dans la régulation de la symbiose, proposez une série d'expériences permettant de comprendre les mécanismes associés aux NCRs ?

Réponse 6. On peut réaliser des expériences de RNA silencing sur les NCRs où isoler des mutants de ces gènes. Le phénotype de plantes en condition d'interaction symbiotique doit être analysé afin de vérifier le rôle des gènes au cours de la symbiose.

En Hongrie, l'équipe du Dr. Peter Kalo, a travaillé sur le mutant de *M. truncatula dnf7* isolé par l'équipe du Pr. Charon Long à Stanford (USA). Le mutant est déficient dans la fixation d'azote, la figure 3 présente l'analyse de l'infection des nodosités par les rhizobia.

Partie II : Travaux dirigés
TD5 Transcriptome et analyses fonctionnelle

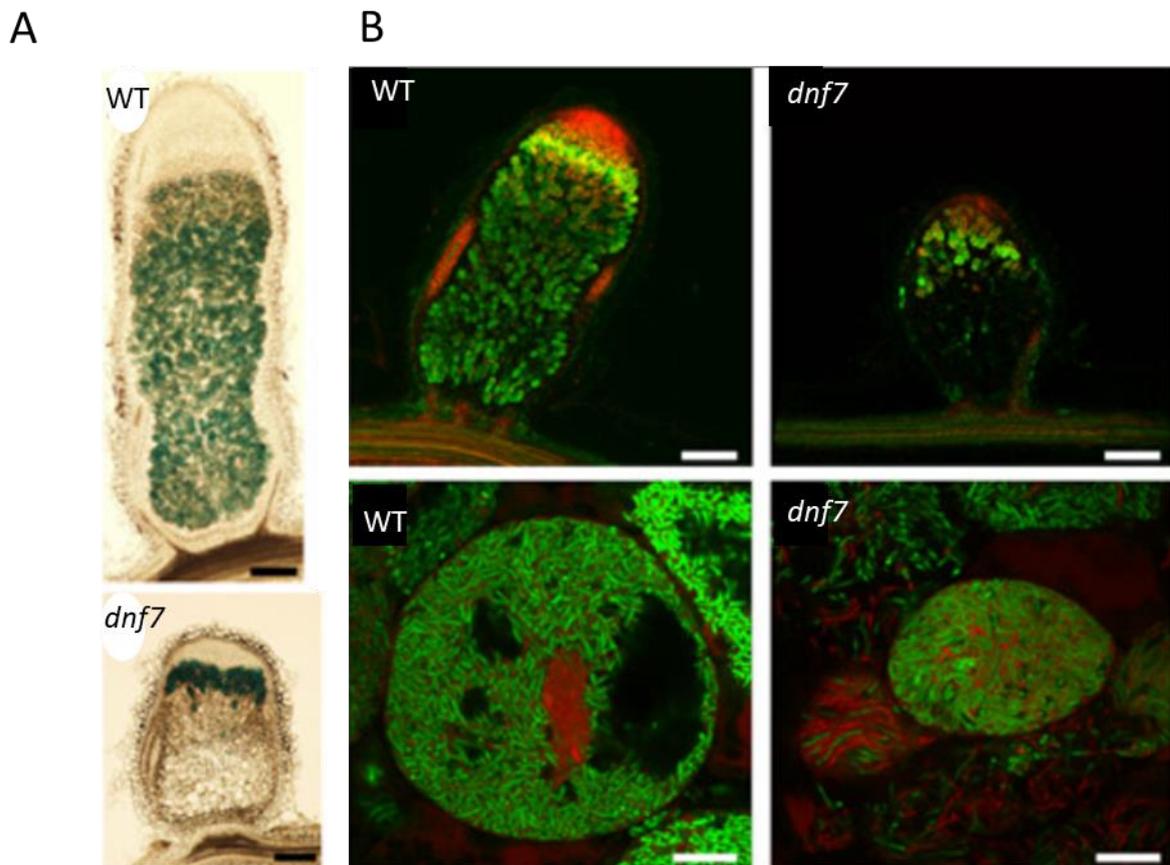


Figure 3. (A) Symbiotic cells do not persist in the nitrogen fixation zone of *dnf7-2* nodules. β -galactosidase activity was detected histochemically in sections of wild-type (WT) and *dnf7* nodules. (B) Live/dead staining in WT and *dnf7* nodules sections.

Question 7. Expliquez la méthodologie employée permettant d'obtenir les résultats de la figure 3 ?

Réponse 7. (a) Des rhizobia exprimant constitutivement la β -galactosidase sont inoculées aux plantes, une fois les nodosités formées, une coupe histologique est réalisée sur ces organes, puis l'activité enzymatique est analysée. (b) les plantes sont inoculées avec des rhizobia, les nodosités sont disséquées, puis une coloration live/dead staining est réalisée, elle révèle les bactéries mortes en rouge grâce à l'iodure de propidium et les bactéries vivantes en vert grâce au CYTO13/CYTO9.

Question 8. Analyser la figure 3, quelle est votre conclusion ?

Partie II : Travaux dirigés

TD5 Transcriptome et analyses fonctionnelle

Réponse 8. Dans la figure A on observe une réduction de la coloration chez le mutant par rapport au WT, ce qui indique que l'infection des tissus de la plante *dnf7* est réduite. Dans la figure B ce résultat est confirmé car les nodosités du mutant présentes un nombre de bactéries mortes plus élevé que chez le WT. Donc le mutant *dnf7* forme des nodosités, mais dans lesquelles les bactéries n'arrivent pas à survivre.

Question 9. En sachant que le gène muté chez *dnf7* correspond au NCR169, qu'elle est votre conclusion quant au rôle de ce gène dans le phénomène symbiotique ?

Réponse 9. Les NCRs sont impliqués dans la symbiose, notamment en prévenant la mort précoce des rhizobia au sein des cellules de l'hôte.

Question 10. L'analyse de l'adressage des NCRs au sein des cellules symbiotiques a été réalisée, proposez 3 expériences permettant d'identifier la localisation des NCRs au sein des cellules symbiotique ?

Réponse 10. 1) fusion traductionnelle entre un NCR et la GFP exprimée dans la plante. Les nodosités formées sont récoltées et disséquées, la fluorescence de la GFP reflète l'adressage de la protéine. 2) Immunohistochimie : Les nodosités de la plante WT sont fixées et un anticorps marqué spécifique du NCR étudié est appliqué sur les coupes de nodosité. 3) Les nodosités de la plante WT sont broyées, les composants du broya sont fractionnés et une identification par spectrométrie de masse est réalisée sur les extraits protéiques des différentes fractions, Les NCR sont ainsi détectés.

La figure 4 présente les résultats de la localisation des NCRs ainsi que la morphologie des rhizoïdes en culture ou issue de cellules symbiotique.

Partie II : Travaux dirigés
TD5 Transcriptome et analyses fonctionnelle

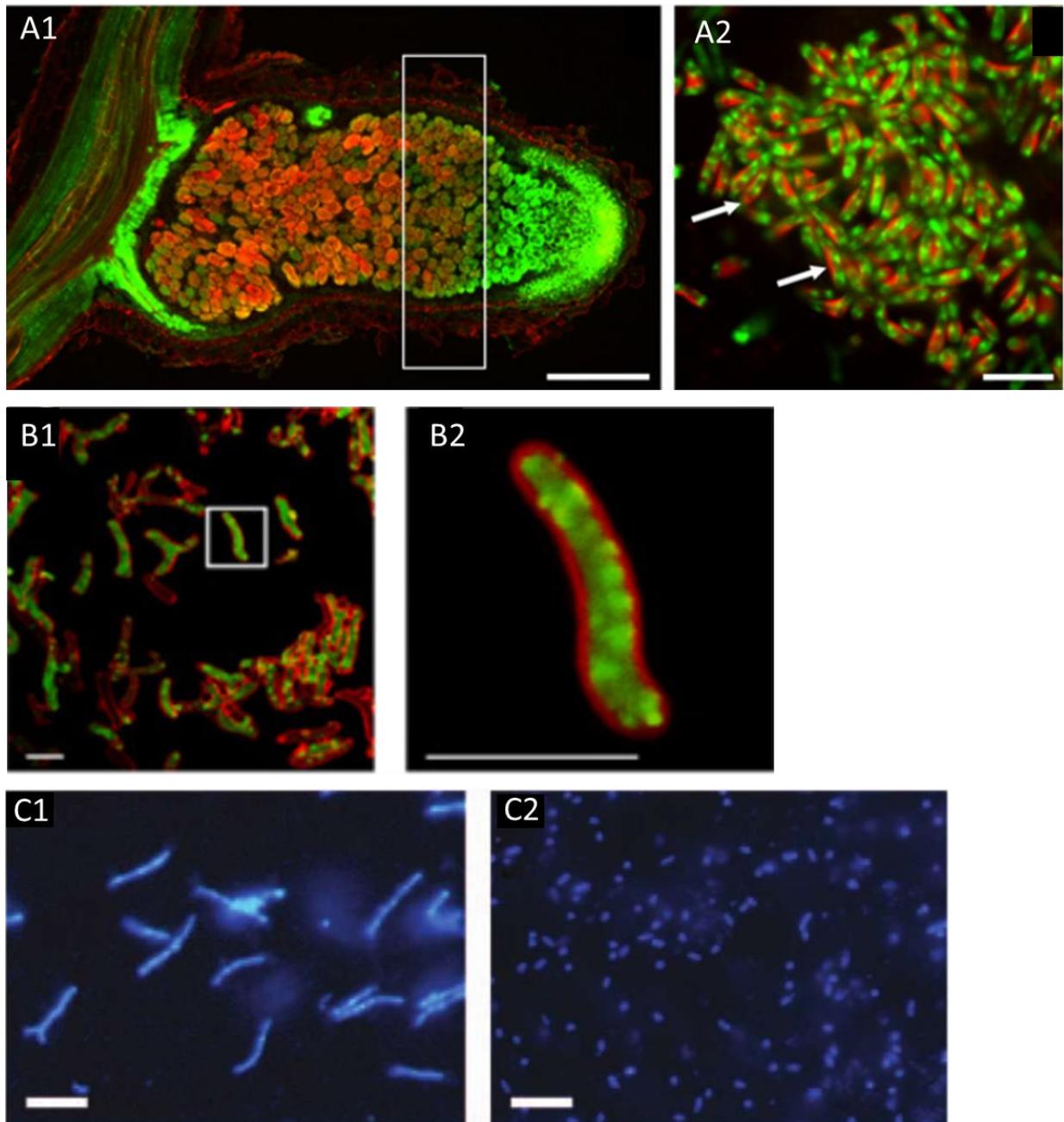


Figure 4. A1 & A2. NCR169-mCherry fusion protein restored the symbiotic phenotype of *dnf7-2* at 21 dpi. In the distal part of Zone III (white rectangle), the green fluorescence of SYTO13 staining of the bacteroids does not overlap with the red mCherry signal surrounding the bacteroids in the peribacteroid space. B1 & B2 *S. meliloti* culture stained with FITC- NCR247 and PI. C1 bacteroid extracted from nodules cells or bacteria isolated from free living culture and stained with DAPI.

Partie II : Travaux dirigés

TD5 Transcriptome et analyses fonctionnelle

Question 11. Analysez la figure 4, quelle est votre conclusion quant à la localisation des NCRs ?

Réponse 11. La figure 4a présente la location du NCR169 par fusion traductionnelle, on observe que le peptide est détecté en zone d'infection/différenciation et zone de fixation, ce qui indique une co-localisation du peptide avec les bactéries. La figure 4b présente la localisation d'un NCR247 on observe une accumulation du peptide au sein de la bactérie. Ces résultats indiquent que les NCRs agissent au niveau des bactéries au cours de la symbiose. La figure 4c présente la morphologie de bactéries issues d'une culture liquide ou d'extraits de nodosités, on observe que les bactéries sont plus grandes quand elles sont dans les nodosités.

Question 12. En sachant que les NCRs affectent la morphologie des bacteroids, proposez une hypothèse quant au rôle potentiel des NCRs dans la variation morphologique des rhizobia ?

Réponse 12. Il est possible que les NCRs agissent dans les nodosités sur les bactéries et provoque l'augmentation de leur taille.

Un mutant de *Sinorhizobium meliloti*, *bacA* a été identifié, la figure 5 présente la caractérisation du mutant.

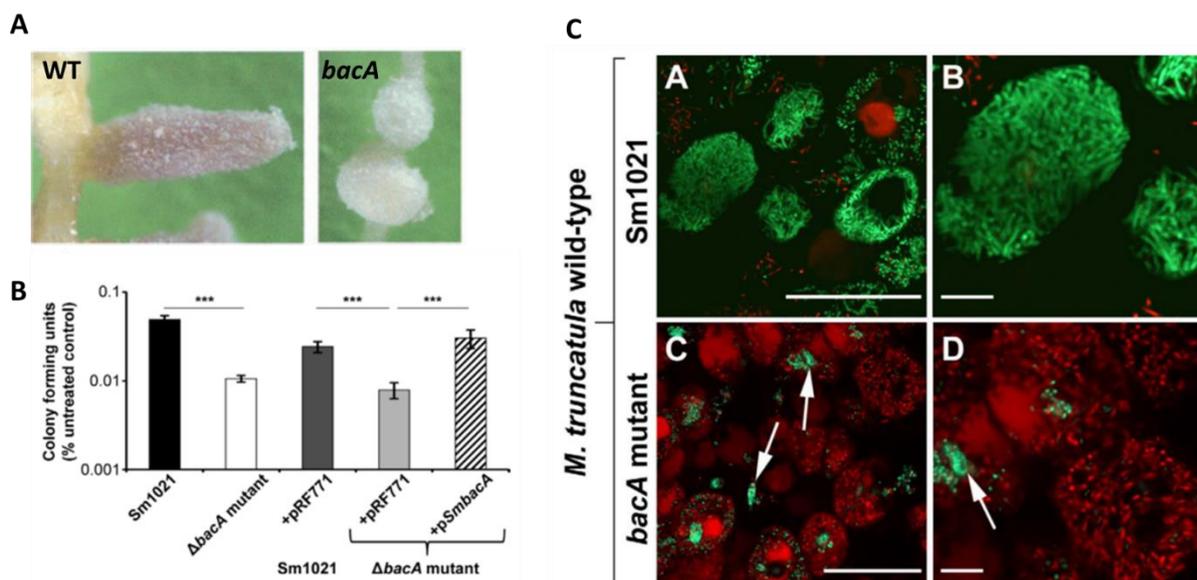


Figure 5. Phenotypic analysis of *bacA* mutant. A. nodule induced by a WT or *bacA*. B. Colony forming ability of the indicated strains was assessed after exposure towards 20 mM of the NCR247 peptide for 3 h. C lived/dead staining of *S. meliloti* or *bacA* in infected cells.

