



HAL
open science

Organisation de la périphérie nucléaire et son influence sur l'expression des gènes chez *Arabidopsis thaliana*

Sarah Mermet

► **To cite this version:**

Sarah Mermet. Organisation de la périphérie nucléaire et son influence sur l'expression des gènes chez *Arabidopsis thaliana*. Génétique des plantes. Université Clermont Auvergne, 2022. Français. NNT : 2022UCFAC123 . tel-04533975

HAL Id: tel-04533975

<https://theses.hal.science/tel-04533975>

Submitted on 5 Apr 2024

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

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

UNIVERSITE CLERMONT AUVERGNE

ECOLE DOCTORALE SCIENCES DE LA VIE, SANTE,
AGRONOMIE, ENVIRONNEMENT

Thèse

Présentée à l'Université Clermont Auvergne
pour l'obtention du grade de

DOCTEUR D'UNIVERSITE

Spécialité : Génétique, Physiologie, Pathologies, Nutrition, Microbiologie Santé,
Innovation

Présentée et soutenue publiquement par

Sarah MERMET

Le 2 Décembre 2022

**Organisation de la périphérie nucléaire et son influence sur
l'expression des gènes chez *Arabidopsis thaliana***

Rapporteurs : Pr Célia BAROUX, Université de Zurich (Suisse)
Dr Cécile RAYNAUD, IPS2 (Paris)
Dr Frederic PONTVIANNE, LGDP (Perpignan)

Examineur Pr Emilie BRASSET, iGrED (Clermont-Ferrand)

Directeur de thèse Pr Christophe TATOUT, iGrED (Clermont-Ferrand)

Encadrant de thèse Dr Emmanuel VANROBAYS, iGrED (Clermont-Ferrand)

Remerciements

Je tiens à remercier tout particulièrement les Docteurs Célia Baroux, Cécile Raynaud et Frédéric Pontvianne d'avoir accepté le rôle de rapporteur ainsi qu'Emilie Brassat pour celui d'examineur de cette thèse afin d'évaluer mon travail.

Je remercie très sincèrement Christophe Tatout et Emmanuel Vanrobays, respectivement mon co-directeur et mon co-encadrant de thèse même si j'aurais préféré dire mes co-directeurs (Manu, tu la passes quand cette HDR ?). Christophe, je tiens à te remercier pour m'avoir donné l'opportunité d'effectuer mon stage de master 2 dans l'équipe puis d'enchaîner sur une thèse. Tu m'as souvent poussé à donner le meilleur de moi-même ainsi que de prendre confiance en moi afin de pouvoir faire ce que je souhaite pour la suite de ma carrière. Merci pour tes conseils et ton aide ainsi que pour ton discours de remerciement qui m'a touché et les propos sont réciproques. Ensuite Manu, je te remercie pleinement pour ton encadrement, tes conseils et ton aide tout au long de ces années. J'ai beaucoup apprécié le binôme que nous formions même si nous n'étions pas d'accord sur les couleurs de mes présentations ; le partage de la paillasse qui finissait très souvent, pour ne pas dire tout le temps, par l'envahissement de ma paillasse ; tes interruptions dans mon bureau tout au long de la journée me manqueront ainsi que les soirées jeux. Enfin, ton discours de remerciements lors de ma thèse m'a réellement touché même si je ne connaissais jamais la fin.

Je remercie l'ensemble des membres de l'équipe Probst-Tatout pour leur bienveillance et leur accueil. La hiérarchie veut que je commence donc par toi Aline, merci pour tes conseils et le temps consacré à mon projet. Un grand merci à Sylvie pour ton aide tout au long de mon projet et notamment dans la réalisation de nombreuses tâches «ingrates» comme les semis et les génotypages pour établir nos 27 nouvelles lignées multi-mutantes, les extractions d'ARN malgré une inversion de colonnes de nos échantillons les plus précieux à envoyer au RNA-seq qui *in fine* n'a pas eu beaucoup d'impact mais aussi pour tes nombreuses venues aux concerts de l'harmonie. Sylviane, les inséparables avec Sylvie, merci pour ton aide pour le seul western blot que j'ai fait durant ma thèse. Merci Sam pour ton aide sur le double-hybride quand Manu n'était pas là et pour ta bonne humeur, toujours en train de fredonner une chanson du matin au soir, le plus souvent Salade de fruit ou le chant des partisans tout en cherchant ta tasse dans tout l'étage. Merci Simon, pour ton aide pour l'immunofluorescence ainsi que pour l'organisation de paris pour l'euro ou la coupe du monde de football (je ne suis pas très douée en pronostic !).

Merci à Sophie et Aurélia pour leur aide dans l'acquisition d'images et leurs traitements. Merci Lauriane pour tes conseils sur la rédaction du manuscrit et la préparation de l'oral de thèse ainsi que pour les démarches de recherche de post-doc. Merci Guillaume, pour ton aide lorsque j'étais bloquée avec les lignes de code et ton fameux « PEBKAC : Problem Exist Between Keyboard And Chair » certes vrai dans la plupart des cas mais pas pour Aradeepopsis qui nous aura occupé de nombreuses heures. Merci Manon de m'avoir aidé et de t'être occupé de mes plantes lors de la préparation de ma présentation de thèse. Merci à Gilles et Léa, même si nous nous sommes peu croisés, bon courage pour votre thèse. Et le meilleur pour la fin, merci Tristan pour tes conseils, les supers moments passés ensemble (souvent salés !), les fameux CTDC et CIDC (je n'écrirai pas ici la signification de ces acronymes que nous avons beaucoup utilisés) et les baignades aux lacs après le boulot.

Je tiens à remercier toutes les personnes qui sont passées par le bureau des étudiants (215), souvent surpeuplé en été, pour les échanges et les moments de rigolade. Je remercie l'ensemble des personnes du 2ème étage pour leurs conseils, discussions et autres, notamment Marie-Jo et Maryse pour les moments de rigolade sur la terrasse et surtout Margaux pour ton soutien, tes conseils, nos pauses sur la terrasse. Je te remercie infiniment pour ton aide pour le rendu de mon manuscrit où rien n'allait mais où tu m'as aidé afin que je puisse partir en vacances sereinement. La vie au labo n'aurait pas été du tout la même sans toi et Tristan.

Merci aux autres thésards de l'institut : Marine, Lisa, Marianne, Guerric, Victor, Ronan, Blandine et tous les autres pour les bons moments passés ensemble, la création des Greduates et leurs fameux jeux tout au long de l'année petit bémol pour le Sumo lors des journées externalisées du GReD qui m'a valu une entorse du pouce et 1 mois et demi sans pouvoir manipuler.

Enfin, je tiens à remercier ma famille qui m'a soutenu durant les moments difficiles et qui a essayé de comprendre en quoi consistait ma thèse, notamment lors des discussions pendant le repas de Noël où je me souviendrai toujours de cette phrase après avoir expliqué ce que je faisais « Sarah, je mange des noyaux de salades ! ». Et surtout un immense merci à Simon, qui m'a soutenu tout au long de ma thèse mais surtout dans la dernière ligne droite où la confiance en moi n'était pas au rendez-vous.

Merci à mes amis, Alex, Cloé, Antho, Marine, Héloïse, Cindy, JB, Julie, Tristan et Margaux qui m'ont soutenu, écouté et souvent permis de décompresser un peu lors de longues soirées de discussions plus ou moins arrosées.

Pour finir, je tiens à remercier mes amis musiciens, notamment le pupitre de trompette/cornet de l'harmonie de Clermont-Ferrand et surtout Laurent et Inge mes deux collègues cornettistes, nous formions le meilleur pupitre de cornet ! Merci pour l'accueil et les bons moments passés durant les répétitions, les concerts et le voyage à Braga. Merci également aux membres de l'harmonie de Chamalières pour leur accueil lors du voyage à Geretsried riche en péripéties et émotions, qui m'a permis de me ressourcer avant d'entamer la dernière étape de rédaction de ma thèse. Un grand merci à Lucile, Téva, David, Sarah, Sarah, Christophe, Pierre, Inge, JC, Marie et Mylène pour le morceau joué lors de ma soutenance qui restera à jamais dans ma mémoire. Un grand merci à Valentine qui a été à l'initiative de cette surprise mais qui n'a malheureusement pas pu jouer à cause du COVID et à Guillaume pour l'arrangement de Victory Célébration réalisé spécialement pour cette occasion.

Résumé

Les eucaryotes organisent et protègent leur génome au sein du noyau, un organite délimité par une enveloppe nucléaire ponctuée par des pores nucléaires qui permettent les échanges nucléocytoplasmiques, et en dessous de laquelle se localise le nucléosquelette, une structure filamenteuse en contact avec le matériel génétique. Enveloppe nucléaire, pores nucléaires et nucléosquelette constituent non seulement des éléments structuraux qui participent à l'établissement de la morphologie du noyau mais sont également suggérés comme constituant la périphérie nucléaire, un domaine nucléaire qui régule l'expression du génome.

Dans ce travail, nous avons utilisé la plante modèle *Arabidopsis thaliana* pour étudier KAKU4 et CROWDED NUCLEI 1-4 (CRWN) qui sont les principales protéines entrant dans la composition du nucléosquelette des plantes. En cherchant à déterminer comment le nucléosquelette se localisait à la périphérie nucléaire, nous avons identifié de courts motifs protéiques dans la région N-terminale de KAKU4 qui sont également retrouvés dans NUP82 et NUP136, deux nucléoporines appartenant au pore nucléaire. Par des méthodes de double hybride, d'expression transitoire et de complémentation fonctionnelle, j'ai montré que ces motifs sont nécessaires à l'interaction physique entre le pore nucléaire et le nucléosquelette et permettent la localisation du nucléosquelette à la périphérie nucléaire. De plus, une analyse phylogénétique suggère que NUP82, NUP136 et KAKU4 dériveraient d'une nucléoporine ancestrale commune à partir de laquelle KAKU4 aurait divergé au cours de l'évolution puis acquis une nouvelle fonction associée au nucléosquelette. J'ai étendu ces liens évolutifs à la protéine CRWN1 qui possède le même motif protéique que KAKU4 que nous avons appelé *peripheral localisation signal* (PLS) et qui détermine pour ces deux protéines l'enrichissement à la périphérie nucléaire. Ces liens évolutifs étroits entre acteurs de la périphérie nucléaire ont été renforcés par la création d'une collection de multi-mutants dont la caractérisation met en évidence l'existence d'interactions génétiques fortes entre le pore nucléaire et le nucléosquelette. Ainsi, les mutations *nup136* et *crwn1* induisent des effets synergiques qui se traduisent par des défauts sévères dans le développement de la plante au stade végétatif et reproductif et une dérégulation forte de gènes impliqués dans la réponse aux stress biotiques et abiotiques, soulignant le rôle essentiel de l'acide salicylique dans les anomalies de développement de nos mutants.

L'ensemble de ces travaux met en évidence une origine évolutive et des fonctions communes entre plusieurs acteurs de la périphérie nucléaire et leur rôle central dans l'intégration des signaux de réponse de la plante lors de stress biotiques ou abiotiques.

Abstract

Eukaryotes organise and protect their genome within the nucleus, an organelle delimited by a nuclear envelope punctuated by nuclear pores that allow nucleocytoplasmic exchanges, and below which is located the nucleoskeleton, a filamentous structure in contact with the genetic material. The nuclear envelope, nuclear pores and nucleoskeleton are not only structural elements that participate in establishing the morphology of the nucleus but are also suggested to constitute the nuclear periphery, a nuclear domain that regulates genome expression.

In this work, we used the model plant *Arabidopsis thaliana* to study KAKU4 and CROWDED NUCLEI 1-4 (CRWN), which are the main proteins that make up the plant nucleoskeleton. In an effort to determine how the nucleoskeleton locates at the nuclear periphery, we identified short protein motifs in the N-terminal region of KAKU4 that are also found in NUP82 and NUP136, two nucleoporins belonging to the nuclear pore. Using two-hybrid, transient expression and functional complementation assays, I showed that these motifs are required for the physical interaction between the nuclear pore and the nucleoskeleton and allow the localisation of the nucleoskeleton at the nuclear periphery. Furthermore, phylogenetic analysis suggests that NUP82, NUP136 and KAKU4 derive from a common ancestral nucleoporin from which KAKU4 diverged during evolution and then acquired a new function associated with the nucleoskeleton. I extended these evolutionary links to the CRWN1 protein, which has the same protein motif as KAKU4, which we called peripheral localization signal (PLS) and which determines enrichment at the nuclear periphery for both proteins. These close evolutionary links between actors in the nuclear periphery have been reinforced by the creation of a collection of multi-mutants whose characterisation highlights the existence of strong genetic interactions between the nuclear pore and the nucleoskeleton. Thus, the *nup136* and *crwn1* mutations induce synergistic effects that result in severe defects in the development of the plant at the vegetative and reproductive stages and a strong deregulation of genes involved in the response to biotic and abiotic stresses, underlining the essential role of salicylic acid in the developmental anomalies of our mutants.

All of this work highlights an evolutionary origin and common functions between several actors of the nuclear periphery and their central role in the integration of plant response signals during biotic or abiotic stress.

Sommaire

REMERCIEMENTS	
RESUME	
ABSTRACT	
SOMMAIRE	
LISTE DES ABREVIATIONS	1
LISTE DES FIGURES	5
LISTES DES TABLEAUX	9
CHAPITRE I : INTRODUCTION BIBLIOGRAPHIQUE	10
I- LE NOYAU	10
I.1) <i>Origine et évolution du noyau des cellules eucaryotes</i>	10
I.2) <i>Le nucléoplasme et son organisation</i>	12
I.2.a) Les corpuscules nucléaires.....	12
I.2.b) La chromatine et son organisation locale.....	15
I.2.b.i) Structure de la chromatine.....	15
I.2.b.ii) Les marques épigénétiques.....	15
I.2.c) L'organisation tri-dimensionnelle de la chromatine au sein du noyau.....	17
I.2.c.i) L'organisation des chromosomes en territoires chromosomiques	17
I.2.c.ii) La compartimentation de la chromatine.....	18
I.2.c.iii) Les domaines de chromatine : TAD, LAD et NAD	20
I.2.c.iv) Les boucles de chromatine et le modèle d'extrusion des boucles	22
I.3) <i>La périphérie nucléaire</i>	23
I.3.a) L'enveloppe nucléaire interne et ses constituants principaux	23
I.3.a.i) Le complexe LINC.....	24
I.3.a.ii) Les protéines NEAP.....	26
I.3.a.iii) La protéine MAN1	26
I.3.a.iv) Les protéines PNET2	27
I.3.b) Le pore nucléaire ou Nuclear Pore Complex (NPC)	28
I.3.b.i) Structure des pores nucléaires	29
I.3.b.ii) Fonctions des pores nucléaires.....	30
I.3.c) Le nucléosquelette.....	32
I.3.c.i) Le nucléosquelette ou lamina chez les animaux	32
I.3.c.ii) Le nucléosquelette chez les plantes.....	33
I.3.c.ii.a) Les protéines CRWN	33
Structure et localisation des protéines CRWN/NMCP	33
Fonction des protéines CRWN.....	35
Interaction des protéines CRWN/NMCP avec la chromatine	37
I.3.c.ii.b) La protéine KAKU4.....	38
I.3.c.iii) Les connexions entre le nucléosquelette et l'enveloppe nucléaire	39

II- RÔLE DES PROTÉINES DE LA PÉRIPHÉRIE NUCLEAIRE DANS L'ORGANISATION DE LA CHROMATINE.....	41
II.3) Rôle du NPC dans la régulation de l'activité transcriptionnelle	41
II.3.a) Premières évidences.....	41
II.3.b) Complexes SAGA et TREX-2	42
II.3.c) Boucles de chromatine.....	43
II.4) Rôle du NPC dans la régulation transcriptionnelle chez <i>Arabidopsis thaliana</i>	43
II.5) Rôle du nucléosquelette.....	46
II.6) Relocalisation des gènes au sein du noyau en fonction de l'environnement	48
III- RÔLE DE LA PÉRIPHÉRIE NUCLEAIRE DANS LA RÉPONSE AUX STRESS CHEZ LA PLANTE	49
III.3) Stress biotique	97
III.3.a) La réponse aux pathogènes chez les plantes.....	49
III.3.b) Voie de réponse médiée par l'Acide Salicylique	50
III.3.c) Implication des protéines CRWN dans la réponse au stress biotique.....	51
III.3.c.i) Les protéines CRWN affectent la voie de l'acide salicylique	51
III.3.c.ii) Les protéines CRWN affectent la voie de l'acide jasmonique	53
III.3.d) Implication du NPC dans la réponse au stress biotique	53
III.4) Stress abiotique	55
III.4.a) Evidences de l'implication du nucléosquelette dans la réponse aux stress abiotiques	56
III.4.a.i) Tolérance à l'excès de cuivre :.....	56
III.4.a.ii) Tolérance à l'excès de sel	56
III.4.a.iii) Tolérance au stress hydrique.....	56
III.4.b) Evidences de l'implication du NPC dans la réponse aux stress abiotiques.....	57
III.4.b.i) Tolérance aux stress hydrique et salin.....	57
III.4.b.ii) Tolérance aux stress thermiques.....	57
III.5) Bilan sur les effets du NPC et du nucléosquelette dans la réponse aux stress	58
IV- PROJET DE RECHERCHE.....	59
CHAPITRE II : RESULTATS	61
I- CONNEXION DU NUCLEOSQUELETTE A LA MEMBRANE NUCLEAIRE INTERNE.....	61
I.1) Article : « Evolutionary conserved protein motifs drive attachment of the plant nucleoskeleton at nuclear pores ».....	62
I.1.a) Résumé de l'article en français.....	62
I.1.b) Article soumis	62
I.1.c) Perspectives	74
I.1.c.i) Analyse transcriptomique des mutants nup136, kaku4 et crwn1	74
I.1.c.ii) Influence du nucléosquelette sur la distribution des pores nucléaires	74
I.1.c.iii) Influence du nuclear basket sur la localisation des composants du nucléosquelette.....	76
I.2) Études de protéines CRWN et leurs connections à périphérie nucléaire.....	77
I.2.a) Structures des protéines CRWN	77
I.2.b) Identification des interactions protéiques impliquant les protéines CRWN	79
I.2.c) Détermination des domaines nécessaires à localisation des protéines CRWN.....	81
I.2.d) Interactions génétiques entre les gènes <i>NUP</i> , <i>KAKU4</i> et <i>CRWN</i>	82
I.2.d.i) Collection de mutants	82

I.2.d.ii) Analyse phénotypique au stade végétatif.....	84
I.2.d.iii) Analyse phénotypique au stade reproductif.....	84
I.2.d.iv) Analyse phénotypique de la morphologie nucléaire.....	86
I.2.d.v) Effets des mutations <i>crwn2</i> , <i>crwn3</i> et <i>crwn4</i>	89
II- IMPACT DES PROTEINES DE LA PERIPHERIE NUCLEAIRE SUR L'EXPRESSION DES GENES	90
II.1) <i>Analyse des mutants des protéines KAKU4, CRWN1 et CRWN4</i>	90
II.1.a) Analyse transcriptomique des mutants <i>kaku4</i> , <i>crwn1</i> et <i>crwn4</i>	90
II.1.b) Caractérisation de la duplication du mutant T-DNA <i>kaku4-2</i>	91
II.2) <i>Comparaison des différents allèles mutants kaku4</i>	92
II.2.a) RNA-seq de trois allèles mutants de <i>kaku4</i> et d'une lignée complémentée	92
II.2.b) Impact sur la morphologie nucléaire des différents allèles de <i>kaku4</i>	96
II.3) <i>Étude des multi-mutants kaku4-2, nup136-2 et crwn1-2</i>	97
II.3.a) Analyse du transcriptome des multi-mutants <i>kaku4-2</i> , <i>nup136-2</i> et <i>crwn1-2</i>	97
II.3.b) Implication de la voie de l'acide salicylique	99
III- ETUDE PRELIMINAIRE DE MAN1, UNE PROTEINE DE LA MEMBRANE INTERNE	101
III.1) <i>Etude de la localisation de MAN1</i>	101
III.2) <i>Caractérisation du mutant man1</i>	103
III.3) <i>Détermination des interactions protéiques de MAN1</i>	104
CHAPITRE III: DISCUSSION ET PERSPECTIVES	106
I- SYNTHÈSE DES RESULTATS ACQUIS AU COURS DE LA THESE	106
II- ROLES DES MOTIFS PROTEIQUES CONSERVES COMMUNS AU NUCLEOSQUELETTE ET AU NPC	106
II.1) <i>Motifs vs domaines protéiques</i>	106
II.2) <i>Rôle des motifs</i>	108
II.2.a) Rôle des motifs dans les interactions protéine-protéine	108
II.2.b) Rôle des motifs dans la localisation à la périphérie nucléaire.....	109
II.2.c) Autres fonctions des motifs	110
II.3) <i>Connexion ente le NPC et le nucléosquelette</i>	110
III- PERIPHERIE NUCLEAIRE ET NUCLEOSQUELETTE : UN ENSEMBLE COMPLEXE DE PHENOTYPES ET D'INTERACTIONS GENETIQUES	111
RÉFÉRENCES.....	116
ANNEXES	131

Liste des abréviations

3C	Chromosome Conformation Capture
3D	Trois dimensions
AA	Acide Aminé
ABA	Abscissic Acid
ABI	ABA Insensitive
ADN	Acide Désoxyribonucléique
AKIP1	A-Kinase Interacting Protein 1
ARN	Acide Ribonucléique
ARN PolII	ARN polymérase II
ARNr	ARN Ribosomique
ARNt	ARN de transfert
BAF	Barrier-of-Autointegration Factor
BiFC	Bimolecular Fluorescence Complementation
BTH	Benzothiadiazole
bZIP	BASIC LEUCINE-ZIPPER
CA	Copper-Associated
CAB	Chlorophyll A/B-Binding
CB	Cajal Body
CBF	C-Repeat-BINDING FACTORS
CC-NB-LRR	Coiled-Coil-NB-LRR
Co-IP	Co-ImmunoPrecipitation
ChIL	Chromatin Integration Labeling
ChIP	Chromatin ImmunoPrecipitation
CM	Chorismate Mutase
CMT	Chromomethylase
COP	Constitutively Photomorphogenic
COPII	Coat Protein Complex II
CPR5	Constitutive expresser of Pathogenesis-Related 5
CR	Cytoplasmic Ring
CRIPSR	Clustered Regularly Interspaced Short Palindromic Repeats
CRWN	CRoWded Nuclei
CT	Chromosome Territories
CTCF	CCCTC-Binding Factor
DamID	DNA adenine methyltransferase identification
DAPI	4', 6-diamidino-2-phénylindole
DCL1	Dicer-Like 1
DET	De-Etiolated
DH	DNase I Hypersensitive
DRM	Domains Rearranged Methylase
EDS	Enhanced Disease Susceptibility
EMS	Méthanesulfonate d'Ethyl

EPS	Enhanced Pseudomonas Susceptibility
ETI	Effector-Triggered Immunity
FISH	Fluorescence <i>in situ</i> hybridization
FRET	Fluorescence Resonance Energy Transfer
FLC	Flowering Locus C
FG	Phénylalanine-Glycine
GAL	Galactokinase
GAR	Glycine-Arginine Rich
GBPL	Guanylate-Binding Protein-Like
GFP	Green Fluorescent Protein
GIP	GCP3-Interacting Protein
GUN	Genomes Uncoupled
HR	Hypersensitive Response
HeH	Helix-extended loop-Helix
HXX	Hexokinase
HYL1	Hyponastic Leaves 1
ICE	Inducer of CBF Expression
ICS	Isochorismate Synthase
IDP	Intrinsically Disordered Protein,
IDR	Intrinsically Disordered Region
IHI	Interactive Heterochromatin Islands
INM	Inner Nuclear Membrane
IR	Inner Ring
IRC	Inner Ring Complex
JA	Jasmonic Acid or Jasmonate
KASH	Klarsicht/Anc1/Syne1 Homology
KD	Knock-Down
KEE	KNOT-Engaged Element
KLS	KNOT Linked Silencing
KO	Knock-Out
LAD	Lamin-Associated Domain
LBR	Lamin B Receptor
LECA	Last Eukaryotic Common Ancestor
LEM	LAP2 β -Emerin-MAN1
LINC	Linker of Nucleoskeleton and Cytoskeleton
LLPS	Liquid-Liquid Phase Separation
LMN	Lamine
LW	Leucine Tryptophane
LWH	Leucine Tryptophane Histidine
LWHA	Leucine Tryptophane Histidine Adénine
MaMYB	Membrane-Anchored MYB
miARN	micro-ARN
MAPK	Mitogen-Activated Protein Kinase
MLKG	Maize Linc Kash Grass-Specific

MMS	Methyl methanesulfonate
MOS	Modifier Of SNC1
MSC	MAN1/SRC1p/C-terminal domain
MT	Microtubules
MTOC	Microtubule Organization Center
MYTH	Membrane-Yeast Two Hybrid
NAD	Nucleolus-Associated Domains
NAP	Nucleosome Assembly Protein
NB-LRR	Nucleotide-Binding site-Leucine-Rich Repeat
NEAP	Nuclear Envelope-Associated Protein
NES	Nuclear Exportation Signal
NLS	Nuclear Localization Signal
NMCP	Nuclear Matrix Constituent Protein 1
NTL	NAC Transcription Factor-Like
NTR	Nuclear Transport Receptors
NOR	Nucleolus Organizer Region
NPR	Nonexpresser of PR genes 1
NPC	Nuclear Pore Complex
NR	Nuclear Ring
NUP	Nucléoporine
ONM	Outer Nuclear Membrane
ORC	Outer Ring Complex
PAMP	Pathogen-Associated Molecular Patterns
PAL	Phenylalanine Ammonia-Lyase
PBS	avrPphB Susceptible
pb	Paire de bases
PC	Plastocyanin
PcG	Polycomb group
PHY	Phytochrome
PIF	Phytochrome Interacting Factors
piRNA	PIWI-interacting RNA
pLAD	plant Lamina-Associated Domains
PL-LFQMS	Proximity Labelling coupled with Label-Free Quantitative Mass Spectrometry
PLS	Peripheral Localisation Signal
PM	Permissive Medium
PML	Promyelocytic Leukemia Protein
PNET	Plant Nuclear Envelope Transmembrane
PNS	Perinuclear Space
PR	Pathogenesis-Related
PRC	Polycomb Repressive complexe
PRR	Pattern Recognition Receptors
pré-ARNm	Précurseur d'ARN messenger
PTI	Pattern-Triggered Immunity

PTM	Post-Translational Modifications
PWO1	PWWP INTERACTOR OF POLYCOMBS1
PWWP	Proline - Tryptophane - Tryptophane – Proline
RAN	RAS-related Nuclear Protein
RAN-GAP	RAN- GTPase Activating Protein
RAN-GEF	RAN-Guanine Nucleotide Exchange Factor
RBSC	Rubisco Small Subunit
RdDM	RNA-directed DNA methylation
RE	Réticulum Endoplasmique
RE-ChIP	Restriction Enzyme mediated-ChIP
ROS	Reactive oxygen species
RNP	Ribonucleoprotein
RS	Arginine-Sérine
SA	Salicylic Acid
SAGA	Spt-Ada-GCN5-acetyltransferase
sARN	small ARN
SID	Salicylic Acid Induction Deficient
siARN	small interfering ARN
SINE	SUN-Interacting Nuclear Envelope Protein
sLiM	Short Linear Motif
SMC	Structural Maintenance Complex
SNC	Suppressor of NPR1-1 Constitutive
SNI	Suppressor of NPR1
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein
snoRNP	small nucleolar ribonucleoprotein
SR	Sérine-Arginine
STED	Stimulated emission depletion
SUN	Sad1/Unc84
SUS	Sucrose Synthase
TAD	Topologically Associating Domain
TCP	Teosinte Branched 1, Cycloidea, And PCF Family
TGA	TGACG Sequence-Specific Binding Protein
TIK	Toll-Interleukin Receptor KASH-domain protein
TIR-NB-LRR	Toll Interleukin-1 Receptor (TIR)-
TM	Transmembrane domain
TREX	TRanscription-coupled Export
WHD	Winged Helix Domain
WIP	WPP-Interacting Protein
WIT	WPP Domain-Interacting Tail-Anchored Protein
WPP	Tryptophane – Proline – Proline
WRKY	WRKY DNA Binding Protein
Y2H	Yeast Two-Hybrid
YFP	Yellow Fluorescent Protein

Liste des Figures

Figure 1.1 : Modèles principaux de l'origine du noyau lors de l'évolution des eucaryotes

Figure 1.2 : Représentation schématique d'un noyau

Figure 1.3 : Principaux corpuscules nucléaires

Figure 1.4 : Structure de base de la chromatine : le nucléosome

Figure 1.5 : Les marques épigénétiques et les états chromatinien

Figure 1.6 : Organisation 3D de la chromatine

Figure 1.7 : Organisation de la chromatine chez l'Homme, *Arabidopsis* et la tomate

Figure 1.8 : L'extrusion de boucles comme modèle pour la formation de TAD.

Figure 1.9 : Représentation schématique de la périphérie nucléaire d'*A.thaliana*

Figure 1.10 : Caractérisation du complexe LINC

Figure 1.11 : Détermination des protéines NEAP

Figure 1.12 : Protéines interagissant avec les lamines chez l'Homme et *Arabidopsis*

Figure 1.13 : Caractérisation des protéines PNET2

Figure 1.14 : Structure des pores nucléaires

Figure 1.15 : Densité et distribution des pores nucléaires

Figure 1.16 : Fonction d'import-export du pore nucléaire

Figure 1.17 : Fonction du NPC dans le développement des plantes

Figure 1.18 : La lamina et ses protéines associées localisées à la périphérie nucléaire

Figure 1.19 : Caractéristiques du nucléosquelette des plantes

Figure 1.20 : Phénotypes développementaux et nucléaires des mutants *crwn*

Figure 1.21 : Caractérisation de la protéine KAKU4

Figure 1.22 : Connexions entre le nucléosquelette et l'enveloppe nucléaire

Figure 1.23 : Conservation des protéines de la périphérie nucléaire au sein du règne végétal

Figure 1.24 : Représentations de la régulation génétique médiée par le pore nucléaire

Figure 1.25 : Le pore nucléaire s'associe avec l'hétérochromatine péri-centromérique

Figure 1.26 : Enrichissement de la périphérie en gènes exprimés

Figure 1.27 : Les protéines CRWN impactent faiblement l'organisation spatiale de la chromatine

Figure 1.28 : CRWN1 s'associe à la chromatine et forme les pLAD

Figure 1.29 : Relocalisation des gènes *CA* et *CAB* à la périphérie nucléaire lors de leur induction transcriptionnelle

Figure 1.30 : Représentation schématique de la réponse immunitaire innée chez la plante

Figure 1.31 : Voie de biosynthèse de l'acide salicylique

Figure 1.32 : Impact de la voie de l'acide salicylique dans les mutants *crwn*

Figure 1.33 : Représentation schématique de la voie de l'acide salicylique chez la plante

Figure 1.34 : Impact de la voie de l'acide salicylique dans les mutants *nup*

Figure 1.35 : Résumé des rôles régulateurs des nucléoporines dans le système immunitaire des plantes

Figure 1.36 : Rôle des protéines CRWN dans la réponse au stress abiotique

Figure 1.37 : Rôle du NPC dans la réponse au stress abiotique

Figure 1.38 : Modèles de régulation de la croissance et de l'immunité par PIF4 en fonction de la température

Figure 1.39 : Résumé des fonctions de la périphérie nucléaire

Figure 2.1 : KAKU4 et CRWN1 dérégulent un ensemble de gènes communs

Figure 2.2 : Essai avec l'anticorps anti-NUP96 (Abcam) d'analyse de la distribution des pores nucléaires

Figure 2.3 : Etude de la répartition nucléaire de NUP136 et KAKU4

Figure 2.4 : Localisation de CRWN1 et KAKU4 dans des protoplastes racinaires d'*Arabidopsis thaliana*

Figure 2.5 : Analyse MEME des protéines CRWN d'*Arabidopsis thaliana*

Figure 2.6 : Interactions de CRWN1 avec elle-même, les autres protéines CRWN, KAKU4 et NUP

Figure 2.7 : Identification de motifs conservés dans les protéines CRWN chez *Arabidopsis thaliana*

Figure 2.8 : Localisation de CRWN1 et présence de motifs conservés communs dans les protéines CRWN, KAKU4, NUP82 et NUP136

Figure 2.9 : Etude de la taille de la rosette dans les différents mutants *nup136*, *kaku4* et *crwn1*

Figure 2.10 : Taille des différents mutants à un stade tardif du développement

Figure 2.11 : Etude de la taille des siliques dans les différents mutants *nup136*, *kaku4* et *crwn1*

Figure 2.12 : Etude du nombre de graines et de leur taille des différents mutants *nup136*, *kaku4* et *crwn1*

Figure 2.13 : Etude de la viabilité du pollen des différents mutants *nup136*, *kaku4* et *crwn1*

Figure 2.14 : Morphologie nucléaire des mutants *nup136*, *crwn1* et *nup136 crwn1*

Figure 2.15 : Organisation de la chromatine dans les différents mutants *nup136*, *kaku4* et *crwn1*

Figure 2.16 : Taille des différents mutants à un stade tardif du développement

Figure 2.17 Phénotypes des mutants impliquant les composants du nucléosquelette

Figure 2.18 : Impact des protéines KAKU4, CRWN1 et CRWN4 dans la réponse aux stress.

Figure 2.19 : Présence d'une duplication sur le chromosome 4 dans le mutant *kaku4-2*

Figure 2.20 : Analyse transcriptomique des différents allèles de *kaku4*

Figure 2.21 : Comparaison des gènes dérégulés dans les mutants *kaku4-2* dans deux expériences différentes

Figure 2.22 : Analyse transcriptomique des différents allèles de *kaku4-2* et du mutant complémenté *kaku4-2 ; pKAKU4-KAKU4-tRFP*

Figure 2.23 : Morphologie nucléaire des différents allèles de *kaku4*

Figure 2.24 : Analyse transcriptomique des multi-mutants impliquant les gènes *NUP136*, *KAKU4* et *CRWN1*

Figure 2.25 : Gene ontology des gènes communs aux mutants *nup136*, *kaku4* ou *crwn1*

Figure 2.26 : Comparaison des gènes dérégulés entre mutants

Figure 2.27 : Comparaison des gènes dérégulés entre nos mutants et ceux d'un traitement à l'acide salicylique

Figure 2.28 : Implication de la voie de l'acide salicylique dans le mutant *nup136 crwn1*

Figure 2.29 : Complémentation partielle du phénotype du mutant *nup136 crwn1* par des mutants de la voie du SA

Figure 2.30 Localisation de la protéine MAN1

Figure 2.31 : Etude du développement des mutants CRISPR *man1*

Figure 2.32 : Impact de MAN1 sur la morphologie nucléaire

Figure 3.1 Classification fonctionnelle des SLiM

Figure 3.2 Schéma récapitulatif de la fonction des motifs des protéines KAKU4, NUP136 et NUP82

Figure 3.3 : Etapes du crible EMS en cours sur le mutant *nup136 crwn1*

Listes des Tableaux

Tableau 1 : Récapitulatif des interactions double-hybride entre CRWN1 (appât) et les l'ensemble des protéines CRWN et leurs dérivés (proies)

Tableau 2 : Récapitulatifs des interactions double-hybride entre CRWN1 et KAKU4-NUP

Tableau 3 : Récapitulatifs des interactions double-hybride entre CRWN1 et KAKU4-NUP

Tableau 4 : Tableau récapitulatif de la localisation des motifs conservés retrouvés chez les protéines CRWN-NUP-KAKU4 d'*A.thaliana*

Tableau 5 : Récapitulatif des lignées mutantes créées

Tableau 6 : Liste des 24 gènes dérégulés communs dans les différents allèles de *kaku4*

Tableau 7 : Bilan du crible MYTH avec la banque d'ADNc CD4-10 sur les appâts MAN1

Tableau 8 : Récapitulatif des interactions de MYTH entre MAN1 et CRWN-MAN1

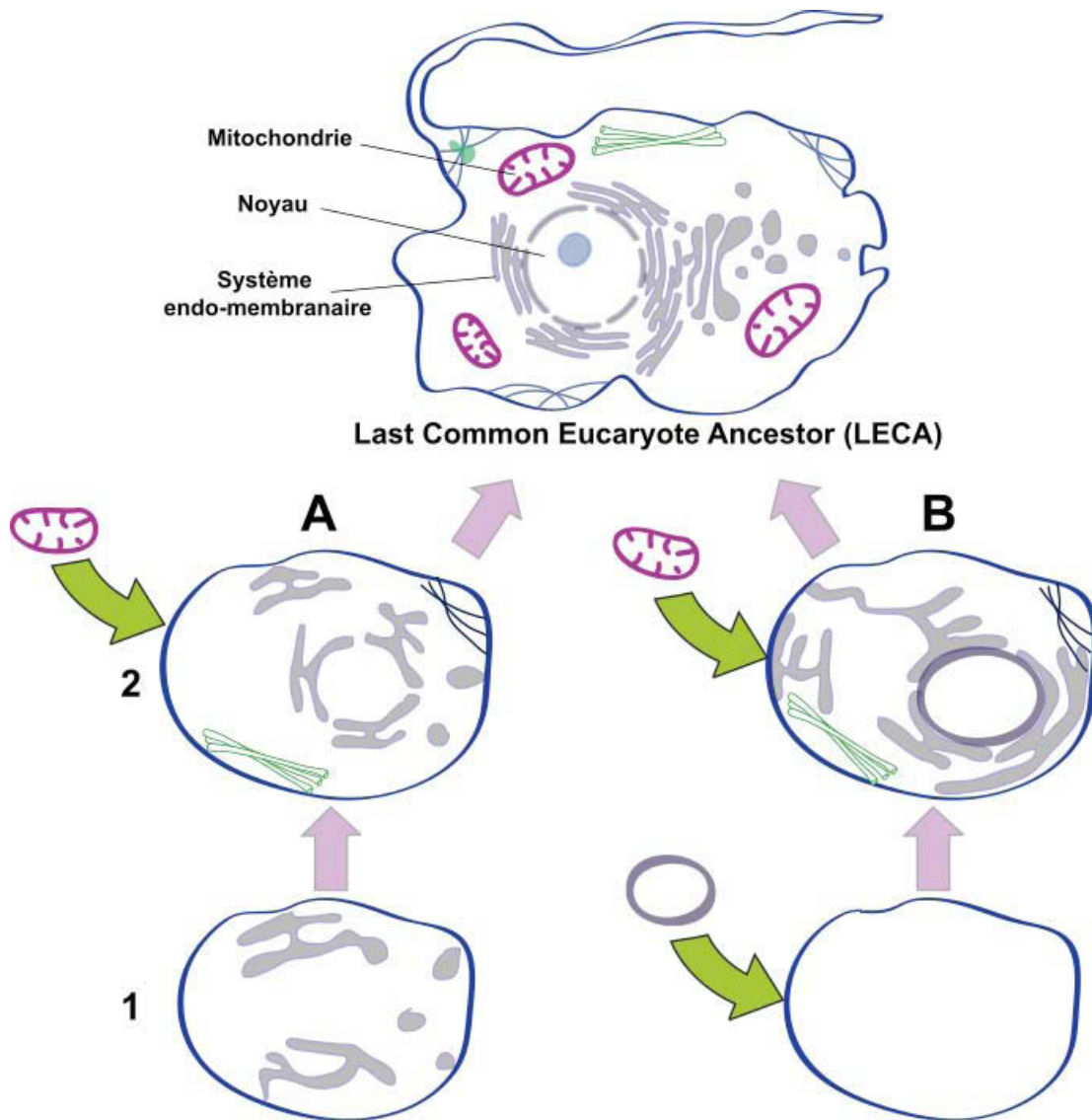


Figure 1.1 : Modèles principaux de l'origine du noyau lors de l'évolution des eucaryotes

A. Modèle autogène : à partir d'une lignée proto-eucaryotique ancestrale, une différenciation progressive du système endomembranaire (**A1**) permet de générer le système membranaire endoplasmique, puis l'enveloppe nucléaire suite à une accumulation de vésicules endoplasmiques autour du matériel génétique (**A2**). **B.** Modèle endosymbiotique : le noyau est issu de la symbiose et l'assimilation d'une archée par une lignée proto-eucaryotique ancestrale (**B1**), le système endomembranaire évolue ensuite pour transporter les enzymes hydrolytiques de l'archée vers le périplasme (**B2**). Dans les deux cas, les mitochondries proviennent de l'assimilation d'une alpha-protéobactérie probablement endosymbiote, permettant finalement l'apparition de l'eucaryote ancestral : le *last eukaryotic common ancestor* (LECA).

Chapitre I : Introduction Bibliographique

I- Le noyau

I.1) Origine et évolution du noyau des cellules eucaryotes

Le noyau constitue le critère principal différenciant les organismes eucaryotes et procaryotes. Alors que les organismes procaryotes ne possèdent pas de structure spécialisée pour stocker et protéger leur matériel génétique, laissé alors libre dans le cytoplasme, les eucaryotes organisent et protègent leur génome au sein du noyau délimité par une double membrane (interne : *Inner Nuclear Membrane*, INM et externe : *Outer Nuclear Membrane*, ONM) ponctuée par des pores nucléaires et dont la membrane externe est en continuité avec le réticulum endoplasmique (RE).

L'apparition du noyau correspond à une évolution majeure de l'organisation de la cellule eucaryote, les nombreux scénarii proposés afin d'expliquer l'origine de l'enveloppe nucléaire et par extension du noyau, peuvent être divisés en deux modèles, endosymbiotique et autogène (Devos, Gräf and Field, 2014; Baum, 2015; López-García and Moreira, 2015; Martin, Garg and Zimorski, 2015) (**Figure 1.1**). Les hypothèses basées sur le modèle endosymbiotique, envisagent une origine bactérienne du noyau, *via* l'internalisation au sein du cytoplasme, d'une archée avec laquelle elle entretenait une relation endosymbiotique, c'est-à-dire une coopération mutuellement bénéfique aux deux organismes. Des hypothèses similaires expliquent l'apparition des mitochondries et des chloroplastes par endosymbiose. Ces théories sont étayées par le fait que la double membrane des mitochondries et des chloroplastes présente des similitudes avec la structure des alpha-protéobactéries dont elles seraient issues (Martin, Garg and Zimorski, 2015). Toutefois, que la membrane nucléaire soit une double membrane, n'explique pas l'origine du système endomembranaire eucaryote (réticulum endoplasmique) et sa continuité avec l'enveloppe nucléaire ainsi que la présence des pores nucléaires. Une alternative est le modèle autogène qui repose sur des modifications structurelles par étapes à partir d'un proto-eucaryote ancestral (Baum, 2015). Une différenciation progressive du système endomembranaire, initialement restreinte au niveau de la membrane plasmique aurait donné naissance au système membranaire endoplasmique, puis à l'enveloppe nucléaire suite à une accumulation de vésicules endoplasmiques autour du matériel génétique (Cavalier-Smith, 2002). Cette hypothèse repose sur la continuité entre ces deux systèmes endomembranaires et suppose que l'enveloppe nucléaire a été créée suite à la fusion de vésicules endoplasmiques.

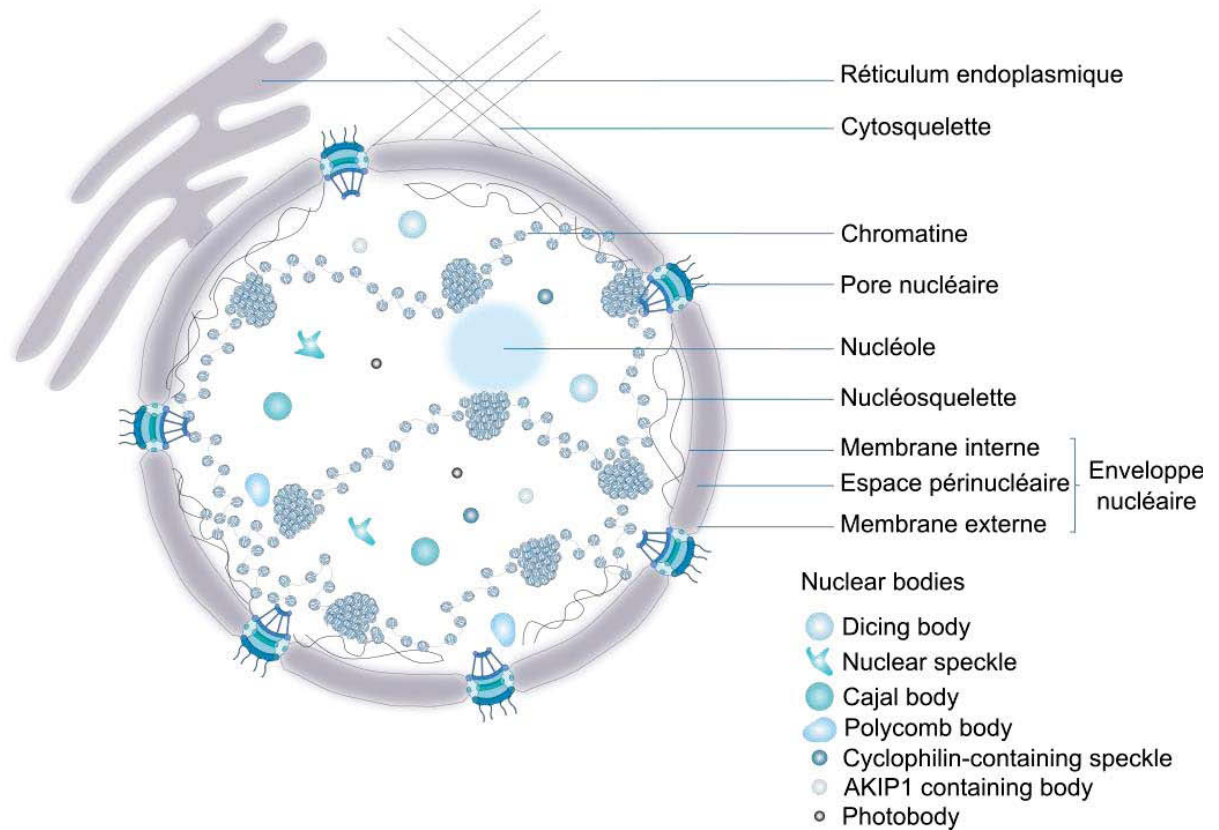


Figure 1.2 : Représentation schématique d'un noyau

L'enveloppe nucléaire est constituée d'une membrane interne (INM) et externe (ONM). Les deux membranes sont séparées par l'espace périnucléaire (PNS). L'enveloppe nucléaire est ponctuée de nombreux pores nucléaires (*Nuclear Pore Complex*, NPC). L'ONM est en continuité avec le réticulum endoplasmique (RE). A l'intérieur du noyau, le nucléosome est localisé à la périphérie nucléaire en contact avec, d'une part la chromatine et d'autre part l'INM. Dans le nucléoplasme, différents corpuscules nucléaires ou *nuclear bodies* sont présents assurant des fonctions diverses comme l'épissage ou la maturation de l'ARN. Dans le cytoplasme, le cytosquelette constitué de microtubules, de microfilaments d'actine et de filaments intermédiaires, assure la connexion entre l'ONM et la membrane plasmique (non représentée).

Cela permet également d'expliquer l'apparition des pores nucléaires, qui dériveraient de protéines de recouvrement vésiculaires de type COPII (*Coat Protein Complex II*) qui auraient participées au regroupement de «macro-vésicules» constituant une première étape dans l'assemblage de l'enveloppe nucléaire (Schwartz, 2013). En résumé, chacun des deux modèles semble pouvoir expliquer en partie la formation des cellules eucaryotes avec une origine endosymbiotique pour les mitochondries et les chloroplastes alors que l'enveloppe nucléaire en continuité avec le réticulum endoplasmique provient d'une différenciation de la membrane plasmique.

Schématiquement, le noyau comme on le trouve chez les eucaryotes actuels, peut être divisé en deux grandes régions, le nucléoplasme qui constitue la région interne du noyau et la périphérie nucléaire constituée de l'enveloppe nucléaire, de pores nucléaires et d'un réseau de protéines proches de l'enveloppe nucléaire appelé le nucléosquelette (**Figure 1.2**). Au sein du nucléoplasme, le matériel génétique est organisé sous forme de chromatine qui constitue l'enroulement de l'ADN avec les protéines histones (cf paragraphe I.2.b). Lors de la division cellulaire la chromatine est très compactée mais lors de l'interphase elle peut adopter deux grands états selon son niveau de compaction visibles par microscopie après coloration de l'ADN. Le premier état est un état condensé appelé hétérochromatine qui se localise principalement proche de l'enveloppe nucléaire et du nucléole (S. Feng *et al.*, 2014; Pontvianne *et al.*, 2016; Poulet, Duc, *et al.*, 2017a), ces régions de chromatine ont une accessibilité réduite de l'ADN aux facteurs de transcription, se traduisant par une faible activité transcriptionnelle. Quant à lui, le second état est constitué de chromatine décondensée, l'euchromatine, située à l'intérieur du noyau et accessible aux facteurs de transcription permettant ainsi une transcription plus active. La localisation préférentielle de l'hétérochromatine au sein du noyau est proche de la périphérie nucléaire ce qui suggère que cette proximité pourrait influencer l'organisation de la chromatine en permettant la formation ou le maintien de l'hétérochromatine et, de ce fait, d'impacter son expression (cf paragraphe I.2.c).

Outre son rôle dans la protection du matériel génétique, le noyau permet de compartimenter certains processus fondamentaux de la cellule en son sein comme la transcription, l'épissage des ARN, la réplication ou la réparation de l'ADN et d'assurer leur bon déroulement. La présence du noyau chez les eucaryotes a donc permis de créer un espace cellulaire délimité au sein duquel se déroulent des fonctions spécifiques dédiées au contrôle de la « machinerie ADN/ARN ». Toutefois, les échanges nucléo-cytoplasmiques sont primordiaux pour la vie de la cellule, en particulier lors du transfert des ARNm en vue de leurs traductions dans le

cytoplasme ou lors de l'importation nucléaire de facteurs de régulation de la transcription ou de la réplication. Afin d'assurer ces échanges, l'enveloppe nucléaire est ponctuée de plusieurs milliers de pores nucléaires (*Nuclear Pore Complex*, NPC), au travers desquels certaines molécules sont transportées d'un compartiment cellulaire à un autre (nucléoplasme/cytoplasme).

Selon les organismes et le type cellulaire, le noyau présente une taille variable de quelques micromètres à plusieurs dizaines de micromètres. Sa forme est également variable allant de petits noyaux plutôt sphériques souvent associés à des tissus en division active, à des noyaux plus gros et plus allongés présents dans certains tissus différenciés. Ces variations de formes et de tailles sont contrôlées par des forces exercées par les protéines du cytosquelette (filaments dans le cytoplasme) et du nucléosquelette (filaments dans le noyau) sur l'enveloppe nucléaire.

Dans un premier temps, je vais décrire les différents composants du nucléoplasme et l'organisation de la chromatine puis les protéines qui composent la périphérie nucléaire, l'enveloppe nucléaire, les pores nucléaires et enfin les protéines du nucléosquelette.

[I.2\) Le nucléoplasme et son organisation](#)

Le nucléoplasme constitue la région interne et fluide du noyau dans laquelle se trouve la chromatine et les protéines nécessaires au métabolisme de l'ADN et de l'ARN formant des agrégats appelés corpuscules nucléaires.

[I.2.a\) Les corpuscules nucléaires](#)

Au sein du nucléoplasme, de nombreux corpuscules nucléaires (nuclear bodies) dépourvus de membrane et aux fonctions très spécifiques sont retrouvés tels que le nucléole ou les corps de Cajal (**Figure 1.2**). Deux modèles permettent d'expliquer la formation des corpuscules nucléaires, le premier est le modèle d'assemblage stochastique ou aléatoire et le deuxième est le modèle d'assemblage ordonné (Dundr and Misteli, 2010; Mao, Zhang and Spector, 2011). Le modèle stochastique consiste en l'assemblage dans un ordre aléatoire des différents composants du corpuscule alors que le modèle ordonné permet l'assemblage des différents composants dans un ordre précis et en absence de l'une des étapes, le corpuscule ne peut pas être formé (Dundr and Misteli, 2010; Mao, Zhang and Spector, 2011).

Le nucléole est le corps nucléaire le plus gros et le plus visible du noyau des cellules eucaryotes (**Figure 1.3**). Cet espace particulier du noyau est composé de protéines et d'ARN ribosomiques et est le siège d'une transcription très active notamment des gènes d'ARNr 45S,

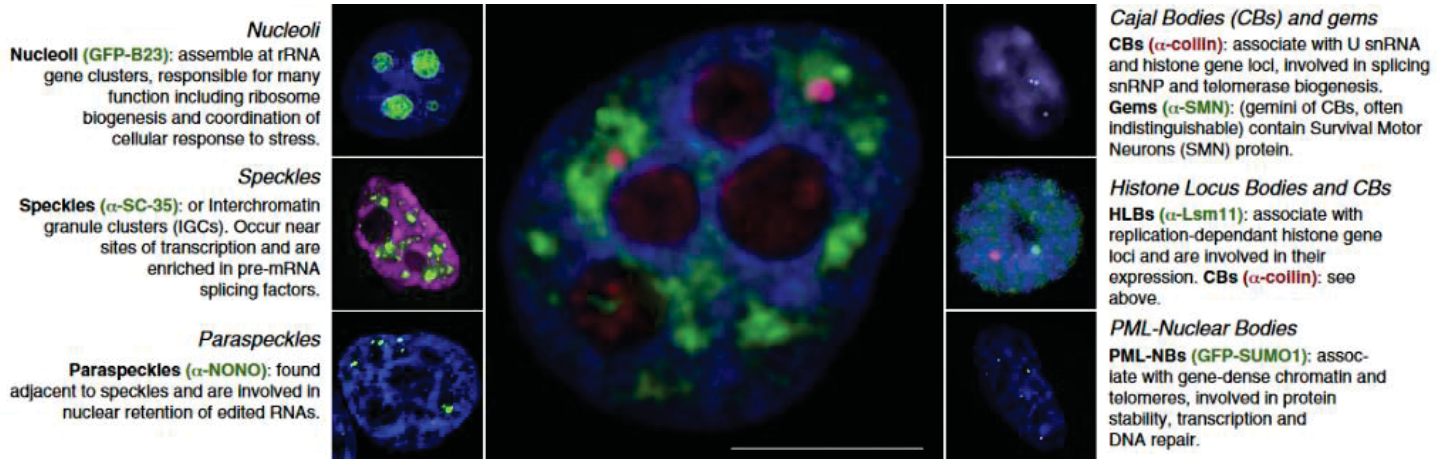


Figure 1.3 : Principaux corpuscules nucléaires (Sleeman and Trinkle-Mulcahy, 2014)

Le panneau central montre l'organisation spatiales de différentes structures dans le noyau d'une cellule HeLa. La chromatine est colorée avec du DAPI (bleu) dans une cellule exprimant une *small nuclear ribonucleoprotein* (snRNP) appelée SmB qui est marquée avec la YFP (vert). Les composants fibrillaires denses des nucléoles sont colorés avec des anticorps contre la fibrillarine (rouge), tandis que l'anti-fibrillarine et l'anti-coiline colorent tous deux les corps de Cajal (rose). Barre = 10 μm.

Les panneaux latéraux plus petits montrent le nucléole dans les cellules HeLa (avec un marqueur pour le composant granulaire), les *speckles* dans les cellules épithéliales du cristallin humain, les *paraspeckles* dans les cellules HeLa, les corps de Cajal (CB) dans les cellules de neuroblastome SH-SY5Y, les corps de locus d'histone (HLB) et les corps de Cajal dans les tissus de drosophile et les corps nucléaires de *PROMYELOCYTIC LEUKEMIA PROTEIN* (PML) dans les cellules HeLa.

regroupés sur des régions chromosomiques particulières appelées les régions NOR (*Nucleolus Organizer Regions* ou Régions Organisatrices du Nucléole) retrouvées proche du nucléole et certainement à l'origine de la formation du nucléole. La fonction principale du nucléole est la transcription des pré-ARNr par l'ARN polymérase I, formant alors l'ARN précurseur 45S, qui sera ensuite clivé pour former les ARNr matures 25S, 18S et 5.8S chez la plante. Le nucléole constitue le premier lieu de maturation des ARNr et d'assemblage des différentes sous-unités des ribosomes, ces dernières étant ensuite exportées dans le cytoplasme *via* les pores nucléaires. En dehors du nucléole, les corpuscules nucléaires sont retrouvés chez l'ensemble des eucaryotes mais varient par leur type et/ou leur fonction selon les organismes. Le nucléole, les corps de Cajal (Cajal bodies), les *nuclear speckles* et les *Polycomb bodies* sont les principaux corpuscules nucléaires (**Figure 1.3**).

Chez les plantes, les corps de Cajal forment 1 à 2 foyers par noyau, ils sont dynamiques, ils peuvent se diviser ou fusionner ou encore être inclus dans le nucléole. Ils jouent un rôle important dans le métabolisme de l'ARN et la formation des particules ribonucléoprotéiques (RNP) impliquées dans la transcription, l'épissage, la biogenèse des ribosomes et le maintien des télomères. Le marqueur principal des corps de Cajal est la coiline qui possède une affinité forte et non spécifique pour les ARN dont les *small nuclear RNA* (snRNA). En plus de la coiline, les corps de Cajal contiennent, des protéines impliquées dans la formation de petits ARN associés à des protéines nucléaires (snRNP) nécessaires à l'épissage des pré-ARNm, de petits ARN associés à des protéines nucléolaires (snoRNP) impliquées dans la maturation des ARNr, ARNt, snARN comme la fibrillarine (**Figure 1.3**) (Shaw and Brown, 2004; Love *et al.*, 2016).

Les *nuclear speckles* sont des amas de protéines disposées à l'intérieur du noyau et sont principalement constitués de protéines impliquées dans l'épissage des ARN comme certains facteurs d'épissage qui contiennent des motifs de liaison à l'ARN et des régions riches en sérine-arginine (SR) ou en arginine-sérine (RS) (**Figure 1.3**). Les protéines SR sont nécessaires à la reconnaissance de l'intron, à la définition de l'exon et à l'assemblage du spliceosome. Chez les animaux mais également chez les plantes, ces protéines sont stockées dans les *nuclear speckles* avant d'être relocalisées à l'intérieur du noyau vers les régions de transcription active, ce qui suggère que les *nuclear speckles* sont des sites de stockage et d'assemblage des composants du spliceosome (Shaw and Brown, 2004).

Les protéines du groupe Polycomb (PcG) forment deux types de complexes appelés complexes répressifs Polycomb 1 et 2 (PRC1 et PRC2) impliqués dans la répression

transcriptionnelle des gènes et régulant ainsi l'expression d'une multitude de gènes. Les protéines PcG sont organisées en domaines nucléaires distincts appelés *Polycomb* ou *PcG bodies*, qui sont souvent localisés près de l'hétérochromatine péri-centromérique (Mao, Zhang and Spector, 2011), il a également été observé des *speckles* d'H3K27me3, suggérant qu'il y a un regroupement des gènes pour ou après le dépôt de cette marque (Y. Huang *et al.*, 2021).

Il existe également des corpuscules nucléaires spécifiques des plantes comme les *Dicing-bodies* ou *D-bodies* qui sont spécialisés dans la biosynthèse des microARN (miRNA) et contiennent les protéines DICER-LIKE1 (DCL1), une ribonucléase et HYPONASTIC LEAVES 1 (HYL1) une protéine de liaison à l'ARN double brin. Il y a entre 1 à 4 structures par noyau de 0,2 à 0,8 µm de diamètre. Les *phytochromes nuclear bodies*, comme leur nom l'indique, sont composés de phytochromes qui sont les photorécepteurs de la lumière rouge et rouge lointain régulant de nombreux aspects du développement des plantes en transformant les signaux lumineux en changements de l'expression génétique. La lumière induit l'importation des phytochromes dans le noyau, où ils forment des complexes protéiques et s'accumulent dans des domaines nucléaires en forme de corpuscules. La fonction biologique de ces domaines est inconnue mais leur formation est liée à la réponse à la lumière (Shaw and Brown, 2004; Liu, Shi and Fang, 2012; Liu *et al.*, 2014). Il existe deux autres types de corpuscules spécifiques des plantes : les *AKIP1 containing nuclear body* et les *Cyclophilin-containing domains*, leurs fonctions sont encore peu connues mais il semblerait qu'ils soient impliqués dans la protection des ARNm face au stress cellulaire pour le premier et, pour le second, dans le stockage des protéines cyclophilines, jouant ainsi un rôle dans le repliement des protéines et un rôle de chaperonne en se liant à des séquences riches en proline et en catalysant des réarrangements structuraux notamment pour les protéines SR (Shaw and Brown, 2004).

Ces différents corpuscules nucléaires présents dans le nucléoplasme, jouent des rôles variés aussi bien dans la transcription par l'ARN PolIII, la maturation des ARNr par le nucléole, l'épissage, la protection des ARNm ou encore la synthèse de miRNA. Les mécanismes impliqués dans la formation de ces corpuscules sont encore largement méconnus mais pourraient faire intervenir des propriétés d'agrégation de certaines protéines capables d'effectuer de la transition de phase, en isolant les protéines impliquées du nucléoplasme mais sans formation de membrane. Cette compartimentation des corpuscules dans l'espace nucléoplasmique est encore un vaste champ d'investigation mais pourrait donner des idées également sur les modalités de formation d'autres compartiments cette fois formés par la chromatine (Pontvianne and Liu, 2020).

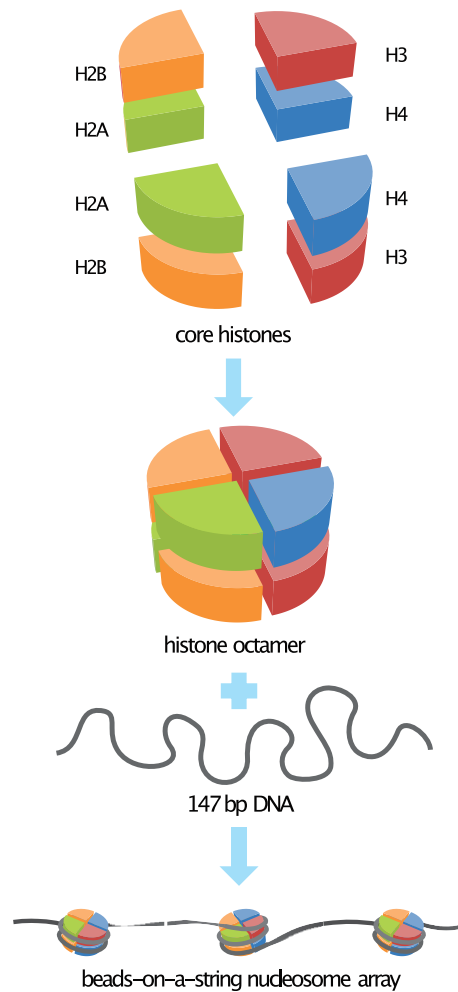


Figure 1.4 : Structure de base de la chromatine : le nucléosome (adapté de Morgan, 2007).

Le nucléosome est constitué d'un octamère d'histone composé d'un tétramère d'histone H2A-H2B et d'un autre formé avec les histones H3-H4, autour duquel s'enroule 147 bp d'ADN. L'enchaînement des nucléosomes forme la chromatine, un nucléofilament similaire à un « collier de perles » qui peut ensuite être plus ou moins compacté.

I.2.b) La chromatine et son organisation locale

I.2.b.i) Structure de la chromatine

Au sein du noyau, l'ADN est organisé sous forme de chromatine, ayant comme unité de base le nucléosome qui forme le premier niveau de compaction de l'ADN. Le nucléosome est constitué d'un octamère d'histones composé des histones H2A, H2B, H3 et H4, chacune en double exemplaire formant ainsi le cœur du nucléosome autour duquel s'enroule 146 à 147 paires de bases (pb) d'ADN (**Figure 1.4**). L'histone H1 ou *histone-linker*, se fixe sur les ~30 pb d'ADN libre entrant et sortant d'un même nucléosome afin de limiter les mouvements de ce dernier. L'enchaînement des nucléosomes forme un nucléofilament initialement décrit comme un « collier de perles » qui peut ensuite être plus ou moins compacté (**Figure 1.4**). Le niveau de compaction de l'ADN modifie son accessibilité à la machinerie transcriptionnelle constituant ainsi l'un des mécanismes de contrôle de l'expression des gènes. La chromatine condensée ou hétérochromatine contient des régions réprimées alors que la chromatine décondensée ou euchromatine contient des régions transcrites (**Figure 1.5A**). La compaction de la chromatine et donc la régulation de la transcription est également régulée par des marques épigénétiques, modifications réversibles et transmissibles, comme la méthylation ADN au niveau des cytosines, des modifications post-traductionnelles des histones ou encore l'incorporation de variants d'histones. Ces marques épigénétiques peuvent impacter localement la stabilité du nucléosome et / ou l'accessibilité des facteurs de transcription à la chromatine et donc à l'expression des gènes (cf paragraphe I.2.c).

I.2.b.ii) Les marques épigénétiques

Il existe différentes modifications de la chromatine influençant l'expression des gènes qui peuvent être répressives ou activatrices/permisives pour la transcription. Ces marques vont moduler les interactions entre l'ADN et les histones, changeant la stabilité des nucléosomes et par conséquent, la compaction de la chromatine. Les marques épigénétiques peuvent être déposées au niveau de l'ADN ou des histones.

La méthylation de l'ADN, chez les eucaryotes, s'effectue majoritairement au niveau des cytosines, et est déposée par différentes protéines méthyltransférases en fonction du contexte dans lequel se trouve la cytosine (CG, CHG ou CHH avec H une base autre que G). Elle est principalement retrouvée dans les séquences répétées comme les éléments génétiques mobiles (les éléments transposables) ainsi que dans les promoteurs des gènes où elle a plutôt un effet répressif sur la transcription.

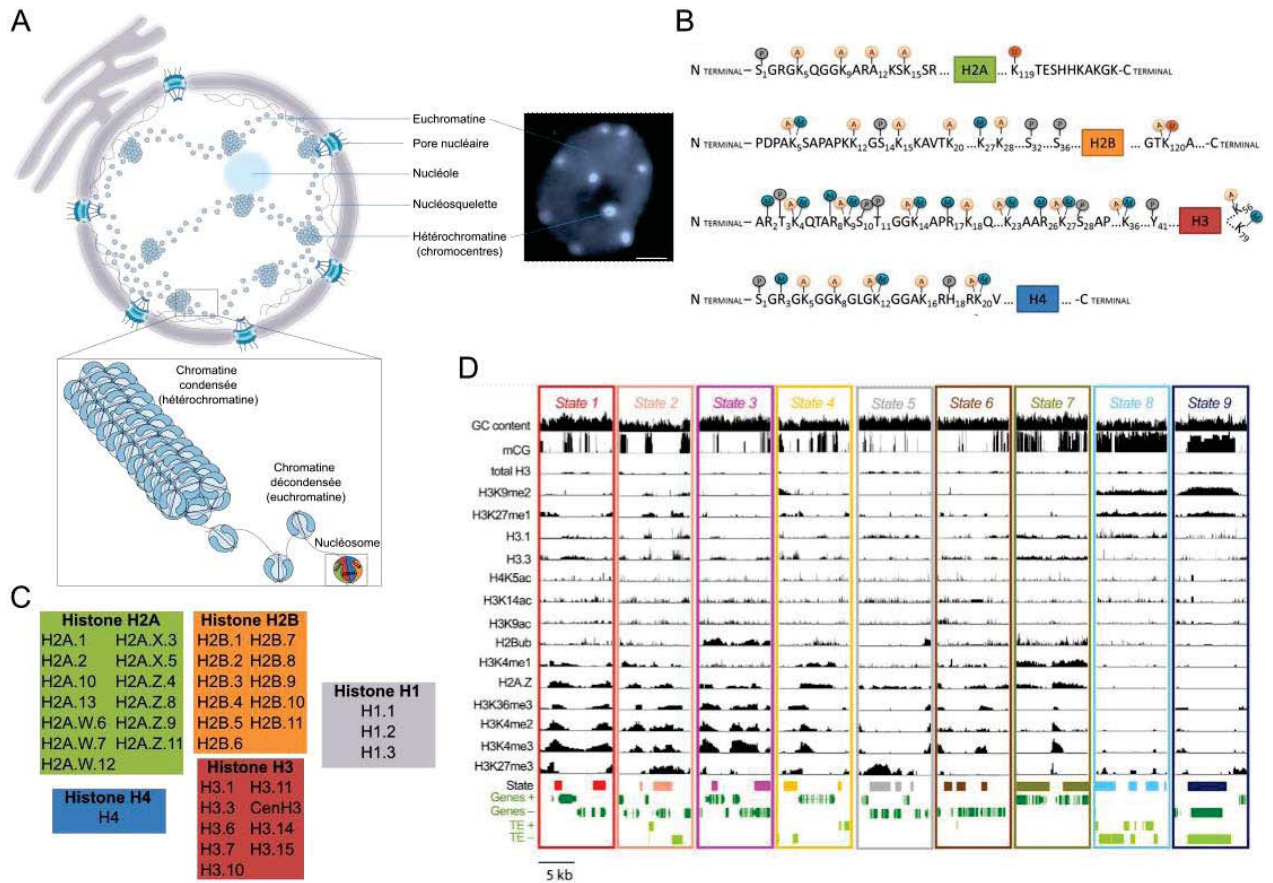


Figure 1.5 : Les marques épigénétiques et les états chromatinien

A. Représentation schématique du noyau avec la chromatine condensée ou hétérochromatine qui s'organise en chromocentres chez *Arabidopsis thaliana*. Les chromocentres sont facilement visibles lors d'un marquage des noyaux avec un intercalant de l'ADN et forme des zones de forte intensité alors que la chromatine décondensée ou euchromatine apparaît sous forme de régions d'intensité homogène. Le nucléole correspond à la zone sans marquage au centre du noyau. **B.** Représentation des différentes modifications post-traductionnelles des histones H2A, H2B, H3 et H4 (Xu *et al.*, 2013). **C.** Tableaux récapitulants les différents variants d'histones présents chez *A. thaliana*. **D.** Représentation des neuf états chromatinien « définis » en fonction des marques épigénétiques qui les composent (Sequeira-Mendes *et al.*, 2014).

Toutefois, elle est également retrouvée dans les gènes exprimés de façon constitutive cette fois plutôt dans le corps du gène et en contexte CG, et agira dans ce cas comme un activateur transcriptionnel (Zhang *et al.*, 2006).

Au niveau des histones, les marques épigénétiques correspondent à des modifications post-traductionnelles. La cible de ces changements se situe essentiellement dans la queue N-terminale des histones ; elles peuvent être méthylées au niveau des lysines ou arginines, acétylées sur les lysines, phosphorylées sur des sérines et thréonines ou encore ubiquitinées (**Figure 1.5B**). Sur une même histone, différentes marques peuvent être trouvées. Chez *A. thaliana*, l'acétylation des histones comme H3 va diminuer la charge positive de celles-ci réduisant ainsi leur interaction avec l'ADN chargé négativement induisant une chromatine plus ouverte, favorable à la transcription. Il existe d'autres modifications associées à une transcription active comme H3K4me², H3K4me³, H3K36me³ et H2Bub. *A contrario*, la méthylation des histones comme H3K27me³ ou H3K9me² est associée à une chromatine plus compactée et répressive pour la transcription.

Outre les modifications post-traductionnelles des histones, l'incorporation de variants d'histones peut également modifier la régulation de la transcription. Chez *A. thaliana*, de nombreux variants d'histones ont été identifiés : l'histone H3 est codée par 15 gènes : 5 codant pour H3.1 (histones canoniques), 3 pour H3.3 et 7 autres variants dont cenH3, l'histone H4 ne possède pas de variant, l'histone H2A avec 4 H2A (1, 2, 10 et 13), 2 H2A.X, 4 H2A.Z et 3 H2A.W, l'histone H2B avec H2B1 à 11 et enfin l'histone H1 avec H1.1 à 1.3 (Probst, Desvoyes and Gutierrez, 2020) (**Figure 1.5C**). Les différents variants d'histones sont incorporés dans la chromatine, comme cenH3 qui est exclusivement localisé dans les régions centromériques répétées ou H2A.W présente au sein des régions péri-centromériques. L'incorporation des variants peut également avoir lieu à des moments différents du cycle cellulaire ou en fonction des tissus et du stade de développement (Probst, Desvoyes and Gutierrez, 2020).

La compilation de ces différentes marques (modifications post-traductionnelles, variants d'histone et méthylation de l'ADN) ainsi que leur localisation a permis de définir 9 états chromatiniens (**Figure 1.5D**) (Sequeira-Mendes *et al.*, 2014).

En dehors des modifications locales, la chromatine est également sujette à des changements de localisation au niveau spatial (position dans le noyau) pouvant modifier l'accessibilité de la chromatine et donc aussi influencer la transcription.

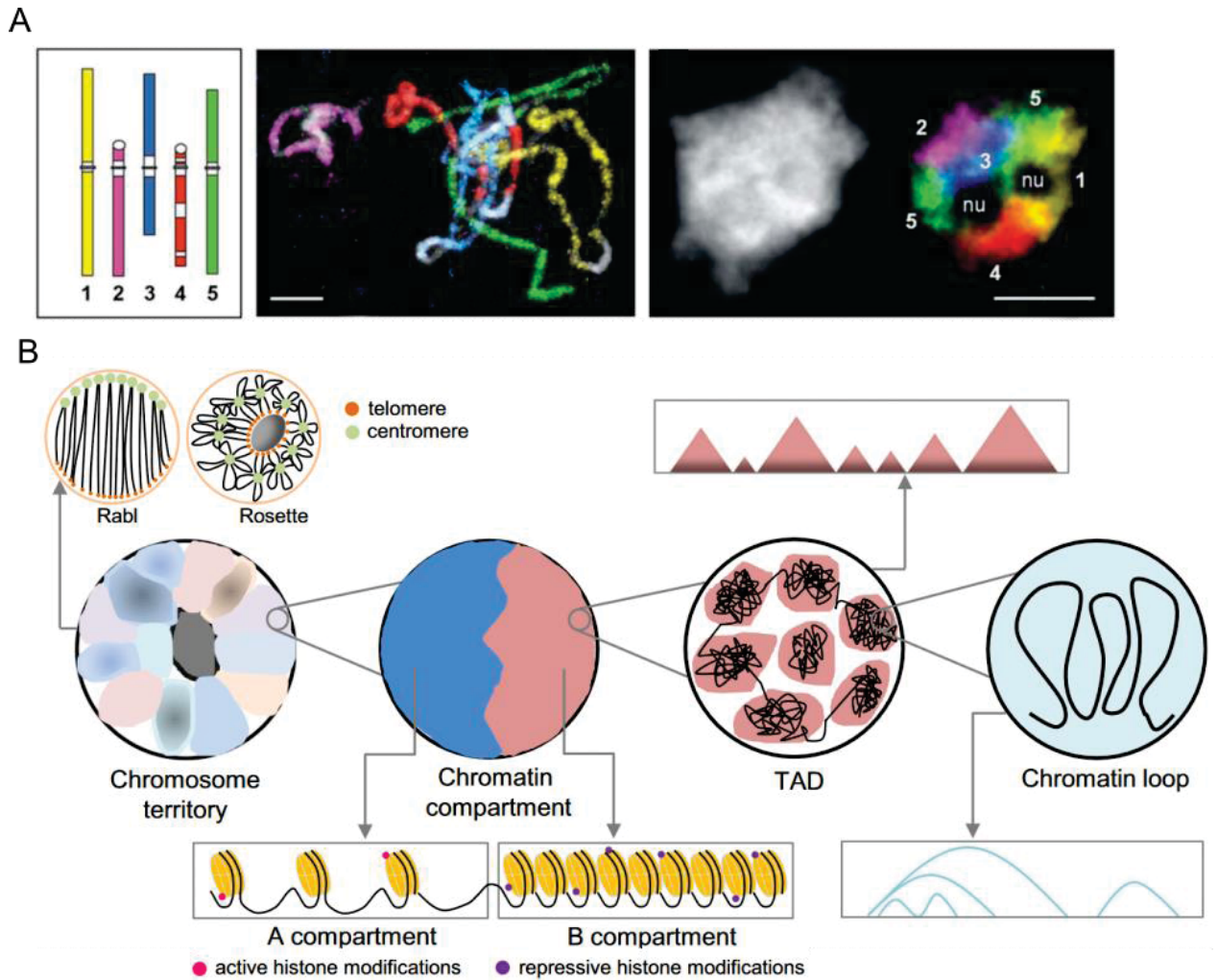


Figure 1.6 : Organisation 3D de la chromatine

A. Le génome d'*Arabidopsis thaliana* est réparti en 5 chromosomes (à gauche), chacun coloré par une couleur différente lors de la réalisation de la technique de chromosome painting au stade pachytène (au centre). A droite, hybridation *in situ* des 5 chromosomes d'*A. thaliana* grâce à des sondes marquées et spécifiques de chaque chromosome permettant ainsi la visualisation des territoires chromosomiques (*Chromosome Territories* ; CT); nu : nucléole (Pecinka *et al.*, 2004) **B.** Représentation schématique de l'organisation 3D de la chromatine avec l'organisation des chromosomes selon la conformation Rabl ou selon le modèle en rosette et formant les territoires chromosomiques qui peuvent être divisés en deux compartiments chromatiniens différents par leurs marques chromatiniennes et leur activité transcriptionnelle (compartiment A : actif et compartiment B : répressif), eux-mêmes constitués de TAD (*Topologically Associating Domains*). Chaque TAD contient un ensemble de boucles de chromatine (Pei *et al.*, 2021).

I.2.c) L'organisation tri-dimensionnelle de la chromatine au sein du noyau

L'une des découvertes importantes de l'étude de l'organisation tri-dimensionnelle de la chromatine a été la mise en évidence de l'organisation non aléatoire des chromosomes dans le noyau des cellules eucaryotes. L'organisation 3D de la chromatine peut être observée à différentes échelles du génome en commençant par les territoires chromosomiques (à l'échelle du chromosome) puis les compartiments (à l'échelle de quelques mégabases), puis les domaines (à l'échelle de la mégabase) et enfin les boucles (à l'échelle du kilobase à la mégabase) (**Figures 1.6A-B**). Ces différents niveaux d'organisation de la chromatine vont être développés ci-dessous.

I.2.c.i) L'organisation des chromosomes en territoires chromosomiques

L'organisation de la chromatine a, dans un premier temps, été étudiée par cytologie couplée à la microscopie en utilisant l'hybridation *in situ*. Cette première analyse a permis de montrer que chaque chromosome occupe une région nucléaire qui lui est propre sans trop se mélanger avec d'autres chromosomes. Ces régions occupées par chaque chromosome sont appelées territoires chromosomiques (CT) (Lichter *et al.*, 1988; Pinkel *et al.*, 1988; Fransz *et al.*, 2002; Fransz and de Jong, 2011).

Chez *A. thaliana*, les 5 chromosomes forment 5 CT distincts qui n'ont pas de localisation préférentielle au sein du noyau à part pour les chromosomes 2 et 4 portant les NOR qui se localisent préférentiellement proches du nucléole (**Figure 1.6A**) (Pecinka *et al.*, 2004).