Partie II : Travaux dirigés

TD5 Transcriptome et analyses fonctionnelle

Question 13. Analysez la figure 5, quelle est votre conclusion quant au phénotype du mutant *bacA* ?

Réponse 13. La figure 5A présente le phénotype de plantes WT inoculées avec la bactérie WT et son mutant *bacA*. On observe que les nodosités formées par le mutant sont plus petites et blanches ce qui suggère un dysfonctionnement de la symbiose chez ce dernier. La figure 5B présente le nombre de colonies formées après traitement ou non au NCR247. On observe que le mutant *bacA* forme moins de colonies après ajout du NCR. Par ailleurs l'ajout de la copie sauvage du gène *BacA* chez le mutant provoque une restauration du phénotype. On conclue que l'inactivation du gène *bacA* provoque une augmentation de la sensibilité au NCR de *Sinorhizobium*. La figure 5C présente la survie de la souche WT ou *bacA* dans les nodosités, on observe que le mutant est incapable de survivre dans la nodosité.

Question 14. Quel est votre conclusion par rapport à l'effet des NCRs sur *bacA* ?

Réponse 14. A partir des résultats observés on peut conclure que le gène *BacA* est indispensable pour la survie des bactéries dans les nodosités. Il protège la bactérie de l'action des NCRs dans les cellules de l'hôte.

Le mutant *dnf1* de *M. truncatula* présente un défaut d'adressage de protéine au bacteroid dans les cellules symbiotiques, la figure 6 présente des tests *in vitro* d'effet du NCR247 sur les bactéries ainsi que le phénotype du mutant *bacA* inoculé au mutant *dnf1*.

Question 15. Analysez la figure 6, quelle est vos conclusions sur l'effet des NCRs ?

Réponse 15. La figure 6A présente le phénotype de la bactérie après ajout d'un NCR, on observe une variation de morphologie par rapport au WT ce qui confirme l'hypothèse formulée en réponse 12. La figure 6B présente la survie de la souche WT ou du mutant *bacA* au sein de nodosités formées chez la plante WT ou le mutant *dnf1*. On observe une mort de *bacA* chez le WT, mais pas chez *dnf1*. Ce qui suggère que l'inactivation de *dnf1* annule l'effet de la mutation sur *BacA*.

Question 16. Expliquez le phénotype du mutant *bacA* sur le mutant *dnf1* ? Proposez une hypothèse reliant les NCR aux phénotypes du mutant *bacA* ?

Partie II : Travaux dirigés
TD5 Transcriptome et analyses fonctionnelle

Réponse 16. Il est possible que chez *dnf1* les NCRs ne sont pas adressés vers les bactéries et donc leur effet antimicrobien est annulé au sein de la cellule de l'hôte.

Question 17. Présentez un schéma synthétique de l'ensemble des résultats et des hypothèses formulés ?

Réponse 17. Les étudiants doivent réaliser un travail de groupe afin de restituer les points clés de l'étude. Une note est attribuée en fonction de la qualité de la restitution et intègre l'évaluation continue

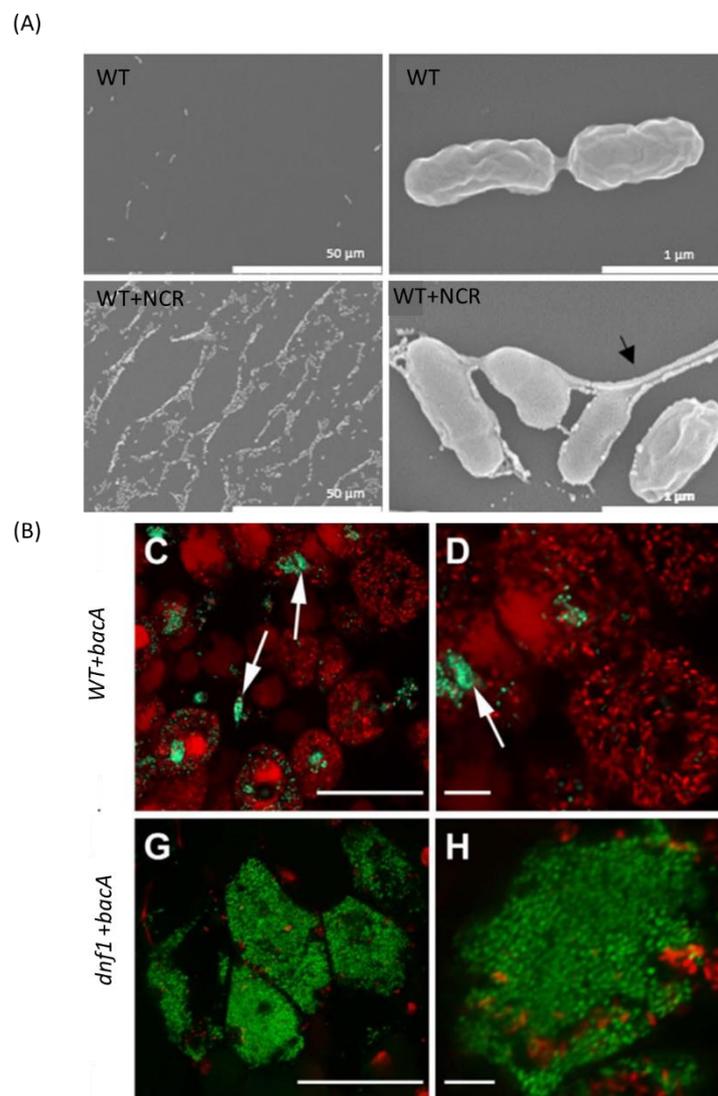


Figure 6. (A) The effect of peptides NCR on the morphology of *S. meliloti*, (B) live/dead staining of *bacA* mutant inoculated in WT or *dnf1*

**Partie III : Travaux pratique
TP1**

**Université Ziane Achour
Faculté des sciences de la nature et de la vie
Département de biologie**

Matière : génie génétique

Nature des enseignements dispensés : Travaux pratiques

Niveau : troisième année biologie, spécialité biologie moléculaire

Année universitaire : 2018/2019

Responsable de la matière : Dr. Berrabah Fathi

**TP1 : Caractérisation transcripionnelle de la réponse
de *M. truncatula* à *R. solanacerum* et étude de la
fonction du gène *MtSickle* dans la régulation de la
réponse.**

Partie III : Travaux pratique

TP1

Introduction au projet de recherche

La bactérie *Ralstonia Solanacearum* est un pathogène responsable de la pourriture bactérienne chez de nombreuses plantes chez lesquelles elle infecte les apex racinaires et se propage dans toute la plante via le système vasculaire. Elle est impliquée chaque année dans des pertes agricole considérables.

Les TPs se proposent de caractériser l'interaction entre *R. Solanacearum* et l'un de ces hôtes, *Medicago truncatula*, afin de mieux comprendre les mécanismes moléculaires associés à la réponse de l'hôte à la bactérie.

Les TPs sont structurés en projet de recherche réalisés par binôme et organisés comme ci-dessous :

TP1 : Caractérisation transcripionnelle de la réponse de *M. truncatula* à *R. solanacerum* et étude de la fonction du gène *MtSickle* dans la régulation de la réponse.

TP2 : Identification de gènes importants pour la réponse de *M. truncatula* à *R. solanacerum* et leurs caractérisations d'un point de vue fonction et bibliographie.

TP3 : Analyse phylogénétique et analyse de l'expression du gène candidat dans des contextes de pathogénie et de symbiose chez l'espèce modèle *M. truncatula*.

TP4 : Synthèse des données obtenues et préparation de la présentation orale pour l'évaluation du TP.

Evaluation du TP/TD : à la fin de la série de TPs vous devrais faire une présentation orale en binôme/trinôme reprenant l'ensemble de vos résultats. La note du TP/TD compte pour 40% de la note finale et comprend 30% sur l'analyse des résultats et 10% pour le rendu oral.

Introduction au TP 1

Les auteurs de l'étude ont réalisés un transcriptome par puce à ADN sur des échantillons de la plante modèle *Medicago truncatula* écotype A17 inoculée avec ou sans l'agent pathogène *R. solanacerum*.

Partie III : Travaux pratique

TP1

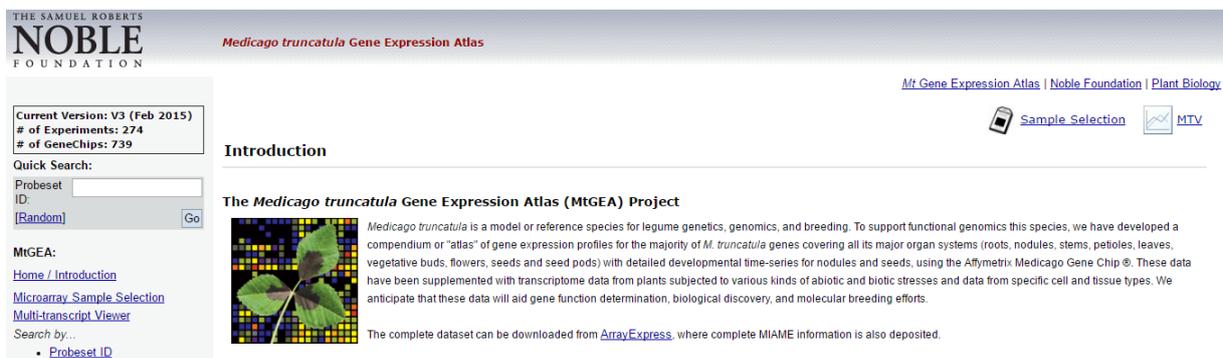
Après inoculation, les auteurs ont incubés les plantes 12 heures et 72 heures en présence du pathogène, puis ils ont récupérés les apex racinaires et les ARNs des plantes ont été extraits. Suite à cela les auteurs ont procédé à la préparation des ADNc, puis ils réalisent leurs hybridations sur des puces à ADN représentant le génome de *M. truncatula*.

Le gène *MtSICKLE*, code un co-récepteur de l'éthylène, une hormone connue pour activer la réponse immunitaire chez les plantes. Afin de vérifier si *Mtsickle* pouvait être impliqué dans l'interaction entre *M. truncatula* et *R. solanacearum*, l'équipe a ajoutée à l'étude le mutant *sickle*, ce dernier est testé dans les mêmes conditions que la plante sauvage (A17).

Le TP1, se propose d'analyser de manière simple les transcriptomes obtenus et de comparer la réponse globale de l'hôte pour certains paramètres (métabolisme, réponse immunitaire, recyclage des protéines et voies hormones).

1- Préparation des outils nécessaire à l'étude : vérifiez que les logiciels suivants sont bien installés : Excel, Document texte, MAPMAN. Aller sur le bureau, faire cliquer droit de la souris, aller sur nouveau, puis cliquer sur nouveau dossier, renommer le dossier avec les initiales de votre binôme, toutes vos données/résultats seront stockées dans ce dossier.

2- Aller sur le site *Medicago truncatula* Gene Expression Atlas (<http://mtgea.noble.org/v3/>).



THE SAMUEL ROBERTS
NOBLE
FOUNDATION

Medicago truncatula Gene Expression Atlas

Current Version: V3 (Feb 2015)
of Experiments: 274
of GeneChips: 739

Quick Search:
Probeset ID:
[Random] [Go]

MtGEA:
[Home / Introduction](#)
[Microarray Sample Selection](#)
[Multi-transcript Viewer](#)
Search by...
• [Probeset ID](#)

[Mt Gene Expression Atlas](#) | [Noble Foundation](#) | [Plant Biology](#)

[Sample Selection](#) [MTV](#)

Introduction

The Medicago truncatula Gene Expression Atlas (MtGEA) Project

Medicago truncatula is a model or reference species for legume genetics, genomics, and breeding. To support functional genomics this species, we have developed a compendium or "atlas" of gene expression profiles for the majority of *M. truncatula* genes covering all its major organ systems (roots, nodules, stems, petioles, leaves, vegetative buds, flowers, seeds and seed pods) with detailed developmental time-series for nodules and seeds, using the Affymetrix Medicago Gene Chip®. These data have been supplemented with transcriptome data from plants subjected to various kinds of abiotic and biotic stresses and data from specific cell and tissue types. We anticipate that these data will aid gene function determination, biological discovery, and molecular breeding efforts.

The complete dataset can be downloaded from [ArrayExpress](#), where complete MIAME information is also deposited.

3- Dans la page d'accueil cliquer sur « directly download the whole dataset » pour accéder à l'ensemble des données.

4- Dans « Balzergue et al., unpublished » chercher les données correspondant à :

[Root Tip A17 Ohpi \(Ralstonia\)](#), [Root Tip A17 12hpi \(Ralstonia\)](#), [Root Tip A17 72hpi \(Ralstonia\)](#),
[Root Tip sickle Ohpi \(Ralstonia\)](#), [Root Tip sickle 12hpi \(Ralstonia\)](#), [Root Tip sickle 72hpi \(Ralstonia\)](#).

Partie III : Travaux pratique TP1

5- Pour chaque valeur cliquer sur “Mean” afin de sélectionner la moyenne de valeurs correspondantes aux trois répétitions biologiques.