Outre l'organisation en CT, les chromosomes présentent également une organisation soit de type Rabl (Rabl C., 1885), soit de type Rosette (Fransz *et al.*, 2002) dans les noyaux, qui dépend de la manière dont les bras des chromosomes, les centromères et les télomères sont positionnés et entrent en contact les uns avec les autres (**Figure 1.6B**). Dans la configuration Rabl, les chromosomes sont disposés de manière polarisée, c'est-à-dire que les centromères sont regroupés à un pôle du noyau et les télomères à l'autre pôle. Dans la configuration de type Rosette comme il était postulé d'exister chez *A. thaliana*, l'hétérochromatine est fortement concentrée et forme des chromocentres, structures denses visibles par marquage au DAPI de l'ADN. Les loops euchromatiques des régions proximales du chromocentre émanent de ces structures et sont ancrés dans les chromocentres tandis que les télomères se localisent au niveau du nucléole (Fransz *et al.*, 2002; S. Feng *et al.*, 2014). La présence de ces différentes organisations chromosomiques n'est pas encore bien comprise même si une tendance est que

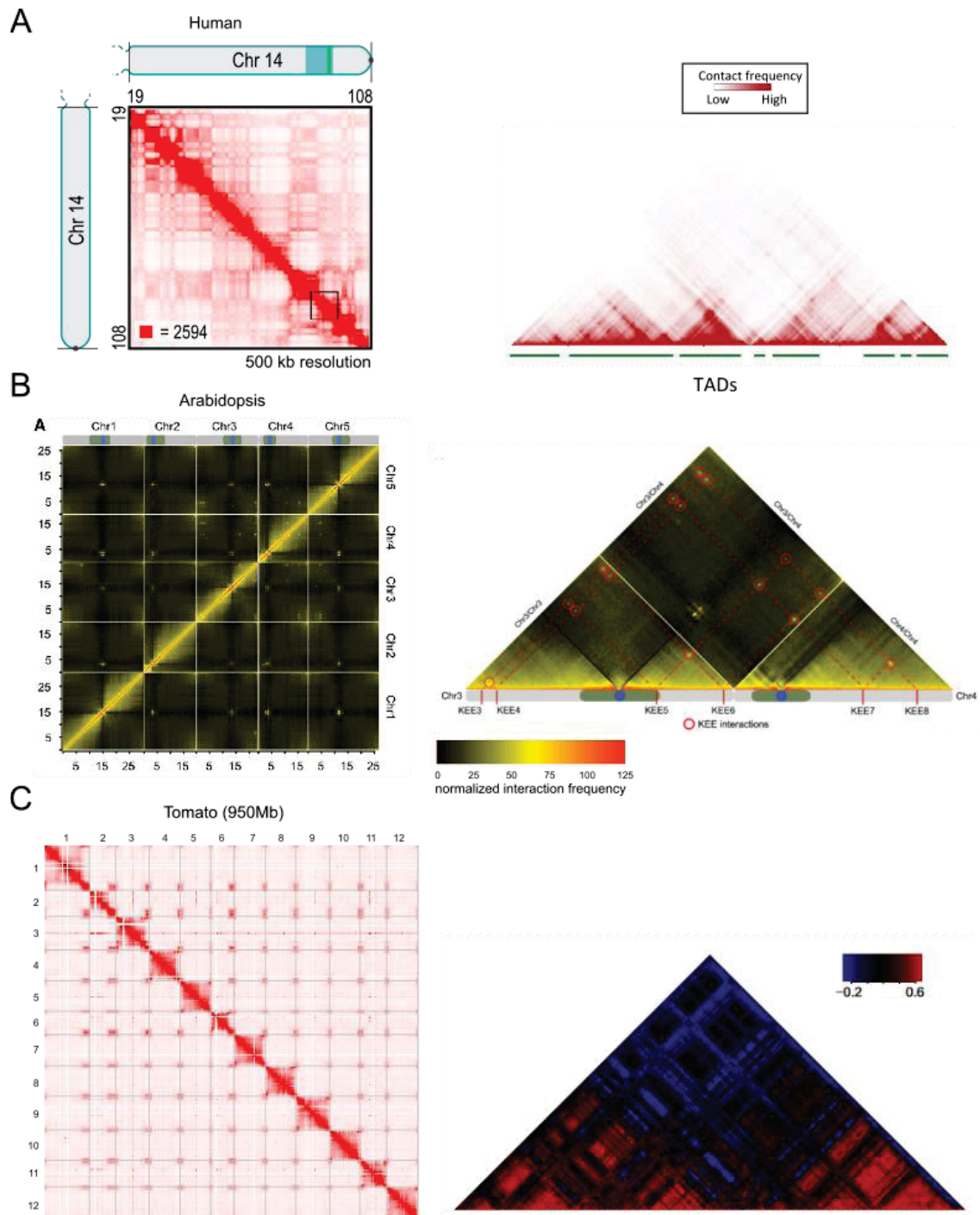


Figure 1.7 : Organisation de la chromatine chez l'Homme, Arabidopsis et la tomate

A-C. Cartes de contacts chromatiniens générés par la technique de Hi-C à partir d'un chromosome (**A**) ou de l'ensemble des chromosomes (**B-C**) avec à droite une représentation permettant de visualiser les TAD. **A.** Cartes Hi-C du chromosome 14 de l'Homme à gauche (Rao *et al.*, 2014) et de 3 Mb du chromosome 15 de l'Homme à droite (Sotelo-Silveira *et al.*, 2018). **B.** Cartes Hi-C du génome à gauche et entre les chromosomes 3 et 4 d'*A. thaliana* à droite (Grob, Schmid and Grossniklaus, 2014). Les cercles rouges représentent les interactions fortes entre chromosomes appelées *Knot Engaged Element* (KEE). **C.** Cartes Hi-C du génome à gauche et du bloc 5 (14 Mb : 84M200k-98M440k) du chromosome 1 de la tomate à droite (Dong *et al.*, 2017).

les plantes à grand génome comme l'orge, le blé et le seigle présentent la configuration Rabl (Santos and Shaw, 2004); alors que les plantes à petit génome comme *A. thaliana*, possèdent souvent une configuration en Rosette (Fransz *et al.*, 2002). Toutefois les données récentes montrent que, pour une même espèce comme l'orge, certains types cellulaires peuvent avoir une conformation Rabl alors que d'autres non (Ales Pecinka, données non publiées) et un contre-exemple, concernant la relation taille du génome et conformation Rabl, est donné par la levure dont le très petit génome de seulement 12Mb est organisé selon la conformation Rabl (Jin *et al.*, 1998).

Ces premières observations indiquent que la chromatine ne s'organise pas de façon aléatoire dans le noyau mais sous forme de CT avec une organisation qui peut différer d'un organisme à l'autre, voire d'un tissu à un autre. Chez *A. thaliana*, l'hétérochromatine est préférentiellement localisée à la périphérie nucléaire ou proche du nucléole (Pontvianne *et al.*, 2016; Poulet, Duc, *et al.*, 2017b), tandis que l'euchromatine occupe plutôt l'espace interne dans le noyau. Ce positionnement pourrait avoir un rôle dans la régulation transcriptionnelle des gènes, comme cela a été constaté chez les animaux (Fransz *et al.*, 2002; Fransz and de Jong, 2011).

I.2.c.ii) La compartimentation de la chromatine

Les territoires peuvent également être visualisés par une approche moléculaire appelée Hi-C, une technique basée sur la méthode du Chromosome Conformation Capture (3C) permettant l'identification des domaines/régions de chromatines interagissant les unes avec les autres. La méthode d'Hi-C permet d'identifier les interactions entre des régions de la chromatine qui sont distantes physiquement, que ce soit à l'échelle d'un gène (par exemple entre une séquence régulatrice et un promoteur), d'un chromosome (interactions intrachromosomiques) ou entre chromosomes (interactions interchromosomiques). Les résultats de Hi-C se traduisent par des cartes de contact représentées sous forme de matrice 2D (chaque axe représentant les deux régions pour lesquelles on cherche des interactions) et où les régions en forte interaction sont identifiées par des carrés rouges ou jaunes et celles de faible interaction par des carrés blancs ou noirs (**Figures 1.7A-C panels de gauche**). Cette méthode montre que les interactions les plus fréquentes sont de type intrachromosomique.

Chez les animaux, des expériences de Hi-C ont montré que la chromatine n'est pas disposée de manière aléatoire à l'intérieur du noyau des mammifères, et que les régions actives et répressives sont spatialement séparées (**Figure 1.7A**).

En effet, la chromatine active s'associe préférentiellement avec les autres régions actives et la chromatine répressive plutôt avec elle-même. Ainsi, selon la technique du Hi-C et les interactions entre régions de la chromatine, le génome se subdiviserait en deux compartiments différents dans l'espace nucléaire (Lieberman-Aiden *et al.*, 2009). Ces compartiments actifs et répressifs sont appelés respectivement, compartiment A et B. Le compartiment A des mammifères est activement transcrit, enrichi en chromatine ouverte présentant des marques épigénétiques actives des histones telles que H3K4me3 et H3K27ac. Le compartiment B est, au contraire, enrichi en marques d'histones répressives telles que H3K9me3. Enfin, les séquences composant le compartiment A sont plutôt riches en GC alors que celles du compartiment B sont riches en AT, une caractéristique rappelant respectivement l'euchromatine riche en gènes et l'hétérochromatine riche en séquences répétées (Lieberman-Aiden *et al.*, 2009). La définition des compartiments basée sur le Hi-C est cohérente avec les études microscopiques antérieures. Avec la technique Hi-C, les mêmes compartiments ont été trouvés dans différentes espèces de plantes comme Arabidopsis, la tomate ou le maïs. Le compartiment A est caractérisé par une forte densité de gènes et des modifications épigénétiques associées à une activité transcriptionnelle active, soit d'euchromatine, tandis que le compartiment B présente une plus forte densité en éléments transposables et en modifications épigénétiques répressives soit d'hétérochromatine comme chez les mammifères (**Figures 1.7A-C**) (Dong et al., 2017 ; Xie et al., 2019).

Indépendamment des études réalisées avec le Hi-C, une seconde technique appelée le *DNA adenine methyltransferase identification* (DamID) et visant à déterminer les régions de chromatine localisées à la périphérie du noyau proche du composant majeur du nucléosquelette (les lamines chez les animaux, cf paragraphe I.3.c.i) ont montré que la chromatine active réside plutôt au centre du noyau, alors que l'hétérochromatine est enrichie vers la périphérie nucléaire et s'associe aux lamines (Bickmore et van Steensel, 2013). Le DamID permet d'identifier d'autres caractéristiques de la chromatine qui selon cette technique s'organise en *Lamin associated Domains* (LAD) enrichis à la périphérie nucléaire et en non-LAD distants de la périphérie du noyau (voir paragraphe I.2.c.iii).

L'arrivée des nouvelles techniques moléculaires comme le DamID et le Hi-C permettent d'ajouter de la précision quant à l'organisation de la chromatine en CT. Le compartiment B est enrichi en hétérochromatine et également enrichi en séquences LAD proche de la périphérie nucléaire alors que le compartiment A correspond plus à l'euchromatine.

Cependant, la compartimentalisation n'est pas statique, des loci peuvent passer du compartiment A vers B ou de LAD à non-LAD et inversement en fonction du tissu, du stade de développement ou des stimuli auxquels seront soumis les cellules.

I.2.c.iii) Les domaines de chromatine : TAD, LAD et NAD

La méthode de Hi-C a permis de faire une seconde découverte dans l'organisation spatiale de la chromatine. En effet, à plus petite échelle, le génome est organisé sous forme de domaines de chromatine plus communément appelés *topologically associated domains* (domaines associés topologiquement, TAD) chez les mammifères. Les TAD sont des domaines 3D le long des chromosomes. Un TAD est défini comme un groupe de séquences consécutives, en interaction distantes, qui forment un domaine chromatinien : les séquences d'ADN à l'intérieur du TAD interagissent plus fréquemment entre elles qu'avec les séquences des TAD adjacents. Cette caractéristique des TAD les rend facilement identifiables comme des sous régions d'interaction préférentielle le long de la diagonale des cartes de contact Hi-C (Stam, Tark-Dame and Fransz, 2019; Dong *et al.*, 2020; Grob, 2020; Pei *et al.*, 2021) (**Figures 1.7A-C panels de gauche**). La subdivision du génome en TAD définit ainsi des « blocs isolants » au sein desquels les éléments régulateurs sont confinés et permettent de coréguler un ensemble de gènes inclus dans le même TAD. Les bordures de la plupart des TAD sont enrichies en cohésine, en facteur de liaison CCCTC (CTCF) ayant une fonction d'insulateur et en modifications épigénétiques associées avec la transcription active comme H3K4me3 et H3K9ac (Stam, Tark-Dame and Fransz, 2019; Dong *et al.*, 2020; Grob, 2020; Pei *et al.*, 2021).

La présence de TAD chez les plantes n'a que très récemment été étudiée et varie en fonction de l'espèce. En effet, *A. thaliana* possède peu de TAD par rapport aux mammifères et ils sont également plus petits et avec de plus faibles interactions, ils sont donc moins visibles sur les cartes d'Hi-C (**Figures 1.7B**). Alors que d'autres espèces à plus grand génome comme le riz, le coton, la tomate et le maïs possèdent des TAD clairement définis (**Figure 1.7C**). Ceci pourrait s'expliquer par une distribution plutôt uniforme de l'activité transcriptionnelle le long des chromosomes et par la faible densité en éléments transposables par rapport à d'autres espèces comme le maïs qui possède une faible densité en gènes et donc beaucoup plus d'éléments transposables intercalés entre des îlots de gènes (Rowley and Corces, 2016; Stam, Tark-Dame and Fransz, 2019). Une différence majeure avec les animaux est l'absence de la protéine CTCF bien que les bordures des TAD soient, comme chez les animaux, délimitées par la présence de marques épigénétiques actives.

Chez les plantes, ce serait la cohésine et d'autres protéines, comme les facteurs de transcription plante spécifique TEOSINTE BRANCHED 1, CYCLOIDEA, AND PCF (TCP) et les protéines BASIC LEUCINE-ZIPPER (bZIP) qui définiraient la bordure des TAD (Stam, Tark-Dame and Fransz, 2019).

D'autres domaines de chromatine comme les *lamina-associated domains* (LAD) ou les *nucleolus-associated domains* (NAD) correspondent à une spécialisation des TAD. En effet, les LAD sont des TAD retrouvés à la périphérie du noyau, il s'agit de domaines de chromatine interagissant avec les lamines (Pickersgill *et al.*, 2006; Guelen *et al.*, 2008a). Chez *A. thaliana*, les LAD ont été identifiés par ChIP-Seq en utilisant un anticorps dirigé contre certaines protéines de la périphérie nucléaire comme NUP136 (un composant du pore nucléaire) ou CRWN1 (un homologue des lamines animales) (cf paragraphe [1.3.b](#)) et [1.3.c](#)) (Bi *et al.*, 2017; Hu *et al.*, 2019). Les NAD sont quant à eux des TAD composés de grandes régions génomiques présentant des caractéristiques hétérochromatiques. Chez la souris, deux types de NAD ont été décrits, les « NAD de type I » sont des régions d'hétérochromatine enrichies en marque H3K9me3 s'associant à la fois au nucléole et à la lamina. En revanche, les « NAD de type II » ne s'associent qu'au nucléole, sont plus permissifs à l'expression génique et plus enrichis en H3K27me3 que les NAD de type I (Vertii *et al.*, 2019). Chez *A. thaliana*, les NAD sont des régions génomiques composées d'éléments transposables, de régions télomériques et sub-télomériques ainsi que de gènes actifs d'ARNr et de tout le bras court du chromosome 4 qui leur est adjacent et enrichies en marques chromatiniennes répressives (Pontvianne *et al.*, 2016). Globalement, les TAD des plantes partagent la plupart des caractéristiques des TAD des animaux, même si leur rôle fonctionnel reste encore largement à explorer.

Outre la formation de domaines le long des chromosomes *via* des interactions intrachromosomiques, des interactions physiques interchromosomiques ont également été trouvées chez *A. thaliana* ainsi que chez le riz entre des petites région hétérochromatiques situées dans l'euchromatine. Ces régions ont été appelées *Interactive Heterochromatin Islands* (IHI) ou *KNOT Engaged Elements* (KEE) et consistent en des régions riches en éléments transposables enrichis en H3K9me² et associées aux petits RNA (sRNA) (S. Feng *et al.*, 2014; Grob, Schmid and Grossniklaus, 2014). Ces loci sont au nombre de 10 répartis sur les 5 chromosomes et sont trouvées à la fois au sein des bras chromosomiques et dans les régions péri-centromériques (**Figure 1.7B panel de droite**). Dans l'ensemble des 10 KEE, il a un enrichissement en éléments transposables *VANDAL6* et *ATLANTYS3*, la présence de séquences

identiques favoriserait les interactions de ces régions entre-elles (Grob, Schmid and Grossniklaus, 2014).

De plus, les KEE semblent former des sites d'insertions préférentiels pour les TE, d'où la spéculation qu'ils sont impliqués dans un mécanisme de défense nucléaire semblable au système immunitaire pour protéger le génome contre les éléments d'ADN mobiles similaires au locus *flamenco*, un piRNA cluster chez la drosophile. En effet, les KEE sont impliqués dans l'extinction d'éléments d'ADN invasifs, tels que les transgènes, par un phénomène appelé *KNOT Linked Silencing* (KLS) qui n'a pas pu être associé à des mécanismes de répressions canoniques bien étudiés impliquant les sRNA et la méthylation de l'ADN, ce qui pourrait indiquer que le KLS représente un mécanisme de défense nucléaire indépendant et inexploré à ce jour (Grob and Grossniklaus, 2019; Grob, 2020).

I.2.c.iv) Les boucles de chromatine et le modèle d'extrusion des boucles

A une échelle encore plus petite, au niveau du kilobase, la chromatine forme des boucles (*chromatin loops*), résultant d'interactions plus locales de la chromatine. Ces boucles permettent notamment aux éléments *cis*-régulateurs comme les enhanceurs, qui peuvent être éloignés de leurs promoteurs cibles, d'interagir efficacement avec ces derniers afin d'exercer leur rôle régulateur. Les interactions entre les enhanceurs et les promoteurs jouent un rôle crucial dans la régulation de l'expression des gènes et sont souvent tissu-spécifiques. La première boucle de chromatine découverte chez les plantes a été décrite pour le promoteur et l'enhanceur du locus *b1* du maïs et entre les régions 5' et 3' du locus *FLC* chez *A. thaliana* (Louwers *et al.*, 2009; Crevillén *et al.*, 2013). Dans ces deux cas, les interactions permettent l'activation transcriptionnelle des gènes *b1* et *FLC*. Chez les plantes, des boucles de chromatine à courte portée seraient associées à des gènes constitutivement actifs (Li *et al.*, 2019).

Une découverte importante de ces dernières années est le modèle « d'extrusion des boucles de chromatine » expliquant comment un complexe protéique permet progressivement d'agrandir les boucles de chromatine, d'un contexte très local, à une organisation plus large en domaines jusqu'aux CT, voire à la forme la plus condensée de la chromatine dans le chromosome métaphasique (Banigan and Mirny, 2020). L'extrusion des boucles est assurée par des complexes de maintien structurel des chromosomes (Structural Maintenance Complex ; SMC), incluant notamment des protéines de type cohésine (pendant la division cellulaire) et condensine (pendant l'interphase). Ces protéines se lient de façon aléatoire à la chromatine et extrudent les boucles de chromatine jusqu'à ce qu'elles se décrochent. Pendant l'interphase, les

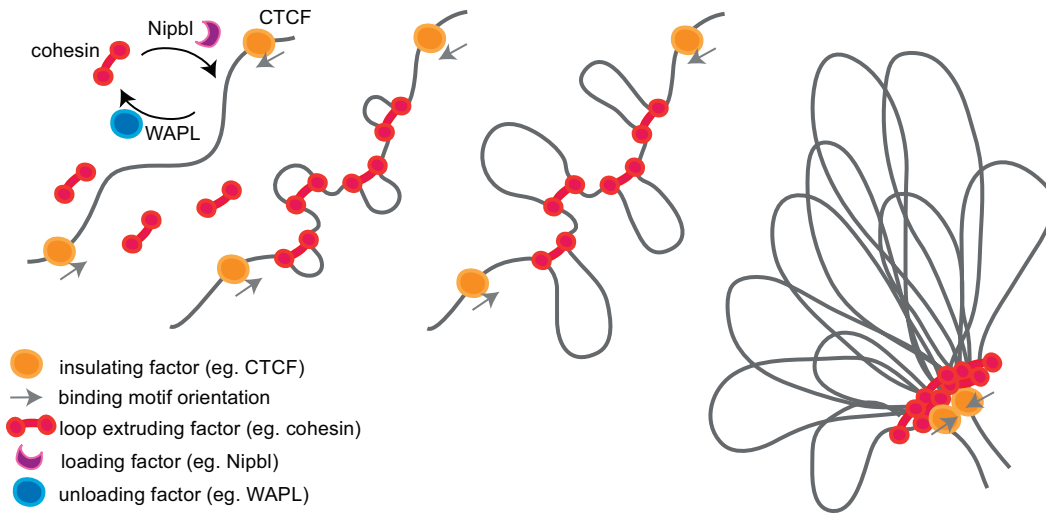


Figure 1.8 : L'extrusion de boucles comme modèle pour la formation de TAD.

Les complexes d'extrusion de boucles impliquant des protéines de type condensine et cohésine (rouge) sont continuellement chargés et déchargés de la chromatine interphasique et extrudent (action d'un moteur protéique qui lors de sa migration sur la chromatine fait progresser la taille de la boucle) des boucles de manière bidirectionnelle lorsqu'ils sont engagés sur la chromatine. Pendant l'interphase, ces complexes peuvent former des domaines d'association topologique (TAD) en extrudant des boucles jusqu'à ce qu'elles rencontrent une protéine CTCF (orange) orientée dans la bonne direction (c'est-à-dire flèche pointant vers la cohésine) (Sikorska and Sexton, 2019).

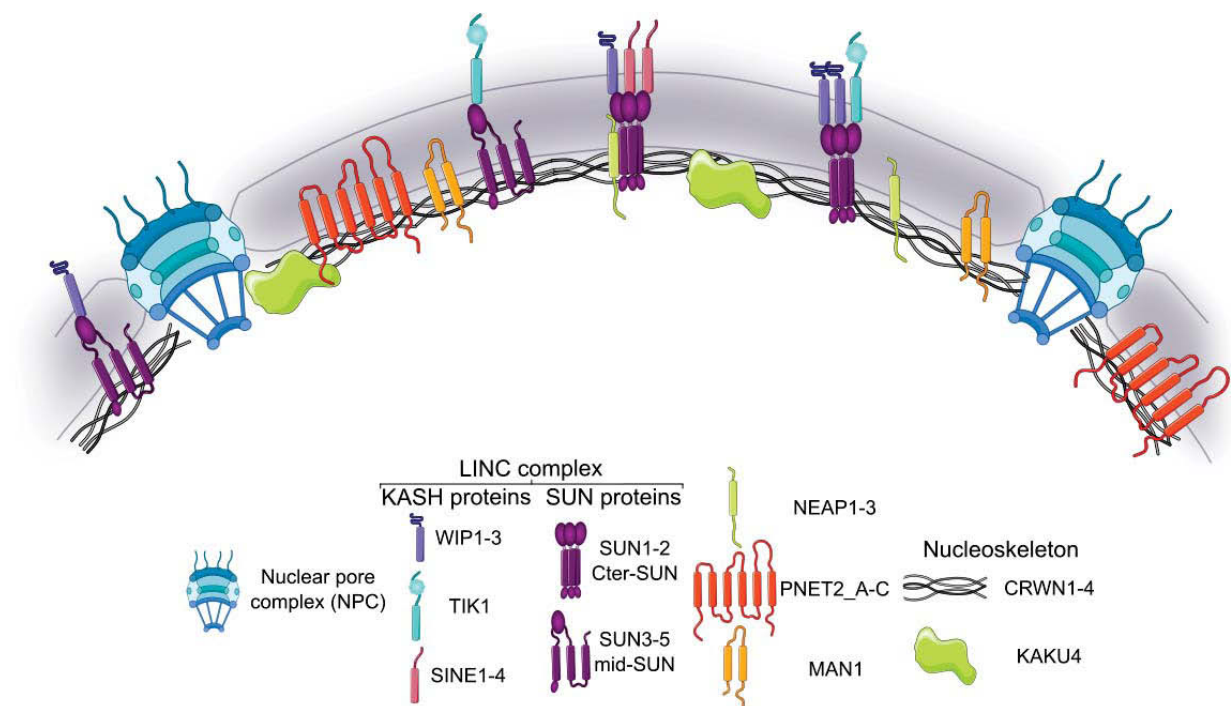


Figure 1.9 : Représentation schématique de la périphérie nucléaire d'*A.thaliana*

cohésines fabriquent des boucles et forment les TAD en extrudant des boucles jusqu'à ce qu'elles rencontrent une protéine CTCF orientée dans la bonne direction (**Figure 1.8**). Sans doute l'une des expériences clefs a été l'utilisation de mutations affectant le complexe des cohésines notamment les *cohesin loading factors* *Scs4*, *Nipb* ou *Rad21* qui abolissent l'organisation en TAD chez la souris (Beagan and Phillips-Cremins, 2020). Très peu de choses sont connues chez les plantes, cependant il semblerait que ce soit la cohésine, les facteurs TCP et bZIP qui définiraient la bordure des TAD, sans toutefois jouer le rôle d'insulateur des protéines CTCF (Liu *et al.*, 2017; Stam, Tark-Dame and Frasz, 2019).

Pour conclure, l'organisation tri-dimensionnelle de la chromatine est assurée par la formation de boucles de chromatine de plus ou moins grande taille qui s'explique assez bien aujourd'hui avec la caractérisation du rôle des cohésines et des condensines dans le modèle d'extrusion des boucles. Il s'agira ensuite d'inclure dans cette vision « en boucles » un aspect d'organisation spatiale de la chromatine dans l'espace nucléaire et locale avec les marques épigénétiques. Ces différents niveaux d'organisation de la chromatine sont sans doute nécessaires à la régulation fine de l'expression des gènes et constituent un vaste champ de recherche pour les années à venir. La suite de ce manuscrit se concentre sur notre compréhension des acteurs qui pourraient participer au rapprochement des régions de la chromatine comme les LAD au niveau de la périphérie nucléaire. J'essayerai de montrer dans la suite de cette introduction comment la périphérie nucléaire pourrait jouer un rôle important dans la régulation de la transcription (cf paragraphe [II-](#)).

I.3) La périphérie nucléaire

Nous appellerons périphérie nucléaire, la région du noyau regroupant l'ensemble des protéines ancrées dans la membrane nucléaire, les pores nucléaires ainsi que le nucléosquelette (**Figure 1.9**). La périphérie nucléaire est plutôt bien décrite chez la levure et les animaux avec 300 à 400 protéines identifiées (Gomez-Cavazos and Hetzer, 2012; Smoyer *et al.*, 2016) alors que chez la plante, les composants de la périphérie sont moins bien décrits avec environ 200 protéines répertoriées (Tang, Huang and Gu, 2020). Je vais donc décrire les différents composants de la périphérie nucléaire en commençant par l'enveloppe nucléaire puis les pores nucléaires et enfin le nucléosquelette.

[I.3.a\) L'enveloppe nucléaire interne et ses constituants principaux](#)

L'enveloppe nucléaire est une double membrane composée de phospholipides et de protéines jouant le rôle de barrière protectrice pour le génome. Elle est constituée d'une

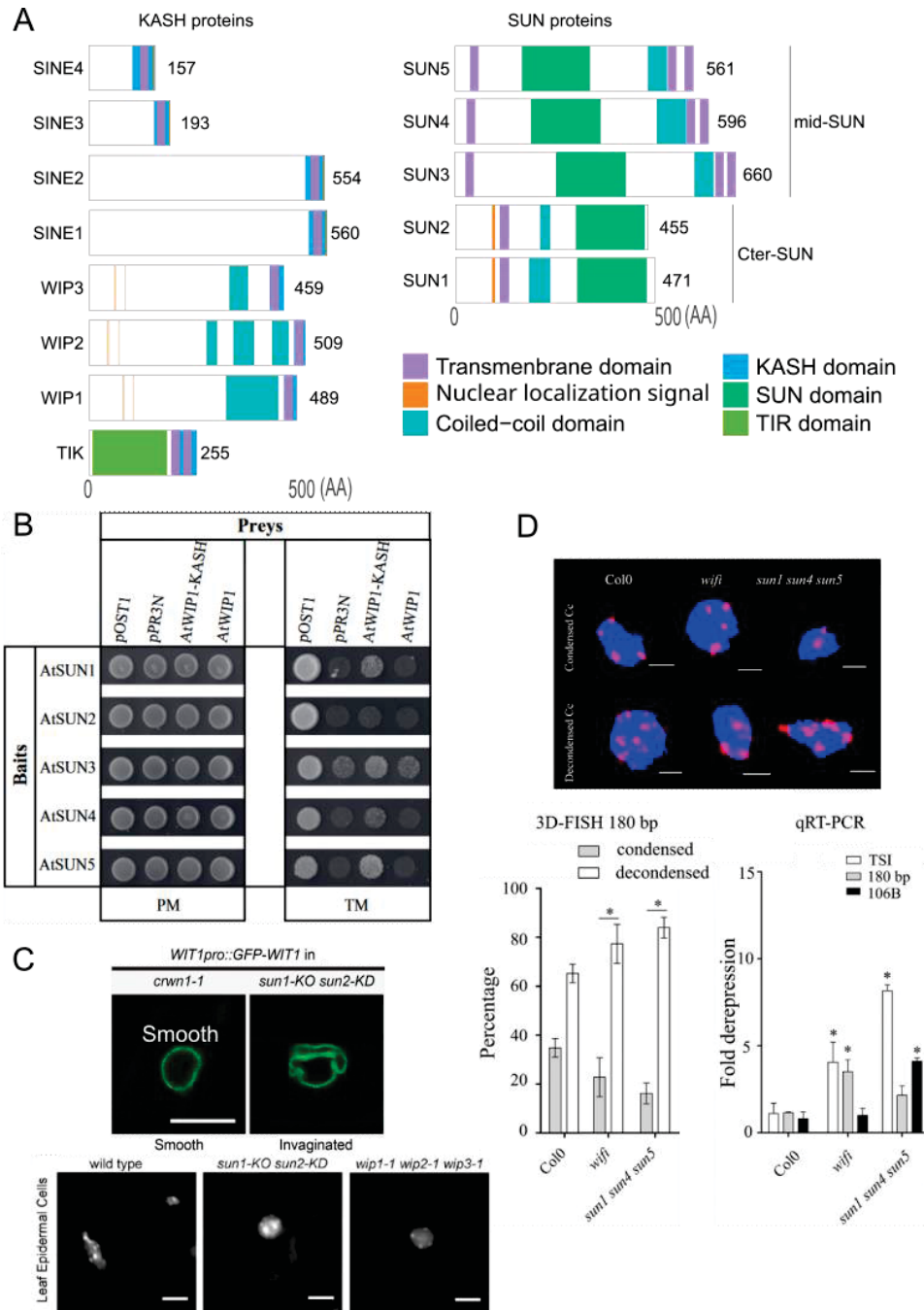


Figure 1.10 : Caractérisation du complexe LINC

A. Représentation schématique des protéines SUN et KASH d'*Arabidopsis thaliana*. **B.** Interaction des protéines SUN1-5 avec la protéine KASH appelée WIP1 entière (WIP1) ou avec seulement son domaine KASH de WIP1 (WIP1-KASH) via la technique du MYTH (*Membrane Yeast Two-Hybrid*) ; PM : permissive medium, TM : test medium (Graumann *et al.*, 2014a). **C.** Morphologie de l'enveloppe nucléaire dans des poils racinaires des mutants *crwn1-1* et *sun1-KO sun2-KD* via l'expression de WIT1-GFP (Zhou, Groves and Meier, 2015) et morphologie nucléaire des mutants pour les protéines SUN, *sun1-KO sun2-KD*, et KASH, *wip1-1 wip2-1 wip3-1* de noyaux de l'épiderme de feuilles (Zhou *et al.*, 2012a) **D.** Condensation des régions péri-centromériques (séquences répétées TSI, 180 bp et 106B) chez des plantes dépourvues des protéines WIP et SUN corrélant avec la réactivation des séquences répétées (Poulet, Duc, *et al.*, 2017a).

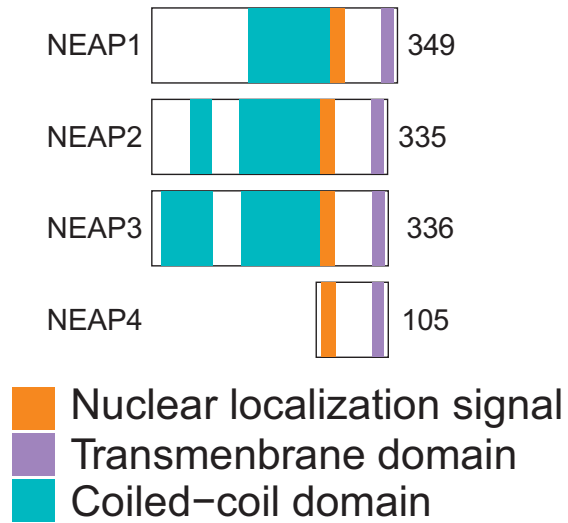
membrane interne (INM) et d'une membrane externe (ONM) en continuité avec le réticulum endoplasmique. Ces deux membranes délimitent un espace périnucléaire, également appelé espace périplasmique (**Figure 1.2**). Cet espace est maintenu par le complexe Linker of Nucleoskeleton and Cytoskeleton (LINC) connectant le cytosquelette au nucléosquelette, *via* un pont protéique entre l'INM et l'ONM. Les deux membranes se connectent au niveau des pores nucléaires ponctuant l'enveloppe nucléaire. Les protéines présentes dans l'enveloppe nucléaire possèdent au moins un domaine transmembranaire et peuvent être regroupées dans la famille des protéines transmembranaires de l'enveloppe nucléaire (PLANT NUCLEAR ENVELOPE TRANSMEMBRANE, PNET) rassemblant environ 200 protéines chez *A. thaliana* incluant les composants du NPC, du complexe LINC et les protéines ancrées dans la membrane nucléaire comme les protéines NUCLEAR ENVELOPE-ASSOCIATED PROTEINS (NEAP) ainsi que la protéine MAN1. Parmi ces protéines, 16 ont été nommées PNET1-14 et deux protéines homologues de PNET2 ont été identifiées par la suite, elles ont donc été renommées PNET2-A à C (**Figure 1.9**) (Tang, Huang and Gu, 2020; Tang *et al.*, 2021).

I.3.a.i) Le complexe LINC

Le complexe LINC est composé de deux familles de protéines : les protéines à domaine SAD1/UNC84 (SUN) et les protéines à domaine KLARSICHT/ANC-1/SYNE-1 HOMOLOGY (KASH) respectivement ancrées dans l'INM et l'ONM (**Figure 1.8**) (Van Damme *et al.*, 2004; Tatout *et al.*, 2014). Elles interagissent entre-elles *via* les domaines SUN et KASH dans l'espace périnucléaire. Chez les eucaryotes, le complexe LINC est l'un des composants de l'enveloppe nucléaire le plus conservé avec les pores nucléaires (Poulet, Probst, *et al.*, 2017). La cristallographie des protéines SUN et KASH humaines montre que ces protéines interagissent sous forme de trimère (Sosa *et al.*, 2013) et, du fait de la bonne conservation de ces protéines au sein des eucaryotes, ceci pourrait également être le cas chez les plantes.

Les protéines SUN sont caractérisées par la présence d'un domaine SUN. En fonction de la position de ce domaine dans la protéine, on distingue deux sous-familles : les Cter-SUN où le domaine SUN est en position C-terminale et les mid-SUN dans lesquelles le domaine SUN est central (Graumann, Runions and Evans, 2010; Graumann *et al.*, 2014b; Zhou *et al.*, 2012b; Zhou and Meier, 2014) (**Figure 1.10A**). Chez *A. thaliana*, les Cter-SUN (SUN1 et SUN2) sont localisées à la périphérie nucléaire alors que les protéines mid-SUN (SUN 3 à 5) se localisent à la fois dans l'INM et/ou l'ONM et dans la membrane du réticulum endoplasmique (Graumann, Runions and Evans, 2010; Graumann *et al.*, 2014b).

A



B

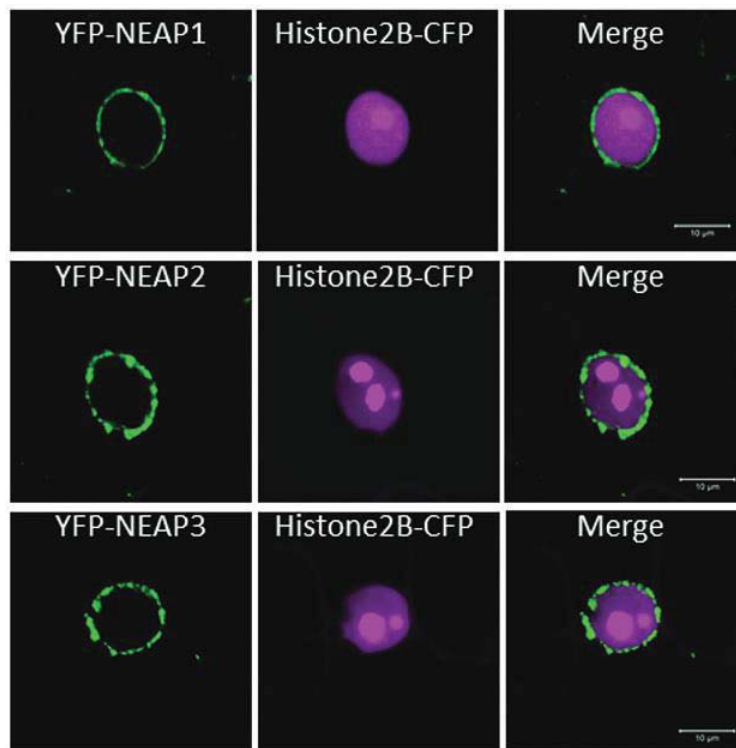


Figure 1.11 : Détermination des protéines NEAP

A. Représentation schématique des protéines NEAP d'*A. thaliana*. **B.** Co-expression transitoire chez le tabac des protéines NEAP1-3 avec l'histone H2B, montrant la localisation périphérique des NEAP ; échelle : 10 μm (Pawar *et al.*, 2016).

Outre le domaine SUN, ces protéines possèdent également un domaine coiled-coil (sauf SUN3 qui en possède deux) et un ou trois domaines transmembranaires, respectivement pour les Cter-SUN et les mid-SUN (**Figure 1.10A**). Les protéines SUN interagiraient entre-elles sous forme de trimère *via* leur domaine coiled-coil d'après des études réalisées chez l'Homme (Li, McKee-Muir and Gilbert, 2018) et peuvent ainsi former des combinaisons de complexes différents (Graumann *et al.*, 2014b). Le double mutant *sun1-KO sun2-KD* (KO : *Knock-Out* et KD : *Knock-Down*) présente des invaginations au niveau de son enveloppe nucléaire ainsi que de petits noyaux sphériques tout comme le triple mutant *sun1 sun4 sun5* (Zhou *et al.*, 2012b; Zhou, Groves and Meier, 2015; Poulet, Duc, *et al.*, 2017b). Ces résultats suggèrent que les protéines SUN jouent un rôle dans le maintien de la morphologie nucléaire. Elles participent également à l'organisation de la chromatine étant donné que certains mutants *sun* entraînent une décompaction partielle des chromocentres ainsi qu'une levée du silencing de certaines séquences répétées (**Figure 1.10D**) (Poulet, Duc, *et al.*, 2017b).

Le deuxième composant du complexe LINC correspond aux protéines KASH. Chez *A. thaliana*, il existe plusieurs sous-familles de protéines KASH : les WPP (Tryptophane-Proline-Proline) INTERACTING PROTEIN (WIP) 1 à 3 (Zhou *et al.*, 2012b), les SUN INTERACTING NUCLEAR ENVELOPE PROTEINS (SINE) 1 à 4 (Zhou *et al.*, 2014) et les TOLL INTERLEUKIN RECEPTOR DOMAIN KASH PROTEIN (TIK) présentes uniquement dans quelques espèces de Brassicacées (**Figure 1.10A**) (Graumann *et al.*, 2014b). Une dernière sous-famille de protéine KASH a été découverte chez le maïs et est présente uniquement chez les Poacées, il s'agit des MAIZE LINC KASH GRASS-SPECIFIC (MLKG) 1 et 2 (Gumber *et al.*, 2019). Ces protéines sont caractérisées par la présence d'un domaine KASH, constitué d'une région transmembranaire suivie du motif XXPT à l'extrémité C-terminale de la protéine. Ce motif localisé dans la région transmembranaire permet l'interaction avec le domaine SUN des protéines SUN pour former le complexe LINC. Chez les plantes, ce domaine a été décrit dans un premier temps chez les protéines WIP, capables d'interagir avec les protéines SUN1 et SUN2 (**Figure 1.10B**) ; et puis a permis, par des méthodes *in silico*, d'identifier d'autres protéines KASH comme les SINE (Graumann *et al.*, 2014b; Zhou *et al.*, 2014). Tout comme les protéines SUN, les protéines KASH jouent également un rôle dans la morphologie nucléaire comme le mutant *wifi* impliquant les protéines WIP et WPP DOMAIN-INTERACTING TAIL-ANCHORED PROTEIN (WIT) (quintuple mutant *wip1-1 wip2-1 wip3-1 wit1-1 wit2-1*) avec des noyaux plus petits et sphériques ainsi qu'une décondensation des chromocentres liée à une

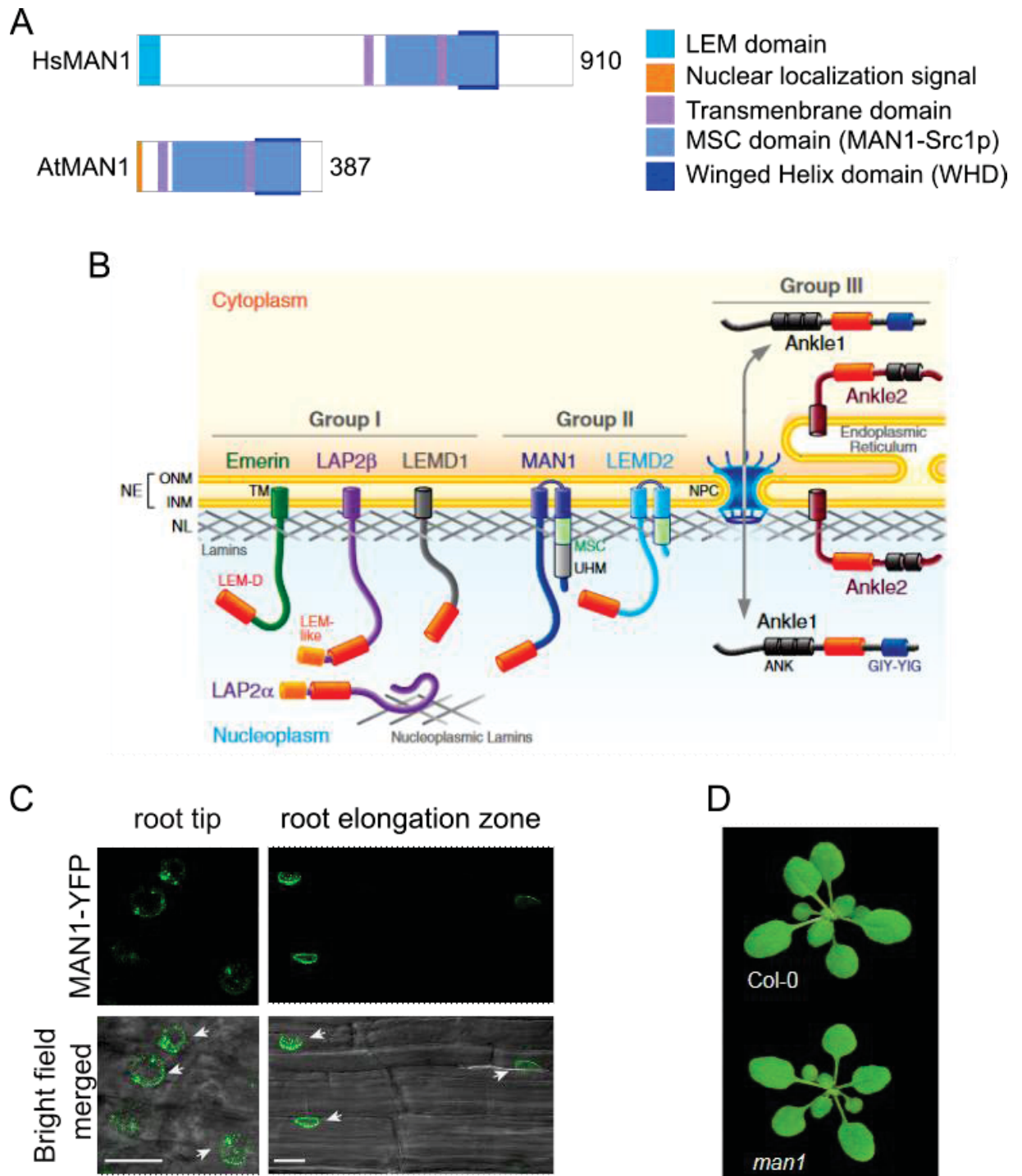


Figure 1.12 : Protéines interagissant avec les lamines chez l'Homme et *Arabidopsis*

A. Représentation schématique des protéines MAN1 de l'Homme et d'*A. thaliana* (At5g46560). **B.** Schéma représentatif de la périphérie nucléaire chez l'Homme avec les différentes protéines interagissant avec les lamines (Barton, Soshnev and Geyer, 2015). **C.** Images de cellules racinaires d'*Arabidopsis* transgénique p35S:MAN1-YFP, flèches : noyaux. (Tang *et al.*, 2021). **D.** Photo du simple mutant d'insertion *man1-1*. (Tang *et al.*, 2021).

dérégulation de certaines régions répétées (**Figures 1.10C-D**) (Zhou *et al.*, 2012b; Poulet, Duc, *et al.*, 2017b).

Chez *A. thaliana*, outre le nombre important de chaque type de protéine du complexe LINC (5 protéines SUN et 8 protéines KASH à ce jour identifiées), le fait que celui-ci puisse constituer un trimère de protéines SUN et de même pour les protéines KASH, suggère une grande diversité de complexes LINC possibles en fonction du type cellulaire, du stade de développement voire même certaines combinaisons pourraient permettre des réponses spécifiques à certaines étapes du développement ou à des stimuli environnementaux.

I.3.a.ii) Les protéines NEAP

Une autre famille de protéines est présente dans l'enveloppe nucléaire interne, les protéines NEAP pour NUCLEAR ENVELOPE-ASSOCIATED PROTEINS, spécifiques des plantes. Chez *A. thaliana*, trois protéines NEAP1 à 3 et un pseudogène NEAP4 ont été identifiées par des méthodes bio-informatiques. Elles possèdent un signal de localisation nucléaire (NLS), des domaines coiled-coil ainsi qu'un domaine transmembranaire dans leur partie C-terminale (**Figure 1.11A**) (Pawar *et al.*, 2016). Les protéines NEAP se localisent à la périphérie nucléaire (**Figure 1.11B**) et sont ancrées dans l'INM où elles interagissent avec les protéines SUN1 et SUN2 (Pawar *et al.*, 2016). La protéine NEAP1 serait également capable d'interagir avec le facteur de transcription basic-LEUCINE ZIPPER 18 (bZIP18) (Pawar *et al.*, 2016).

I.3.a.iii) La protéine MAN1

Pendant longtemps, il a été considéré que les plantes ne possédaient pas de lamines, le composant du nucléosquelette chez les mammifères, et par conséquent ne possédaient pas non plus de protéines liant les lamines comme les protéines LAMIN B RECEPTOR (LBR) ou les protéines LAP2 β / Emerine / MAN1 (collectivement appelées protéines LEM) permettant l'ancrage des lamines à la périphérie nucléaire. Cependant, il existe une protéine MAN1 potentielle, retrouvée par recherche d'homologie chez *A. thaliana*, comme chez les mammifères ou la levure. MAN1 possède deux domaines transmembranaires, un domaine C-terminal MAN1/SRC1p/C-terminal (MSC) conservé, qui est caractérisé comme un médiateur de la liaison à l'ADN chez l'homme ainsi qu'un domaine Winged Helix (WHD) connu chez la levure pour interagir avec l'anneau interne du pore nucléaire (**Figure 1.12A**) (Caputo *et al.*, 2006; Borah *et al.*, 2021). La protéine MAN1 des plantes ne possède pas de domaine LEM permettant l'interaction avec la protéine BARRIER-TO-AUTOINTEGRATION FACTOR (BAF) qui interagit avec la chromatine, caractéristique des protéines LEM des animaux et des levures.

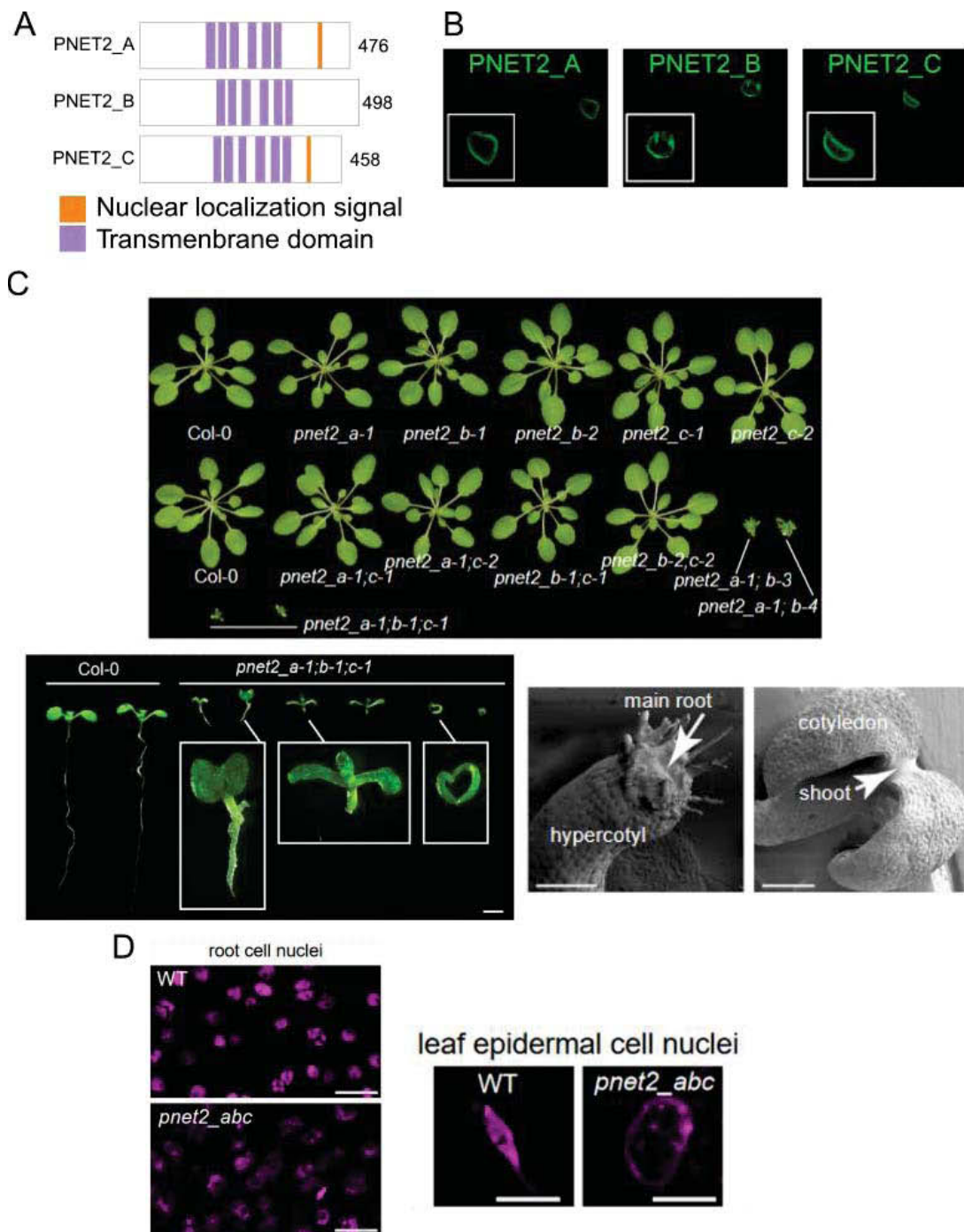


Figure 1.13 : Caractérisation des protéines PNET2 (Tang *et al.*, 2021)

A. Représentation schématique des protéines PNET2 d'*A. thaliana*. **B.** Localisation des protéines PNET2_A-C fusionnées à la YFP en expression transitoire chez le tabac. **C.** Photos des différents mutants *pnnet2_a-c* (les allèles 1 sont ceux obtenus par la délétion d'une base créée par CRISPR et les allèles 2 sont des mutants d'insertion de T-DNA ; les allèles *b-3* et *b-4* ont été générés directement par CRISPR dans le mutant *pnnet2_a-1*) âgés de 4 semaines sur terreau (haut) et de 4 jours sur milieu MS (bas à gauche) avec une image par microscopie électronique à balayage des défauts au niveau de la racine et du méristème apical des triples mutants *pnnet2_abc* (bas à droite). **D.** Marquage au DAPI des noyaux de la racine et de l'épiderme de feuille de plantules de 7 jours du sauvage et du mutant *pnnet2_abc*.

Chez les animaux, les homologues de la protéine MAN1 sont ancrées dans la membrane interne de la membrane nucléaire grâce à la présence de ses deux domaines transmembranaires, la partie centrale de ces protéines se localisant dans l'espace intermembranaire, tandis que les régions N et C-terminales sont retrouvées du côté nucléoplasmique (**Figure 1.12B**).

Chez *A. thaliana*, la protéine MAN1 se localise également à la périphérie nucléaire et le mutant d'insertion *man1-1* ne possède pas de phénotype visible comparé au sauvage (**Figures 1.12C-D**) (Tang *et al.*, 2021), contrairement à la majorité des mutations des protéines LEM animales notamment celle de MAN1 entraînant une létalité embryonnaire chez la souris (Ishimura *et al.*, 2006; Cohen, Kosti and Stewart, 2007). De plus, la protéine HEH2, homologue de MAN1 chez la levure *Schizosaccharomyces pombe*, interagit avec les pores nucléaires et impacte leur distribution, en l'absence de HEH2 les pores nucléaires forment des agrégats (Borah *et al.*, 2021). A ce jour, une seule étude a analysé la fonction de la protéine MAN1 indiquant qu'elle serait capable d'interagir avec certaines protéines du nucléosquelette (Tang *et al.*, 2021) (cf paragraphe I.3.c.iii).

I.3.a.iv) Les protéines PNET2

Les protéines PNET2 sont retrouvées à la fois chez les animaux et les plantes mais avec une divergence de séquence assez importante. Chez *A. thaliana*, trois protéines PNET2 ont été identifiées, elles ont donc été renommées PNET2-A à C (Tang, Huang and Gu, 2020; Tang *et al.*, 2021). Les PNET2 possèdent 6 domaines transmembranaires dans la partie centrale de la protéine ainsi qu'un domaine hélice-tour-hélice en C-terminal. PNET2_A et PNET2_C possèdent également un signal de localisation nucléaire (NLS) (**Figure 1.13A**). Ces trois protéines se localisent à la périphérie nucléaire (**Figure 1.13B**), et sont capables de former des homomères *via* leur région C-terminale (Tang *et al.*, 2021).

Les proximitomes de PNET2_A/B ont été réalisés par la technique de protéomique par marquage de proximité adaptée à partir de Kim *et al.*, 2016, en construisant une protéine de fusion de la protéine d'intérêt, ici PNET2_A/B, avec la biotine ligase BioID2. La biotinylation des protéines proches par BioID2, ont permis d'identifier les protéines CROWDED NUCLEI (CRWN)1, CRWN4 et KAKU4, trois protéines du nucléosquelette (cf paragraphe I.3.c.ii), ainsi que des histones H2A et H2B. Ce résultat suggère que les protéines PNET2 seraient des composants de la membrane nucléaire interne (INM) qui interagissent avec la chromatine ainsi qu'avec le nucléosquelette (cf paragraphe I.3.c.iii) (Tang *et al.*, 2021). Les protéines PNET2 jouent un rôle dans le développement des plantes comme le montre le double mutant *pnet2_a/b*

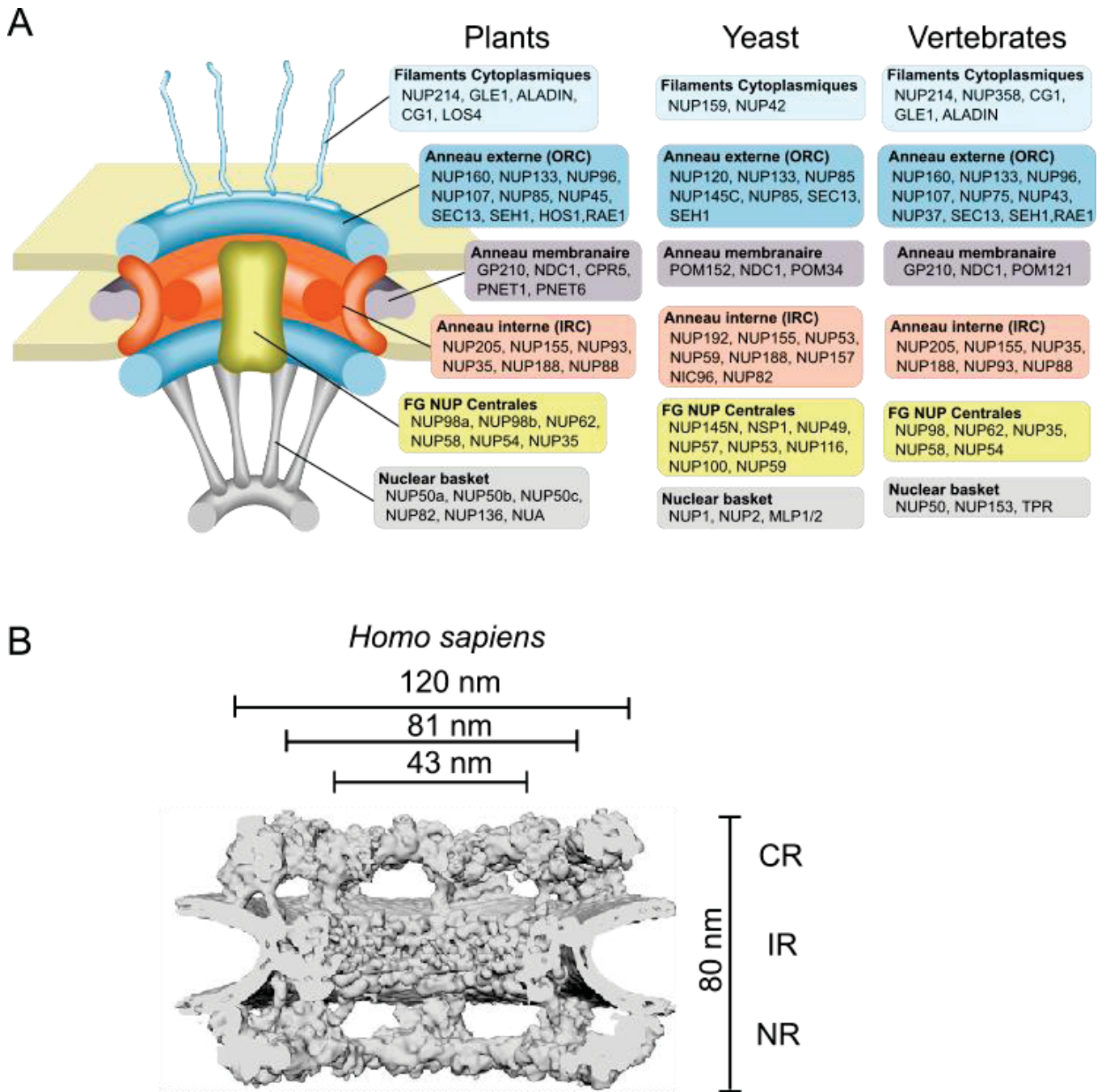


Figure 1.14 : Structure des pores nucléaires

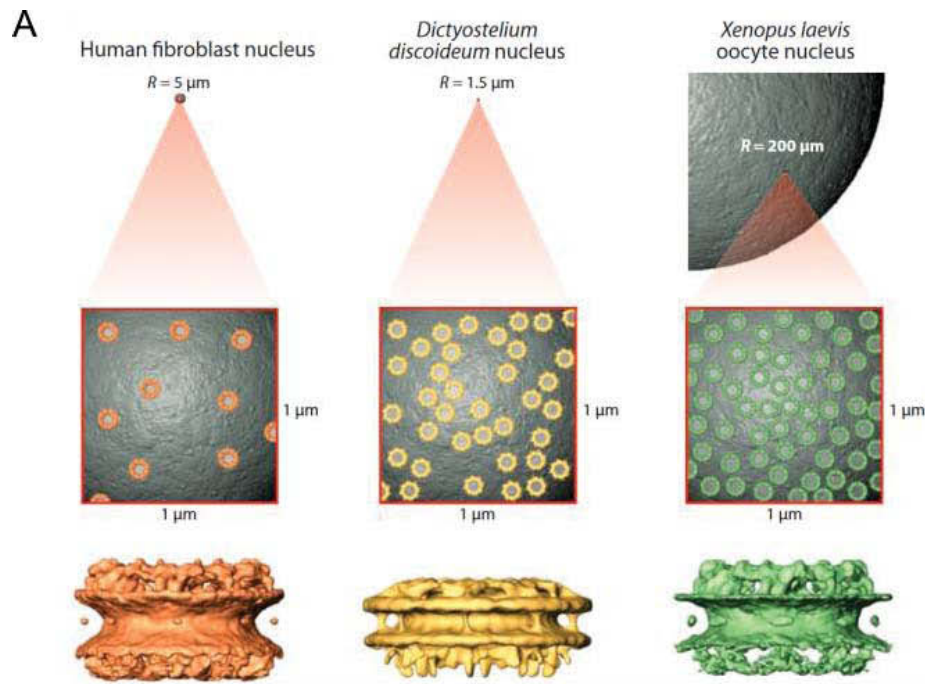
A. Représentation schématique d'un pore nucléaire et de ses différents composants chez les plantes, les levures et les vertébrés (schéma adapté d'Hiromasa ONO) **B.** Structure d'un pore nucléaire humain observé en microscopie électronique. Les chiffres indiquent le diamètre (au dessus) ou la taille à travers l'enveloppe nucléaire (à droite) des différentes structures du pore (CR *Cytoplasmic Ring*, IR *Inner Ring*, and NR *Nuclear Ring*) (Mosalaganti *et al.*, 2018).

induisant des petites plantes et le triple mutant *pnet2_a/b/c* (mutants CRISPR de délétion d'une base entraînant un décalage du cadre de lecture et un codon stop prématuré) qui provoque l'absence de méristème apical et/ou celle de la racine principale induisant une létalité au stade plantule (**Figure 1.13C**). Les PNET2 sont donc requises, notamment PNET2-A ou B pour le développement de la plante et sa survie. L'étude du transcriptome a montré que les doubles mutants *pnet2_a/c* et *pnet2_b/c* ont un faible impact sur la régulation des gènes, moins de 100 gènes dérégulés en RNA-seq alors que le triple mutant *pnet2_a/b/c* dérégule 7457 gènes dont 3836 sont sous-exprimés. Une grande partie des gènes sous-exprimés sont impliqués dans des processus liés au cycle cellulaire et dans la réponse aux phytohormones activatrices de la division cellulaire, ce qui concorde bien avec les observations phénotypiques d'absence de croissance et du développement des plantules. Alors que ce triple mutant sur-exprime de nombreux gènes impliqués dans la réponse immune innée et les différentes voies de signalisations des stress biotiques et abiotiques (acide salicylique : SA, acide jasmonique : JA, acide abscissique : ABA et éthylène : ET). Les PNET2 sont donc essentielles pour maintenir l'homéostasie de l'expression des gènes du développement et de croissance par rapport à ceux du stress pendant le développement précoce des plantules. De plus, les PNET2 impactent la morphologie du noyau en induisant chez le triple mutant des noyaux plus gros et plus ronds (**Figure 1.13D**). Bien que la formation des chromocentres et le contenu global de l'ADN ne semblent pas être affectés, la coloration des noyaux par le DAPI est largement diminuée dans les cellules *pnet2_a/b/c* provenant de divers tissus, ce qui suggère une désorganisation de la chromatine (**Figure 1.13D**) (Tang *et al.*, 2021).

En résumé les protéines PNET2 sont ancrées dans l'INM et interagissent avec le nucléosquelette. Elles impactent la croissance et le développement de la plante ainsi que la morphologie nucléaire et la perte de ces protéines induit une reprogrammation globale de la transcription des gènes ainsi qu'une désorganisation de la chromatine.

[I.3.b\) Le pore nucléaire ou Nuclear Pore Complex \(NPC\)](#)

Les pores nucléaires (en anglais *Nuclear Pore Complex*, NPC) sont des complexes multiprotéiques situés dans la membrane nucléaire traversant de part en part l'enveloppe nucléaire, permettant les échanges entre le cytoplasme et le nucléoplasme.



B

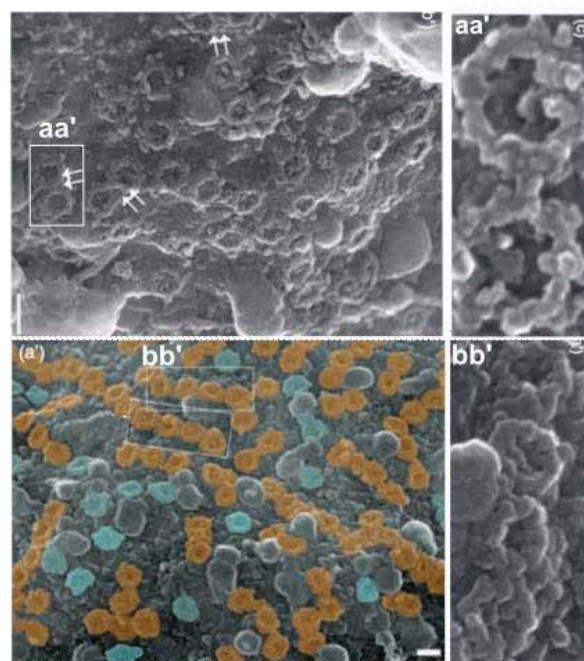


Figure 1.15 : Densité et distribution des pores nucléaires

A. Densité et structure en microscopie électronique des pores nucléaires chez l'Homme, l'unicellulaire *Dictyostelium discoideum* et le Xénope (Grossman, Medalia and Zwerger, 2012).
B. Organisation des pores nucléaires de cellules de tabac âgées de 3 jours sous forme de doublets (en haut) ou âgées de 10 jours qui dans ce cas forme des lignes de pores pouvant aller jusqu'à 30 NPC (en bas). Les NPC se touchant sont représentés en marron tandis que les NPC isolés sont représentés en bleu (Fiserova, Kiseleva and Goldberg, 2009).

I.3.b.i) Structure des pores nucléaires

Chaque NPC forme une structure cylindrique composée de huit sous-unités, chaque sous-unité étant constituée d'une trentaine de protéines différentes appelées nucléoporines (NUP) très conservées au sein des eucaryotes (Tamura *et al.*, 2010) (**Figure 1.14A**). Ce complexe est divisé en sous-structures : *i*) les anneaux externes et internes présents à la fois du côté cytoplasmique et nucléoplasmique du NPC ainsi que les anneaux transmembranaires ou luminaux, structurant le NPC et permettant son ancrage dans l'enveloppe nucléaire ; *ii*) le cœur central, reliant les NUP transmembranaires aux NUP centrales, assure la cohésion du NPC ; *iii*) des NUP ayant des projections dans le cytoplasme formant les filaments cytoplasmiques afin d'importer des protéines du cytoplasme dans le noyau ; *iv*) les NUP centrales riches en répétitions Phénylalanine-Glycine (FG) assurant la translocation des protéines au travers du NPC et *v*) le "nuclear basket", composé de filaments émanant dans le nucléoplasme qui convergent pour former un anneau distal (**Figure 1.14A**).

Les nucléoporines peuvent également être subdivisées en deux catégories : les nucléoporines structurales et les nucléoporines de transport contenant des répétitions FG et nommées nucléoporines FG. Les nucléoporines structurales forment des sous-complexes composant les différents anneaux du NPC dont le diamètre externe est de ~120 nm et le diamètre interne de ~43 nm (**Figure 1.14B**) (Mosalaganti *et al.*, 2018). Quant à elles, les nucléoporines FG sont présentes dans le canal central ainsi que dans les filaments cytoplasmiques et le nuclear basket formant une barrière sélective bloquant le passage de certaines molécules alors que d'autres seront transportées activement à travers le NPC.

La densité et la répartition des pores nucléaires chez les plantes ont été étudiées en cellules de tabac BY-2 par microscopie électronique. Il y aurait 40 à 50 NPC/ μm^2 pour les cellules de 3 jours et environ 50 NPC/ μm^2 pour celles de 10 jours (Fiserova, Kiseleva and Goldberg, 2009). Ces densités sont parmi les plus élevées mesurées chez les eucaryotes avec une densité comprise entre 10 et 20 NPC/ μm^2 pour la plupart des vertébrés comme l'homme ou la grenouille léopard (Maul *et al.*, 1980; Winey *et al.*, 1997; Görlich and Kutay, 1999; Grossman, Medalia and Zwerger, 2012) et seuls les ovocytes de Xénope et les noyaux d'*Acetabularia* (algues vertes unicellulaires) présentent des densités de NPC comparables (plus de 60 NPC par μm^2 ; Gerace et Burke, 1988) (**Figure 1.15A**). La répartition des NPC varie en fonction de l'âge de la cellule ; en effet, dans les cellules de 3 jours les pores sont dispersés au sein de la membrane nucléaire avec 32,7% des pores qui sont physiquement liés par 2, formants ainsi des doublets (**Figure**

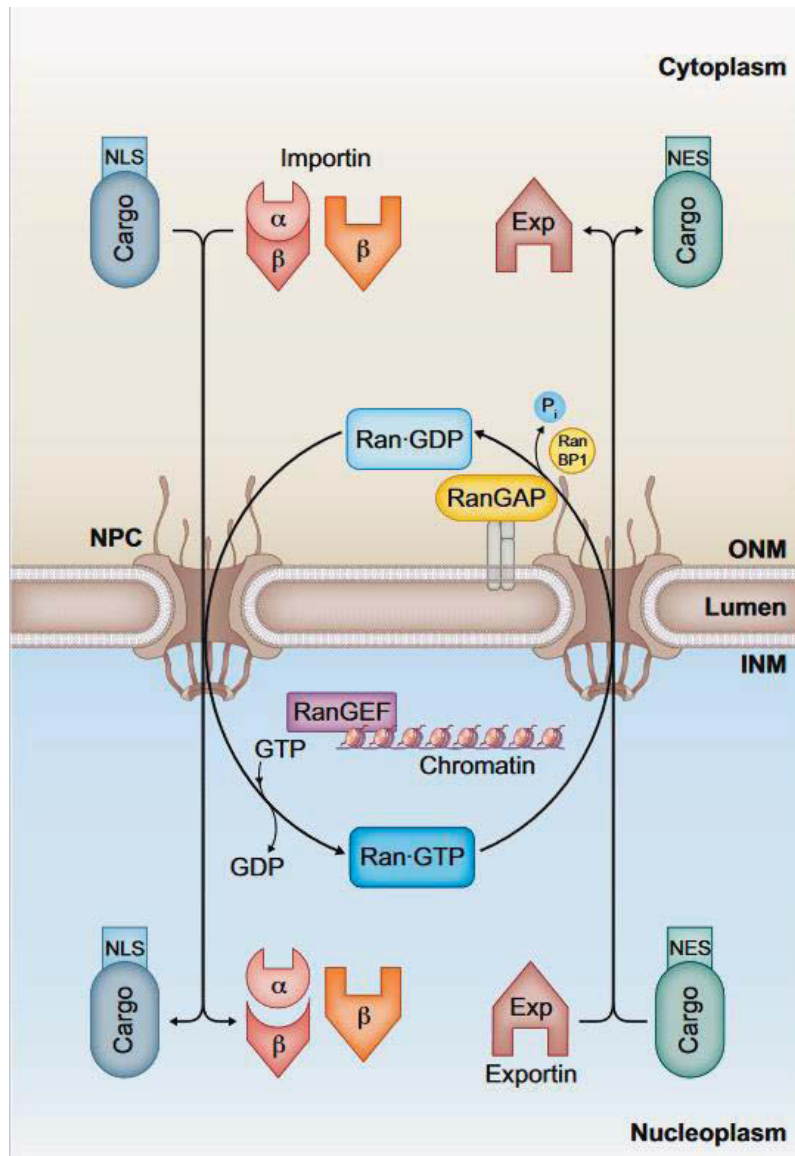


Figure 1.16 : Fonction d'import-export du pore nucléaire (Lüdke *et al.*, 2022)

Les molécules transportées *via* le pore nucléaire par des importines ou exportines (protéines, ARNm,...) sont nommées *cargo*. Les molécules *cargo* possédant un signal de localisation nucléaire (NLS) qui leur permet d'être prises en charge par les importines afin d'être transportées vers le noyau à travers le pore nucléaire *via* des nucléoporines riches en répétitions phénylalanine-glutamine (FG). La conversion de la protéine RAN-GDP en RAN-GTP par la molécule RAN-GEF (RAN GUANINE NUCLEOTIDE EXCHANGE FACTOR) associée à la chromatine permet la dissociation de la molécule *cargo* de l'importine. Les molécules *cargo* exportées du noyau possèdent un signal d'export nucléaire (NES) et sont reconnues par des exportines et forment un complexe avec RAN-GTP permettant ainsi leur export *via* le pore nucléaire. Une fois dans le cytoplasme, l'enzyme RAN-GAP (RAN-GTPase ACTIVATING PROTEIN) hydrolyse RAN-GTP en RAN-GDP ce qui conduit à la dissolution de ce complexe et au relargage des molécules *cargo* dans le cytoplasme.

1.15B), alors que dans les cellules de 10 jours, plus de 73% des NPC sont liés physiquement et forment des lignes de 5 à 30 NPC (**Figure 1.15B**) (Fiserova, Kiseleva and Goldberg, 2009).

I.3.b.ii) Fonctions des pores nucléaires

La fonction principale des NPC est d'importer et d'exporter des macromolécules (protéines, ARNt, ribosomes, ARNm, ...) du cytoplasme vers le noyau et *vice versa*. Les molécules à transporter ou molécules *cargo*, possèdent généralement un signal de localisation nucléaire (Nuclear Localisation Signal, NLS) pour être importées dans le noyau ou un signal d'exportation nucléaire (Nuclear Export Signal, NES) pour être exportées du noyau. Les petites molécules, elles transitent à travers le pore par diffusion passive alors que les molécules de taille supérieure à 40-60 kDa nécessitent des transporteurs (*Nuclear Transport Receptors*, NTR) comme les karyophérines (importines et exportines). Le transport via les pores nucléaires est également dépendant d'un gradient de la protéine RAS-RELATED NUCLEAR PROTEIN-GDP/GTP (RAN-GDP/GTP), et médié par l'interaction des nucléoporines riches en répétitions FG (filaments cytoplasmiques, nuclear basket et NUP centrales) avec les transporteurs (**Figure 1.16**) (Lüdke *et al.*, 2022).

Le NPC des plantes est moins bien caractérisé que celui de la levure ou des mammifères et de nouvelles nucléoporines potentielles continuent d'être identifiées comme PNET1 et PNET6 puis leurs interactions avec les autres NUP sont ensuite décrites (Tang, Huang and Gu, 2020). Au niveau protéique, PNET1 est capable d'interagir physiquement avec des NUP dont CPR5, une NUP transmembranaire qui interagit également avec PNET 6 (**Figure 1.17A**). Au niveau génétique, le double mutant *pnet1 nup160* (anneau externe) est létal au stade embryonnaire, le double mutant *pnet1 hos1* (anneau externe) est stérile alors que les simples mutants n'induisent pas ces phénotypes (**Figure 1.17B**). Ceci suggère une interaction génétique fonctionnelle entre PNET1 et le NPC avec des conséquences sur le développement de la plante. Ainsi, les nouvelles NUP sont décrites via des études d'interactions génétiques et protéiques qui suggèrent pour PNET1 et PNET6 que ce sont de nouveaux composants du NPC qui dans ce cas entrent dans la constitution des anneaux transmembranaires.

Le panier du NPC ou nuclear basket est la partie nucléoplasmique du NPC ; chez *A. thaliana*, il est composé des protéines : NUP50a, NUP50b, NUP50c, NUA, NUP1/NUP136 (nommée par la suite NUP136) et NUP82 (**Figure 1.14A**). Une étude récente de l'interactome de NUP82 par la technique de marquage de proximité couplé à la spectrométrie de masse quantitative sans étiquette (*Proximity Labelling coupled with Label-Free Quantitative Mass*

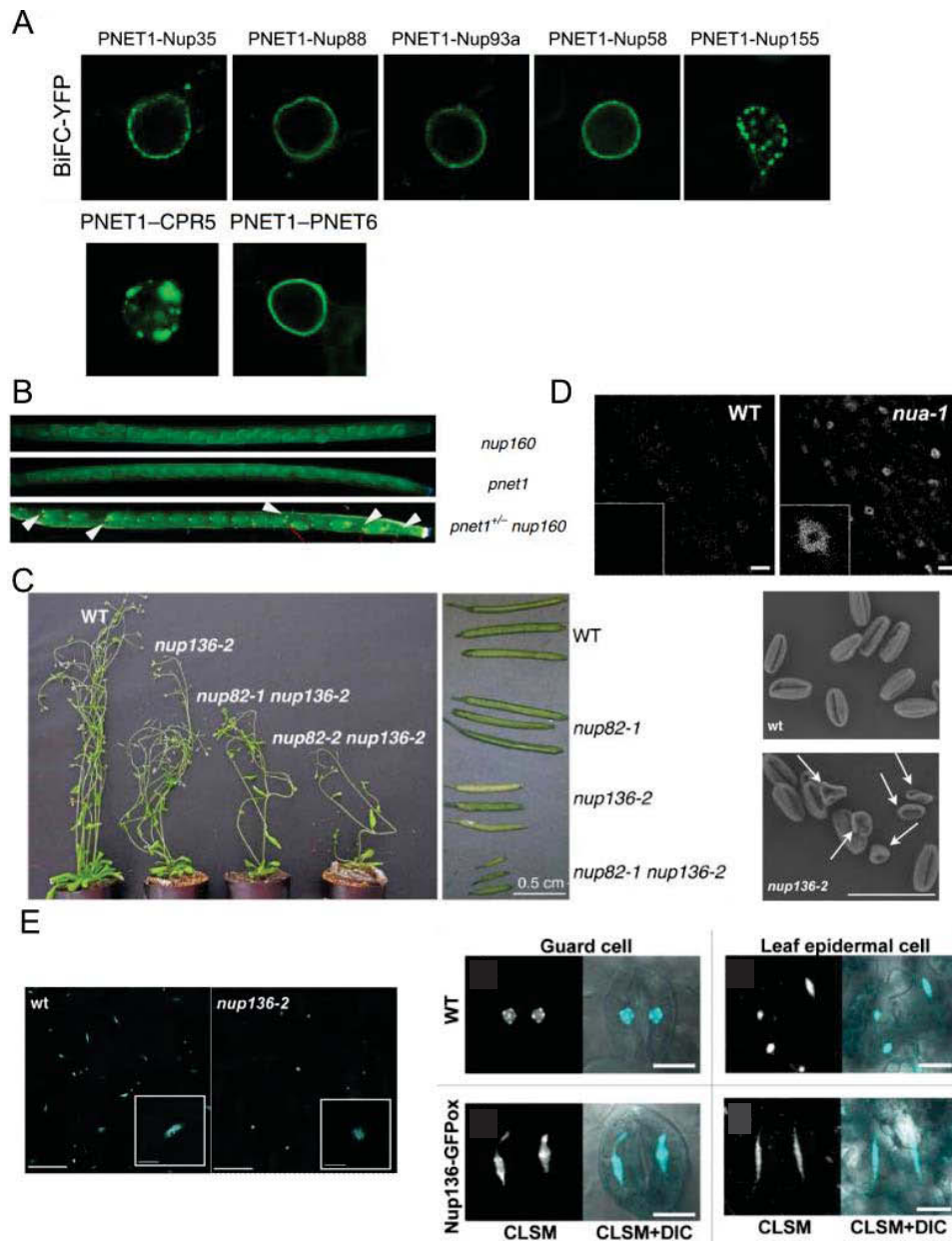


Figure 1.17 : Fonction du NPC dans le développement des plantes

A. Interaction de la protéine PNET1 avec différents composants du pore nucléaire lors de l'utilisation de la technique de *Bimolecular Fluorescence Complementation* (BiFC) (Tang, Huang and Gu, 2020) **B.** Phénotype des mutants *pnet1* et *nup160* au niveau des siliques ; flèches : double mutants *pnet1 nup160* léthal au niveau embryonnaire (Tang, Huang and Gu, 2020). **C.** Photos des phénotypes de développement (plante entière, silique et grain de pollen) des mutants *nup136* et *nup82* ; flèches indiquant les grains de pollen difformes non viables (Tamura *et al.*, 2010, 2017). **D.** Accumulation des ARNm au sein du noyau dans le mutant *nua* par rapport au génotype sauvage (Xu *et al.*, 2007) **E.** Impact de NUP136 sur la morphologie nucléaire avec à gauche le mutant *nup136* induisant des noyaux petits et ronds et à droite la surexpression de NUP136 qui provoque la formation de noyaux allongés dans les différents types cellulaires présents dans l'épiderme de feuille (Tamura *et al.*, 2010; Tamura and Hara-Nishimura, 2011).

d'autres composants du nuclear basket comme NUP136, NUA et NUP50a ainsi que des NUP des anneaux internes (Nup93 a et b, et NUP155) et externes (NUP106, NUP133, NUP96, NUP107 et HOS1) (Tang, Huang and Gu, 2020). NUP82 et NUP136, deux protéines paralogues (Tamura *et al.*, 2017), jouent un rôle dans la défense des plantes aux pathogènes dépendante de l'acide salicylique (cf paragraphe [III-](#)). Les mutants de ces protéines ont des petites siliques et des défauts de formation du pollen (**Figure 1.17C**) (Tamura *et al.*, 2017). Le nuclear basket a également été montré comme interagissant avec le complexe TRanscription-coupled Export (TREX) notamment via NUP136 qui interagit avec THP1, un composant du complexe d'initiation de la transcription Spt-Ada-GCN5-acetyltransferase (SAGA) (Tamura and Hara-Nishimura, 2011, 2013; Tamura, 2020) (cf paragraphe [II-](#)). De plus, les protéines NUA et NUP136 jouent un rôle dans l'export des ARNm (**Figure 1.17D**) (Xu *et al.*, 2007; Tamura *et al.*, 2010). Ces différentes études suggèrent que le NPC joue un rôle dans la régulation de la transcription en maintenant les régions de chromatine devant être activement transcrites proches des NPC, théorie du gene gating (Blobel, 1985; Tamura, 2020).

Outre leur rôle dans les échanges de macromolécules et la régulation de la transcription, les NPC contribuent également au maintien de la morphologie du noyau. En effet chez *A. thaliana*, la perte de fonction du gène *NUP136* entraîne l'apparition de noyaux sphériques alors qu'à l'inverse des noyaux plus gros et très allongés sont observés lors de sa surexpression, (**Figure 1.17E**) (Tamura and Hara-Nishimura, 2011, 2013). Les changements de morphologie nucléaire provoqués par certaines NUP pourraient s'expliquer par une altération du nucléosquelette (cf paragraphe [I.3.c](#)). En effet, NUP153 chez le Xénope qui est un probable homologue fonctionnel de NUP136, interagit avec le nucléosquelette (Smythe, Jenkins and Hutchison, 2000). Chez les plantes, l'interaction NPC/nucléosquelette n'est pas décrite dans la littérature, toutefois en microscopie électronique chez le tabac, des filaments situés sous la membrane nucléaire interne relient les NPC entre eux. Ces filaments pourraient correspondre au nucléosquelette (Fiserova, Kiseleva and Goldberg, 2009).

Pour conclure, les pores nucléaires sont conservés chez l'ensemble des eucaryotes confortant leur rôle essentiel dans plusieurs mécanismes de régulation comme le contrôle des échanges nucléocytoplasmiques, la transcription et la morphologie nucléaire. Le NPC participe activement à la définition de la périphérie nucléaire notamment via sa région nucléoplasmique, le nuclear basket, impactant la régulation de la transcription, l'export des ARNm en interagissant avec les complexes TREX-SAGA2 et interagit avec le nucléosquelette.

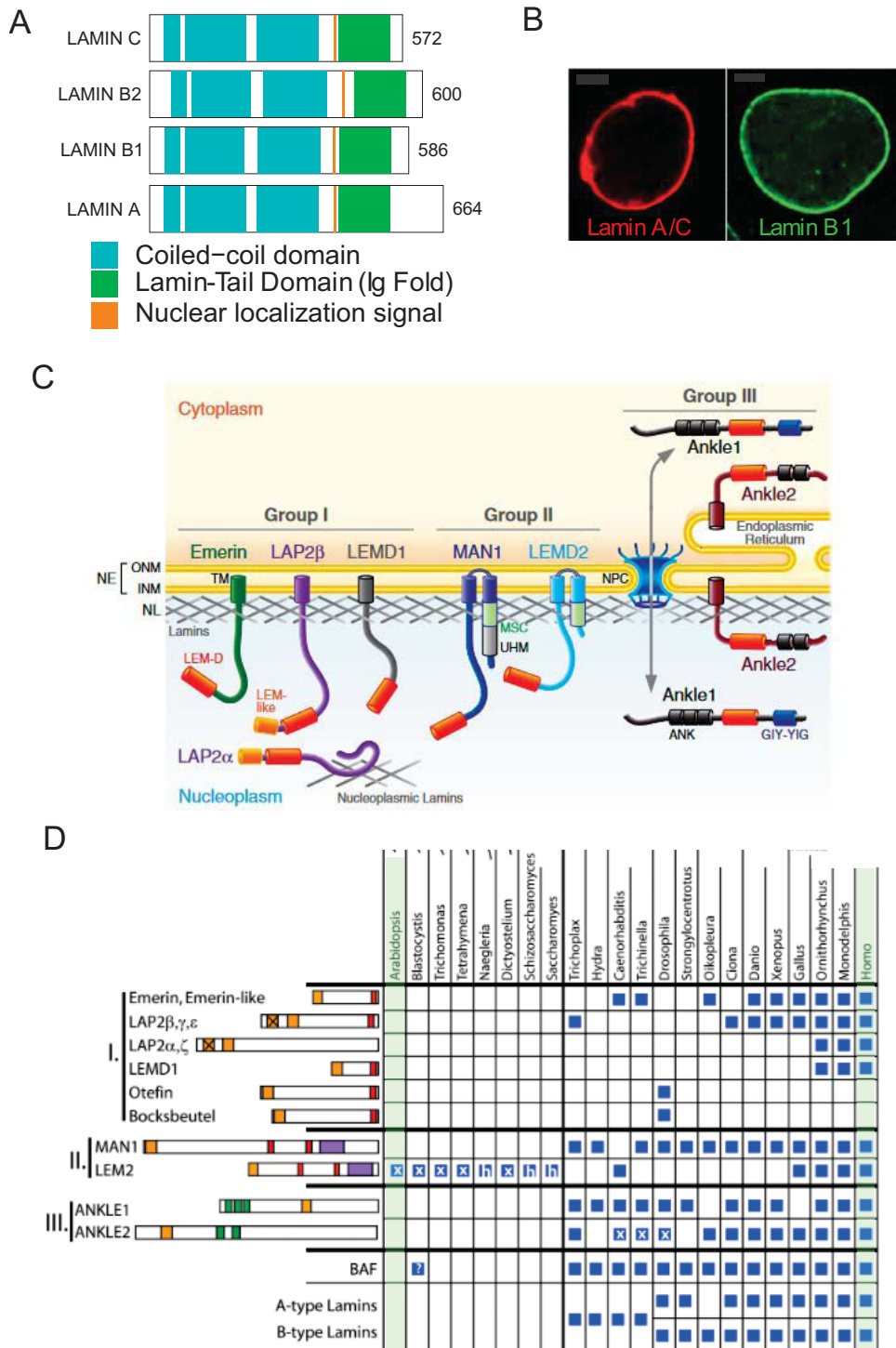


Figure 1.18 : La lamina et ses protéines associées localisées à la périphérie nucléaire

A. Représentation schématique des lamines présentes chez la souris. **B.** Localisation des lamines A/C et B1 visualisées par immunofluorescence dans des cellules humaines (Dittmer and Misteli, 2011). **C.** Schéma représentatif de la périphérie nucléaire chez l'Homme avec les différentes protéines interagissant avec les lamines (Barton, Soshnev and Geyer, 2015). **D.** Analyse phylogénétique des différentes protéines LEM et des lamines chez les eucaryotes ; carré : présence, x : absence domaine LEM, h : présence du motif HeH/SAP (helix-extension-helix/SAF-A/B Acinus and PIAS : structure proche du domaine LEM) à la place du domaine LEM (Brachner and Foisner, 2011).

I.3.c) Le nucléosquelette

Le nucléosquelette est l'un des composants centraux de la périphérie nucléaire. Il est formé d'un réseau de filaments situés sous la membrane nucléaire interne. Le nucléosquelette joue un rôle dans le maintien de la structure du noyau ainsi que la régulation de l'expression des gènes, deux fonctions également décrites pour le NPC.

I.3.c.i) *Le nucléosquelette ou lamina chez les animaux*

Chez les animaux, le nucléosquelette est appelé la lamina, il est composé majoritairement des Lames A/B et C qui possèdent de longs domaines coiled-coil permettant leur polymérisation en un réseau de filaments enrichi à la périphérie nucléaire. Aucun domaine transmembranaire n'a été décrit dans les lames et leur enrichissement à la périphérie du noyau doit être assuré par des *Lamin-Binding Proteins* qui sont ancrées à l'INM et interagissent avec les lames (**Figure 1.18A**) (cf paragraphe I.3.c.iii). Parmi ces protéines, on retrouve les protéines LBR et les protéines LEM (cf paragraphes [I.3.a.iii](#)) et I.3.c.iii). La localisation des lames n'est pas strictement restreinte à la périphérie nucléaire ; en effet, si les lames A/C sont localisées en périphérie du noyau, la lame B est également nucléoplasmique, suggérant une certaine mobilité et suggérant une fonction différente comme l'interaction avec la chromatine présente à l'intérieur du noyau (**Figures 1.18B-C**) (Dittmer and Misteli, 2011; Adam and Goldman, 2012). Les lames sont impliquées dans un large éventail de fonctions nucléaires, notamment l'organisation spatiale et locale de la chromatine (cf paragraphe I.2.c.iii), la transcription, la réplication, la réparation de l'ADN et constitueraient un réseau de filaments relativement rigide qui maintiendrait la morphologie de la membrane nucléaire.

Les mutations des lames sont à l'origine de maladies humaines connues collectivement sous le nom de laminopathies comme le syndrome de Hutchinson-Gilford ou progeria dont le phénotype s'exprime sous forme d'un vieillissement précoce, une altération de l'organisation de la chromatine notamment de l'hétérochromatine et des noyaux déformés (Dechat *et al.*, 2010), ce qui souligne leur importance fonctionnelle. Un autre rôle de lames est d'interagir avec certaines régions de la chromatine, les *Lamin-Associated Domains* (LAD) révélées grâce à la technique de DamID (Pickersgill *et al.*, 2006; Guelen *et al.*, 2008b) (cf paragraphe I.2.c.iii). Les LAD couvrent environ un tiers du génome chez les mammifères et sont enrichis en séquences dont les caractéristiques rappellent celles de l'hétérochromatine, c'est-à-dire des régions de faible densité en gènes, peu transcrites et avec des marques épigénétiques répressives H3K9me2, H3K9me3 ainsi que H3K27me3 en bordure (Pickersgill *et al.*, 2006; van Steensel and Belmont, 2017). La moitié des LAD est constitutive et est toujours associée à la

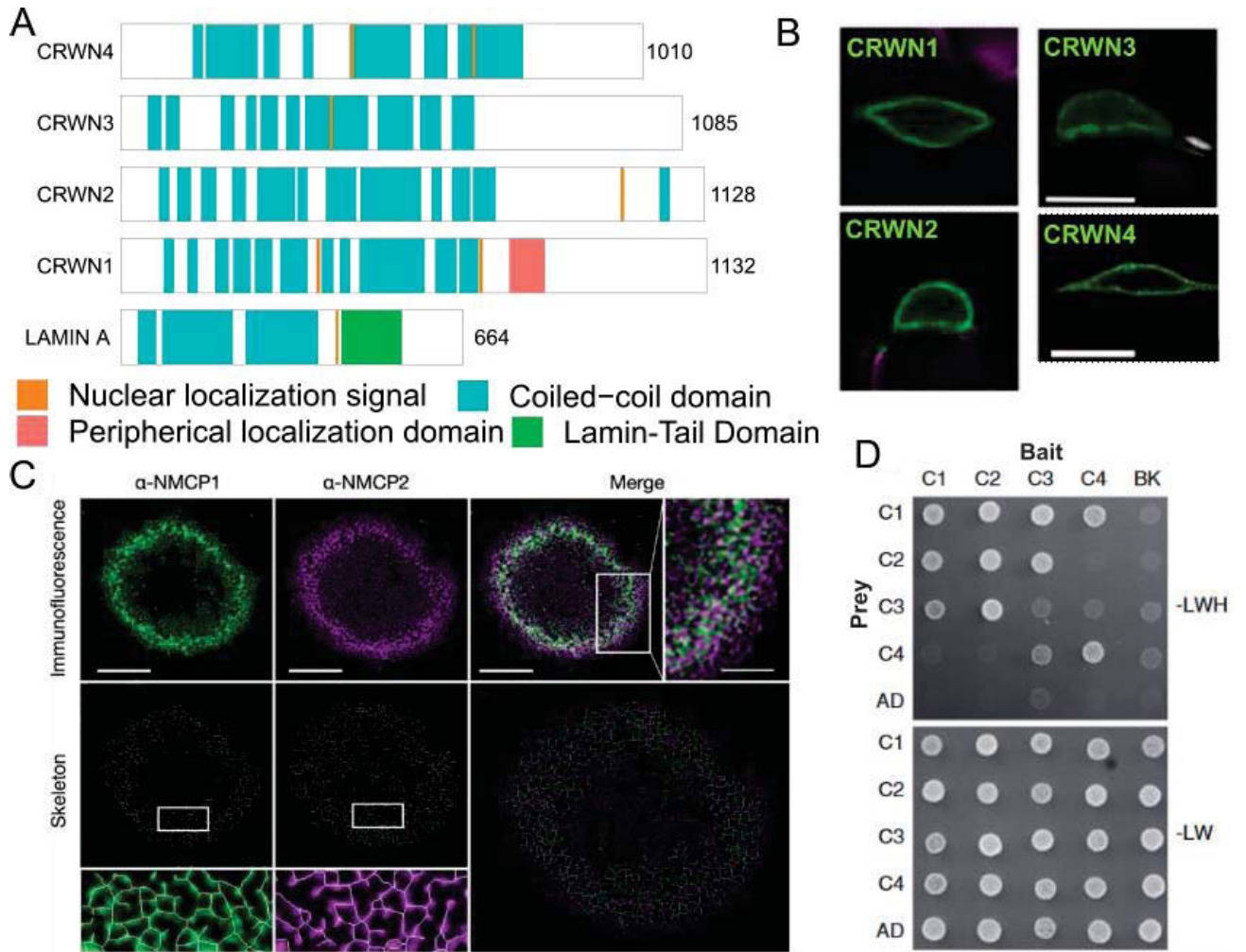


Figure 1.19 : Caractéristiques du nucléosquelette des plantes

A. Représentation schématiques des protéines CRWN d'*A. thaliana* et de la lamine A de la souris. **B.** Localisation des protéines CRWN1 à 4 par expression transitoire chez le tabac (Sakamoto and Takagi, 2013; Sakamoto *et al.*, 2020). **C.** Immunofluorescence des protéines NMCP1 et NMCP2 dans des cellules AG3 (*Apium graveolens*) observée par microscopie STED. Les panneaux inférieurs montrent le squelette (skeleton) segmentés à partir des signaux d'immunofluorescence, les grossissements du squelette sont superposés à l'image d'immunofluorescence ; échelle : 2,5 μm (images noyaux entiers) et 1 μm (grossissements) (Masuda, Hikida and Fujino, 2021). **D.** Interactions en double-hybride des protéines CRWN1-4 (C1 à C4) d'*A. thaliana* entre-elles ; -LW : milieu sans leucine ni tryptophane (milieu permissif), -LWH : milieu sans leucine, tryptophane et histidine (milieu test à faible stringence) (Sakamoto *et al.*, 2020).

lamina, tandis que l'autre moitié est facultative, et montre des changements dynamiques de leur position par rapport à la lamina et ce en fonction du type cellulaire, du stade de différenciation ou de stimuli externes.

I.3.c.ii) *Le nucléosquelette chez les plantes*

Chez les plantes, il n'existe pas d'orthologues aux lamines, ni aux protéines permettant de les ancrer à l'enveloppe nucléaire comme les LBR ou les LEM. Seule la protéine MAN1 a récemment été décrite (cf paragraphe I.3.a.iii) mais ne possède pas le domaine LEM comme son homologue chez les animaux (**Figure 1.18D**). Cependant, une famille de protéines a été identifiée sur la base de similitudes fonctionnelles avec les lamines, notamment sur des similitudes structurelles comme la présence de longs domaines coiled-coil et des similitudes sur leur implication dans la morphologie nucléaire et sur l'organisation de la chromatine. Ces protéines appelées NUCLEAR MATRIX CONSTITUENT PROTEIN 1 et 2 (NMCP1 et 2) ont initialement été découvertes chez *Daucus carota* (Masuda *et al.*, 1997; Kimura *et al.*, 2014) puis identifiées chez *A. thaliana*, appelées dans un premier temps, protéines LINC pour LITTLE NUCLEI dû au phénotype des mutants de petits noyaux mais pour éviter toute confusion avec le complexe LINC, elles ont été renommées CROWDED NUCLEI (CRWN) (Dittmer *et al.*, 2007). Les protéines NMCP/CRWN sont des protéines spécifiques du règne végétale présentes à partir des Charophytes, les algues vertes, jusqu'aux Angiospermes, les plantes à fleurs (Poulet, Probst, *et al.*, 2017; Ciska *et al.*, 2019).

I.3.c.ii.a) Les protéines CRWN

Structure et localisation des protéines CRWN/NMCP

Chez *A. thaliana*, il existe quatre protéines CRWN (CRWN1 à 4) et l'analyse phylogénétique de ces protéines montre une séparation en deux sous-familles : les protéines CRWN1 à 3 se trouvent dans le groupe monophylétique des protéines de type NMCP1 alors que la protéine CRWN4 est apparentée à NMCP2 (Wang, Dittmer and Richards, 2013).

Bien qu'à la différence des lamines, les CRWN ne possèdent pas de domaine Lamin-Tail (Ig domain), elles partagent avec elles la présence de plusieurs longs domaines coiled-coil dans leur région N-terminale (**Figure 1.19A**). Des fusions des protéines CRWN1 à 4 avec la GFP ou YFP (Green/Yellow Fluorescent Protein) se localisent préférentiellement à la périphérie nucléaire et dans certaines études CRWN2 et CRWN3 se localisent à la fois périphérie nucléaire et dans le nucléoplasme (**Figure 1.19B**) (Dittmer *et al.*, 2007; Sakamoto and Takagi, 2013; Sakamoto *et al.*, 2020). Lors de la division cellulaire en tissus somatiques, la protéine CRWN1 d'*A. thaliana* se localise sur la chromatine pendant la prométaphase et l'anaphase, alors que les

autres protéines CRWN sont localisées de façon diffuse dans le cytoplasme. Puis au début de la télophase, lorsque l'enveloppe nucléaire s'assemble, CRWN1 est transféré de la chromatine vers l'enveloppe nucléaire. Les protéines CRWN2 et CRWN3 sont, quant à elles transférées du cytoplasme vers la chromatine proche de l'enveloppe nucléaire alors que CRWN4 est assemblée en structures ponctuées dans le cytoplasme avant d'être relocalisée vers l'enveloppe en fin de télophase (Sakamoto and Takagi, 2013). Ceci suggère que les quatre protéines CRWN pourraient avoir des fonctions spécifiques comme l'illustre leur dynamique de liaison à la chromatine et d'enrichissement à la périphérie nucléaire lors de la mitose.

Tout comme les lamines, les protéines CRWN se localisent à la périphérie nucléaire malgré l'absence de domaine transmembranaire. Deux études différentes ont été réalisées afin d'identifier la région de CRWN/NMCP nécessaire à cet enrichissement. La première étude menée sur la protéine NMCP1 chez *D. carota* a identifié *in silico* quatre NLS dans NMCP1 dont deux, NLSm3 et NLSm4 (NLS monopartite, 1004–1008AA et 1023–1026AA) sont fonctionnels *in vivo*, c'est-à-dire impliqués dans la localisation de NMCP1 à la périphérie nucléaire. Ces deux NLS sont conservés entre les espèces végétales et présents uniquement dans NMCP1/CRWN1-3 (Kimura *et al.*, 2014; Blunt *et al.*, 2020). La présence de la région N-terminale 1-141 acides aminés (AA) avec les NLSm3 et NLSm4 permet de localiser NMCP1 à la périphérie nucléaire, cependant la localisation correcte de NMCP1 est dépendante de l'hexapeptide R/QYNLRR situé en amont et chevauchant NLSm4 (Kimura *et al.*, 2014). Toutefois, chez *A. thaliana*, la région de 751-820 AA de CRWN1 participe à son enrichissement périphérique bien qu'une part non négligeable de la protéine reste dans le nucléoplasme, suggérant que plusieurs régions pourraient être nécessaires (Guo *et al.*, 2017). Ces deux études suggèrent que la région protéique responsable de la localisation des protéines NMCP/CRWN à la périphérie nucléaire pourrait varier d'une espèce à une autre.

In vivo, les protéines NMCP1 ou NMCP2 de *D. carota* s'associent en homomères formant des filaments d'environ 100 nm de long et 32 nm de diamètre qui ensuite s'assemblent pour former des faisceaux de filaments. Seule, NMCP2 ne semble pas être capable d'assembler des filaments de taille supérieure à 300nm, suggérant que la formation de ces filaments nécessite la coopération de NMCP1 et NMCP2 et ainsi qu'elles possèdent des fonctions différentes (Masuda, Hikida and Fujino, 2021). De plus, chez *Apium graveolens*, l'utilisation de microscopie à haute résolution de type *Stimulated emission depletion* (STED) a permis de détecter les filaments du nucléosquelette *in vivo* au niveau de la périphérie nucléaire (**Figure 1.19C**) (Masuda, Hikida and Fujino, 2021). Bien que chez *A. thaliana* ces mêmes expériences

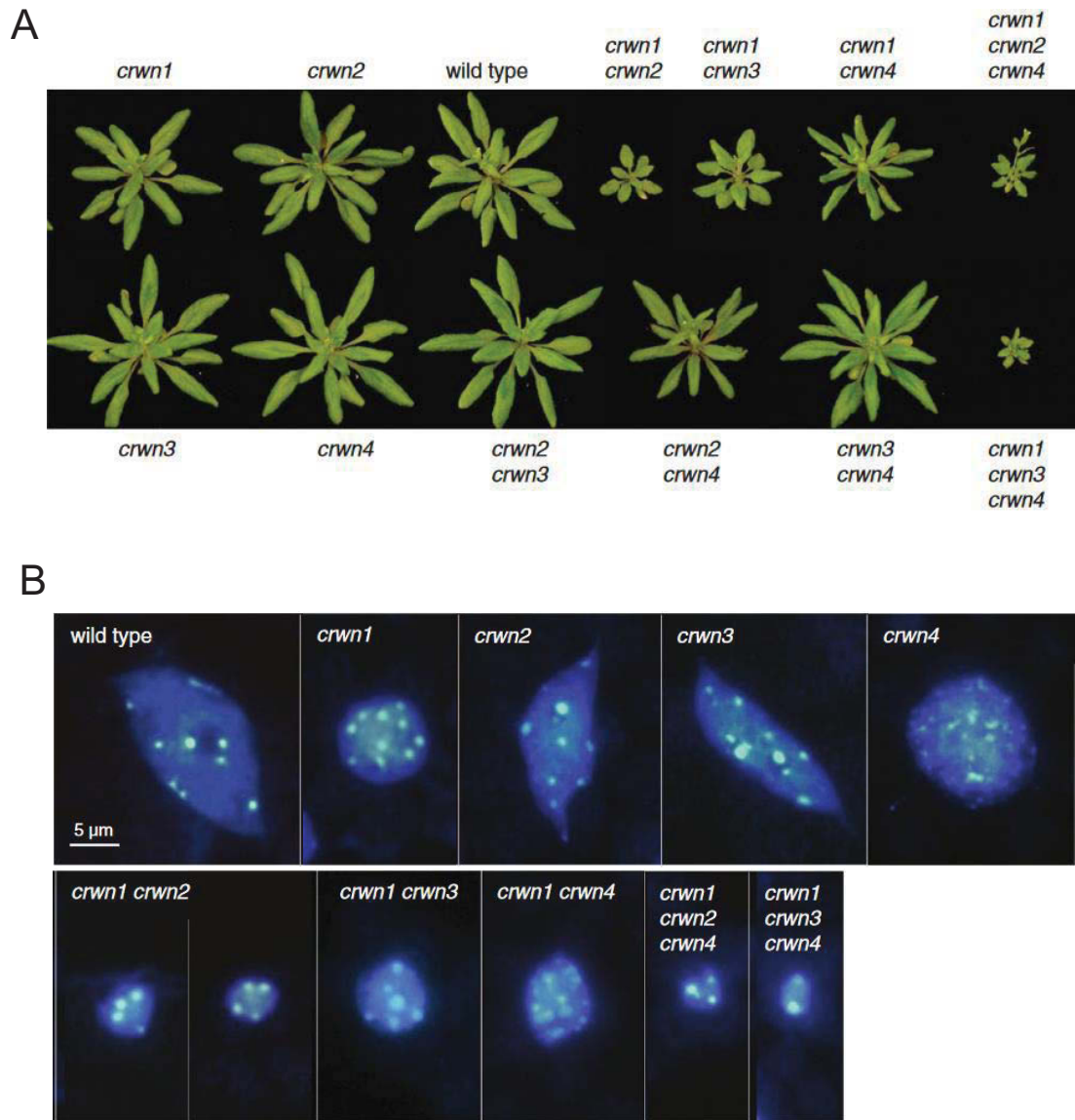


Figure 1.20 : Phénotypes développementaux et nucléaires des mutants *crwn* (Wang, Dittmer and Richards, 2013)

A. Photos des différents mutants *crwn* âgés d'un mois juste après l'initiation de la floraison, le diamètre de la rosette sauvage est d'environ 8 cm. **B.** Images de noyaux isolés représentatifs de cellules de feuilles adultes colorées au DAPI des différents mutants *crwn*.

n'aient pas encore été réalisées, chaque protéine CRWN est capable d'interagir avec elle-même en double-hybride et/ou Co-IP, mais également d'interagir entre elles par exemple, les protéines CRWN1, CRWN2 et CRWN3 sont capables d'interagir entre-elles et seules CRWN1 et CRWN3 interagissent avec CRWN4 (**Figure 1.19D**) (Sakamoto *et al.*, 2020).

L'étude du proxitome de CRWN1 *via* la technique de protéomique par marquage de proximité a montré que CRWN1 est proche de CRWN4 et de KAKU4 (cf paragraphe I.3.c.iii) mais aussi des deux protéines CterSUN : SUN1 et SUN2 et de NUP50c un composant du nuclear basket. Contrairement à d'autres protéines de la périphérie nucléaire comme PNET2 ou KAKU4, aucune histone ou protéine liant l'ADN n'est retrouvée à proximité de CRWN1 (cf paragraphes [I.3.a.iv](#)) et I.3.c.ii) (Tang *et al.*, 2021). Cependant, il existe une interaction faible en double-hybride sur milieu à faible stringence entre les protéines CRWN1 et PWWP INTERACTOR OF POLYCOMBS1 (PWO1) qui est associée au complexe polycomb PRC2 qui lie la chromatine (Mikulski *et al.*, 2019).

Fonction des protéines CRWN

L'étude phénotypique sur le développement des plantes portant des mutations pertes de fonctions des gènes *CRWN* a permis de montrer que les simples mutants *crwn* n'affectent pas la taille de la rosette des plantes (**Figure 1.20A**) (Wang, Dittmer and Richards, 2013). Les combinaisons de mutants *crwn2 crwn3* et *crwn3 crwn4* ont également un phénotype similaire au sauvage alors que les doubles mutants *crwn2 crwn4* et *crwn1 crwn4* présentent des rosettes légèrement plus petites, tandis que les autres doubles mutants, *crwn1 crwn2* et *crwn1 crwn3*, sont encore plus affectés (**Figure 1.20A**). Parmi les triples mutants, seules les combinaisons *crwn1 crwn2 crwn4* et *crwn1 crwn3 crwn4* ont pu être obtenus, tous deux sont extrêmement petits et ne donnent que peu de graines ; aucun quadruple mutant n'a pu être obtenu (**Figure 1.20A**) (Wang, Dittmer and Richards, 2013). L'ensemble des résultats obtenus chez *Arabidopsis* suggèrent que la famille de protéines CRWN est essentielle à la viabilité de la plante et que CRWN1 et CRWN4 auraient chacune des fonctions essentielles non redondantes alors CRWN2 et CRWN3 seraient partiellement redondantes avec CRWN1 et CRWN4. Cependant, chez *Marchantia polymorpha* qui ne possède qu'une seule protéine NMCP, le mutant perte de fonction est viable et induit seulement une diminution de croissance (Wang *et al.*, 2021). Les protéines CRWN pourraient avoir acquis d'autres fonctions au cours de l'évolution les rendant maintenant essentielles chez les dicotylédones.

L'analyse de différents mutants perte de fonction *crwn* réalisée par Sakamoto and Takagi, 2013 et Wang et al., 2013, montre que seules les protéines CRWN1 et CRWN4 jouent un rôle dans la morphologie du noyau. En effet, les mutants *crwn1* et *crwn4* possèdent de petits noyaux relativement ronds comparés à la lignée sauvage (**Figure 1.20B**) (Wang, Dittmer and Richards, 2013). Là encore tout comme pour la viabilité, les protéines CRWN2 et CRWN3 auraient des fonctions partiellement redondantes avec CRWN1 et CRWN4 avec par exemple le double mutant *crwn1 crwn2* formant de très petits noyaux ronds plus petits que ceux de *crwn1* seul alors que *crwn2* seul n'a pas d'impact sur la taille des noyaux (**Figure 1.20B**).

Outre l'impact des protéines CRWN sur la croissance des plantes et la forme du noyau, les protéines CRWN ont également un rôle lors de la germination des graines dont l'un des acteurs est l'acide abscissique (ABA). Cette hormone joue un rôle vital dans l'inhibition de la germination des graines (la dormance) et dans l'établissement des plantules après la germination. La voie de réponse à l'ABA s'exerce au travers de différentes protéines dont le facteur de transcription ABI5 (ABA INSENSITIVE 5), régulant négativement certains gènes nécessaires à la germination. L'inhibition médiée par ABI5 est levée lors de la germination par une dégradation protéasome dépendante qui s'effectue dans des corps nucléaires (voir paragraphe [1.2.a](#)). Or, le mutant *crwn1 crwn3* est hypersensible à l'ABA c'est-à-dire qu'en présence d'ABA, son taux de germination est extrêmement faible (diminution de 58 % du taux de germination contre 23% pour le sauvage). Cette hypersensibilité s'expliquerait par une accumulation plus importante de la protéine ABI5 dans le double mutant *crwn* par rapport au sauvage (Zhao *et al.*, 2016). Enfin, la protéine CRWN3 est retrouvée dans les mêmes corps nucléaires où la protéine ABI5 est dégradée. L'ensemble de ces résultats suggère que les protéines CRWN réguleraient négativement l'accumulation de la protéine ABI5 en participant à la dégradation d'ABI5 par le protéasome dans des corps nucléaires spécialisés (Zhao *et al.*, 2016).

Plus tard dans le développement de la plante, les doubles mutants *crwn1 crwn3*, *crwn2 crwn3* et *crwn2 crwn4* présentent une sénescence foliaire plus précoce, une mort cellulaire foliaire accrue et une suraccumulation de ROS dans des conditions de croissance normales. Dans des conditions de traitement par méthane sulfonate de méthyl (MMS) à 0,01 % induisant des dommages à l'ADN, les mutants *crwn*, en particulier les trois doubles mutants cités ci-dessus, accumulent plus de *Reactive Oxygen Species* (ROS) et leur ADN subit des dommages plus graves que le sauvage (Wang *et al.*, 2019).

Dans l'ensemble, ces données démontrent clairement que les protéines de la famille CRWN jouent des rôles importants dans la diminution de l'accumulation des ROS et dans la protection de l'ADN contre les dommages oxydatifs excessifs causés par le MMS.

Les protéines CRWN affectent également l'organisation de la chromatine et la régulation transcriptionnelle des gènes. Le mutant *crwn4* présente une dispersion des chromocentres suggérant un rôle dans l'organisation de l'hétérochromatine (**Figure 1.20B**) (Wang, Dittmer and Richards, 2013). Le mutant *crwn1 crwn2* possède un nombre réduit de chromocentres qui semblent fusionner (**Figure 1.20B**) (Dittmer *et al.*, 2007; Wang, Dittmer and Richards, 2013). Ces observations sont assez surprenantes puisque ces changements dans l'hétérochromatine n'affectent pas la transcription des éléments transposables qui sont localisés préférentiellement dans les régions péri-centromériques chez *Arabidopsis* (Choi, Strickler and Richards, 2019). A l'inverse les mutants *crwn* affectent de nombreux gènes impliqués dans la réponse aux pathogènes via la réponse à l'acide salicylique (SA) (cf paragraphe [III.3.c](#)) (Guo *et al.*, 2017; Choi, Strickler and Richards, 2019). Ceci suggère que les protéines CRWN joueraient un rôle dans la régulation transcriptionnelle des gènes tels que ceux impliqués dans la réponse au stress biotique comme les gènes *PATHOGENESIS-RELATED (PR)*.

Interaction des protéines CRWN/NMCP avec la chromatine

Ces changements d'expression pourraient s'expliquer par des modifications de position de certains gènes vis-à-vis de la périphérie nucléaire comme cela a été montré pour CRWN1 lors d'une étude sur un cluster de gènes localisé sur le chromosome 5 chez *Arabidopsis* (Sakamoto *et al.*, 2020). Ce cluster contient un ensemble de 11 gènes qui sont induits transcriptionnellement lorsque la plante est soumise à un excès de cuivre (*copper-associated genes*, CA). Des analyses de *Chromatin Integration Labeling* (ChIL) ont permis d'obtenir les sites de liaisons entre la protéine CRWN1 et la chromatine et montre que CRWN1 interagit avec le cluster de gènes CA et recrute ce cluster à la périphérie nucléaire dans des conditions d'excès de cuivre induisant l'expression des gènes CA. Dans le double mutant *crwn1 crwn4*, cette relocalisation est abolie et les gènes CA sont sous-exprimés conduisant à une faible tolérance à l'excès de cuivre (Sakamoto *et al.*, 2020) (cf paragraphe [II.4](#)) et [III.4.a.i](#)).

Une étude plus globale à l'échelle du génome par une analyse de ChIP-seq réalisée pour CRWN1 montre son interaction avec de nombreuses régions chromatiniennes dont l'une des caractéristiques est leur localisation préférentielle à la périphérie nucléaire (Hu *et al.*, 2019).

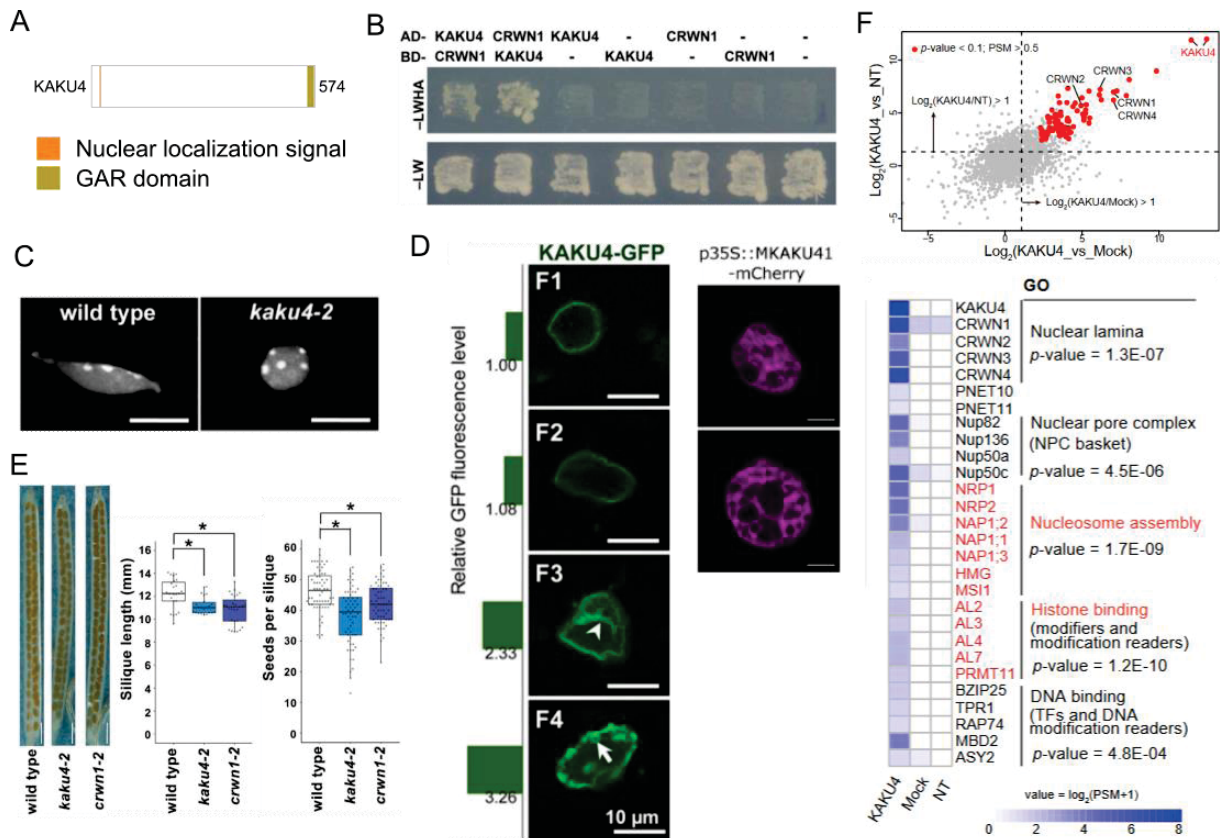


Figure 1.21 : Caractérisation de la protéine KAKU4

A. Représentation schématique de la protéine KAKU4 d'*A. thaliana*. **B.** Interactions en double-hybride des protéines KAKU4 et CRWN1 ; -LW : milieu sans leucine ni tryptophane (milieu permissif), -LWHA : milieu sans leucine, tryptophane, histidine et adénine (milieu test à forte stringence) (Goto *et al.*, 2014). **C.** Morphologie des noyaux de poils racinaires marqués au Hoechst 33342 ; échelle : 10 μm (Goto *et al.*, 2014). **D.** Expression transitoire de KAKU4 d'*A.thaliana* à gauche et du maïs à droite chez le tabac. Plus KAKU4 est exprimée et plus les noyaux possèdent des invaginations ou structures membranaires dans le noyau (Goto *et al.*, 2014; McKenna *et al.*, 2021). **E.** Taille des siliques et nombre de graines par siliques des mutants *kaku4-2* et *crwn1-2* (Goto *et al.*, 2020). **F.** Ensemble des protéines obtenues dans le proxitome de KAKU4 réalisé par protéomique par marquage de proximité (Tang *et al.*, 2021).

Ces régions rappellent les LAD identifiés dans les cellules animales par la technique du DamID (voir paragraphe I.3.c.i).

Chez *Marchantia polymorpha* ne contenant qu'une copie de NMCP, la mutation *nmcp* entraîne la délocalisation vers une position plus centrale dans le noyau de séquences répétées présentes en multicopies sur le chromosome 5 qui sont proches de la périphérie nucléaire chez le sauvage (Wang *et al.*, 2021). Ce changement de position entraîne une décondensation de cette région, mais à ce jour, aucune étude d'expression n'a été effectuée pour voir si cette décondensation était associée à une réactivation transcriptionnelle. Bien que l'association de NMCP de *Marchianta* puisse étayer le rôle de cette famille de protéines dans le recrutement de régions comme les LAD, il convient de noter que la perte de fonction de NMCP n'a pas d'effet drastique sur l'organisation globale du génome ou sur les interactions de la chromatine à la périphérie nucléaire (Wang *et al.*, 2021).

Les protéines CRWN semblent donc jouer un rôle dans le positionnement de certains gènes, séquences répétées particulières voire de régions chromosomiques à la périphérie nucléaire afin de réguler leur expression. Ce recrutement pourrait avoir un effet répressif ou activateur en fonction des séquences considérées (cf paragraphe II.5).

I.3.c.ii.b) La protéine KAKU4

Lors d'un crible génétique cherchant à identifier des mutants qui affectent la morphologie nucléaire réalisé par l'équipe de Kentaro Tamura au Japon, plusieurs mutants *KAKU* (noyau en japonais) ont été isolés notamment un nouvel allèle de *CRWN1* initialement appelé *kaku2-1* (allèle qui dans ce manuscrit sera nommé *crwn1-2*) et un mutant du gène codant une nouvelle protéine : KAKU4 (Goto *et al.*, 2014). KAKU4 est une protéine spécifique des plantes à fleurs (angiospermes), qui possède un domaine Riche en Glycine et en Arginine (domaine GAR) (**Figure 1.21A**). Le domaine GAR est généralement impliqué dans la liaison aux ARN ou dans l'interaction avec d'autres protéines possédant un domaine GAR (McBride *et al.*, 2009). Une expérience de GFP-TRAP suivie d'un séquençage en spectrométrie de masse a permis de montrer que KAKU4 est associée aux protéines CRWN1 et CRWN4. L'interaction avec CRWN1 a par la suite été confirmée par double-hybride (Goto *et al.*, 2014) montrant que KAKU4 est une *CRWN-Associated Protein* (**Figure 1.21B**). Les protéines KAKU4 et CRWN se localisent à la périphérie nucléaire malgré l'absence de domaine transmembranaire et des expériences préliminaires réalisées par le groupe japonais n'ont pas permis d'expliquer par

quel(s) mécanisme(s) ces deux protéines sont enrichies à la périphérie du noyau (Goto *et al.*, 2014).

KAKU4 partage plusieurs propriétés avec les protéines CRWN, elle affecte la morphologie nucléaire et sa perte de fonction conduit à une perte de l'élongation nucléaire dans les tissus différenciés et inversement lors de sa sur-expression, elle provoque des invaginations dans le noyau (**Figures 1.21C-D**) (Goto *et al.*, 2014). Les mêmes défauts de morphologie nucléaire ont également été observés chez le maïs avec MKAKU41 l'un des deux homologues de KAKU4 du maïs. Les invaginations observées lors de la surexpression de MKAKU41 possèdent également des marqueurs de protéines de l'INM (SUN2), de l'ONM (MLKP1) et du RE/ONM (CNX) ce qui suggère que ce ne sont pas seulement des invaginations d'une seule protéine mais de l'ensemble de l'enveloppe nucléaire (**Figure 1.21D**) (McKenna *et al.*, 2021).

La fonction précise de la protéine KAKU4 reste inconnue, cependant KAKU4 est fortement exprimée dans le grain de pollen mature et notamment dans le noyau végétatif dont la forme irrégulière devient plus régulière et ronde dans le mutant. Dans un contexte sauvage, après la germination du grain de pollen, le noyau végétatif migre dans le tube pollinique suivi par les noyaux spermatiques dans 90 % des cas, mais seulement dans 50% des cas pour le mutant *kaku4*. De plus, le pollen du mutant *kaku4* est moins compétitif pour la fertilisation que le pollen sauvage (Goto *et al.*, 2020). Ceci suggère que la forme du noyau végétatif des grains de pollen affecterait la migration ordonnée des différents noyaux du pollen dans le tube pollinique. D'autres anomalies du développement reproductif sont observées comme une légère diminution de la taille des siliques ainsi que du nombre de graines (**Figure 1.21E**) (Goto *et al.*, 2020). Le mutant d'insertion *mkaku41* du maïs montre aussi une diminution de la viabilité du pollen (McKenna *et al.*, 2021).

Plus récemment, l'équipe d'Yangnan Gu a décrit le proxitome de KAKU4 qui suggère une proximité avec les protéines CRWN, PNET-10 et -11, des protéines du NPC basket (NUP82, NUP136, NUP50a, NUP50c), des protéines associées au nucléosome (famille *Nucleosome Assembly Protein*, NAP), des protéines liant les histones et l'ADN (Facteurs de transcription et *DNA modification reader*) (**Figure 1.21F**) (Tang *et al.*, 2021). Ceci suggère que comme PNET2, KAKU4 pourrait interagir avec la chromatine.

I.3.c.iii) Les connexions entre le nucléosquelette et l'enveloppe nucléaire

Chez les animaux, les Lamines A/B et C interagissent avec des protéines de la membrane nucléaire interne regroupées sous le terme générique de *Lamin-Binding Proteins* et assurent l'enrichissement des lamines à la périphérie nucléaire (Wilson and Foisner, 2010).

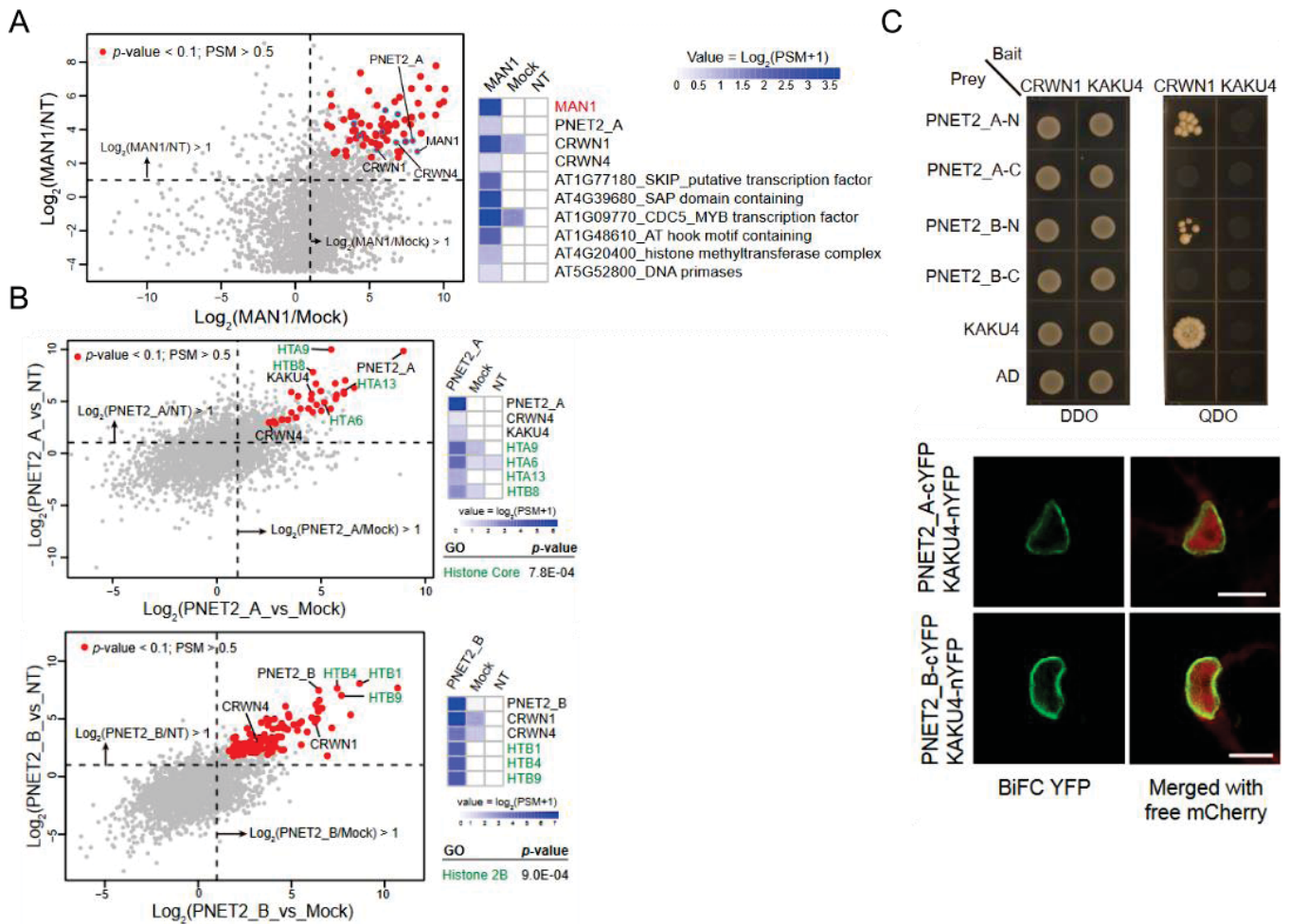


Figure 1.22 : Connexions entre le nucléosquelette et l'enveloppe nucléaire (Tang *et al.*, 2021)

A-B. Résultats du proxitome réalisé par protéomique par marquage de proximité de MAN1 (**A**) et PNET2_A et B (**B**). **C.** Interactions de PNET2_A et B avec CRWN1 en double-hybride (région N-ter ou Cter des PNET) (en haut) et avec KAKU4 en *bimolecular fluorescence complementation* (BiFC) (en bas).

Ces protéines incluent les protéines SUN, les protéines LBR et les protéines LEM, interagissant également avec la chromatine répressive. Les lamines sont également connectées avec le NPC notamment via la protéine NUP153 localisée dans le nuclear basket (Smythe, Jenkins and Hutchison, 2000).

Chez les plantes, mis à part la protéine MAN1, il n'existe pas d'orthologues aux protéines LBR et LEM. La protéine MAN1 a été étudiée par l'équipe d'Yannan Gu qui a réalisé le proximateur de MAN1 *via* la protéomique par marquage de proximité et montré que MAN1 est proche de PNET2, des protéines CRWN1 et CRWN4 ainsi que d'autres protéines associées à la chromatine comme des facteurs de transcriptions ou des modificateurs d'histones (**Figure 1.22A**) (Tang *et al.*, 2021). Ceci suggère que malgré l'absence de domaine LEM, MAN1 est retrouvée proche de constituants du nucléosquelette et ou de la chromatine, suggérant des interactions possibles avec ces facteurs. L'étude récente des protéines PNET2 présentes dans l'INM a permis de montrer la présence des protéines du nucléosquelette dans leur proximateur (**Figure 1.22B**). L'interaction des protéines PNET2_A et PNET2_B avec le nucléosquelette a ensuite été démontrée *via* leur interaction physique en double-hybride (Y2H) avec CRWN1 et *via* leur interaction en *Bimolecular Fluorescence Complementation* (BiFC) avec KAKU4 (**Figure 1.22C**) (Tang *et al.*, 2021). PNET2A et B ne sont pas les seules protéines à interagir avec les CRWN puisque les protéines SUN1 et SUN2 ont été montrées d'interagir par *fluorescence resonance energy transfer* (FRET) avec la protéine CRWN1 et elles participeraient également à son ancrage à la périphérie nucléaire (Graumann, 2014). La coexpression de PNET2_A ou PNET2_B avec CRWN1 ou KAKU4 induit des déformations de la membrane nucléaire avec un effet synergique entre KAKU4 et PNET2_B. L'INM ne serait donc pas la seule à interagir avec le nucléosquelette puisque les protéines CRWN1, CRWN3 et CRWN4 sont retrouvées dans le proximateur de NUP82, un composant du nuclear basket (Tang, Huang and Gu, 2020), suggérant des interactions possibles. Ainsi certains composants du NPC, les protéines SUN, MAN1 et PNET2 participeraient à l'ancrage du nucléosquelette à la périphérie nucléaire. Finalement, tout comme les lamines, les protéines CRWN constitueraient un réel hub d'interaction à la périphérie nucléaire (Wilson and Foisner, 2010).

Pour conclure, chez les eucaryotes, la périphérie nucléaire joue principalement les mêmes fonctions entre les plantes et les animaux alors que celle-ci diffère grandement ; en effet, seuls le complexe LINC, les protéines MAN1 et PNET2 et le pore nucléaire sont des composants conservés à la périphérie nucléaire entre les animaux et les plantes (**Figure 1.23**). Ce qui suggère que ces deux règnes ont des protéines différentes qui ont acquises les mêmes fonctions.

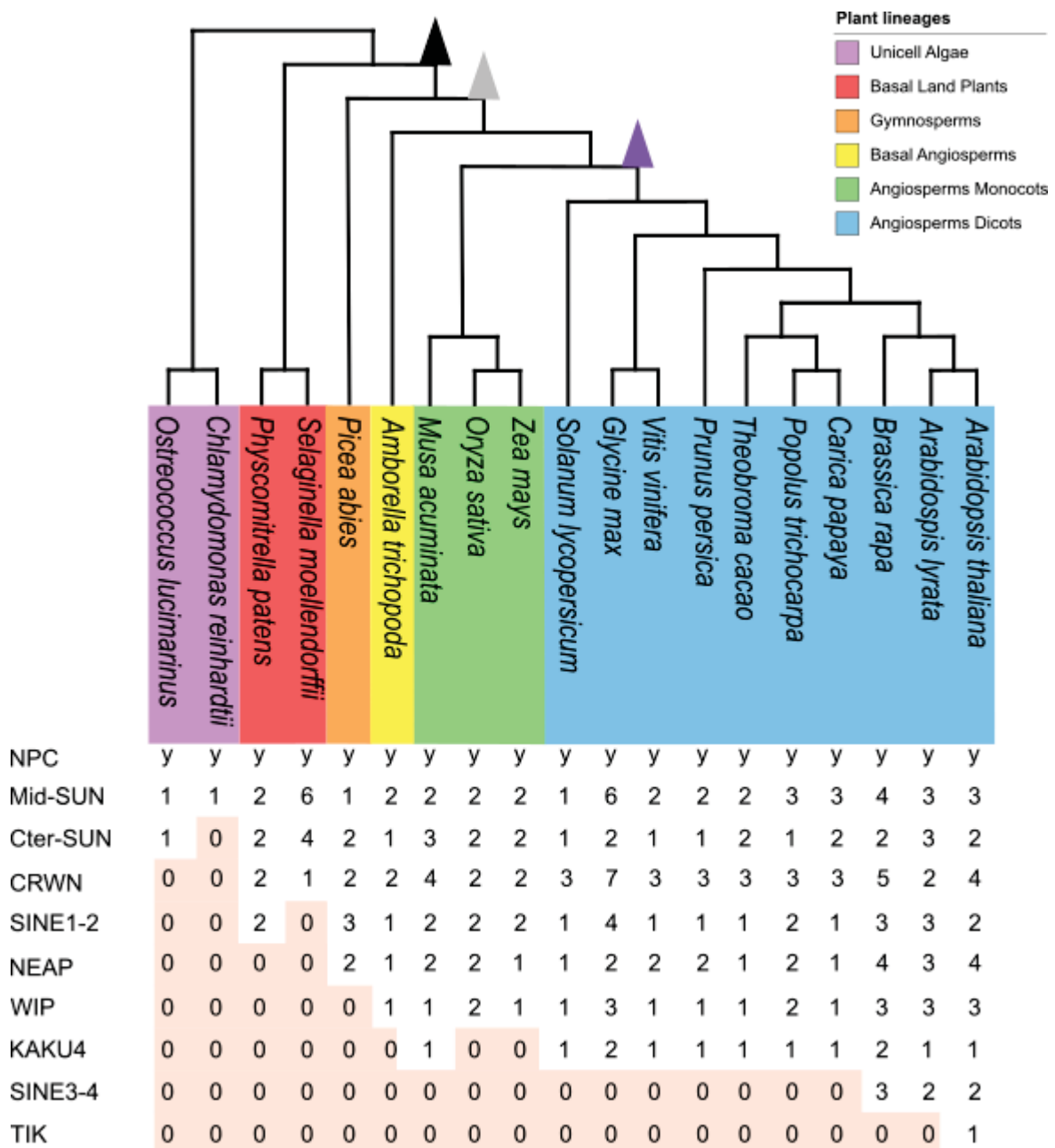


Figure 1.23 : Conservation des protéines de la périphérie nucléaire au sein du règne végétal (adapté de Poulet, Probst, *et al.*, 2017).

Arbre composé d'espèces représentatives du règne végétal avec en dessous un tableau récapitulatif de la présence des pores nucléaires ainsi que le nombre de protéines trouvées dans ces espèces pour les autres composants de la périphérie nucléaire ; parme : algues, rouge : mousses, orange : gymnospermes, jaune : angiospermes primitifs, vert : angiospermes monocotylédones, bleu : angiospermes dicotylédones ; y : yes.

Le meilleur exemple est le cas du nucléosquelette qui est composé des lamines A à C chez les mammifères et des protéines CRWN1 à 4 ainsi que de la protéine KAKU4 chez les plantes. La localisation des lamines à la périphérie nucléaire est principalement médiée par leurs interactions avec les protéines LBR, LEM, SUN, NUP153 ; les protéines CRWN et KAKU4 interagissant quant à elles avec les protéines PNET2 et SUN et son proche de MAN1.

II- Rôle des protéines de la périphérie nucléaire dans l'organisation de la chromatine

J'ai montré dans la première partie de mon introduction d'une part que le noyau permet d'organiser l'information génétique localement mais également spatialement grâce à des domaines de chromatine, cette organisation ayant des conséquences sur la régulation de l'expression des gènes (cf paragraphe I.2). J'ai également synthétisé nos connaissances sur la périphérie nucléaire et ses différents composants qui permettent de définir un des sous-domaines nucléaires pouvant également participer à la régulation de l'expression du génome (cf paragraphe I.3). Dans cette seconde partie, je souhaiterais confronter ces deux premiers thèmes afin de démontrer qu'effectivement la périphérie du noyau est un acteur important de la structuration de l'organisation de la chromatine chez les plantes.

L'une des premières observations de l'organisation de la chromatine remonte sans doute aux analyses cytologiques faites par Emile Heitz (1928) qui ont été depuis confirmées par d'autres techniques comme la *Fluorescence in situ Hybridization* FISH, Hi-C ou DamID. L'ensemble de ces méthodes suggèrent toutes un enrichissement des régions hétérochromatiques, répressives pour la transcription, au voisinage de la périphérie nucléaire (Pickersgill *et al.*, 2006; Pontvianne *et al.*, 2016; Bi *et al.*, 2017; Poulet, Duc, *et al.*, 2017b).

Je vais donc développer ci-dessous, les interactions entre la chromatine et la périphérie nucléaire, en me concentrant sur les interactions connues entre les pores nucléaires, le nucléosquelette et la chromatine.

II.3) Rôle du NPC dans la régulation de l'activité transcriptionnelle

II.3.a) Premières évidences

De longue date les chercheurs s'intéressent au rôle du NPC dans la régulation transcriptionnelle. Ainsi, chez la drosophile, les nucléoporines NUP50, NUP62 ou encore NUP98 sont capables d'interagir directement avec la chromatine et le plus souvent ces

interactions impliquent des régions transcriptionnellement actives qui présentent des marques épigénétiques activatrices comme H3K4me2 et H4K16Ac (Capelson *et al.*, 2010; Kalverda *et al.*, 2010).

L'interaction du NPC n'est cependant pas uniquement avec des régions transcriptionnellement actives comme le suggère d'autres études chez la levure avec la nucléoporine NUP157/NUP170 qui interagit avec des séquences hétérochromatiques et est requise pour le silencing des gènes sub-téломériques (Van de Vosse *et al.*, 2013).

La proximité du NPC a été associée à la possibilité d'une induction rapide de la transcription de certains gènes. Ce mécanisme initialement proposé par Gunter Blobel (Blobel, 1985), appelé le *gene gating*, propose que les gènes devant être activement transcrits seraient recrutés à la périphérie nucléaire par le biais d'interactions avec les nucléoporines. Il est aujourd'hui connu que certaines nucléoporines interagissent avec des régulateurs de transcription tels que les complexes SAGA, TREX2 et les protéases SUMO et participent ainsi à la régulation de la transcription (Parry 2015). De nombreux facteurs de transcription participeraient également au rapprochement des gènes cibles des NPC, favorisant ainsi leur expression (Brickner *et al.*, 2019). A titre d'exemple, chez la levure, des gènes inductibles tels que l'*HEXOKINASE 1* (*HXK1*) et la *GALACTOKINASE 1* (*GAL1*) interagissent avec les composants du NPC pour leur activation transcriptionnelle (**Figures 1.24A-B**) (Grossman, Medalia and Zwerger, 2012; Parry, 2015).

[II.3.b\) Complexes SAGA et TREX-2](#)

Les interactions entre les composants du NPC et la région promotrice de chaque gène seraient des événements physiologiques précoces nécessaires à l'induction de la transcription (Schmid *et al.*, 2006). L'un des candidats possibles pour la médiation de l'interaction entre la région promotrice et le NPC pourrait être le complexe SAGA. SAGA est un complexe coactivateur transcriptionnel qui régule de nombreux processus cellulaires par la coordination de multiples modifications post-traductionnelles des histones. Les composants du complexe SAGA de la levure lient physiquement la chromatine avec le nuclear basket du NPC pendant l'activation de la transcription (Rodriguez-Navarro *et al.* 2004). Le NPC permet également le recrutement et le maintien de la machinerie transcriptionnelle à la périphérie nucléaire *via* le complexe TREX-2. Le complexe TREX-2 est un complexe d'exportation d'ARNm chez la levure (Kohler et Hurt 2007), et est attaché à la face interne du NPC par la nucléoporine NUP1 et NUP60, deux orthologues potentiels de NUP136 chez les plantes

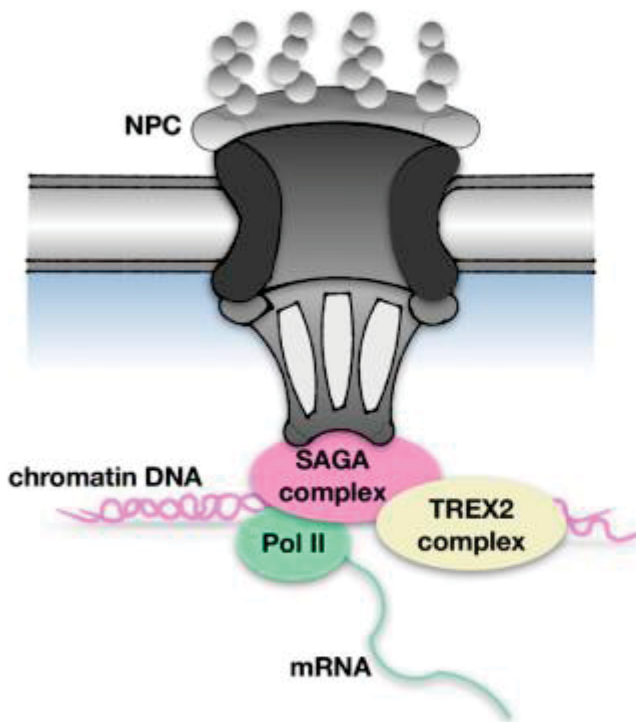
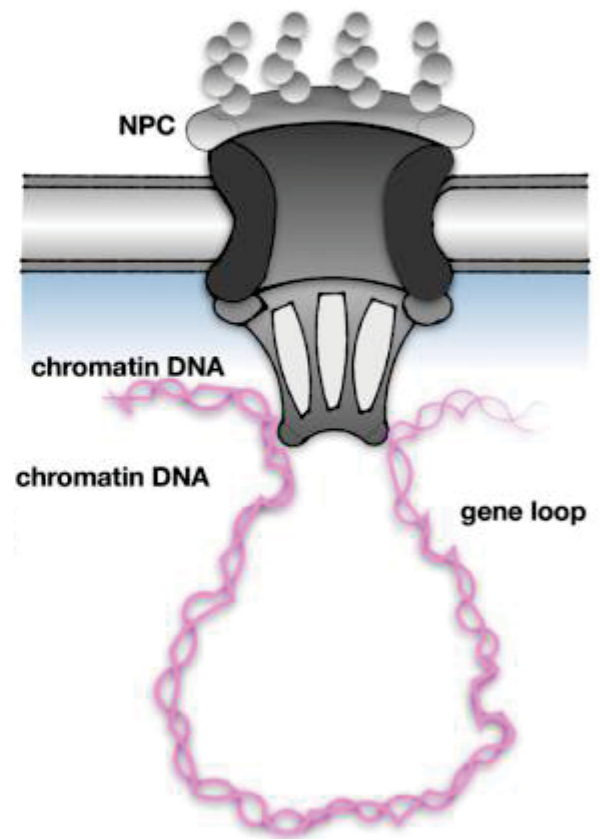
A**B**

Figure 1.24 : Représentations de la régulation transcriptionnelle médiée par le pore nucléaire (adapté de Tamura, 2020)

A. Le NPC fournit une plateforme pour l'assemblage de la machinerie transcriptionnelle, comprenant les complexes SAGA, TREX-2, et l'ARN polymérase II (Pol II). **B.** Le nuclear basket du NPC est impliquée dans le maintien des boucles de chromatine, qui permettent la réactivation rapide des gènes en réponse à des stimuli externes comme les stress biotiques et abiotiques.

(Tamura and Hara-Nishimura, 2013). TREX-2 et le complexe SAGA partagent tous deux un composant commun, la protéine SUCROSE SYNTHASE 1 (SUS1), suggérant un couplage fonctionnel de ces deux complexes pour l'exportation des ARNm du noyau (Rodríguez-Navarro *et al.*, 2004). SUS1 est d'abord recruté dans la région de la séquence d'activation en amont *via* SAGA lors de l'induction du gène. Lors du processus d'activation de la transcription, les complexes SAGA, TREX-2 ainsi que l'ARN polymérase II phosphorylée favorisent le recrutement de SUS1 (**Figures 1.24A**) (García-Oliver, García-Molinero and Rodríguez-Navarro, 2012).

II.3.c) Boucles de chromatine

Les NPCs modulent l'expression des gènes en favorisant l'assemblage et le maintien des boucles de gènes permettant ainsi la juxtaposition des régions promotrices et terminatrices des gènes (Tan-Wong, Wijayatilake and Proudfoot, 2009). Chez la levure, les boucles de gènes sont liées à la mémoire transcriptionnelle, de sorte que les gènes peuvent être réactivés rapidement s'ils ont été récemment exprimés. En effet, MLP1, un composant du nuclear basket de la levure, est important pour le maintien des boucles géniques après la répression des gènes et pour la réactivation rapide de leur transcription comme les HXK1 et GAL10 inductibles par le galactose (Tan-Wong, Wijayatilake and Proudfoot, 2009). Cette réactivation rapide s'effectue grâce à la formation et au maintien de boucles de gènes au niveau du NPC (**Figure 1.24B**).

Les interactions entre le NPC et la chromatine sont donc de deux ordres. Le premier consiste en l'interaction du NPC avec des régions répressives notamment les régions péri-centromériques. Le second est d'interagir avec des régions chromatiniennes actives transcriptionnellement *via* les complexes SAGA et TREX-2 ou la formation de boucles de gènes (**Figures 1.24A-B**). Le positionnement de la chromatine au niveau des NPC reste encore flou et nécessite d'autres études afin de mieux comprendre les mécanismes impliqués.

II.4) Rôle du NPC dans la régulation transcriptionnelle chez *Arabidopsis thaliana*

Plusieurs évidences expérimentales montrent que la périphérie du noyau et la proximité de la chromatine vers le NPC modifieraient la transcription chez *Arabidopsis*. Ainsi, une étude réalisée par Smith *et al.* montre que le rapprochement d'un gène rapporteur avec les composants des NPC affecte son niveau d'expression (Smith *et al.*, 2015). La démonstration formelle d'un changement de position du gène rapporteur vers le NPC lors de son activation transcriptionnelle reste à démontrer.

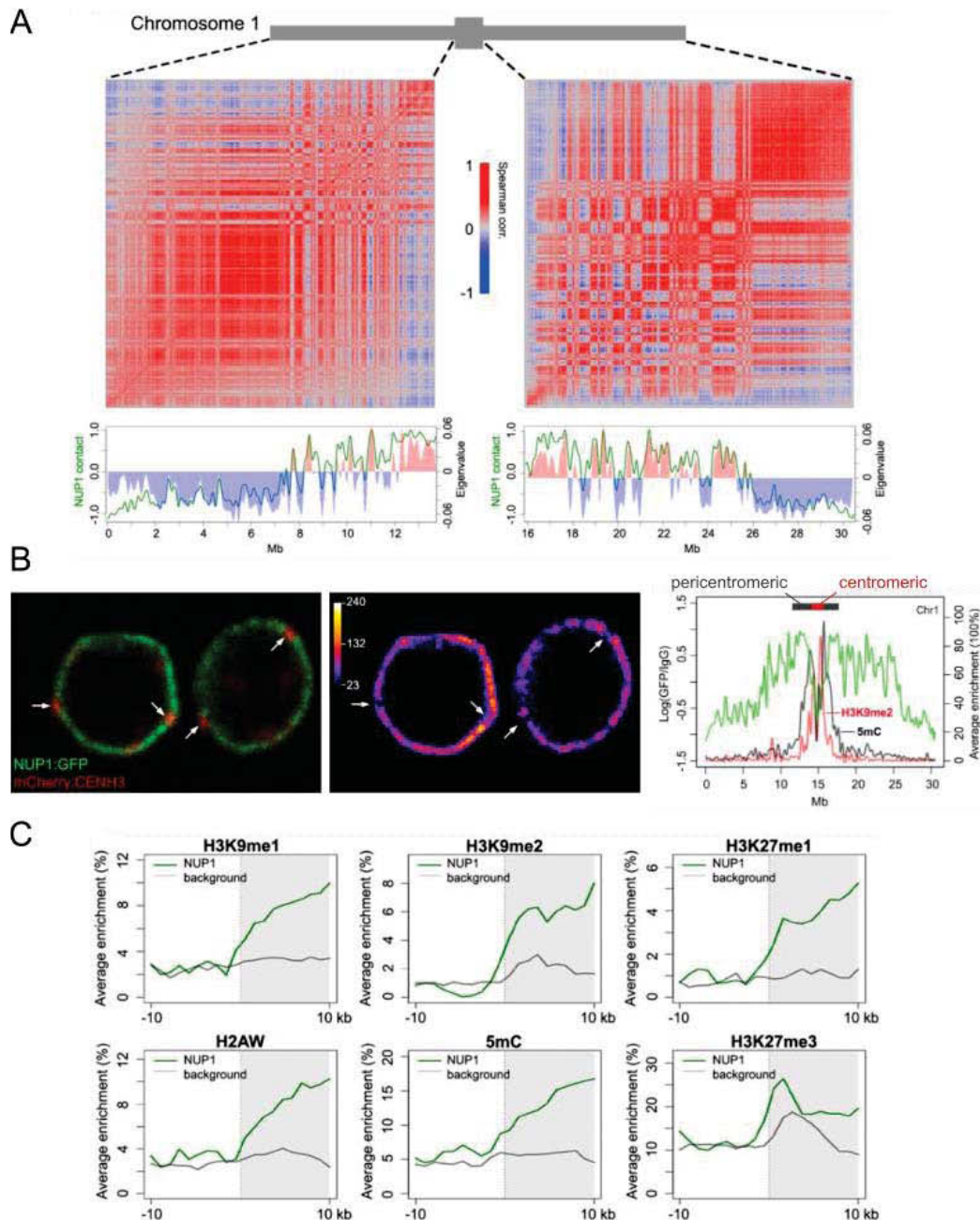


Figure 1.25 : Le pore nucléaire s’associe avec l’hétérochromatine péri-centromérique (Bi *et al.*, 2017)

A. Comparaison des données de Hi-C et de RE-ChIP de NUP136:GFP (ici indiqué sous le nom NUP1 :GFP). Matrices Hi-C des bras gauche et droit du chromosome 1, présentées sous forme de cartes de contacts entre régions de la chromatine (**en haut**). Régions d’interactions définies par Hi-C ont été regroupées en compartiment A (en bleu) et compartiment B (en rouge) et représentées simultanément avec le profil de liaison de NUP1/NUP136 avec la chromatine obtenue par RE-ChIP (vert) (**en bas**). **B.** Image confocale de noyaux racinaires coexprimant NUP136 (vert) et CEN-H3 (rouge) (**à gauche**) et l’intensité de pixel (**au centre**) avec les flèches indiquant les centromères. Graphique représentant la distribution des signaux de NUP136 obtenue par RE-ChIP ainsi que des marques épigénétiques sur le chromosome 1 (**à droite**). **C.** Répartitions des marques épigénétiques enrichies dans les régions immunoprécipitées à partir de NUP136:GFP (fond grisé) *versus* régions non enrichies (fond blanc).

L'une des premières études visant à déterminer quelles sont les régions chromatiniennes interagissant avec le pore nucléaire a été réalisée sur une lignée mutante *nup136* complétée avec un transgène exprimant NUP1/NUP136:GFP en utilisant une technique originale nommée restriction enzyme mediated ChIP (RE-ChIP) (Bi *et al.*, 2017). Cette technique consiste en une digestion ménagée de la chromatine suivie d'une immunoprécipitation avec un anticorps anti-GFP afin d'identifier les régions de chromatine liées par NUP136 donc proche du nuclear basket (voir paragraphe I.3.b). Dans cet article, les données de RE-ChIP sont confrontées à celles obtenue par Hi-C. Les données de Hi-C montrent que les régions télomériques, sub-télomériques ainsi que les bras chromosomiques éloignés des péri-centromères correspondent essentiellement au compartiment A (chromatine active) alors que les régions péri-centromériques et les bras chromosomiques proches des péri-centromères correspondent au compartiment B (régions répressives). Les données de RE-ChIP obtenues avec NUP136 montrent que les interactions NUP136/chromatine correspondent essentiellement à celles coïncidant avec le compartiment B (**Figure 1.25A**) (Bi *et al.*, 2017). Toutefois, la coexpression de NUP136 avec CENH3, un marqueur centromérique, montre que NUP136 ne colocalise pas avec CENH3 excluant les régions centromériques des régions de chromatine interagissant avec NUP136 (**Figure 1.25B**). Les régions chromosomiques interagissant avec NUP136 ne sont pas particulièrement enrichies en A/T (une marque des séquences répétées de type éléments transposables) mais sont enrichies en cytosines méthylées et en marques répressives comme H3K9me1 et me2 et H3K27me3 ou encore avec le variant d'histone péri-centromérique H2A.W, renforçant les observations précédentes qui montrent que NUP136 interagit préférentiellement avec le compartiment B (**Figure 1.25C**). Plusieurs tissus ont été analysés et ce travail montre que NUP136 représente 10 à 20 % du génome en fonction des tissus. Ces séquences correspondent à des gènes, des pseudogènes et des éléments transposables. Les gènes liant NUP136 ont des niveaux d'expression plus faible que les gènes n'interagissant pas avec NUP136 alors que le niveau d'expression des éléments transposables n'est pas significativement différent selon qu'ils interagissent ou non avec NUP136 (**Figures 1.26A-B**). La conclusion générale de cette étude est donc que la périphérie nucléaire, ici définie comme l'ensemble des régions de chromatine liant NUP136, correspond à un ensemble de séquences globalement associées aux marques répressives pour la transcription. Cette conclusion peut paraître surprenante considérant que la proximité du NPC est en général plutôt associée à des régions transcriptionnellement actives ou qui peuvent être rapidement activées en cas de stress. Toutefois, même si les auteurs se sont concentrés sur l'étude des régions réprimées liant NUP136, ils indiquent, dans leur discussion et les tableaux supplémentaires, l'existence de

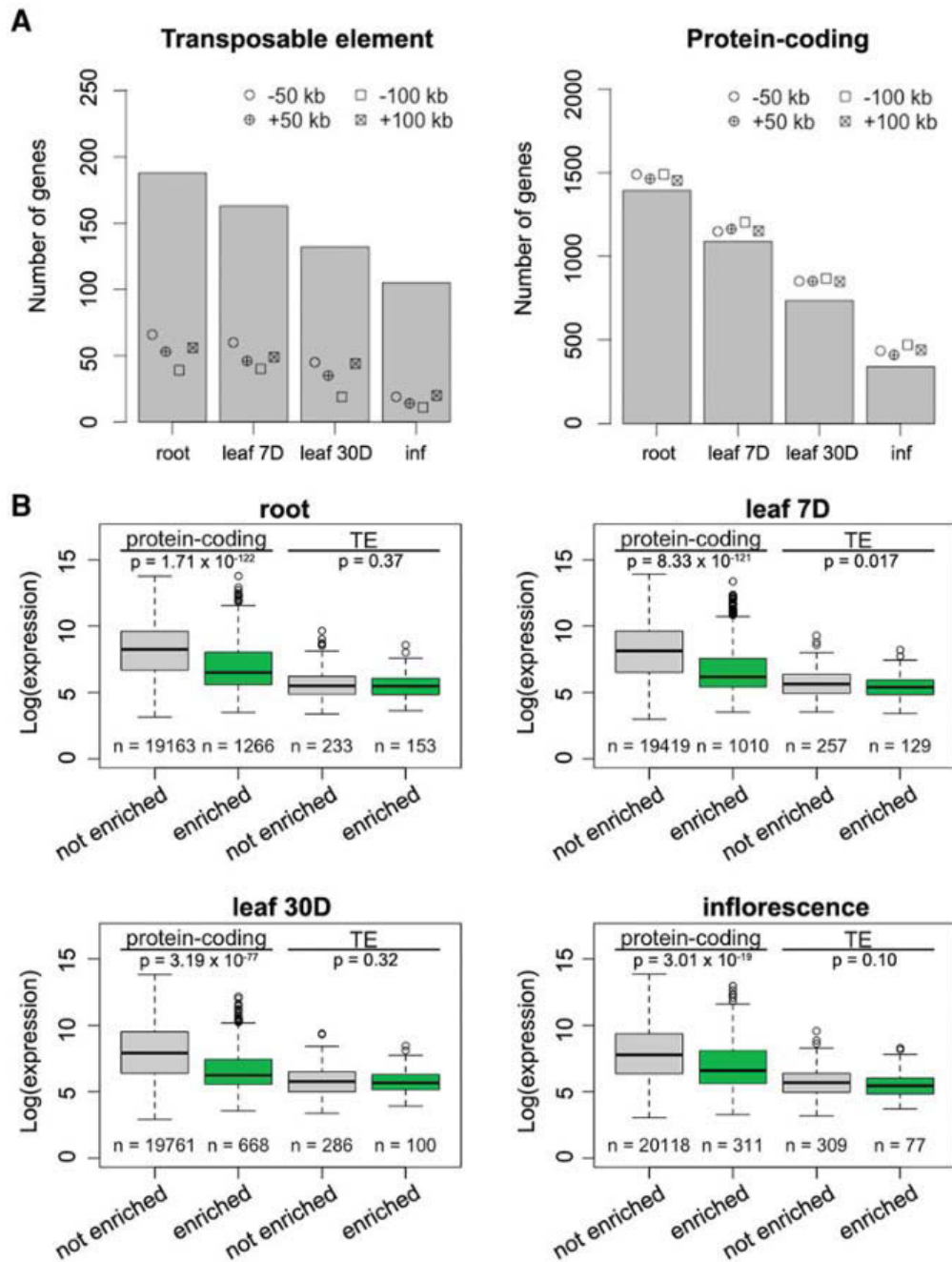


Figure 1.26 : Enrichissement de la périphérie en gènes exprimés (Bi *et al.*, 2017)

A. Nombre d'éléments transposables (à gauche) et de gènes codant pour des protéines (à droite) enrichis dans différents tissus. Pour chaque colonne, le nombre observé de gènes est significativement différent ($P < 0,001$) par rapport à 3000 gènes, pris aléatoirement, présentant le même profil de distribution que celui des gènes enrichis en NUP136 ; inf : inflorescence. **B.** Comparaison des niveaux d'expression des gènes enrichis (vert) ou non enrichis (gris) dans le RE-ChIP de NUP136:GFP dans différents tissus ; p-value indiquant résultats du test U de Mann-Whitney.

gènes fortement exprimés dans les gènes liant NUP136. Malheureusement, aucune proportion gènes actifs/gènes réprimés n'est donné dans cet article.

En conclusion, la proximité de NUP136 ne peut pas se restreindre à une seule classe de séquences qui pourrait être soit active soit inactive transcriptionnellement même si cet article nous indique que les séquences enrichies en marques répressives et donc plutôt inactives transcriptionnellement sont enrichies au voisinage de NUP136. Ceci montre qu'il ne faut pas tirer de conclusions trop hâtives sur l'aspect répressif ou activateur de la périphérie nucléaire.

Chez les plantes, les composants du complexe SAGA sont également impliqués dans l'induction de l'expression des gènes en réponse à des signaux développementaux et environnementaux comme le froid (Tamura, 2020). Le complexe TREX-2 est ancré au NPC *via* l'interaction de NUP136 et de la protéine THP1. La perte des composants du complexe TREX-2 a révélé diverses fonctions : un effet protecteur contre le silencing des gènes par des mécanismes épigénétiques, le maintien de la stabilité de l'hétérochromatine et l'exportation d'ARNm (Lu *et al.*, 2009). Cependant, peu de choses sont connues sur le rôle que joue le complexe TREX-2-NPC des plantes dans le recrutement de gènes spécifiques pour réguler leur expression.

Chez *A. thaliana*, il existe également des évidences concernant la formation de boucles de gènes positionnées au voisinage du NPC et liées à la mémoire transcriptionnelle en réponse au stress. Une étude sur l'induction de la mémoire transcriptionnelle lors du choc thermique, montre que le NPC pourrait jouer un rôle important dans la réactivation des gènes régulés par la chaleur médiée par la présence d'éléments de réponse au choc thermique dans leur promoteur (Liu *et al.*, 2018). Cela suggère que chez les plantes une formation de boucles de gènes dynamiques est impliquée dans la mémoire transcriptionnelle lors de changements environnementaux comme chez les levures. Il serait intéressant de savoir si les composants des NPC des plantes jouent des rôles importants dans le maintien des boucles génétiques pour réguler l'expression des gènes et la mémoire transcriptionnelle. D'autres études sont encore nécessaires pour identifier les composants qui contrôlent la régulation dynamique de l'organisation du génome dans le noyau des plantes.

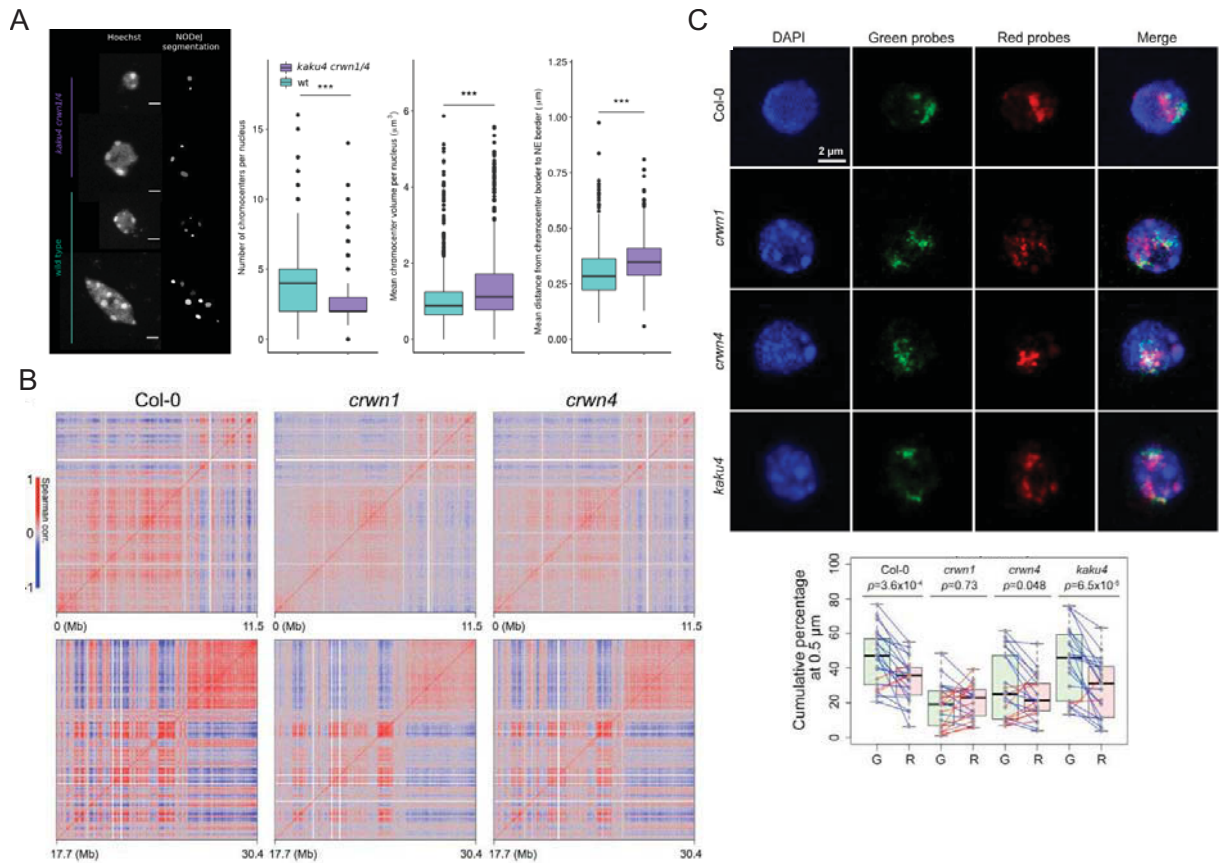


Figure 1.27 : Les protéines CRWN impactent faiblement l'organisation spatiale de la chromatine

A. Projection en Z de noyaux de l'épiderme du type sauvage et du mutant *kaku4 crwn1 crwn4*, coloré au Hoechst ainsi que le résultat de segmentation des chromocentres par NODEJ (**à gauche**) (Dubos *et al.*, 2022). Résultats de l'analyse NODEJ avec le nombre de chromocentres, le volume moyen des chromocentres par noyau et la distance moyenne entre la bordure du chromocentre et l'enveloppe nucléaire par noyau (**à droite**), test U de Mann-Whitney p-value : * ≤ 0.05 , *** ≤ 0.001 (Dubos *et al.*, 2022). **B.** Cartes de contacts Hi-C des bras gauche et droit du chromosome 1 en contexte sauvage et mutant *crwn1* et *crwn4* (Hu *et al.*, 2019). **C.** Images confocales représentatives de noyaux 2C de type sauvage et mutant *crwn1*, *crwn4* et *kaku4* avec des sondes s'hybridant sur des régions du chromosome 1 enrichies à la périphérie nucléaire (vert) ou non (rouge) d'après le RE-ChIP de NUP136:GFP. Quantification du nombre de sondes vertes et rouges présentes à moins de 0,5 μm de la périphérie nucléaire. Les sondes rouges et vertes de chaque noyau sont reliées par une ligne bleue (nb sondes vertes > nb sondes rouges) ou rouge (nb sondes vertes < nb sondes rouges) dans cette région du noyau (Hu *et al.*, 2019).