Root Tip A17 0hpi (Ralstonia): Organ: root. Treatment Type: biological treatment. Treatment: WT - 0h post-infection (Ralstonia)
Selection: <input type="checkbox"/> RT_Tip_A17_0hpi_1 <input type="checkbox"/> RT_Tip_A17_0hpi_2 <input type="checkbox"/> RT_Tip_A17_0hpi_3 <input checked="" type="checkbox"/> Mean (RT_Tip_A17_0hpi)

Root Tip A17 12hpi (Ralstonia): Organ: root. Treatment Type: biological treatment. Treatment: WT - 12h post-infection (Ralstonia)
Selection: <input type="checkbox"/> RT_Tip_A17_12hpi_1 <input type="checkbox"/> RT_Tip_A17_12hpi_2 <input type="checkbox"/> RT_Tip_A17_12hpi_3 <input checked="" type="checkbox"/> Mean (RT_Tip_A17_12hpi)

Root Tip A17 72hpi (Ralstonia): Organ: root. Treatment Type: biological treatment. Treatment: WT - 72h post-infection (Ralstonia)
Selection: <input type="checkbox"/> RT_Tip_A17_72hpi_1 <input type="checkbox"/> RT_Tip_A17_72hpi_2 <input type="checkbox"/> RT_Tip_A17_72hpi_3 <input checked="" type="checkbox"/> Mean (RT_Tip_A17_72hpi)

6- A la fin de la page web en bas, cliquer sur « Download » pour télécharger les données sélectionnées.

7- Une nouvelle page apparait, cliquer sur « Right click here and select 'Save link as... » Pour faire apparaitre les données dans une nouvelle page.

8- Un nouvelle page apparait, attendre que toutes les données soit télécharger pour poursuivre la manœuvre (astuce : attendre l'apparition à la fin de la page de la ligne qui commence par Mtr.9999.1.S1_at).

9- Un fois les données télécharger sur la page web, cliquer avec le bouton gauche de la souris sur une zone textuelle de la page, appuyer sur les boutons « ctrl+A » pour sélectionner l'ensemble des données, puis cliquer sur les boutons « ctrl+C » pour copier.

10- Aller sur le bureau, ouvrez votre dossier, cliquer sur le bouton droit de la souris, allez dans nouveau, cliquer sur fichier texte, nommez votre fichier « transdata + initial de votre binôme ».

11- Ouvrir le fichier texte et cliquer sur « ctrl+V » pour coller les données dans le fichier texte. Cliquer sur les boutons « ctrl+S » pour sauvegarder.

12- A partir du fichier texte refaite l'étape (9), cliquer avec le bouton gauche de la souris sur une zone textuelle du fichier, appuyer sur les boutons « ctrl+A » pour sélectionner l'ensemble des données puis cliquer sur les boutons « ctrl+C » pour copier.

Partie III : Travaux pratique

TP1

13- Toujours dans votre dossier sur le bureau, cliquer sur le bouton droit de la souris, allez dans nouveau, cliquer sur «feuille de calcul », nommez votre fichier « transdata + initial de votre binôme ».

14- Ouvrir la feuille de calcul, sélectionner la case 1A et cliquer sur « ctrl+V » pour coller les données dans le fichier Excel. Cliquer sur les boutons « ctrl+S » pour sauvegarder, vous disposer ainsi de vos donnée et vous pouvez les analysées.

15- Vérifier que vous avez bien : 7 colonnes (A à G) et 50 901 lignes. La ligne 1 correspond aux différentes conditions, assurée vous que les conditions sont organiser comme suite :

B: RT_Tip_A17_0hpi, C: RT_Tip_A17_12hpi, D: RT_Tip_A17_72hpi, E: RT_Tip_sickle_0hpi, F: RT_Tip_sickle_12hpi, G: RT_Tip_sickle_72hpi.

Supprimez les guillemets et ce qu'elles contiennent pour chaque colonne. La colonne A contient l'annotation des probsets (sondes de la puce).

16- Sélectionner l'ensemble des données des colonnes **B** à **G**, cliquer sur « ctrl+F » afin de rechercher les éléments dans les colonnes, cliquer sur « remplacer », dans « recherche » mettre un point (.) et dans « remplacer par » mettre une virgule (,) enfin cliquer sur « remplacer tous »

Attention : ne pas remplacer les points des probsets.

17- Ajouter deux nouvelles colonnes en position E, nommer la nouvelle colonne (E : Log2 A17 12hpi/0hpi) et la nouvelle colonne (F : Log2 A17 72hpi/0hpi). Ajouter à la colonne (J : Log2 sickle 12hpi/0hpi) et dans la colonne (K : Log2 sickle 72hpi/0hpi).

18- Dans la case **E2** introduire la formule suivante : **$=\text{LOG}((\text{C2}/\text{B2}); 2)$** , dans la case **F2** introduire la formule suivante : **$=\text{LOG}((\text{D2}/\text{B2}); 2)$** , dans la case **J2** introduire la formule suivante : **$=\text{LOG}((\text{H2}/\text{G2}); 2)$** et dans la case **K2** introduire la formule suivante : **$=\text{LOG}((\text{I2}/\text{G2}); 2)$** .

19- Pour chaque case (E2, F2, J2, K2) cliquer deux fois sur le petit point en bas à droite de la case afin d'étendre le calcul au reste des cases pour chaque colonne.

Partie III : Travaux pratique TP1

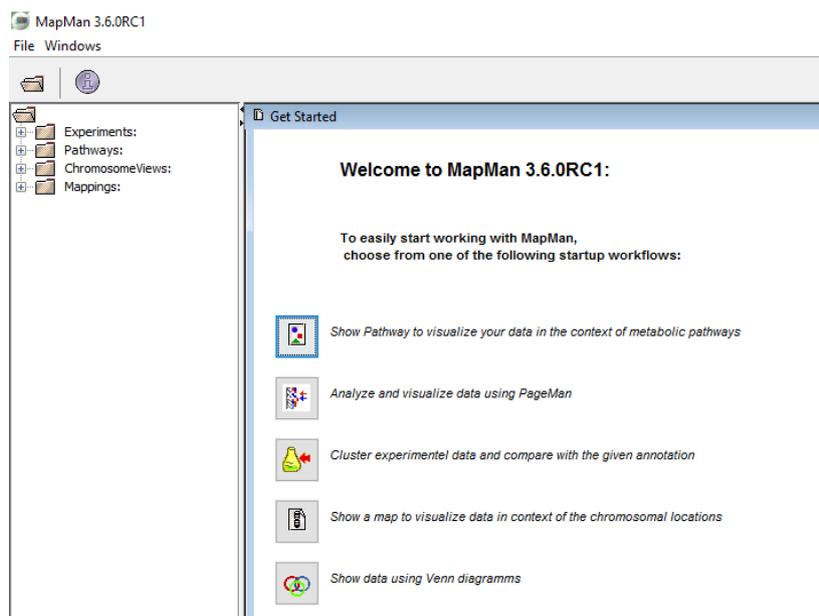
	Log2 A17 12hpi/0hpi	Log2 A17 72hpi/0hpi
7	0,363395884	0,198188465

20- Sélectionner les colonnes : **A** (probset), **E** (Log2 A17 12hpi/0hpi), **F** (Log2 A17 72hpi/0hpi), **J** (Log2 sickle 12hpi/0hpi), **K** (Log2 sickle 72hpi/0hpi), Puis appuyer sur les touches « ctrl+C » afin de copier les éléments sélectionnés.

21- Crée un nouvelle feuille de calcules dans votre dossier et nommée la : analysdata + initial de votre binôme, ouvrir le nouvelle feuille de calcul, sélectionner la case A1 puis appuyer sur la touche « ctrl+V » afin de coller les éléments sélectionnés.

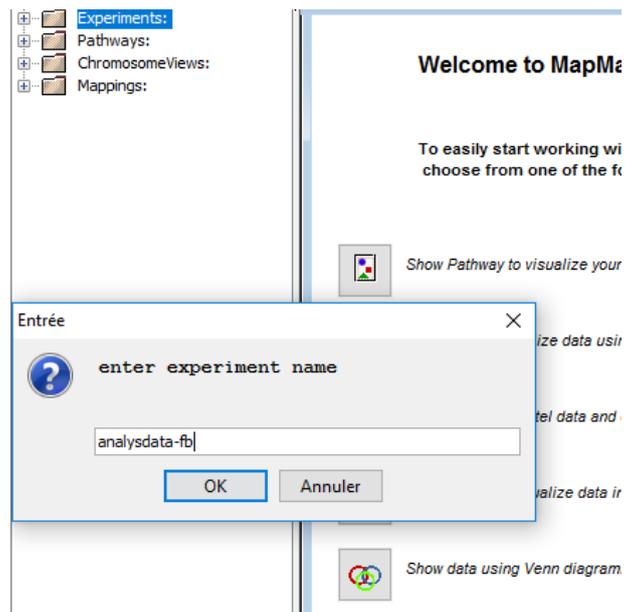
22- Enregistrer une copie de la feuille de calcules sous le format : classeur Excel 97-2003.

23- Ouvrir le logiciel MAPMAN sur votre bureau, une interface apparait,



23- Faire un clic droit sur « experiment » puis sélectionner « new folder », on vous propose de nommer le nouveau dossier, ajouter le nome : analysdata+nom du binôme.

Partie III : Travaux pratique TP1



24- Ouvrir le fichier « expermient », votre dossier apparait, faire un clic droit sur votre fichier et sélectionner « add data ».

25- Chercher votre dossier et sélectionner la feuille Excel sous format Excel 97-2003, une nouvelle fenêtre apparait, cliquer sur « OK ».

26- Vos données apparaissent dans votre fichier dans la partie « expermiment ».

27- Dans la page d'accueil cliquer sur « show pathway », dans «choose pathway » sélectionner « metabolisme overview », dans « choose maping » selectioner «Mt_AFFY_Mt3.0_0510 » et dans « shoose experimental data set » sélectionner la première condition expérimentale (Log2 A17 12hpi/0hpi). Vous avez ainsi une représentation des différentes classifications des gènes induit/réprimer.

28- Régler le seuil « scale » à 1 pour représenter les gènes induits plus de deux fois, en bleu sont représentés les gènes induits et en rouge les gènes réprimés, les gènes non induits sont représentés en case vide. Augmenter la taille des cases à 6 sur « Size ».

29- Aller dans votre fichier sur le bureau, cliquer sur le bouton droit de votre souris, allez sur nouveau puis cliquer sur nouvelle présentation Microsoft PowerPoint, nommé le : Resultats_TP1_initiales du binôme.

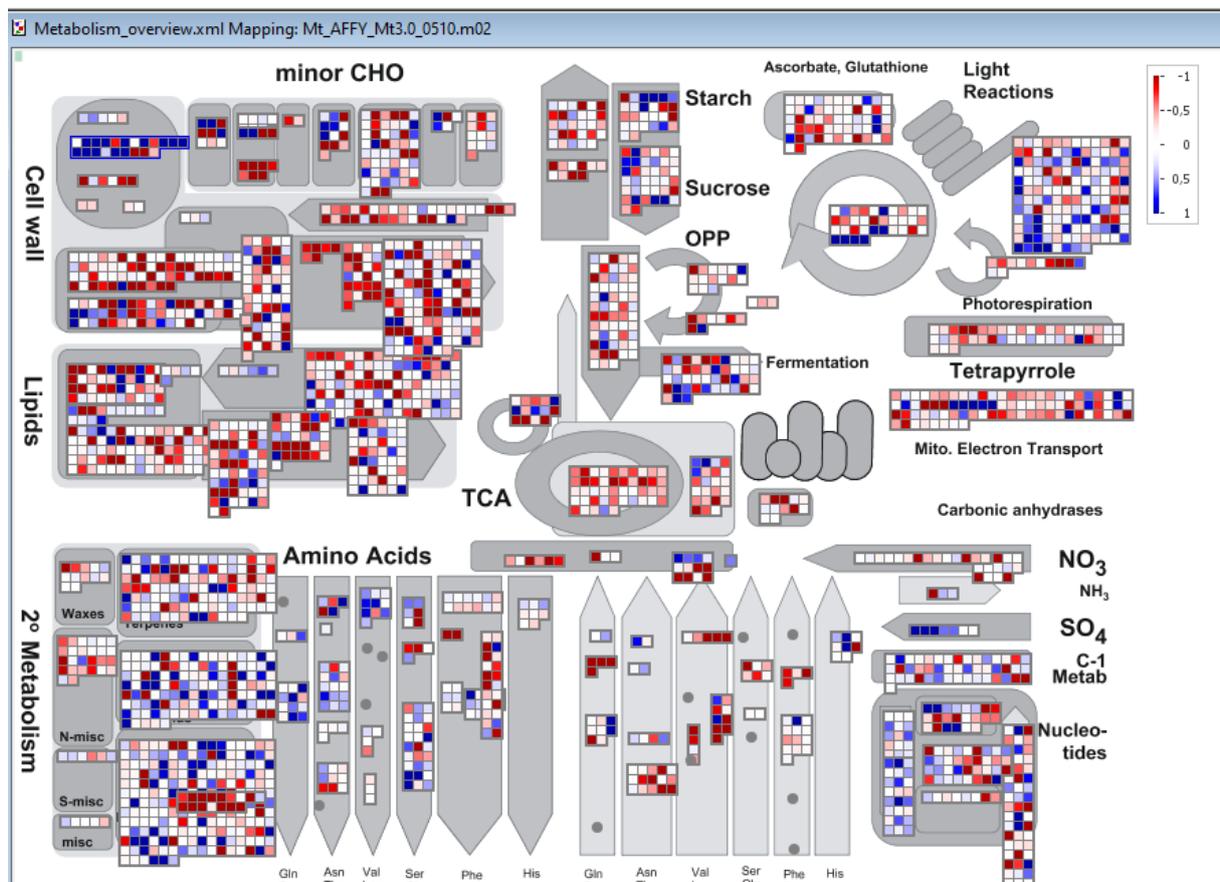
30- Créé une nouvelle diapo de présentation.

Partie III : Travaux pratique TP1

31- Sur MAPMAN, afficher la page de donnée au maximum de sa taille, ajustez le schéma, cliquer sur la première condition puis appuyer sur « imprime écran ».