II.5) Rôle du nucléosquelette

Chez les animaux, la chromatine localisée à la périphérie nucléaire et interagissant avec les lamines forme les LAD (paragraphe I.2.c.iii). Les LAD sont des régions d'hétérochromatine de faible densité en gènes, peu transcrites et enrichies en marques épigénétiques répressives (Pickersgill *et al.*, 2006; Guelen *et al.*, 2008a; van Steensel and Belmont, 2017). Une partie de ces régions chromatiniennes localisées à la périphérie nucléaire montre des changements dynamiques de leur position par rapport à la lamina et ce en fonction du type cellulaire, du stade de différenciation ou de stimuli externes.

Chez la plante, des marquages de l'ADN par un agent intercalant ont permis de montrer que les mutants du nucléosquelette impactent le nombre et la condensation des chromocentres visualisés par microscopie. En effet, les mutants *crwn* et *kaku4*, notamment *crwn1 crwn2*, *crwn1 crwn4* et *crwn1 crwn4 kaku4*, possèdent moins de chromocentres et d'un volume supérieur par rapport au génotype sauvage (**Figure 1.27A**) (Dittmer *et al.*, 2007; Wang, Dittmer and Richards, 2013; Poulet, Duc, *et al.*, 2017b; Dubos *et al.*, 2022; Sakamoto *et al.*, 2022). A l'inverse, le simple mutant *crwn4* provoque la décondensation et la fragmentation des chromocentres (Wang, Dittmer and Richards, 2013). Précédemment, nous avons vu que les NPC interagissent avec les régions péri-centromériques et tendent à exclure les régions centromériques (**Figure 1.27**), une étude récente a montré chez *Arabidopsis* que la condensine II, inhibant le regroupement des chromocentres, est capable d'interagir avec la protéine SUN2 un composant du complexe LINC ancré dans l'INM (paragraphe I.3.a.i) (Sakamoto *et al.*, 2022). Des interactions entre les protéines SUN et CENH3 ont également été montrée en Y2H et Co-IP, ce qui suggère que le complexe LINC *via* les protéines SUN est capable d'interagir avec la chromatine et notamment avec les centromères (Sakamoto *et al.*, 2022). De plus, les protéines CRWN1 et CENH3 interagissent et co-localisent à la périphérie nucléaire. Le complexe LINC et les protéines CRWN sont donc capables d'interagir avec les centromères (CENH3) à la périphérie nucléaire. Dans cette étude, un nouveau mécanisme de distribution des centromères dans les noyaux a été découvert et peut se subdiviser en deux étapes : (i) dispersion des centromères dépendante de leurs interactions avec le complexe Condensine II-LINC pendant la mitose (anaphase à télophase) ; et (ii) stabilisation des centromères dépendante de leurs interactions avec le nucléosquelette (CRWN) tout au long de l'interphase. Ce processus en deux étapes contribue à la distribution spatiale des centromères chez *A. thaliana* et confirme que les centromères sont en « répulsion » les uns par rapport aux autres (Arpòn *et al.*, 2021; Sakamoto *et al.*, 2022).

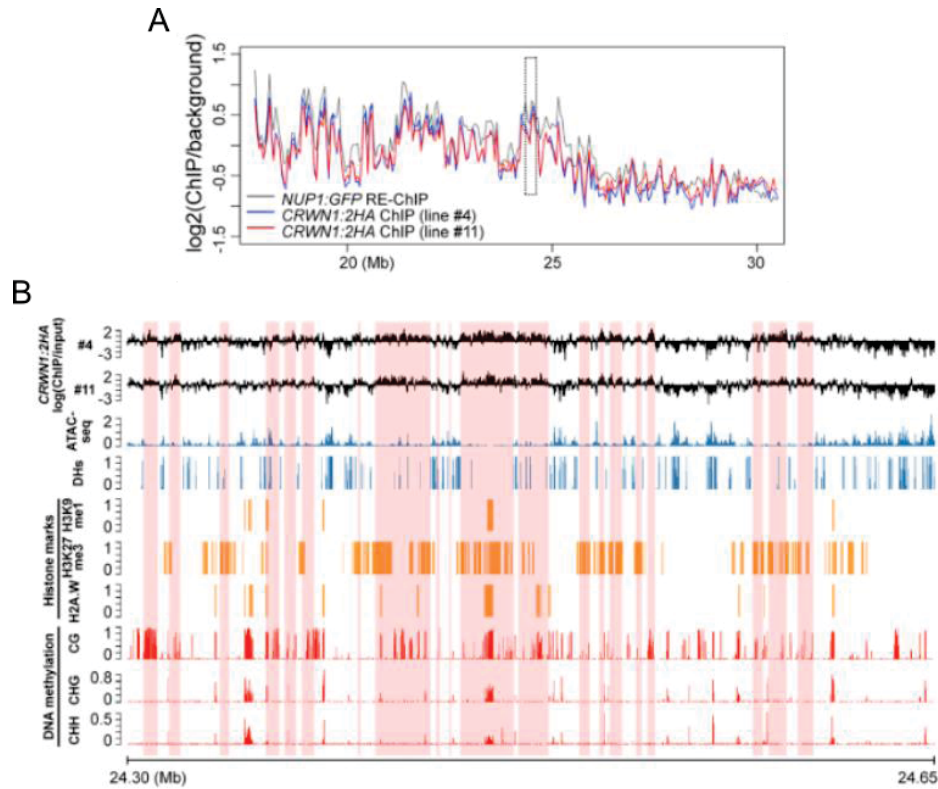


Figure 1.28 : CRWN1 s’associe à la chromatine et forme les pLAD (Hu *et al.*, 2019)

A. Comparaison des profils d’association de la chromatine avec la périphérie nucléaire étudiées par deux approches différentes, RE-ChIP de NUP136:GFP (en noir, identique à la courbe verte de la **Figure 1.25A**) et ChIP-seq de CRWN1 :HA (rouge et bleu). Ce graphique montre le bras droit du chromosome 1, une région de 350 kb mise en évidence par une boîte en pointillés est également représentée en **B**. **B.** Comparaisons des signaux de ChIP de CRWN1:HA avec différentes caractéristiques structurales et marques épigénétiques. Les régions surlignées en rose représentent les LAD d’Arabidopsis ; DH : *DNase I hypersensitive sites*.

De façon assez surprenante, l'analyse Hi-C des mutants *crwn1*, *crwn4* et *crwn1 crwn4* montre qu'il n'y a pas de changement important dans l'organisation de la chromatine (Hu *et al.*, 2019; Sakamoto *et al.*, 2022). Seule une compartimentalisation plus faible de la chromatine est observée avec une diminution des interactions intra-compartiment et une augmentation de celles entre différents compartiments comme entre chromosomes ou entre compartiment A/B d'un même chromosome ainsi que celles entre les péri-centromères des différents chromosomes (**Figure 1.27B**). Cela indique que la chromatine est moins ordonnée au sein des mutants *crwn*, cependant, ils ne présentent pas d'altérations massives de l'accessibilité de la chromatine ou du transcriptome (Choi, Strickler and Richards, 2019; Hu *et al.*, 2019; Sakamoto *et al.*, 2022). Toutefois des expériences de FISH montrent que les mutants *crwn1* et, dans une moindre mesure, *crwn4* affectent la position des régions de chromatine proches de la périphérie nucléaire définies lors de l'analyse RE-ChIP réalisée avec NUP136-GFP (Bi *et al.*, 2017) (**Figure 1.27C**) (Hu *et al.*, 2019). Ces changements ne seraient pas dus à des changements de morphologie nucléaire puisqu'aucune modification de la sphéricité des noyaux n'a été observée dans cette expérience. Rappelons ici que les noyaux utilisés sont issus d'un tri et se limitent à des noyaux diploïdes, or l'effet des mutants *crwn* s'observe surtout sur les noyaux allongés qui sont le plus souvent des noyaux polyplloïdes. Les deux voies de méthylation de l'ADN non-CG, c'est-à-dire CHG et CHH impliquant les protéines DOMAINS REARRANGED METHYLASE 1 (DRM1), DRM2, CHROMOMETHYLASE 2 (CMT2) et CMT3, seraient requises pour le recrutement de la chromatine à la périphérie nucléaire indépendamment de la présence de la marque répressive H3K9me2 (Hu *et al.*, 2019).

En plus de l'étude de NUP136, la même équipe s'est intéressée aux régions de chromatine interagissant avec CRWN1 en réalisant un ChIP-seq avec pCRWN1:CRWN1:2HA. CRWN1 s'associe donc directement à des régions chromatiniennes de taille variable, appelées *plant Lamina-Associated Domains* (pLAD), qui chevauchent largement avec les domaines de chromatine enrichis à la périphérie nucléaire identifiés via le RE-ChIP de NUP136 (**Figure 1.28A**). Ces pLAD ne sont pas enrichis en marques chromatiniennes répressives, contrairement à ce qui a été montré dans l'article de Bi *et al.*, 2017, et ne possèdent pas de profil particulier au niveau des caractéristiques structurelles et des marques épigénétiques à part qu'elle serait peu accessible (**Figure 1.28B**). En formant des contacts directs avec la chromatine peu accessible, les protéines CRWN1 médient l'arrimage de la chromatine à la périphérie nucléaire chez les plantes, tout comme les lamines des animaux qui forment les LAD.

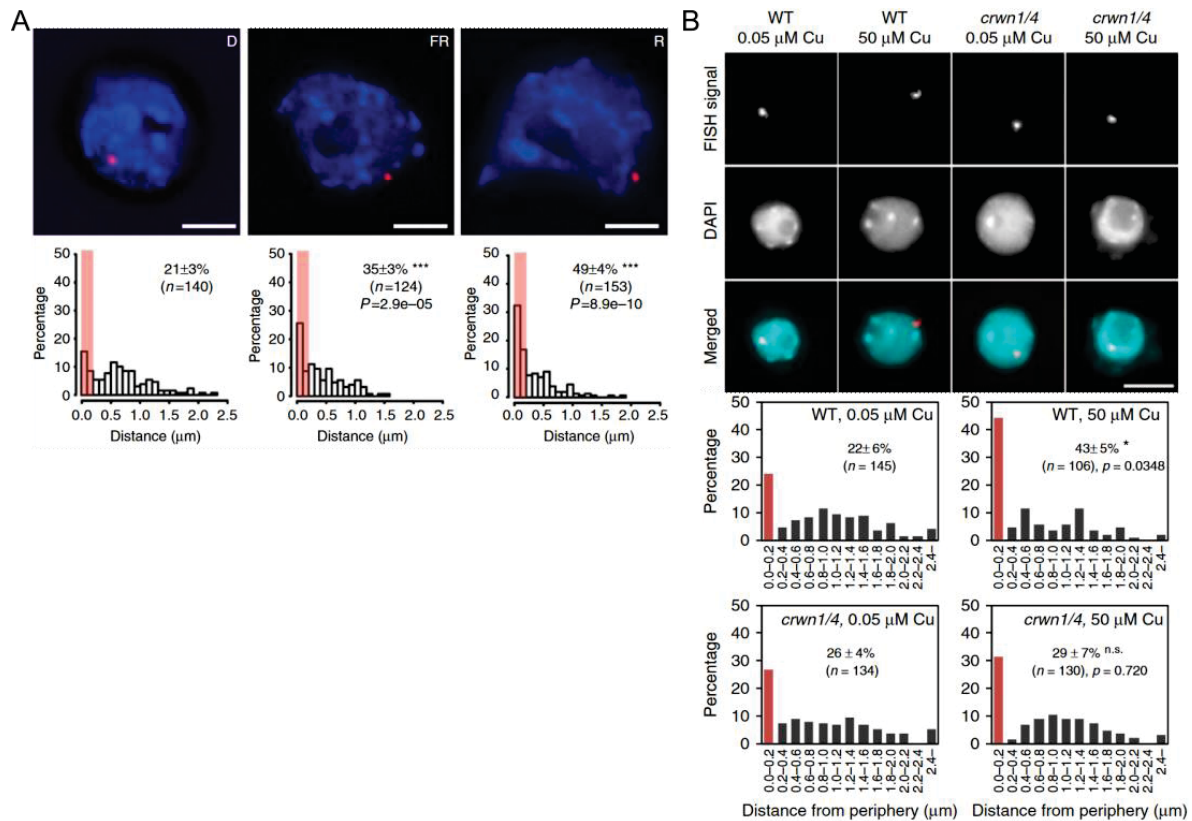


Figure 1.29 : Relocalisation des gènes *CAB* et *CA* à la périphérie nucléaire lors de leur induction transcriptionnelle

A. Visualisation du locus *CAB* par FISH dans des noyaux des plantules cultivées à l'obscurité (D), en rouge lointain (FR) et en lumière rouge (R), échelle : 2 μm (en haut). Histogrammes représentant la distance entre le locus *CAB* et le bord nucléaire avec les zones rouges représentant la zone de la périphérie nucléaire (en bas) (C.-M. Feng *et al.*, 2014). **B.** Visualisation du locus *CA* dans le noyau par FISH en présence ou en absence de cuivre chez le sauvage et en contexte mutant *crwn1 crwn4*, échelle : 5 μm (en haut). Histogrammes représentant la distance entre le locus *CA* et le bord nucléaire, les barres rouges de l'histogramme représentent la périphérie nucléaire (en bas) (Sakamoto *et al.*, 2020).

II.6) Relocalisation des gènes au sein du noyau en fonction de l'environnement

Au cours des deux sections précédentes, la mobilité de certaines régions de la chromatine a été observée vis-à-vis soit du NPC soit du nucléosquelette. Deux exemples de changements de position associés à des modulations transcriptionnelles sont documentés chez les plantes : les gènes inductibles par la lumière comme les gènes *CHLOROPHYLL A/B-BINDING (CAB)*, *RUBISCO SMALL SUBUNIT (RBCS)*, *GENOMES UNCOUPLED 5 (GUN5)* et *PLASTOCYANIN (PC)* et les gènes associés au cuivre (*COPPER-ASSOCIATED, CA*).

L'induction de l'expression des gènes de réponse à la lumière *CAB*, *RBCS*, *GUN5* et *PC* se traduit effectivement par un changement de localisation vers la périphérie nucléaire. En effet, à l'obscurité, le locus *CAB* se localise dans 21% des cas à la périphérie nucléaire alors qu'il y est retrouvé respectivement, dans 35% et 49% des cas lors d'une exposition à la lumière rouge lointain et rouge ; cela est corrélé avec l'augmentation de leur expression dans ces deux derniers contextes (**Figure 1.29A**) (C.-M. Feng *et al.*, 2014). L'activation ainsi que le repositionnement des gènes *CAB* sont régulés positivement par les photorécepteurs à la lumière rouge lointain et rouge, respectivement *PHYA* et *PHYB* et négativement par les gènes répresseurs de la signalisation des *PHY*, *CONSTITUTIVELY PHOTOMORPHOGENIC (COP1)*, *DE-ETIOLATED 1 (DET1)* et les *PHYTOCHROME INTERACTING FACTORS (PIF)* (C.-M. Feng *et al.*, 2014).

Dans le cas des gènes *CA* formant un cluster de 11 gènes sur le chromosome 5, la relocalisation nucléaire a pu être associée aux protéines CRWN grâce à la technique du ChIL. CRWN1 interagit avec le locus des gènes *CA* en condition d'excès de cuivre notamment avec les gènes *CA3*, *CA5* et *CA8* alors qu'aucune interaction n'est montrée avec ce locus en condition contrôle (Sakamoto *et al.*, 2020). Cette interaction semble liée à la perte de la marque répressive H3K27me3 et au gain en H4ac, marque épigénétique activatrice, au niveau de ces gènes en contexte mutant *crwn1 crwn4* lors d'un excès de cuivre. Ceci suggère que la décondensation de la chromatine au niveau du locus *CA* est nécessaire à leur induction dans des conditions d'excès de cuivre. En condition sauvage, dans 22% des cas le locus *CA* se localise à la périphérie nucléaire alors qu'en excès de cuivre, ce locus est retrouvé à la périphérie nucléaire dans 43% des cas mais cette position ne change pas dans le mutant *crwn1 crwn4* quel que soit les conditions (**Figure 1.29B**) (Sakamoto *et al.*, 2020). Ainsi les protéines CRWN, en interagissant avec le locus *CA*, permettent son changement de localisation et ainsi d'induire leur expression *via* une décompaction de la chromatine lors d'un excès de cuivre dans le milieu.

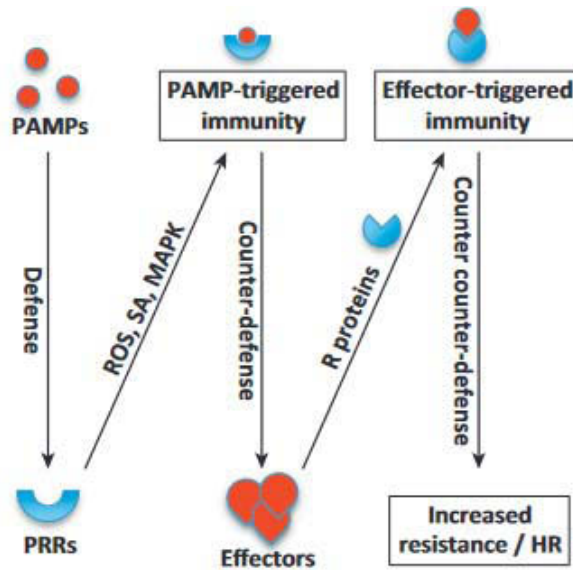


Figure 1.30 : Représentation schématique de la réponse immunitaire innée chez la plante (Incarbone and Dunoyer, 2013)

Les *Pathogen-Associated Molecular Patterns* (PAMP) sont des composés/motifs antigéniques issus des pathogènes qui sont reconnus par des récepteurs transmembranaires (*pattern recognition receptors*, PRR). Les PAMP déclenchent l'immunité induite par les motifs (*pattern/PAMP-triggered immunity*, PTI). Les agents pathogènes sécrètent des protéines (effecteurs) qui interfèrent avec la PTI, entraînant une « susceptibilité » accrue de la plante au pathogène. Certains effecteurs sont détectés par des protéines de résistance aux maladies codées par des gènes de résistance (gènes R), ce qui entraîne le déclenchement de l'immunité induite par les effecteurs (*effector-triggered immunity*, ETI) accélérant et amplifiant la PTI. Il en résulte une résistance à la maladie mais aussi une réponse d'hypersensibilité (HR) au site de l'infection qui se traduit par une mort cellulaire. ROS : espèces réactives de l'oxygène, SA : acide salicylique, MAPK : protéines kinases activées par des agents mitogènes.

Pour conclure, la périphérie nucléaire joue un rôle important dans l'organisation de la chromatine avec des rôles distincts pour le pore nucléaire et le nucléosquelette. Le NPC, notamment les composants du nuclear pore basket comme NUP136, serait capable à la fois d'interagir avec la chromatine active transcriptionnellement *via* les complexes SAGA et TREX-2 mais également avec des régions répressives mais exclue les interactions avec les centromères comme le montre les travaux de Bi *et al.* Le nucléosquelette jouerait quant à lui un rôle dans le maintien des chromocentres (et des centromères) à la périphérie nucléaire mais participerait également à la compartimentation (compartiment B) des domaines de chromatine enrichis en régions réprimées du génome (gènes et éléments transposables), formant les pLAD. Enfin, notons que l'organisation des gènes dans l'espace nucléaire est dynamique et peut varier notamment au cours du développement, de la différenciation cellulaire ou en réponse à des stress ou des stimuli externes comme la lumière. Certaines protéines présentes à la périphérie nucléaire, comme CRWN1, participent à ce mécanisme de régulation de la transcription en modifiant l'organisation de la chromatine et son environnement spatial en les repositionnant à la périphérie nucléaire afin de moduler leur expression.

III- Rôle de la périphérie nucléaire dans la réponse aux stress chez la plante

Il existe deux grands types de stress, le stress biotique correspondant à un stress provoqué par un autre être vivant et le stress abiotique induit quant à lui par un changement de l'environnement. Chez la plante, le stress biotique peut-être déclenché par un organisme vivant pathogène comme un champignon, une bactérie, un virus, un animal ou une autre plante alors que le second est déclenché par des changements d'environnement comme la température, la lumière ou encore les nutriments. Au cours des dernières années, de nombreuses études se sont intéressées chez les plantes aux rôles des protéines de la périphérie nucléaire dans la réponse à ces deux types de stress.

III.3) Stress biotique

III.3.a) La réponse aux pathogènes chez les plantes

Contrairement aux animaux, les plantes sont des organismes sessiles qui ne possèdent pas de cellules immunitaires spécialisées ni de système immunitaire adaptatif. Les cellules végétales s'appuient sur leur système immunitaire inné pour se défendre contre les agents pathogènes impliquant notamment la reprogrammation du transcriptome afin de réguler les réponses de défense. Les plantes possèdent deux types d'immunité innée, l'immunité induite par des motifs (*Pattern-Triggered Immunity*, PTI) et celle provoquée par les effecteurs

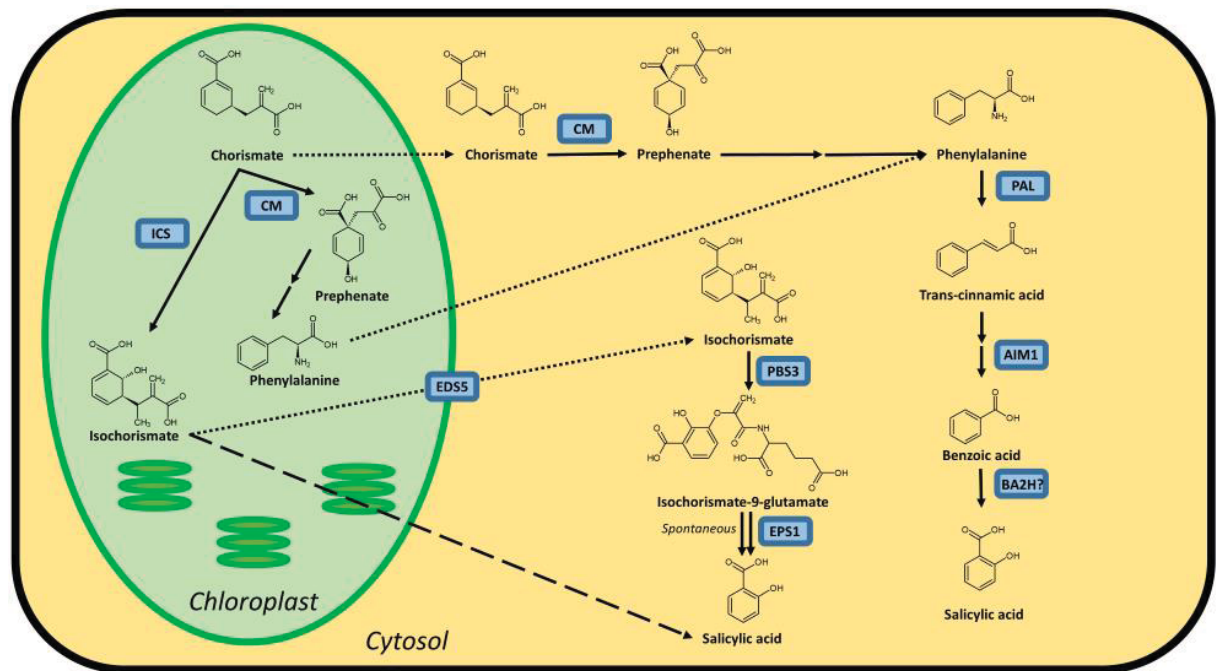


Figure 1.31 : Voie de biosynthèse de l'acide salicylique. (Lefevre, Bauters and Gheysen, 2020)

Le SA est synthétisé par deux voies distinctes l'une dans le chloroplaste (vert) et l'autre cytoplasmique (jaune). Les abréviations des différentes enzymes impliquées sont : ICS : ISOCHORISMATE SYNTHASE ; CM : CHORISMATE MUTASE ; PAL : PHENYLALANINE AMMONIA-LYASE ; AIM1 : ABNORMAL INFLORESCENCE MERISTEM1 ; BA2H : BENZOIC ACID 2-HYDROXYLASE ; ED5 : ENHANCED DISEASE SUSCEPTIBILITY 5 ; PBS3, *avrPphB* SUSCEPTIBLE 3 ; EPS1 : ENHANCED PSEUDOMONAS SUSCEPTIBILITY 1. Notons que chez *Arabidopsis*, l'ISOCHORISMATE SYNTHASE 1 (ICS1) porte de nombreux autres acronymes, notamment SALICYLIC ACID INDUCTION DEFICIENT 2 (**SID2**) qui sera l'un des mutants utilisés dans mon travail de thèse.

(*Effector-Triggered Immunity*, ETI) (**Figure 1.30**). Par analogie au système immunitaire animal, la PTI repose sur la reconnaissance d'antigènes décrits dans la littérature comme des motifs moléculaires associés aux pathogènes (*Pathogen-Associated Molecular Patterns*, PAMP) via des récepteurs de reconnaissance de motifs (*Pattern Recognition Receptors*, PRR). Ces récepteurs rappellent les *Toll-like Receptors* des animaux présents sur la membrane plasmique qui lors de leur activation entraînent une cascade de signalisation impliquant les *Mitogen-Activated Protein Kinase* (MAPK), le calcium Ca^{2+} et les espèces réactives de l'oxygène (*Reactive Oxygen Species*, ROS) pour activer l'immunité de la plante (**Figure 1.30**). Afin de contrer ces barrières immunitaires, certains pathogènes ont développé des protéines effectrices, qui sont sécrétées par le pathogène dans les cellules végétales pour détourner certaines voies métaboliques au profit du pathogène. Pour contrer cela, les plantes ont développé un groupe de récepteurs immunitaires intracellulaires appartenant à la superfamille TIR-NB-LRR (*Toll interleukin-1 receptor* (TIR)-*nucleotide-binding site* (NB)-*leucine-rich repeat* (LRR)) pour reconnaître les effecteurs et ainsi activer une réponse immunitaire forte et robuste, *via* des « gènes de résistance » (**Figure 1.30**). Les gènes de la plante codant les protéines ciblées dans ces voies métaboliques sont des « gènes de sensibilité ». Comme on le voit, il existe une réelle coévolution des interactions hôte/pathogène chacun s'adaptant à l'évolution de l'autre.

La PTI et l'ETI ont une signalisation précoce distincte mais partagent de nombreux régulateurs immunitaires communs en aval pour coordonner la reprogrammation transcriptionnelle vers l'activation de la défense. Cette réponse aux pathogènes passe principalement par les hormones végétales comme l'acide salicylique (SA), l'acide jasmonique et l'éthylène (Verma, Ravindran and Kumar, 2016). Ces hormones sont des intégrateurs des réponses physiologiques régulant à la fois la croissance des plantes et la réponse aux pathogènes, compromis inévitable et crucial pour la survie des plantes qui sont exposées à divers agents pathogènes dans l'environnement.

III.3.b) Voie de réponse médiée par l'Acide Salicylique

L'acide salicylique (SA) est sans doute l'une des hormones la plus étudiée dans la réponse au stress biotique. Elle joue un rôle essentiel dans les réponses de résistance locale (au niveau de la cellule) et systémique (à distance via une diffusion par le phloème) aux pathogènes. Le SA est synthétisé à partir du chorismate à la fois *via* la voie de l'ISOCHORISMATE SYNTHASE (ICS) et celle de la PHENYLALANINE AMMONIA-LYASE (PAL) (**Figure 1.31**) (Lefevre, Bauters and Gheysen, 2020). Chez *Arabidopsis*, la voie ICS est majoritaire et explique jusqu'à

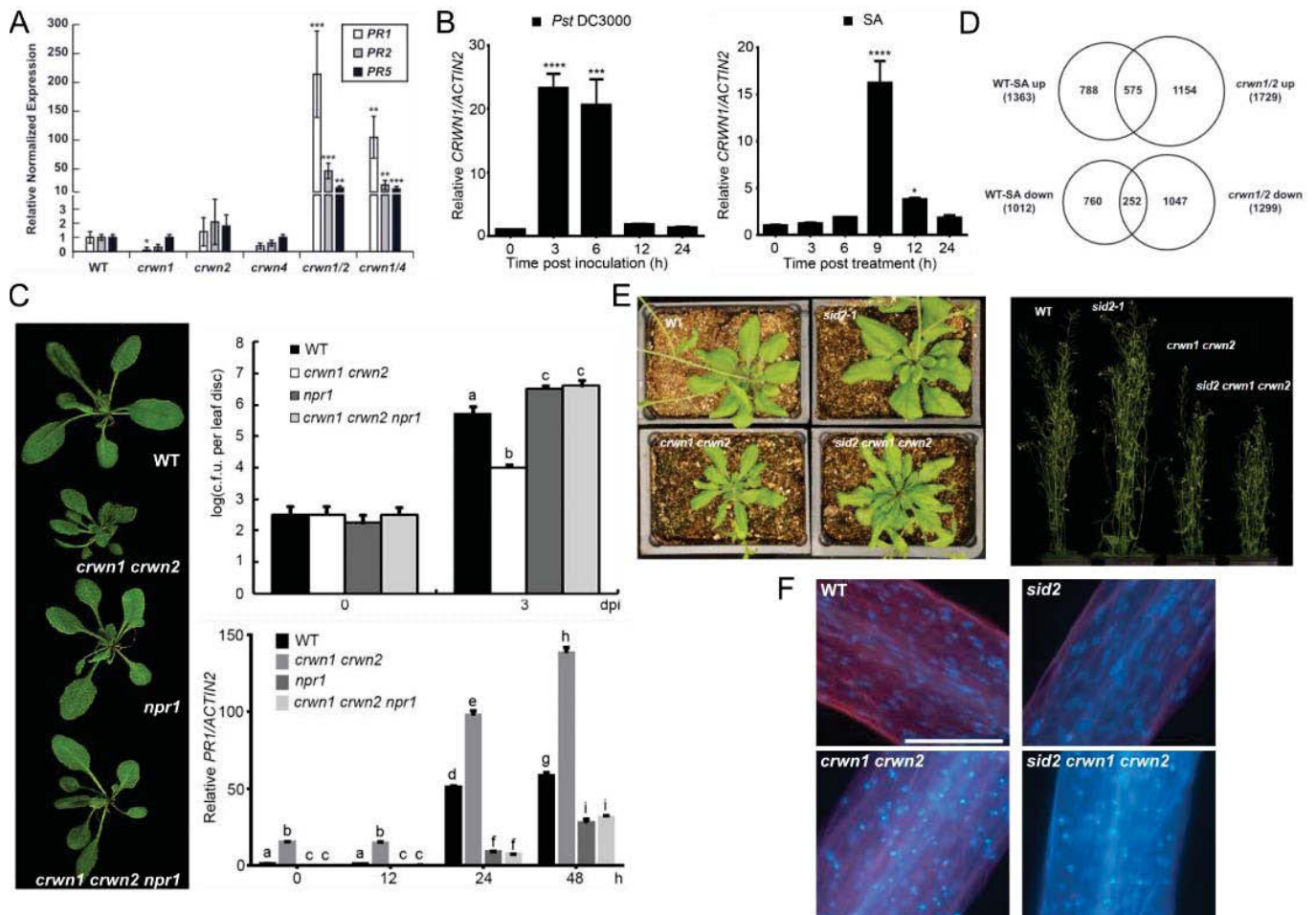


Figure 1.32 : Impact de la voie de l'acide salicylique dans les mutants *crwn*

A. Données de RT-qPCR montrant l'expression de trois gènes *PR* dans des plantes âgées de 27 jours des différents mutants *crwn*. (Choi, Strickler and Richards, 2019). **B.** Analyse de l'expression de *CRWN1* dans des feuilles de plantes sauvages de 4 semaines après traitement par *Pseudomonas* ou SA (Guo *et al.*, 2017). **C.** Photos des mutants impliquants *crwn1 crwn2* et *npr1* âgés de 3 semaines et étude de la sensibilité des différents mutants ainsi que de l'expression de *PR1* après traitement par *Pseudomonas* (Guo *et al.*, 2017). **D.** Comparaison des gènes dérégulés dans le mutant *crwn1 crwn2* déterminés par RNA-seq avec les gènes dérégulés 3h après un traitement au SA 1mM, déterminé par microarray (Fold Change>2 et p-value<0.05) (Choi, Strickler and Richards, 2019). **E-F.** Photos des phénotypes développementaux (**E**) et de la morphologie nucléaire dans les cellules des filets des étamines (**F**) dans les mutants impliquant *crwn1 crwn2* et *sid2* (Choi, Strickler and Richards, 2019).

90% de la production de SA (Chen *et al.*, 2009). Une attaque de pathogènes provoque l'augmentation du SA dans la plante et induit l'expression de nombreux gènes de défense dont le gène *PATHOGENESIS RELATED 1 (PR1)*, marqueur de l'induction de la défense. Cette réponse peut également être reproduite en appliquant un traitement exogène apportant du SA. La transcription du gène *PR1* est régulée positivement par NONEXPRESSOR OF PR GENES 1 (NPR1), un régulateur de la résistance aux pathogènes, contrôlant l'expression des gènes de défense en s'associant soit avec les répresseurs de transcription WRKY DNA BINDING PROTEIN (WRKY) soit avec les activateurs de transcription TGACG SEQUENCE-SPECIFIC BINDING PROTEIN (TGA). NPR1 est un facteur de transcription central et important dans la régulation de la réponse induite par le SA. Notons que *PR1* est également régulé négativement par NAC TRANSCRIPTION FACTOR-LIKE 9 (NTL9), un facteur de transcription participant à de multiples réponses immunitaires chez la plante, en s'associant avec la protéine SUPPRESSOR OF NPR1-1 (SNI1).

III.3.c) Implication des protéines CRWN dans la réponse au stress biotique

III.3.c.i) Les protéines CRWN affectent la voie de l'acide salicylique

L'analyse transcriptomique des mutants des 4 gènes *CRWN* montre que les gènes impliqués dans la réponse aux stimuli ainsi que ceux impliqués dans la réponse aux pathogènes sont sur-exprimés (Guo *et al.*, 2017; Choi, Strickler and Richards, 2019), le mutant *crwn4* ayant le plus d'effet parmi les simples mutants. Les doubles mutants *crwn* ont une résistance accrue à l'infection par la bactérie *Pseudomonas syringae* pv *maculicola* (Guo *et al.*, 2017) ou *Pseudomonas syringae* pv *tomato* (Choi, Strickler and Richards, 2019) vraisemblablement due à une augmentation de l'expression basale du gène *PR1* mais également des gènes *PR2* et *PR5*, gènes cibles de la voie de l'acide salicylique, dans les doubles mutants *crwn1 crwn2* et *crwn1 crwn4* (**Figure 1.32A**). Cette étude a également montré que CRWN1 est régulée aussi bien au niveau transcriptionnel que post-transcriptionnel par une infection par *Pseudomonas* ou un traitement au SA ; en effet, dans ces deux cas l'expression de CRWN1 est induite rapidement sur une courte durée (**Figure 1.32B**), mais une dégradation rapide de la protéine CRWN1 est observée par le protéasome via sa région Cter (751-1132 AA) (Guo *et al.*, 2017). Ceci rappelle les observations faites sur CRWN3 qui est associé à des corps nucléaires enrichis en enzymes du protéasome (Zhao *et al.*, 2016). Ceci suggère que les protéines CRWN sont sans doute régulées de façon post-traductionnelle mais également qu'elles pourraient aussi réguler les protéines avec lesquelles elles interagissent via le protéasome. De plus, CRWN1 est capable d'interagir avec NTL9 et SNI1, deux protéines inhibant l'expression de *PR1*, l'interaction entre

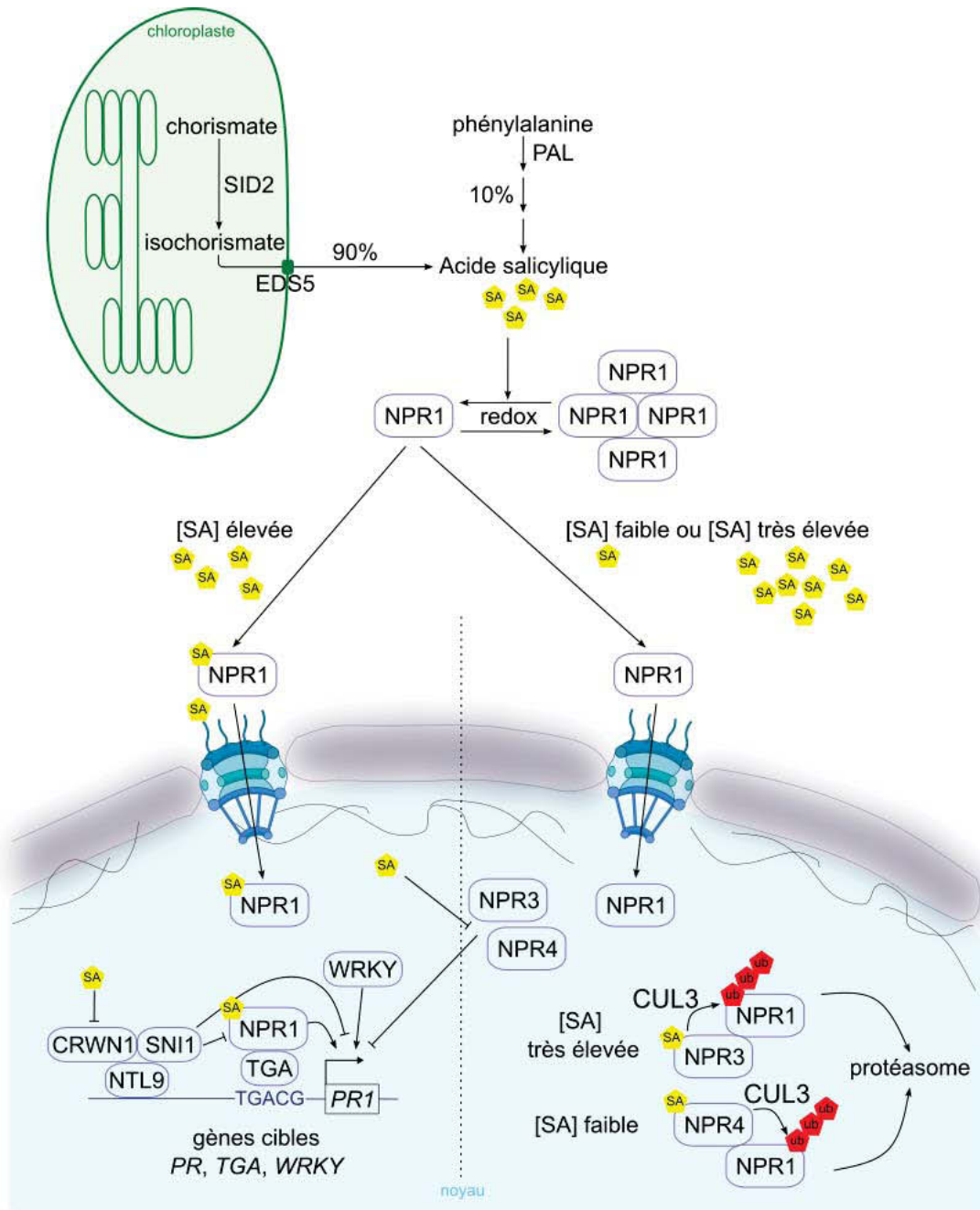


Figure 1.33 : Représentation schématique de la voie de l'acide salicylique chez la plante

L'acide salicylique est fortement produit dans la plante lors de l'infection par un pathogène, favorisant la forme monomérique de NPR1 permettant sa translocation dans le noyau où il active les gènes cibles *PR* en s'associant aux facteurs de transcription TGA ou WRKY (en haut). L'interaction entre le facteur de transcription NTL9 avec CRWN1 et SNI1 induit la fixation de ce complexe sur le promoteur de *PR1*. Ce complexe inhibe l'action de NPR1 et prévient l'activation du gène *PR1* (en bas à gauche). Quand la concentration en SA est faible, NPR1 est préférentiellement retrouvée sous forme de complexe et le peu de monomères présents et transloquant dans le noyau vont être dégradés par le protéasome suite à leur ubiquitinylation par CUL3 recrutée par NPR4 lors de son interaction avec NPR1 (en bas à droite). Lorsque la concentration en SA est très élevée, NPR1 est dégradée par le protéasome par le même mécanisme en faisant intervenir NPR3 au lieu de NPR4.

CRWN1 et NTL9 promeut la fixation de NTL9 sur le promoteur de *PR1* inhibant ainsi son expression (Guo *et al.*, 2017). CRWN1 régule également l'expression de *PR1* via des voies indépendantes de NTL9 comme celle de NPR1 (**Figure 1.33**).

Phénotypiquement les plantules *crwn1 crwn2 npr1* sont similaires à celles des simples mutants *npr1* et légèrement moins grandes que le sauvage, la mutation de *npr1* supprimant le phénotype de petite taille de *crwn1 crwn2* (**Figure 1.32C**). Ce triple mutant a une moins bonne résistance aux infections par *Pseudomonas* que le double mutant *crwn1 crwn2* ou le sauvage, dû fait de la faible expression de *PR1* dans ce mutant (**Figure 1.32C**). La résistance accrue des doubles mutants *crwn* contre les pathogènes dépend donc de NPR1 (Guo *et al.*, 2017).

Les doubles mutants *crwn1 crwn2* et *crwn1 crwn4* présentent un phénotype foliaire de mort cellulaire programmée avec une incidence plus élevée pour *crwn1 crwn2*, alors que ce phénotype de mort cellulaire n'est pas observé dans les simples mutants *crwn* (Choi, Strickler and Richards, 2019). Ces résultats sont en accord avec le niveau d'expression des gènes *PR* et la résistance aux pathogènes observés dans ces mutants. Le nombre de cellules mortes augmente avec le temps tout comme l'expression de *PR1* au sein des plantules. Les doubles mutants et notamment *crwn1 crwn2* possèdent un taux élevé de SA qui est corrélé à la forte expression du gène *SID2* impliqué dans la voie principale de synthèse du SA (**Figure 1.30**). Il existe un fort recoupement entre le transcriptome d'une plante traitée au SA et le transcriptome du double mutant *crwn1 crwn2* : 40 % des gènes surexprimés et 25% des gènes sous-exprimés (**Figure 1.32D**) (Choi, Strickler and Richards, 2019). Afin de déterminer si les phénotypes des mutants *crwn* sont liés à une accumulation forte du SA, la mutation *sid2* a été introduite dans le double mutant *crwn1 crwn2*. Comme pour NPR1 (Guo *et al.* 2017), la mutation de *sid2* supprime partiellement le phénotype de petite plante du mutant *crwn1 crwn2* suggérant que l'accumulation de SA est responsable en grande partie du phénomène de petite plante des plantes (**Figure 1.32E**) (Choi, Strickler and Richards, 2019). À l'inverse, l'analyse de la morphologie des noyaux montre que le phénotype petits noyaux observés dans le mutant dans le mutant *crwn1 crwn2* est également retrouvé dans le mutant *sid2 crwn1 crwn2* (**Figure 1.32F**). Cette expérience élégante d'interaction génétique permet de dissocier le phénotype de morphologie nucléaire (indépendant du SA) et le phénotype de petite plante (dépendant du SA) observés lors de la perte de fonction de CRWN1 et CRWN2. Les gènes sur-exprimés dans le mutant *crwn1 crwn2* sont également retrouvés dans les cibles de H3K27me3 (Mikulski *et al.*, 2019). Or, la diminution de cette marque dans le double mutant *crwn1 crwn2* conduit à la perte de la répression transcriptionnelle du gène *PR1* et des facteurs de transcription de la voie du

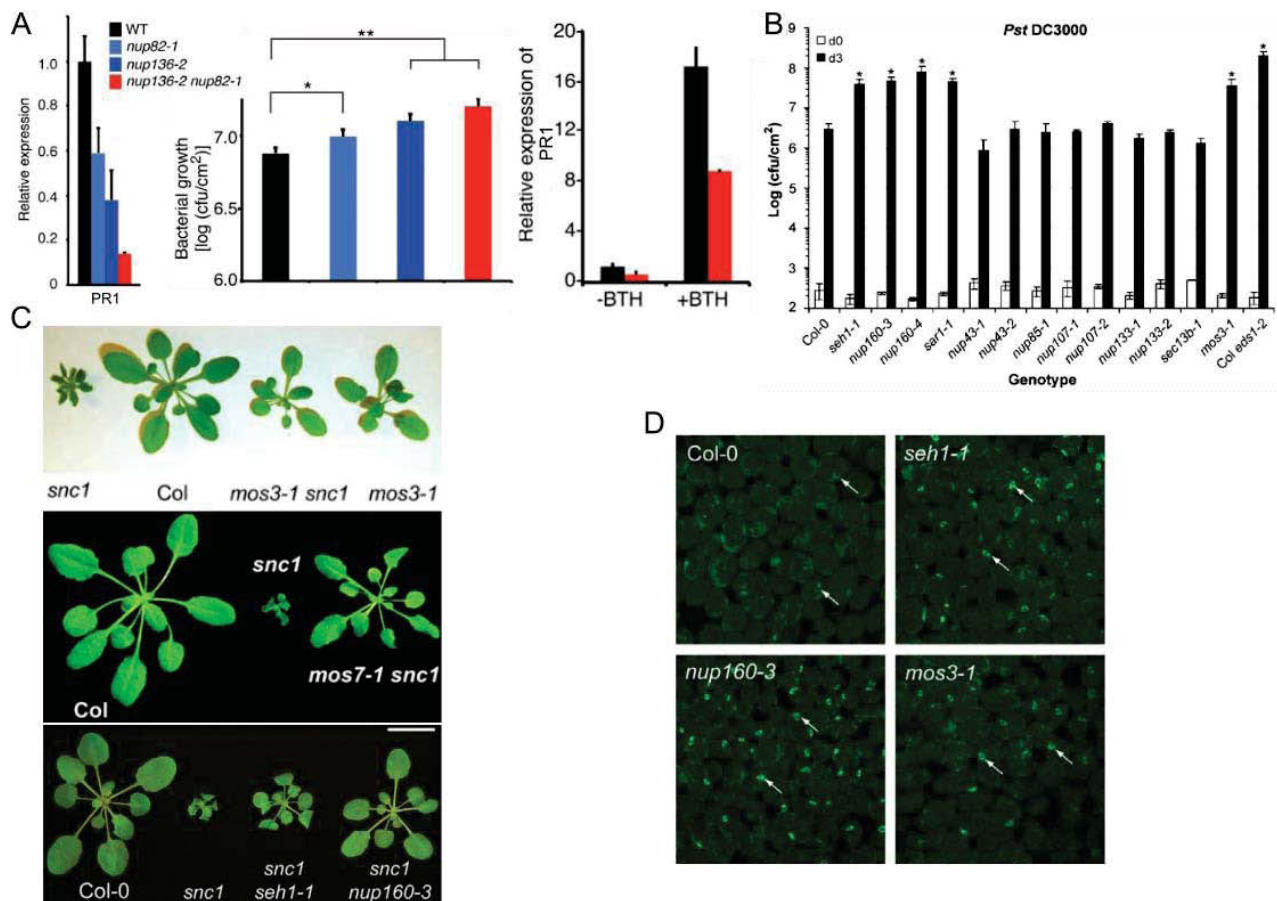


Figure 1.34 : Impact de la voie de l'acide salicylique dans les mutants *nup*

A. Expression de *PR1* et analyse de la sensibilité à *Pseudomonas* des mutants *nup82*, *nup136* et *nup136 nup82* (à gauche). Etude de l'expression de *PR1* après traitement au benzothiadiazole (BTH), analogue au SA, dans le sauvage et le mutant *nup136 nup82* (à droite) (Tamura *et al.*, 2017). **B.** Analyse de la sensibilité à *Pseudomonas* de différents mutants du NPC (Wiermer *et al.*, 2012). **C.** Analyse phénotypique des mutants impliquant des composants du NPC (anneau externe, ORC), *nup96* également appelé *mos3* (*modifier of snc1-3*) et *seh1* sont deux mutants connus pour être des modificateurs du mutant gain de fonction *snc1* (*suppressor of npr1-1, constitutive 1*) (Zhang and Li, 2005; Cheng *et al.*, 2009; Wiermer *et al.*, 2012). **D.** Etude de la localisation sub-cellulaire des ARNpoly(A) (ARNm) par hybridation *in situ* avec une sonde oligo(dT) dans différents mutants du NPC, les flèches indiquent les noyaux (Wiermer *et al.*, 2012).

SA. Les protéines CRWN1 et CRWN2 pourraient donc participer au dépôt de la marque H3K27me3 à certains loci (Choi and Richards, 2019).

III.3.c.ii) Les protéines CRWN affectent la voie de l'acide jasmonique

Une autre étude a montré que la protéine CRWN1 joue un rôle important dans la signalisation de la PTI en promouvant l'activation rapide des MAPK immunitaire et en permettant ainsi l'expression de gènes impliqués dans cette réponse immunitaire et l'inhibition de ceux impliqués dans la croissance de la plante (Jarad *et al.*, 2020). Lorsque le mutant *crwn1* est infecté par une souche peu virulente de *Pseudomonas*, les gènes de réponse au SA et à l'éthylène sont sous-exprimés, alors que les gènes liés à l'acide jasmonique (JA) ou impliqués dans les fonctions liées au cycle cellulaire et au métabolisme sont sur-exprimés. CRWN1 joue donc un rôle de régulateur positif des réponses PTI ainsi que dans la signalisation du JA.

L'implication des CRWN dans la réponse aux stress est également vraie chez d'autres espèces comme *Marchantia* ; chez *Marchantia*, il existe 68% de recoupement entre les gènes dérégulés dans le mutant *nmcp* (l'homologue des protéines CRWN d'*Arabidopsis*) et ceux retrouvés lors d'une infection à *Phytophthora palmivora* et 30% de recoupement de ce même mutant avec le transcriptome du mutant *jaz* (Wang *et al.*, 2021). Comme JAZ est un répresseur de la voie du JA, ceci suggère que la voie du JA est activée dans le mutant *nmcp* (Wang *et al.*, 2021).

Pour conclure, les protéines CRWN jouent donc un rôle dans la réponse aux stress biotiques et ceci passant par un fin contrôle de l'expression des gènes de défense aux stress et notamment ceux impliqués dans la réponse aux hormones comme le SA ou le JA. Même si les effets sur la voie du SA sont le mieux décrits, il reste encore à élucider comment la perte de fonction des CRWN induit une surexpression de *SID2*.

III.3.d) Implication du NPC dans la réponse au stress biotique

Tout comme pour les protéines CRWN, de plus en plus d'études s'intéressent au rôle du NPC dans la réponse au stress et notamment au stress biotique. Une étude par *microarray* du double mutant *nup82 nup136*, deux protéines paralogues du nuclear basket, a montré une sous-expression des gènes impliqués dans la réponse de défense aux pathogènes (Tamura *et al.*, 2017). L'expression des gènes ayant la plus forte dérégulation en *microarray*, dont *PR1*, a été étudiée dans les simples mutants et a montré un plus fort impact pour le mutant *nup136* mais moindre que le double mutant (**Figure 1.34A**).

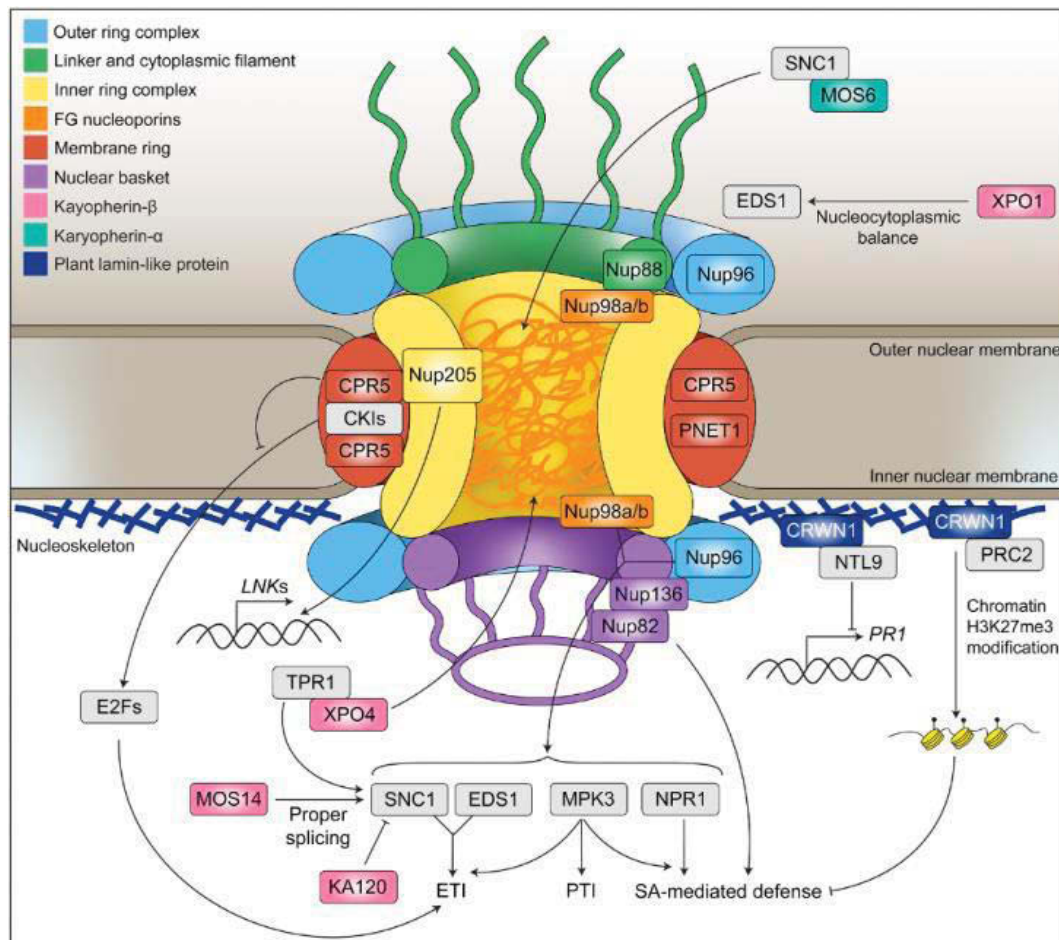


Figure 1.35 : Résumé des rôles régulateurs des nucléoporines dans le système immunitaire des plantes (Fang and Gu, 2021)

Importine : MOS6 (*importin α3*) est une karyophérine permettant le transport de *cargo* comme SNC1 impliqué dans la réponse au SA.

Nucléoporines : Les nucléoporines NUP98 et NUP88 sont des nucléoporines de l'anneau central à répétitions FG qui permettent le transport puis l'accumulation nucléaire de SNC1 mais également d'autres *cargo* comme EDS1, NPR1 et MPK3. La régulation de SNC1, EDS1, NPR1 et MPK3 implique également les facteurs XPO4, MOS14, TPR1 et KA120.

Les nucléoporines du nuclear basket NUP82 et NUP136 sont également impliquées dans la réponse au SA.

CPR5 est une nucléoporine de l'anneau membranaire empêchant l'activation d'E2F qui est essentiel pour l'activation de l'ETI.

Nucléosquelette : La protéine CRWN1 joue un rôle dans la suppression des réponses au SA en interagissant avec le facteur de transcription NTL9 qui réprime l'expression du gène de défense *PR1*. CRWN1 interagirait avec le complexe polycomb PRC2 responsable du dépôt de la marque H3K27me3 au niveau des gènes comme *PR1* favorisant la synthèse du SA.

Concordant avec ces résultats, ces mutants résistent moins aux pathogènes (*Pseudomonas*) que le sauvage, *nup136* étant le plus affecté parmi les simples mutants et une sensibilité encore plus importante est constaté chez le double mutant (Tamura *et al.*, 2017) (**Figure 1.34A**). Le traitement avec un analogue du SA, le benzothiadazole, ne restaure que faiblement la résistance à l'infection de *Pseudomonas* et la transcription des gènes principaux impliqués dans la réponse aux pathogènes comme *PR1* (**Figure 1.34A**). Ces résultats suggèrent que NUP82 et NUP136 ont des fonctions partiellement redondantes dans la réponse immunitaire dépendante du SA (Tamura *et al.*, 2017) (**Figure 1.35**). L'effet semble ici opposé aux mutants *crwn* qui provoque une surexpression de *PR1* et une meilleure résistance des plantes à *Pseudomonas*. Le NPC pourrait exercer son rôle, soit en modifiant le transport de protéines *cargo* impliquée dans l'expression des gènes de défense, soit interagissant physiquement avec les gènes de défense afin de réguler leur expression.

La nucléoporine CONSTITUTIVE EXPRESSOR OF PATHOGENESIS-RELATED GENES 5 (CPR5) joue également un rôle dans l'activation de l'ETI (*effector-triggered immunity*) chez *Arabidopsis*. CPR5 passe de l'état d'oligomère à celui de monomère lors de la reconnaissance des effecteurs par les récepteurs cytoplasmiques TIR-NB-LRR et CC-NB-LRR (coiled-coil-NB-LRR). Cette interaction libère les inhibiteurs de la kinase cycline-dépendante SIAMESE et SIAMESE-RELATED 1, qui activent le facteur de transcription E2F induisant à son tour l'expression de gènes de défense (Wang *et al.*, 2014; Gu *et al.*, 2016) (**Figure 1.35**).

Beaucoup de découvertes ont également été faites grâce au mutant gain de fonction *snc1* (*SUPPRESSOR OF NPR1-1 CONSTITUTIVE 1*) qui d'un point de vue moléculaire correspond à une activation constitutive d'un récepteur de type TIR-NB-LRR déclenchant la voie de réponse au SA. Le mutant *snc1* provoque un phénotype de petite plante comme celui observé dans les mutants *crwn* (voir paragraphe III.3.c.i). De nombreux cribles génétiques ont été effectués pour rechercher des modificateurs de *snc1* en se basant sur l'aggravation (*enhancers*) ou la suppression (*suppressors*) du phénotype de petite plante (**Figure 1.34C**) (Cheng *et al.*, 2009). Ainsi, le mutant *mos3* (*modifier of snc1-3*) correspond à une mutation de *NUP96* (ORC du NPC) (Zhang and Li, 2005) et le mutant *mos7* à une mutation de *NUP88* (ORC du NPC) (Yu *et al.* 2009) ont été identifiés comme suppresseurs de *SNC1* et montrés comme impactant l'export des ARNm, ces protéines sont donc nécessaires à l'induction de la réponse au SA. D'autres nucléoporines ont ensuite été identifiées comme *NUP160* et *SEH1*, *nup160* et *seh1* étant également des *su(snc1)* (suppresseur de *snc1*) (**Figures 1.34B-C**) et provoquent une

réretention nucléaire des ARNm impliqués dans l'immunité (**Figure 1.34D**) (Zhang and Li, 2005; Wiermer *et al.*, 2012). Ces découvertes suggèrent que l'anneau externe du NPC participant à l'export des ARNm pourrait jouer un rôle plus spécifique dans les échanges nucléo-cytoplasmiques des signaux immunitaires (**Figure 1.35**). Dans le cas de *nup160*, un appauvrissement considérable d'EDS1, un régulateur positif clé de la résistance basale et de la réponse au SA, a été observé (Wiermer *et al.*, 2012). Cet effet est également observé pour l'allèle hypomorphe de *nup88* nommé *mos7* qui provoque une diminution de l'accumulation nucléaire de régulateurs immunitaires importants comme EDS1, NPR1 et MPK3 (**Figure 1.35**).

Le pore nucléaire joue donc un rôle important dans l'immunité de la plante notamment en régulant les transports nucléo-cytoplasmiques liés à la réponse aux pathogènes. Les cribles de *su(snc1)* ont ainsi mis en évidence que le transport *via* le NPC est un des facteurs de régulation de la réponse aux pathogènes. Les composants du NPC, *cpr5*, *mos3 (nup96)*, *mos7 (nup88)*, *nup136*, *nup82*, mais également des karyophérines comme *mos6 (importin α 3)* et *mos14 (importin β)* participent à ce transport *via* le NPC. Ces données suggèrent que la modulation spécifique des concentrations nucléaires de certaines protéines de défense et notamment leur accumulation au sein du noyau régule la défense. La périphérie nucléaire *via* le NPC et le nucléosquelette permet donc de contrôler la réponse de la plante en modulant l'expression des gènes de défense aux pathogènes (**Figure 1.35**).

III.4) Stress abiotique

Outre les attaques de pathogènes, la plante doit également faire face aux changements de l'environnement qui peuvent impacter sa survie. Ces dernières années, de plus en plus d'études s'intéressent à l'impact des différents stress abiotiques comme le froid, la chaleur, la sécheresse ou le sel sur le développement de la plante (Zhang *et al.*, 2020; Wang *et al.*, 2021; Liu *et al.*, 2022). Chaque stress abiotique induit une cascade de signalisation qui lui est propre et qui influence ainsi l'expression des gènes afin d'apporter la réponse voulue. Le rôle de la périphérie nucléaire dans les réponses aux stress est moins bien documenté que la réponse aux stress biotiques mais plusieurs études suggèrent son rôle dans la réponse aux stress abiotiques.

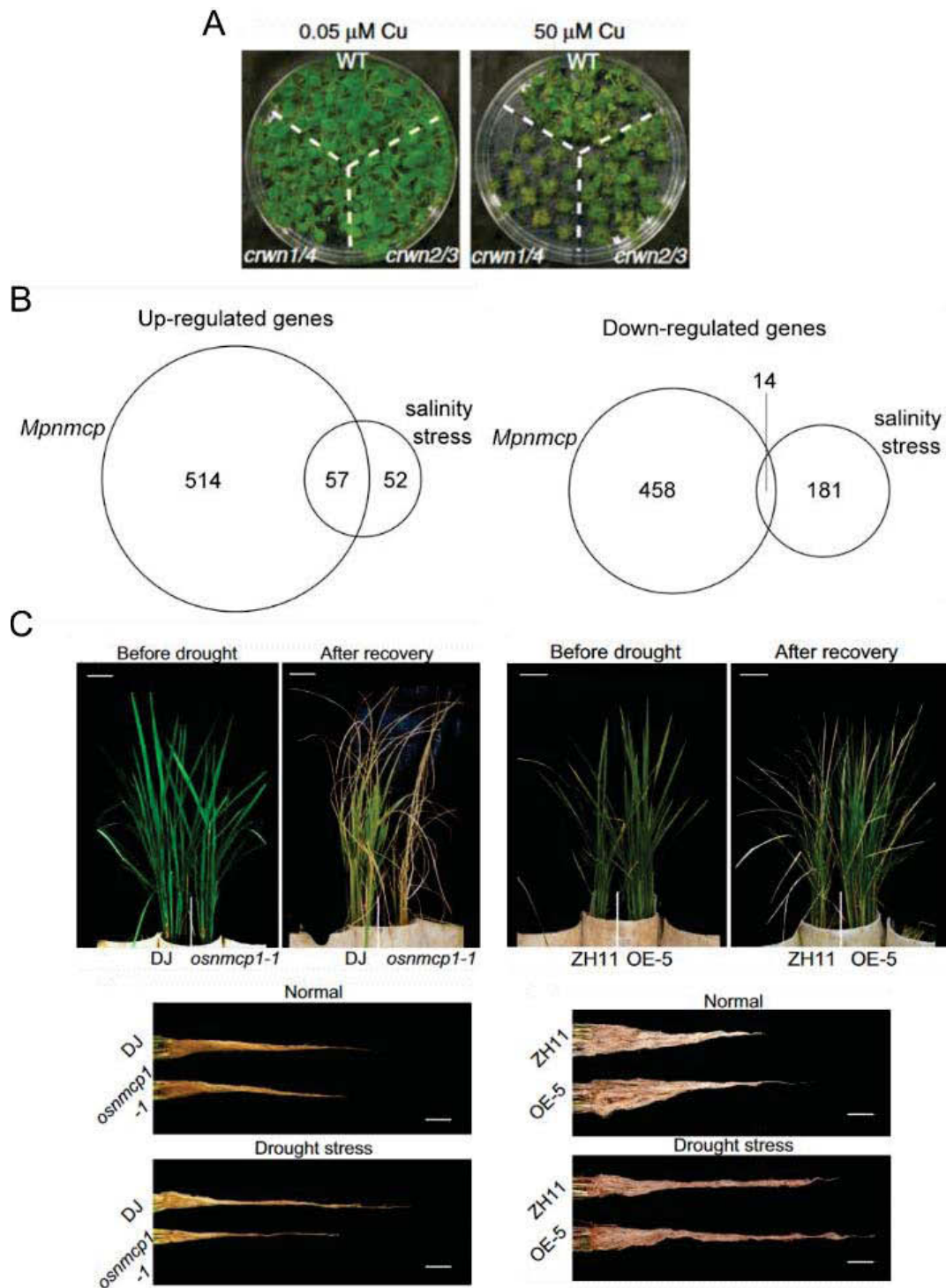


Figure 1.36 : Rôle des protéines CRWN dans la réponse au stress abiotique

A. Etude de la croissance des mutants *crwn* sur milieu normal ou en excès de cuivre. Un effet est principalement observé avec le double mutant *crwn1 crwn4* (à droite) (Sakamoto *et al.*, 2020). **B.** Comparaisons des gènes dérégulés dans le mutant *nmcp* avec ceux impliqués dans la réponse au stress salin chez *Marchantia* (Wang *et al.*, 2021). **C.** Etude phénotypique de l'appareil caulinaire et racinaire de plantes de riz sauvage, mutant pour *nmcp1* ou sur-exprimant NMCP1 au stade reproductif après un stress de sécheresse (Yang *et al.*, 2020).

III.4.a) Evidences de l'implication du nucléosquelette dans la réponse aux stress abiotiques

III.4.a.i) Tolérance à l'excès de cuivre :

Chez *A. thaliana*, le double mutant *crwn1 crwn4* provoque la sous-expression des gènes associés au cuivre (*COPPER-ASSOCIATED, CA*) présents sous forme d'un cluster de 11 gènes sur le chromosome 5 (Sakamoto *et al.*, 2020). Le double mutant *crwn1 crwn4* entraîne la sous-expression des gènes *CA7, 8* et *10*, devenant hyper sensible à la présence excessive de cuivre dans le milieu et ce phénotype est complémenté par l'expression de *CRWN1* (**Figure 1.36A**).

De façon intéressante, cet effet n'est pas observé pour le double mutant *crwn2 crwn3* suggérant que *CRWN1* et *CRWN4* mais pas *CRWN2* et *CRWN3* régulent l'expression des gènes *CA*. Enfin, la réponse au cuivre de la plante est associée à un changement de position des gènes *CA* médié par *CRWN1*. Le nucléosquelette joue donc un rôle actif dans la réponse au cuivre, modulant l'expression des gènes *CA* en changeant leur localisation nucléaire (cf paragraphe II.6).

III.4.a.ii) Tolérance à l'excès de sel

L'étude du transcriptome du mutant *nmcp* chez *Marchantia* montre que de nombreux gènes dérégulés sont impliqués dans la réponse aux stress abiotiques. En réalisant une étude approfondie, plus de la moitié des gènes impliqués dans la réponse au stress salin sont sur-exprimés dans le mutant *nmcp* (**Figure 1.36B**) (Wang *et al.*, 2021). Cela suggère que la protéine *NMCP* joue un rôle dans la régulation de la réponse au stress salin.

III.4.a.iii) Tolérance au stress hydrique

Chez le riz, la protéine *NMCP1* est induite lors de stress de sécheresse (Ning *et al.*, 2010). Le mutant *nmcp1* est plus sensible à la sécheresse alors que la surexpression de *NMCP1* entraîne la formation d'un système racinaire plus profond et plus épais, avec pour conséquence une résistance accrue à la sécheresse par rapport au sauvage (**Figure 1.36C**) (Yang *et al.*, 2020). De plus, la protéine *NMCP1* est capable d'interagir avec *SWI3C*, une sous-unité du complexe *SWI/SNF* capable de modifier la structure nucléosomale de la chromatine. Or *SWI3C* est connue pour inhiber la réponse à la sécheresse en réduisant l'accessibilité de la chromatine au niveau des gènes de réponse ainsi que ceux impliqués dans la croissance des racines (Yang *et al.*, 2020). Ceci suggère que dans des conditions de sécheresse, cette interaction pourrait provoquer la libération de *SWI3C* du complexe *SWI/SNF* ou de l'ensemble du complexe, modifiant ainsi l'accessibilité de la chromatine des gènes liés à la croissance racinaire et à la

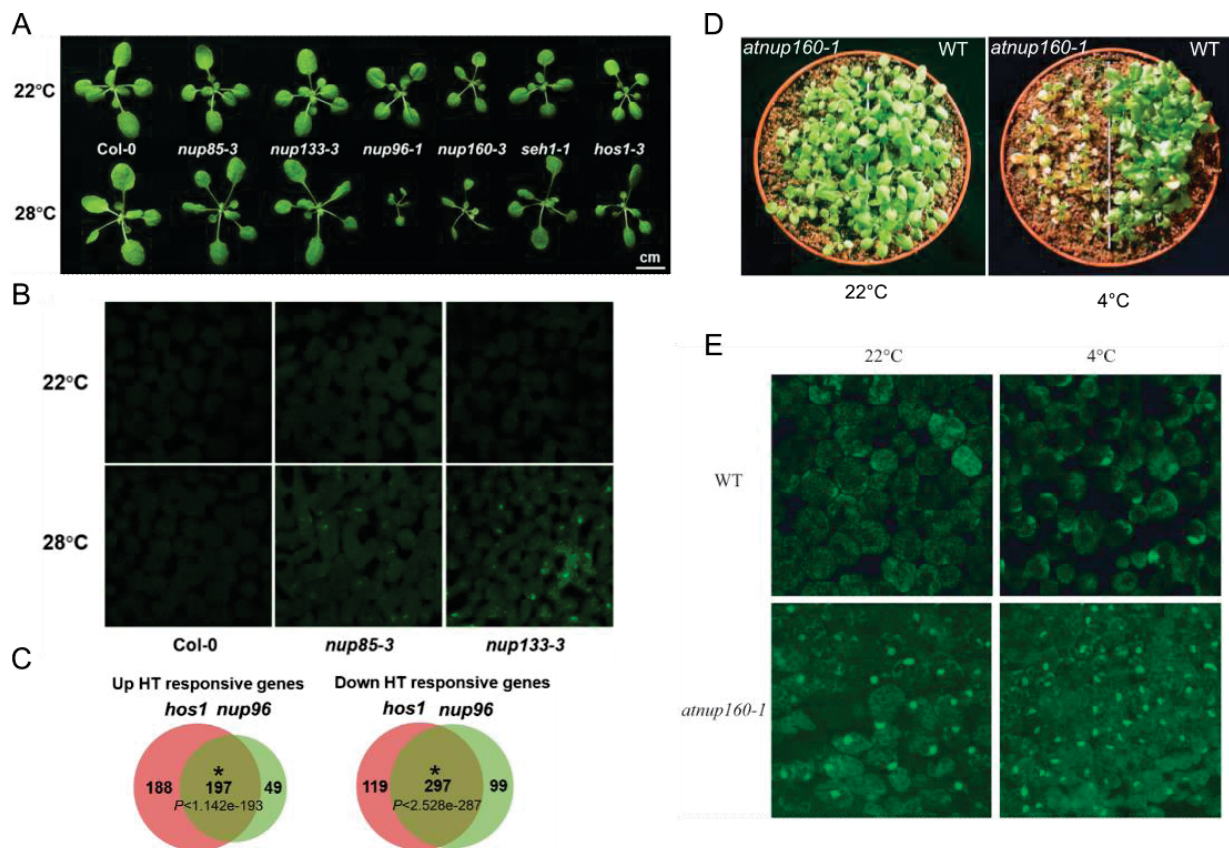


Figure 1.37 : Rôle du NPC dans la réponse au stress abiotique

A. Phénotypes des différents mutants du NPC à température normale (22°C) et élevée (28°C). *nup96* et *nup160* qui ont des tailles de plantes réduites à 22°C sont très impactées par l'augmentation de température (28°C) alors qu'inversement *nup133* à une taille augmentée à 28°C (Zhang *et al.*, 2020). **B.** Localisation des ARNpoly(A) (ARNm) par hybridation *in situ* avec une sonde oligo(dT) dans différents mutants du pore nucléaire à différentes températures (Zhang *et al.*, 2020). **C.** Comparaison des gènes dérégulés (sur-exprimé à gauche et sous-exprimés à droite) des mutants *hos1* et *nup96* à 28°C (Zhang *et al.*, 2020). **D.** Phénotypes du mutant *nup160* à température normale et au froid (Dong, Hu, *et al.*, 2006). **E.** Etude de la localisation des ARNpoly(A) (ARNm) par hybridation *in situ* avec une sonde oligo(dT) du mutant *nup160* à différentes températures (Dong, Hu, *et al.*, 2006).

résistance à la sécheresse. NMCP1 induit une résistance à la sécheresse accrue *via* une meilleure accessibilité de la chromatine au niveau des gènes liés à la résistance à la sécheresse ainsi qu'à la croissance racinaire.

III.4.b) Evidences de l'implication du NPC dans la réponse aux stress abiotiques

III.4.b.i) Tolérance aux stress hydrique et salin

Chez le maïs, la nucléoporine NUP58, NUP centrale contenant des répétition FG, joue un rôle dans la réponse à la sécheresse mais aussi au stress salin. Le mutant *nup58* est plus sensible à ces 2 stress comparé au sauvage alors que sa surexpression induit une résistance *via* l'augmentation de la teneur en chlorophylle ainsi que des activités des enzymes antioxydantes dans ces conditions (Liu *et al.*, 2022). Chez le maïs, NUP58 joue donc un rôle de régulateur positif de la résistance à la sécheresse et à la concentration en sel par la régulation de l'expression de phytohormone comme le JA ou l'ABA.

III.4.b.ii) Tolérance aux stress thermiques

Les nucléoporines NUP85, NUP133 et HOS1 ainsi que NUP96, NUP160, SEH1 (qui jouent également un rôle dans le stress biotique) ont une fonction dans la réponse au stress thermique. En contexte mutant, ces nucléoporines modifient la croissance des plantes à température élevée (28°C) avec un effet opposé entre les mutants *nup96* et *nup160* (croissance réduite) et le mutant *nup133* (croissance augmentée) (**Figure 1.37A**) (Zhang *et al.*, 2020). Les altérations de la croissance de la plante lors de fortes températures sont directement liées à la suppression de l'immunité à température élevée. PHYTOCHROME INTERACTING FACTOR 4 (PIF4) régule positivement la croissance et le développement et régule négativement l'immunité (Gangappa, Berriri and Kumar, 2017). Il existe donc un lien entre la croissance thermosensible et l'immunité chez les plantes. Ceci pourrait permettre la sélection de plantes résistantes aux maladies liées à la température.

Au niveau moléculaire, NUP96, NUP85 et NUP133 affectent l'export des ARNm (**Figure 1.37B**) (Zhang *et al.*, 2020). L'analyse des gènes dérégulés par la température des mutants *nup96* et *hos1* à 28°C montre que près de 70% des gènes dérégulés sont communs entre ces deux mutants (**Figure 1.37C**). Ces gènes sont en grande partie des cibles du facteur de transcription PHYTOCHROME INTERACTING FACTOR 4 (PIF4), régulateur central des réponses aux températures élevées mais également de la réponse aux pathogènes (**Figure 1.38**) (Gangappa, Berriri and Kumar, 2017). La mutation *snc1* entraîne une activation constitutive des réponses de défense et de graves défauts de croissance (cf paragraphe III.3.d), ces deux

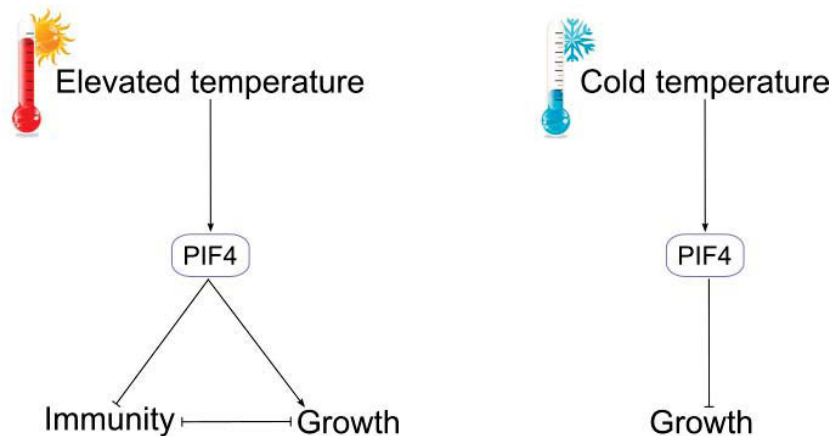


Figure 1.38 : Modèles de régulation de la croissance et de l'immunité par PIF4 en fonction de la température

A température élevée, PHYTOCHROME INTERACTING FACTOR 4 (PIF4) active la croissance et inhibe l'immunité (à gauche) alors qu'à faible température PIF4, inhibe la croissance de la plante (à droite) (adapté de Gangappa, Berriri and Kumar, 2017).

phénomènes étant complètement supprimés à une température plus élevée. Cet effet serait médié par NUP96 et HOS1 en augmentant fortement l'accumulation de PIF4 au sein du noyau lors de températures élevées (Zhang *et al.*, 2020).

Outre son rôle dans la réponse à des températures élevées, il a également été montré que NUP160 et HOS1 jouent un rôle dans la réponse au froid. En effet, le mutant *nup160* est sensible aux températures basses (**Figure 1.37D**) et altèrent l'expression de gènes impliqués dans la réponse au froid comme les facteurs de transcription C-REPEAT-BINDING FACTORS (CBF) régulant l'expression des principaux gènes de tolérance au froid ainsi qu'en entraînant l'accumulation des ARNm poly(A) dans le noyau (**Figure 1.37E**) (Dong, Hu, *et al.*, 2006). Alors que le mutant *hos1* est résistant à ces basses températures, en effet puisque HOS1 interagit avec INDUCER OF CBF EXPRESSION 1 (ICE1), un facteur de transcription régulant positivement l'induction par le froid des facteurs CBF. Cette interaction provoque l'ubiquitination d'ICE1 et ainsi sa dégradation (Dong, Agarwal, *et al.*, 2006). NUP160 joue donc un rôle positif dans la réponse au froid contrairement à HOS1 qui joue un rôle répressur.

De plus, il a été montré chez la tomate que le facteur de transcription PIF4 régule négativement la tolérance au froid *via* l'inhibition de l'expression des facteurs CBF ainsi que de la biosynthèse et des signaux de réponses de certaines hormones végétales comme l'acide abscissique, le jasmonate et la gibbérelline, en réponse aux basses températures (Wang *et al.*, 2020; Pan *et al.*, 2021). L'ensemble de ces résultats suggèrent que le NPC joue un rôle central dans plusieurs types de stress en régulant la concentration nucléaire des facteurs régulateurs comme PIF4 (**Figure 1.38**).

III.5) Bilan sur les effets du NPC et du nucléosquelette dans la réponse aux stress

Aussi bien le nucléosquelette que le pore nucléaire jouent un rôle important dans les réponses aux stress qu'ils soient biotiques ou abiotiques. L'action du NPC s'exercerait via la régulation des échanges nucléo-cytoplasmiques notamment au niveau de l'import des signaux de défenses, de l'export des ARNm ainsi qu'en interagissant avec des facteurs de transcription. Alors que le nucléosquelette via les protéines CRWN et notamment CRWN1, interagirait avec des facteurs de transcription afin de moduler l'accessibilité de la chromatine et la position des gènes dans l'architecture nucléaire.

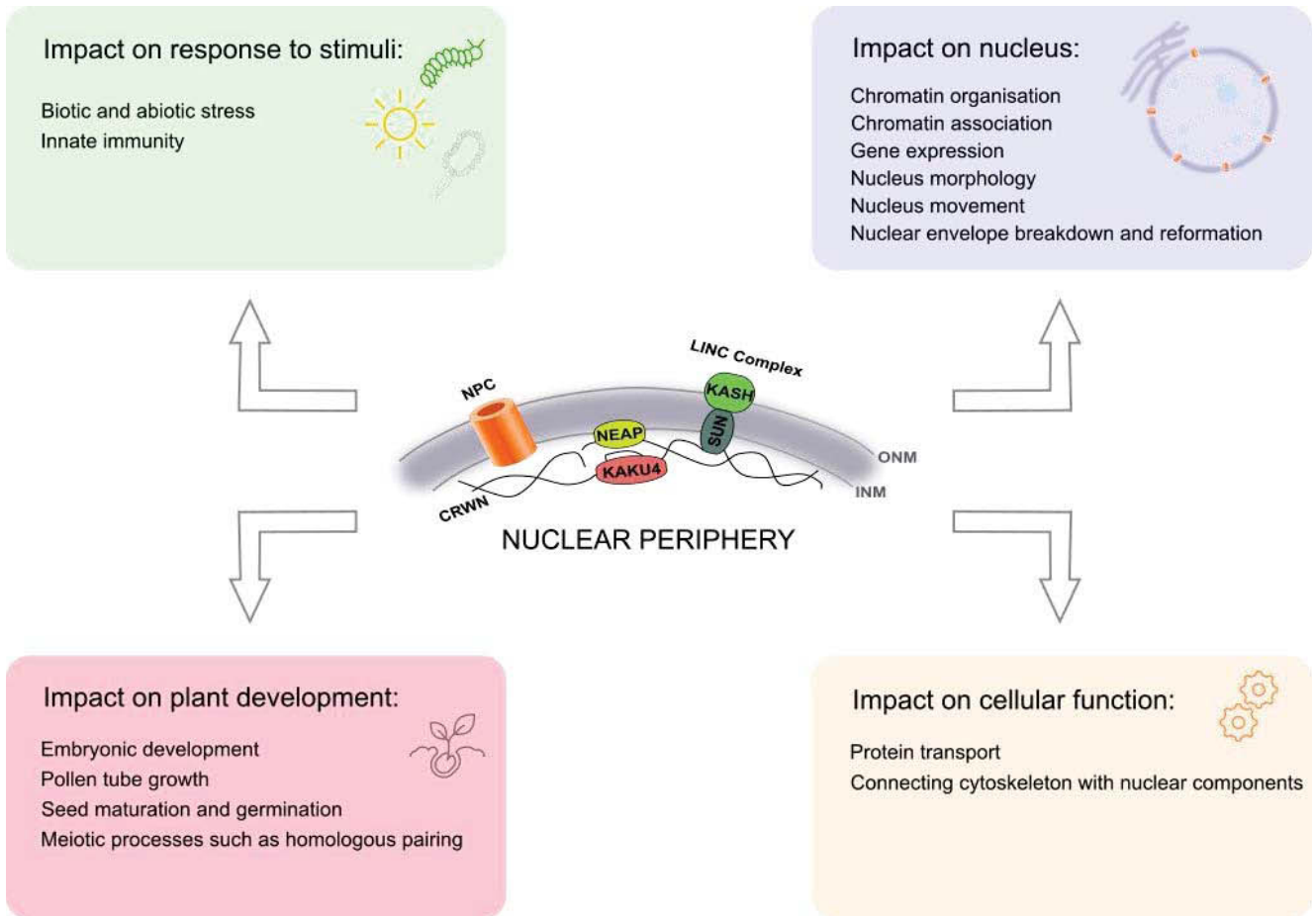


Figure 1.39 : Résumé des fonctions de la périphérie nucléaire (Evans, Mermet and Tatout, 2020)

IV- Projet de recherche

Cette introduction regroupe un ensemble d'évidences suggérant que la chromatine est organisée de façon structurée par chromosome en territoires chromosomiques, par régions de chromatine qui se regroupent en compartiments A et B, par des sous domaines chromatiniens comme les TAD, les LAD ou les NAD dans lesquels il existerait des gènes corégulés. Cette organisation contrôlée permet de réguler finement l'expression des gènes à l'intérieur du noyau principalement actif transcriptionnellement alors que la périphérie nucléaire est plutôt réprimée même s'il existe de nombreuses exceptions. Les protéines présentes à la périphérie nucléaire comme les nucléoporines et les protéines CRWN sont essentielles pour mettre en place ces régulations et pour permettre à la plante une réponse adaptée aux stress biotiques et abiotiques. La périphérie nucléaire joue donc diverses fonctions que j'ai pu résumer lors de l'écriture d'une revue avec David Evans et Christophe Tatout (**Figure 1.39**) (Revue disponible en Annexe 2) (Evans, Mermet and Tatout, 2020).

Mon projet de thèse a débuté en octobre 2019 au sein de l'équipe co-dirigée par Aline Probst et Christophe Tatout, dans laquelle nous nous intéressons à la régulation de la transcription et notamment à la contribution des protéines de la périphérie nucléaire dans cette régulation. En partant des protéines du nucléosquelette, CRWN et KAKU4, j'ai essayé de répondre à plusieurs questions : comment ces protéines se localisent à la périphérie nucléaire, quelles sont les interactions entre ces protéines et d'autres protéines de la périphérie nucléaire ou quelles sont leurs rôles dans la régulation de la transcription.

Pour cela, j'ai étudié l'interactome de ces protéines et plus précisément les régions de ces protéines nécessaires à leur interaction ainsi qu'à leur localisation à la périphérie nucléaire, en me focalisant dans un premier temps sur la protéine KAKU4 puisque dans le génome d'*Arabidopsis*, il existe qu'une seule copie de ce gène et que c'est une protéine liant les protéines CRWN. Elle joue à la fois un rôle dans la morphologie nucléaire mais aussi dans l'assemblage de la membrane nucléaire (Goto *et al.*, 2014). De plus, mon équipe d'accueil avait initié une collaboration avec Kentaro Tamura qui avait identifié KAKU4 quelques années auparavant et qui nous avait en quelque sorte donné carte blanche pour poursuivre les investigations sur cette protéine encore mal connue.

Je montrerai que KAKU4 possède de petits motifs protéiques conservés permettant son homomérisation, son interaction avec les protéines CRWN ainsi que sa localisation à la périphérie nucléaire mais également que ces motifs sont retrouvés dans deux nucléoporines

NUP82 et NUP136 présentent dans le nuclear basket. Ce résultat inattendu nous conduira à proposer que KAKU4, NUP82 et NUP136 dérivent d'un ancêtre commun qui serait une nucléoporine à FG repeats proche de NUP136. Il conduira également à se poser la question de l'importance de cette interaction entre le nucléosquelette et le NPC. Ces résultats font l'objet d'un article en cours de révision au journal *The Plant Cell*. J'ai également approfondi la caractérisation des CRWN, et pu restreindre les régions de CRWN1 nécessaires à sa localisation et à ses interactions, impliquant très vraisemblablement à nouveau de courts motifs protéiques. Enfin, j'apporterai des évidences sur le rôle de KAKU4, CRWN1, 2, 3 et 4, NUP82, NUP136 dans la morphologie nucléaire, le développement et la croissance des plantes et sur la régulation de l'expression des gènes.

Chapitre II : Résultats

Mon projet de thèse a pour but de déterminer comment les protéines du nucléosquelette, CRWN et KAKU4, se localisent à la périphérie nucléaire et leur rôle dans la régulation de l'expression des gènes. Pour cela, j'ai étudié l'interactome de ces protéines et plus précisément recherché les régions protéiques nécessaires à leurs interactions ainsi qu'à leur localisation à la périphérie nucléaire, leur rôle sur la morphologie nucléaire, sur le développement et la croissance des plantes ainsi que sur la régulation de l'expression des gènes. Nous avons choisi dans un premier temps de nous concentrer sur la protéine KAKU4. Il s'agit d'une protéine codée par un gène unique et dont la fonction était encore mal définie, les résultats de cette analyse font l'objet d'un article soumis en septembre 2022 au journal *The Plant Cell* (cf paragraphe I.1). L'étude que nous avons initiée sur les protéines CRWN sera présentée dans le paragraphe I.2). Une partie importante de mon travail a été de déterminer, par une approche de RNA-Seq, l'impact des protéines KAKU4 et CRWN sur la régulation de la transcription en utilisant une collection de multi-mutants (cf paragraphe II-). Enfin, j'ai participé à la caractérisation de la protéine MAN1, un candidat potentiel pour expliquer le recrutement des protéines CRWN et KAKU4 à la périphérie du noyau (cf paragraphe III-).

I- Connexion du nucléosquelette à la membrane nucléaire interne

Des études précédentes ont permis de montrer que les protéines CRWN et KAKU4 sont localisées à la périphérie nucléaire malgré l'absence de domaine transmembranaire, et qu'elles interagissent entre-elles ainsi qu'avec d'autres composants de la périphérie nucléaire (cf Introduction I.3)). Au début de ma thèse, seules les interactions entre KAKU4, CRWN1 et CRWN4 ainsi que celles entre CRWN1 et les protéines SUN1 et SUN2 étaient connues (Goto *et al.*, 2014; Graumann, 2014). Les interactions entre les protéines PNET2, CRWN1 et KAKU4, les connections potentielles entre NUP82, un composant du nuclear basket du pore nucléaire, et les protéines CRWN ou celles de MAN1 avec les protéines CRWN1 et CRWN4 n'étaient pas encore décrites et ont été publiées plus récemment (Tang, Huang and Gu, 2020; Tang *et al.*, 2021). Notons également que Guo *et al.* avaient montré que la localisation périphérique des protéines CRWN dépendrait d'une région protéique située à la position 751 à 820 AA (Guo *et al.*, 2017). Cette région pourrait donc être nécessaire à l'interaction avec d'autres protéines ancrées à l'INM. Nous avons donc cherché à comprendre quelles sont les régions nécessaires des protéines du nucléosquelette, CRWN et KAKU4, permettant leur localisation à la périphérie

nucléaire et leurs interactions entre-elles ainsi qu'avec d'autres composants de la périphérie nucléaire. Pour cela, j'ai recherché la présence de motifs conservés dans ces protéines et réalisé différentes constructions afin d'étudier leurs localisations par expression transitoire chez le tabac et leurs interactions par double-hybride.

[I.1\) Article : « Evolutionary conserved protein motifs drive attachment of the plant nucleoskeleton at nuclear pores »](#)

[I.1.a\) Résumé de l'article en français](#)

Le nucléosquelette forme un réseau de filaments sous l'enveloppe nucléaire et contribue à la régulation de la forme du noyau et de l'expression des gènes. Pour comprendre comment le nucléosquelette d'*Arabidopsis* se connecte physiquement à la périphérie nucléaire, nous avons étudié la protéine du nucléosquelette KAKU4 et cherché des régions fonctionnelles responsables de sa localisation à la périphérie nucléaire. Nous avons identifié trois motifs peptidiques conservés dans la région N-terminale de KAKU4 qui sont nécessaires pour l'homomérisation de KAKU4, l'interaction avec les protéines du nucléosquelette, CROWDED NUCLEI (CRWN), la localisation à la périphérie nucléaire et l'élongation nucléaire dans les tissus différenciés. Des motifs protéiques similaires ont été découverts dans NUP82 et NUP136, deux nucléoporines spécifiques des plantes qui se localisent dans le panier du pore nucléaire permettant l'interaction physique entre le pore nucléaire et le nucléosquelette. Enfin, alors que NUP82, NUP136 et KAKU4 ont une histoire évolutive commune qui dériverait d'une nucléoporine ancestrale, la protéine KAKU4 constitue un composant du nucléosquelette qui se localise en dehors du NPC, suggérant sa divergence d'une ancienne nucléoporine en un nouveau composant du nucléosquelette.

[I.1.b\) Article soumis](#)

Title: Evolutionary conserved protein motifs drive interactions between the plant nucleoskeleton and nuclear pores

Condensed title: KAKU4 small peptide motifs

Authors and affiliations

Sarah Mermet¹, Maxime Voisin¹, Joris Mordier¹, Tristan Dubos¹, Sylvie Tutois¹, Pierre Tuffery², Célia Baroux³, Kentaro Tamura⁴, Aline V. Probst¹, Emmanuel Vanrobays^{1, *}, Christophe Tatout^{1, *}

¹ iGReD, Université Clermont Auvergne, CNRS, INSERM, 63001 Clermont-Ferrand, France

² Université de Paris, CNRS UMR 8251, INSERM ERL U1133, Paris, France.

³ Department of Plant and Microbial Biology, Zürich-Basel Plant Science Center, University of Zürich, Zürich, Switzerland

⁴ Department of Environmental and Life Sciences, University of Shizuoka, Shizuoka 422-8526, Japan

* To whom correspondence should be addressed: emmanuel.vanrobays@uca.fr (ORCID: [0000-0001-5215-2338](https://orcid.org/0000-0001-5215-2338)) christophe.tatout@uca.fr (ORCID: [0000-0002-3209-6743](https://orcid.org/0000-0002-3209-6743)) iGReD, Université Clermont Auvergne, CNRS, INSERM, 63001 Clermont-Ferrand, France

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Pr Christophe Tatout (christophe.tatout@uca.fr)

Abstract

The nucleoskeleton forms a filamentous meshwork under the nuclear envelope and contributes to the regulation of nuclear shape and gene expression. To understand how the Arabidopsis nucleoskeleton physically connects to the nuclear periphery, we investigated the nucleoskeleton protein KAKU4 and sought for functional regions responsible for its localization at the nuclear periphery. We identified three conserved peptide motifs within the N-terminal region of KAKU4 that are required for homomerization of KAKU4, interaction with the nucleoskeleton proteins CROWDED NUCLEI (CRWN), localization at the nuclear periphery and nuclear elongation in differentiated tissues. Similar protein motifs were discovered in NUP82 and NUP136, two plant specific nucleoporins from the Nuclear Pore Complex (NPC) basket allowing the physical interaction between the NPC and the nucleoskeleton. Finally, whilst NUP82, NUP136 and KAKU4 have a common evolutionary history predating non-vascular land plants, KAKU4 mainly localizes outside the NPC suggesting its divergence from an ancient nucleoporin into a new nucleoskeleton component.

Introduction

The nuclear envelope has emerged as one of the key structures of the eukaryotic cell. It separates the nucleus from the cytoplasm, creating two functional

compartments with replication and transcription taking place inside the nucleus and translation in the cytoplasmic compartment (Meier et al., 2017).

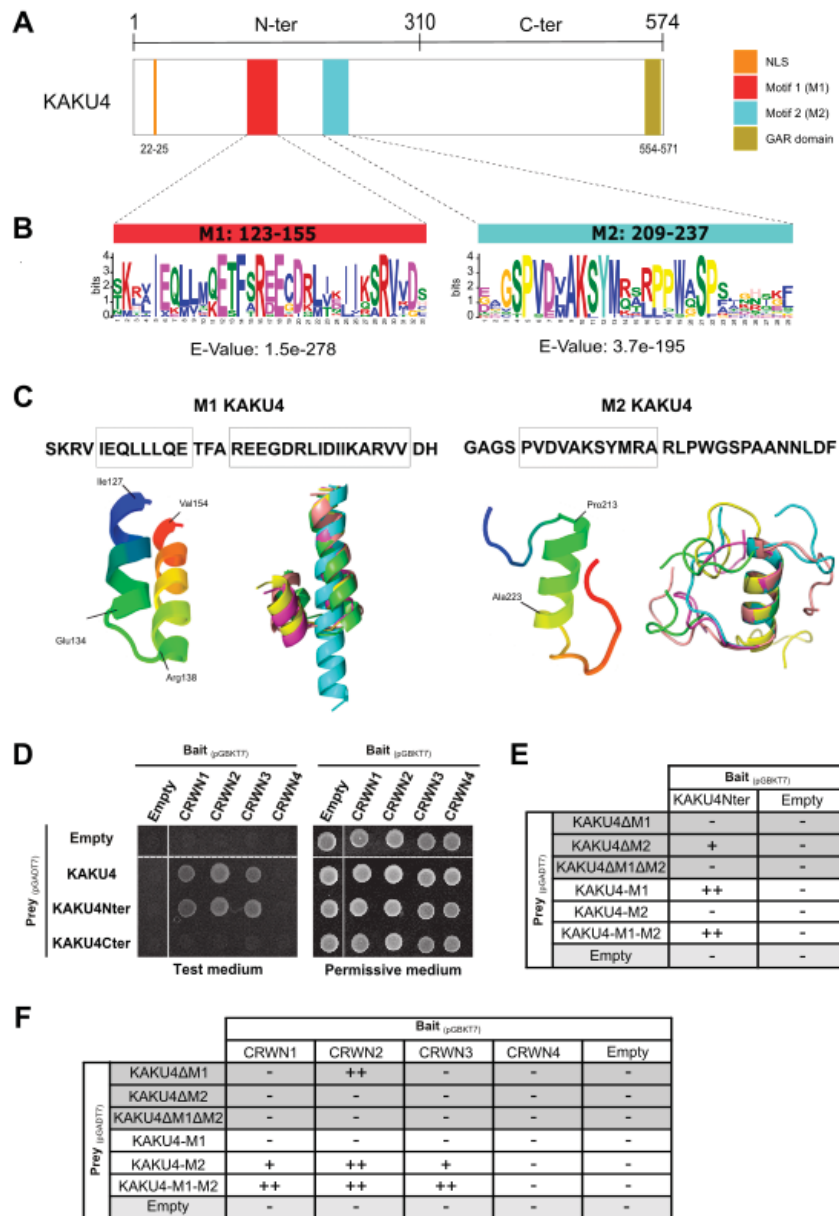


Figure 1: *Arabidopsis thaliana* KAKU4 contains two conserved motifs required for protein-protein interactions. (A) Domains and motifs organization of KAKU4: NLS (22-25), glycine arginine rich domain (GAR domain, 554-571) and two conserved motifs named M1 (123-155) and M2 (209-237) identified by MEME analysis. **(B)** Motifs identified by MEME. Motif logos correspond to the probability of each possible amino acid at each position in the peptide sequence. E-values are indicated under each logo. **(C)** Ribbon diagram of the predicted folded structures using PEP-FOLD3 with the peptide sequence used for prediction (top) and the predicted helices (boxes). Best model (left) and superposition of the five best models (right) of the helical organization of M1 and M2 (bottom). **(D-F)** Yeast two-hybrid assay on test (SD/–Leu/–Trp/–Ade/–His) (left) or permissive media (SD/–Leu/–Trp) (right). Empty vector pGBKT7 and pGADT7 are used respectively as bait and prey negative controls. **(D)** Results from Yeast-Two Hybrid assays testing interactions between KAKU4 deletion constructs and CRWN1-4. **(E-F)** Summary of Yeast-Two Hybrid interactions between M1 and M2 motifs from KAKU4, M1 and M2 deletions and the N-terminal region of KAKU4 **(E)** or the CRWN proteins **(F)**. Strong interaction (++) , interaction (+) and no interaction (-).

The nuclear envelope is made of a double membrane that is however more than a simple barrier. The outer nuclear membrane is connected to the cytoskeleton and the inner nuclear membrane to the nucleoskeleton (Meier et al., 2017). Functional components such as the Linker of Nucleoskeleton and Cytoskeleton (LINC) complex made of Klarsicht/ANC-1/Syne Homology (KASH) and Sad1/UNC84 homology (SUN) domain proteins bridge the two nuclear membranes and propagate extranuclear signals across the nuclear envelope into the nucleus to ultimately influence gene expression (Evans et al., 2020; Poulet et al., 2017). The nuclear envelope also contains several hundreds of Nuclear Pore Complexes (NPCs) that regulate the trafficking of macromolecules inside and outside the nucleus, and which play significant roles in transcription regulation (Lisitsyna et al., 2020; Tamura, 2020). Finally, the nuclear membrane is also tightly linked with the nucleoskeleton, a layer of filamentous structures maintaining the nuclear shape and tethering specific chromatin domains to the nuclear periphery (Dechat et al., 2008; Ciska and Moreno Díaz de la Espina, 2014). There is increasing evidence that the nuclear envelope, NPCs and the nucleoskeleton are closely connected to perform their function in a coordinated manner, but these interactions and their functions remain poorly understood in plants.

The plant nucleoskeleton is made by four plant specific proteins called CROWDED NUCLEI (CRWN1-4) that share long coiled-coil regions with their functional homologs, the lamins, in animal cells (Dittmer et al., 2007). Another plant specific component of the nucleoskeleton that interacts with CRWNs is KAKU4 (Goto et al., 2014; Tang et al., 2022). Despite their different composition, the nucleoskeletons in animals and plants share common functions such as maintenance of nuclear shape (Dittmer et al., 2007; Goto et al., 2014; Shumaker et al., 2006) and creation of distinct chromatin domains called Lamin Associated Domains (LADs) at the nuclear periphery (van Steensel and Belmont, 2017; Hu et al., 2019). Plant LADs have been identified by ChIP for chromatin

domains enriched in CRWN1 containing repressed genes and transposable elements belonging to the repressed B compartment (Hu et al., 2019). Chromatin regions located at the nuclear periphery have also been characterized by RE-ChIP with the nuclear pore component NUP136 (Bi et al., 2017) with a considerable overlap between the regions identified with CRWN1 raising the possibility of a possible dual contribution of NPC and nucleoskeleton to promote contact between repressed chromatin and the nuclear envelope. Interestingly, nuclear shape is also affected in a similar way in CRWN (Wang et al., 2013) and NUP136 mutants (Tamura and Hara-Nishimura, 2011) supporting a possible connection between NUP136 and the nucleoskeleton. More recently, combining subtractive proteomics and the proximity-dependent biotinylation identification technology termed BioID2, the protein interactome of the Arabidopsis nuclear periphery was characterized and includes up to 200 Plant Nuclear Envelope Transmembrane (PNET) proteins. Among them, PNET2 interacts with CRWN1 and KAKU4 and KAKU4 was found close to the CRWNs, PNET10 and 11, NPC basket proteins (NUP82, NUP136, NUP50a, NUP50c), nucleosome assembly proteins, histone H2A and H2B, histone chaperones and DNA binding proteins. These new finding suggest that PNET2, CRWN and KAKU4 proteins are situated in close proximity to chromatin (Tang et al., 2022).

Here, we dissected the protein interactions that occur at the nuclear periphery using the nucleoskeleton protein KAKU4 as an entry point. While KAKU4 neither possess any identifiable protein domain nor transmembrane domain that could account for its enrichment at the nuclear periphery, we searched for conserved regions and identified three short peptide motifs in its N-terminus. We show that these motifs are required to localize KAKU4 at the nuclear periphery, involved in protein interactions with the CRWN protein family and mandatory for elongation of the nuclear envelope in differentiated tissues. KAKU4 shares these motifs with NUP82 and NUP136, two nucleoporins located in the NPC basket allowing the two nucleoporins to interact

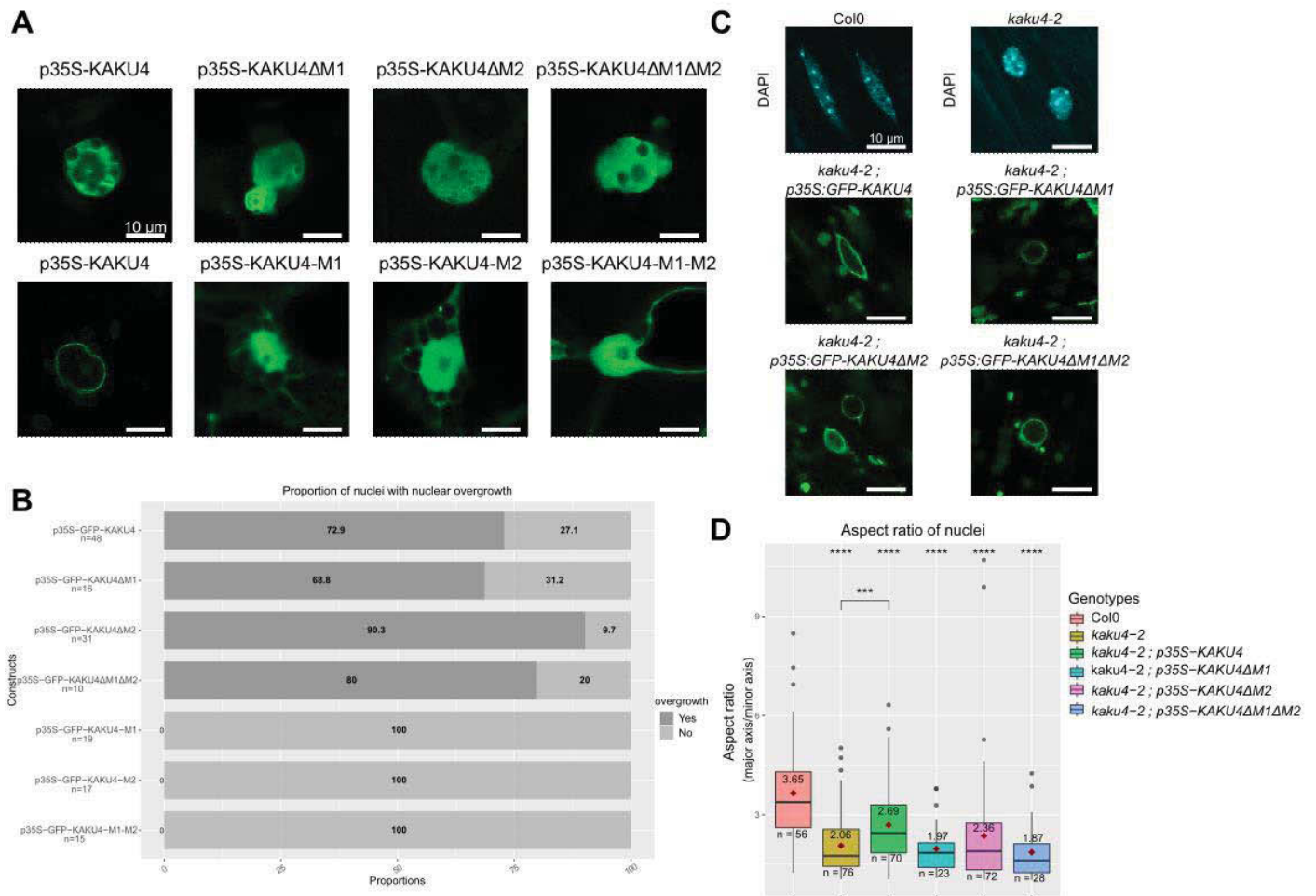


Figure 2: M1 and M2 motifs of KAKU4 contribute to nuclear shape.

(A) Assay of nuclear localization of KAKU4 truncation constructs and membrane overgrowth in transient expression assays in *Nicotiana benthamiana* using KAKU4 deletion derivatives expressed as GFP fusions. **(B)** Quantification of nuclear membrane overgrowth induced by KAKU4 deletion derivatives. **(C-D)** Characterization of nuclear shape using transgenic Arabidopsis lines complemented with KAKU4 deletion derivatives. **(C)** Cross sections of nuclei from elongated anther filament cells using confocal microscopy. Wild-type (Col0) and *kaku4-2* mutant stained with DAPI (in blue, top) and representative *kaku4-2* mutant nuclei complemented with different constructs of KAKU4 coupled with GFP (in green, middle and lower panels). **(D)** Quantification of nuclear shape using aspect ratio (major axis/ minor axis of each nucleus); Wilcoxon test, *** p-value < 0.001, **** p-value < 0.0001. Scale = 10 μ m. n = number of nuclei.

with the CRWN protein family hence demonstrating the existence of a physical connection between the plant nucleoskeleton and the NPC basket. Genetic interactions support this connection between the NPC and the nucleoskeleton with *KAKU4* mutation enhancing the *NUP136* phenotypes in vegetative and reproductive stages. Finally, predictions gained from phylogenetic analyses suggest that the small peptide motifs diverged from an ancestral plant-specific nucleoporin, which had the ability to interact with the nucleoskeleton, and led to the emergence of *KAKU4* as a new component of the plant nucleoskeleton.

Results

The N-terminal region of *KAKU4* is required for protein interaction and contains two conserved peptide motifs

Previous experiments revealed that Arabidopsis *KAKU4* and *CRWN1* interact directly to form a complex together with *CRWN4* (Goto et al., 2014). However, none of these proteins including *KAKU4* (**Figure 1A**) contain transmembrane domains that could explain their enrichment at the nuclear periphery. As complexes involving *KAKU4* were recently described (Tang et al., 2020, 2022), enrichment at the nuclear periphery could be explained by protein-protein interaction (PPI) with proteins anchored within the inner nuclear membrane. To discover possible interacting regions within *KAKU4*, we deployed the MEME suite (Bailey et al., 2015), to predict novel motifs conserved in 16 *KAKU4* orthologs from basal to recent angiosperms (Materials and Methods). This approach revealed conserved regions in the N-terminal domain of *KAKU4*. The two most conserved motifs (M) were selected and numbered according to their E-value ranking (M1, M2) (**Figure 1B**). To identify possible secondary structures adopted by the short motifs we used PEP-FOLD3 *de novo* structure modeling engine (Lamiable et al., 2016). The five best models predicted by PEP-FOLD3 for M1 and M2 sequences were superposed and suggest the occurrence of alpha helices, more specifically, a helix region (I127-V154) possibly broken around residues 135-TFA-137 was defined for M1, and a shorter helix

(P213-A223) was predicted for M2 (**Figure 1C**). Alpha-helices were also predicted by QUARK and PHYRE2 algorithms (Kelley et al., 2015; Xu and Zhang, 2012) (**Supplemental Figure S1**). As alpha helices are known to frequently be involved in PPI (Jochim and Arora, 2009), our analyses suggest that the N-terminal region of *KAKU4* could interact with other proteins via the M1 and M2 motifs.

As a first step, we tested the importance of the N-terminal tail for known protein-protein interactions with itself and with *CRWN* proteins. Several constructs expressing either the full-length protein (amino-acids 1-574), the N-terminus (1-310) or the C-terminus (311-574) of *KAKU4* were designed (Materials and Methods and **Supplemental Figure S2A**). In combinatorial yeast-two-hybrid (Y2H) experiments, the N-terminal region of *KAKU4* is sufficient to enable both homomerization (**Supplemental Figure S3A**) and interaction with *CRWN1*, *CRWN2* and *CRWN3* but not with *CRWN4* (**Figure 1D**). We also found that the N-terminal domain of *KAKU4* was sufficient for its localization at the nuclear periphery (**Supplemental Figure S4**). The N-terminal region of *KAKU4* thus plays a central role in allowing *KAKU4* to interact with its known protein partners and to ensure its localization to the nuclear periphery. We investigated M1 and M2 properties by designing two complementary sets of constructs, the first one expresses short, coding sequences restricted to M1 and/or M2 motifs while the second set expresses full-length *KAKU4* versions, from which M1 and/or M2 motifs were deleted (Materials and Methods and **Supplemental Figure S2B** and **S2C**). We tested the ability of these constructs to interact with *KAKU4* and *CRWN* proteins by Y2H. The results showed that the M1 motif interacts with *KAKU4*, while a *KAKU4* protein deleted of M1 loses this property, hence the M1 motif is required to promote interaction of *KAKU4* with itself (**Figure 1E** and **Supplemental Figure S3B**) and that M2 is necessary and sufficient for interactions with *CRWN1*, *CRWN2* and *CRWN3* proteins (**Figure 1F** and **Supplemental Figure S3C**). These results support the functional importance of the two conserved M1 and M2 motifs in the

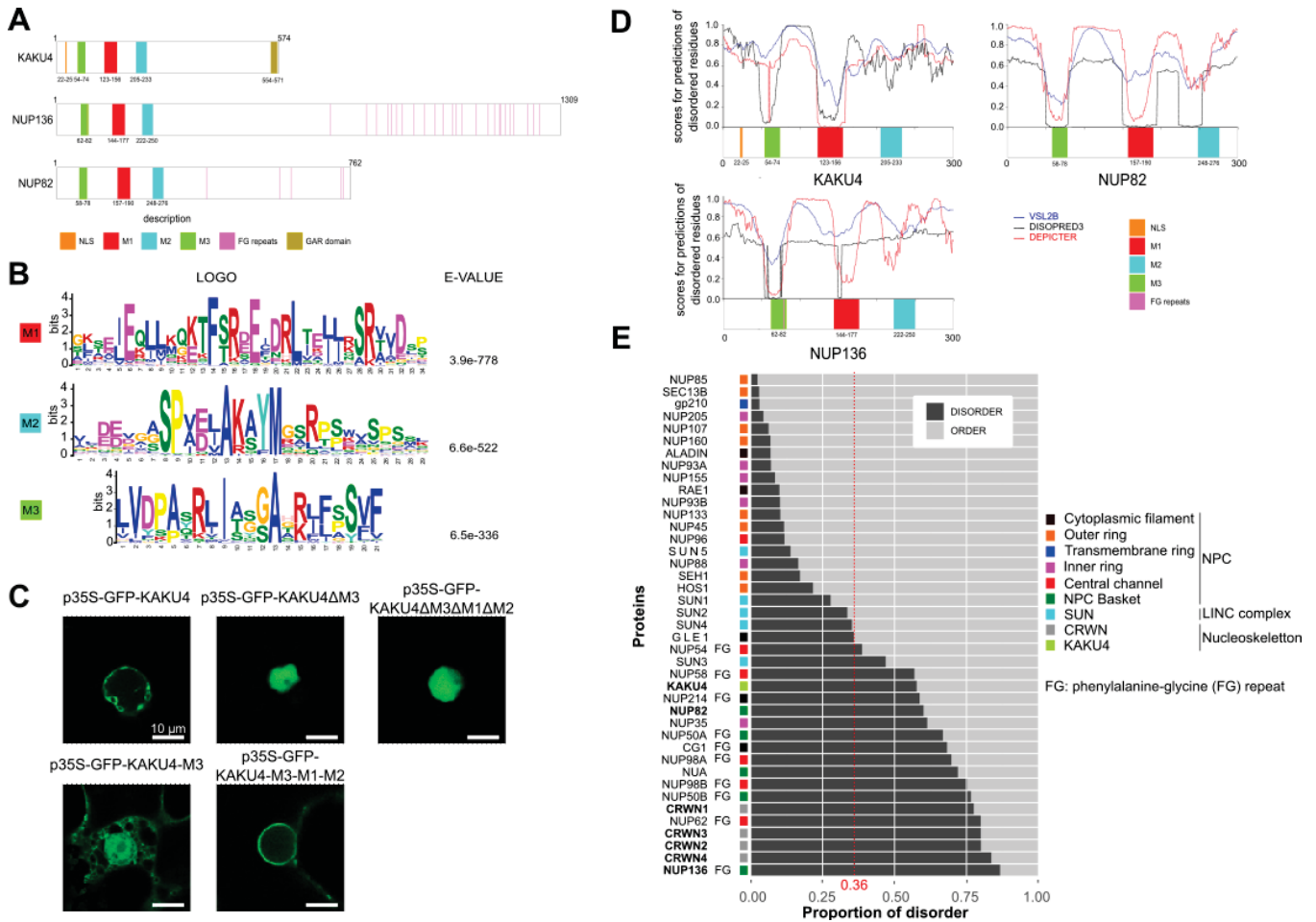


Figure 3: M1 and M2 motifs are conserved in NUP136 and NUP82.

(A) Schematic representation of *A. thaliana* KAKU4, NUP136 and NUP82 proteins. **(B)** MEME logos and associated E-value for three motifs conserved between KAKU4, NUP82 and NUP136 (Materials and Methods). **(C)** Assay of nuclear localization by confocal imaging of GFP-KAKU4 and KAKU4 deletion derivatives in *N. benthamiana*. Scale = 10 μ m. **(D)** Profiles of intrinsically disordered residues (IDR) predicted by VSL2B (blue), DISOPRED3 (black) and DEPICTER (red) within the N-terminal region (1-300) of KAKU4, NUP82 and NUP136 containing M3, M1 and M2 peptide motifs indicated in green, red and blue respectively. **(E)** Proportion of intrinsically disordered regions (dark grey) and ordered regions (light grey) in NUP, CRWN, KAKU4 and SUN proteins from *A. thaliana*. The median value (0.36) of disorder proportion of the Arabidopsis proteome is indicated as a red dashed line. FG-NUP and NPC substructures are indicated by FG and different colours respectively.

N-terminal region of KAKU4 in PPI to form a complex with CRWN proteins.

The short peptide motifs of KAKU4 are required for localization and the nuclear periphery and nuclear membrane elongation

The different KAKU4 constructs were then transiently expressed in *Nicotiana benthamiana* to record the localization of the different truncated versions of KAKU4 (**Figure 2**). In the absence of M1, M2 or both motifs, KAKU4 loses its specific localization at the nuclear periphery (**Figure 2A**) suggesting that both homomerization of KAKU4 and interaction with CRWN proteins are required to localize KAKU4 at the nuclear periphery. However, expression of M1 and M2 alone is not sufficient to anchor the protein at the periphery (**Figure 2A**).

Moreover, several molecular phenotypes were described to be caused by change in KAKU4 expression. These include excessive nuclear membrane formation and deformation (hereafter collectively referred as membrane overgrowth) when KAKU4 is overexpressed as well as reduced nuclear membrane elongation in a *kaku4* loss of function mutant (Goto et al., 2014; Tang et al., 2022). To determine the contribution of the short motifs to these two phenotypes, we first overexpressed using a p35S promoter full-length KAKU4 and M1 and/or M2 deletion derivatives in tobacco leaves by transient expression assays. While full length KAKU4 induced membrane overgrowth in 72% of the observed nuclei, so did the constructs lacking the short motifs, indicating that they are not required to induce membrane overgrowth. In agreement, no membrane overgrowth was observed in nuclei expressing only M1, M2 or M1 and M2 (**Figure 2A** and **2B**, **Supplemental Figure S2B** and **2C**). We then investigated the capacity of KAKU4 M1 and M2 deletions to complement the reduced nuclear membrane elongation observed in *kaku4* loss of function. For that purpose, we used *kaku4-2* mutant plants to establish a set of transgenic lines complemented with full length or truncated versions of KAKU4. Phenotypic analysis was performed on anther filament cells,

which contain elongated nuclei in wild type plants (Dittmer et al., 2007; Choi et al., 2019). We confirmed that *kaku4-2* mutants show reduced nuclear membrane elongation, partially complemented upon expression of full length KAKU4. However, when M1, M2 or M1 and M2 were deleted, no complementation was observed indicating that both short motifs are involved in nuclear membrane elongation (**Figure 2C** and **2D**, **Supplemental Figure S2**). Taken together, while M1 and M2 motifs are dispensable for membrane overgrowth induced by KAKU4 overexpression, they are required for KAKU4 localization at the nuclear periphery and its role in nuclear membrane elongation.

The short peptide motifs of KAKU4 are conserved in NUP82 and NUP136

To determine whether the short peptide motifs M1 and M2 are present in other proteins, we used them as query in a BLASTp search (Altschul et al., 1990) or FIMO (Grant et al., 2011). Regions of similarity were found in NUP82 and NUP136, two paralog proteins described as components of the NPC basket (Tamura et al., 2010, 2017). Motif prediction was refined by performing a new MEME analysis for the three protein families. With this strategy, three common conserved motifs were identified (**Figure 3A** and **3B**) among a set of 52 orthologs representing the *NUP82*, *NUP136* and *KAKU4* gene families (Materials and Methods and **Supplemental Table S1**): M1 and M2 previously identified in KAKU4 and a third motif, M3 located N-terminally of the two other motifs (**Figure 3B**). The discovery of this new conserved motif prompted us to test whether M3 contributes to target Arabidopsis KAKU4 to the nuclear periphery. When testing the requirement of this motif for KAKU4 localization in transient expression assays, we found that M3 alone was not sufficient, but requires the presence of M1 and M2 to mediate KAKU4 localization at the nuclear periphery (**Figure 3C**). Thus, the enrichment of KAKU4 at the nuclear periphery is mediated by the three short motifs.

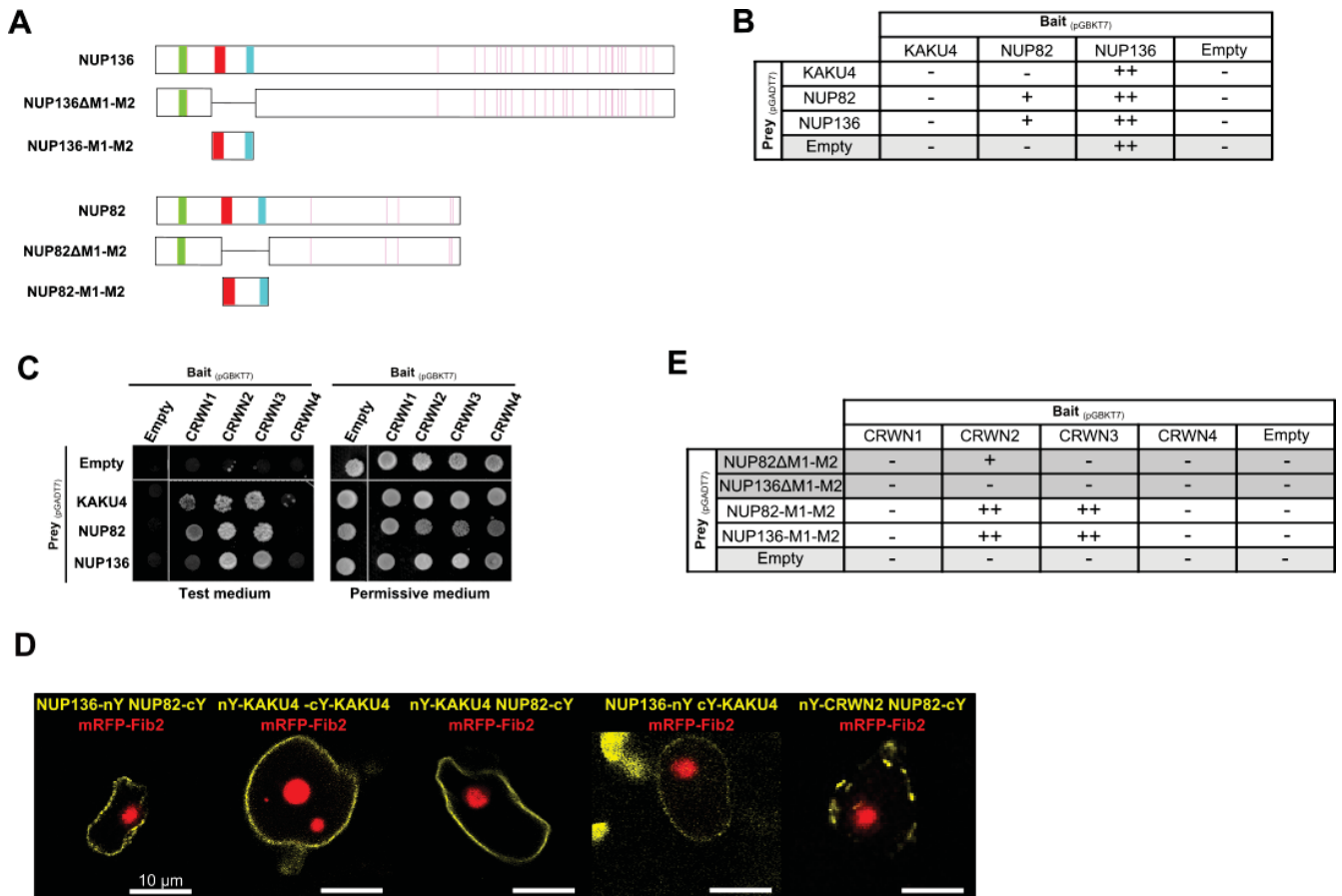


Figure 4: NUP82 and NUP136 interact with CRWN proteins via M1 and M2.

(A) Schematic representation of NUP136, NUP82 and mutant constructs. **(B-C, E)** Yeast two-hybrid assay on test (SD/-Leu/-Trp/-Ade/-His) or permissive (SD/-Leu/-Trp) media. Empty vector pGBKT7 and pGADT7 are used respectively as bait and prey negative controls. Interactions (strong interaction (++) , interaction (+) and no interaction (-)) between KAKU4, NUP136 and NUP82 proteins **(B)** and CRWN proteins **(C)**. **(D)** Protein protein interaction assay in *N. benthamiana* using Bimolecular fluorescence complementation (BiFC). Different proteins fused with the N-terminal (nY) or C-terminal (cY) half of YFP (yellow fluorescent protein) are co-expressed with mRFP-Fibrillarlin2 (Fib2, red fluorescent signal) as a positive marker for transformed cells. Scale = 10 μ m. **(E)** Contribution of M1 and M2 motifs of NUP136 and NUP82 to interact with CRWN proteins in Yeast-Two Hybrid assay.

Structure of the short peptide motifs

Short peptide motifs are often found in proteins with intrinsically disordered regions (IDR) that do not autonomously fold into a defined tertiary structure (Tompa et al., 2014). However, when protein binding is initiated at short peptide motifs, IDRs can undergo a disorder-to-order transition (van der Lee et al., 2014). We took advantage of VSL2B (Peng et al., 2006) and the meta predictors DISOPRED3 (Jones and Cozzetto, 2015) and DEPICTER (Barik et al., 2020) to predict IDRs in the N-terminal part of KAKU4, NUP82 and NUP136 (amino-acids 1 to 300). Several IDRs interspersed with structured regions overlapping the M1, M2 and M3 peptide motifs were discovered (**Figure 3D**). This prompted us to investigate more proteins located at the NPC and its vicinity. For that purpose, we took advantage of DescribePROT, a database providing access to pre-computed predictions from VSL2B for the *Arabidopsis thaliana* proteome (Zhao et al., 2021). For the whole proteome, VSL2B predicts a median value of 36% of IDRs consistent with previous predictions (Pietrosemoli et al., 2013) while NUPs from the NPC basket such as NUP50a, NUP50b, NUP82, NUP136 and NUA all contain a higher proportion of disordered regions with a median value of 72.5% (**Figure 3E, Supplemental Tables S2 and S3**). Such an enrichment in IDR was expected for NUP50a, NUP50b and NUP136 as they belong to Phenylalanine-Glycine (FG)-NUPs known to be intrinsically disordered proteins (Lemke, 2016) but this enrichment was also observed for NUP82 and NUA as well as for CRWNs and KAKU4, a feature which was overlooked so far in these proteins (**Figure 3E**).

These predictions suggest that KAKU4, NUP82, and NUP136 may not fold in stable secondary structures in their monomeric form except within their N-terminal part that contains the conserved M1, M2 and M3 motifs, which may dock the proteins within multimeric complexes.

The short peptide motifs of NUP82 and NUP136 also drive protein interactions

Since the conserved motifs M1 and M2 of KAKU4 are driving PPI, we tested if

these motifs in the two NPC basket proteins would enable their interaction with KAKU4 and CRWN proteins (**Figure 4**). We generated different constructs expressing either the full-length proteins, deletions for their corresponding M1 and M2 motifs or expressing only these two motifs (Materials and Methods and **Figure 4A, Supplemental Figure S2D and S2E**). We first performed Y2H assays between KAKU4, NUP82 and NUP136 full-length proteins. While NUP136, when expressed as bait, auto-activated the reporter construct impeding to test its ability to homomerize, NUP82 can interact with itself (**Figure 4B, Supplemental Figure S3D**). NUP82 and NUP136 interact physically, and interaction was confirmed by bimolecular fluorescence complementation (BiFC) as previously shown (Tamura et al., 2017) as was the interaction of KAKU4 with itself (**Figure 4D**). In Y2H assay using high stringency medium, we did not detect interaction between NUP82 or NUP136 with KAKU4 (**Figure 4B, Supplemental Figure S3D**), but recorded interaction of the two NUPs with KAKU4 *in planta* using BiFC assays (**Figure 4D and Supplementary Table S4**). We then tested the capacity of the two NUPs to interact with CRWN proteins, and showed that similar to KAKU4, the two NUPs interact directly with CRWN1, CRWN2 and CRWN3 proteins but not with CRWN4 in Y2H assays (**Figure 4C**). Interaction between NUP82 and CRWN2 was confirmed *in planta* by BiFC at the nuclear periphery (**Figure 4D**). NUP expressing constructs lacking the M1 and M2 motifs or containing only the two motifs were also tested for their ability to interact with CRWN proteins. In Y2H assays, M1 and M2 of NUP82 or NUP136 are sufficient to induce interactions with the CRWNs proteins (**Figure 4E, Supplemental Figure S3E**). Stronger interactions were observed with CRWN2 and CRWN3 and deletion of the two motifs completely impaired these interactions for NUP136.

All together these results suggest that the NPC basket connects the NPC with the CRWN proteins of the nucleoskeleton via the two motifs M1 and M2 present in NUP82 and NUP136.

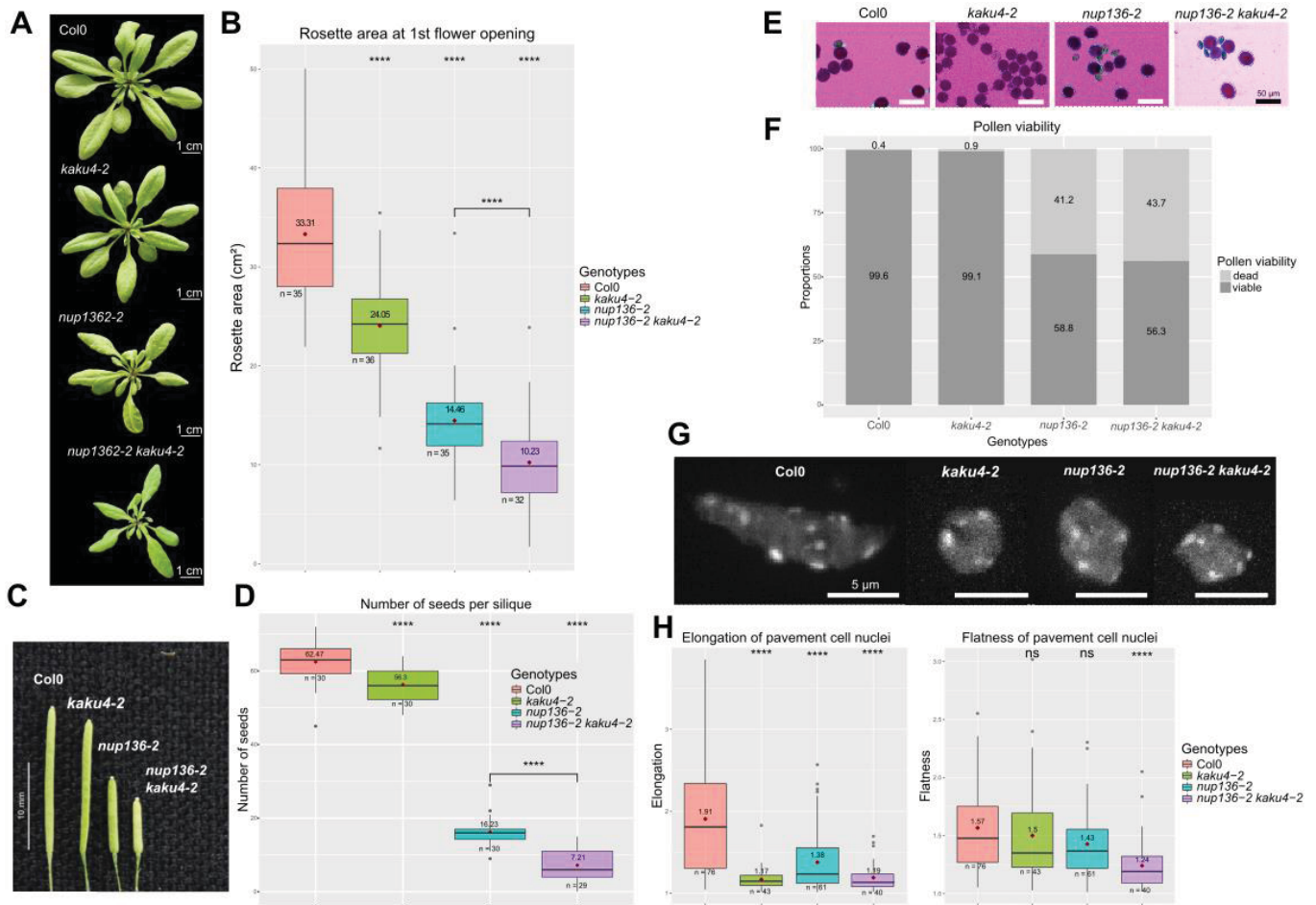


Figure 5: Genetic interaction between *NUP136* and *KAKU4*.

(A) Pictures and (B) quantification of rosette area of wild-type (Col0) and different mutants (*kaku4-2*, *nup136-2* and *nup136-2 kaku4-2*) at the first day of flowering, scale = 1 cm. (C) Size of mature siliques of different mutants, scale = 10 mm. (D) Number of seeds per silique of different mutants. (E) Microscopy observation of Alexander's staining of viable pollen in dark pink and dead pollen in blue. Scale = 50 μ m. (F) Quantification of pollen viability in different mutants. (G) Characterization of nuclear shape of DAPI stained pavement cell nuclei from young seedlings. Scale = 5 μ m. (H) Quantification of nuclear shape parameters (flatness and elongation) by *NucleusJ 2.0* software. Wilcoxon tests, ns: not significant (p-value > 0.05), **** p-value < 0.0001.

Genetic interaction occurs between *KAKU4* and *NUP136*

As single mutations, both *KAKU4* and *NUP136* loss of function mutations affect nuclear shape, impact vegetative development and reduce seed production (Tamura et al., 2010; Goto et al., 2014, 2020). Since we identified that *KAKU4* and *NUP136* proteins are able to interact, we then investigated their genetic interaction *in planta*. For that purpose, we generated the *nup136-2 kaku4-2* double mutant and recorded the phenotypic defects (**Figure 5**).

First, we observed an additive effect for the double mutant *nup136-2 kaku4-2* both at the vegetative (**Figure 5A-B**) and reproductive tissues (**Figure 5C-D**). Regarding the reproductive stage, we observed a strong decrease in silique size (**Figure 5C**) and seed number (**Figure 5D**) in the *nup136-2 kaku4-2* double mutant. As silique size has a direct impact on the number of seeds produced, we have chosen here to calculate the number of seeds per millimeter of silique to compare the different mutant. We observed a strong decrease in the number of seeds when combining *nup136-2* and *kaku4-2* (1.12 seeds/mm of silique) while *nup136-2*, *kaku4-2* mutants and wild type plants produced 2.20, 3.99 and 4.05 seeds/mm of silique respectively. In addition, pollen viability scored by the Alexander staining assay indicate that 41% of pollen grains in *nup136-2* and 44% in *nup136-2 kaku4-2* are unviable, while loss of *KAKU4* does not affect pollen viability (**Figure 5E-F**). All together, these results argue in favor of a major effect of *nup136-2* that decrease the pollen viability leading to a reduced number of seeds.

To document further the observed effects in *nup136-2 kaku4-2* double mutants, we then studied the impact on nuclear shape. In 14-day-old seedlings we observed a similar decrease in nuclear elongation for the *nup136-2*, *kaku4-2* and *nup136-2 kaku4-2* mutants compared to the wild type (**Figure 5G**). The *nup136-2 kaku4-2* mutant also display a unique nuclear flatness phenotype indicating in this case an additive effect of the two mutations (**Figure 5H**).

Finally, the phenotypic effects observed *in planta* are in good agreement with the

molecular interactions observed *in yeast* and *in tobacco* between *NUP136* and *KAKU4*.

The short peptide motifs of *NUP82*, *NUP136* and *KAKU4* originate from an ancestral nucleoporin

The structural similarities in the N-terminal region of *KAKU4*, *NUP82* and *NUP136* led us to investigate whether they have a common evolutionary origin (**Figure 6**). The presence of the three proteins was sought in a set of representative species of the green lineage selected for their fully sequenced genomes (Materials and Methods). We observed that *NUP136* is present in non-vascular land plants *Marcantiophyta*, conserved in *Bryophyta* and in all Angiosperms, which suggests that according to TimeTree (Kumar et al., 2017), *NUP136* appeared about 593 Million years ago (Mya). *KAKU4* and *NUP82* emerged more recently. They are both encoded by the basal Angiosperm, *Amborella trichopoda* genome, but the analysis including plant species before and after basal Angiosperm diversification (Zhang et al., 2020), did not allow resolving the exact timing of *KAKU4* and *NUP82* emergence. This identified *NUP136* as the most ancestral gene family with one to four paralogs in each of the analyzed species (**Figure 6A**).

To illustrate the relationship between the three gene families, we then constructed a phylogenetic tree including all homologs (**Figure 6B**). The tree includes 52 proteins organized in three connected clades with different evolutionary rates for each clade. Evolution rates were quantified by estimating the sum of the number of amino-acid substitutions per site for each clade (**Figure 6B**). The lowest value was observed for the *NUP136* clade (9.4 substitutions/site), and the highest for the *KAKU4* clade (12.1 substitutions/site). These values may argue in favor of a higher conservation of *NUP136* protein sequence, which may reflect the preservation of its original function within the NPC basket while *KAKU4* diverged faster. We then used the codon substitution modeling to determine if the observed divergence occurs randomly or if selection applied at nonsynonymous changes (Yang, 2007).

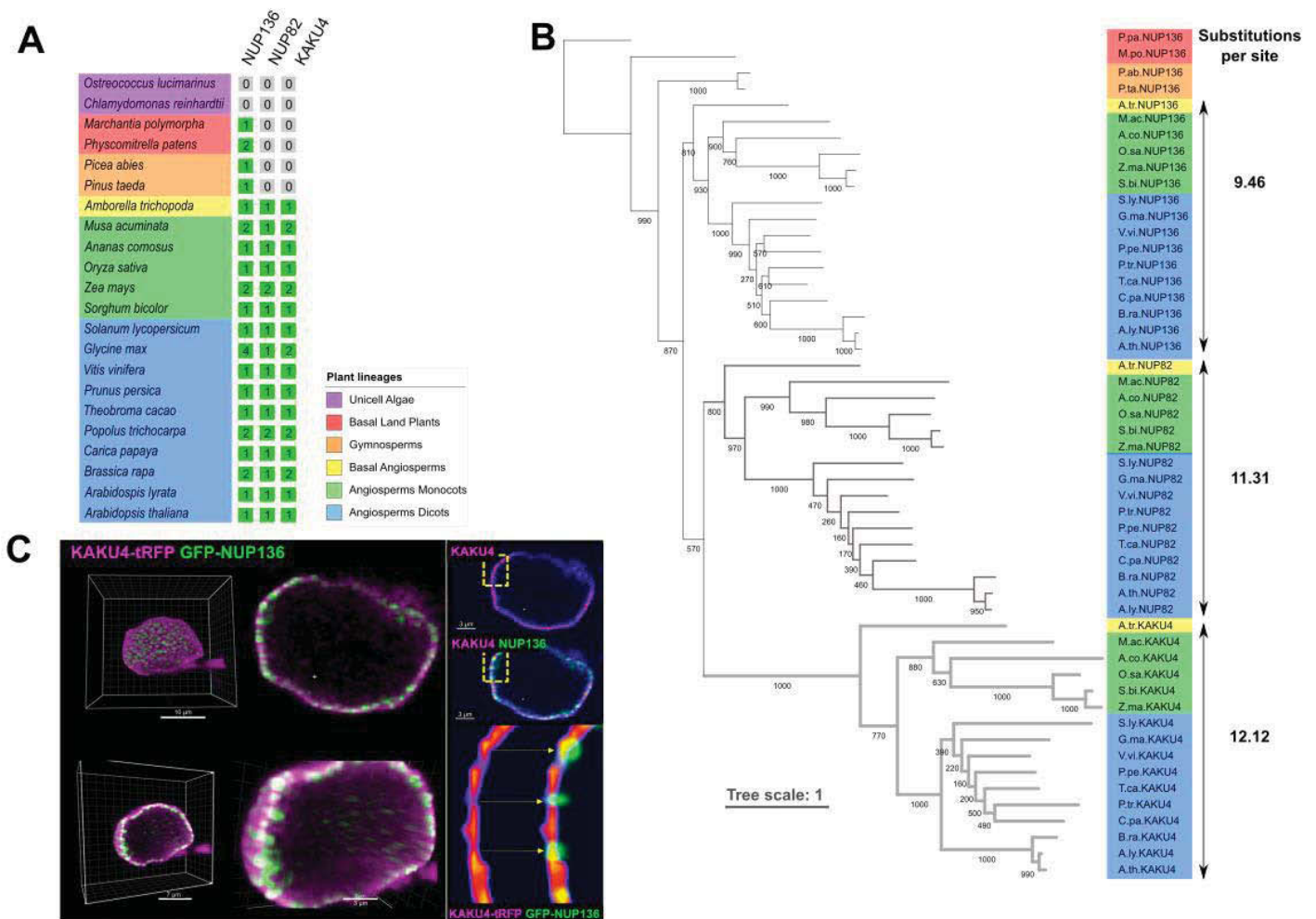


Figure 6: Phylogenetic relationships among KAKU4, NUP82 and NUP136.

(A) Presence or absence of KAKU4, NUP82 and NUP136 proteins in the green lineage. The 22 plant species used in this study (left) and numbers of paralogs per species (right) for *NUP136*, *NUP82* and *KAKU4* genes are indicated. **(B)** Maximum likelihood tree constructed from an alignment of amino-acid sequences of orthologs of *NUP136*, *NUP82* and *KAKU4*. Bootstrap values are indicated on each branch. Clades are distinguished by thin dark branches (*NUP136*), thick dark grey (*NUP82*) and thick light grey (*KAKU4*). Substitutions per site are given at the right of the tree for each clade. **(C)** Co-expression between KAKU4 and NUP136 in transient expression assay in *Nicotiana benthamiana* leaves. Rendering of 3D view using the Imaris software showing the localization of KAKU4-tRFP between spots of NUP136-GFP expression. The right-bottom panel shows a zoom in a specific region showing alternating NUP136 (green spots) and KAKU4 (red segments between the green spots).

For that purpose, we used Codeml *NSsites* on the whole cDNA or only the N-terminus and observed that selection indeed occurs in the three gene families (**Supplemental Figure S5A**). Finally, we compare in more detail the localization of NUP82, NUP136 and KAKU4 at the nuclear periphery. GFP-NUP136 shows a spotty pattern as expected for a NPC component (Tamura et al., 2010), while GFP-NUP82 has a more regular pattern as KAKU4-tRFP or forms large clusters at the nuclear periphery that might potentially be attributed to the overexpression in a heterologous system (**Figure 6C** and **Supplemental Figure S5B**). Detailed image reconstruction using IMARIS shows that KAKU4-tRFP fusion proteins adopt a more uniform distribution at the nuclear periphery, interspersed by GFP-NUP136 foci (**Figure 6C**). Even though the three proteins share the three conserved motifs, KAKU4 shows a different localization pattern than the two NUPs.

Collectively, our evolutionary analyses suggest a scenario of functional diversification of KAKU4, NUP82 and NUP136, whereas we propose *NUP136* as the parental gene ancestor with a biological function at the NPC, while NUP82 and KAKU4 diversified in protein composition and likely biological function.

Discussion

Our initial investigation of the KAKU4 protein led us to the discovery of three conserved motifs within its N-terminal region called M1, M2 and M3. These motifs range from 20 to 30 amino acids according to the MEME prediction or even smaller if only residues involved in secondary structures predicted by PEP-FOLD3 are considered. Here we could show that these motifs share at least three biological functions including protein-protein interaction (M1, M2), targeting KAKU4 at the nuclear periphery (M1, M2 and M3) and nuclear envelope elongation (M1, M2). M1 and M2 are further involved in different types of interactions, including homomerization (M1) and multimeric complex formation with the CRWN proteins (M2). These motifs are not only present in

KAKU4 homologs but also in the N-terminal part of NUP82 and NUP136, two nucleoporins of the NPC basket. Short peptide motifs have been described in NUPs from other species such as NUP93/Nic96, NUP98/Nup145N and NUP35/Nup53 (human/yeast, respectively) where they mediate key interactions needed for the assembly of the inner ring complex of the nuclear pore basket with NUP188/Nup188, NUP205/Nup192 and NUP155/Nup170, respectively (Beck and Hurt, 2017; Lin et al., 2016). Interestingly, although the M1, M2 and M3 motifs of the plant NUP82, NUP136 and KAKU4 do not share sequence similarities with the motifs in yeast and human NUPs, they are functionally similar in promoting protein-protein interactions. Extending the approach described in this study to all plant NUPs may help unravel the structural basis for protein assembly at the plant NPC.

We also observed that NUP82, NUP136 and KAKU4 proteins contain large IDRs, which are interspersed with structured regions containing the identified peptide motifs in the N-terminal tail. IDRs may function as flexible linkers favoring the exposure of the motifs to facilitate interactions with other proteins while also providing a conformational separation of these protein-docking regions compatible with the formation of a multimeric complex. This scenario is reminiscent of the entropic chain-class of IDRs found for example in the 70 kDa subunit of replication protein A (RPA70), a eukaryotic protein that is essential for replication, recombination, and DNA repair (van der Lee et al., 2014). RPA70 contains an intrinsically unstructured linker region tethering two structured domains located at the center of the protein to a third structured DNA binding domain located at its N-terminus, called DBD F. The flexibility of the RPA70 allows DBD F to interact with DNA or with other proteins and mutations in the linker region of yeast RPA70 have been shown to impair its response to UV-induced DNA damage (Daughdrill et al., 2007). We hypothesize that for NUP82, NUP136 and KAKU4, conformational transitions between fully disordered and partially structured folding states induced/triggered by motif-mediated protein-protein

interaction, may allow alternating between NPC and nucleoskeleton-specific functions. The occurrence of motifs with functional significance within disordered proteins may be more widespread in plants. While it is well documented in other species (van der Lee et al., 2014), few examples are known in plants (Hatos et al., 2020), but efforts to identify them are starting to raise awareness on the role played by these proteins notably in nuclear functions (this study and (Pontvianne and Liu, 2020)).

Here we revealed that KAKU4 shares the peptide motifs M1, M2 and M3 with the nucleoporins NUP82 and NUP136, where these motifs are involved in direct interactions with nucleoskeleton components CRWN and KAKU4. Possibly, in the absence of transmembrane domains in CRWN and KAKU4, these protein-protein interactions with nucleoporins may help to anchor the nucleoskeleton at the nuclear envelope. Besides nuclear pores, other anchoring points of the nucleoskeleton may exist, such as the inner nuclear envelope proteins, SUN. SUN1 was shown to bind CRWN1 (Graumann, 2014). Furthermore, using subtractive proteomics between NE enriched and NE depleted extracts as well as proximity-labeling proteomics, CRWNs, KAKU4, PNET2 were found to be in close proximity at the nuclear envelope (Tang et al., 2020, 2022). As PNET2 proteins are sitting within the inner nuclear membrane through their transmembrane domains, it can be expected that they contribute to anchor the nucleoskeleton at the nuclear periphery. Following this line of thought, KAKU4 may localize to the nuclear periphery not only through its interaction with CRWN1, CRWN2 and CRWN3 but also with complexes involving PNET2. Experimentally, it will remain challenging to test what drives KAKU4 localization at the nuclear periphery as both PNET2 and CRWN protein families display redundant functions and because depletion of several members of each gene family strongly impairs plant viability (Wang et al., 2013; Tang et al., 2022).

An alternative, but not exclusive, role for the interactions between nucleoporins and the nucleoskeleton may be to achieve a regular distribution of NPCs within the

nuclear envelope. NUP136 was proposed to be the functional counterpart of the human nucleoporin NUP153 (Neumann et al., 2007). NUP153 interacts with lamin A and B and this interaction is required to recruit and maintain NUP153 at the NPC (Al-Haboubi et al., 2011; Smythe et al., 2000). In mouse embryonic fibroblasts, removal of all lamins leads to clustering of the NPCs, which can be rescued by re-expression of either A or B-type lamins (Guo and Zheng, 2015). This remains to be investigated in plants.

The role of mechanical transduction in regulating gene expression is increasingly discussed (Goswami et al., 2020b; Fal et al., 2017). Since CRWN proteins also interact with genomic regions (Bi et al., 2017), and chromatin modifiers influencing gene expression (Hohenstatt et al., 2018; Mikulski et al., 2019), complexes associating CRWN proteins with the nucleoskeleton and/or the NPC may provide a molecular mechanism to the hypothesized rheostat function of the nuclear envelope (Goswami et al., 2020a). Beyond the mechanotransduction hypothesis, the molecular connection between the nucleoskeleton and the NPC may provide a mechanism to induce or enhance gene expression upon external stimuli. Indeed, rapid responses involving one or several of the NUPs, KAKU4 and CRWN proteins have been reported during hyperosmotic stress (Goswami et al., 2020b), in response to copper in the medium (Sakamoto et al., 2020) or upon biotic stress (Choi et al., 2019). Consistently, several examples suggest that gene expression in plants is influenced by the positioning of genes in respect to the nuclear periphery (Smith et al., 2015; Feng et al., 2014; Sakamoto et al., 2020; Nützmann et al., 2020). As such, the nucleoskeleton may play a central role by recruiting specific chromatin domains at the nuclear periphery (Mikulski et al., 2019; Hu et al., 2019). NPCs are also known to participate in stress response (Yang et al., 2017). More specifically NUP82 and NUP136 control the expression of defense-related genes. This can be achieved either by direct interaction between NUPs and target genes or by modulating the nuclear import of specific regulators (Tamura, 2020;

Tamura et al., 2017). Thus, the physical link between NPC basket and the nucleoskeleton may contribute to rapid transitions between an inactive (*i.e.* near the nucleoskeleton) and active states (*i.e.* near the NPCs) in response to stress, as suggested during the formation of gene loops in yeast and animals (Yang et al., 2017). Investigation of sub-nuclear localization of NUP82, NUP136 and CRWN may help to determine in the future if chromatin domains with different transcription activity can be found at the NPC-nucleoskeleton boundary.

Finally, our phylogenetic analysis traced NUP136 back to the basal land plants and suggested that NUP82 and KAKU4 emerged from an ancestral duplication. One interesting hypothesis could be that two subsequent polyploidy events termed ξ (~ 319 Mya) and ϵ (~192 Mya), which according to TimeTree occurred respectively soon before (~355 Mya) and after Angiosperm's divergence (~170 Mya) were responsible for these new duplication events (Kumar et al., 2017). During evolution KAKU4 diverged into a new component of the nucleoskeleton, but still conserved some ability to interact with the NPC. This work highlights the tight relationship between the nuclear envelope, nuclear pores and nucleoskeleton. As part of this network the KAKU4 protein represents an intriguing case of neofunctionalization of an ancestral nucleoporin into a new component of the nucleoskeleton.

Methods

Gene and Protein sequences

Orthologs of KAKU4, NUP136 and NUP82 were collected from 22 plant species that best represent the evolutionary history of the green lineage: *Arabidopsis thaliana* (A.th.), *Arabidopsis lyrata* (A.ly.), *Brassica rapa* (B.ra.), *Carica papaya* (C.pa.), *Glycine max* (G.ma.), *Theobroma cacao* (T.ca.), *Vitis vinifera* (V.vi.), *Populus trichocarpa* (P.tr.), *Prunus persica* (P.pe.), *Solanum lycopersicum* (S.ly.), *Oryza sativa* (O.sa.), *Zea mays* (Z.ma.), *Sorghum bicolor* (S.bi.), *Ananas comosus* (A.co.), *Musa acuminata* (M.ac.), *Amborella trichopoda* (A.tr.), *Picea abies* (P.ab.), *Pinus taeda* (P.ta.), *Marchantia polymorpha* (M.po.),

Physcomitrella patens (P.pa.), *Oestrococcus lucimarinus* (O.lu.) and *Chlamydomonas reinhardtii* (C.re.). Most genomes of these plant species are available at Phytozome 7.0 (Goodstein et al., 2012) and Monocots or Dicots PLAZA 4.5 (Van Bel et al., 2018). The gymnosperm genomes were collected from conGenIE.org (Conifer Genome Integrative Explorer) (Sundell et al., 2015). Gene and protein accession numbers used in this study are listed in **Supplemental Table S1**.

Protein structure predictions

Protein structure of KAKU4, NUP82 and NUP136 were defined using Uniprot (The UniProt Consortium, 2019) and pfam (El-Gebali et al., 2019) databases for domain predictions.

Secondary structures were predicted using PHYRE2, QUARK and PEP-FOLD3 (Kelley et al., 2015; Lamiable et al., 2016; Xu and Zhang, 2012). Intrinsically disordered regions (IDRs) were predicted using DEPICTER, DISOPRED3 and VSL2B (Barik et al., 2020; Jones and Cozzetto, 2015; Peng et al., 2006) (**Supplemental Table S2**). VSL2B is the baseline predictor of VSL2 that is using only the amino-acid composition-based features that can be calculated directly from the protein sequence (Peng et al., 2006). DescribePROT database was used to collect pre-computed IDR predictions for 27,466 proteins of the *Arabidopsis thaliana* proteome including those predicted by VSL2B (Zhao et al., 2021). VSL2B predictions were extracted for 31 NUPs, 5 nucleoskeleton proteins and 5 inner nuclear membrane proteins (**Supplemental Table S3**). All schemes of protein structures were drawn to scale using the drawProteins R package (Brennan, 2018).

Phylogenetic studies

Phylogenetic tree was constructed using protein sequences collected from 52 species using MAFFT 7.407 (Kato and Standley, 2013) for multiple alignment and PhyML 3.1 (Guindon et al., 2010) with the LG substitution model for tree with 500 bootstrap replicates. Trees were refined using the Interactive Tree Of Life (ITOL) (Letunic and Bork, 2011). The amino-acid substitution rate was computed as the sum

of the branch lengths (*i.e.* sum of the amino-acid substitutions per site) of the phylogenetic tree described above and by considering separately KAKU4, NUP82 and NUP136 clades. Only the branch lengths corresponding to the 16 species shared by the clades were considered.

Domain and motif search in proteins

The orthologs of KAKU4, NUP82 and NUP136 were identified using the Arabidopsis proteins as a query by performing an all-against-all protein sequence similarity search using BLASTP with an E-value cut-off of 10^{-10} . Ortholog occurrence in distant species was refined using sequences from *Amborella trichopoda* (basal angiosperm species) with the same approach. The MEME 5.1.1 (Multiple Em for Motif Elicitation) suite was used for *de novo* motif predictions (Bailey et al., 2015). From a list of orthologous proteins, MEME was parameterized to define 10 motifs, each with a maximum length of 100 amino-acids. Protein sequences from each clade (KAKU4, NUP82 and NUP136) were submitted to MEME either individually, or in combination of 2 or 3 with the assumption that the three protein families contain common conserved regions.

Nonsynonymous (dN) and synonymous (dS) substitution rates were calculated to determine if selection or genetic drift occurred within the *KAKU4*, *NUP82* or *NUP136* gene families. To avoid underestimation of dS, the analysis was restricted to 9 close relative species from Brassicaceae (*Arabidopsis thaliana*, *Arabidopsis lyrata*, *Camelina sativa*, *Capsella rubella*, *Eutrema salsugineum*, *Brassica napus*, *Brassica oleraceae*, *Brassica rapa*, *Raphanus sativus*). The Codeml program of the Phylogenetic Analysis by Maximum Likelihood (PAML) package (v.4.4c) (Yang, 2007) was used with the variable *NSsites* to compare the observed dN/dS ratio for KAKU4, NUP82 and NUP136 phylogenetic tree with the null model M0 where all tree branches have the same dN/dS set to 1 (no selection). The M0 model was rejected if P-value < 0.05.

Constructs and cloning

cDNAs of *CRWN1* (AT1G67230), *CRWN2* (AT1G13220), *CRWN3* (AT1G68790), *CRWN4* (AT5G65770), *KAKU4* (AT4G31430), *NUP136* (At3g10650), *NUP82* (AT5G20200) and *FIBRILLARIN2* (AT4G25630) were cloned in pDONR vectors using gene-specific primer sequences (**Supplemental Table S5**). Motifs were amplified from the corresponding cDNA using specific primers (**Supplemental Table S5**) and cloned in pDONR vectors, while the domain deletions were built from initial cDNA constructs using the site-directed mutagenesis kit QuikChange™ (Agilent Technologies) and adequate primers (**Supplemental Table S5**). All cloning procedures rely on Gateway technology®. After initial cloning, full-length coding sequences and other constructs were cloned into appropriate expression vectors for yeast-two hybrid assays (bait vector pDEST-GBKT7 or prey vector pDEST-GADT7), expression *in planta* (pGWB606-GFP, pGWB555-RFP) or BiFC (pB4cYGW, pB4nYGW, pB4GWnY and pB4GWcY). The list of all the plasmids used in this study can be found in **Supplemental Table S6** and **S7**.

Plant Materials

All mutants used in this study are T-DNA insertion alleles in the Columbia-0 (Col-0) genetic background. The mutant combinations were produced by crosses using *kaku4-2* (SALK_076754) and *nup136-2* (SAIL_796_H02) alleles described in previous work (Tamura et al., 2017; Goto et al., 2014). Plants were grown in long-day light conditions (16 h of light/8 h of dark) at 23°C on soil in Arabilab growth chambers. Genotyping of individual T-DNA alleles was performed by standard PCR using allele-specific primers (**Supplemental Table S5**).

Yeast Two-Hybrid Assay

Yeast cultures were grown at 30°C on YPD or on selective SD media. *Saccharomyces cerevisiae* strains AH109 Gold and Y187 (Clontech, MATCHMAKER GAL4 Two-Hybrid System) were transformed according to the classical protocol (Daniel Gietz and Woods, 2002) separately by the bait vector pDEST-

GBKT7 or the prey vector pDEST-GADT7 (Rossignol et al., 2007) and grown on appropriate medium (SD-Trp or SD-Leu plates respectively). After mating on YPD and selection of diploids on SD-Leu-Trp medium, interactions were recorded on stringent medium SD-Leu-Trp-His-Ade. Empty pDEST-GBKT7 or pDEST-GADT7 vectors were used as negative controls.

Transient expression of fusion proteins in tobacco

Vectors pGWB606 (GFP Nter), pGWB555 (mRFP Nter) (Nakagawa et al., 2007) or pB4GWnY (Nter of YFP / nYFP Cter), pB4nYGW (Nter of YFP / nYFP Nter) pB4GWcY (Cter of YFP / cYFP Cter) and pB4cYGW (Cter of YFP / cYFP Nter) (Kamigaki et al., 2016) were used for transient expression studies. Expression vectors were transformed in strain GV3101 of *Agrobacterium tumefaciens*. The p19 suppressor of gene silencing was used to enhance expression (Norkunas et al., 2018). Transient co-expression of the expression vectors and p19 constructs was performed by infiltration *Agrobacterium* cultures at OD₆₀₀ of 0.1 and 0.05, respectively, into 5-6-week-old leaves of *Nicotiana benthamiana* as described before (Sparkes et al., 2006).

Image acquisition, storage and analysis

Fluorescence images were obtained using an inverted confocal laser-scanning microscope (LSM800; Carl Zeiss). The 488-nm line of a 40-mWAr/Kr laser, and the 544-nm line of a 1-mW He/Ne laser were used to excite GFP/YFP, and RFP, respectively. Images were acquired with 40x or 63x oil immersion objectives. All images are stored in an *in house* OMERO server and figures were produced using OMERO.figure (Allan et al., 2012).

Nuclear shape analysis

3D images were obtained from whole mount preparations of 13 dag cotyledons as previously described (Poulet et al., 2017; Dubos et al., 2020). Briefly, pairs of cotyledons were collected and fixed using 1% formaldehyde and 10% DMSO in 1× PBS with 6.7 mM EGTA (pH 7.5) under vacuum for 5 min and incubated for 25 min at room temperature. Tissues were then

washed with methanol and ethanol washes, to obtain transparent tissue preparations. DNA in whole-mount preparations was stained with Hoechst (Hoechst 33 258 overnight at 1 µg/ml final at 4°C).

Microscopic observations were performed by structured illumination microscopy to produce wide-field stacks using an Optigrid module (Leica Microsystems MAAF DM 16000B). All stacks were acquired using an X63 oil N.A. 1.4 objective and a digital CMOS Camera (ORCA-Flash4.0 V2 C11440-22 CU – Hamamatsu) at an optimal resolution such that lateral and axial resolution were respectively XY = 0.103 µm and Z = 0.2 µm. The ImageJ plugin NucleusJ2.0 was used to characterize nuclear shape (Dubos et al., 2020).

Pollen viability assay

Pollen viability was studied via Alexander staining (ethanol 10% v/v, malachite green 0.01% w/v, distilled water 50% v/v, glycerol 25% v/v, phenol 5.5%v/v, chloral hydrate 5% w/v, fuchsin acid 0.05% w/v, orange G 0.005% w/v and glacial acetic acid 2% v/v). For this purpose, buds with their white petals starting to appear were dissected to recover the stamens with their mature unopened anthers. The stamens were then placed in Alexander stain and the anthers were crushed to release the pollen grains. The slides were observed under a Zeiss Imager M2 microscope in white light (bright field with an RGB camera) with X10 and/or X20 objectives and at least 1000 pollen grains per genotype from different plants were studied.

Stable expression of KAKU-GFP fusions and nuclear shape complementation

Full length and deletion derivatives of KAKU4 were expressed under the 35S promotor as stable transformants in the *kaku4-2* mutant using similar constructs as in transient expression (pGWB506, GFP Nter). Observation of nuclear shape was performed in filament cells of stamens which are non-pigmented and contain elongated nuclei in wild-type plants (Dittmer et al., 2007; Choi et al., 2019). For Col0 and *kaku4-2* mutant, anther filaments were fixed with ethanol:acetic acid solution (3:1,

v/v) for 10 min, then washed in PBS 1X for 10 min and mounted in Vectashield with DAPI. For each construct, two independent transgenic lines were selected (T2, T3 and T4 transgenic plants). After collection, filaments were placed in PBS 1X and mounted in PBS 1X and observed using confocal microscopy (Zeiss LSM 800). The images (0.31x0.31x1) were analyzed with the Fiji plugin shape descriptors to compute shape parameters.

Acknowledgments

The authors declare no competing financial interests. This work was supported by Centre National de la Recherche Scientifique (CNRS), Institut National de la Santé et de la Recherche Médicale (INSERM), Université Clermont Auvergne (UCA), 16-IDEX-0001 CAP20-25 challenge 1, Pack Ambition Recherche project *Noyau-HD* from the Region Auvergne Rhone-Alpes and the COST-Action INDEPTH (CA16212) to SM, EV, JM, ST, TD, AVP and CTa. This work was also supported by Grants-in Aid from the Ministry of Education, Culture, Sports,

References

- Al-Haboubi, T., Shumaker, D.K., Köser, J., Wehnert, M., and Fahrenkrog, B. (2011). Distinct association of the nuclear pore protein Nup153 with A- and B-type lamins. *Nucleus* 2: 500–509.
- Allan, C. et al. (2012). OMER0: flexible, model-driven data management for experimental biology. *Nature Methods* 9: 245–253.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J Mol Biol* 215: 403–410.
- Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME Suite. *Nucleic Acids Res* 43: W39–W49.
- Barik, A., Katuwawala, A., Hanson, J., Paliwal, K., Zhou, Y., and Kurgan, L. (2020). DEPICTER: Intrinsic Disorder and Disorder Function Prediction Server. *Journal of Molecular Biology* 432: 3379–3387.
- Beck, M. and Hurt, E. (2017). The nuclear pore complex: understanding its function through structural insight. *Nat Rev Mol Cell Biol* 18: 73–89.
- Bi, X., Cheng, Y.-J., Hu, B., Ma, X., Wu, R., Wang, J.-W., and Liu, C. (2017). Nonrandom domain organization of the Arabidopsis genome at the nuclear periphery. *Genome Res.* 27: 1–12.
- Brennan, P. (2018). drawProteins: a Bioconductor/R package for reproducible and programmatic generation of protein schematics. *F1000Res* 7: 1105.
- Choi, J. and Richards, E.J. (2020). The role of CRWN nuclear proteins in chromatin-based regulation of stress response genes. *Plant Signaling & Behavior* 15: 1694224.
- Science and Technology of Japan (18K06283), and by the Human Frontier Science Program (RGP0009/2018 to KT) from the International Human Frontier Science Program Organization to KT. All pictures were acquired on the CLIC microscopy facility (CLermont Imagerie Confocale). CTa would like to thank Lukasz Kurgan for his advices on the use of DescribeProt and Antoine Molaro for critical reading.