32- Retourner dans PowerPoint, ajouter le nom de la condition comme titre à votre diapo (exemple : Metabolisme Log2 A17 12hpi/0hpi) puis faire « ctrl+V » pour coller l'image.

33- En cliquant deux fois sur l'image et dans l'onglet « rogner » en haut à droite, vous pouvez ajuster votre image.

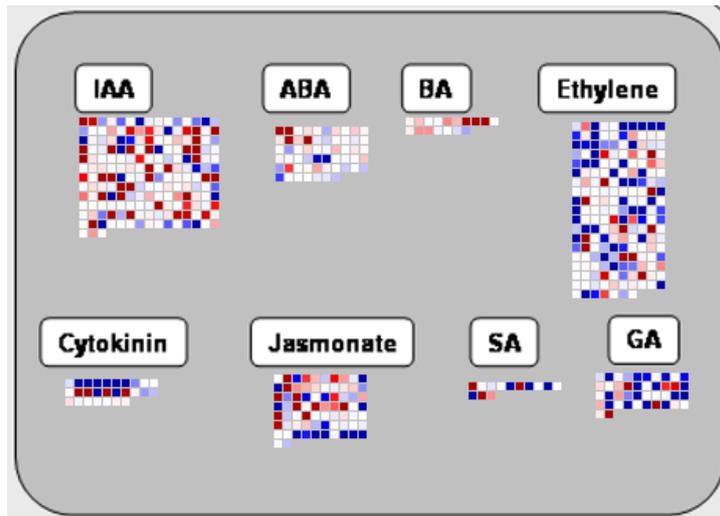


34- Refaite les procédures 30, 31, 32 et 33 pour chaque conditions, puis cliquer dans PowerPoint sur « ctrl+S » pour sauvegarder les documents.

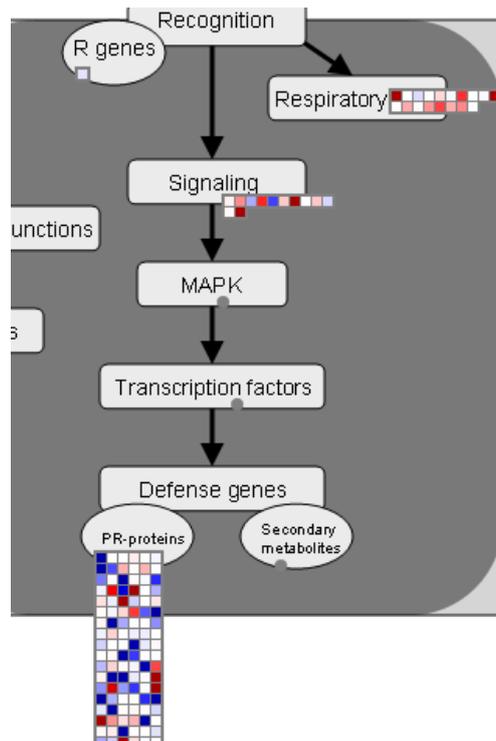
35- Refaite la procédure (étape 30, 31, 32, 33, 34) pour chaque conditions en utilisant les pathways suivant :

- Régulation overview, la partie hormone (**N.B : toujours régler le « sacle » sur 1**).

Partie III : Travaux pratique TP1

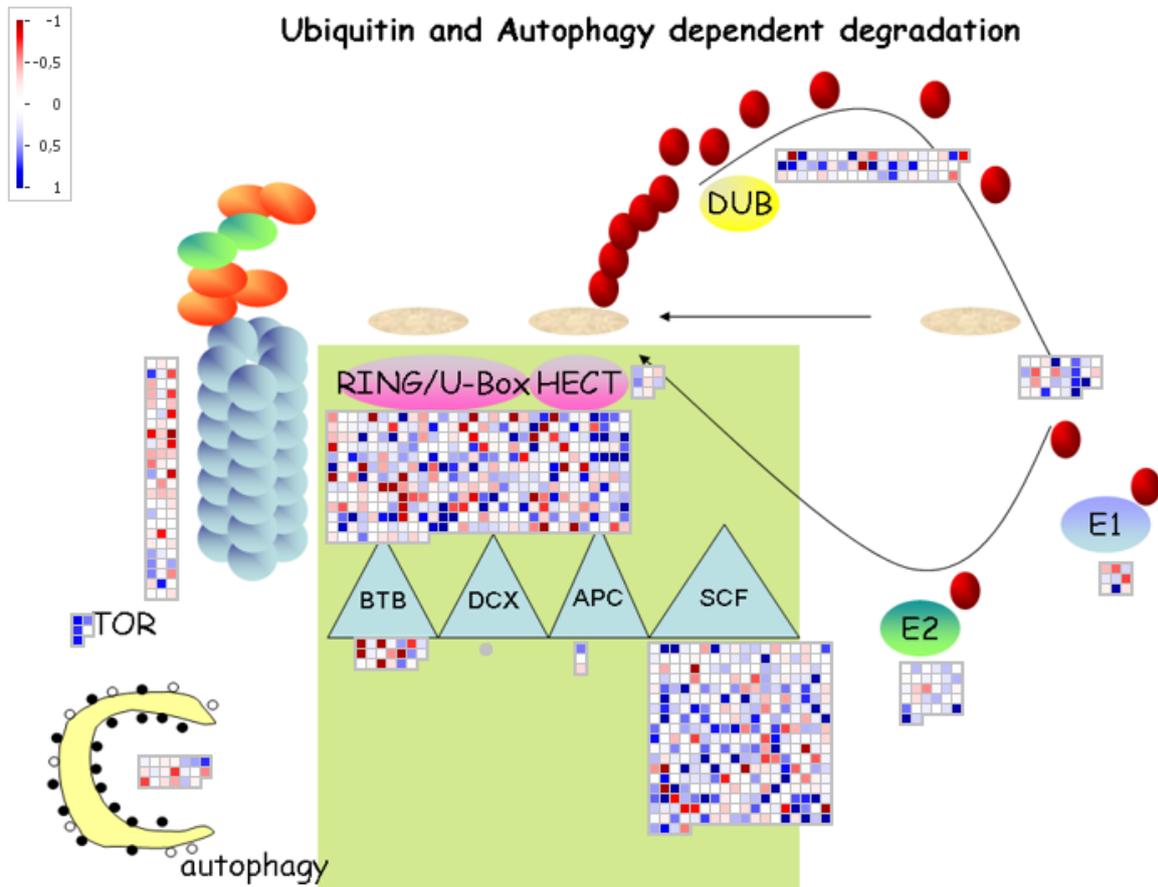


- Biotic stress, les parties : respiratory, signaling et PR protein (N.B : toujours régler le « sacle » sur 1).

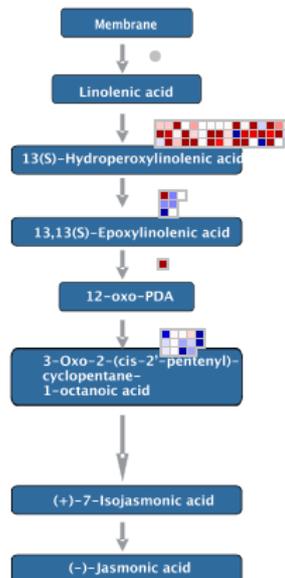


- Ubiquitine and autophagy dependent degradation. (N.B : toujours régler le « sacle » sur 1).

Partie III : Travaux pratique TP1



- Hormone, JA synthesis.



36- Sauvegarder vos données, fermer toutes les pages et transférait votre dossier sur clé USB.

Partie III : Travaux pratique TP1

Objectif du TP

Chez vous, vous devrais traiter et analyser les données que vous avez obtenu, vous vous focaliserais sur les questions ci-dessous :

- Comparer chez A17 la réponse de manière qualitative et quantitative pour les paramètres étudiés, comparer entre 12 hpi et 72 hpi. Quelles sont vos conclusions ?
- Comparer de manière qualitative et quantitative la réponse de *sickle* à celle d'A17 à 12hpi et 72hpi. Quelles sont vos conclusions ?

**Partie III : Travaux pratique
TP2**

**Université Ziane Achour
Faculté des sciences de la nature et de la vie
Département de biologie**

Matière : génie génétique

Nature des enseignements dispensés : Travaux pratiques

Niveau : troisième année biologie, spécialité biologie moléculaire

Année universitaire : 2018/2019

Responsable de la matière : Dr. Berrabah Fathi

**TP2 : Identification de gènes importants dans la
réponse de *M. truncatula* à *R. solanacerum* et leurs
caractérisation d'un point de vue fonction et
bibliographie.**

Partie III : Travaux pratique TP2

Introduction au TP 2

Après avoir analysé dans le TP1 la réponse globale de *Medicago truncatula* écotype A17 à *Ralstonia solanaceum* et l'effet potentiel que produit la mutation de perte de fonction du gène *MtSickle* sur la réponse de *M. truncatula*. Le TP2 se propose d'analyser de manière plus fine la réponse de la plantes hôte, notamment en examinant un gène fortement induit chez *M. truncatula* en réponse à la bactérie pathogène.

Le TP est divisé en plusieurs parties :

Partie A : Identification et sélection d'un gène d'intérêt et récupérer la séquence du CDS.

Partie B : Recherche d'homologues de votre protéine et identification de la fonction potentielle.

Partie C : Recherche de peptide signale et de domaines transmembranaires.

Partie D : Recherche de domaine important (fonctionnels) pour la protéine.

Partie E : Recherche bibliographique sur votre gène/protéine ou la famille à laquelle il appartient.

Partie III : Travaux pratique TP2

Partie A : identification et sélection d'un gène d'intérêt et récupérer la séquence du CDS

- 1- Aller sur votre dossier de TP (sur le bureau ou votre Clé USB), ouvrez votre fichier **analysdata + initial de votre binôme**.
- 2- Sélectionner les colonnes : A, B, C, D et E, dans l'anglet « accueil » en haut à droite, cliquer sur « filtrer et trier » puis sur « filtrer ».
- 3- Cliquer sur la flèche de la colonne B, puis cliquer sur « trier du plus grand au plus petit », vous obtenez les gènes les plus induits après 12 heures d'incubation.
- 4- Cliquer sur la flèche de la colonne C, puis cliquer sur « trier du plus grand au plus petit ». Vous obtenez les gènes les plus induits après 72 heures d'incubation.
- 5- Chaque binôme/trinôme va sélectionner un probset selon le tableau ci-dessous :

Numéro binôme	Probeset
1	Mtr.24731.1.S1_at
2	Mtr.42985.1.S1_at
3	Mtr.18380.1.S1_at
4	Mtr.7244.1.S1_at
5	Mtr.39593.1.S1_at
6	Mtr.8884.1.S1_at
7	Mtr.1633.1.S1_x_at
8	Mtr.10494.1.S1_x_at
9	Mtr.43627.1.S1_at
10	Mtr.27355.1.S1_s_at
11	Mtr.37859.1.S1_at
12	Mtr.51360.1.S1_at

- 6- Une fois votre probset identifié, sélectionner toute la ligne de votre probset ainsi que la première ligne qui donne les correspondance des valeurs de chaque case, appuyé sur « ctrl+C » puis aller dans la feuille 2 de calcul de votre fichier Excel, sélectionner la case 1A puis appuyé sur «ctrl+V ».
- 7- Sélectionner l'ensemble des données de la feuille 2, aller dans l'anglet « insertion » et cliquer sur « graphique recommandé » puis sélectionner « histogramme groupé », vous

Partie III : Travaux pratique TP2

obtenez ainsi un histogramme représentant le niveau d'expression de votre gène dans le fond génétique WT et *sickle* en réponse à *R. solanacearum* (Voir l'exemple dans la figure 1)

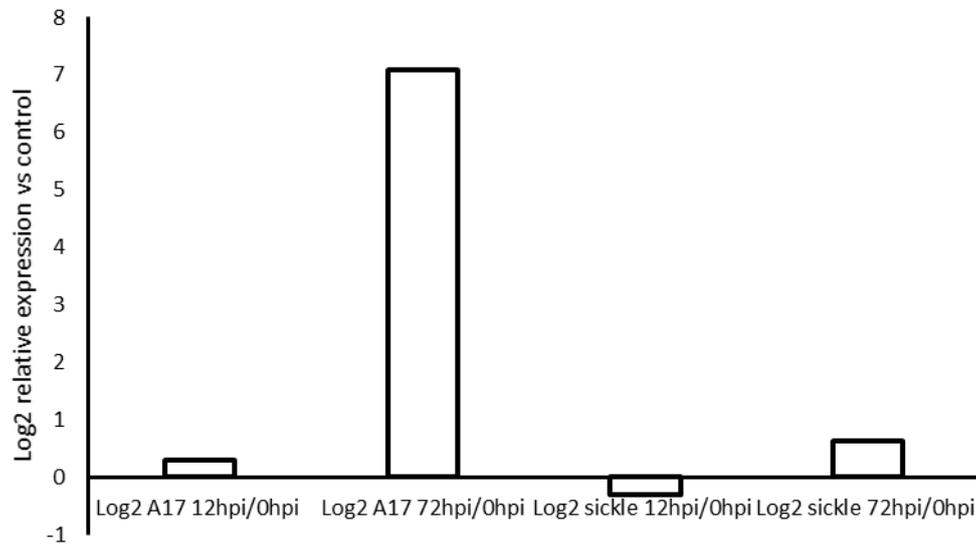
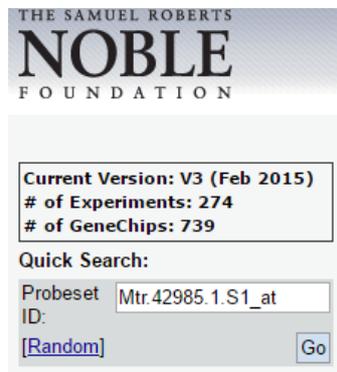


Figure 1. Expression relative du gène X chez la plante WT et le mutant *sickle* 12 heures ou 72 heures après inoculation avec *R. solanacearum*

8- Analyser et interpréter vos résultats (à faire chez vous).