Author contributions

Emmanuel Vanrobays and Christophe Tatout supervised the study. Sarah Mermet, Emmanuel Vanrobays, Sylvie Tutois and Maxime Voisin performed experiments. Joris Mordier and Tristan Dubos performed the bio-informatics work on evolution and IDR. Pierre Tuffery performed the PEPS-FOLD3 analysis. Kentaro Tamura provided the initial plant lines and constructs used in this study. Célia Baroux performed the IMARIS analysis. Christophe Tatout, Sarah Mermet, Emmanuel Vanrobays and Aline Probst wrote the manuscript.

- El-Gebali, S. et al. (2019). The Pfam protein families database in 2019. *Nucleic Acids Res* 47: D427–D432.
- Evans, D.E., Mermet, S., and Tatout, C. (2020). Advancing knowledge of the plant nuclear periphery and its application for crop science. *Nucleus* 11: 347–363.
- Fal, K., Asnacios, A., Chabouté, M.-E., and Hamant, O. (2017). Nuclear envelope: a new frontier in plant mechanosensing? *Biophys Rev* 9: 389–403.
- Feng, C.-M., Qiu, Y., Van Buskirk, E.K., Yang, E.J., and Chen, M. (2014). Light-regulated gene repositioning in *Arabidopsis*. *Nat Commun* 5: 3027.
- Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D.S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res* 40: D1178–D1186.
- Goswami, R., Asnacios, A., Hamant, O., and Chabouté, M.-E. (2020a). Is the plant nucleus a mechanical rheostat? *Current Opinion in Plant Biology* 57: 155–163.
- Goswami, R., Asnacios, A., Milani, P., Graindorge, S., Houlné, G., Mutterer, J., Hamant, O., and Chabouté, M.-E. (2020b). Mechanical Shielding in Plant Nuclei. *Current Biology* 30: 2013–2025.e3.
- Goto, C., Tamura, K., Fukao, Y., Shimada, T., and Hara-Nishimura, I. (2014). The Novel Nuclear Envelope Protein KAKU4 Modulates Nuclear Morphology in *Arabidopsis*. *Plant Cell* 26: 2143–2155.
- Goto, C., Tamura, K., Nishimaki, S., Maruyama, D., and Hara-Nishimura, I. (2020). KAKU4 determines the migration order of the vegetative nucleus and sperm cells in pollen tubes. *Journal of Experimental Botany*: eraa367.
- Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27: 1017–1018.
- Graumann, K. (2014). Evidence for LINC1-SUN Associations at the Plant Nuclear Periphery. *PLoS ONE* 9: e93406.
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst Biol* 59: 307–321.
- Guo, T., Mao, X., Zhang, H., Zhang, Y., Fu, M., Sun, Z., Kuai, P., Lou, Y., and Fang, Y. (2017). Lamin-like Proteins Negatively Regulate Plant Immunity through NAC WITH TRANSMEMBRANE MOTIF1-LIKE9 and NONEXPRESSOR OF PR GENES1 in *Arabidopsis thaliana*. *Molecular Plant* 10: 1334–1348.
- Guo, Y. and Zheng, Y. (2015). Lamins position the nuclear pores and centrosomes by modulating dynein. *MBoC* 26: 3379–3389.
- Hatos, A. et al. (2020). DisProt: intrinsic protein disorder annotation in 2020. *Nucleic Acids Res* 48: D269–D276.
- Hohenstatt, M.L., Mikulski, P., Komarynets, O., Klose, C., Kycia, I., Jeltsch, A., Farrona, S., and Schubert, D. (2018). PWWP-DOMAIN INTERACTOR OF POLYCOMBS1 Interacts with Polycomb-Group Proteins and Histones and Regulates *Arabidopsis* Flowering and Development. *The Plant Cell* 30: 117–133.
- Hu, B., Wang, N., Bi, X., Karaaslan, E.S., Weber, A.-L., Zhu, W., Berendzen, K.W., and Liu, C. (2019). Plant lamin-like proteins mediate chromatin tethering at the nuclear periphery. *Genome Biology* 20: 87.
- Jochim, A.L. and Arora, P.S. (2009). Assessment of helical interfaces in protein–protein interactions. *Mol. Biosyst.* 5: 924–926.
- Jones, D.T. and Cozzetto, D. (2015). DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* 31: 857–863.
- Kamigaki, A., Nito, K., Hikino, K., Goto-Yamada, S., Nishimura, M., Nakagawa, T., and Mano, S. (2016). Gateway Vectors for Simultaneous Detection of Multiple Protein–Protein Interactions in Plant Cells Using Bimolecular Fluorescence Complementation. *PLoS ONE* 11: e0160717.
- Katoh, K. and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30: 772–780.
- Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* 10: 845–858.
- Kumar, S., Stecher, G., Suleski, M., and Hedges, S.B. (2017). TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol Biol Evol* 34: 1812–1819.
- Lamiabie, A., Thévenet, P., Rey, J., Vavrusa, M., Derreumaux, P., and Tufféry, P. (2016). PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res* 44: W449–W454.
- van der Lee, R. et al. (2014). Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.* 114: 6589–6631.
- Lemke, E.A. (2016). The Multiple Faces of Disordered Nucleoporins. *Journal of Molecular Biology* 428: 2011–2024.
- Letunic, I. and Bork, P. (2011). Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucl. Acids Res.* 39: W475–W478.
- Lin, D.H. et al. (2016). Architecture of the symmetric core of the nuclear pore. *Science* 352: aaf1015–aaf1015.
- Lisitsyna, O.M., Kurnaeva, M.A., Arifulin, E.A., Shubina, M.Y., Musinova, Y.R., Mironov, A.A., and Sheval, E.V. (2020). Origin of the nuclear proteome on the basis of pre-existing nuclear localization signals in prokaryotic proteins. *Biology Direct* 15: 9.
- Meier, I., Richards, E.J., and Evans, D.E. (2017). Cell Biology of the Plant Nucleus. *Annual Review of Plant Biology* 68: 139–172.
- Mikulski, P., Hohenstatt, M.L., Farrona, S., Smaczniak, C., Stahl, Y., Kalyanikrishna, Kaufmann, K., Angenent, G., and Schubert, D. (2019). The Chromatin-Associated Protein PWO1 Interacts with Plant Nuclear Lamin-like Components to Regulate Nuclear Size. *Plant Cell* 31: 1141–1154.
- Nakagawa, T. et al. (2007). Improved Gateway Binary Vectors: High-Performance Vectors for Creation of Fusion Constructs in Transgenic Analysis of Plants. *Bioscience, Biotechnology, and Biochemistry* 71: 2095–2100.
- Neumann, N., Jeffares, D.C., and Poole, A.M. (2007). Outsourcing the Nucleus: Nuclear Pore Complex Genes are no Longer Encoded in Nucleomorph Genomes. *Evol Bioinform Online* 2: 23–34.
- Norkunas, K., Harding, R., Dale, J., and Dugdale, B. (2018). Improving agroinfiltration-based transient

- gene expression in *Nicotiana benthamiana*. *Plant Methods* 14: 71.
- Nützmann, H.-W., Doerr, D., Ramírez-Colmenero, A., Sotelo-Fonseca, J.E., Wegel, E., Stefano, M.D., Wingett, S.W., Fraser, P., Hurst, L., Fernandez-Valverde, S.L., and Osbourn, A. (2020). Active and repressed biosynthetic gene clusters have spatially distinct chromosome states. *PNAS* 117: 13800–13809.
- Peng, K., Radivojac, P., Vucetic, S., Dunker, A.K., and Obradovic, Z. (2006). Length-dependent prediction of protein intrinsic disorder. *BMC Bioinformatics* 7: 208.
- Pietrosemoli, N., García-Martín, J.A., Solano, R., and Pazos, F. (2013). Genome-Wide Analysis of Protein Disorder in *Arabidopsis thaliana*: Implications for Plant Environmental Adaptation. *PLoS One* 8: e55524.
- Pontvianne, F. and Liu, C. (2020). Chromatin domains in space and their functional implications. *Current Opinion in Plant Biology* 54: 1–10.
- Poulet, A., Duc, C., Voisin, M., Desset, S., Tutois, S., Vanrobays, E., Benoit, M., Evans, D.E., Probst, A.V., and Tatout, C. (2017). The LINC complex contributes to heterochromatin organisation and transcriptional gene silencing in plants. *J Cell Sci* 130: 590–601.
- Rossignol, P., Collier, S., Bush, M., Shaw, P., and Doonan, J.H. (2007). *Arabidopsis* POT1A interacts with TERT-V(18), an N-terminal splicing variant of telomerase. *Journal of Cell Science* 120: 3678–3687.
- Sakamoto, Y. et al. (2020). Subnuclear gene positioning through lamina association affects copper tolerance. *Nat Commun* 11: 5914.
- Shumaker, D.K., Dechat, T., Kohlmaier, A., Adam, S.A., Bozovsky, M.R., Erdos, M.R., Eriksson, M., Goldman, A.E., Khuon, S., Collins, F.S., Jenuwein, T., and Goldman, R.D. (2006). Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proceedings of the National Academy of Sciences* 103: 8703–8708.
- Smith, S., Galinha, C., Desset, S., Tolmie, F., Evans, D., Tatout, C., and Graumann, K. (2015). Marker gene tethering by nucleoporins affects gene expression in plants. *Nucleus* 6: 471–478.
- Smythe, C., Jenkins, H.E., and Hutchison, C.J. (2000). Incorporation of the nuclear pore basket protein Nup153 into nuclear pore structures is dependent upon lamina assembly: evidence from cell-free extracts of *Xenopus* eggs. *The EMBO Journal* 19: 3918–3931.
- Sparkes, I.A., Runions, J., Kearns, A., and Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protocols* 1: 2019–2025.
- van Steensel, B. and Belmont, A.S. (2017). Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* 169: 780–791.
- Sundell, D., Mannapperuma, C., Netotea, S., Delhomme, N., Lin, Y.-C., Sjödin, A., Peer, Y.V. de, Jansson, S., Hvidsten, T.R., and Street, N.R. (2015). The Plant Genome Integrative Explorer Resource: PlantGenIE.org. *New Phytologist* 208: 1149–1156.
- Tamura, K. (2020). Nuclear pore complex-mediated gene expression in *Arabidopsis thaliana*. *J Plant Res* 133: 449–455.
- Tamura, K., Fukao, Y., Hatsugai, N., Katagiri, F., and Hara-Nishimura, I. (2017). Nup82 functions redundantly with Nup136 in a salicylic acid-dependent defense response of *Arabidopsis thaliana*. *Nucleus* 8: 301–311.
- Tamura, K., Fukao, Y., Iwamoto, M., Haraguchi, T., and Hara-Nishimura, I. (2010). Identification and Characterization of Nuclear Pore Complex Components in *Arabidopsis thaliana*. *The Plant Cell Online* 22: 4084–4097.
- Tamura, K. and Hara-Nishimura, I. (2011). Involvement of the nuclear pore complex in morphology of the plant nucleus. *Nucleus* 2: 168–172.
- Tang, Y., Dong, Q., Wang, T., Gong, L., and Gu, Y. (2022). PNET2 is a component of the plant nuclear lamina and is required for proper genome organization and activity. *Developmental Cell* 57: 19–31.e6.
- Tang, Y., Huang, A., and Gu, Y. (2020). Global profiling of plant nuclear membrane proteome in *Arabidopsis*. *Nat. Plants* 6: 838–847.
- The UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research* 47: D506–D515.
- Tomba, P., Davey, N.E., Gibson, T.J., and Babu, M.M. (2014). A Million Peptide Motifs for the Molecular Biologist. *Molecular Cell* 55: 161–169.
- Van Bel, M., Diels, T., Vancaester, E., Kreft, L., Botzki, A., Van de Peer, Y., Coppens, F., and Vandepoele, K. (2018). PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Res* 46: D1190–D1196.
- Wang, H., Dittmer, T.A., and Richards, E.J. (2013). *Arabidopsis* CROWDED NUCLEI (CRWN) proteins are required for nuclear size control and heterochromatin organization. *BMC plant biology* 13: 200.
- Xu, D. and Zhang, Y. (2012). Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins: Structure, Function, and Bioinformatics* 80: 1715–1735.
- Yang, Y., Wang, W., Chu, Z., Zhu, J.-K., and Zhang, H. (2017). Roles of Nuclear Pores and Nucleocytoplasmic Trafficking in Plant Stress Responses. *Front. Plant Sci.* 8: 574.
- Yang, Z. (2007). PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol Biol Evol* 24: 1586–1591.
- Zhang, L. et al. (2020). The water lily genome and the early evolution of flowering plants. *Nature* 577: 79–84.
- Zhao, B. et al. (2021). DescribePROT: database of amino acid-level protein structure and function predictions. *Nucleic Acids Res* 49: D298–D308

Les données supplémentaires sont disponibles en Annexe 3

Ensemble des gènes dérégulés

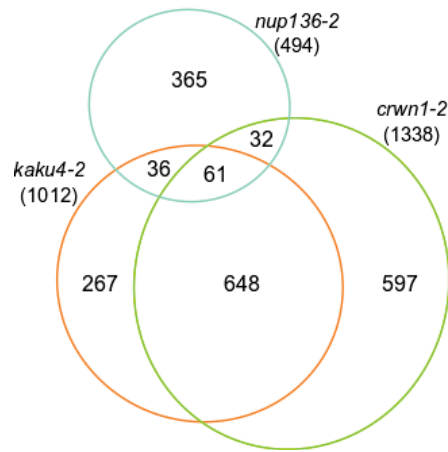
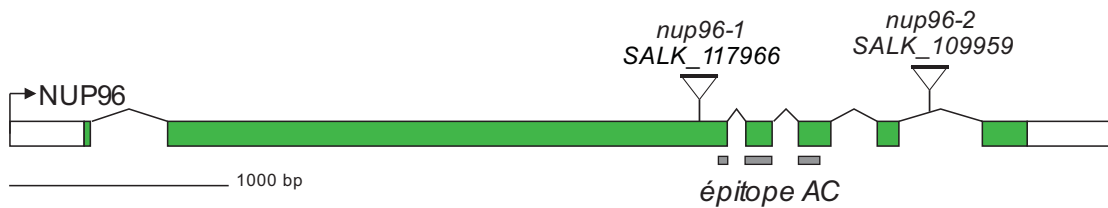


Figure 2.1 : KAKU4 et CRWN1 dérégulent un ensemble de gènes communs

Diagramme de Venn de l'ensemble des gènes dérégulés dans les mutants *kaku4-2*, *crwn1-2* et *nup136-2* par rapport au sauvage. La taille de chaque ellipse est proportionnelle au nombre de gènes dérégulés dans nos données de RNA-seq ($FC > 2$ ou $FC < 0.5$; $padj < 0.05$).

A



B

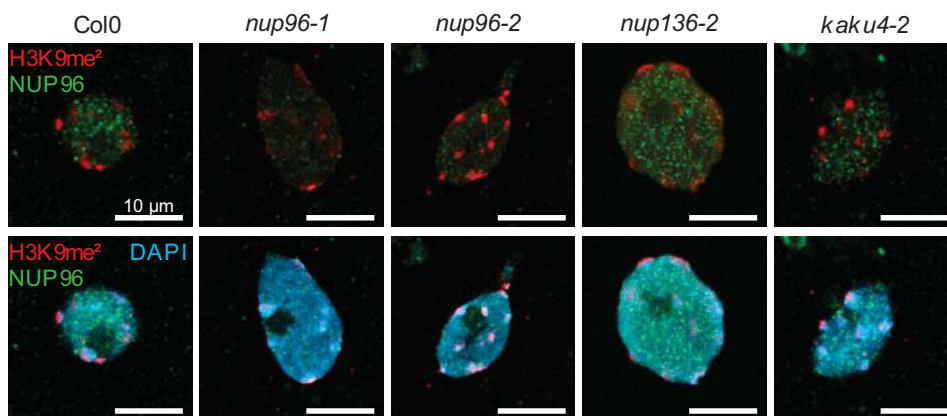


Figure 2.2 : Essai avec l'anticorps anti-NUP96 (Abcam) d'analyse de la distribution des pores nucléaires

A. Localisation des insertions des T-DNA des différents mutants *nup96* ainsi que de la région contenant l'épitope de l'anticorps anti-NUP96 (abcam124980) utilisé en immunofluorescence sur le gène *NUP96*. **B.** Immunofluorescence de noyaux isolés de plantules de 7 jours avec les anticorps anti-NUP96 (vert) pour marquer les pores nucléaires et anti-H3K9me² (ab1220, rouge) pour marquer l'hétérochromatine et marquage de l'ADN au DAPI (cyan), anticorps primaire utilisé au 200^e et anticorps secondaire utilisé au 1000^e ; échelle : 10μm.

I.1.c) Perspectives

I.1.c.i) Analyse transcriptomique des mutants *nup136*, *kaku4* et *crwn1*

Dans notre article nous avons montré qu'il existe une connexion entre le nucléosquelette et le pore complexe et que les mutants de ces composants jouent un rôle dans le développement de la plante et la morphologie nucléaire et les récents articles de l'équipe de Chang Liu démontrent que certaines régions de chromatine sont localisées préférentiellement à la périphérie du noyau que ce soit à proximité de NUP136 (Bi *et al.*, 2017) ou des protéines CRWN (Hu *et al.*, 2019). Nous avons donc cherché à savoir quel est l'impact des protéines NUP136, KAKU4 et CRWN1 sur la régulation de l'expression des gènes, en réalisant plusieurs analyses transcriptomiques. J'ai choisi de présenter ici notre dernière analyse réalisée sur les mutants *nup136-2*, *kaku4-2* et *crwn1-2* via un séquençage 150 bp en paired-end, des ARN polyA extraits à partir de la partie caulinaire de plantules de 14 jours poussées sur milieu synthétique. Les données ont été alignées avec HiSat2 sur le génome de TAIR10 et les gènes différentiellement exprimés par rapport au sauvage (Col0) ont été déterminés avec Deseq2. Nous avons considéré les gènes dérégulés comme ayant un Fold Change supérieur à 2 pour les gènes sur-exprimés et inférieur à 0,5 pour les gènes sous-exprimés avec une p-value inférieure à 0,05 par rapport au sauvage. Les mutants *kaku4-2* et *crwn1-2* possèdent de nombreux gènes dérégulés en commun mais peu avec le mutant *nup136-2* (**Figure 2.1**), suggérant que les protéines CRWN1 et KAKU4 sont impliquées dans des fonctions associées au nucléosquelette, alors que NUP136 est impliquée dans des fonctions propres au pore nucléaire. Cette analyse sera développée dans le paragraphe [II.3](#)).

I.1.c.ii) Influence du nucléosquelette sur la distribution des pores nucléaires

L'une des principales conclusions de notre article est qu'il existe une connexion entre le nucléosquelette et le pore complexe. Quel pourrait-être le rôle fonctionnel de cette connexion ? Une première hypothèse serait que ces interactions seraient nécessaires pour organiser la distribution des pores nucléaires à la surface de l'enveloppe nucléaire. Une seconde hypothèse serait que cette connexion participerait à l'ancrage du nucléosquelette à la périphérie nucléaire.

Afin de tester notre première hypothèse, nous avons décidé en première approche, de réaliser de l'immunofluorescence sur des noyaux isolés de plantules de 7 jours avec l'anticorps commercial d'abcam (ab124980), qui est un anticorps dirigé contre la région 1600 à 1700 AA de la protéine humaine NUP98 qui correspond en réalité à la protéine NUP96 d'*Arabidopsis thaliana*, en effet chez l'Homme ces deux protéines sont produites par un même gène dont le

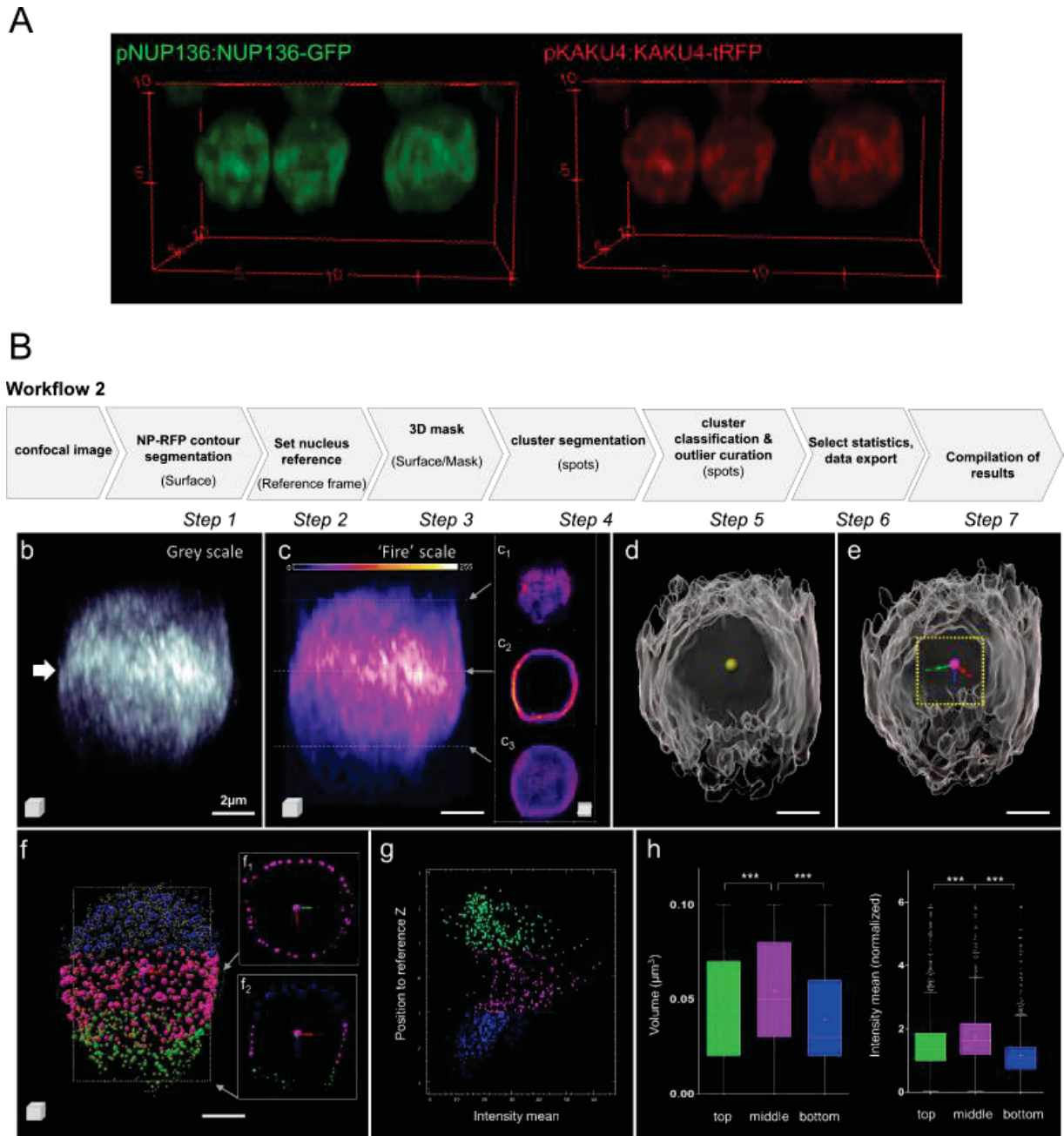


Figure 2.3 : Etude de la répartition nucléaire de NUP136 et KAKU4

A. Vue 3D de noyaux de l'apex racinaire d'*A. thaliana*, co-exprimant pNUP136-NUP136-GFP (vert) et pKAKU4-KAKU4-tRFP (rouge). **B.** Description du workflow imaris permettant l'analyse de la répartition du signal selon l'axe z des images en 3D. Les spots sont répartis en trois zones de taille similaire selon l'axe Z du noyau : haute, centrale et basse, déterminées à partir d'un repère placé au centre du noyau. La zone centrale correspond à la région allant de $-2\ \mu\text{m}$ à $+2\ \mu\text{m}$ par rapport à ce repère et tout ce qui est au-dessus de $+2\ \mu\text{m}$ et en dessous de $-2\ \mu\text{m}$ sont définis comme zone haute et zone basse respectivement. Le volume de ces spots ainsi que leur intensité peuvent être quantifiés en fonction des différentes zones et ainsi montrer que la zone centrale diffère des deux autres en ayant des spots plus gros et plus intenses cf Annexe 4.

produit protéique est clivé pour former 2 nucléoporines distinctes NUP98 et NUP96, ces 2 facteurs étant produits *via* des gènes distincts chez les plantes (Fontoura, Blobel and Matunis, 1999). Nous avons réalisé des tests de cet anticorps commercial sur des noyaux isolés de deux mutants *nup96*, *nup96-1* et *nup96-2* ; l'épitope de l'anticorps n'étant pas connu (malgré nos requêtes auprès du fournisseur), nous avons essayé d'estimer sa zone d'hybridation sur la protéine NUP96 d'*Arabidopsis* présente dans l'anneau externe du NPC (**Figure 2.2A**). Nous avons certes observé moins de signal dans le mutant *nup96-1* que dans le mutant *nup96-2*, cependant le signal d'hybridation se localise aussi bien à l'intérieur du noyau qu'à la périphérie nucléaire (**Figure 2.2B**). Cela a également été observé dans d'autres mutants comme *nup136-2* et *kaku4-2*, ce qui suggère que cet anticorps ne serait pas spécifique du pore nucléaire mais reconnaîtrait également une ou plusieurs protéine(s) nucléoplasmique(s) chez *A. thaliana* (**Figure 2.2B**). A ce stade, nous n'avons donc pas pu vérifier la validité de notre hypothèse et il nous faudrait tester des anticorps reconnaissant plus efficacement le pore complexe.

Une autre manière d'étudier la distribution des pores nucléaires est d'observer la localisation de la protéine NUP136. Dans un premier temps, nous avons développé une routine d'analyse d'images en utilisant le logiciel IMARIS avec l'aide de Célia Baroux (Université de Zurich) afin de quantifier l'intensité et le nombre de spots dans un noyau 3D exprimant NUP136 ou KAKU4 sous forme de protéines fluorescentes. Au cours des expériences d'expression transitoire chez le tabac de la protéine NUP136 coexprimée avec KAKU4 (p35S:GFP-NUP136 et pKAKU4:KAKU4-tRFP), lors des acquisitions en 3D nous observons que les spots de NUP136 localisés à la périphérie nucléaire sont plus nombreux et plus intenses sur le plan équatorial des noyaux formant un anneau que nous avons nommé anneau équatorial (**Figure 2.3A**). La routine IMARIS mise en place procède en trois étapes distinctes : i) segmentation du noyau ii) segmentation des spots de fluorescence et iii) analyse de la répartition des spots en trois zones de taille similaire selon l'axe Z du noyau : haute, centrale et basse, déterminées à partir d'un repère placé au centre du noyau (**Figure 2.3B**). Cette routine fait l'objet d'un article en cours de révision dans la revue *Nucleus* (cf **Annexe 4**). Cette analyse d'image initialement réalisée sur quelques noyaux de tabac exprimant transitoirement la protéine *NUP136-GFP*, a maintenant été appliquée avec l'aide d'Aurélia Stevens sur un plus grand nombre de noyaux issus de l'apex racinaire dans des lignées *Arabidopsis thaliana nup136-2 ; pNUP136:NUP136-GFP*.

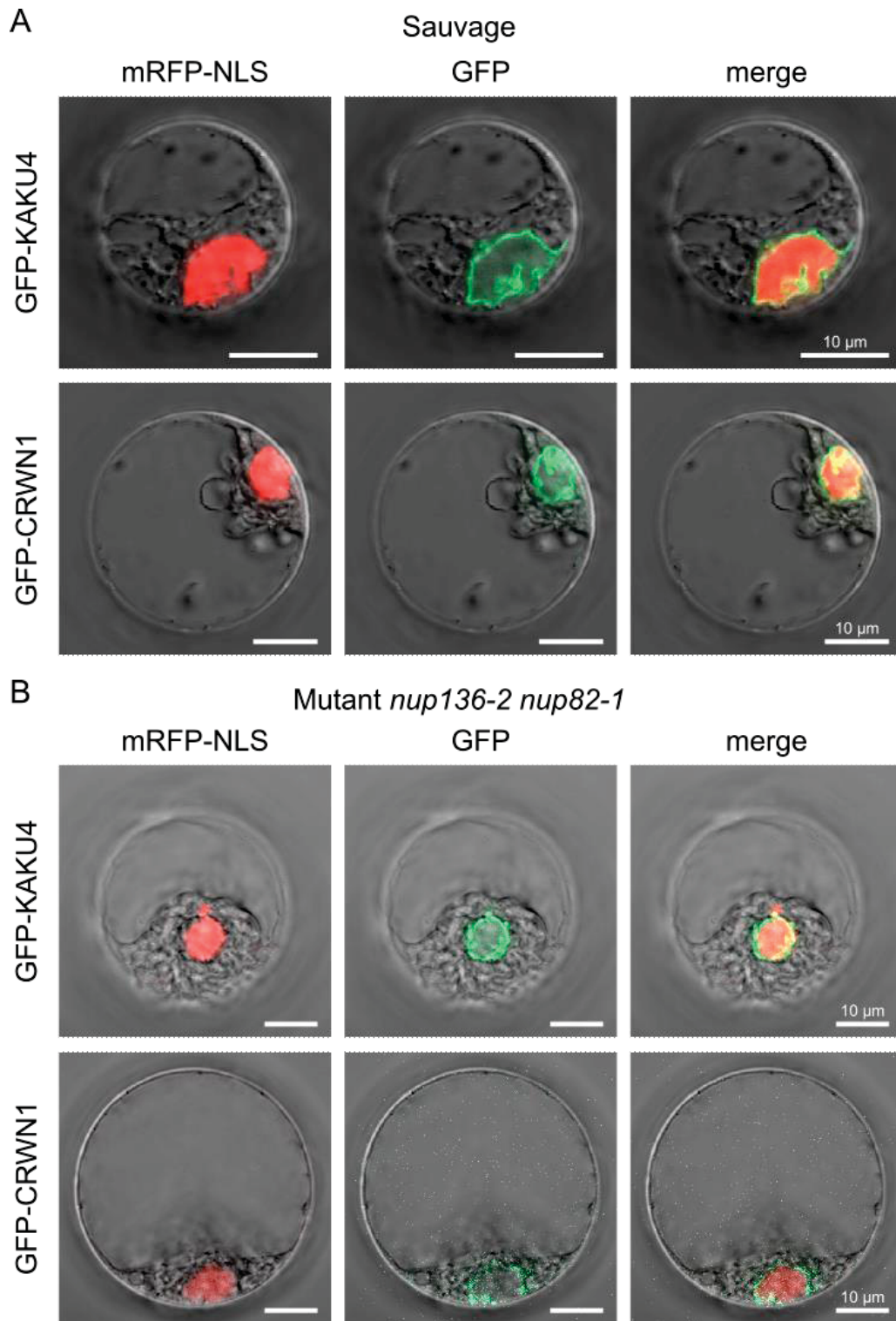


Figure 2.4 : Localisation de CRWN1 et KAKU4 dans des protoplastes racinaires d'*Arabidopsis thaliana*

Co-expression de KAKU4 ou CRWN1 (en vert) avec la NLS (en rouge) dans des protoplastes racinaires sauvages (**A**) et mutants *nup136-2 nup82-1* (**B**) ; vecteurs utilisés pGWB606 et pGWB555 permettant, sous le contrôle d'un promoteur 35S, des fusions traductionnelles en N-terminal, respectivement avec la GFP ou la mRFP, échelle : 10µm.

Dans les deux expériences, nous observons le même enrichissement de l'intensité de fluorescence de la construction NUP136::GFP au niveau de la zone centrale correspondant à l'anneau équatorial. Cet anneau équatorial pourrait correspondre à un enrichissement du nucléosquelette nécessaire à la division cellulaire. En effet, il est connu que l'enveloppe nucléaire joue le rôle d'organisateur des microtubules (*Microtubule Organization Center*, MTOC) (Tatout 2014, Chr research) et que les microtubules forment un anneau au moment de la mitose avant la rupture de l'enveloppe nucléaire. Cette organisation équatoriale pourrait donc rappeler l'organisation particulière des microtubules. Pour tester cela, nous pourrions donc essayer de perturber l'organisation des microtubules ou plus simplement, d'imager des noyaux situés plus haut dans la racine lorsque les cellules ne se divisent plus.

I.1.c.iii) Influence du nuclear basket sur la localisation des composants du nucléosquelette

Afin de tester notre seconde hypothèse, comme quoi la connexion entre le nucléosquelette et le pore nucléaire participerait à l'ancrage du nucléosquelette à la périphérie nucléaire, nous avons cherché à savoir si l'absence de nos deux nucléoporines NUP82 et NUP136 impacteraient la localisation des protéines CRWN et KAKU4.

Nous avons décidé d'étudier la localisation des protéines CRWN1 et KAKU4 dans un contexte mutant pour *NUP136* et *NUP82* via l'expression transitoire dans des protoplastes racinaires. Pour cela, j'ai appris, dans le laboratoire de Martina Dvořáčková (Université Masaryk de Brno, République Tchèque, à isoler et à transformer des protoplastes de racines d'*Arabidopsis* selon la méthode utilisée par (Kolářová *et al.*, 2021). J'ai appliqué cette méthode au laboratoire en exprimant les protéines CRWN1 et KAKU4 sous forme de protéines fluorescentes dans des protoplastes racinaires de plantes sauvages et mutantes pour *NUP136* et *NUP82*, de 7 jours cultivées sur milieu synthétique. En contexte sauvage, les protéines KAKU4 et CRWN1 se localisent à la périphérie nucléaire en formant des invaginations dans le noyau et c'est également le cas dans le mutant *nup136-2 nup82-1* (**Figures 2.4A-B**). Bien que nous ayons identifié des interactions entre les deux nucléoporines et les composants du nucléosquelette, ces interactions ne semblent pas, à elles seules, expliquer la localisation du nucléosquelette à la périphérie. L'hypothèse la plus vraisemblable est qu'il existe d'autres interactions du nucléosquelette avec la périphérie nucléaire, comme par exemple avec les protéines SUN (Graumann, 2014) ou d'autres protéines ancrées à l'INM comme MAN1 ou PNET2 (Tang *et al.*, 2021).

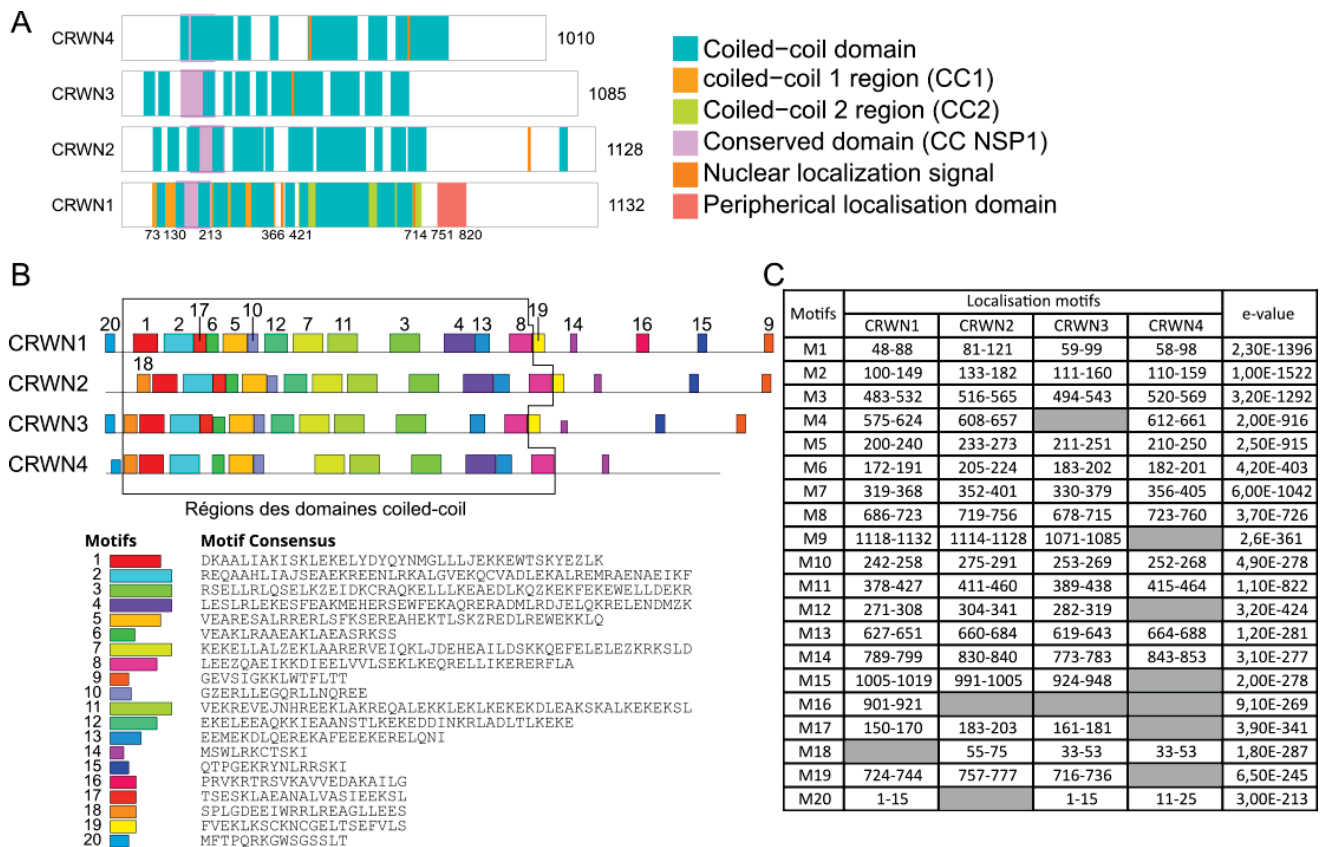


Figure 2.5 : Analyse MEME des protéines CRWN d'*Arabidopsis thaliana*

A. Représentation schématique des protéines CRWN. **B.** Localisation des motifs conservés trouvés lors de l'analyse MEME réalisée avec l'ensemble des orthologues des quatre protéines CRWN. Les séquences des motifs consensus obtenues pour chaque motif conservé sont représentées en dessous. **C.** Tableau récapitulatif de la localisation des différents motifs conservés retrouvés chez les quatre protéines CRWN ainsi que leur e-value.

Nos deux hypothèses suggérant que les interactions entre le NPC et le nucléosquelette seraient nécessaires pour organiser la distribution des pores nucléaires à la surface de l'enveloppe nucléaire et que cette connexion participerait l'ancrage du nucléosquelette à la périphérie nucléaire, se sont avérées difficiles à tester expérimentalement. L'une par manque d'outils (anticorps) et l'autre du fait de la multiplicité des interactions impliquant les protéines du nucléosquelette. A ce stade de notre travail, nous ne pouvons pas inclure ces résultats dans notre publication soumise à *The Plant Cell*.

I.2) Études de protéines CRWN et leurs connections à périphérie nucléaire

Afin d'approfondir sur les acteurs participant à la localisation du nucléosquelette à la périphérie nucléaire, nous avons cherché à cartographier les régions nécessaires aux interactions protéiques des protéines CRWN ainsi qu'à la localisation de ces protéines et identifier de nouveaux acteurs qui eux pourraient être ancrés à l'INM. Pour compléter ce travail nous avons testé les interactions génétiques entre nos gènes d'intérêt en créant une large collection de mutants impliquant *CRWN1*, *KAKU4* et *NUP136* puis en réalisant une étude phénotypique.

I.2.a) Structures des protéines CRWN

Les protéines CRWN sont les homologues fonctionnelles des lamines des mammifères et tout comme elles, possèdent des domaines coiled-coil dans leur région N-terminale (cf Introduction I.3.c.ii.a). La protéine CRWN1 contient des domaines coiled-coil dans sa partie N-terminale, entre 73 et 714 AA, qui ont été répartis en deux régions coiled-coil (coiled-coil 1 et 2, CC1 : 73-366 AA et CC2 : 421-714 AA, en orange et vert respectivement) (**Figure 2.5A**). Au sein de la région coiled-coil 1, le domaine de 130-213 AA correspond à un domaine coiled-coil retrouvé chez la nucléoporeine NSP1 de la levure *S. cerevisiae*, représenté en rose (**Figure 2.5A**). Ce domaine est également présent chez les autres protéines CRWN. La protéine CRWN1 possède également une région impliquée dans sa localisation à la périphérie nucléaire dans sa partie C-terminale au niveau des acides-aminés 751 à 820 (Guo *et al.*, 2017). Afin de déterminer les régions conservées au sein des protéines CRWN et ainsi d'affiner les régions pouvant être impliquées dans leur localisation ou leurs interactions, j'ai réalisé une analyse MEME de 20 motifs sur l'ensemble des orthologues des protéines CRWN présentes dans les mêmes espèces que celles utilisées dans notre publication pour les protéines KAKU4 et NUP (cf paragraphe [I.1](#)). Treize des motifs obtenus dans cette analyse correspondent à des domaines coiled-coil conservés (**Figures 2.5B-C**). Parmi les autres motifs, trois semblent jouer un rôle dans les

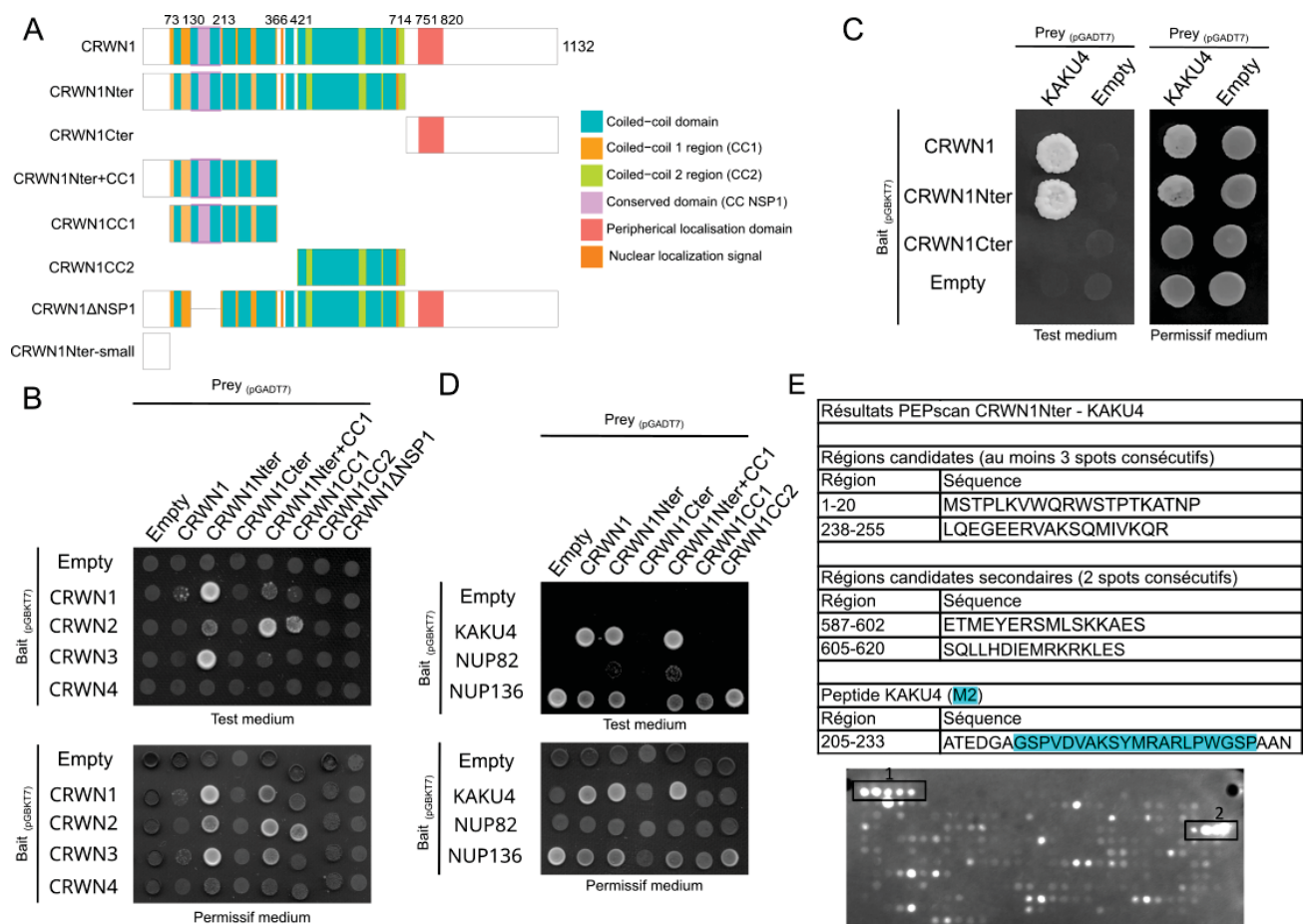


Figure 2.6 : Interactions de CRWN1 avec elle-même, les autres protéines CRWN, KAKU4 et NUP

A. Représentation schématique des différentes constructions de CRWN1 utilisées pour les différentes expériences. **B-D.** Test en goutte de double-hybride sur milieu test (SD/-Leu/-Trp/-Ade/-His) ou permissif (SD/-Leu/-Trp). Les vecteurs vides pGBKT7 et pGADT7 sont utilisés comme témoins négatifs appât et proie respectivement. Interactions entre CRWN1 et les protéines CRWN (**B**), KAKU4 (**C**) et KAKU4, NUP82 et NUP136 (**D**). Dans ces expériences, le vecteur pGBKT7-NUP136 a été démontré comme se comportant comme un auto-activateur. **E.** Tableau récapitulatif et résultat brut (en dessous) des résultats obtenus par la méthode de PEPscan réalisée entre la région N-terminale de CRWN1 (1-714) immobilisée sur la membrane et le motif M2 de KAKU4 (205-233) marqué par un fluorochrome. Les deux principales régions d'interactions (notées 1 et 2) sont indiquées en encadré sur le résultat d'hybridation.

Tableau 1 : Récapitulatif des interactions double-hybride entre CRWN1 (appât) et les l'ensemble des protéines CRWN et leurs dérivés (proies)

		Prey (pGADT7)											
		Empty	CRWN1	CRWN2	CRWN3	CRWN4	CRWN1 Nter	CRWN1Cter	CRWN1 Nter+CC1	CRWN1CC1	CRWN1CC2	CRWN1 Nter-small	CRWN1Δ NSP1
Bait (pGBKT7)	Empty	-	-	-	-	-	ND	ND	-	-	-	ND	ND
	CRWN1	-	++	++	-	-	ND	ND	++	+/-	-	ND	ND
	CRWN2	-	+/-	++	-	-	ND	ND	++	++	-	ND	ND
	CRWN3	-	++	++	-	-	ND	ND	++	-	-	ND	ND
	CRWN4	-	-	-	-	++	ND	ND	+/-	+/-	-	ND	ND
	CRWN1 Nter	-	-	-	-	-	ND	ND	++	++	-	ND	ND
	CRWN1 Cter	-	-	-	-	-	ND	ND	-	-	-	ND	ND
	CRWN1 Nter+CC1	-	-	-	-	-	ND	ND	++	++	-	ND	ND
	CRWN1 CC1	-	-	-	-	-	ND	ND	++	++	-	ND	ND
	CRWN1 CC2	-	-	-	-	-	ND	ND	-	-	-	ND	ND
	CRWN1 Nter-small	++	++	++	++	++	ND	ND	++	++	++	++	++
	CRWN1Δ NSP1	-	+/-	++	-	-	ND	ND	-	-	-	ND	++

++ : interaction forte, + : interaction, +/- : interaction faible, - : pas d'interaction, ND : interaction non-déterminée ; le vecteur pGBKT7-CRWN1 Nter-small est auto-activateur (cases grisées).

Tableau 2 et 3 : Récapitulatifs des interactions double-hybride entre CRWN1 et KAKU4-NUP

		Prey (pGADT7)								
		Empty	CRWN1	CRWN1 Nter	CRWN1 Cter	CRWN1 Nter+CC1	CRWN1 CC1	CRWN1 CC2	CRWN1 Nter-small	CRWN1Δ NSP1
Bait (pGBKT7)	Empty	-	-	-	-	-	-	-	ND	-
	KAKU4	-	++	++	-	++	+/-	-	ND	-
	NUP136	++	++	++	++	++	++	++	ND	++
	NUP82	-	+	++	-	-	+	-	ND	ND

		Prey (pGADT7)			
		Empty	KAKU4	NUP136	NUP82
Bait (pGBKT7)	Empty	-	-	-	-
	CRWN1	-	ND	+	ND
	CRWN1Nter	-	ND	ND	ND
	CRWN1Cter	-	ND	ND	ND
	CRWN1Nter+CC1	-	ND	ND	ND
	CRWN1CC1	-	ND	ND	ND
	CRWN1CC2	-	ND	ND	ND
	CRWN1Nter-small	++	++	++	++
	CRWN1ΔNSP1	-	ND	ND	ND

++ : interaction forte, + : interaction, +/- : interaction faible, - : pas d'interaction, ND : interaction non-déterminée ; les vecteurs pGBKT7-NUP136 et pGBKT7-CRWN1 Nter-small sont auto-activateurs

interactions des protéines CRWN ou leur localisation, il s'agit des motifs M10 et M20 pour les interactions et du motif M14 pour la localisation (cf paragraphe [I.2.c](#)) Les autres motifs identifiés n'ont pas été étudiés.

[I.2.b](#) Identification des interactions protéiques impliquant les protéines CRWN

Nous avons vu précédemment, que les 4 protéines CRWN interagissent entre-elles et avec la protéine KAKU4 pour former le nucléosquelette, ainsi qu'avec les deux nucléoporines NUP82 et NUP136 (cf paragraphe [I.1.b](#)) (Sakamoto and Takagi, 2013; Goto *et al.*, 2014; Sakamoto *et al.*, 2020; Mermet *et al.*, submitted). Afin de cartographier les régions de CRWN nécessaires à ces interactions, j'ai choisi la technique du double-hybride et réalisé différentes constructions à partir de la protéine CRWN1 (**Figure 2.6A**). J'ai montré que la région N-terminale de CRWN1 (1-714 AA) est nécessaire pour l'interaction de CRWN1 avec elle-même ainsi qu'avec CRWN2 et CRWN3 et ceci semble être médié par la région coiled-coil 1 de CRWN1 (1-366 AA) (**Figure 2.6B**). Malheureusement, aucun des croisements de levure portant les constructions CRWN1 délétées du domaine conservé NSP1 (CRWN1 Δ NSP1) n'ont fonctionné, cela aurait pu permettre d'affiner la région d'interaction au sein de la région coiled-coil 1 (**Figure 2.6B et Tableau 1**).

J'ai utilisé la même stratégie pour déterminer quelle région de CRWN1 est nécessaire aux interactions avec les protéines KAKU4, NUP82 et NUP136. Dans un premier temps, j'ai montré que la région N-terminale de CRWN1 (1-714 AA) est nécessaire à son interaction avec KAKU4 (**Figure 2.6C**). Cette même région est également essentielle pour interagir avec NUP82 (**Figure 2.6D et Tableaux 2-3**) cependant, dans ce cas-ci, la région coiled-coil 1 seule n'est pas suffisante, la région N-terminale en amont (CRWN1 Nter+CC1, 1-366 AA) est requise pour l'interaction avec NUP82 et KAKU4 (**Figure 2.6D et Tableaux 2-3**).

Pour affiner ces résultats, nous avons également réalisé une expérience de PEPscan entre le motif M2 de KAKU4, connu pour interagir avec la protéine CRWN1, et la région Nter de CRWN1 que nous avons défini par double hybride comme essentielle à cette interaction. Cette méthode permet de tester des interactions protéine-protéine en découpant une protéine (ici CRWN1) en une série de fragments peptidiques chevauchants de taille constante, fixés sur une membrane de cellulose afin de tester la liaison de chaque fragment avec une protéine partenaire marquée par un fluorochrome (ici M2 de KAKU4) (Rebollo, Savier and Tuffery, 2021). Cette expérience a permis d'identifier deux régions candidates principales de CRWN1 interagissant avec le motif M2 de KAKU4 (au moins trois spots consécutifs), la première présente dans la

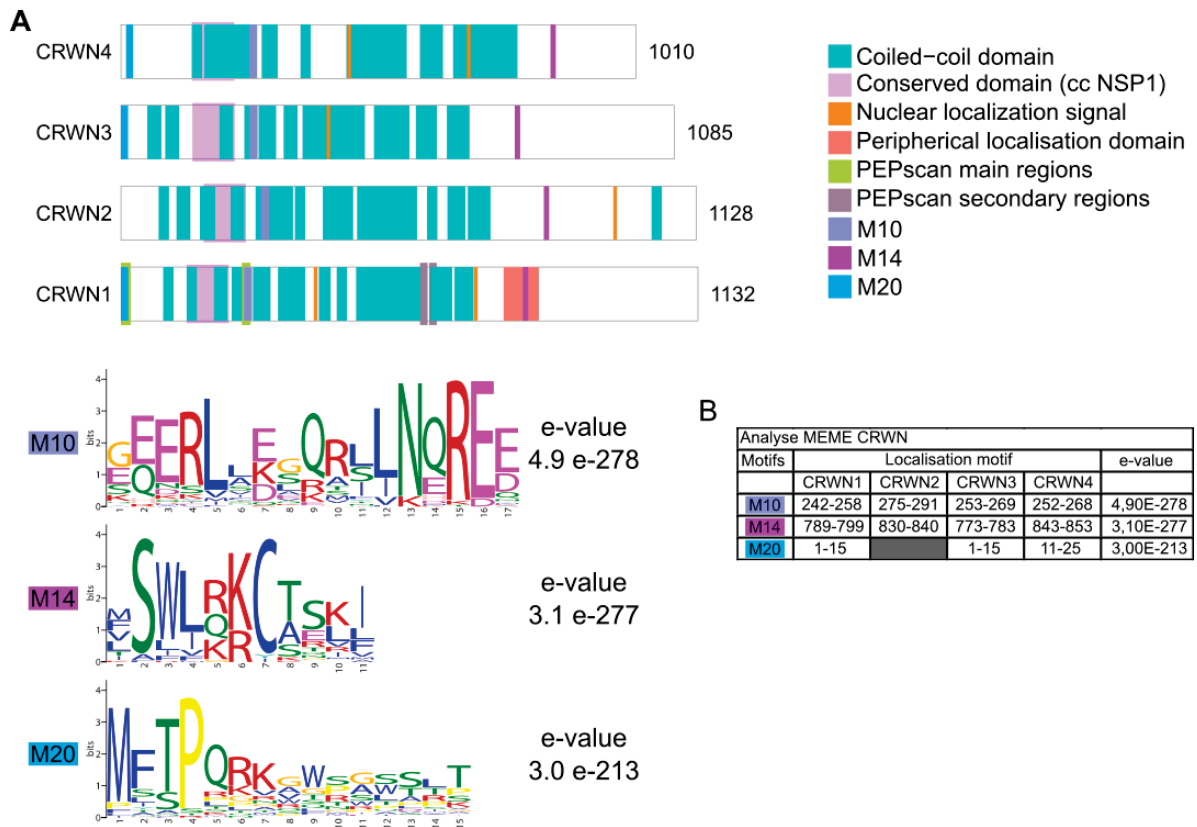


Figure 2.7 : Identification de motifs conservés dans les protéines CRWN chez *Arabidopsis thaliana*

A. Représentation schématique des protéines CRWN avec les régions identifiées par le PEPscan et certains des motifs conservés identifiés par l'analyse MEME. Les séquences des motifs consensus ainsi que la e-value obtenues pour chaque motif conservé sont représentées en dessous. **B.** Tableau récapitulatif de la localisation des motifs conservés retrouvés chez les quatre protéines CRWN, extrait du tableau de la Figure 2.4C.

Deux autres régions avec de plus faibles affinités ont également été identifiées et classées comme régions candidates secondaires (deux spots consécutifs), elles se trouvent dans la région coiled-coil 2 (587-602 AA et 605-620 AA) (**Figure 2.6E**).

La comparaison entre les régions de CRWN1 interagissant avec KAKU4 identifiées par le PEPscan et les motifs conservés de l'analyse MEME, a permis de mettre en évidence la présence d'un motif dans chacune des régions candidates déterminées lors du PEPscan. En effet, le motif M20 (1-15 AA) se situe dans la première région principale du PEPscan (1-20 AA) et le motif M10 (242-258 AA) est trouvé dans la seconde (**Figures 2.7A-B**). Ces deux motifs sont également retrouvés chez les autres protéines CRWN (sauf pour le premier motif qui n'est pas présent chez CRWN2). Ce qui suggère que ces deux motifs conservés joueraient un rôle dans les interactions entre les protéines CRWN et KAKU4.

Toutefois, les expériences de double-hybride, réalisées avec seulement la région en amont de la région coiled-coil 1 (CRWN1 Nter-small, 1-72 AA) n'ont pas permis de confirmer ces résultats. En effet, cette construction est auto-activatrice dans le vecteur pGBKT7 et les croisements de levures avec le vecteur pGADT7 n'ont pas fonctionné dans le cas des interactions avec CRWN1 (aucun zygote obtenu) et n'ont pas encore été réalisés pour les interactions avec KAKU4 et les NUP (**Tableaux 1-3**). Cela aurait permis de déterminer si le motif M20, présent dans cette région, joue un rôle dans les interactions des protéines CRWN1 avec KAKU4 et NUP82 et si les deux motifs, M20 : 1-15 AA et M10 : 242-258 AA, sont nécessaires pour qu'elles aient lieu. De plus, un autre motif conservé, le motif M4 (575-624 AA), a également été identifié au niveau des régions secondaires du PEPscan (587-602 AA et 605-620 AA) et les inclut. Tout comme nous avons identifié les motifs conservés de KAKU4 nécessaires à l'interaction avec les protéines CRWN (motif M2), nous pourrions poursuivre cette caractérisation en définissant les motifs de CRWN nécessaires à l'interaction avec KAKU4.

Nous avons donc montré que la région coiled-coil 1 est nécessaire aux interactions de CRWN1 avec elle-même et avec CRWN2 et CRWN3 et que cette région ainsi que la petite région N-ter en amont et notamment les deux motifs conservés (M10 et M20) joueraient un rôle dans les interactions des protéines CRWN avec KAKU4, NUP82 et très probablement NUP136 mais les expériences avec cette dernière n'ont pas été concluantes (**Tableau 3**). Les analyses des protéines KAKU4, NUP82, NUP136 et CRWN1 à 4 ont montré l'émergence de motifs conservés nécessaires aux interactions protéine-protéine, reste à savoir s'ils sont présents dans d'autres protéines ainsi que leur provenance.

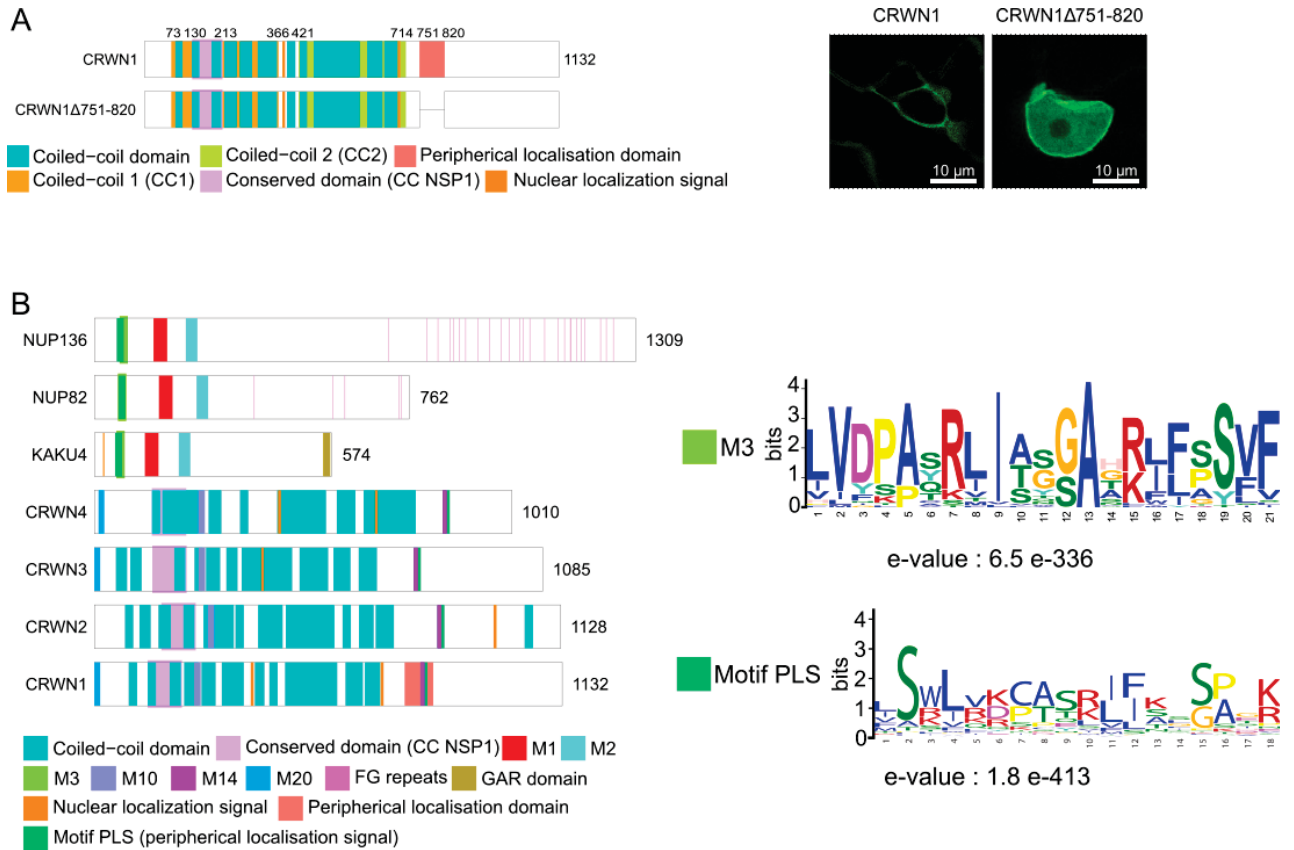


Figure 2.8 : Localisation de CRWN1 et présence de motifs conservés communs dans les protéines CRWN, KAKU4, NUP82 et NUP136

A. Représentation schématique des différentes constructions de CRWN1 utilisées pour étudier la localisation de CRWN1 en expression transitoire chez le tabac (*Nicotiana benthamiana*), clonées dans le vecteur pGWB606-p35S:GFP. Images confocales des différentes constructions, échelle : 10 μ m. **B.** Représentation schématique des protéines KAKU4, NUP82, NUP136 et CRWN1-4 d'*A.thaliana* et de leurs motifs conservés trouvés lors de l'analyse MEME réalisée avec l'ensemble des orthologues de ces protéines. Les séquences des motifs consensus M3 (KAKU4, NUP82 et NUP136) et PLS (KAKU4, NUP82, NUP136 et CRWN1-4) ainsi que la e-value obtenues pour chaque motif conservé sont représentées à droite.

Tableau 4 : Tableau récapitulatif de la localisation des motifs conservés retrouvés chez les protéines CRWN-NUP-KAKU4 d'*A.thaliana*

Motifs	Localisation motif							e-value
	KAKU4	NUP82	NUP136	CRWN1	CRWN2	CRWN3	CRWN4	
M3	54-74	58-78	62-82					6,5e-336
M14				789-799	830-840	773-783	843-853	3,10e-277
PLS	51-68	55-72	59-76	789-806	830-847	773-790	843-860	1,8e-413

Case grisée : analyse non réalisée

I.2.c) Détermination des domaines nécessaires à localisation des protéines CRWN

Les protéines CRWN se localisent à la périphérie nucléaire (Sakamoto and Takagi, 2013; Sakamoto *et al.*, 2020), cette localisation impliquerait la région 751-820 AA (Guo *et al.*, 2017). Lors de leur étude Guo *et al.* ont exprimé différentes régions de CRWN1 (1-769 AA, 1-820 AA ou encore 751-820 AA) en expression transitoire chez le tabac, mais les expériences réciproques de délétion n'ont pas été réalisées pour prouver sans ambiguïté le rôle essentiel de la région 751-820 AA pour la localisation à la périphérie nucléaire de CRWN1. Pour compléter ces résultats, j'ai réalisé une délétion de cette région de CRWN1 (751-820AA) impliquée dans sa localisation périphérique et montré qu'en absence de cette région, CRWN1 est nucléoplasmique avec tout de même un enrichissement à la périphérie nucléaire, établissant que la région identifiée par Guo *et al.* (751-820 AA) est impliquée dans la localisation périphérique de CRWN1 (**Figure 2.8A**).

Lors de la caractérisation de la localisation de KAKU4 (cf paragraphe I.1.b), nous avons montré que le motif conservé M3, également retrouvé chez les protéines NUP136 et NUP82, est impliqué dans la localisation à la périphérie nucléaire de KAKU4 (Mermet *et al.*, submitted). Pour mieux caractériser les régions de CRWN1 nécessaires à sa localisation, je me suis appuyée sur les résultats du MEME des protéines CRWN (cf paragraphe I.2.a) que j'ai comparés avec la région déterminée par Guo *et al.*, 2017. De manière intéressante, le motif M14 (789-799 AA) se situe dans cette région et est également présent chez les autres protéines CRWN (**Figures 2.7A-B**). La localisation de KAKU4 et des protéines CRWN semble chacune reposer sur un motif conservé, pour aller plus loin, j'ai effectué une analyse MEME avec les orthologues des protéines CRWN, KAKU4, NUP82 et NUP136. Cette analyse a été réalisée en stipulant que les motifs doivent être retrouvés dans toutes les protéines (oops : *one occurrence per sequence*) afin d'éviter d'avoir des motifs spécifiques à chaque famille de protéines. De manière surprenante, dans cette analyse, le motif M3 de KAKU4, NUP82 et NUP136 est également retrouvé chez les protéines CRWN au niveau du motif M14 identifié dans l'analyse des protéines CRWN, seules ces limites ont changé passant de 789-799 AA à 789-806 AA (**Figure 2.8B + Tableau 4**). Du fait de son implication dans la localisation périphérique de KAKU4, nous pensons que ce motif permettrait aux protéines de la périphérie nucléaire, CRWN, KAKU4, NUP82 et NUP136, de s'y localiser. Ainsi, j'ai renommé ce motif commun aux protéines KAKU4, NUP et CRWN, motif PLS (*Peripheral Localisation Signal*) ou signal de localisation à la périphérie nucléaire. Pour étudier le rôle de ce motif dans la localisation des

Tableau 5 : Récapitulatif des lignées mutantes créées

Croisement 1 : <i>nup136-2 nup82-1 x kaku4-2</i>
<i>nup136-2 kaku4-2</i>
<i>nup82-1 kaku4-2</i>
<i>nup136-2 nup82-1 kaku4-2</i>
Croisement 2 : <i>nup136-2 nup82-1 x kaku4-2ad crwn1-2</i>
<i>nup136-2 crwn1-2</i>
<i>nup82-1 crwn1-2</i>
<i>nup136-2 nup82-1 crwn1-2</i>
<i>nup136-2 kaku4-2ad crwn1-2</i>
<i>nup82-1 kaku4-2ad crwn1-2</i>
<i>nup136-2 nup82-1 kaku4-2ad crwn1-2</i>
Croisement 3 : <i>nup136-2 nup82-1 kaku4-2 x crwn2-1</i>
<i>nup136-2 crwn2-1</i>
<i>nup82-1 crwn2-1</i>
<i>nup136-2 nup82-1 crwn2-1</i>
<i>nup136-2 kaku4-2 crwn2-1</i>
<i>nup82-1 kaku4-2 crwn2-1</i>
<i>nup136-2 nup82-1 kaku4-2 crwn2-1</i>
Croisement 4 : <i>nup136-2 nup82-1 kaku4-2 x crwn3-1</i>
<i>nup136-2 crwn3-1</i>
<i>nup82-1 crwn3-1</i>
<i>nup136-2 nup82-1 crwn3-1</i>
<i>nup136-2 kaku4-2 crwn3-1</i>
<i>nup82-1 kaku4-2 crwn3-1</i>
<i>nup136-2 nup82-1 kaku4-2 crwn3-1</i>
Croisement 5 : <i>nup136-2 nup82-1 kaku4-2 x crwn4-1</i>
<i>nup136-2 crwn4-1</i>
<i>nup82-1 crwn4-1</i>
<i>nup136-2 nup82-1 crwn4-1</i>
<i>nup136-2 kaku4-2 crwn4-1</i>
<i>nup82-1 kaku4-2 crwn4-1</i>
<i>nup136-2 nup82-1 kaku4-2 crwn4-1</i>

ad : avec duplication (une duplication de 53kb sur le chromosome 4 a été identifiée dans le mutant *kaku4-2*).

protéines CRWN, la délétion du motif PLS (789-806 AA) a été initiée mais n'est pas finalisée pour le moment.

En conclusion, la localisation de la protéine CRWN1 nécessite la région de 751 à 820 AA dans laquelle a été identifié le motif PLS, commun aux protéines KAKU4, NUP82, NUP136 et CRWN et chevauchant le motif M3 de KAKU4, NUP82 et NUP136 impliqué dans la localisation à la périphérie nucléaire de KAKU4, reste à démontrer son rôle pour les protéines CRWN.

I.2.d) Interactions génétiques entre les gènes *NUP*, *KAKU4* et *CRWN*

Nous avons précédemment montré que les composants du nuclear basket et ceux du nucléosquelette interagissent physiquement les uns avec les autres (Mermet *et al.*, submitted). La suite logique est donc de déterminer si ces facteurs sont impliqués dans les mêmes processus et de ce fait s'ils interagissent fonctionnellement. Pour cela, nous avons étudié les interactions génétiques entre ces protéines en réalisant des multi-mutants et en étudiant leur impact sur le développement de la plante à différents stades ainsi que sur la morphologie nucléaire.

I.2.d.i) Collection de mutants

Nous disposons au laboratoire d'une collection de mutants d'insertions de T-DNA de l'ensemble des gènes codant pour ces facteurs, obtenus soit à partir du stock center, soit envoyés par Eric Richards (Institut Boyce Thompson, USA) ou par Kentaro Tamura (Université de Shizuoka, Japon) (**Annexe 1**). Les allèles sélectionnés sont tous dans le fond génétique Col-0. A partir des mutants initiaux, j'ai réalisé différents croisements afin d'obtenir l'ensemble des doubles, triples et quadruples mutants impliquant les gènes *NUP136*, *NUP82*, *KAKU4* et *CRWN1* à *CRWN4*, représentant un ensemble de 27 mutants différents (**Tableau 5**). A noter qu'à l'issue de notre première expérience de RNA-Seq, nous avons détecté dans le mutant *kaku4-2* une duplication présente sur le chromosome 4 (la caractérisation de cette duplication est présentée en paragraphe II.1.b). Cette duplication a, à notre connaissance, peu d'impact phénotypique, Kentaro Tamura a observé une senescence précoce. Afin de s'affranchir d'un éventuel effet de cette duplication, nous avons par croisement sélectionné une nouvelle lignée *kaku4-2* sans duplication. Certaines combinaisons de mutants plus anciennes possèdent encore cette duplication et sont notées *kaku4-2 ad* (avec duplication). Parmi les lignées obtenues, le double mutant *nup136-2 kaku4-2* a fait l'objet d'une publication (cf paragraphe [I.1.b](#)). A ce stade, nous n'avons réalisé qu'une étude phénotypique complète sur les lignées combinant les mutations des gènes *NUP136*, *KAKU4* et *CRWN1*.

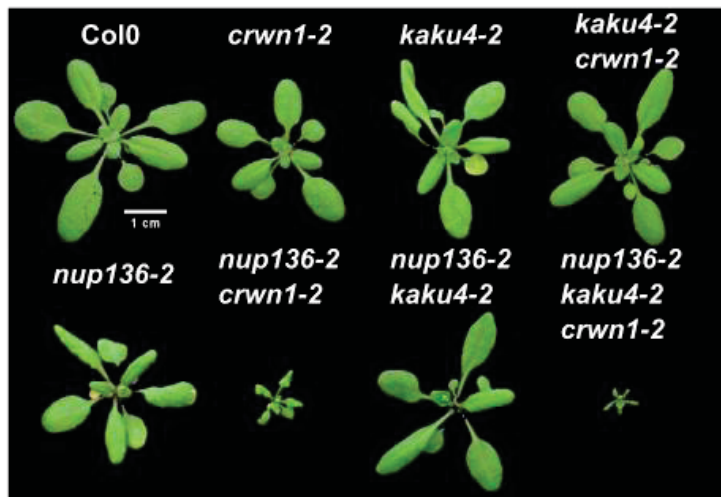
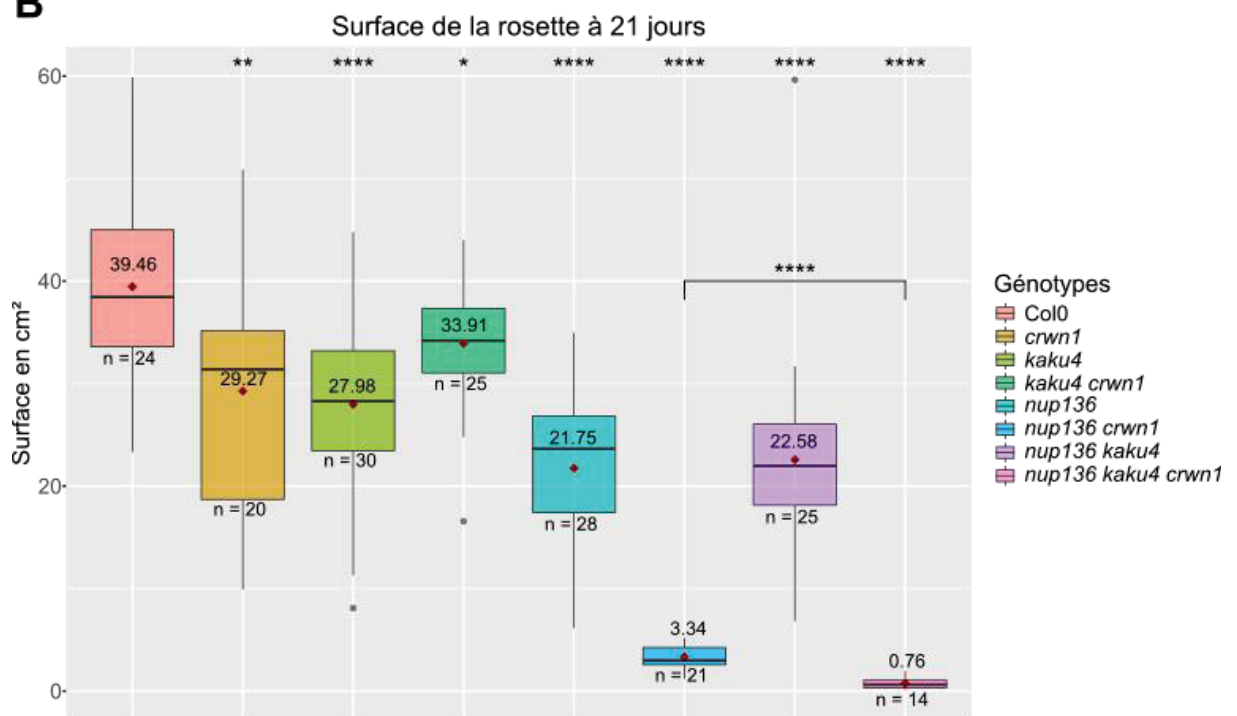
A**B**

Figure 2.9 : Etude de la taille de la rosette dans les différents mutants *nup136*, *kaku4* et *crwn1*

A. Photos de la rosette des différents mutants à 21 jours, échelle : 1 cm. **B.** Boxplot de la surface de la rosette à 21 jours à partir des données obtenues après traitement des photos par le logiciel Aradeepopsis (<https://github.com/Gregor-Mendel-Institute/aradeepopsis>, Hüther *et al.*, 2020) avec la moyenne affichée et représentée par un losange rouge ; n : nombre d'individus analysés par génotype ; Wilcoxon test, ns : non significatif (p -value > 0.05), * < 0.5, ** < 0.01, *** < 0.001, **** < 0.0001.

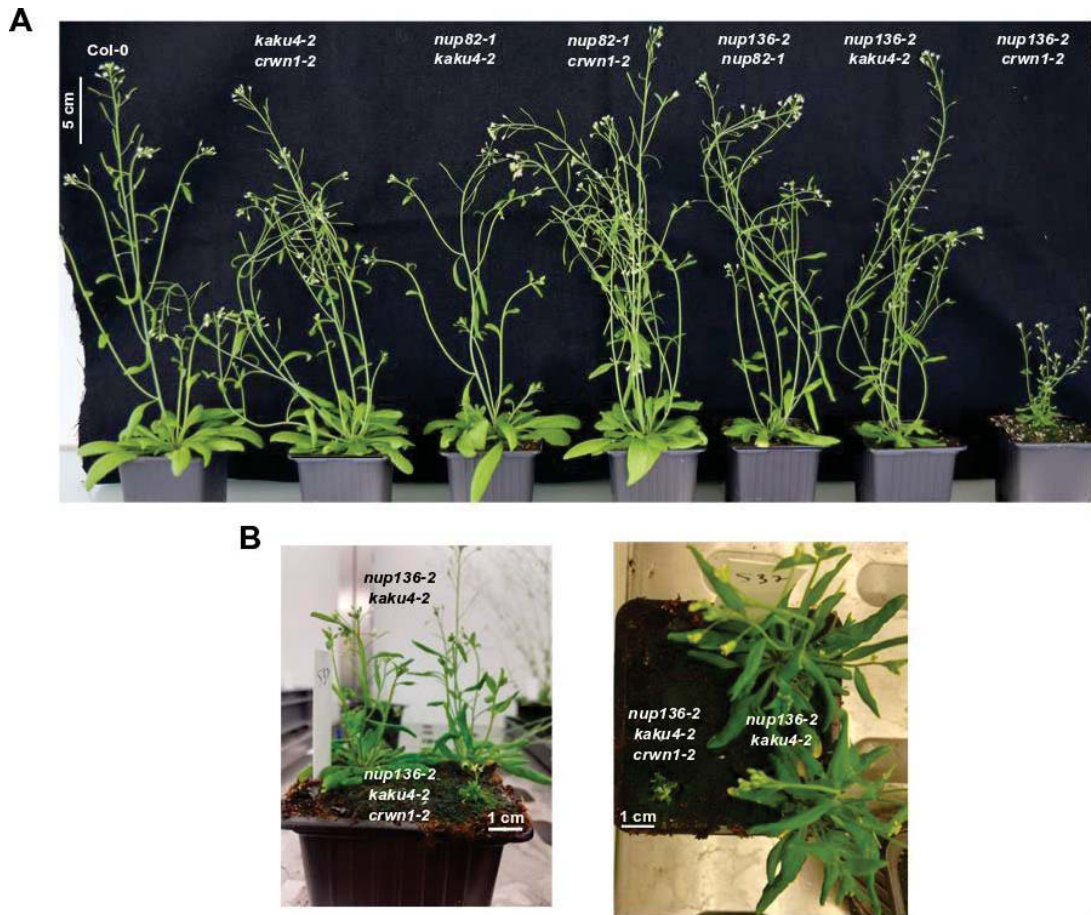


Figure 2.10 : Taille des différents mutants à un stade tardif du développement

Photos des différents doubles mutants (A) et du triple mutant en (B) ; échelle : 5 cm (A), 1 cm (B).

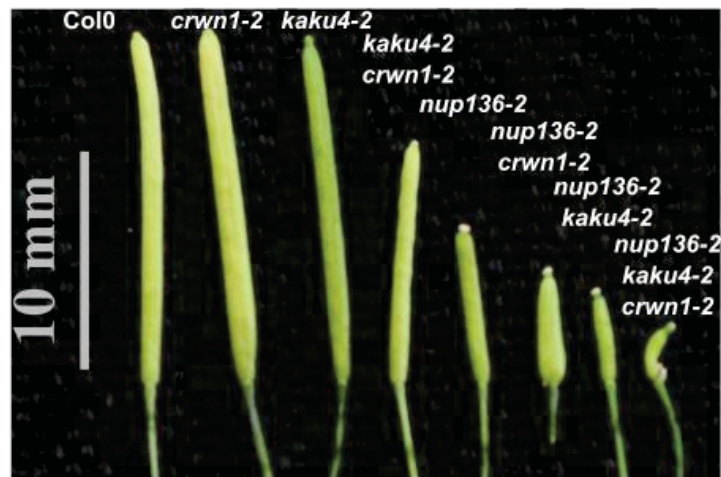
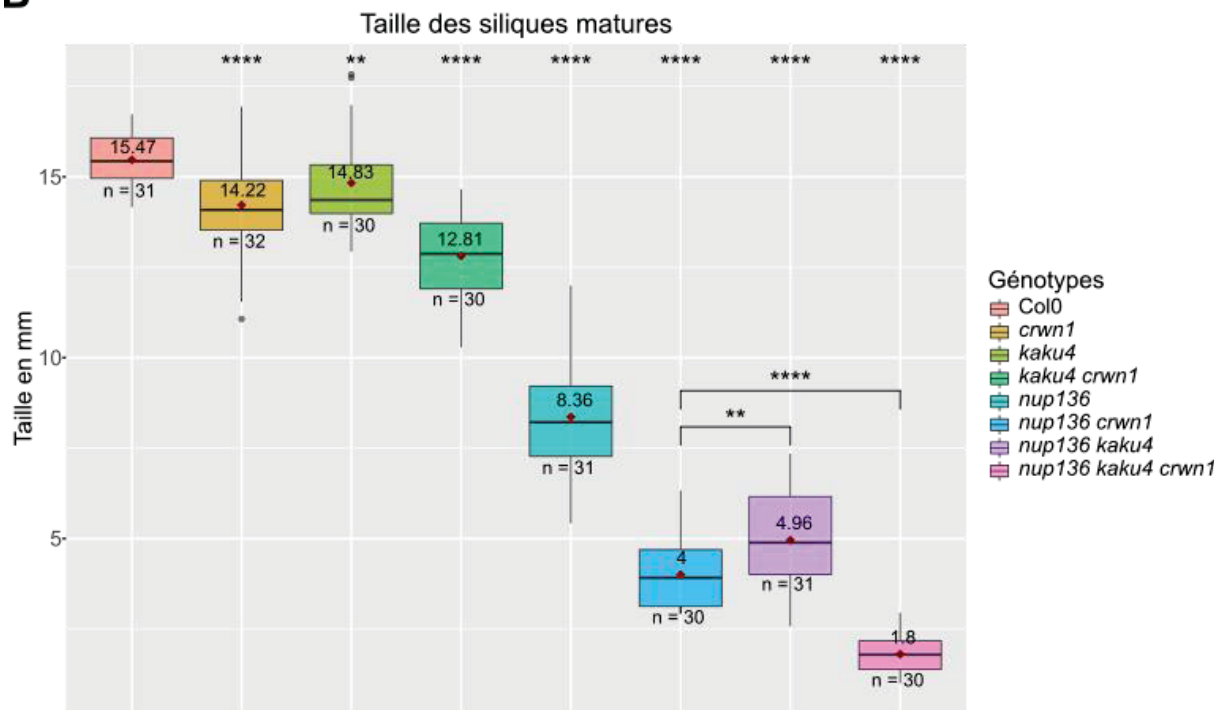
A**B**

Figure 2.11 : Etude de la taille des siliques dans les différents mutants *nup136*, *kaku4* et *crwn1*

A. Photo d'une silique des différents mutants, échelle : 10mm. **B.** Boxplot de la taille des siliques matures après mesure des siliques en utilisant le logiciel d'analyse d'images Fiji avec la moyenne affichée et représentée par un losange rouge ; n : nombre d'individus analysés par génotype ; Wilcoxon test, ns : non significatif (p -value > 0.05), *<0.5, **<0.01, ***<0.001, ****<0.0001.

Pour réaliser cette étude phénotypique, j'ai suivi le développement des plantes à deux stades, le stade végétatif avec la taille de la rosette et le stade reproductif avec la taille des siliques, le nombre de graines par siliques, la taille des graines et la viabilité du pollen.

I.2.d.ii) Analyse phénotypique au stade végétatif

L'étude du stade végétatif a été réalisée sur des plantes de 21 jours et a permis de montrer que l'ensemble des mutants ont des rosettes plus petites que le génotype sauvage et cela est très accentué pour les mutants impliquant *NUP136* et notamment pour le double mutant *nup136-2 crwn1-2* et le triple mutant *nup136-2 kaku4-2ad crwn1-2*, la rosette passant d'environ 40 cm² pour le sauvage à 3 cm² pour le double mutant et 0,76 cm² pour le triple mutant (**Figures 2.9A-B**). De plus, nous avons remarqué que les mutants impliquant *NUP136* possèdent des feuilles courbées vers la face interne plus ou moins visibles en fonction des plantes à 21 jours et induisent une transition florale précoce par rapport aux plantes sauvages. A un stade plus tardif, les mutants *nup136-2 crwn1-2* et *nup136-2 kaku4-2ad crwn1-2* présentent des petites plantes avec une taille respective de 5 cm et de 1 cm de haut, au lieu de 30cm (**Figures 2.10A-B**).

I.2.d.iii) Analyse phénotypique au stade reproductif

Au niveau reproductif, l'ensemble des mutants analysés possèdent de plus petites siliques que le sauvage et comme pour la taille de la rosette, les mutants impliquant *NUP136* sont plus impactés (**Figures 2.11A-B**). En effet, la taille des siliques du simple mutant *nup136-2* est réduite de moitié avec une taille moyenne de 8,36 mm (contre 15,47 mm pour le sauvage) et encore réduite de moitié pour les doubles mutants *nup136-1 crwn1-2* et *nup136-1 kaku4-2* avec respectivement 4 et 4,96 mm allant jusqu'à une taille d'environ 2 mm pour le triple mutant *nup136-2 kaku4-2ad crwn1-2* (**Figures 2.11A-B**). La diminution de la taille induit une réduction du nombre moyen de graines par silique, passant d'une soixante pour le sauvage à une vingtaine pour le simple mutant *nup136-2* et à environ 6 pour le double mutant *nup136-2 kaku4-2* et à environ 3 pour *nup136-2 crwn1-2* alors que le triple mutant est quasiment stérile avec seulement trois graines observées au sein de trente siliques matures (**Figure 2.12A**). Comme il est connu que lorsque le nombre de graines dans les siliques est réduit, celles-ci sont plus grosses (Herridge *et al.*, 2011; Cheng *et al.*, 2014), j'ai mesuré la taille des graines des différents mutants (à l'exception du triple mutant où le nombre de graines est insuffisant). La taille des graines change peu dans les mutants *kaku4-2ad*, *crwn1-2* et *kaku4-2ad crwn1-2* alors qu'elle augmente dans les mutants *nup136-2*, *nup136-2 kaku4-2* et *nup136-2 crwn1-2* par rapport au sauvage, passant de 88 μm² pour ce dernier, à 112-119 μm² pour les mutants impliquant *NUP136* (**Figure 2.12B**).

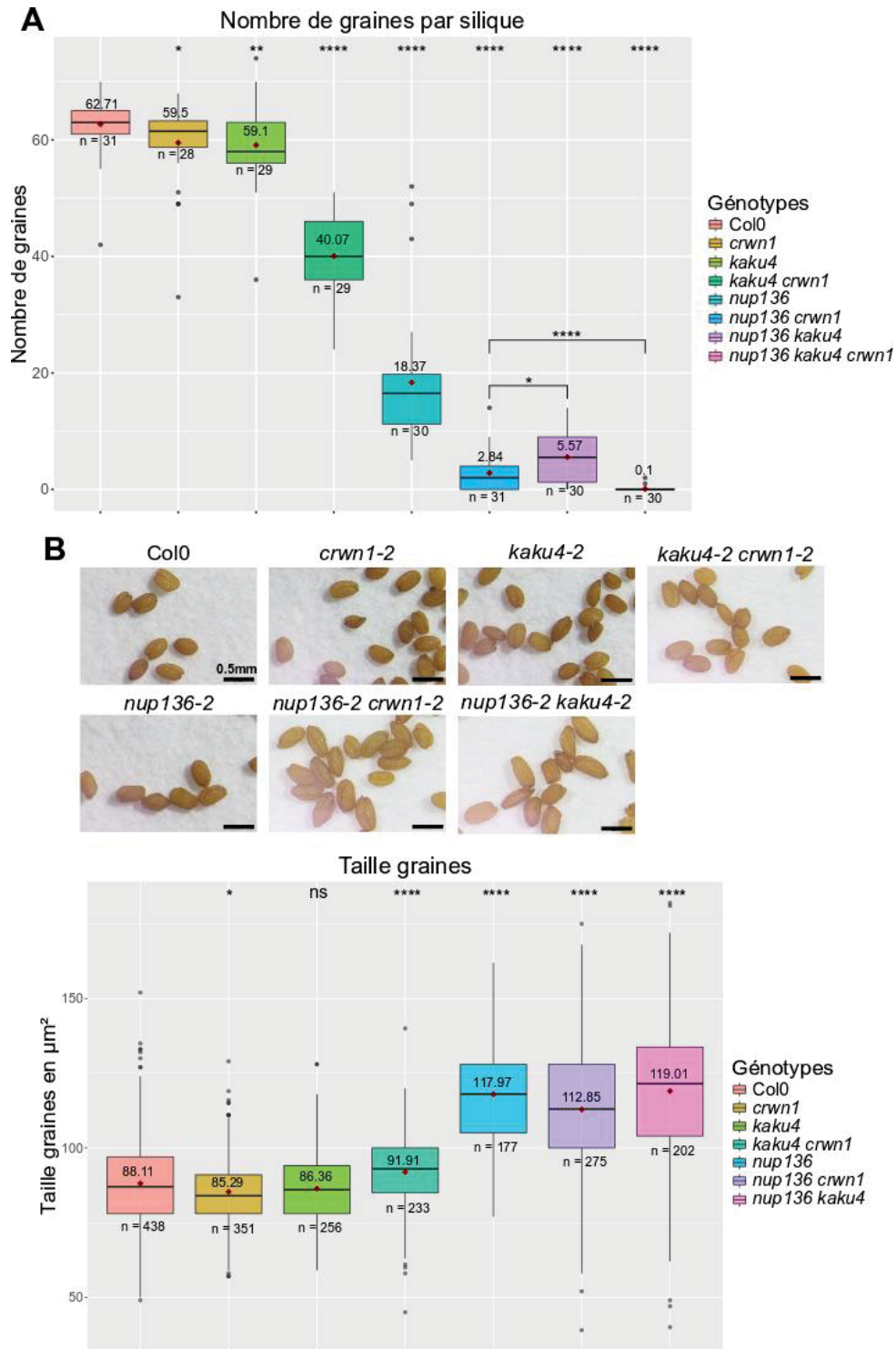


Figure 2.12 : Etude du nombre de graines et de leur taille des différents mutants *nup136*, *kaku4* et *crwn1*

A. Boxplot du nombre de graines par silique après décoloration des siliques et comptage avec la moyenne affichée et représentée par un losange rouge. **B.** Photos des graines sèches des différents mutants, échelle : 0.5mm. **C.** Boxplot de la taille des graines sèches après mesure de l'aire des graines en utilisant le logiciel d'analyse d'images Fiji avec la moyenne affichée et représentée par un losange rouge ; n : nombre d'individus analysés par génotype ; Wilcoxon test, ns : non significatif (p -value > 0.05), * < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001.

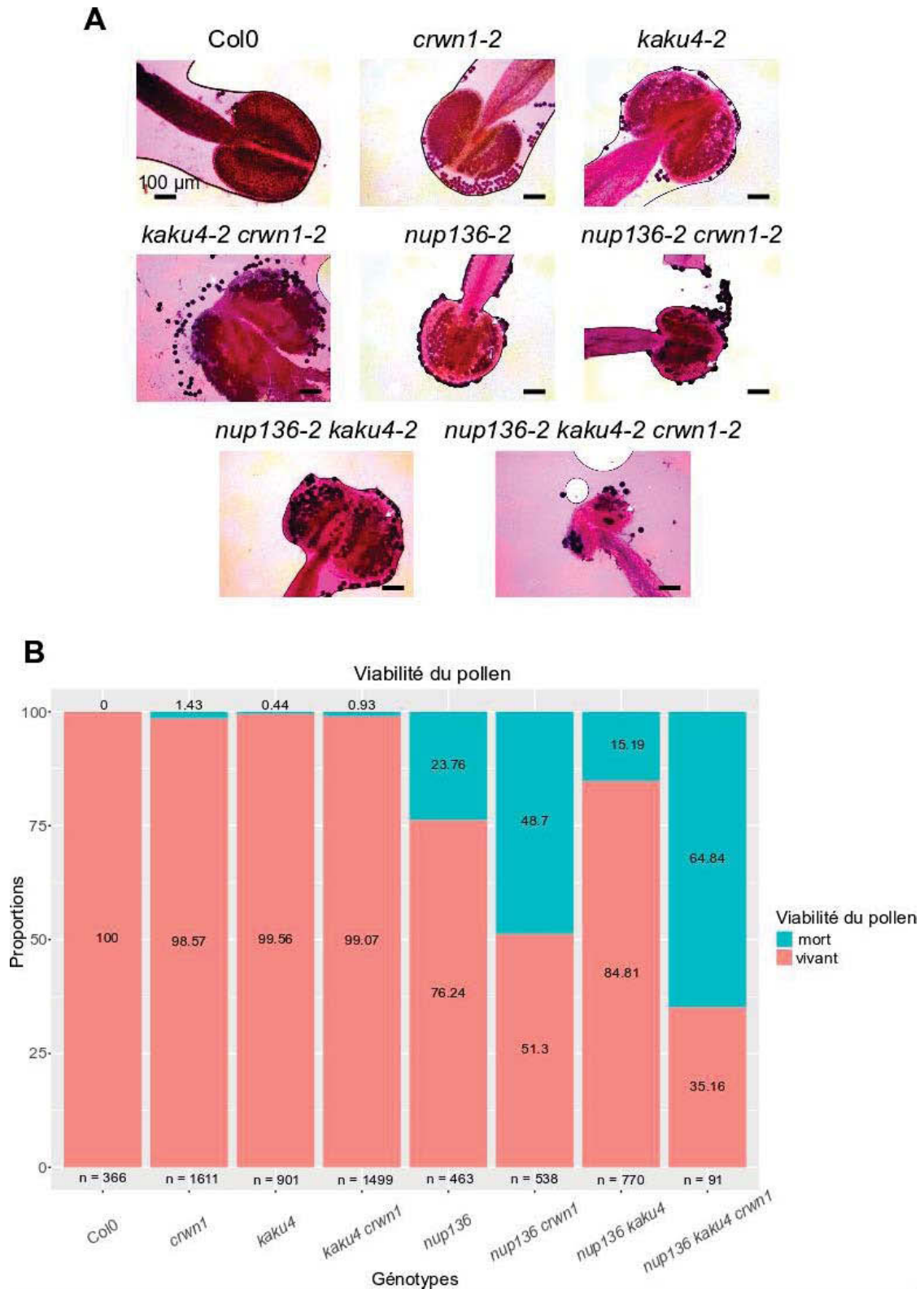


Figure 2.13 : Etude de la viabilité du pollen des différents mutants *nup136*, *kaku4* et *crwn1*

A. Photos d'étamines des différents mutants, colorée par la coloration d'Alexander (bleu : mort et rose : vivant) ; flèches blanches : pollen mort, échelle : 100μmm. **B.** Histogramme de la viabilité du pollen après comptage avec les pourcentages de chaque type affiché ; n : nombre de grains de pollen analysés par génotype.

Le nombre de graines par siliques est réduit d'un facteur 3 pour le mutant *nup136-2* et d'un facteur 20 pour le mutant *nup136-2 crwn1-2* alors que la taille des siliques n'est divisée que par 2 et 4 respectivement et que la taille des graines est seulement multipliée par 1,3 dans ces mutants. La taille des siliques et des graines ne sont donc pas les seuls facteurs induisant la réduction du nombre de graines.

Cette différence dans le nombre de graines peut s'expliquer par un défaut au niveau de la fécondation pouvant être provoqué notamment par la plus faible viabilité du pollen observée chez le mutant *nup136-2* (Tamura *et al.*, 2017). Pour vérifier cette hypothèse, j'ai testé la viabilité du pollen pour ces différents mutants *via* la coloration d'Alexander d'étamines entières (cf Matériel et Méthodes de notre article). En accord avec les observations de Tamura *et al.*, le mutant *nup136-2* a une viabilité du pollen plus faible que le sauvage alors que les autres simples mutants et le mutant *kaku4-2ad crwn1-2* ont peu d'impact sur cette viabilité (**Figures 2.13A-B**). Nous avons observé dans le double mutant *nup136-2 crwn1-2*, un effet synergique sur la viabilité du pollen en la réduisant d'environ 50% et cet effet est encore plus fort dans le triple mutant *nup136-2 kaku4-2ad crwn1-2* où seul 35% du pollen est viable (**Figure 2.13A-B**). De plus, dans le triple mutant il y a très peu de grains de pollen par anthère diminuant encore plus les chances de pollinisation. Nous observons une nette différence de viabilité du pollen entre cette expérience et celle présentée dans notre article pour les mutants *nup136-2* et *nup136-2 kaku4-2* (cf paragraphe I.1.b), cela s'explique d'une part par le fait que moins de grains de pollen ont été comptés dans cette expérience et avec de grandes disparités en fonction des génotypes ainsi que par le comptage principalement au sein de l'anthère alors que dans notre publication, le pollen a été libéré des anthères afin de faciliter le comptage notamment des grains de pollen morts.

I.2.d.iv) Analyse phénotypique de la morphologie nucléaire

Outre les phénotypes observés au niveau du développement et de la croissance de la plante, nous avons également étudié la morphologie de leur noyau. En effet, les mutations affectant les gènes *NUP136*, *KAKU4* et *CRWN1* impactent tous la morphologie du noyau en réduisant l'élongation de la membrane nucléaire ce qui aboutit à des noyaux plus ronds (Tamura and Hara-Nishimura, 2011; Wang, Dittmer and Richards, 2013; Goto *et al.*, 2014). Pour déterminer si la combinaison de ces mutations impacte la morphologie nucléaire en 3D, nous avons utilisé le plugin NucleusJ2 disponible sur Fiji (Dubos *et al.*, 2020). Notons que cette analyse réalisée pour *KAKU4* et *CRWN* a été publiée dans des articles antérieurs de l'équipe (Dubos *et al.*, 2020, 2022) et montrent que le triple mutant *kaku4-2 crwn1-2 crwn4-1* possède

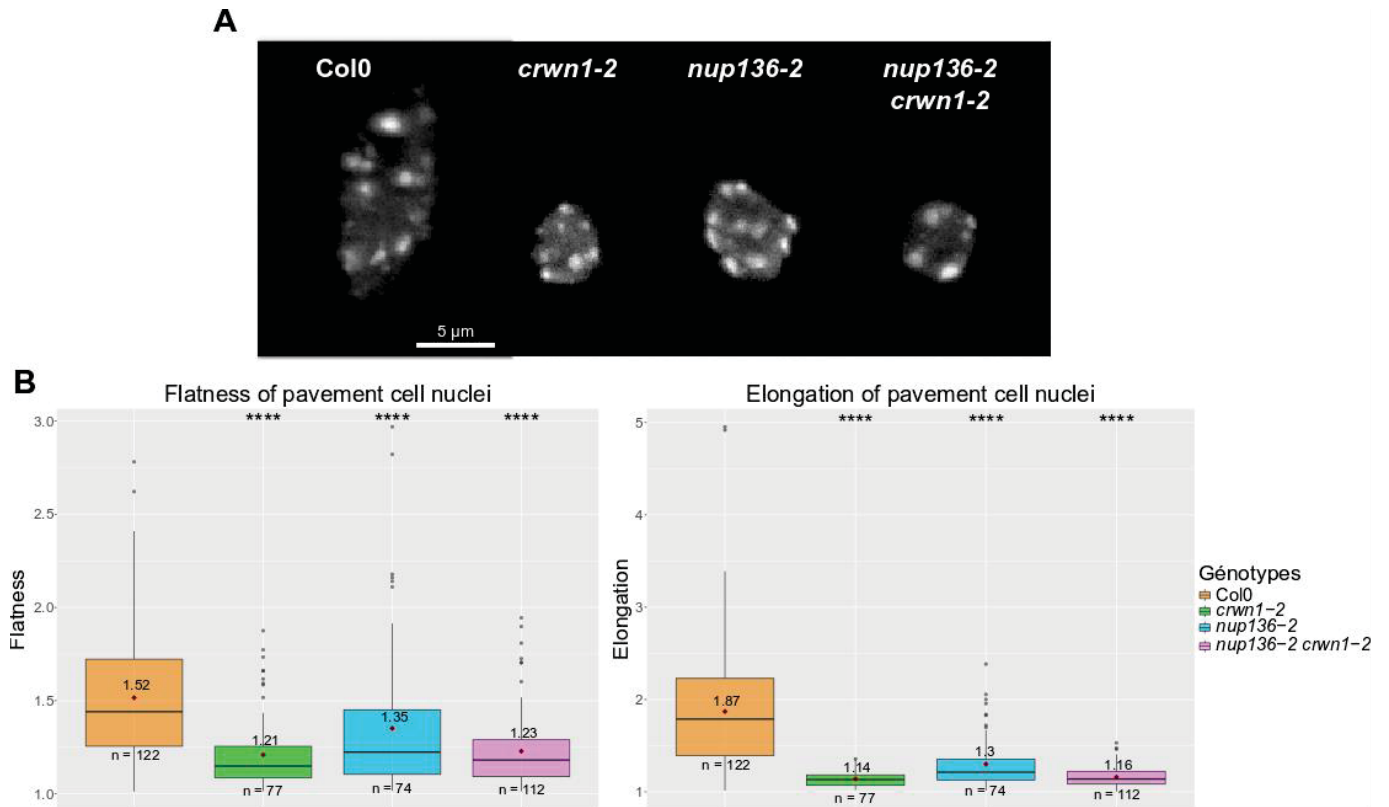


Figure 2.14 : Morphologie nucléaire des mutants *nup136*, *crwn1* et *nup136 crwn1*

A. Images des différents mutants, sélectionnées pour leurs paramètres proches des valeurs moyennes pour les différents critères morphologiques. Projection en Z de noyaux bruts de cellules pavimenteuses des différents mutants, échelle : 5µm. **B.** Boxplots de la flatness et de l'élongation des noyaux de cellules pavimenteuses après analyse avec le logiciel NucleusJ2 (Dubos *et al.*, 2020) avec la moyenne affichée et représentée par un losange rouge ; n : nombre d'individus analysés par génotype ; Wilcoxon test, ns : non significatif (p-value > 0.05), *<0.5, **<0.01, ***<0.001, **** <0.0001.

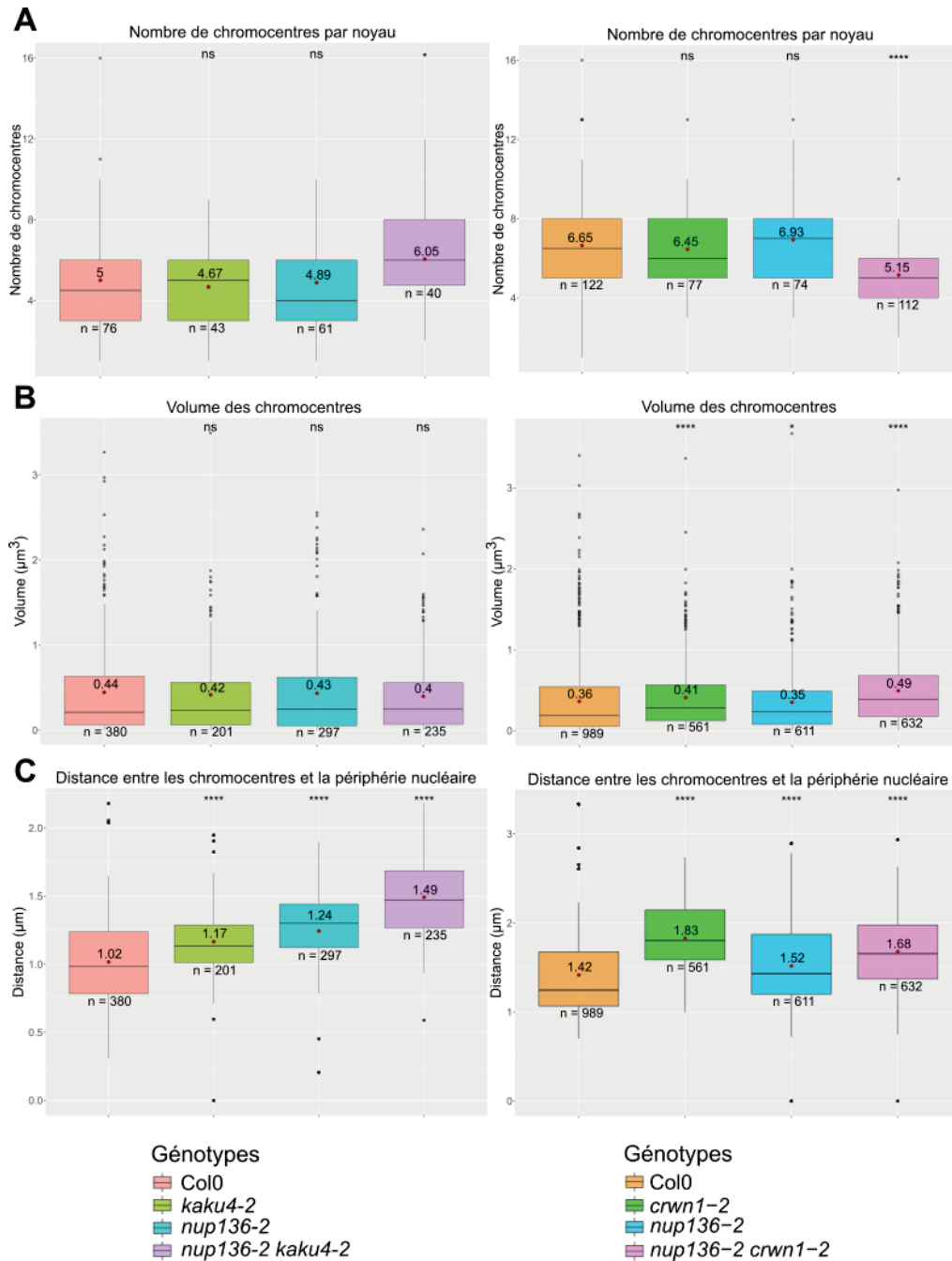


Figure 2.15 : Organisation de la chromatine dans les différents mutants *nup136*, *kaku4* et *crwn1*

Boxplots du nombre de chromocentres (A), de leur volume (B) et de la distance entre les chromocentres et la périphérie nucléaire (C) dans les noyaux de cellules pavimenteuses après analyse avec le plugin Fiji, NODeJ (Dubos *et al.*, 2022) avec la moyenne affichée et représentée par un losange rouge ; n : nombre d'individus analysés par génotype (noyaux en A et chromocentres en B) ; Wilcoxon test, ns : non significatif (p -value > 0.05), * < 0.5, ** < 0.01, *** < 0.001, **** < 0.0001.

des noyaux plus petits, sphériques ou le nombre de chromocentres est réduit et leurs positions plus éloignées de l'enveloppe nucléaire que dans le sauvage. Dans ce triple mutant, les loci d'ADNr 5S habituellement proches du nucléole, se retrouvent plus proches de la périphérie nucléaire alors qu'à l'inverse la position des loci de 180bp identifiant les chromocentres, ont tendance à s'éloigner de la périphérie du noyau. Ceci suggère une réorganisation importante de l'organisation 3D de la chromatine qu'il faudrait décrire par exemple avec des expériences de Hi-C.

Les données pour NUP136 et KAKU4 ont été publiées dans notre article de cette année (cf paragraphe I.1.b) et j'ai également étudié la morphologie nucléaire des mutants impliquant NUP136 et CRWN1. Les mutants *crwn1-2*, *nup136-2* et *nup136-2 crwn1-2* possèdent des noyaux sphériques avec notamment une diminution de l'élongation et de l'aplatissement (flatness) avec un plus fort impact du mutant *crwn1-2* par rapport au mutant *nup136-2* et des valeurs semblables entre le mutant *crwn1-2* et *nup136-2 crwn1-2*, suggérant que les effets observés dans le double mutant ont pour origine le mutant *crwn1-2* (**Figures 2.14A-B**). J'ai déterminé le nombre de chromocentres, leur volume ainsi que leur position par rapport à la périphérie nucléaire *via* le plugin NODEJ disponible sur Fiji (Dubos *et al.*, 2022), dans les mêmes mutants que la morphologie nucléaire impliquant *NUP136*, *KAKU4* et *CRWN1*. Le nombre de chromocentres par noyau est semblable entre les simples mutants et le sauvage alors que dans le double mutant *nup136-2 kaku4-2*, le nombre de chromocentres augmente avec une légère tendance à la diminution de leur volume (**Figures 2.15A-B**). Quant au double mutant *nup136-2 crwn1-2*, son nombre de chromocentres diminue induisant l'augmentation de leur volume (**Figures 2.15A-B**). Cependant, l'ensemble des mutants, que ce soit les simples mutants ou les doubles mutants, impactent la distance entre les chromocentres et la périphérie nucléaire avec des chromocentres plus éloignés de la périphérie nucléaire (**Figure 2.15C**). Ce qui suggère que les protéines NUP136, KAKU4 et CRWN1 jouent un rôle sur la compaction de l'ADN et sur le maintien des chromocentres à la périphérie nucléaire. Afin d'étudier plus précisément l'impact de ces protéines sur l'organisation de la chromatine, nous avons initié l'étude des loci 5S et 180 pb par DNA-FISH dans les mutants *nup136-2 crwn1-2* et *nup136-2 kaku4-2*.

Nos résultats montrent que les protéines KAKU4, NUP136 et CRWN1 interagissent fonctionnellement, que la mutation de *NUP136* affecte la transition florale qui est plus précoce, entraîne la formation de siliques de petite taille et impacte la viabilité du pollen. L'association de cette mutation avec celle de *CRWN1* entraîne une forte réduction de la taille des plantes, des siliques, du nombre de graines et de la viabilité du pollen et impacte la morphologie nucléaire

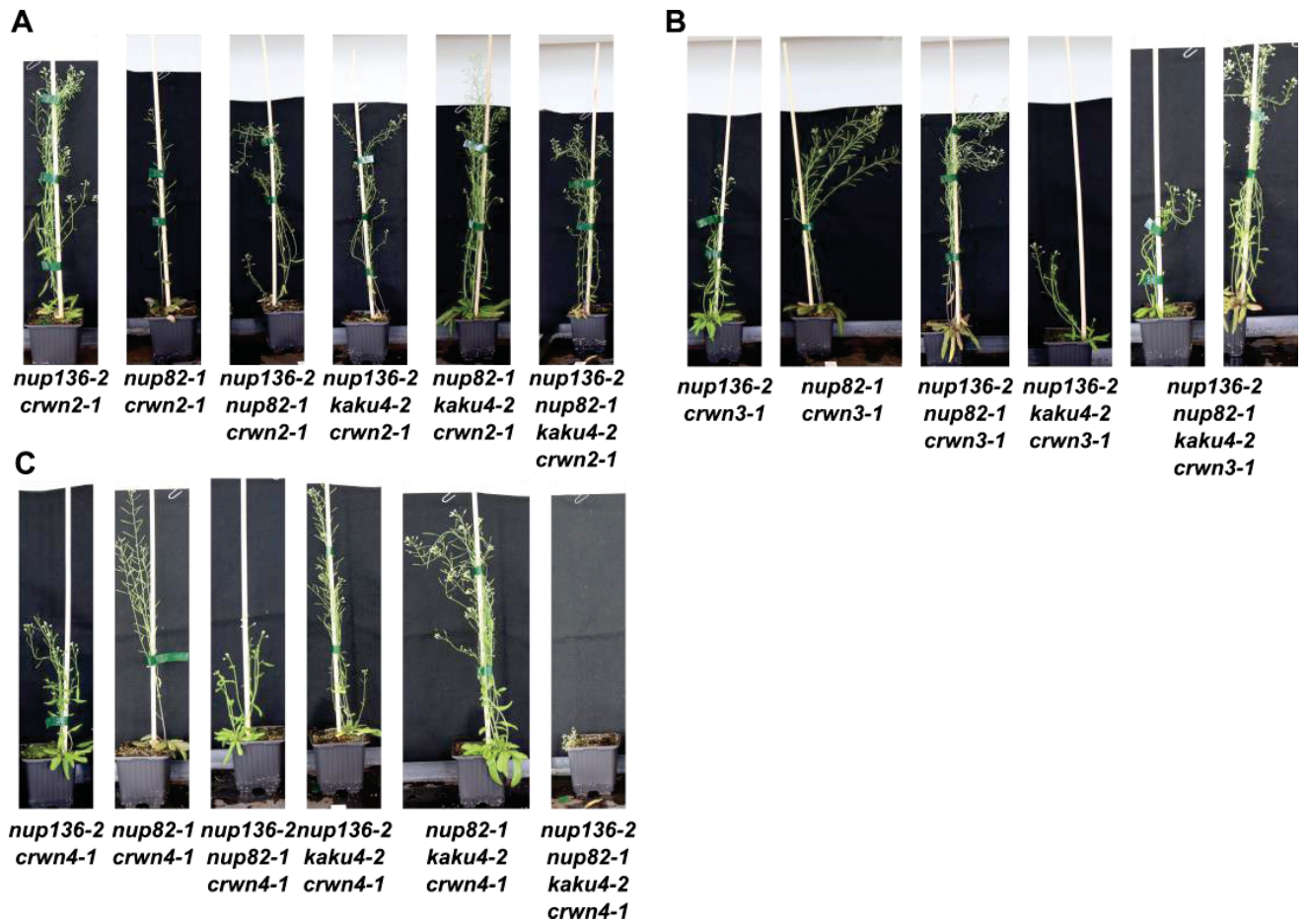


Figure 2.16 : Taille des différents mutants à un stade tardif du développement

Photos des différentes combinaisons de mutants *nup136-2*, *nup82-1*, *kaku4-2* et *crwn2-1* âgés de 7 semaines (A), *nup136-2*, *nup82-1*, *kaku4-2* et *crwn3-1* (B) et *nup136-2*, *nup82-1*, *kaku4-2* et *crwn4-1* (C).

en induisant des noyaux sphériques, suggérant un effet synergique entre CRWN1 et NUP136. L'ajout de la mutation de *KAKU4* aux deux précédentes n'a qu'un faible impact additionnel, en réduisant un peu plus la taille de la plante, des siliques, du nombre de graines et de la viabilité du pollen, cependant son association avec la mutation *NUP136* seule n'a pas d'impact drastique sur le phénotype des plantes très similaire au simple mutant *nup136-2* sauf pour ce qui est de la morphologie nucléaire où c'est le phénotype de *kaku4-2* qui a le plus fort impact en induisant des noyaux sphériques. Je n'ai pas inclus les résultats des combinaisons avec la mutation de *NUP82*, cette dernière n'impactant pas ou que très partiellement les phénotypes observés (**Figure 2.10A**).

I.2.d.v) Effets des mutations *crwn2*, *crwn3* et *crwn4*

Pour tester les interactions génétiques entre les protéines KAKU4, NUP136 et les 3 autres protéines CRWN : CRWN2, CRWN3 et CRWN4, nous avons utilisé la même stratégie et construit les différentes combinaisons de mutants. Les analyses ont été initiées, cependant pour le moment bien moins complètes qu'avec *crwn1*. Ainsi, l'association du mutant *crwn2* avec *kaku4* et *nup136* n'impacte que très faiblement le développement de la plante (**Figure 2.16A**), alors que la mutation *crwn3* couplée à celle de *nup136* induit des plantes de petite taille légèrement plus grandes que le mutant *nup136-2 crwn1-2* (**Figure 2.16B**). Ce phénotype est accentué par l'ajout de la mutation *kaku4* mais le quadruple mutant *nup136-2 nup82-1 kaku4-2 crwn3-1* n'impacte pas plus la taille de la plante que le triple mutant *nup136-2 kaku4-2 crwn3-1* (**Figure 2.16B**). Tandis que la mutation *crwn4* associée à celle de *nup136* induit des petites plantes semblables au double mutant *nup136-2 crwn3-1*, l'ajout de la mutation de *KAKU4* engendre une plante de taille sauvage alors que le quadruple mutant *nup136-2 nup82-1 kaku4-2 crwn4-1* provoque de très petites plantes semblables au triple mutant *nup136-2 kaku4-2 crwn1-2* et au quadruple mutant *nup136-2 nup82-1 kaku4-2 crwn1-2* (**Figure 2.16C**).

Ces premiers tests suggèrent que les quatre protéines CRWN n'auraient pas des fonctions complètement redondantes en ce qui concerne le développement de la plante, ce qui pouvait être attendu au vu des travaux précédents sur les mutants *crwn* (Wang, Dittmer and Richards, 2013). Nos résultats suggèrent que génétiquement, les effets observés avec les mutants des protéines CRWN1 et CRWN4 avec NUP136 et KAKU4, sont plus forts que ceux observés avec CRWN2 et CRWN3. Cette distinction permet de séparer fonctionnellement les protéines CRWN en deux sous-groupes avec des effets majeurs observés pour CRWN1 et CRWN4.

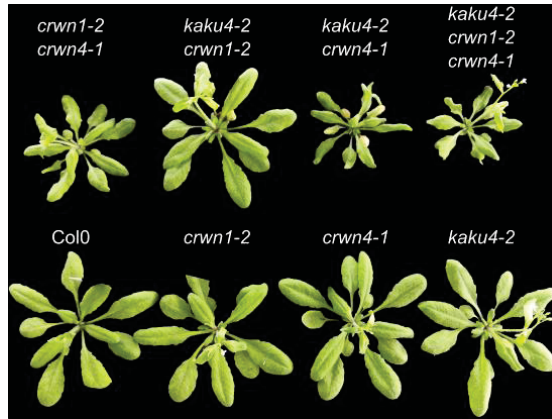


Figure 2.17 Phénotypes des mutants impliquant les composants du nucléosquelette
Photo des différents mutants impliquant *KAKU4*, *CRWN1* et *CRWN4* âgés de 24 jours.

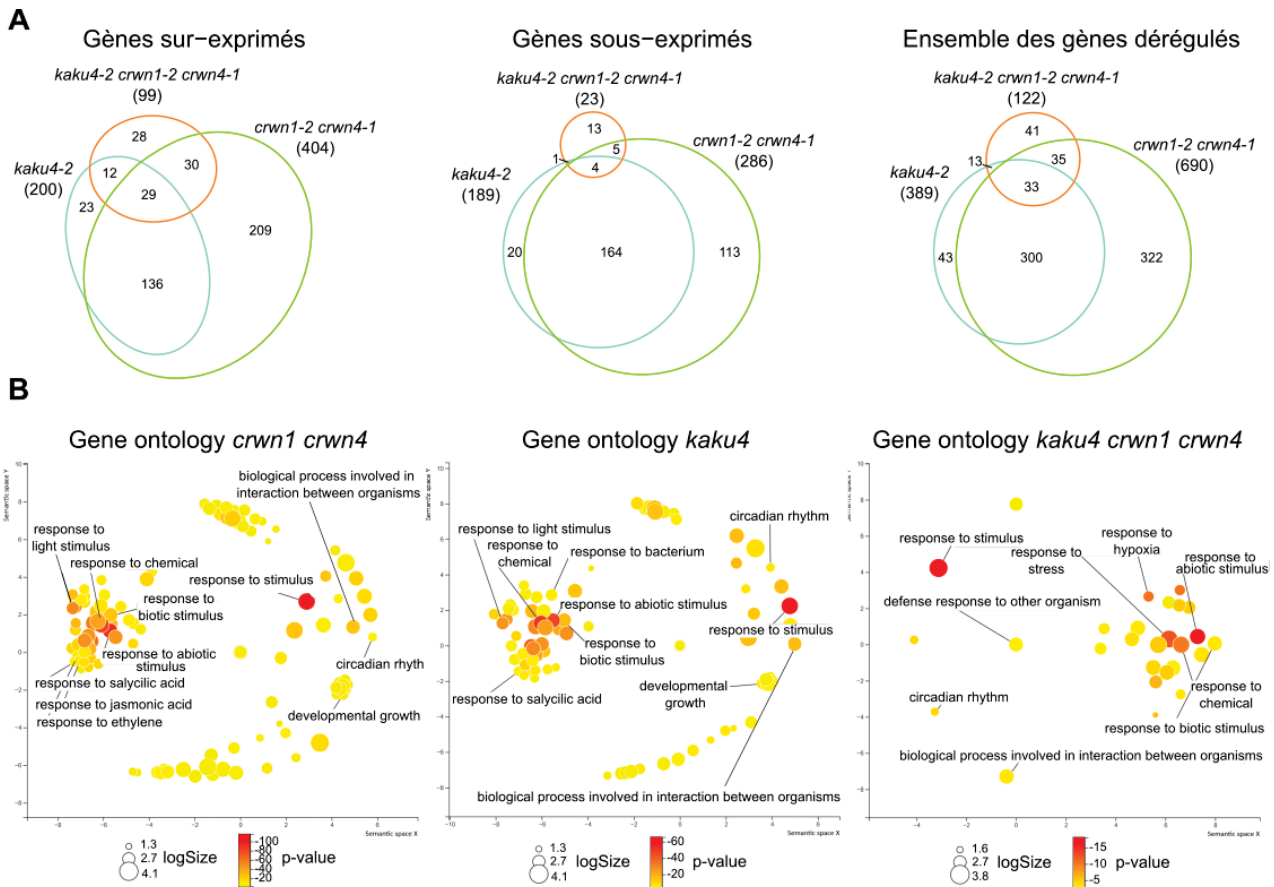


Figure 2.18 : Impact des protéines *KAKU4*, *CRWN1* et *CRWN4* dans la réponse aux stress.

A. Diagrammes de Venn des gènes sur-exprimés, sous-exprimés et totaux dérégulés dans les mutants *kaku4-2*, *crwn1-2 crwn4-1* et *kaku4-2 crwn1-2 crwn4-1* par rapport au sauvage. La taille de chaque ellipse est proportionnelle au nombre de gènes dérégulés dans nos données de RNA-seq ($FC > 1.7$ ou $FC < 0.58$; $p_{adj} < 0.05$). **B.** Gene ontology des gènes dérégulés dans les différents mutants. Analyse réalisée avec LAGO (Logically Accelerated GO Term Finder, <https://go.princeton.edu/cgi-bin/LAGO>) puis visualisée avec REVIGO (reduce + visualize gene ontology, Supek *et al.*, 2011).

II- Impact des protéines de la périphérie nucléaire sur l'expression des gènes

II.1) Analyse des mutants des protéines KAKU4, CRWN1 et CRWN4

II.1.a) Analyse transcriptomique des mutants *kaku4*, *crwn1* et *crwn4*

Les protéines du nucléosquelette et notamment les protéines CRWN jouent un rôle dans la régulation de l'expression des gènes. Afin d'étudier leurs rôles dans ce processus, nous avons réalisé une analyse RNA-seq sur les mutants *kaku4-2*, *crwn1-2 crwn4-1* et *kaku4-2 crwn1-2 crwn4-1* (séquençage 150 bp en paired-end sur des plantules de 14 jours poussées sur milieu synthétique). L'analyse des données a été réalisée comme décrit dans le paragraphe [I.1.c.i](#), sauf que dans ce cas, j'ai considéré les gènes dérégulés comme ayant un Fold Change supérieur à 1,7 pour les gènes sur-exprimés et inférieur à 0,58 pour les gènes sous-exprimés avec une p-value inférieure à 0,05 par rapport au Col-0. Contrairement à ce que nous attendions au vu du phénotype développemental assez important (**Figure 2.17**), le triple mutant *kaku4-2 crwn1-2 crwn4-1* a très peu de gènes dérégulés par rapport aux deux autres mutants. En effet, seuls 122 gènes sont dérégulés alors que le mutant *kaku4-2* en dérégule 389 et le mutant *crwn1-2 crwn4-1*: 690. Pour analyser ces résultats, j'ai comparé les gènes sur-exprimés, sous-exprimés ou l'ensemble des gènes dérégulés dans nos trois mutants *via* des diagrammes de Venn et observé que 85% des gènes dérégulés dans le mutant *kaku4-2* le sont également dans le mutant *crwn1-2 crwn4-1*, soit 48% des gènes de ce dernier (**Figure 2.18A**). Les changements de l'organisation de la chromatine dans ces mutants du nucléosquelette pouvaient laisser supposer qu'ils dérégulent les éléments transposables. L'équipe d'Eric Richards a montré cependant, que les simples mutants *crwn* et les doubles mutants *crwn1 crwn2* et *crwn1 crwn4* ne dérégulent pas les éléments transposables, les seuls éléments dérégulés sont ceux présents dans des gènes qui eux-mêmes sont dérégulés (Choi, Strickler and Richards, 2019) dans les analyses de nos mutants nous n'observons pas non plus de dérégulation des éléments transposables.

Pour aller plus loin dans les fonctions biologiques affectées dans ces mutants, j'ai effectué une analyse de gene ontology des gènes dérégulés. Celle-ci a permis de montrer que nos mutants dérégulent significativement des gènes impliqués dans la réponse aux stress aussi bien biotiques qu'abiotiques ainsi que le rythme circadien et le développement de la plante (**Figure 2.18B**). L'implication des protéines CRWN dans la réponse aux stress biotiques a été montré précédemment *via* son implication dans l'induction de la voie de l'acide salicylique (Choi, Strickler and Richards, 2019). Cependant mon analyse montre que KAKU4 n'induit pas de dérégulation pour des gènes connus pour avoir des effets majeurs dans les différentes voies de

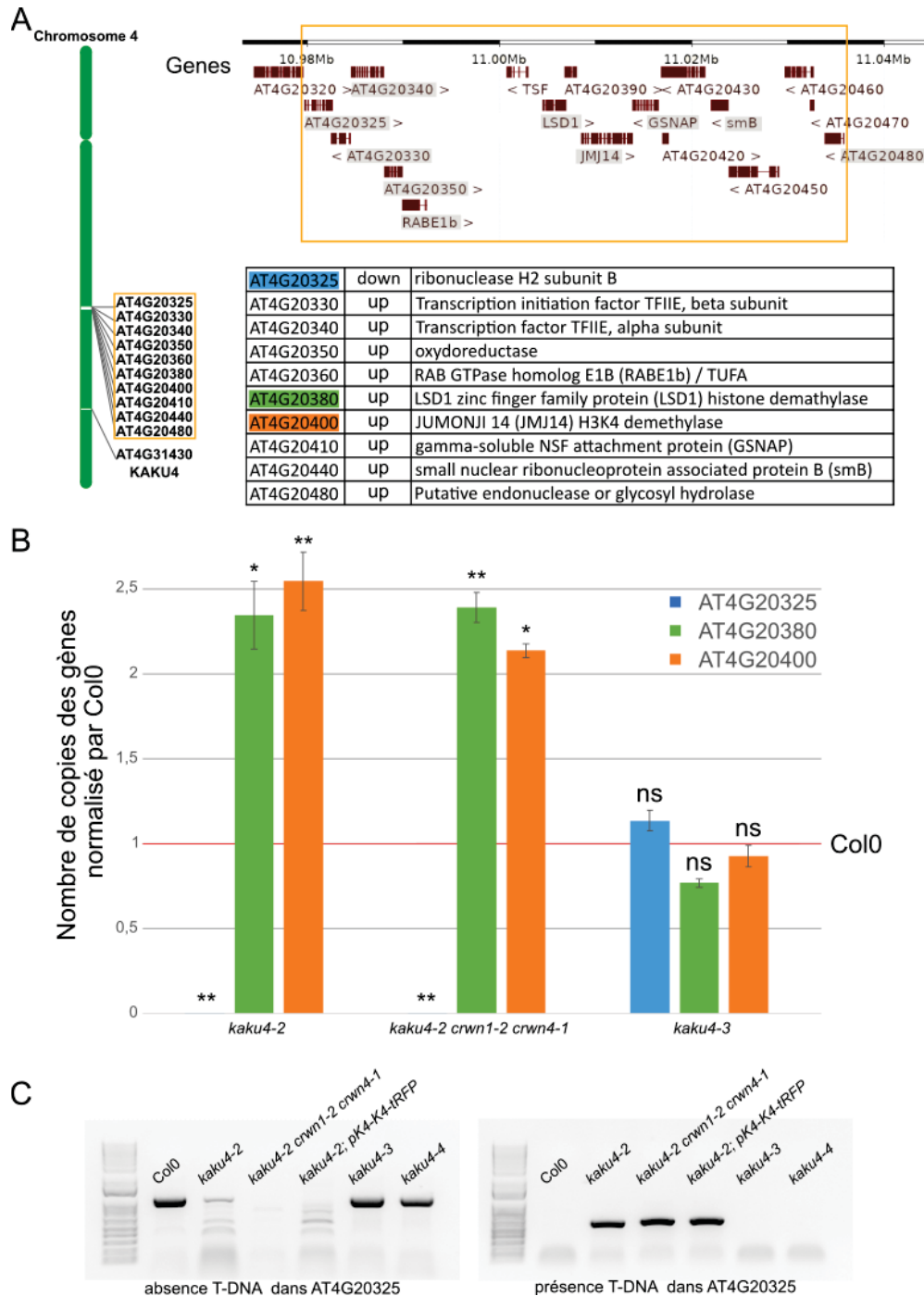


Figure 2.19 : Présence d'une duplication sur le chromosome 4 dans le mutant *kaku4-2*

A. Position du cluster de gènes dérégulés identifiés sur le chromosome 4 (position : 10 979 741-11 035 878) dans le mutant *kaku4-2* ainsi que du gène *KAKU4* (AT4G31430) et tableau de ces gènes avec leur nom et de leur dérégulation : sur-exprimés (up) ou sous-exprimés (down). Les gènes surlignés dans le tableau sont ceux utilisés en qPCR (**B**) avec en bleu le gène *AT4G20325*, en vert le gène *LSD1* (AT4G20380) et en orange le gène *JMJ14* (AT4G20400). **B.** Analyse qPCR du nombre de copie des gènes *AT4G20325*, *AT4G20380* et *AT4G20400* dans les mutants *kaku4-2*, *kaku4-2 crwn1-2 crwn4-1* et *kaku4-3* normalisé par rapport à Col0 (ligne rouge) ; test de Student, ns : non significatif (p -value > 0.05), * < 0.5, ** < 0.01. **C.** Analyse de génotypage dans différents mutants de *kaku4* montrant que le gène *AT4G20325* contient une insertion d'un T-DNA seulement dans l'allèle *kaku4-2*.

réponse aux stress biotiques ou abiotiques comme par exemple les gènes *PATHOGENESIS RELATED* qui sont dérégulés dans les mutants des gènes *CRWN* (Choi, Strickler and Richards, 2019). Ces résultats suggèrent un effet modéré de *KAKU4* dans la régulation de l'expression des gènes.

II.1.b) Caractérisation de la duplication du mutant T-DNA *kaku4-2*

L'une de nos hypothèses concernant le rôle des composants du nucléosquelette dans la régulation des gènes est que ces protéines pourraient recruter à la périphérie nucléaire des régions chromatiniennes et ainsi former des clusters de gènes corégulés. Au cours de mes analyses, j'ai observé un cluster de 10 gènes présents sur le chromosome 4 entre les positions +10 982 517 et +11 035 878 (soit ~20 cM du locus *KAKU4*) dérégulés dans les mutants *kaku4-2* et *kaku4-2 crwn1-2 crwn4-1* (**Figure 2.19A**). L'ensemble de ce cluster est sur-exprimé sauf le premier gène qui est sous-exprimé (**Figure 2.19A**). Dans un premier temps, j'ai cherché à lier l'expression des gènes de ce cluster à une coexpression due à un facteur de transcription commun via l'outil GeneMANIA (<https://genemania.org/>), mais aucune coexpression connue entre ces gènes ni aucun facteur commun n'est sorti de l'analyse. J'ai ensuite étudié la possibilité que cette dérégulation soit liée à un réarrangement chromosomique dû à l'insertion du T-DNA comme cela a été montré pour d'autres mutants (Jupe *et al.*, 2019). De plus, le Fold Change de ces gènes étant souvent proche de 2, nous avons émis l'hypothèse que cela pourrait être lié à une duplication de cette région.

J'ai donc réalisé une qPCR afin de déterminer le nombre de copies présentes dans le génome pour deux des gènes sur-exprimés dans le cluster, *AT4G20380* (*LSD1*) et *AT4G20400* (*JMJ14*) ainsi que pour le seul gène sous-exprimé situé au début du cluster, *AT4G20325*, dans le sauvage et dans les combinaisons de mutants construites à partir de la lignée *kaku4-2* : *kaku4-2*, *kaku4-2 crwn1-2 crwn4-1* ainsi qu'un allèle mutant indépendant de *kaku4* : *kaku4-3*. Nos résultats démontrent sans ambiguïté que le mutant *kaku4-3* ne possède qu'une seule copie de ces 3 gènes (comparé au sauvage, Col0) alors que nous observons 2 copies pour les gènes *AT4G20380* et *AT4G20400* dans les mutants construits à partir de *kaku4-2*, et aucune copie pour le gène *AT4G20325* qui lui était sous exprimé (**Figure 2.19B**). Une recherche de la lignée SALK_076754 contenant la mutation *kaku4-2* dans la base de données TAIR (The Arabidopsis Information Resource) décrit, qu'outre l'insertion T-DNA présente dans le gène *KAKU4* (allèle *kaku4-2*), cette lignée possède également une autre insertion additionnelle dans le gène *AT4G20325* au niveau des amorces choisies lors des expériences de qPCR.

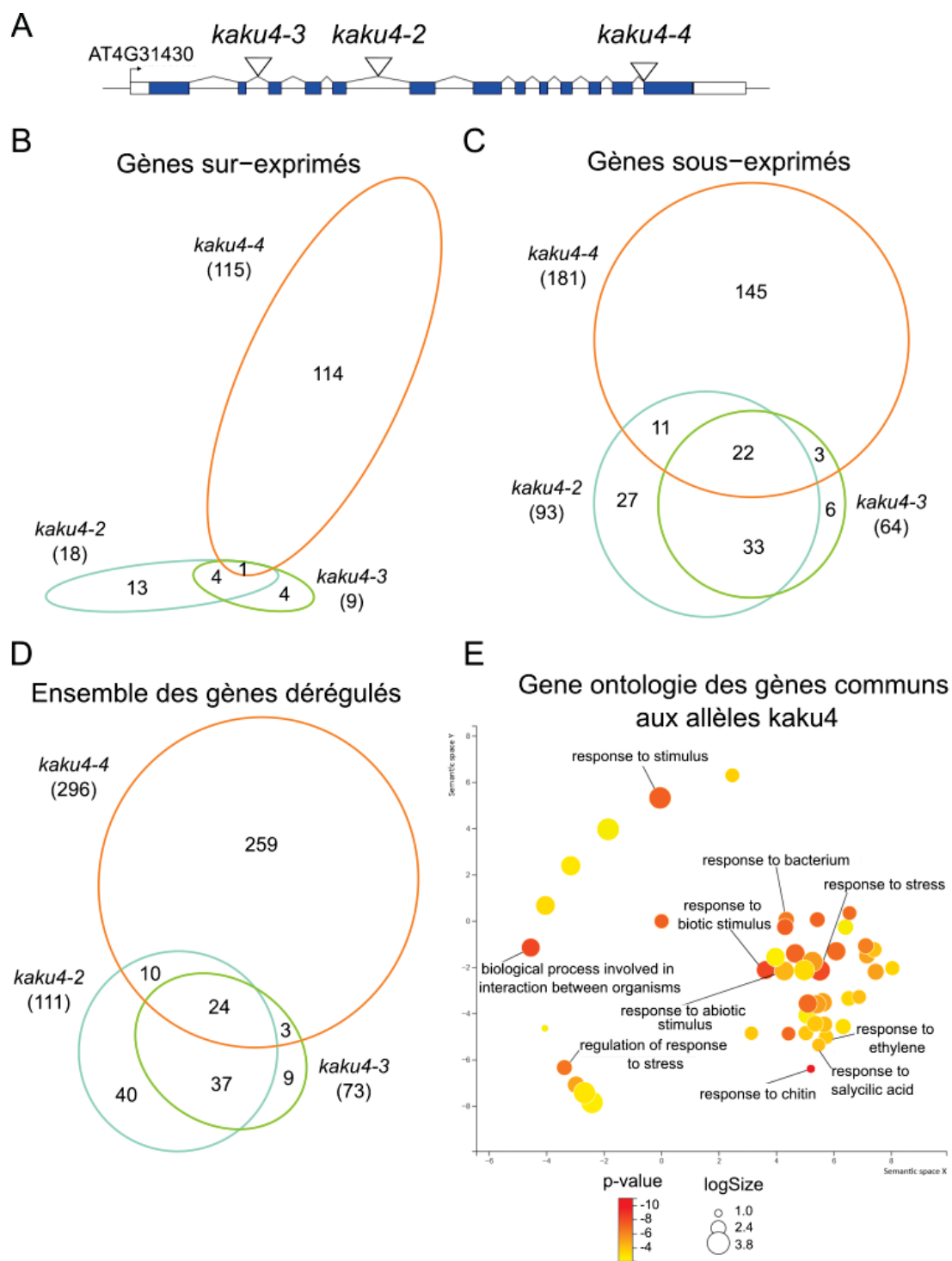


Figure 2.20 : Analyse transcriptomique des différents allèles de *kaku4*

A. Localisation des différents allèles mutants sur le gène *KAKU4* décrits comme nuls par Goto *et al.* mais après vérification présence de transcrits avant ou après les insertions T-DNA en fonction des mutants (cf **Figure 2.29** et **Annexe1**) **B-D.** Diagrammes de Venn des gènes sur-exprimés (**B**), sous-exprimés (**C**) et totaux (**D**) dérégulés dans les mutants *kaku4-2*, *kaku4-3* et *kaku4-4* par rapport au sauvage. La taille de chaque ellipse est proportionnelle au nombre de gènes dérégulés dans nos données de RNA-seq ($FC > 1.7$ ou $FC < 0.58$; $padj < 0.05$). **E.** Gene ontology des gènes dérégulés communs aux différents mutants de *kaku4*. Analyse réalisée avec LAGO puis visualisée avec REVIGO (cf figure 2.18).

A partir de ces données, j'ai pu définir des amorces pour détecter la présence de cette seconde insertion localisée dans le gène *AT4G20325*, et, par PCR, confirmer ou non la présence de ce T-DNA associé à la duplication dans nos différentes combinaisons mutantes. Il convient ici de préciser que *KAKU4* correspond au gène *AT4G31430* qui se trouve à quelques 20 cM de cette autre insertion. L'ensemble des mutants contenant l'allèle *kaku4-2*, c'est-à-dire *kaku4-2*, *kaku4-2 crwn1-2 crwn4-1* ainsi que le mutant complémenté pour *KAKU4*, *kaku4-2 ; pKAKU4-KAKU4-tRFP*, possèdent l'insertion au niveau du gène *AT4G20325* ainsi que la duplication alors que le sauvage et tous les autres allèles mutants indépendants de *kaku4*, *kaku4-3* et *kaku4-4* (**Figure 2.20A**) ne la possèdent pas. Cette lignée T-DNA dans l'allèle *kaku4-2* a également une duplication de la région 10 982 517 à 11 035 878 du chromosome 4 associée à une autre insertion, celle-ci pourrait donc avoir un impact sur les gènes dérégulés dans la lignée mutante *kaku4-2*.

Pour conclure, notre analyse transcriptomique nous a permis de confirmer que les protéines du nucléosquelette CRWN1, CRWN4 et KAKU4 ne dérégulent pas l'expression des éléments transposables. Ce résultat est surprenant car les mutants *crwn* induisent une dispersion et une délocalisation des chromocentres, suggérant leur implication dans le maintien de la compaction de l'hétérochromatine et dans sa localisation à la périphérie nucléaire. Les protéines du nucléosquelette dérégulent donc uniquement des régions géniques qui sont impliquées, pour la plupart, dans la réponse aux stimuli environnementaux comme la défense face aux stress biotiques et abiotiques. L'implication des protéines CRWN dans ces processus a déjà été étudiée par différentes équipes et notamment leurs implications dans la voie de l'acide salicylique (Guo *et al.*, 2017; Choi, Strickler and Richards, 2019), alors que le rôle de KAKU4 dans la régulation de la transcription n'a, pour le moment, pas été étudié. Afin de restreindre le nombre de gènes cibles en l'absence de master gène régulateur, gène impliqué dans la régulation de nombreux autres, nous avons décidé de réaliser une nouvelle analyse transcriptomique mais de différents mutants *kaku4* (cf paragraphe II.3) pour essayer de réduire le nombre de gènes cibles et sélectionner ceux communs aux différents allèles de *KAKU4* pour une plus grande spécificité.

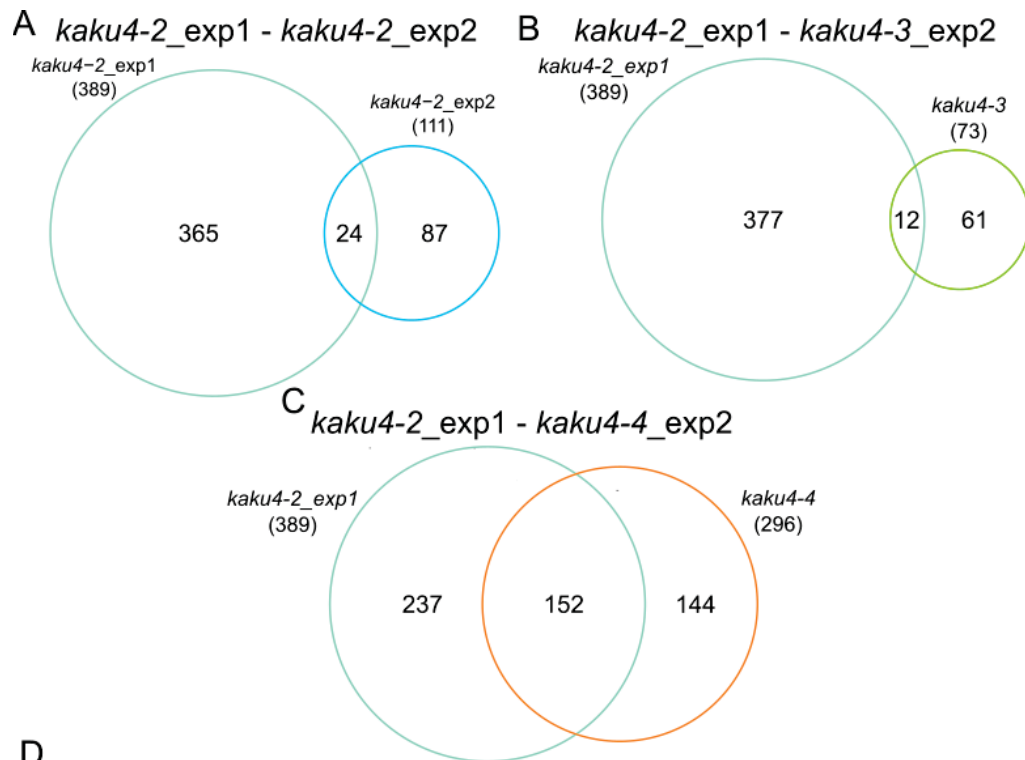
II.2) Comparaison des différents allèles mutants *kaku4*

II.2.a) RNA-seq de trois allèles mutants de *kaku4* et d'une lignée complémentée

Afin de pallier l'impact potentiel de la duplication de la lignée portant l'allèle *kaku4-2* et de déterminer les gènes dérégulés communs aux différents mutants de *KAKU4*, nous avons réalisé

Tableau 6 : Liste des 24 gènes dérégulés communs dans les différents allèles de *kaku4*

ID TAIR	Dérégulation	Nom
AT3G14210	up	EPITHIOSPECIFIER MODIFIER 1 (ESM1)
AT1G12610	down	DWARF AND DELAYED FLOWERING 1 (DDF1)
AT1G18300	down	NUDIX HYDROLASE HOMOLOG 4 (NUDT4)
AT1G33760	down	ETHYLENE RESPONSE FACTOR 022 (ERF022)
AT1G35140	down	PHOSPHATE-INDUCED 1 (PHI-1)
AT1G66500	down	Pre-mRNA cleavage complex II
AT1G73540	down	NUDIX HYDROLASE HOMOLOG 21 (NUDT21)
AT1G74450	down	ROH1D
AT2G27080	down	NDR/HIN1-LIKE 13 (NHL13)
AT2G30020	down	AP2C1
AT3G44260	down	CCR4- ASSOCIATED FACTOR 1A (CAF1A)
AT3G46620	down	RING AND DOMAIN OF UNKNOWN FUNCTION 1117 1 (RDUF1)
AT3G55840	down	HS1PRO-1 PROTEIN
AT3G61190	down	BON ASSOCIATION PROTEIN 1 (BAP1)
AT4G25470	down	DRE/CRT-BINDING PROTEIN 1C (DREB1C)
AT4G25810	down	XYLOGLUCAN ENDOTRANSGLYCOSYLASE 6 (XTR6)
AT4G27652	down	hypothetical protein
AT4G27657	down	hypothetical protein
AT4G29780	down	affected by multiple stresses
AT4G31430	down/up	KAKU4
AT5G17350	down	PADRE protein up-regulated after infection by <i>S. sclerotiorum</i>
AT5G22250	down	CCR4- ASSOCIATED FACTOR 1B (CAF1B)
AT5G59820	down	RESPONSIVE TO HIGH LIGHT 41 (RHL41)
AT5G64310	down	ARABINOGALACTAN PROTEIN 1 (AGP1)



D

ID TAIR	déregulation exp1/exp2	Nom
AT1G35140	down	PHOSPHATE-INDUCED 1 (PHI-1)
AT1G72910	up/down	TIR-NBS8 (TN8)
AT1G72920	up/down	TIR-NBS9 (TN9)
AT1G80840	up/down	WRKY DNA-BINDING PROTEIN 40 (WRKY40)
AT2G01422	up	
AT2G26530	up/down	AR781
AT2G35490	down	FIBRILLIN2 (FBN2)
AT2G41640	down	GT61
AT3G01345	down	Expressed protein
AT4G11280	up/down	1-AMINOCYCLOPROPANE-1-CARBOXYLIC ACID (ACC) SYNTHASE 6 (ACS6)
AT4G20325	down	ribonuclease H2 subunit B
AT4G20330	up	Transcription initiation factor TFIIIE, beta subunit
AT4G20380	up	LSD1 zinc finger family protein (LSD1) histone demethylase
AT4G20400	up	JUMONJI 14 (JM14) H3K4 demethylase
AT4G20410	up	gamma-soluble NSF attachment protein (GSNAP)
AT4G20440	up	small nuclear ribonucleoprotein associated protein B (smB)
AT4G20480	up	Putative endonuclease or glycosyl hydrolase
AT4G25810	down	XYLOGLUCAN ENDOTRANSGLYCOSYLASE 6 (XTR6)
AT4G31430	down	KAKU4
AT4G39640	down	GAMMA-GLUTAMYL TRANSPEPTIDASE 1 (GGT1)
AT5G04340	up/down	COLD INDUCED ZINC FINGER PROTEIN 2 (CZF2)
AT5G17350	up/down	PADRE protein up-regulated after infection by <i>S. sclerotiorum</i>
AT5G35935	up	transposable_element_gene
AT5G57560	down	XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASE 22 (XTH22)

Insertion
T-DNA
Duplication

Figure 2.21 : Comparaison des gènes dérégulés dans les mutants *kaku4-2* dans deux expériences différentes

A-C. Diagrammes de Venn de l'ensemble des gènes dérégulés par rapport au sauvage entre le mutant *kaku4-2* de l'expérience 1 et les différents mutants *kaku4* de l'expérience 2. Comparaison entre *kaku4-2_exp1* et *kaku4-2_exp2* (**A**), *kaku4-2_exp1* et *kaku4-3* (**B**) et *kaku4-2_exp1* et *kaku4-4* (**C**). La taille de chaque ellipse est proportionnelle au nombre de gènes dérégulés dans nos données de RNA-seq ($FC > 1.7$ ou $FC < 0.58$; $padj < 0.05$). **D.** Tableau des 24 gènes dérégulés par le mutant *kaku4-2* dans les deux expériences ; gènes surlignés en vert représentant les gènes également dérégulés entre les mutants *kaku4-2*, *kaku4-3* et *kaku4-4* de l'expérience 2.

une analyse transcriptomique des différents allèles de *kaku4*, c'est-à-dire les mutants d'insertion *kaku4-2* (4^{ème} intron), *kaku4-3* (2^{ème} intron) et *kaku4-4* (13^{ème} exon) ainsi que la complémentation de l'un de ces allèles, *kaku4-2 ; pKAKU4 :KAKU4-tRFP* (nommé ensuite *kaku4-2 ; pK4:K4-tRFP*, qui contient la duplication) (**Figure 2.20A**). L'expérience de RNA-seq a été réalisée selon les conditions de notre première expérience (cf paragraphe II.1.a) de même que pour l'analyse permettant de déterminer les changements d'expression (sur-exprimés : Fold Change supérieur à 1,7 ; sous-exprimés : Fold Change inférieur à 0,58 ; avec une p-value inférieure à 0,05 par rapport au Col0). Lors du séquençage de nos ARN, suite à un problème technique sur la plateforme de séquençage, le séquençage a été obtenu en 150 bp en single-end. Etant donné la bonne qualité et profondeur du séquençage, nous avons décidé d'analyser les données obtenues. Dans cette expérience, le nombre de gènes dérégulés par le mutant *kaku4-2* est plus faible que lors du premier séquençage. De plus, très peu de gènes sont trouvés dérégulés dans le mutant *kaku4-3*, avec seulement 73 gènes dérégulés alors que le mutant *kaku4-4* en dérégule 296. J'ai, dans un premier temps, recoupé les listes de gènes dérégulés dans ces trois mutants *kaku4* via des diagrammes de Venn afin de déterminer les gènes communs aux différents mutants et donc spécifiques à *kaku4* (**Figures 2.20B-D**). Seuls 24 gènes sont dérégulés dans les trois allèles de *kaku4*, ce qui représente 22% des gènes du mutant *kaku4-2*, 33 % pour le mutant *kaku4-3* et 8% pour *kaku4-4*, ce qui est faible pour trois allèles du même gène (**Figures 2.20B-D et Tableau 6**). L'analyse d'enrichissement de gene ontology m'a permis de montrer que les trois allèles mutant de *KAKU4* dérégulent principalement des gènes impliqués dans la réponse aux stimuli aussi bien biotiques qu'abiotiques (**Figure 2.19E**). Les différences dans le nombre de gènes dérégulés entre les différents allèles de *kaku4* peuvent signifier que *KAKU4* n'a pas un effet majeur sur le transcriptome et induit seulement une variation de l'expression des gènes sensibles à l'environnement.

Le nombre de gènes dérégulés pour le même mutant *kaku4-2* a grandement varié entre les deux expériences avec un total de 111 gènes dérégulés dans cette analyse alors qu'il en avait 389 dans la première, avec seulement 24 gènes en commun entre les deux expériences (**Figure 2.21A**). En comparant les gènes dérégulés entre le mutant *kaku4-2*_expérience1 avec ceux des mutants *kaku4-3* ou *kaku4-4*, 16 et 152 gènes sont respectivement en commun (**Figures 2.21B-C**). Parmi ces 24 gènes communs entre les deux mutants *kaku4-2*, 29% sont inversement dérégulés (7/24 gènes) et seulement 4 sont également retrouvés dans les gènes dérégulés entre les différents allèles de *kaku4* (**Figure 2.21D**). A nouveau, ceci suggère que *KAKU4* n'a pas

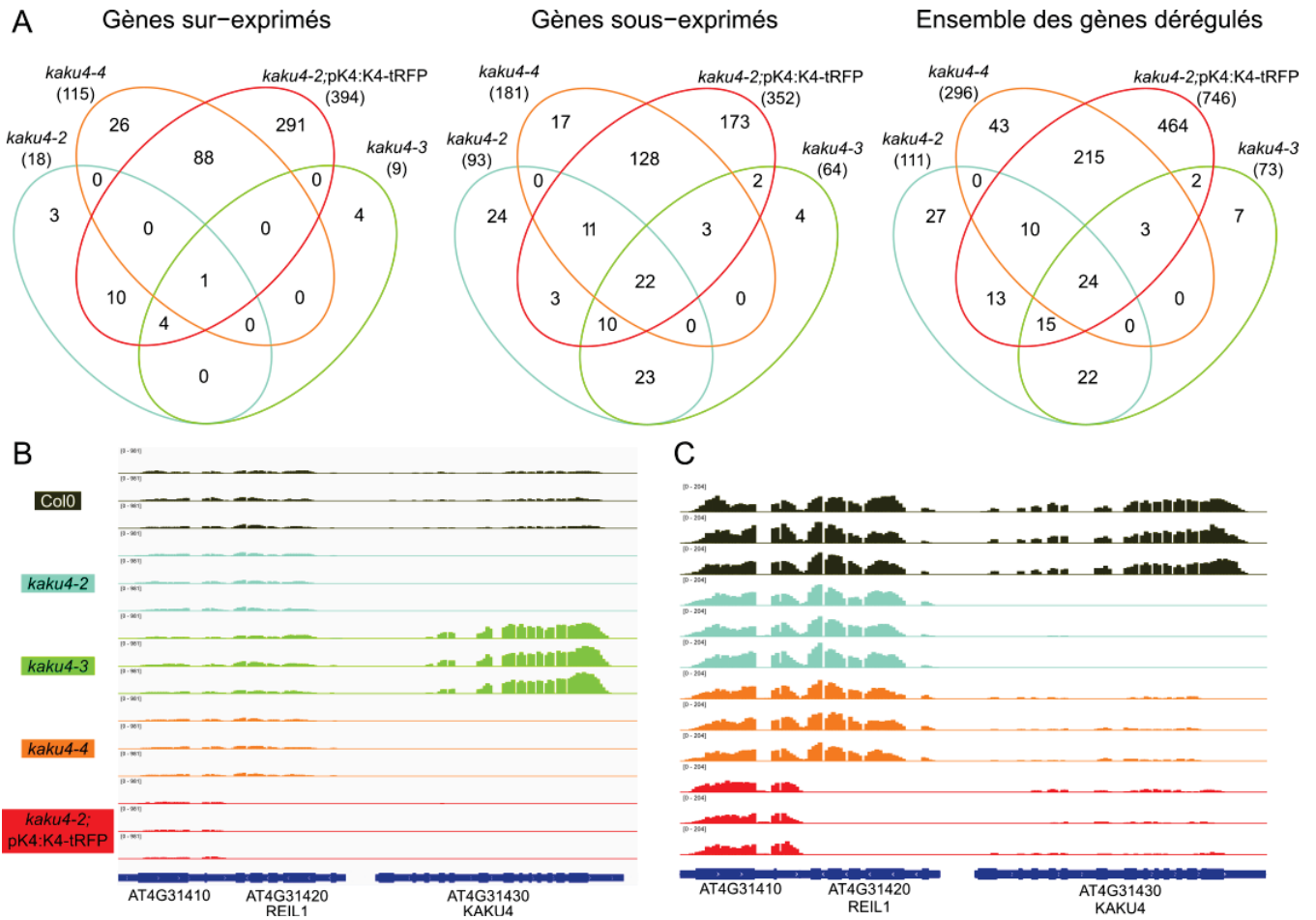


Figure 2.22 : Analyse transcriptomique des différents allèles de *kaku4-2* et du mutant complémenté *kaku4-2 ; pKAKU4-KAKU4-tRFP*

A. Diagrammes de Venn des gènes sur-exprimés, sous-exprimés et totaux dérégulés dans les mutants *kaku4-2*, *kaku4-3*, *kaku4-4* et *kaku4-2 ; pKAKU4-KAKU4-tRFP* par rapport au sauvage à partir de nos données de RNA-seq ($FC > 1.7$ ou $FC < 0.58$; $padj < 0.05$). **B-C.** Visualisation des reads au niveau du gène *KAKU4* avec IGV (Integrative Genomics Viewer, Robinson *et al.*, 2011) des différents génotypes analysés (**B**) et sans *kaku4-3* (**C**).

d'effet majeur sur le transcriptome mais pourrait induire de faibles variations du transcriptome et une plus grande sensibilité aux variations environnementales. L'analyse du transcriptome en conditions de stress en contexte sauvage et mutant pour KAKU4 pourrait être une piste pour étudier l'impact de KAKU4 sur le transcriptome.

Pour aller plus loin, j'ai ensuite, comparé les gènes dérégulés par rapport au sauvage, au sein de ces trois mutants de *kaku4* avec ceux du mutant complémenté *kaku4-2 ; pK4-K4-tRFP* (dans ce cas la lignée complémentée contient la duplication). Cette stratégie avait pour objectif de ne sélectionner que les gènes différentiellement dérégulés dans nos trois allèles mutants par rapport au mutant complémenté. De façon surprenante, le nombre total de gènes dérégulés dans la lignée complémentée est élevé par rapport aux autres mutants *kaku4*, avec 394 gènes, ce qui fait de ce mutant celui qui a le plus de différence avec le génotype sauvage dans cette expérience. De plus, les gènes dérégulés dans les trois allèles *kaku4* le sont également dans la lignée complémentée (**Figure 2.22A**).

L'analyse d'alignement des reads du RNA-Seq sur le génome d'*A. thaliana* en utilisant l'outil Integrative Genomics Viewer (IGV) nous a permis de faire des découvertes intéressantes sur l'expression de KAKU4. En effet, dans le mutant *kaku4-3*, l'expression de KAKU4 qui est 5 fois supérieure à celle du sauvage, s'initie au niveau du 3^{ème} exon après l'insertion du T-DNA (**cf Figure 2.22A et Annexe 1**). Cette expression pourrait correspondre à l'expression d'une protéine tronquée qui serait dépourvue du motif M3 (**Figure 2.22B**). Le mutant *kaku4-2* ne possède pas ou très peu de transcrit au niveau du gène *KAKU4*, le mutant *kaku4-4*, lui possède peu de transcrits sur l'ensemble du gène sauf pour le dernier exon au niveau de l'insertion du T-DNA (**Figure 2.23C**), il est donc possible qu'une protéine tronquée soit faiblement produite. Quant au mutant complémenté, *kaku4-2 ; pK4:K4-tRFP*, *KAKU4* s'exprime mais très faiblement. Cette faible expression pose des questions quant à la réelle complémentation de l'absence de *KAKU4*. Ceci peut s'expliquer par le fait que la région promotrice choisie par l'équipe de Kentaro Tamura pour exprimer *KAKU4* contient aussi le gène *AT4G31420 (REIL1)* en orientation antisens par rapport à la transcription de *KAKU4*. Or, notre analyse RNA-Seq montre que le gène *REIL1* est silencé dans la lignée complémentée alors qu'il ne l'est pas dans le mutant *kaku4-2*, suggérant que le transgène *pK4:K4-tRFP* et le gène *REIL1* pourraient tous les deux être soumis à du silencing (**Figure 2.22C**). Ceci expliquerait la faible expression du transgène qui s'est sans doute accentuée au cours des générations successives.

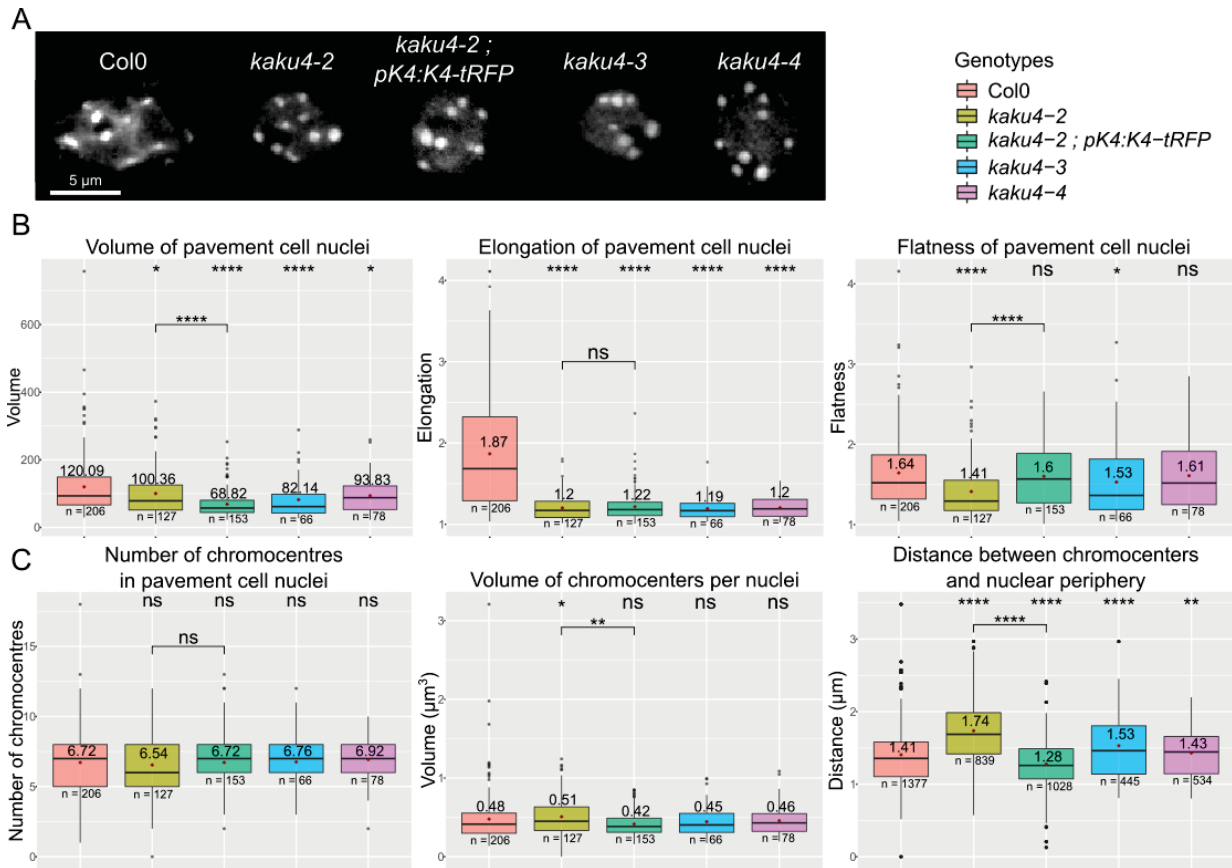


Figure 2.23 : Morphologie nucléaire des différents allèles de *kaku4*

A. Projections en Z de noyaux bruts de cellules pavimenteuses des différents allèles de *kaku4*, sélectionnés pour leurs paramètres proches des valeurs moyennes pour les différents critères morphologiques ; échelle : 5µm. **B.** Boxplots du volume, de l'aplatissement (flatness) et de l'élongation des noyaux de cellules pavimenteuses après analyse avec le plugin Fiji, NucleusJ2 (Dubos *et al.*, 2020). **C.** Boxplots du nombre de chromatocentres, de leur volume ainsi que de leur distance à la périphérie nucléaire dans les noyaux de cellules pavimenteuses après analyse avec le plugin Fiji, NODeJ (Dubos *et al.*, 2022) ; avec la moyenne affichée et représentée par un losange rouge ; n : nombre d'individus analysés par génotype ; Wilcoxon test, ns : non significatif (p-value > 0.05), * < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001.