9- Sauvegarder votre feuille en cliquant sur « ctrl+S ».

10- Sélectionner votre probset, cliquer sur « ctrl+C », aller sur le site du MtGEA (<http://mtgea.noble.org/v3/index.php>) et coller votre probset dans l'anglet « quick search » en cliquant sur « ctrl+V » et cliquer sur « Go ».



11- Une nouvelle fenêtre s'ouvre qui contient les informations suivantes :

- Le profil d'expression du gène : elle correspond au signal d'hybridation de la sonde pour différents type d'expériences.

Partie III : Travaux pratique TP2

- Consensus, target, and probe sequences : elle correspond à la séquence cible ainsi que les régions d'hybridation des sondes sens et anti-sens.
- Chip and genome coordinates : elle correspond à la séquence des sondes et permet d'identifier leur zone d'hybridation sur le génome *M. truncatula*.
- Initial annotation during chip design: correspond à l'annotation initial du gène.
- Gene (re)mapping & annotations : correspond aux nouvelles annotations du gène.
- Legume specific gene : indique si le gène est spécifique des légumineuses.
- Homolog transcripts : permet d'obtenir le(s) CDS orthologue(s) du gène.
- GO annotations : Gene Ontology, elle est l'équivalent des COG (Cluster of Orthologous Groups) chez les procaryotes et elle correspond aux classes fonctionnelles (kinase, hormone, métabolisme, biological process...) à laquelle appartient le gène, elle correspond aux voies biologiques dans lesquelles intervient potentiellement le gène.

12- Pour obtenir la séquence du CDS, vous disposez de deux méthodes (A & B), les deux méthodes vont être présentées mais seule la **méthode B** sera retenue car elle est plus complète.

Méthode A :

13- vous vous basés sur la séquence mise à disposition sur le site du MtGEA, dans « Consensus, target, and probe sequences » vous sélectionnez la séquence et vous la copiez.

>Mtr.42985.1.S1_at

```
GGACAGCAACATTCAATTCTACCAAGAACTCAAACATATATACATACAAAAATGGGAGTTTTTTACTTC
ATCGTTGCCTTATTGGCTTAGCATCCTCTGTTGGTTTTGCCTATGATCCCAGCCCATTGCAAGACTTTT
GTGTTGCAATCAATGATACTAAAAGTGGTGTGTTGTGAATGGAAAATTTGCAAAGATCCTAAGCTTG
CAAATGCGGATGATTTCTTCTTTCAAGGATTGGGACCTGGGAATACTTCAAATCCACTCGGATCAAAG
TGACTCCAGTGACGGTTAATGAAATATTAGGACTCAACACACTTGGTATATCTTTGGCTCGCGTCGATT
TTGCACAAAAGGTCTAAATCCTCCTCACACCCACCAAGAGGCACCGAGATTCTTGATTTTTGGAAG
GCACACTTTACGTTGGATTTGTTTCCTCAAATCAAGACAATAACCGTCTCTTACCAAAGTATTGAACAA
AGGTGATGTATTTGTGTTTCAATCGGTCTCATTCACTTCCAATTGAACGTTGGTTATGGGGAGGCTAT
TGCCATTGCTGGTCTTAGTAGCCAAAACCTGGAGTGATTACCATTGCAAATGCTGTGTTTGGATCTAA
ACCACCAATTTCTTTGGAAGTTCTCACCAAAGCCTTTCAAGTGGATAAAAATGTAATCGACTATCTTCAA
AAGCAATTCTGGTACAACAATAGTTAATTGTGCTTGAACGTTGAACACGTGATTTCTTAAGTTTGCCT
```

Partie III : Travaux pratique TP2

CATGTTGAGTTGAGTCTAAGAAAATTTGACTAATATAGCAAATAAATGCATGAGGATATGCAGTATGG
TTGTTTTCTTTTGTAACTTTGTTTTGTGTGACTTTTAACTTGTTTGATTTATCATGCAAGTATAATAAAT
AGTATTTTGTCTGCTTTGGTACTTC

Methode B :

14- Aller dans "Medicago v4.0 annotation" ce qui correspond à la dernière version de l'annotation du gène. Copier l'ID du gène (Medtr + numéro g + numéro).

Medicago v4.0 annotation	
Transcript ID	Description
Medtr1g079490.1	germin family protein

15- Aller sur le site de Graig Venter (<http://jcv.org/medicago/>), dans l'onglet « Resources », dans « Search » cliquer sur « Locus Search ».

16- Copier l'ID du gène dans « Gene Curation Search » puis cliquer sur « Search ».

Gene Curation Search

Obtain a report of the predicted coding region by chromosome locus.

Search by locus id only (e.g., Medtr5g024510, Medtr2g011160)

17- Cliquer sur « view sequences », la séquence génomique, le CDS ainsi que la séquence de la protéine apparaissent, copier la séquence du CDS et de la protéine (ctrl+C).

18- Aller dans votre dossier, créer un nouveau document texte, nommé le : **candidatgene_nom du binôme**, ouvrir le fichier, coller (ctrl+V) les séquences et cliquer « ctrl+S » pour sauvegarder.

CDS

>Medtr1g079490.1

ATGGGAGTTTTTACTTCATCGTTGCCTTATTGGCTTTAGCATCCTCTGTTGGTTTTGCCTATGATCCCA
GCCATTGCAAGACTTTTGTGTTGCAATCAATGATACTAAACTGGTGTGTTTGTGAATGGAAAATTT
TGCAAAGATCCTAAGCTTGCAAATGCGGATGATTTCTTCTTTCAAGGATTGGGACCTGGGAATACTTC
AAATCCACTCGGATCAAAGTGACTCCAGTGACGGTTAATGAAATATTAGGACTCAACACACTTGGT
ATATCTTTGGCTCGCGTGCATTTTGCACAAAAGGTCTAAATCCTCCTCACACCCACCCAAGAGGCAC
CGAGATTCTGTAGTTTTGGAAGGCACACTTTACGTTGGATTGTTTCCTCAAATCAAGACAATAACC

Partie III : Travaux pratique TP2

```
GTCTCTTCACCAAAGTATTGAACAAAGGTGATGTATTTGTGTTTCCAATCGGTCTCATTCACTTCCAAT  
TGAACGTTGGTTATGGGGAGGCTATTGCCATTGCTGGTCTTAGTAGCCAAAACCCTGGAGTGATTAC  
CATTGCAAATGCTGTGTTTGGATCTAAACCACCAATTTCTTTGGAAGTTCTACCAAAGCCTTTCAAGT  
GGATAAAAATGTAATCGACTATCTTCAAAGCAATTCTGGTACAACAATAGTTAA
```

Protein

>Medtr1g079490.1

```
MGVYFIVALLALASSVGFAYDPSPLQDFCVAINDTKTGVFVNGKFKDKPKLANADDFFFQGLGPGNTSN  
PLGSKVTPVTVNEILGLNLTGLISLARVDFAPKGLNPPHPRGTEILVVLEGLTYVGFVSSNQDNNRLLFTKV  
LNKGDVVFVPIGLIHFQLNVGYGEAIAIAGLSSQNPVITIANAVFGSKPPISLEVLTKAFQVDKNVIDYLQK  
QFWYNNNS
```

Partie B : recherche d'homologues de votre protéine et identification de la fonction potentielle.

19- A partir de votre fichier texte **candidatgene_nom du binôme** sélectionner la séquence de la protéine correspondante et copier la (ctrl+C).

20- Aller sur le site internet Blast de la National Center for Biotechnology Information (NCBI : <https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

21- Sélectionner « Proteine blast », coller la séquence de la protéine dans « Enter accession number(s), gi(s), or FASTA sequence(s) »

Standard Protein BLAST

blastn blastp blastx tblastn tblastx

BLASTP programs search protein databases using a protein query. [more...](#)

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s)

>Medtr1g079490.1
MRRRPVAIIFSDQMEKSTTKSHNSRIFFLASLFAFFWLLLLLYFHFILPHQQSSTTVNIQ
QPPPRKHGFPDIKKTDPCCGGRYIYVHNLPSRFNQDMLRECKTLLSLWTNICKFTTNAAGL
PPLENIDGVFSDTGHVATNQFAVDVIFANRMKQYECLTNDSSIAAAVFPFYAGFDIARY
LNGYNISMRDAASVELVDWLMKRPEWGMNGRDHFLVAGRITNDFRRLSEDEKDWGNKLL

Clear Query subrange

From

To

BLAST Search database Non-redundant protein sequences (nr) using Blastp (protein-protein BLAST)

Show results in a new window

22- En bas de la page, cliquer sur « Show results in a new window » puis sur blast.

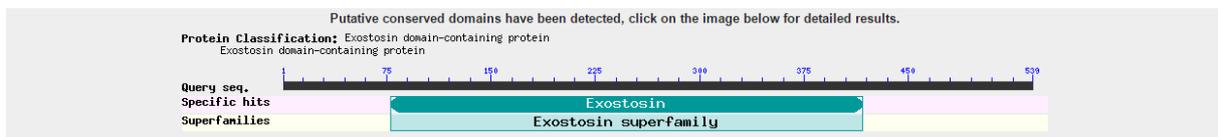
23- Un nouvelle page s'affiche, correspondant aux résultats du blast, Les informations suivantes sont disponibles :

Partie III : Travaux pratique TP2

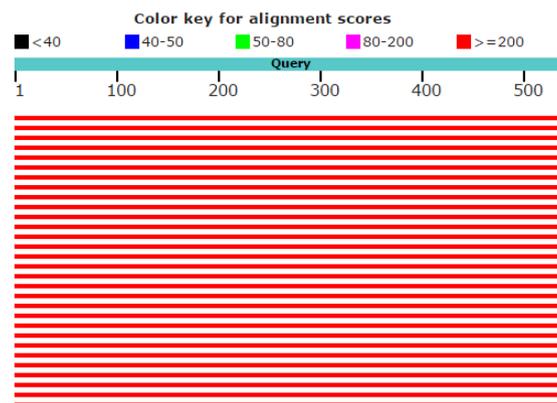
- Les informations sur la protéine (cadrant rouge) et les informations sur la banque de donnée utilisé pour le blaste (cadrant vert).

Job title: Medtr7g098650.1 RID 8FDBRA9E014 (Expires on 01-26 03:13 am) Query ID lcl Query_94151 Description Medtr7g098650.1 Molecule type amino acid Query Length 539	Database Name nr Description All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples from WGS projects Program BLASTP 2.6.0+ Citation
---	--

- Les informations sur les domaines putatifs que contient la protéine.



- Le résultat du score, le code couleur permet d'évaluer la force du score pour chaque identification.



- Les résultats détaillés des alignements, la première ligne correspond à votre protéine, puis les autres lignes à des protéines homologues.

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected: 0

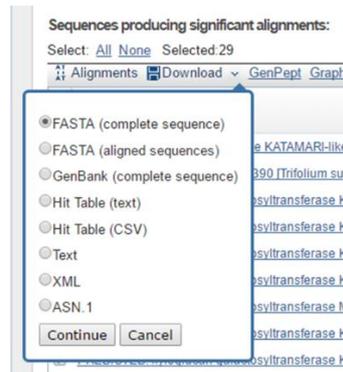
Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> xyloglucan galactosyltransferase KATAMARI-like protein [Medicago truncatula]	1129	1129	100%	0.0	100%	XP_003625392.1
<input type="checkbox"/> hypothetical protein TSUD_336390 [Trifolium subterraneum]	1014	1014	100%	0.0	88%	GAI119352.1
<input type="checkbox"/> PREDICTED: xyloglucan galactosyltransferase KATAMARI1-like [Cicer arietinum]	1008	1008	100%	0.0	89%	XP_004493835.1
<input type="checkbox"/> PREDICTED: xyloglucan galactosyltransferase KATAMARI1 [Arachis ipaensis]	983	983	99%	0.0	82%	XP_016207374.1
<input type="checkbox"/> PREDICTED: xyloglucan galactosyltransferase KATAMARI1 [Arachis duranensis]	975	975	99%	0.0	79%	XP_015968108.1
<input type="checkbox"/> PREDICTED: xyloglucan galactosyltransferase KATAMARI1 [Glycine max]	967	967	100%	0.0	82%	XP_003553511.2
<input type="checkbox"/> PREDICTED: xyloglucan galactosyltransferase MUR3-like [Vigna angularis]	960	960	100%	0.0	82%	XP_017411149.1
<input type="checkbox"/> PREDICTED: xyloglucan galactosyltransferase KATAMARI1-like [Vigna radiata var. radiata]	958	958	100%	0.0	82%	XP_014496344.1
<input type="checkbox"/> PREDICTED: xyloglucan galactosyltransferase KATAMARI1 [Glycine max]	952	952	100%	0.0	79%	XP_003536950.2
<input type="checkbox"/> PREDICTED: xyloglucan galactosyltransferase KATAMARI1-like [Cicer arietinum]	947	947	99%	0.0	78%	XP_004495055.1
<input type="checkbox"/> Xyloglucan galactosyltransferase KATAMARI1 [Glycine soja]	939	939	100%	0.0	77%	KHN48947.1

Dans notre exemple, la protéine est une xyloglucan galactosyltransferase KATAMARI-like protein de *Medicago truncatula*.

Partie III : Travaux pratique TP2

24- En cochant sur la case devant chaque protéine, sélectionné maximum les 25 premiers candidats avec une identité minimum de plus de 75%.

25- Une fois sélectionner, cliquer sur « download » et sélectionner « FASTA complete sequence».



26- Cliquer sur continue, un fichier texte est télécharger, ouvrir le fichier texte.

27- Enregistrer une copie du fichier téléchargé sur votre dossier, puis renommer le : **homologue_nom du binôme**, vous utiliserez ces informations dans le TP3.

Quelle est votre protéine, à quelle famille elle appartient ?

Partie C : recherche de peptide signale et de domaines transmembranaires

28- Copier la séquence de votre protéine à partir du fichier texte **candidatgene_nom du binôme**.

29- Pour vérifier s'il y'a un peptide signale, aller sur le site internet « SignalP 4.1 Server» (<http://www.cbs.dtu.dk/services/SignalP/>), puis coller votre séquence dans le cadran enfin cliquer sur « submit».

Partie III : Travaux pratique TP2

SignalP 4.1 Server

SignalP 4.1 server predicts the presence and location of signal peptide cleavage sites in amino acid sequences from different organisms: Gram-positive prokaryotes, Gram-negative prokaryotes, and eukaryotes. The method incorporates a prediction of cleavage sites and a signal peptide/non-signal peptide prediction based on a combination of several artificial neural networks.

View the [version history](#) of this server. All the previous versions are available on line, for comparison and reference.

NEW: The portable version of SignalP 4.1, previously only available for Mac (Darwin), Linux, and IRIX, is now also available for Windows systems. Academic users: select the "CYGWIN" option at the [download page](#). [Cygwin](#) or [MobaXterm](#) is required to install SignalP under Windows. For details, read the [installation instructions](#).

FAQ	Article abstracts	Instructions	Output format	Performance	Data
-----	-------------------	--------------	---------------	-------------	------

SUBMISSION

Paste a single amino acid sequence or several sequences in [FASTA](#) format into the field below:

```
MRRRPVAAIFSDQMEKSTTKSHISRIFFLASLSAFFNLLLLYFHFIIIPHQSSSTTVNIQPPPRKH  
FGFPDIKKTDPCCG  
GRVYIVHNLPSRFNQWHLRECKTLSLWNTMCKFTTNAGLGPPLENIQVFSDTGIYATIQFAVDVIF
```

Submit a file in [FASTA](#) format directly from your local disk:
[Choisissez un fichier](#) | Aucun fichier choisi

Organism group ([explain](#))

- Eukaryotes
- Gram-negative bacteria
- Gram-positive bacteria

Output format ([explain](#))

- Standard
- Short (no graphics)

D-cutoff values ([explain](#))

- Default (optimized for correlation)
- Sensitive (reproduce SignalP 3.0's sensitivity)
- User defined:
0.45 D-cutoff for SignalP-noTM networks
0.50 D-cutoff for SignalP-TM networks

Method ([explain](#))

- Input sequences may include TM regions
- Input sequences do not include TM regions

Graphics output ([explain](#))

- No graphics
- PNG (inline)
- PNG (inline) and EPS (as links)

Positional limits ([explain](#))

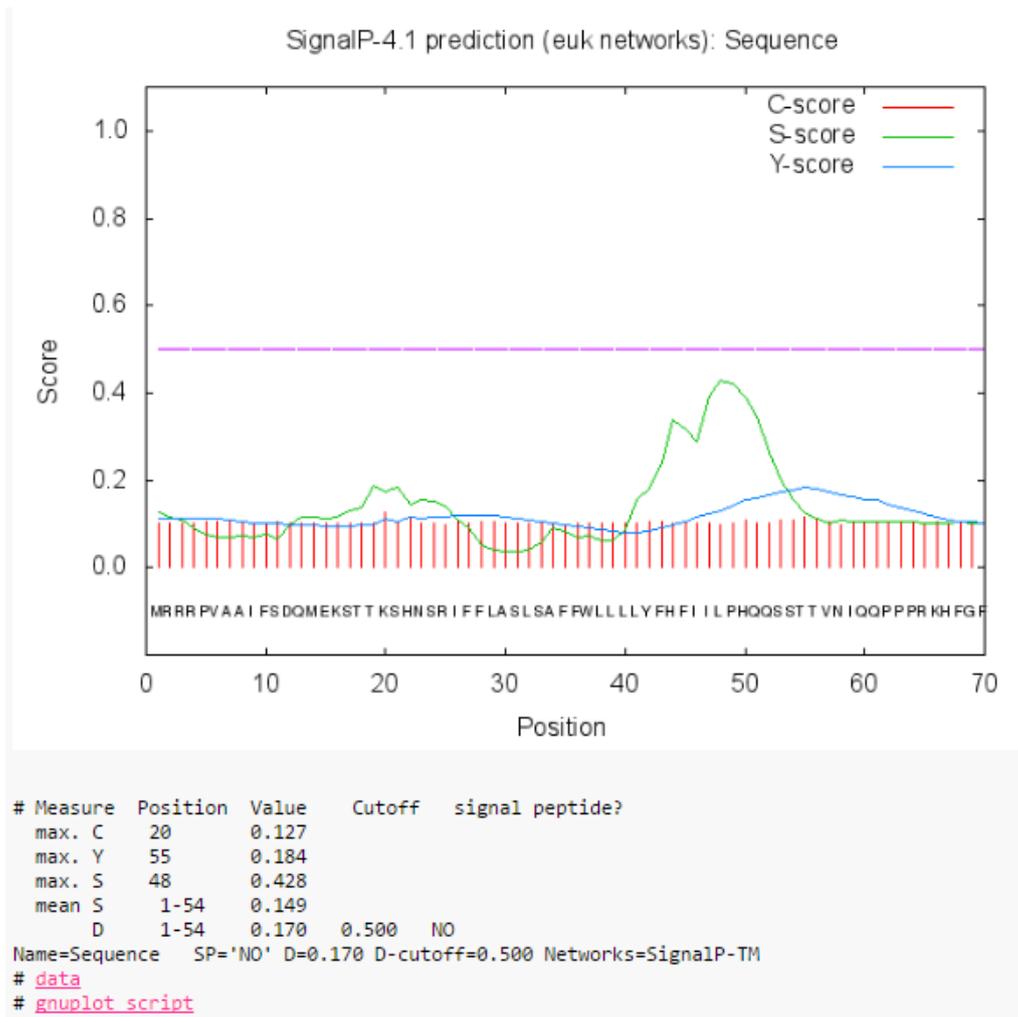
Minimal predicted signal peptide length. *Default: 10*

N-terminal truncation of input sequence (0 means no truncation).
Default: Truncate sequence to a length of 70 aa

30- Le résultat apparait, il indique la probabilité de présence d'un peptide signal et sa position potentielle. Dans notre exemple, la protéine ne présente pas de peptide signale.

Astuce : Afin de pouvoir lire vos résultats, veuillez consulter la page suivante :
<http://www.cbs.dtu.dk/services/SignalP/output.php>.

Partie III : Travaux pratique TP2



31- Pour prédire les sites de domaines transmembranaire et le peptide signal, aller sur le site internet « Phobius » (<http://phobius.sbc.su.se/>), puis coller votre séquence dans le cadrant enfin cliquer sur « visualiser ».



Phobius

A combined transmembrane topology and signal peptide predictor



POST NEBULA PHOBIUS

[Normal prediction](#)
 [Constrained prediction](#)
 [PolyPhobius](#)
 [Instructions](#)
 [Download](#)
 [Mirror site at KU](#)

Normal prediction

Paste your protein sequence here in Fasta format:

```

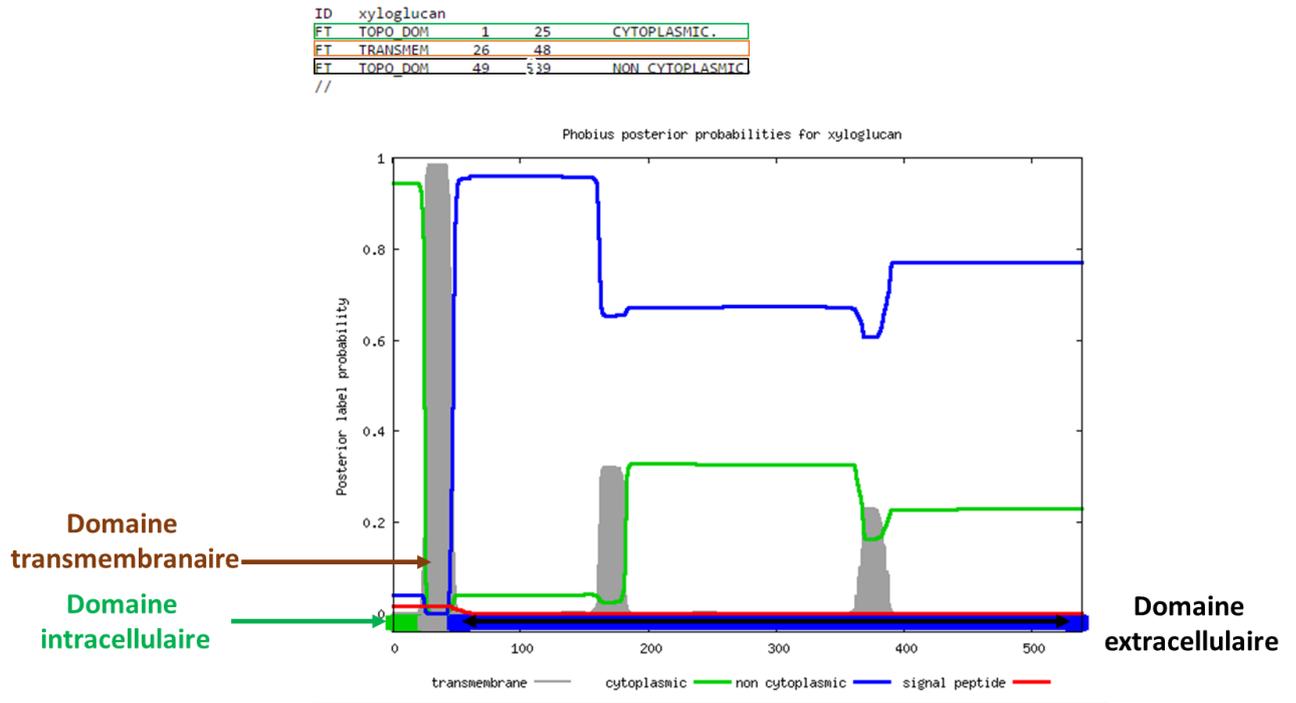
>xyloglucan galactosyltransferase KATAMARI-like protein
MRRRVAIAIFSDQHEKSTTKSHSRIFFLASLFAFWLLLLYFHFILPHQSSSTTVAIQPPPRKHFQFPDIKKT
DPCGGRVIVVHNLPSRFNQDLRECKTSLNLTNCKFTTNAGLGPPLNIDGVESDTGNYATNQFAVDVIFANRHK
QYELTNDSSIAAAVFPFYAGFDIARYLWGYNISMRDAASVELVDNLKRPENGVHNGRDFLVAGRITDFRRL
SEDEKDWGKLLFLPAAKNISHLVVSSPMNANDFGIPYPTVFHPAKDKDVFVQERMRRLERKILFSFAGAPRPG
NAKSIHQIIEQRSS

```

Or: Select the sequence file you wish to use | Choisissez un fichier | Aucun fichier choisi

Partie III : Travaux pratique TP2

32- Un nouvelle fenêtre apparait avec les résultats, dans notre exemple la protéine ne possède pas de peptide signale, elle a un domaine intracellulaire de l'acide aminé 1 à 25, un domaine transmembranaire de l'acide aminé 26 à 48 et un domaine extracellulaire de l'acide aminé 49 à 539.



Esque votre protéine contient un peptide signale ? Quelle est votre conclusion sur l'adressage de votre protéine.

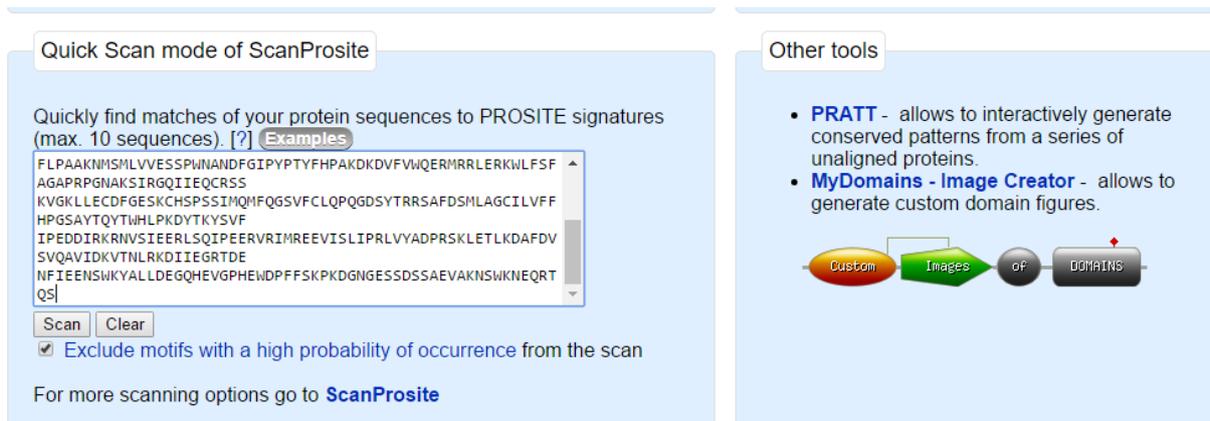
Esque votre protéine contient un/des domaine(s) transmembranaire ? Quelle est votre déduction quant à la localisation de votre protéine ?

Partie III : Travaux pratique TP2

Partie D : recherche de domaine important pour la protéine

33- Copier la séquence de votre protéine à partir du fichier texte **candidatgene_nom du binôme**.

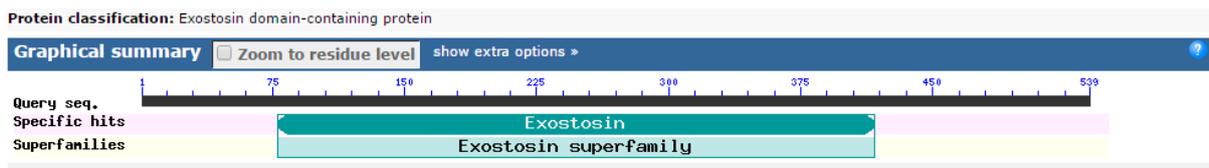
34- Pour prédire les domaines potentielle de la protéine, aller sur le site internet « Prosite» (<http://prosite.expasy.org/>) puis coller votre séquence dans le cadrant de la rubrique « Quick scan mode of ScanProsite ».



The screenshot shows the Prosite website interface. On the left, under 'Quick Scan mode of ScanProsite', there is a text input field containing a protein sequence: FLPAAKNMSMLVVESSPINANDFGIPYPTYFHPAKDKDVFVWQERMRRLERKWLFSF AGAPRPGNAKSIRGQIIIEQCRSS KVGKLLCEDFGESKCHSPSSIMQMFQGSVFC LQPQGD SYTRRS AFDSMLAGCILVFF HPGSAYTQYTHLPLPKDYTKYSVF IPEDDIRKRNVSIEERLSQIPEERVIRIMREEVISLIPRLVYADPRSKLET LKDAFDV SVQAVIDKVTNLRKDIIEGRTE NFIEENSWKYALLDEGQHEVGPHEWDPFFFSKPKDNGESSDSSAEVAKNSWIKNEQRT QS. Below the input field are 'Scan' and 'Clear' buttons, and a checked checkbox for 'Exclude motifs with a high probability of occurrence from the scan'. A link to 'ScanProsite' is also present. On the right, under 'Other tools', there are two bullet points: 'PRATT - allows to interactively generate conserved patterns from a series of unaligned proteins.' and 'MyDomains - Image Creator - allows to generate custom domain figures.' Below the text is a diagram showing a flow from 'Custom' to 'Images' to 'of' to 'DOMAINS'.

35- Cliquer sur « scan » pour lancer l’analyse, dans notre cas, le site Prosite ne prédit pas de domaine important.

36- Une autre manière de trouver des domaines important est d’utiliser le blast protéine de NCBI et ainsi de réitérer les étapes : 19 à 23.



Dans notre exemple, la protéine avec une recherche sur NCBI blast proteine, contient un **domaine Exostosin**.

Es-que votre protéine contient des domaines important ?

Partie III : Travaux pratique TP2

Partie E : recherche bibliographique sur votre gène/protéine ou la famille à laquelle il appartient.

35- En se basant sur les informations obtenus au cours du TD1, chercher sur pubmed (<https://www.ncbi.nlm.nih.gov/pubmed/>) et Google Scholar (<https://scholar.google.fr/>) des publications/revue en rapport avec :

- L'interaction de *Ralstonia* avec ces hôtes.
- Des publications sur votre gène/protéine ou la famille à laquelle il appartient.
- Des publications sur les domaines composant la protéine et leurs fonctions.

Décrire l'interaction entre *Ralstonia* et ces hôtes ?

Quelle est la fonction potentielle de votre protéine ?

En prenant en considération l'expression de votre protéine, proposer une hypothèse quant à son rôle dans l'interaction entre *M. truncatula* et *R. solanacearum* ?

Quelle est le rôle potentiel des domaines de votre protéine dans sa fonction hypothétique dans l'interaction entre *M. truncatula* et *R. solanacearum* ?

N.B : si vous n'arrivez à télécharger un article, veuillez contactez le responsable de l'U.E.

**Partie III : Travaux pratique
TP3**

**Université Ziane Achour
Faculté des sciences de la nature et de la vie
Département de biologie**

Matière : génie génétique

Nature des enseignements dispensés : Travaux pratiques

Niveau : troisième année biologie, spécialité biologie moléculaire

Année universitaire : 2018/2019

Responsable de la matière : Dr. Berrabah Fathi

**TP3 : Analyse phylogénétique et analyse de
l'expression du gène candidat dans des contextes de
pathogénie et de symbiose chez l'espèce modèle *M.
truncatula*.**

Partie III : Travaux pratique TP3

Introduction au TP3

Au cours du TP2 un gène candidat potentiellement important dans la réponse de *M. truncatula* à *R. solanacearum* a été identifié. Vous avez étudié la fonction putative du gène et caractérisé certains aspects structurels de la protéine correspondante. Le TP3 se propose d'aller plus loin dans la caractérisation de votre gène candidat et cela en analysant l'aspect évolutif de la protéine correspondante et d'analyser le profil d'expression du gène en réponse à des pathogènes et des symbiotes.

Le TP est divisé en deux parties :

Partie A : Analyse phylogénétique de la protéine.

Partie B : Analyse de l'expression du gène correspondant.

Partie III : Travaux pratique

TP3

Partie A : analyse phylogénétique de la protéine

1- Dans votre dossier, ouvrir votre fichier texte : **homologue_nom du binôme**, les séquences des protéines ne sont pas organisées, veuillez les organiser en réduisant la taille du titre de la séquence (conserver : >code + le nom de l'espèce, exemple : >XP_003625392.1 *Medicago truncatula*), séparer les séquences des différentes protéines selon le modèle ci-dessous :

>XP_003625392.1 *Medicago truncatula*

```
MRRRPVAAIFSDQMEKSTTKSHNSRIFFLASLSAFFWLLLLYFHFIILPHQQSSTTVNIQQPPPRKHFGFPDI
KKTDPGGRYIYVHNLPSRFNQDMLRECKTSLWLTNMCKFTTNAGLGPPLNIDGVFSDTGWYATNQF
AVDVIFANRMKQYECLTNDSSIAAAVFVFPFYAGFDIARYLWGYNISMRAASVELVDWLMKRPEWGV
MNGRDHFLVAGRITWDFRRLSEDEKDWGNKLLFLPAAKNMSMLVVESSPWNANDFGIPYPTYFHPAK
DKDVFVWQERMRRLERKWLFSFAGAPRPGNAKSIRGQIIEQCRSSKVGKLECDFGESKCHSPSSIMQM
FQGSVFCLQPQGDSYTRRSAFDSMLAGCILVFFHPGSAYTQYTWHLPKDYTKYSVFIPEDDIRKRNVSIEE
RLSQIPEERVIRIMREEVISLIPRLVYADPRSKLETLKDAFDVSVQAVIDKVTNLRKDIIEGRTDENFIEENSWK
YALLDEGQHEVGPHEWDPFFSKPKDNGESSDSSAEVAKNSWKNEQRTQS
```

>GAU19352.1 *Trifolium subterraneum*

```
MRRRPLPGIPSNQMEKSPIKTPNSRIFFLASLSAFFWLLLLYFHFIILPQSSTTTTTTTPRKNFGFPDIINKNPI
QIHQKNLTFITSLETDOCKGRYIYVHDLPSRFNQDMLRDCKSLSLWLTNMCKFVTNAGLGPPLNVDGVFS
DTGWYATNQFAVDVIFANRMKQYECLTNDSSIAAAVFVFPFYAGFDIARYLWGYNISMRAASIDLVDWL
VKRPEWDVMNGRDHFLVAGRITWDFRRLSKEEKDWGNKLLFLPAAKNMSMLVVESSPWNANDFGIPY
PTYFHPAKDEDVFWQERMRRLERKWLFSFAGAPRPNPKSIRGQIIEQCRSSKVGKLECDFGESKCHS
PSSIMQMFQGSFCLQPQGDSYTRRSAFDSMLAGCIPVFFHPGSAYTQYTWHLPKDYTKYSVFIPEDDIR
KRNVSIEERLSQIPEEQVRIMREEVISLIPRLVYADPRSKLETLKDAFDVSVQAVIDKVTNLRKDIIEGRTDEN
FIEENSWKYALLDEGQREAGPHEWDPFFSKPKDNGESSDSSAEAAKNSWKNEQRTE
```

2- Sélectionner l'ensemble des séquences, copier (ctrl+C).

3- Sur votre bureau, chercher et ouvrir le logiciel : Bioedit, cliquer sur « new alignment » : la petite page blanche en haut à gauche.

4- Aller dans « File » et cliquer sur « import from clipboard », vos séquences sont chargées dans le logiciel.

5- Re-cliquer sur « File » puis sur « save as », sélectionner le format FASTA et sauvegarder le fichier dans votre dossier sous le nom : **homologueFASTA+nom binôme**.

6- Fermer le logiciel BioEdit.

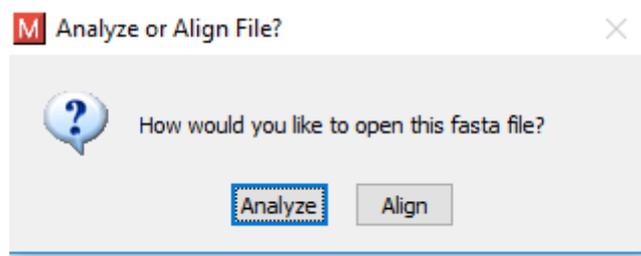
Partie III : Travaux pratique

TP3

7- Dans votre bureau, ouvrir le logiciel Mega7, cliquer sur « Data ; TA » puis sur « open a file/session ».

8- Sélection le fichier « **homologueFASTA+nom binome** », une nouvelle fenêtre s'ouvre.

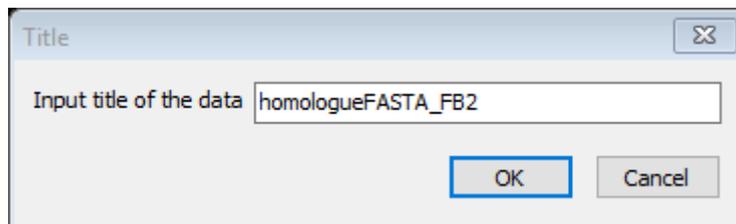
9- cliquer sur « align ».



10- Un fenêtre d'alignement s'ouvre, cliquer sur « W ou aligne by ClusteralW » et choisir les paramètres par défaut.

11- Un alignement est généré, cliquer sur « data » en haut à droite, puis sur « export alignement » et sélectionner « MEGA format ».

12- Aller dans votre dossier et cliquer sur « enregistrer », une nouvelle fenêtre s'ouvre, écrire le nom suivant : **homologueFASTA+nom binome2** puis cliquer sur « OK ».



13- Fermer la fenêtre d'alignement (attention : ne pas fermer le logiciel Mega7)

14- Toujours dans Mega7, cliquer sur « Data ; TA » puis sur « open a file/session » et sélectionnez le document : **homologueFASTA+nom binome2**.

15- cliquer sur « Distance » puis cliquer sur « compute distance pairwise ».

16- Un nouvelle fenêtre apparait, dans « substitutional model » assuré que pour « Model/Metho » la methode sélection soit « p-distance », puis cliquer sur « compute ».

Partie III : Travaux pratique TP3

Option	Selection
Analysis	Distance Estimation
Scope	Pairs of taxa
Estimate Variance	
Variance Estimation Method	None
No. of Bootstrap Replications	Not Applicable
Substitution Model	
Substitutions Type	Amino acid
Model/Method	p-distance
Rates and Patterns	
Rates among Sites	Uniform rates
Gamma Parameter	Not Applicable
Pattern among Lineages	Same (Homogeneous)
Data Subset to Use	
Gaps/Missing Data Treatment	Complete deletion
Site Coverage Cutoff (%)	Not Applicable

17- Un nouvelle fenêtre apparait, elle correspond au résultats du calcul des distances (divergence) entre les différentes protéine.

	1	2	3	4	5	6	7	8	9	10
1. XP 003625392.1 Medicago truncatula										
2. GAU19352.1 Trifolium subterraneum	0.062									
3. XP 004493835.1 Cicer arietinum	0.093	0.089								
4. XP 016207374.1 Arachis ipaensis	0.112	0.105	0.122							
5. XP 015968108.1 Arachis duranensis	0.112	0.105	0.122	0.000						
6. XP 003553511.2 Glycine max	0.126	0.116	0.130	0.110	0.110					
7. XP 017411149.1 Vigna angularis	0.140	0.130	0.138	0.128	0.128	0.089				
8. XP 014496344.1 Vigna radiata var. radiata	0.141	0.134	0.138	0.130	0.130	0.089	0.010			
9. XP 003536950.2 Glycine max	0.124	0.134	0.143	0.101	0.101	0.122	0.138	0.140		
10. XP 004495055.1 Cicer arietinum	0.136	0.136	0.147	0.112	0.112	0.124	0.147	0.149	0.105	

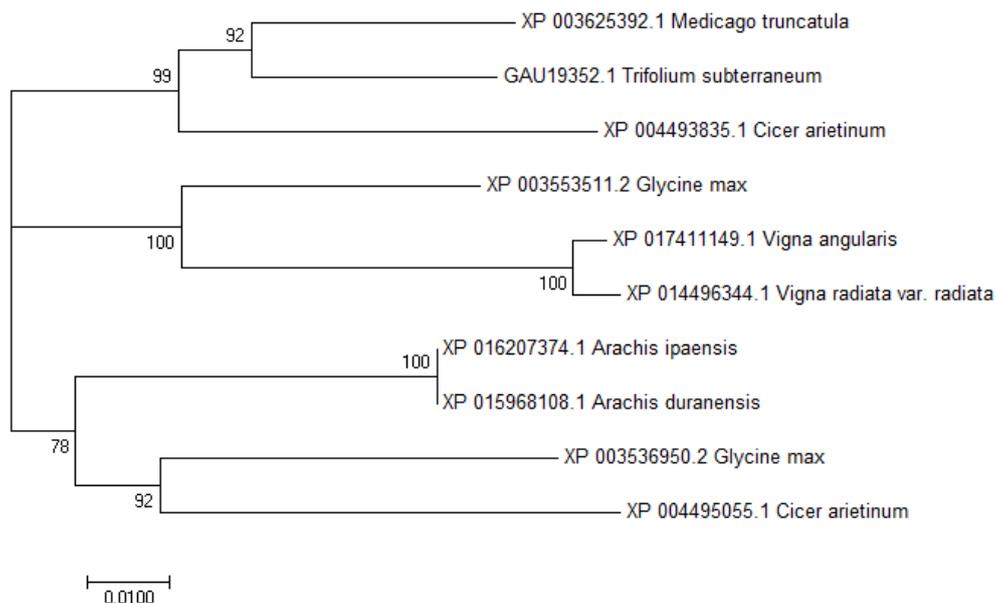
18- retourner dans la page d'accueil de Mega7, puis cliquer sur « Phylogeny » et sélectionner « Construct/test Neighbor-Joining tree.

19- Un nouvelle fenêtre apparait, dans « Phylogeny test », dans « test of phylogeny » sélectionner « bootstrap method », puis dans « substitutional model » dans « Model/Metho » sélectionner « p-distance » enfin cliquer sur « compute ».

Partie III : Travaux pratique TP3

Options Summary	
Option	Selection
Analysis	Phylogeny Reconstruction
Scope	All Selected Taxa
Statistical Method	Neighbor-joining
Phylogeny Test	
Test of Phylogeny	Bootstrap method
<i>No. of Bootstrap Replications</i>	500
Substitution Model	
Substitutions Type	Amino acid
Model/Method	p-distance
Rates and Patterns	
Rates among Sites	Uniform rates
<i>Gamma Parameter</i>	Not Applicable
Pattern among Lineages	Same (Homogeneous)
Data Subset to Use	
Gaps/Missing Data Treatment	Complete deletion
<i>Site Coverage Cutoff (%)</i>	Not Applicable

20- Un arbre phylogenetique apparait avec vos homologues, en bas de la figure apparait la légende de l'arbre phylogenetique qui récapitule la méthode utilisé.



21- Cliquer sur « imprime écran », puis aller dans votre dossier et ouvrez votre document PowerPoint (créée pendant le TP1).

22- Crée une nouvelle diapo et cliquer sur « ctrl+V » afin de coller l'image de votre arbre.

23- Cliquer sur « Ctrl+S » pour sauvegarder.

Analyser et interpréter votre arbre phylogénétique, quelle(s) est/sont votre/vos conclusion(s) ?

Partie III : Travaux pratique TP3

Partie B : analyse de l'expression du gène correspondant.

24- Copier le probset qui correspond à votre gène, pour rappelle voilà la liste des probsets par binôme :

Numéro binôme	Probset
1	Mtr.24731.1.S1_at
2	Mtr.42985.1.S1_at
3	Mtr.18380.1.S1_at
4	Mtr.7244.1.S1_at
5	Mtr.39593.1.S1_at
6	Mtr.8884.1.S1_at
7	Mtr.1633.1.S1_x_at
8	Mtr.10494.1.S1_x_at
9	Mtr.43627.1.S1_at
10	Mtr.27355.1.S1_s_at
11	Mtr.37859.1.S1_at
12	Mtr.51360.1.S1_at

25- Aller sur le site MtGEA (<http://mtgea.noble.org/v3/index.php>), dans « Quick search, probset ID », coller votre probset et cliquer sur « Go ».

26- Un nouvelle fenêtre est ouverte, elle est décrite dans le TP2, étape 11.

27- Cliquer sur « Customize Slides (experiment/sample) », une nouvelle fenêtre apparait, elle permet de sélectionner les expériences que nous souhaitons afficher.

28- Cliquer sur « data source », cela nous permet d'accéder au expériences en fonction de la source dont elles proviennent.

29- Cliquer sur :

sélectionner :

- Balzergue et al

RT_Tip_A17_0hpi
RT_Tip_A17_12hpi
RT_Tip_A17_72hpi

- Benedito et al

RT_0dpi
Nod_4dpi
Nod_10dpi
Nod_14dpi
Nod_28dpi

- Hogekamp Et Al 2011

Root_non_Myc_control_6wk_20_uM_P
Root_28dpi_Myc_G_intraradices_6wk_2

Partie III : Travaux pratique TP3

	Root_28dpi_Myc_G_mosseae_6wk_20
- Mah Et Al., 2012	RT_MPH_0hpi
	RT_MPH_24hpi
	RT_MPH_36hpi
	RT_MPH_48hpi
- Rey Et Al., 2013	RT_A17_mock
	RT_A17_1dpi
- Uppalapati Et Al., 2009	RT_CRR_0hpi
	RT_CRR_72hpi
	RT_CRR_96hpi

30- Cliquer sur « Updeat and stay here ».

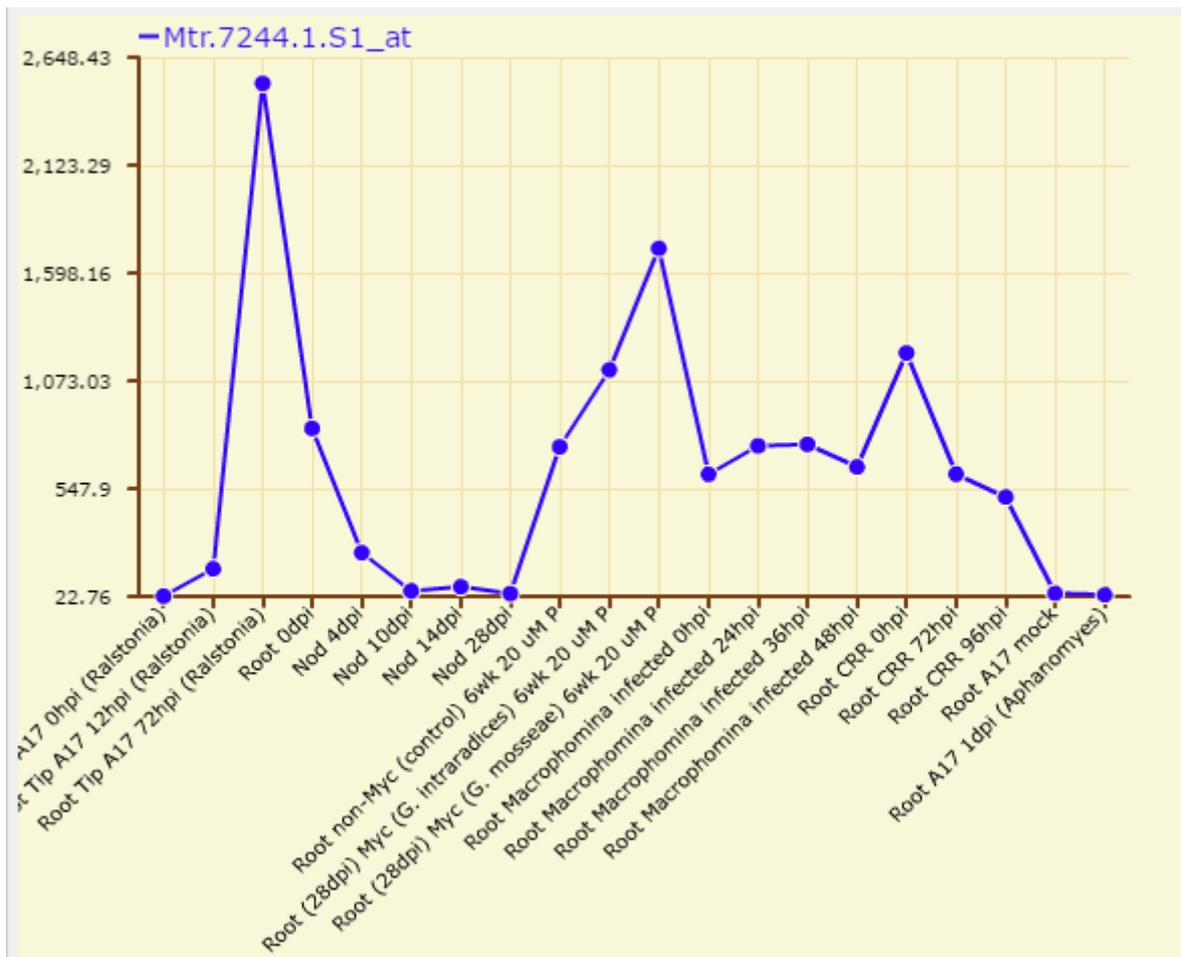
31- retourner sur la page contenant le profil d'expression, puis cliquer sur « rafraichir la page , la flèche circulaire en haut de la page ».

32- le profil d'expression est actualisé et seul apparaissent les conditions sélectionnées.

33- Verifier que toutes les conditions sont représenté, si ce n'ai pas le cas voila la procédure à suivre :

- a- Retourner sur la page de sélection des expériences, désélectionnez toutes les conditions d'une même expérience.
- b- Cliquer sur « updeat and stay here ».
- c- Retournez sur la page contenant le graphe « actualiser la page ».
- d- Revenir sur la page de sélection des expériences.
- e- sélectionner les expériences manquantes.
- f- Cliquer sur « updeat and stay here ».
- g- Retourner sur la page contenant le graphe « actualiser la page ».

**Partie III : Travaux pratique
TP3**



34- Cliquer sur « imprime écran ».

35- Aller sur votre document PowerPoint et créez une nouvelle diapo.

36- Cliquer sur « ctrl+V » pour coller l'image puis cliquer sur « ctrl+S » pour sauvegarder votre résultat.

37- Ci-dessous le tableau qui récapitule le contexte d'interaction (symbiotique ou pathogène) pour chaque expériences :

La condition	Type de microorganisme	Le contrôle est :
Ralstonia	Pathogène	Ralstonia Ohpi
Nod : rhizobium	Symbiote	Roots Odpi
G.intraradices/G. mosseae	Symbiote	roots non-myc (control)
Macrophomina	Pathogène	root Macrophomina Ohpi
CRR : Phymatotricum	Pathogène	Root CRR Ohpi
Aphanomyces	Pathogène	Roots A17 mock

Partie III : Travaux pratique TP3

En se basant sur votre le profil d'expression de votre gène et au vue des données présenter dans le tableau ci-dessus, veuillez :

Analyser vos résultats ?

Qu'en déduisez-vous sur l'expression de votre gène ? Es t-il spécifique de Ralstonia ?
Argumenter votre réponse.

Au vue de l'ensemble des résultats (TP1/TP2/TP3) Présenter une/des hypothèses sur la/les fonction(s) putative du gène et le processus dans lequel il intervient.

**Partie III : Travaux pratique
TP4**

**Université Ziane Achour
Faculté des sciences de la nature et de la vie
Département de biologie**

Matière : génie génétique

Nature des enseignements dispensés : Travaux pratiques

Niveau : troisième année biologie, spécialité biologie moléculaire

Année universitaire : 2018/2019

Responsable de la matière : Dr. Berrabah Fathi

**TP4 : Identification et caractérisation d'un gène
candidat de *R. solanacearum* important pour le
processus infectieux**

Partie III : Travaux pratique TP4

Introduction au TP 4

Après avoir caractérisé au cours des TP1 la réponse de l'hôte *Medicago truncatula* à la bactérie phytopathogène *Ralstonia solanacearum*, identifié et caractérisé au cours des TP2 et 3 un gène candidat de l'hôte pouvant potentiellement participer à l'interaction. Le TP 4 se propose d'identifier un gène candidat de *R. solanacearum* important pour l'interaction avec l'un de ces hôtes, la tomate.

Ci-dessous la description de l'expérience et de la référence bibliographique associée à l'étude :

Summary Investigation of whole genome gene expression level changes in the bacterial wilt pathogen *Ralstonia solanacearum*, strain UW551 at 20°C and 28°C in culture and in planta. The temperate strain UW551 can wilt and cause full-blown disease on tomato plants at 28°C as well as at 20°C.

Overall design A 16 array study using total RNA recovered from the following: 8 separate cultures of *Ralstonia solanacearum* strain UW551 grown in rich medium-CPG (4 grown at 20°C and 4 grown at 28°C) 8 separate in planta samples of *Ralstonia solanacearum* strain UW551 harvested from diseased tomato plants cv. BonnyBest (4 recovered from plants grown at 20°C and 4 from plants grown at 28°C) Each array (4-plex format) measures the expression level of 4318 genes from *Ralstonia solanacearum* strain UW551 with 2 to 6, 40-70 mer probes per gene, with two-fold technical redundancy. The arrays also contain probes for intergenic regions with no technical replicates.

Reference: Jacobs JM, Babujee L, Meng F, Milling A et al. The in planta transcriptome of *Ralstonia solanacearum*: conserved physiological and virulence strategies during bacterial wilt of tomato. *MBio* 2012;3(4). PMID: 22807564

Questions:

- 1- Analysé et résumer la description de l'expérience, pourquoi les auteurs utilisent plusieurs températures ? Quels sont les contrôles auxquels vous allez vous référer ?
- 2- Dans le bureau, vous trouverez un fichier Excel nommé : **Transcriptome Ralstonia**, il contient les éléments suivant :

Partie III : Travaux pratique TP4

- ID du probset.
- La description putative du gène correspondant.
- ORF (Open Reading Fram) qui vous permet d'accéder à la séquence du gène ou de la protéine.
- Les valeurs du signal pour chaque probset par condition (culture des bactéries en milieu liquide à 20°C et 28°C et bactéries prélever sur des plantes infectées à 20°C et 28°C. Les résultats sont en réaliser en 4 réplicas biologiques.

3- Quels est l'intérêt de faire 4 réplicas biologiques ?

4- Faire des calculs des tels sort d'obtenir le log₂ du ratio test/ contrôle.

5- Classer les résultats et sélectionner un gène fortement induit. **Attention il interdit que deux binômes est le même gène.**

6- Avec les outils dont vous disposer identifier la séquence de la protéine et sa fonction putative, quels sont vos hypothèse quant au rôle potentiel de votre gène chez Ralstonia?

7- Chercher les homologues et réaliser un arbre phylogénétique, Analyser et commenter votre résultat.

8- Chercher des domaines important chez votre protéine candidate, analysez et commentez.

9- Chercher un peptide signal et des domaines transmembranaire, analysez et commentez.

Au vue de vos résultats, présentez des hypothèses quant à la fonction de votre gène candidat au cours de l'interaction pathogène-hôte.

Instruction aux auteurs & barème d'évaluation

Chaque binôme devra rédiger un rapport comprenant une introduction (maximum 1 page), matériel et méthode (1 page maximum), les résultats (3 page maximum) et une discussion (1 page maximum). Le TD compte pour 20% de la note finale.

Les questions sont présente pour vous guider, tous vos résultats doivent être sauvegardés et présentés dans votre rapport, n'oublier pas de légénder vos figures. Vous êtes libre de réorganiser les idées et les réponses dans votre rapport temps que cela reste cohérent.

Partie III : Travaux pratique TP4

Vous serais noté sur :

50% sur la qualité des résultats.

25% sur la mise en forme et la présentation des résultats.

25% sur la discussion des résultats.