Cette analyse RNA-seq sur différents allèles de *KAKU4* au lieu de confirmer des gènes cibles potentiels a malheureusement généré encore plus de questions : quels résultats de l'analyse choisir ? quels allèles mutants choisir au début d'une analyse RNA-Seq ? quelle référence choisir pour définir les gènes différentiellement exprimés ? D'une manière générale nos résultats suggèrent que le gène *KAKU4* n'a pas un impact majeur sur la régulation de l'expression des gènes. Nous trouvons bien des dérégulations de gènes liés à la réponse à des changements environnementaux de la plante, mais ceux-ci varient grandement en fonction des expériences.

[II.2.b\) Impact sur la morphologie nucléaire des différents allèles de *kaku4*](#)

Comme nous l'avons vu précédemment, il est possible que les faibles niveaux d'expression de *KAKU4* dans la lignée *kaku4-2 ; pK4:K4-tRFP* ne soit pas suffisants pour une complémentation totale. Nous avons donc cherché à déterminer s'il y avait complémentation des phénotypes de noyaux ronds provoqués par le mutant *kaku4-2* comme cela a été montré par Goto *et al.* Pour cela, j'ai analysé la morphologie nucléaire des différents mutants de *kaku4* ainsi que du mutant complémenté dans des cotylédons de 14 jours colorés au DAPI, les noyaux de l'épiderme ont été imagés en 3D avant d'être analysés par les plugins NucleusJ2 et NODEJ (Dubos *et al.*, 2020, 2022) ; seules les données des noyaux des cellules pavimenteuses ont été étudiées (**Figure 2.23A**). Les différents allèles mutants de *KAKU4* provoquent plus ou moins les mêmes effets avec la diminution de l'élongation et de l'aplatissement des noyaux induisant donc des noyaux sphériques, cependant le nombre de chromocentres et leur volume varient peu alors que leur positionnement au sein du noyau change ; en effet, la distance entre les chromocentres et la périphérie nucléaire augmente (**Figures 2.23B-C**). Quant au mutant *kaku4-2 ; pK4:K4-tRFP*, il restaure l'aplatissement mais pas l'élongation des noyaux et diminue la distance entre les chromocentres et la périphérie nucléaire, ses chromocentres sont plus proches de la périphérie nucléaire que ceux du mutant *kaku4-2* et même que le sauvage (**Figures 2.23B-C**). Ces résultats montrent que les différents allèles mutants *kaku4* ont le même impact sur la morphologie nucléaire, cependant le mutant complémenté ne restaure que très partiellement ce phénotype confirmant que l'expression du transgène n'est pas suffisante pour restaurer la totalité de la fonction de *KAKU4*. Toutefois, ces résultats sont à mesurer puisque durant l'analyse, lors de l'étape de segmentation, de nombreux noyaux ont été éliminés par le logiciel car le contraste du marquage de l'ADN entre l'euchromatine et l'hétérochromatine est très important dans ces noyaux (hétérochromatine très intense et euchromatine faiblement intense) et ne permet pas la segmentation du noyau en entier.

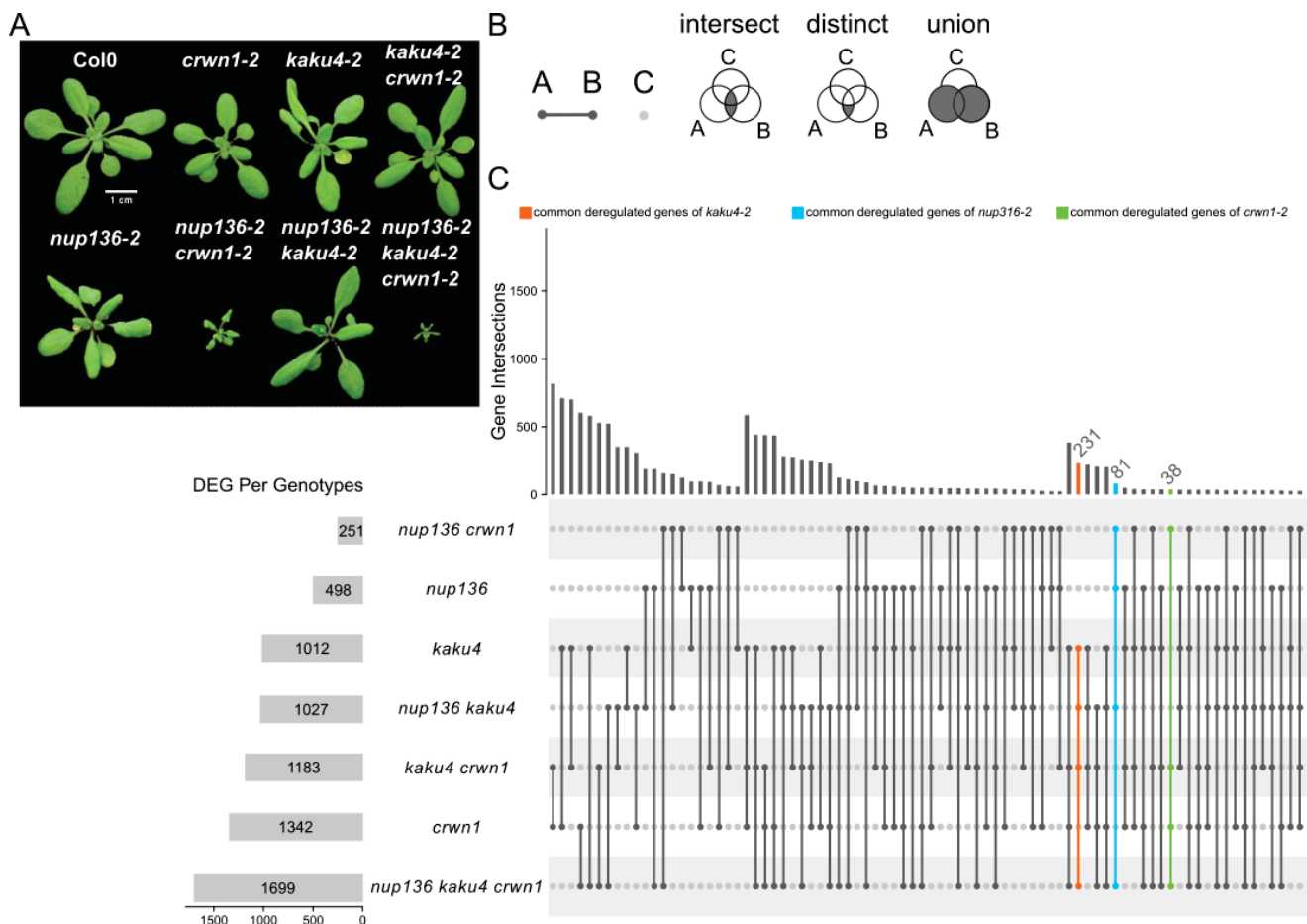


Figure 2.24 : Analyse transcriptomique des multi-mutants impliquant les gènes *NUP136*, *KAKU4* et *CRWN1*

A. Photos de la rosette des différents mutants à 21 jours, échelle : 1cm (similaire à la Figure 2.9A). **B.** Explication des différentes possibilités de recouvrements des gènes dérégulés dans les différents mutants selon le diagramme d'Upset. **C.** Diagramme d'Upset d'intersection des 7 mutants : *nup136-2*, *kaku4-2*, *crwn1-2*, *nup136-2 crwn1-2*, *kaku4-2 crwn1-2*, *nup136-2 kaku4-2* et *nup136-2 kaku4-2 crwn1-2* et visualisation des intersections de l'ensemble des mutants impliquant *kaku4-2* (orange), *nup136-2* (bleu) et *crwn1-2* (vert).

II.3) Étude des multi-mutants *kaku4-2*, *nup136-2* et *crwn1-2*

II.3.a) Analyse du transcriptome des multi-mutants *kaku4-2*, *nup136-2* et *crwn1-2*

Nous avons observé qu'en combinant les mutations *nup136-2*, *kaku4-2* et *crwn1-2*, les plantes sont naines, ont une floraison précoce, des défauts au niveau reproductif de la taille des siliques, du nombre de graines, du pollen ainsi que des noyaux plus sphériques.

Nous avons donc souhaité connaître l'impact de la combinaison de ces mutations sur l'expression des gènes, pour cela nous avons séquencé le transcriptome de trois simples mutants (*nup136-2*, *kaku4-2ad* et *crwn1-2*), de trois doubles mutants (*nup136-2 kaku4-2*, *nup136-2 crwn1-2* et *kaku4-2ad crwn1-2*) et du triple mutant *nup136-2 kaku4-2ad crwn1-2* (**Figure 2.24A**). Les données de transcriptomique ont été obtenues et analysées comme énoncé dans le paragraphe I.1.c.i), les gènes sont considérés comme dérégulés comme ayant un Fold Change supérieur à 2 pour les gènes sur-exprimés et inférieur à 0,5 pour les gènes sous-exprimés avec une p-value inférieure à 0,05 par rapport au génotype sauvage (Col0). Le mutant *nup136-2 crwn1-2* est la combinaison parmi les doubles mutants qui a le plus fort impact sur le développement et la croissance des plantes après le triple mutant, cependant c'est le mutant qui a le moins de gènes dérégulés avec seulement 251 gènes dérégulés comme cela avait été observé avec le triple mutant *kaku4-2 crwn1-2 crwn4-1*, dans la première analyse RNA-seq (cf paragraphe II.1.a). Le triple mutant *nup136-2 kaku4-2 crwn1-2* quant à lui est le mutant qui a le plus de gènes dérégulés de l'analyse avec 1699 gènes dérégulés. Afin de comparer l'ensemble des mutants entre eux dans une même analyse, j'ai réalisé un diagramme d'Upset d'intersection qui permet de recouper les gènes dérégulés dans les différents mutants (**Figures 2.24B-C**).

Dans une première approche, j'ai choisi de définir l'ensemble des gènes dérégulés présents dans les mutants impliquant une même mutation. Par exemple, pour le gène *KAKU4*, j'ai cherché les gènes dérégulés communs entre *kaku4-2*, *nup136-2 kaku4-2*, *kaku4-2 crwn1-2* et *nup136-2 kaku4-2 crwn1-2*. Cette première analyse a permis de définir 231 gènes dérégulés pour *kaku4-2*, 81 pour *nup136-2* et 32 pour *crwn1-2* (**Figure 2.24C**). J'ai ensuite réalisé une étude d'enrichissement des termes de la gene ontology de chaque ensemble de gènes. Cette analyse permet de montrer que les mutants *kaku4-2* et *crwn1-2* dérégulent des gènes impliqués dans la réponse aux stimuli et aux stress particulièrement biotiques pour *crwn1-2* avec notamment la réponse à l'acide salicylique (**Figures 2.25A-B**) et que le mutant *nup136-2* dérégule aussi des gènes impliqués dans la réponse aux stimuli ainsi que dans le transport nucléocytoplasmique dont certains sont engagés dans la réponse aux stimuli (**Figure 2.25C**).

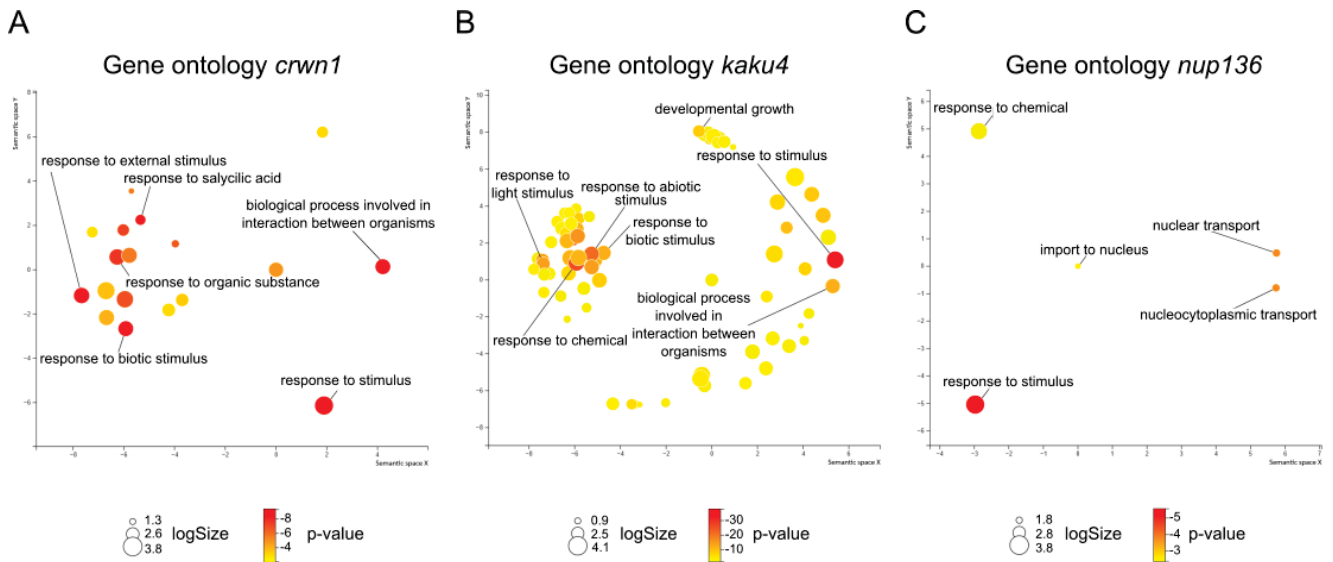


Figure 2.25 : Gene ontology des gènes communs aux mutants *nup136*, *kaku4* ou *crwn1*

A-C. Représentations de la gene ontology des gènes dérégulés communs aux différents mutants impliquant *crwn1* (A), *kaku4* (B) et *nup136* (C). Analyse de la gene ontology réalisée avec LAGO puis visualisée avec REVIGO (cf figure 2.18).

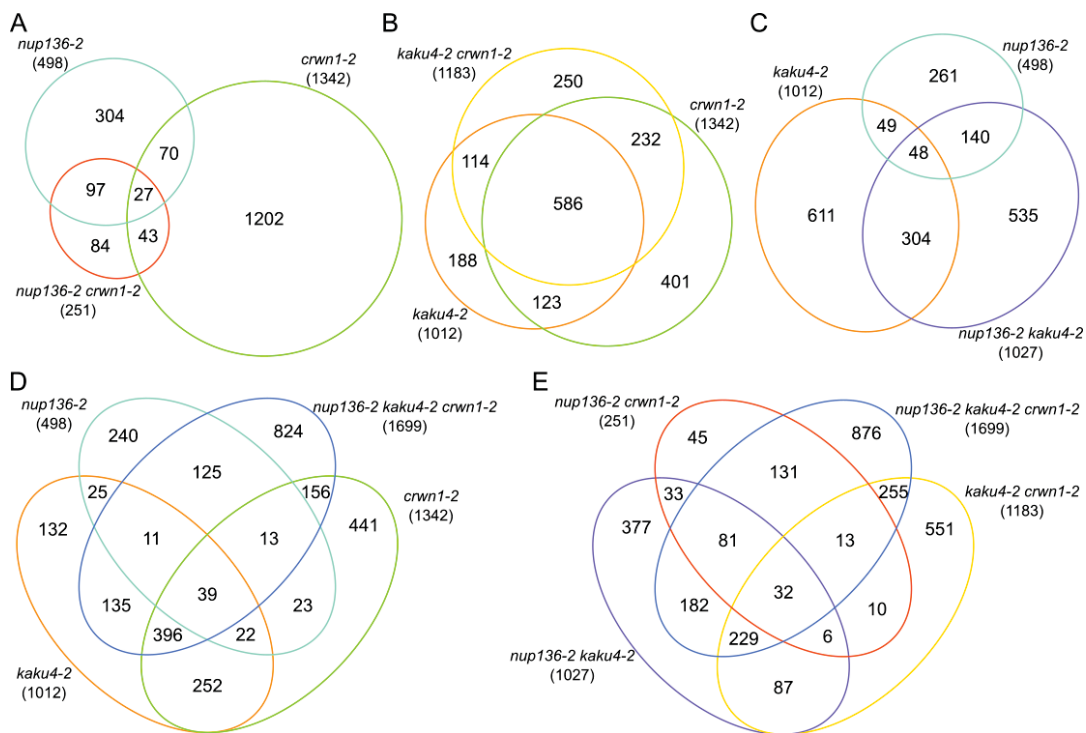


Figure 2.26 : Comparaison des gènes dérégulés entre mutants

A-E. Diagrammes de Venn comparant les gènes dérégulés entre les doubles mutants et les simples mutants les composant (A-C) avec *nup136-2 crwn1-2* (A), *kaku4-2 crwn1-2* (B) et *nup136-2 kaku4-2* (C) et des gènes dérégulés dans le triple mutant *nup136-2 kaku4-2 crwn1-2* avec ceux des simples mutants (D) et des doubles mutants (E). La taille de chaque ellipse (A-C) est proportionnelle au nombre de gènes dérégulés dans nos données de RNA-seq ($FC > 2$ ou $FC < 0.5$; $padj < 0.05$).

Dans un deuxième temps, j'ai déterminé si les doubles mutants dérèglent les mêmes gènes que les simples mutants les composant. Par exemple, j'ai comparé les gènes dérèglés dans le double mutant *nup136-2 crwn1-2* avec ceux des simples mutants *nup136-2* et *crwn1-2*. Dans ce cas, 66% des gènes dérèglés dans *nup136-2* et *crwn1-2* se retrouvent chez le double mutant. L'impact du mutant *nup136-2* est prédominant, ce qui corrèle bien avec nos études phénotypiques où le phénotype du double mutant s'explique principalement par les défauts phénotypiques induits par la mutation de *NUP136* (**Figure 2.26A**). Pour ce qui est du double mutant *kaku4-2 crwn1-2*, 78% de ces gènes dérèglés le sont aussi dans l'un des simples mutants et 49% dans les deux simples mutants et avec une légère prédominance du mutant *crwn1-2*, confirmant la fonction commune de KAKU4 et CRWN1 dans le nucléosquelette (**Figure 2.26B**). Dans le cas du mutant *nup136-2 kaku4-2*, 47% des gènes sont également dérèglés dans les simples mutants mais seulement 4% sont communs aux deux simples mutants et avec une prédominance nette du mutant *kaku4-2* alors qu'au niveau phénotypique de la plante, ce mutant se rapproche plus du mutant *nup136-2* (**Figure 2.26C**). Enfin sur le même principe, j'ai réalisé la comparaison du triple mutant face aux simples et doubles mutants. Dans ce cas, 51% des gènes dérèglés dans le triple mutant *nup136-2 kaku4-2 crwn1-2* le sont dans les trois simples mutants, alors que seulement 48% le sont dans les doubles mutants. Une certaine prédominance est observée pour le simple mutant *crwn1-2* et le double mutant *kaku4-2 crwn1-2* alors que le phénotype de la plante se rapproche plus du double mutant *nup136-2 crwn1-2* (**Figures 2.26D-E**).

Cette nouvelle analyse du transcriptome, nous permet de confirmer que les simples mutants *kaku4-2* et *crwn1-2* dérèglent en grande partie les mêmes gènes. Ces gènes sont distincts des gènes dérèglés observés dans le mutant *nup136-2*. Ceci illustre bien la spécialisation de KAKU4 qui dérive d'une ancienne nucléoporine vers une fonction associée au nucléosquelette (cf paragraphe I.1.c.i). Notre seconde méthode d'analyse montre qu'entre 50 et 80 % des gènes dérèglés dans les doubles mutants le sont aussi dans les simples mutants. Les doubles mutants comportent néanmoins une part non négligeable de gènes dérèglés spécifiques et notamment pour les doubles mutants impliquant *NUP136*. Ceci suggère à nouveau que la mutation d'un composant du NPC (*NUP136*) et de composants du nucléosquelette (*KAKU4*, *CRWN1*) affecte différemment l'expression des gènes. L'un des résultats surprenants de cette analyse est l'observation du petit nombre de gènes dérèglés dans le double mutant *nup136-2 crwn1-2* (251 gènes dérèglés) (**Figure 2.24C**) alors que le triple mutant qui a le plus d'impact phénotypique est aussi celui qui a le plus de gènes dérèglés de l'analyse (1699 gènes dérèglés)

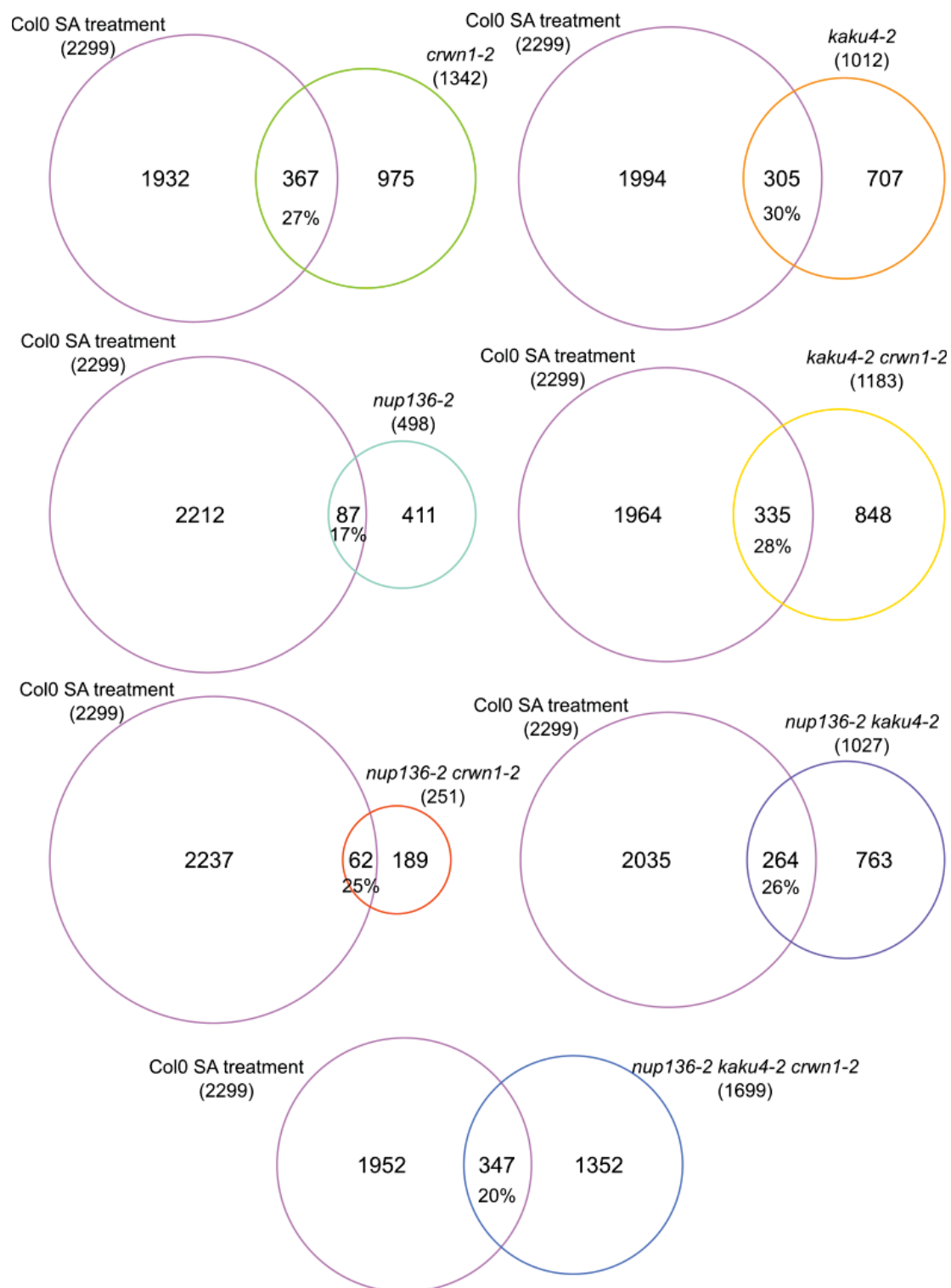


Figure 2.27 : Comparaison des gènes dérégulés entre nos mutants et ceux d'un traitement à l'acide salicylique

Comparaison entre les gènes dérégulés 3h après un traitement au SA (1mM) par *microarray* (Zhou *et al.*, 2015) avec les gènes dérégulés par RNA-seq dans les multi-mutants (sélection des gènes dérégulés de plus de 2 fois avec une p-value ajustée inférieure à 0,05).

(**Figure 2.24C**). Ce résultat suggère une certaine compensation avec moins de gènes dérégulés dans ce double mutant alors que l'impact sur le développement et la croissance des plantes est très fort et presque similaire au triple mutant (**Figure 2.24A**). J'ai recherché si ce résultat pourrait s'expliquer par la dérégulation d'un petit nombre de gènes majeurs mais cette recherche n'a pas donné de conclusion et il faudra sans doute imaginer d'autres hypothèses pour expliquer les résultats obtenus pour ce double mutant. Lors de ces analyses, nous avons remarqué que dans l'ensemble des mutants il y a un plus ou moins faible transcrit correspondant au gène d'intérêt ; or, nous ne savons pas si celui-ci est traduit ou non en protéine et si c'est le cas si cette protéine est fonctionnelle. Afin de remédier à ce problème nous avons initié la création de mutants de délétions des gènes *NUP136*, *NUP82*, *KAKU4* ou *CRWN1* entiers, *via* la technique de CRISPR/Cas9 (Xing *et al.*, 2014).

II.3.b) Implication de la voie de l'acide salicylique

Nos résultats de transcriptome montrent que les mutants des gènes *KAKU4* et *CRWN1* dérégulent la réponse au stress biotique notamment la voie de l'acide salicylique (*Salicylic Acid*, SA) comme cela a déjà été montré pour *CRWN1* et *NUP136* (Tamura *et al.*, 2017; Choi, Strickler and Richards, 2019). J'ai donc décidé de comparer les gènes dérégulés dans nos différents mutants avec ceux dérégulés lors d'un traitement au SA comme cela a été fait pour les mutants *crwn* (Choi, Strickler and Richards, 2019). Les gènes dérégulés par le SA ont été obtenus à partir des données de microarray de Zhou *et al.* disponibles dans la base de données NCBI Gene Expression Omnibus (GSE61059). Il s'agit d'une comparaison entre des plantes sauvages de 23 jours 3h après traitement au SA (1mM) ou à l'eau (Zhou *et al.*, 2015). Les gènes dérégulés de plus de 2 fois avec une p-value ajustée inférieure à 0,05 ont été criblés avec GEO2R en utilisant les options par défaut (Barrett *et al.*, 2013). La liste résultante a été convertie en identifiants de gène TAIR par l'outil TAIR, *Microarray Elements Search and Download*, puis ces gènes ont été comparés à nos données RNA-seq des multi-mutants. Lors de ces comparaisons, j'ai observé un recoupement allant de 17% pour le mutant *nup136-2* à 30% pour le mutant *kaku4-2*, la voie de réponse au SA est donc bien impactée dans nos multi-mutants (**Figure 2.27**). En comparaison, il a été montré un recoupement de 40% pour les gènes sur-exprimés et de 25% pour les sous-exprimés entre le mutant *crwn1-1 crwn2-1* et les gènes dérégulés lors d'un traitement au SA (Choi, Strickler and Richards, 2019). La cible principale de la voie du SA est le gène *PATHOGENESIS-RELATED GENE 1 (PR1)* dont le produit possède une activité antimicrobienne (Pečenková *et al.*, 2022). En présence de SA et du coactivateur transcriptionnel *NONEXPRESSOR OF PATHOGENESIS-RELATED GENES 1*

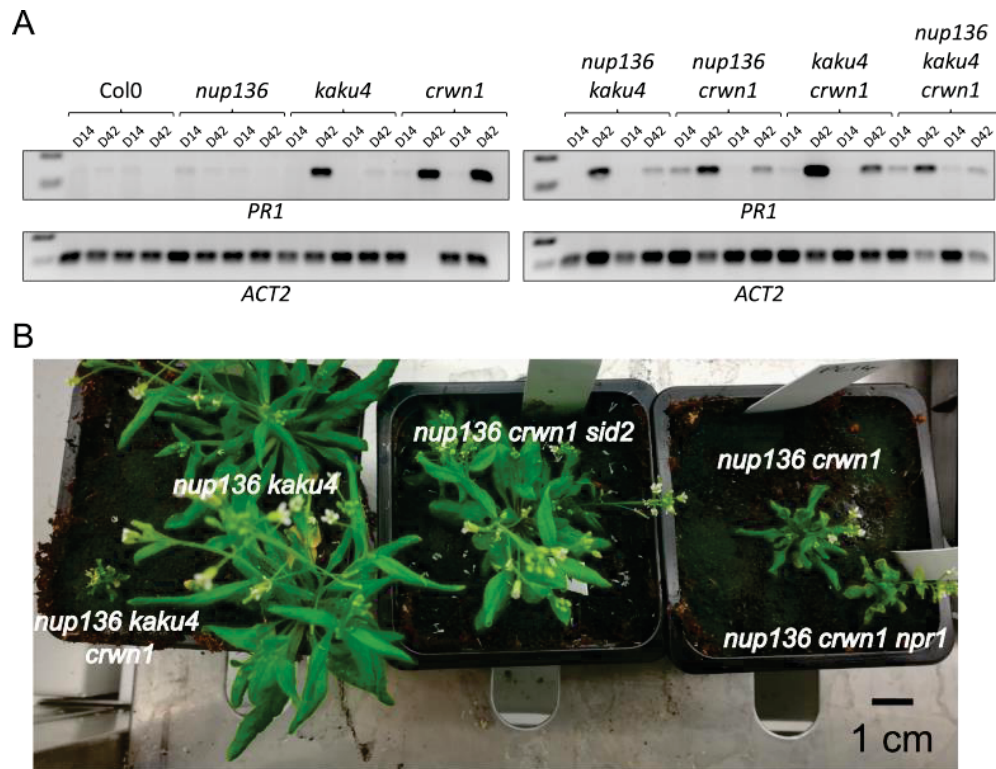


Figure 2.28 : Implication de la voie de l'acide salicylique dans le mutant *nup136 crwn1*

A. RT-PCR du gène *PR1*, gène cible principal de la voie de réponse au SA et de *ACT2* (actine) dans les différents mutants âgés de 14 jours (D14) ou de 42 jours (D42). Deux répliques ont été effectués sur chaque mutant. **B.** Photos des mutants *nup136 crwn1*, *nup136 kaku4*, *nup136 kaku4 crwn1*, *nup136 crwn1 npr1* et *nup136 crwn1 sid2* ; échelle : 1 cm.

(*NPR1*) l'expression du gène *PR1* est activée et augmente au cours du développement végétatif dans les mutants *crwn* (Choi, Strickler and Richards, 2019). Dans nos données de RNA-seq réalisées sur des plantules de 14 jours, ce gène n'est pas dérégulé et nous avons souhaité voir si à un âge plus tardif l'expression augmente également. Pour cela, j'ai réalisé une RT-PCR de *PR1* à partir des ARN envoyés au RNA-seq ainsi que d'ARN de feuilles de plantes âgées de 42 jours. Il y a peu d'expression de *PR1* dans le sauvage et le mutant *nup136-2* mais une forte expression de *PR1* à 42 jours dans l'ensemble des autres mutants confirmant bien la dérégulation de la voie de réponse au SA (**Figure 2.28A**). Cette dérégulation pourrait être quantifiée par une étude en RT-qPCR afin d'être plus précis.

Dans la plante, 90% de l'acide salicylique est produit à partir de la transformation du chorismate en isochorismate par *SALICYLIC ACID INDUCTION DEFICIENT 2 (SID2)* dans le chloroplaste puis transporté dans le cytoplasme pour y être transformé en SA. L'augmentation de SA dans le cytoplasme induit la libération et l'activation de *NPR1* qui va pouvoir se transloquer dans le noyau et ainsi induire l'expression des gènes de réponses comme *PR1*. La mutation de *SID2* ou de *NPR1* induit des plantes de taille sauvage sensibles aux pathogènes alors que la surexpression de *PR1* induit des plantes de petite taille plus résistantes aux pathogènes. Ainsi deux groupes indépendants (Guo *et al.*, 2017; Choi, Strickler and Richards, 2019) ont mis en évidence que le phénotype petite plante observé dans les mutants *crwn* pourrait être lié à la dérégulation de la voie SA. Pour cela ils ont utilisé une mutation dans *sid2* ou *npr1* et montré que ces deux mutations restauraient le phénotype petite plante du double mutant *crwn1 crwn2*. Nous avons donc voulu savoir si le phénotype de petite plante du mutant *nup136-2 crwn1-2* pouvait être supprimé par la mutation de *sid2-1* ou *npr1-1* (mutations ponctuelles induisant une perte de fonction). Le triple mutant *nup136-2 crwn1-2 npr1-1* induit une plante de taille légèrement plus petite que le double mutant *nup136-2 crwn1-2* avec des feuilles plus courbées et plus de senescence visible des feuilles alors que le triple mutant *nup136-2 crwn1-2 sid2-1* restaure totalement la forme des feuilles et partiellement la taille avec une plante proche de celle du double mutant *nup136-2 kaku4-2* (**Figure 2.28B**). Cette observation suggère que le phénotype de petite plante observé pour le mutant *nup136-2 crwn1-2* est médié par la voie du SA mais indépendant de *NPR1*. Les mutants *nup136-2 crwn1-2 npr1-1* et *nup136-2 crwn1-2 sid2-1* restaurent partiellement la taille des siliques et le nombre de graines par silique du mutant *nup136-2 crwn1-2*, se rapprochant ainsi du simple mutant *nup136-2* (**Figures 2.29A-C**). La voie du SA joue donc bien un rôle dans le phénotype du mutant *nup136-2 crwn1-2*. Le blocage de la production du SA restaure partiellement le phénotype de

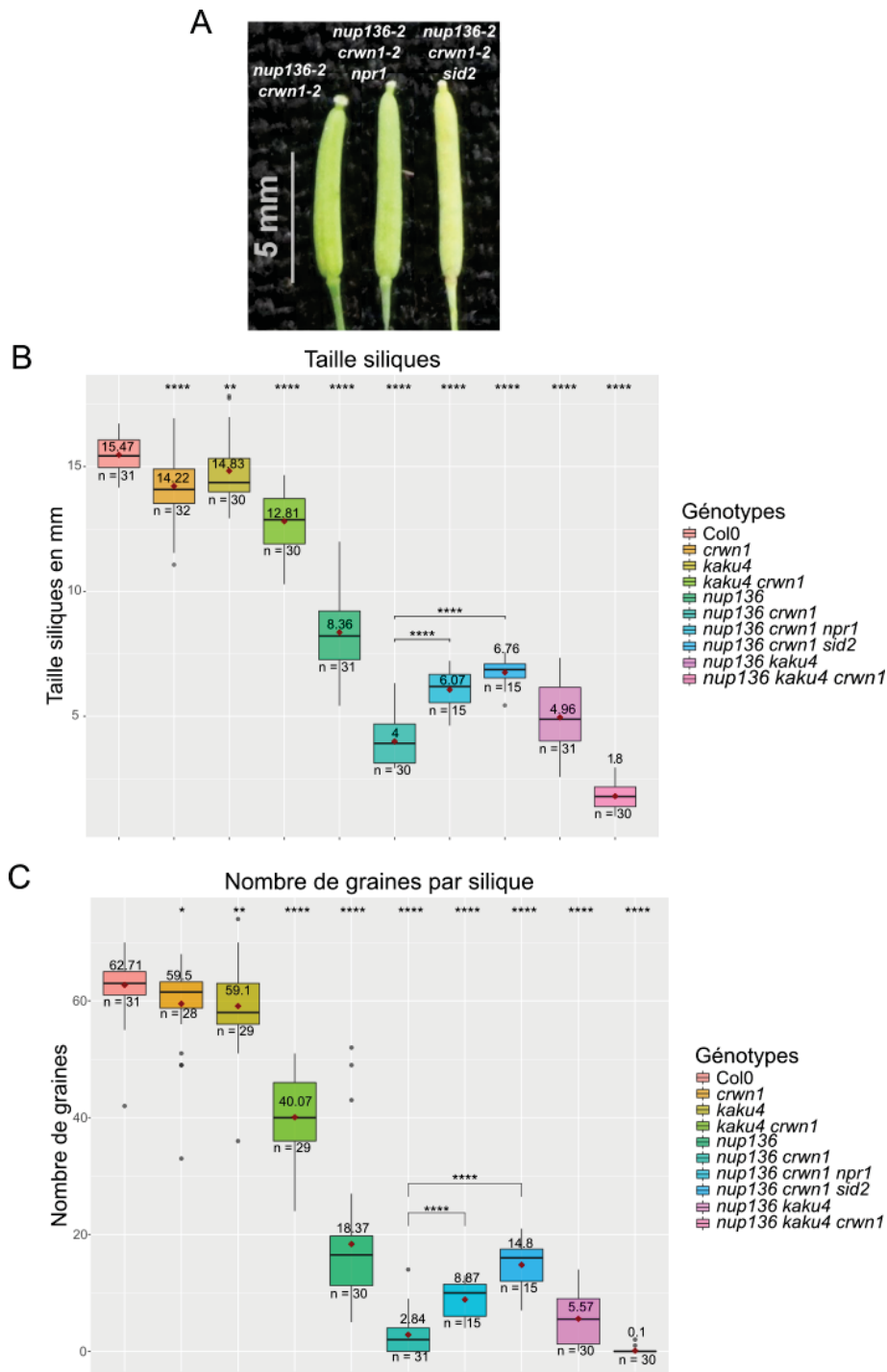


Figure 2.29 : Complémentation partielle du phénotype du mutant *nup136 crwn1* par des mutants de la voie du SA

A. Photo de siliques matures des mutants *nup136 crwn1*, *nup136 crwn1 npr1* et *nup136 crwn1 sid2* ; échelle : 5 mm. **B-C.** Boxplot de la taille des siliques matures après mesure des siliques *via* Fiji (**B**) ou du nombre de graines par silique (**C**) avec la moyenne affichée et représentée par un losange rouge ; n : nombre d'individus analysés par génotype ; Wilcoxon test, ns : non significatif (p -value > 0.05), * < 0.5, ** < 0.01, *** < 0.001, **** < 0.0001.

petite plante. Toutefois, nous observons que ce n'est pas le cas avec NPR1. Une première piste d'investigation pourrait être d'identifier les gènes dérégulés dans le double mutant *nup136-2 crwn1-2* par la voie SA et conduisant au phénotype de petite plante et qui ne seraient pas dérégulés dans le double mutant *crwn1 crwn2*. Ceci pourrait être réalisé en comparant les transcriptomes de ces deux doubles mutants. Une autre hypothèse pourrait être que le transport nucléocytoplasmique est altéré par le mutant *nup136* et en absence de la protéine CRWN1, conduirait à un phénotype de petite plante.

III- Etude préliminaire de MAN1, une protéine de la membrane interne

Pendant longtemps, puisqu'il n'existe pas de lamines chez les plantes, il a été considéré qu'elles ne possédaient pas de protéines LEM, connues pour être impliquées dans le recrutement des lamines et de la chromatine à la périphérie nucléaire chez les animaux. Cependant, une protéine homologue aux protéines LEM, la protéine MAN1 existe chez les plantes, elle ne possède toutefois pas le domaine LEM caractéristique des protéines LEM animales (cf Introduction I.3.a.iii). Cette protéine n'a, pour le moment, été que très peu étudiée chez les plantes, seule une étude partielle impliquant son proxitome, sa localisation à la périphérie nucléaire et le phénotype au stade de la rosette a été publiée (Tang *et al.*, 2021). Afin de caractériser cette protéine MAN1, nous avons essayé de déterminer quelles sont les régions nécessaires à sa localisation et à ses interactions ainsi que d'étudier les phénotypes développementaux et ceux de la morphologie nucléaire de mutants *man1*.

III.1) Etude de la localisation de MAN1

La protéine MAN1 des animaux se localise à la périphérie nucléaire grâce à ses domaines transmembranaires, aussi nous avons cherché à déterminer si la protéine MAN1 d'*Arabidopsis* a la même localisation et quelles sont les régions nécessaires à cette localisation. Ces expériences ont été menées avant la publication de la localisation de la protéine entière réalisée par Tang *et al.* Pour cela nous avons construit des fusions GFP ou mRFP de la protéine MAN1 entière ou déléetée de certaines régions (**Figure 2.30A**). Ces différentes constructions ont été utilisées en expression transitoire chez le tabac et ont permis de montrer que la protéine MAN1 d'*Arabidopsis* se localise à la périphérie nucléaire ; les délétions du signal de localisation nucléaire (NLS) ou du domaine WHD en C-terminal n'affecte pas sa localisation (**Figure 2.30B**). En revanche la délétion de la région centrale incluant les domaines transmembranaires (TM) ainsi qu'une partie du domaine MSC, se localise dans le nucléoplasme et marque très fortement

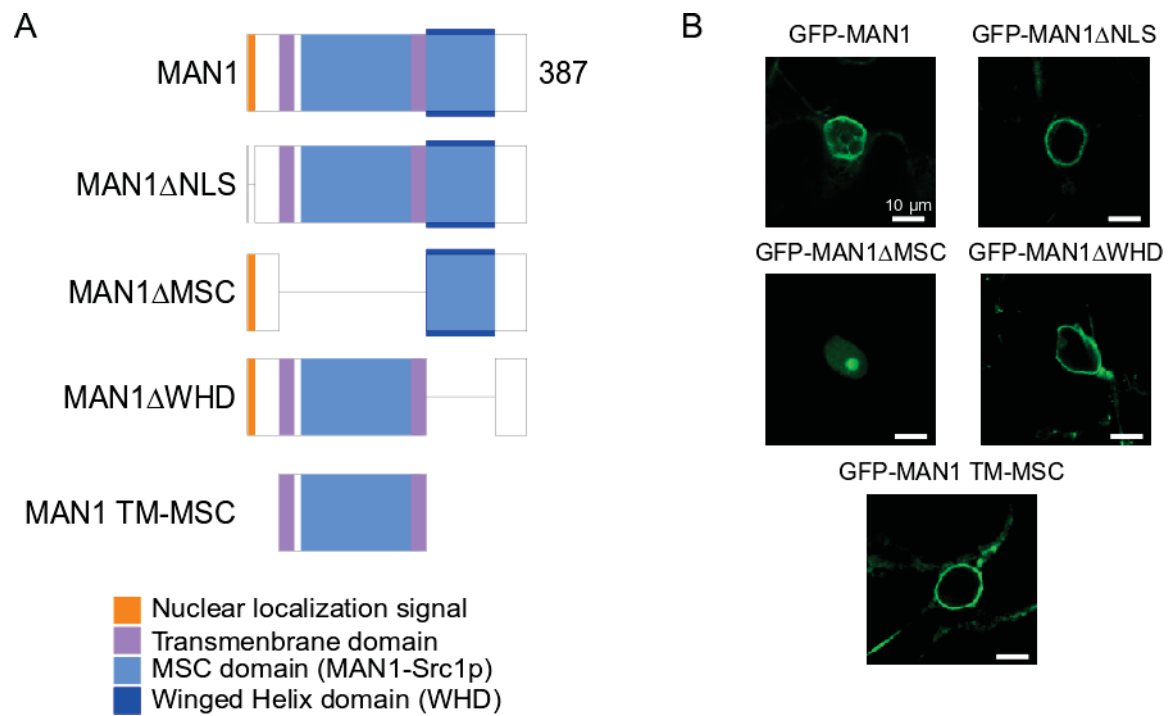


Figure2.30 Localisation de la protéine MAN1

A. Représentation schématique des différentes constructions de la protéine MAN1 utilisées pour étudier sa localisation nucléaire. **B.** Localisation des différentes constructions de MAN1 (pGWB606-p35S:GFP) en expression transitoire chez le tabac ; échelle : 10 μ m.

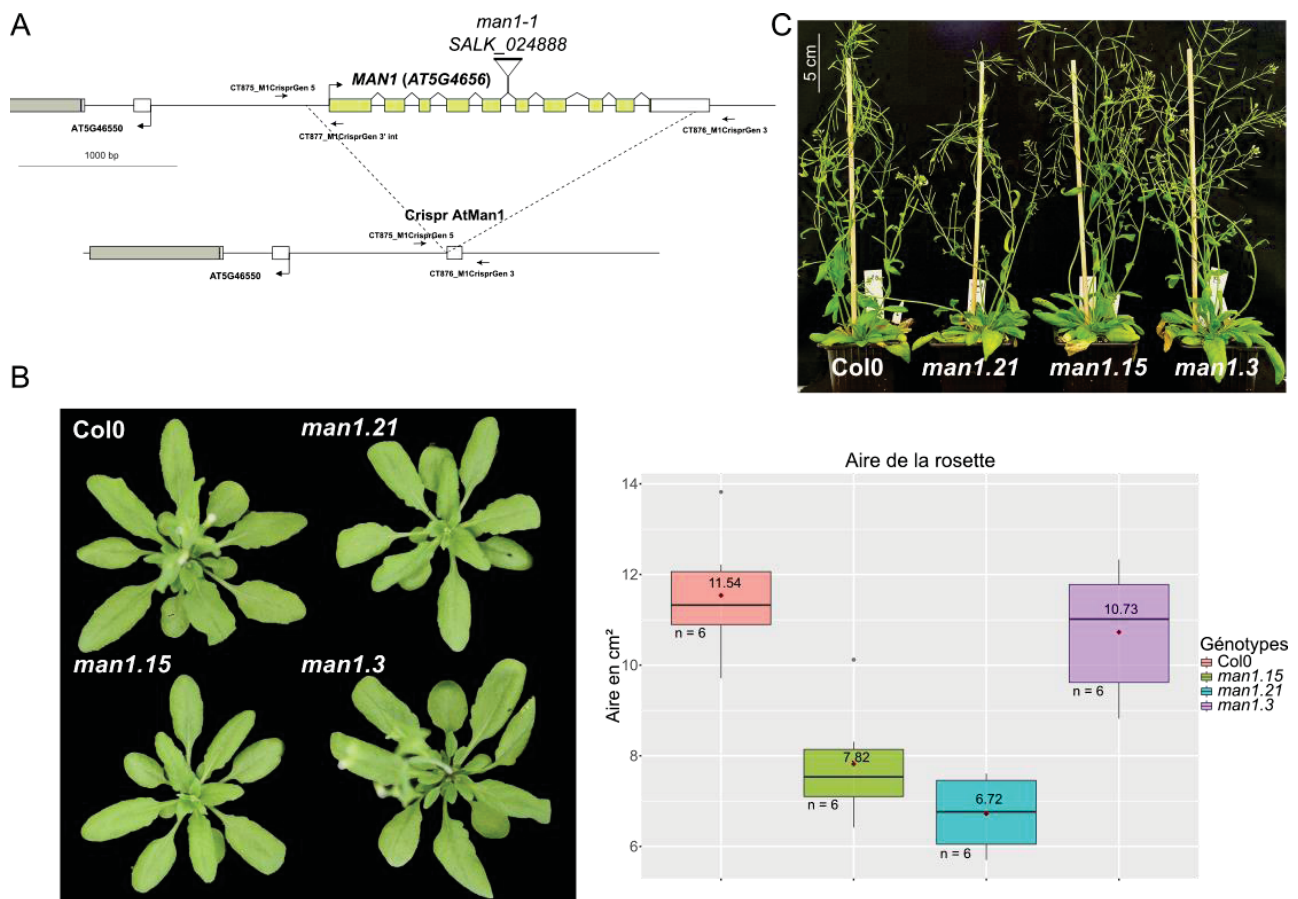


Figure 2.31 : Etude du développement des mutants CRISPR *man1*

A. Représentation schématique du gène *MAN1* pour illustrer la position de l'insertion du T-DNA dans le mutant *man1-1* et la délétion complète du gène effectuée par CRISPR avec les flèches représentant les emplacements des amorces utilisées pour le génotypages de ces lignées.

B. Photos de la rosette des différentes lignées indépendantes obtenues par CRISPR pour *MAN1* à 21 jours et boxplot de la surface de la rosette à 21 jours à partir des données obtenues après traitement des photos par le plugin SIOX disponible sur Fiji avec la moyenne affichée et représentée par un losange rouge ; n : nombre d'individus analysés par génotype.

C. Photos à un stade plus tardif du développement des différentes lignées indépendantes obtenues par CRISPR pour *MAN1*.

le nucléole, l'expression de cette même région, elle, se localise à la périphérie nucléaire (**Figure 2.30B**).

La protéine MAN1 se localise donc bien à la périphérie nucléaire comme nous pouvions l'attendre de par ses homologues avec les protéines de la famille LEM animale et déjà publiée récemment (Tang *et al.*, 2021). Nous avons pu démontrer que la région centrale incluant les domaines TM est suffisante pour cette localisation, il serait maintenant intéressant de définir si, outre les deux domaines TM, il y a des éléments de cette région requis pour cette localisation.

III.2) Caractérisation du mutant *man1*

Pour définir la fonction de la protéine MAN1 chez *A. thaliana*, nous avons choisi de construire un mutant CRISPR délétant l'intégralité du gène (Xing *et al.*, 2014), nous avons sélectionné des lignées mutantes identiques mais indépendantes nommée *man1.3*, *man1.15* et *man1.21* (**Figure 2.31A**). Pour tester l'impact de cette délétion, nous avons réalisé une étude préliminaire du phénotype de ces trois lignées au stade végétatif en comparant la taille des rosettes à 21 jours. Le mutant *man1.3* possède une surface de rosette semblable à celle du sauvage alors que les mutants *man1.15* et *man1.21* sont légèrement plus petites (**Figure 2.31B**). Ces résultats ne permettent pas de montrer un réel effet de cette délétion, cependant les effectifs étudiés sont trop faibles pour réaliser des tests statistiques afin de savoir si ce changement est réellement significatif. De plus, il a récemment été montré que le mutant *man1-1* comportant une insertion T-DNA dans le 6^{ème} intron (**Figure 2.31A**) ne possède pas de défauts de développement au niveau de la rosette (Tang *et al.*, 2021). Plus tardivement, nos trois lignées mutantes *man1* ne possèdent pas de défauts de la taille des hampes ou des siliques (**Figure 2.31C**). Pris ensemble, ces résultats suggèrent que MAN1 ne joue pas un rôle essentiel dans le développement de la plante, bien que nous n'ayons pas analysé de façon précise la taille des siliques ou du nombre de graines par silique aussi si la délétion de MAN1 a un impact, celui-ci est très faible.

Puisque de nombreux facteurs de l'INM ou de la périphérie nucléaire impactent la morphologie nucléaire, nous avons analysé l'impact des délétions de MAN1 sur ce phénotype. De plus, des expériences similaires réalisées sur la protéine LEM2 de mammifère ont montré que dans le mutant Knock-Out (KO), les noyaux avaient un volume accru significativement (Kume *et al.*, 2019). Cette analyse étant assez lourde, nous n'avons utilisé qu'une seule lignée mutante et les noyaux de cotylédons du sauvage et du mutant *man1.3* de 7 jours ont été marqués

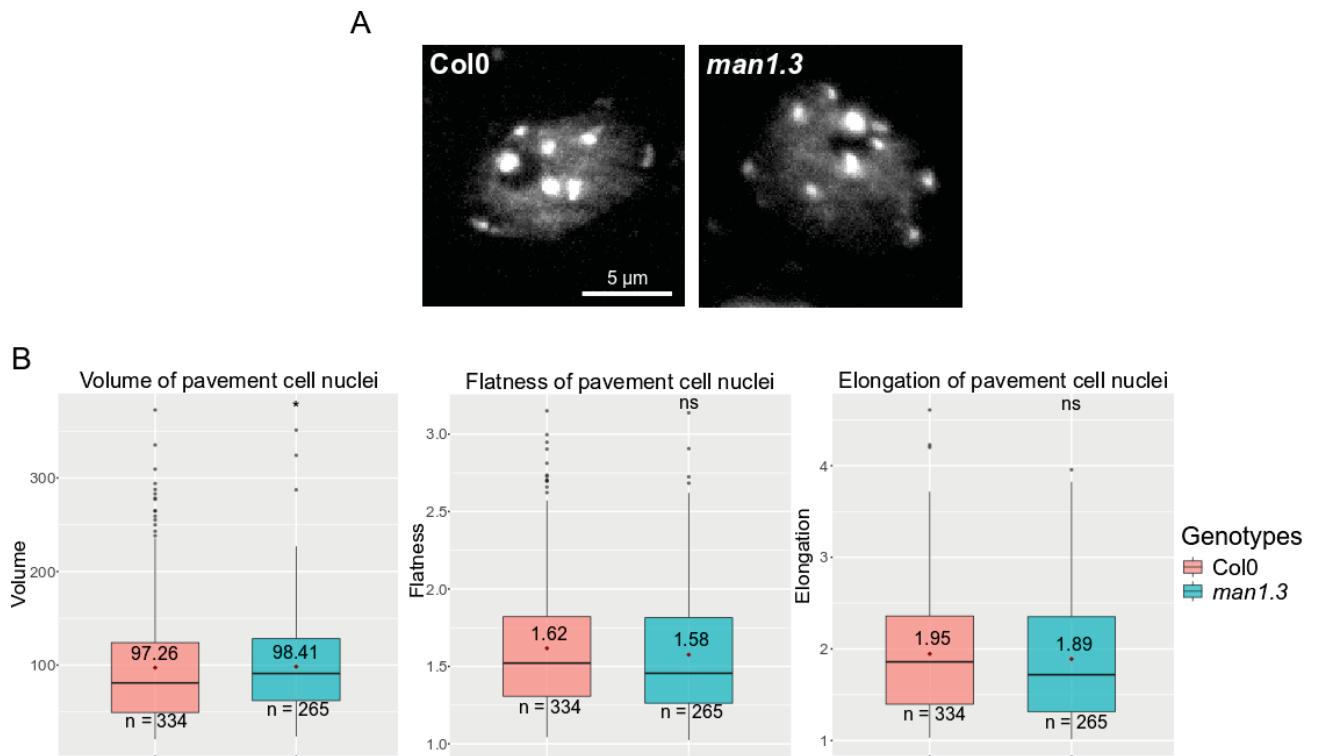


Figure 2.32 : Impact de MAN1 sur la morphologie nucléaire

A. Projections en Z de noyaux bruts de cellules pavimenteuses du sauvage et du mutant *man1.3*, sélectionnées pour leurs paramètres proches des valeurs moyennes des paramètres morphologiques ; échelle : 5 μ m. **B.** Boxplots du volume, de l’aplatissement (flatness) et de l’élargissement des noyaux de cellules pavimenteuses après analyse avec le plugin Fiji, NucleusJ2 (Dubos *et al.*, 2020) avec la moyenne affichée et représentée par un losange rouge ; n : nombre d’individus analysés par génotype ; Wilcoxon test, ns : non significatif (p-value > 0.05), *<0.5, **<0.01, ***<0.001, **** <0.0001.

au DAPI puis analysés avec le plugin NucleusJ2 disponible sur Fiji, seuls les noyaux de cellules de pavement ont été analysés (Dubos et al., 2020). Dans le mutant *man1.3*, le volume des noyaux est très légèrement plus important que celui du sauvage, cependant les autres paramètres comme l'aplatissement ou l'élongation ne changent pas (**Figures 2.32A-B**). Bien que l'impact ne soit pas majeur, nous montrons avec ces données préliminaires qui devront être confirmées que MAN1 joue donc un rôle minime sur la morphologie nucléaire, son KO induit de plus gros noyaux contrairement à la plupart des mutants de la périphérie nucléaire qui ont des noyaux plus petits et sphériques (*CRWN*, *KAKU4*, etc.). Cet impact sur la morphologie n'est cependant pas suffisant pour affecter le développement et la croissance de la plante.

III.3) Détermination des interactions protéiques de MAN1

Récemment l'étude réalisée par Tang *et al.* a caractérisé le proxitome de différentes protéines de l'INM chez *A. thaliana* par protéomique de marquage de proximité, et montré que la protéine MAN1 est proche des protéines PNET2, CRWN1 et CRWN4 (Tang *et al.*, 2021). Nous avons avant la publication de ces résultats initié un crible sans *a priori* contre une banque d'ADNc par double-hybride à partir de MAN1 *via* la technique de MYTH (*Membrane Yeast Two-Hybrid*) (Snider *et al.*, 2010; Voisin, Vanrobays and Tatout, 2018). Cette méthode est dédiée à l'étude des interactions des protéines membranaires. Sur les 50 clones positifs obtenus sur 300 000 criblés, seules 8 protéines candidates ont été retenues (**Tableau 7**). Parmi ces 8 protéines, deux protéines correspondent à des faux positifs déjà isolés dans de très nombreux cribles du laboratoire et cinq sont des nouveaux candidats avec pour certains des domaines transmembranaires mais non dans la membrane nucléaire (**Tableau 7**). Le seul candidat pertinent obtenu, car étant présent dans le noyau et même dans la membrane nucléaire, est la protéine MAN1 elle-même (**Tableau 7**). Nos résultats diffèrent de ceux obtenus dans le proxitome de MAN1, cela peut s'expliquer par le fait que la technique de MYTH permet de montrer des interactions directes avec cependant de nombreux faux positifs contrairement à la protéomique de marquage de proximité utilisée pour faire le proxitome qui permet de montrer les protéines proches et donc des interactions directes et indirectes.

Pour aller plus loin nous avons utilisé les différentes délétions construites pour la localisation de MAN1 (cf paragraphe III.1) pour définir les domaines d'interactions de MAN1 avec elle-même ainsi qu'avec les protéines CRWN. En effet les homologues de MAN1 sont pour la plupart capables de se lier au nucléosquelette, comme démontré chez les animaux, il

Tableau 7 : Bilan du crible MYTH avec la banque d'ADNc CD4-10 sur les appâts MAN1

Clone	TAIR ID	Uniprot	Nom	Localisation prédite	Domaines connus	Croissance		Fonction
						3mM 3-AT	5mM 3-AT	
Candidat(s) déjà connus								
5	AT5G46560	F4KHI7	AtMan1	Membrane/périphérie nucléaire	Domaine TM, Domaine MSC	+	+	Inconnue
Nouveaux candidats								
1.1	AT3G16420	C0Z3F0	PYK10-BINDING PROTEIN 1	Cytoplasme, cytosol, noyau	Jacalin-type lectin	+	+	Chaperonne moléculaire
1.3	AT1G01120	Q0WVC3	KCS1	Membrane/Réticulum	Domaine TM, FAE, ACP_syn III_C	+	+	Métabolisme des lipides
39	AT3G47470	P27521	CAB4, LHCA4	Chloroplastique		+	+	Récepteur de la lumière/ Photosynthèse
44	AT4G30580	Q8GXU8	AtS2	Membrane chloroplastique	Domaine TM	+	+	Métabolisme des lipides
45	AT1G05850	Q9MA41	AtCTL1	Golgi, extra cellulaire	Chitinase-like domain	+	+	Pas d'activité chitinase, importante pour la croissance
Faux positifs déjà connus								
8	AT3G16240	Q41951	AQP1	Membrane vacuolaire	Domaine TM	+	+	Aquaporine
36	AT1G07320	O50061	RIBOSOMAL PROTEIN L4, RPL4	Chloroplastique		+	+	Constituant des ribosomes chloroplastiques

Tableau 8 : Récapitulatif des interactions de MYTH entre MAN1 et CRWN-MAN1

		Prey (pPR3N)	
		MAN1	MAN1 TM-MSC
Bait (pBT3N)	MAN1	++	+
	CRWN1	++	-
	CRWN2	-	
	CRWN3	-	
	CRWN4	-	
	Positive control	++	++
	Empty	-	-

MAN1 TM-MSC : région entre les deux domaines transmembranaires

++ : interaction forte, + : interaction, +/- : interaction faible, - : pas d'interaction, ND : interaction non-déterminée, case grisée : interaction non-testée

nous paraissait donc intéressant de démontrer si ceci était également le cas pour MAN1 et les constituants du nucléosquelette de la plante. Cette question était d'autant plus pertinente que les protéines CRWN font partie de son proximo de MAN1 (Tang *et al.*, 2021).

Nous avons montré que MAN1 est capable d'interagir avec la protéine CRWN1 mais pas CRWN2, CRWN3 et CRWN4 et que cette interaction se fait *via* son domaine MAN1-Src1p (MSC) localisé en position C-terminale (**Tableau 8**). MAN1 interagit avec le nucléosquelette comme c'est le cas chez les animaux, nous n'avons pour le moment pas testé sa capacité à interagir avec KAKU4.

La protéine MAN1 se localise donc à la périphérie nucléaire et est capable de s'homomériser ainsi que d'interagir avec la protéine CRWN1, un constituant majeur du nucléosquelette. MAN1 joue un rôle mineur dans le développement de la plante bien que dans la morphologie nucléaire semble légèrement affectée. Une analyse d'interaction génétique entre les mutants CRISPR de MAN1 et d'autres mutants des constituants du nucléosquelette est en cours pour mieux définir la fonction de cette protéine.

Chapitre III: Discussion et Perspectives

I- Synthèse des résultats acquis au cours de la thèse

Mes travaux de thèse avaient comme contexte général l'acquisition de nouvelles connaissances sur l'organisation de la périphérie nucléaire et son rôle dans la régulation de l'expression des gènes. Plus spécifiquement, je me suis concentrée sur l'étude du nucléosquelette, composé des 4 protéines CRWN et de KAKU4, afin de déterminer les régions nécessaires à leur localisation et leurs interactions mais aussi leur fonction dans la régulation transcriptionnelle et le développement de la plante. Ce travail a permis de découvrir une connexion physique entre le nucléosquelette et le nuclear basket du pore nucléaire. Parmi ces résultats, j'ai choisi d'approfondir deux aspects de mon travail qui m'ont semblés les plus marquants à savoir la caractérisation et la fonction de nouveaux motifs conservés communs au nucléosquelette et au NPC et le lien entre la périphérie nucléaire avec la voie de l'acide salicylique.

II- Rôles des motifs protéiques conservés communs au nucléosquelette et au NPC

II.1) Motifs vs domaines protéiques

Parmi les travaux de ma thèse, j'ai intensivement cherché à identifier de nouveaux domaines et régions fonctionnelles dans nos protéines d'intérêt. Un domaine protéique est défini comme une longue sous-séquence contiguë identifiable d'une protéine, d'au moins 100 AA, qui peut se plier, fonctionner et exister indépendamment du reste de la chaîne ou de la structure protéique. De très nombreux domaines protéiques sont décrits dans les bases de données et assurent des fonctions bien définies comme la liaison à l'ADN, à l'ARN ou à d'autres protéines, cependant, il existe certainement encore des domaines non connus ou des fonctions inconnues. Les domaines protéiques globulaires sont classiquement considérés comme l'archétype des régions protéiques nécessaires à ces interactions, un dogme qui découle de l'affirmation selon laquelle une protéine une fois synthétisée se replie selon une structure tridimensionnelle bien définie pour remplir sa fonction biologique. Si les domaines globulaires fournissent la majorité des interfaces d'interactions protéiques caractérisées à ce jour (Van Roey *et al.*, 2014), ces deux dernières décennies ont révélé que les régions non structurées ou régions intrinsèquement désordonnées (*Intrinsically Disordered Region*, IDR présentes dans les *Intrinsically Disordered Protein*, IDP) constitueraient près de 40% de la séquence protéique (Davey, 2019).

Les IDR jouent un rôle central dans de nombreux processus cellulaires et peuvent conduire à la structuration de la protéine lors d'une interaction. Les mécanismes par lesquels les IDR dirigent et régulent la fonction des protéines ne sont encore pas encore bien documentés et très peu d'exemples existent chez les plantes.

Les modules fonctionnels des IDR sont composés de courtes successions d'acides aminés inférieures à 100AA connus sous le nom de motifs linéaires courts (*Short Linear Motif*, SLiM), également appelés motifs linéaires (*Linear Motifs*, LM) (Davey *et al.*, 2011). Les SLiM sont actuellement classés en *i*) motifs ligands qui servent de médiateurs aux interactions protéine-protéine (*motif/SLiM ligand*) et *ii*) en motifs de modifications post-traductionnelles (*Post-Translational Modification -PTM-motifs/SLiM*), qui sont directement reconnus et ciblés pour une modification post-traductionnelle par des enzymes régulatrices (**Figure 3.1**) (Van Roey *et al.*, 2014). Le site *Eukaryotic Linear Motif* (ELM) est le référentiel le plus complet de SLiMs validés expérimentalement. Or, seuls 79 de ces motifs sont décrits fonctionnellement chez *Arabidopsis* alors que 2200 sont déjà caractérisés chez l'Homme (<http://elm.eu.org/infos/stats.html>) (Kumar *et al.*, 2022). La fonction des SLiM est très large et, en fonction des protéines partenaires qui les reconnaissent, ils peuvent faciliter le ciblage d'une protéine à un emplacement subcellulaire spécifique, déterminer l'état de modification d'une protéine, contrôler sa stabilité ou réguler son activité.

C'est en réalisant différentes études *via* la suite MEME pour trouver des régions conservées dans la protéine KAKU4 qui pourraient expliquer son interaction avec les protéines CRWN et sa localisation à la périphérie nucléaire, que j'ai identifié les premiers motifs conservés de 25-30 AA. Du fait de leur petite taille, ces motifs ressemblent aux SLiM. De plus, nos protéines d'intérêts, NUP136, NUP82, KAKU4 et les protéines CRWN1 à 4, sont toutes prédites comme étant des IDP avec respectivement un pourcentage de régions désordonnées de 86%, 60%, 57% et environ 80% pour les quatre protéines CRWN alors que la valeur médiane pour les protéines de la périphérie nucléaire est seulement de 35%. Les régions ordonnées de ces protéines corrélerent plutôt bien avec nos motifs. De plus, la structure de ces protéines ou de ces motifs, respectivement prédites par Alphafold ou par PEP-FOLD3 montrent que ces motifs forment des hélices alpha, confirmant leur implication dans les fonctions de ces protéines.

Lors de l'étude de la localisation de nos protéines en expression transitoire, nous avons remarqué que la protéine NUP82 forme de gros agrégats. Ceci est observé dans une moindre mesure avec NUP136, une autre protéine du nuclear basket. De surcroît, il est maintenant connu que les protéines désordonnées peuvent former des agrégats qui pourraient former des

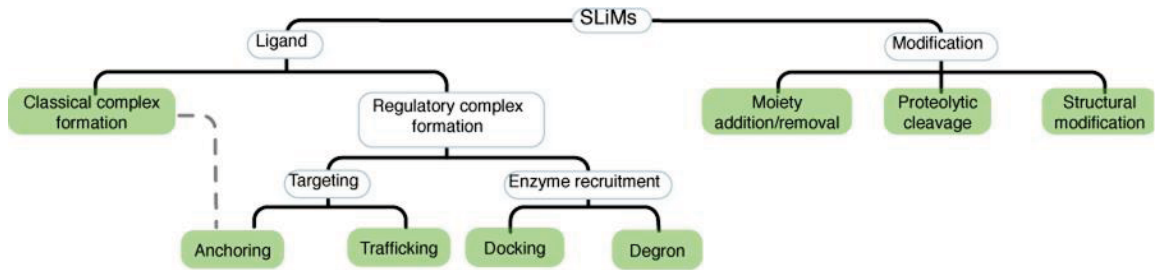


Figure 3.1 Classification fonctionnelle des SLiM (Van Roey *et al.*, 2014).

Les SLiM sont subdivisés en deux groupes : A gauche, **les motifs ligands** sont des motifs régulateurs qui contrôlent la localisation, l'état de modification ou la stabilité des protéines ainsi que des médiateurs permettant l'assemblage de complexes. Ces différentes classes de motifs ligands ne sont pas nécessairement mutuellement exclusives. A droite, les **motifs de modifications post-traductionnelles** qui peuvent fonctionner comme des sites pour l'ajout/le retrait de PTM, le clivage protéolytique ou la modification structurale.

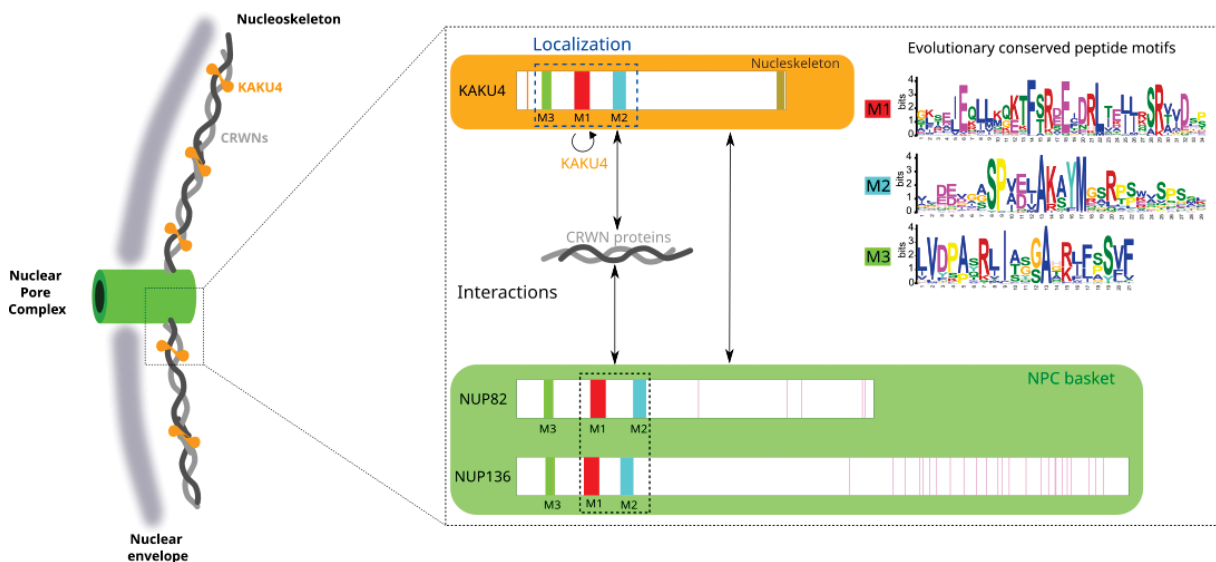


Figure 3.2 Schéma récapitulatif de la fonction des motifs des protéines KAKU4, NUP136 et NUP82

La protéine KAKU4 contient trois motifs conservés nommés M1, M2 et M3, également présents dans les protéines NUP136 et NUP82. M1 est nécessaire à l'homomérisation de KAKU4 alors que le motif M2 est responsable des interactions entre KAKU4 avec les protéines CRWN, ces deux motifs sont également responsables de l'interaction des protéines NUP136 et NUP82 avec les protéines CRWN. Enfin collectivement M3-M1-M2 sont indispensables à la localisation de KAKU4 à la périphérie nucléaire.

corpuscules voir des domaines selon le principe de la séparation de phase liquide-liquide (*liquid-liquid phase separation*, LLPS) (Shin *et al.*, 2017; Uemura *et al.*, 2018; Shen *et al.*, 2021). A nouveau, nos observations de la formation des agrégats en expression transitoire et le fait que nos protéines possèdent sans doute les propriétés des IDP suggèrent que ces protéines pourraient participer au phénomène de LLPS.

Nous proposons que certaines protéines de la périphérie nucléaire, notamment celles du nuclear basket du pore nucléaire et du nucléosquelette, pourraient former un micro-compartiment *via* un réseau d'interactions médié par des SLIM et isolant certaines régions de la chromatine du reste du génome *via* le principe du LLPS. Pour tester cette hypothèse, nous pourrions, en *in vitro*, tester la formation de gouttelettes par nos protéines purifiées dans des milieux ayant différentes concentrations en sel (S. Huang *et al.*, 2021). D'après les différentes expériences que j'ai réalisées, nos motifs s'apparenteraient à des motifs ligands (**Figure 3.1**).

II.2) Rôle des motifs

Sur le plan fonctionnel, plusieurs types de motifs ligands et de motifs de modification post-traductionnelle peuvent être définis (**Figure 3.1**). Les motifs ligands déterminent l'état de modification, la stabilité ou la localisation subcellulaire d'une protéine. Il existe trois types de motifs ligands, *i*) ceux qui permettent le recrutement d'enzyme modificatrices, par exemple pour contrôler la stabilité des protéines ; *ii*) ceux qui permettent la localisation correcte des protéines en intervenant soit dans leur transport (*trafficking motifs*), soit dans la liaison à des complexes spécifiques (motifs d'ancrage) et *iii*) ceux qui n'ont pas de fonction de régulation spécifique mais qui agissent comme simple module de liaison, servant souvent dans l'assemblage de grands complexes macromoléculaires (Van Roey *et al.*, 2014) (**Figure 3.1**).

II.2.a) Rôle des motifs dans les interactions protéine-protéine

L'analyse des motifs la plus aboutie est celle réalisée avec la protéine KAKU4 pour laquelle nous avons identifié et étudié la fonction de trois motifs conservés, les motifs M1, M2 et M3, également présents dans les protéines NUP136 et NUP82. Chronologiquement, M3 a été trouvé après l'identification de M1 et M2 ce qui a nécessité de refaire une seconde série d'expériences incluant ce nouveau motif. Nous avons pu attribuer une fonction spécifique à M1 (homomérisation de KAKU4), M2 (interaction KAKU4/CRWN) et M1-M2-M3 (localisation périphérique) (**Figure 3.2**). Les motifs M1 et M2 semblent donc faire partie des motifs ligands impliqués dans la formation de complexes *via* très certainement une interaction entre eux dans le cas des interactions entre KAKU4 et les NUP et avec d'autres motifs dans les protéines

CRWN. Ce travail a été initié avec deux motifs conservés chez les protéines CRWN : M10 et M20 qui semblent nécessaires aux interactions avec KAKU4 ainsi que M14 qui serait nécessaire à la localisation périphérique de CRWN1, ce dernier motif coïncide avec la région identifiée par Guo et al. (2017).

Pour confirmer le rôle de M10, M20 et M14, plusieurs approches pourraient être mises en place comme de la protéomique de marquage de proximité (BioID) développée récemment par le groupe de Yangnan Gu (Tang et al 2020, 2022), du GFP-trap, du double-hybride ou par des méthodes de prédiction *in silico* par deep learning comme PIPR testée au laboratoire (Chen et al 2019). Cependant certaines des méthodes comme le protéomique de marquage de proximité devraient être adaptées à de courtes séquences protéiques comme les SLIM avant la confirmation des interactions directes par double-hybride ou par BiFC.

II.2.b) Rôle des motifs dans la localisation à la périphérie nucléaire

Le motif M3 des protéines KAKU4, NUP136 et NUP82 présente des similitudes avec le motif M14 des protéines CRWN, nos résultats suggèrent l'implication de ce/ces motif/s pour la localisation de ces protéines à la périphérie nucléaire. J'ai donc nommé ce motif, motif PLS (*Peripheral Localisation Signal*). Le motif PLS serait un motif d'ancrage ou d'attribution, permettant la localisation à la périphérie nucléaire, qui pourrait être présent dans d'autres protéines chez la plante voire chez d'autres organismes que ceux du règne végétal. Ce travail exploratoire pourrait être initié avec les outils de la suite MEME : FIMO (*Find Individual Motif Occurrences*) pour rechercher dans différents organismes ou AME (*Analysis of Motif Enrichment*) pour rechercher dans un ensemble de protéines données comme celles de la périphérie nucléaire chez *A. thaliana*.

Une hypothèse attractive serait que cet ancrage à la périphérie nucléaire serait assuré par des protéines ancrées dans l'INM comme les protéines SUN, MAN1 ou PNET2 *via* des interactions entre motifs PLS ou avec d'autres motifs. La recherche de ligand du motif PLS pourrait-être réalisée par la technique du double-hybride ou une technique qui dériverait du PEPscan. Nous avons exploré le PEPscan entre M2 et la protéine CRWN1 (cf Résultats [I.2.b](#)) en collaboration avec l'équipe de Pierre Tuffery à Paris mais à ce stade la technique ne permet d'étudier qu'une seule protéine (ici CRWN1) découpée en peptide de 20AA avec un petite (M2) marqué par un fluorochrome. Il s'agirait d'évaluer des techniques de screening à plus haut débit c'est-à-dire avec un plus grand nombre de protéines immobilisées sur la membrane comme la technique *Protein Interaction Screen on Peptide Matrix* (PRISMA) (<https://www.jpt.com/>).

II.2.c) Autres fonctions des motifs

Nous avons cherché à savoir si ces motifs étaient impliqués dans d'autres fonctions comme la formation des membranes (cf Résultats [I.1.b](#)), **Figure 2** article soumis), la distribution des NPC dans la membrane nucléaire (cf Résultats [I.1.c.iii](#)), dans les défauts développementaux observés notamment dans les tissus reproducteurs (en cours) ou encore dans les altérations de la morphologie nucléaire (cf Résultats [I.1.b](#)), **Figure 2** article soumis). Cette recherche réalisée ou initiée au cours de ma dernière année de thèse a été souvent infructueuse et parfois techniquement complexe à mettre en place. Toutefois, j'ai pu montrer que les motifs M1 et M2 de KAKU4 sont nécessaires à l'élongation du noyau en utilisant une approche de complémentation de la lignée mutante *kaku4* avec des transgènes exprimant KAKU4 complet ou délété de ses motifs. La difficulté technique a été ici de trouver un tissu adapté à nos analyses. Les filaments des étamines utilisés par l'équipe d'Eric Richards (Choi et al 2019) s'est avéré un tissu approprié car les noyaux sont en majorités allongés et permettent de distinguer le phénotype « petits noyaux ronds » du mutant *kaku4*. Nos résultats suggèrent que les interactions de KAKU4 ou des composants du nuclear basket avec les protéines CRWN sont importantes pour le maintien de la morphologie nucléaire. Des expériences complémentaires incluant le motif M3 sont en cours de réalisation.

De plus, de nombreux mutants de la périphérie nucléaire comme les protéines SUN, PNET2 ainsi que certains composants du NPC sont connus pour modifier la morphologie nucléaire. De nombreux mutants restent à décrire pour la morphologie nucléaire. Nous avons initié la caractérisation de plusieurs mutants du NPC (*nup96*, *hos1*, *nup98*, *nup88*, *nup160* et *nua*) mais il serait intéressant de faire une étude plus systématique. Ceci sera peut-être possible dans le futur grâce à de nouveaux outils d'analyse d'images reposant sur la technologie du *deep learning* qui sont actuellement développés par Guillaume Mougeot, un autre doctorant de mon équipe. Une grande inconnue subsiste : quel rôle joue réellement la morphologie nucléaire sur les changements du transcriptome.

II.3) Connexion ente le NPC et le nucléosquelette

Au cours de ces travaux, l'un de mes résultats les plus excitants est la découverte d'une connexion physique entre le nucléosquelette et le NPC par son nuclear basket, comme cela a déjà été montré chez le Xénope avec NUP153 (homologue fonctionnelle de NUP136) et les lamines (Smythe, Jenkins and Hutchison, 2000).

Un autre résultat marquant est sans doute la suggestion d'une origine évolutive commune entre NUP136 et NUP82 (NPC) et KAKU4 (nucléosquelette). Cette parenté se traduit par la conservation de nos motifs dans la région N-terminale de ces protéines. Lors de l'étude phylogénétique de ces trois protéines, nous avons constaté que NUP136 est présente dans les différentes espèces analysées à partir des mousses alors que NUP82 et KAKU4 apparaissent seulement dans les angiospermes basales. *KAKU4* et *NUP82* seraient donc issus d'une duplication de *NUP136*. Nous pensons que *KAKU4* dériverait d'une nucléoporine ancestrale et aurait ensuite acquis sa nouvelle fonction au sein du nucléosquelette peut-être *via* l'acquisition d'une région C-terminale distincte qui ne contient plus de FG repeats comme dans NUP136. La fonction de la partie C-terminale de *KAKU4* reste à découvrir.

Plus récemment, une connexion entre le nuclear basket et le nucléosquelette a également été montrée avec la protéine GBPL3 (GUANYLATE-BINDING PROTEIN-LIKE 3) (Yangnan Gu, communication personnelle). GBPL3 serait un nouveau composant du nuclear basket et, en utilisant la technique protéomique par marquage de proximité, l'équipe de Y. Gu a pu montrer que GBPL3 était proche des protéines *KAKU4* et *CRWN1* à 4. Par analogie aux connaissances accumulées sur *KAKU4*, *CRWN1-4*, *NUP82* et *NUP136* ; *GBPL3* et *GBPL1* un autre membre de la famille des protéines *guanylate-binding protein-like* sont capables de recruter des complexes de remodelage de la chromatine, des facteurs de transcription et la machinerie transcriptionnelle, ceci est probablement favorisé par de la transition de phase. Cela semble essentiel pour le contrôle de la réponse transcriptionnelle lors d'un stress biotique. Ceci suggère une régulation complexe à la périphérie nucléaire impliquant un réseau de protéines interagissant entre elles mais également avec la chromatine et participant à la réponse aux stress.

III- Périphérie nucléaire et nucléosquelette : un ensemble complexe de phénotypes et d'interactions génétiques

Notre laboratoire s'intéressant à la morphologie du noyau et ses conséquences possibles sur la transcription, l'une de mes priorités était de définir l'impact de notre collection de multi-mutants sur la morphologie nucléaire. C'est une analyse fastidieuse nécessitant de nombreuses images de chaque mutant puis une analyse informatique. Même si de nouveaux développements bio-informatique devraient nous permettre d'aller plus vite dans le futur (Mougeot et al 2022), j'ai pour ma part utilisé notre outil NucleusJ2.0 (Dubos et al 2020) dont certaines des étapes nécessitent l'intervention de l'expérimentateur pour définir des seuils lors de la segmentation des noyaux. En utilisant cet outil, j'ai pu montrer que nos simples, doubles et triples mutants pour *KAKU4*, *NUP136* et *CRWN1* ont tous pour la plupart des cellules de pavement avec des

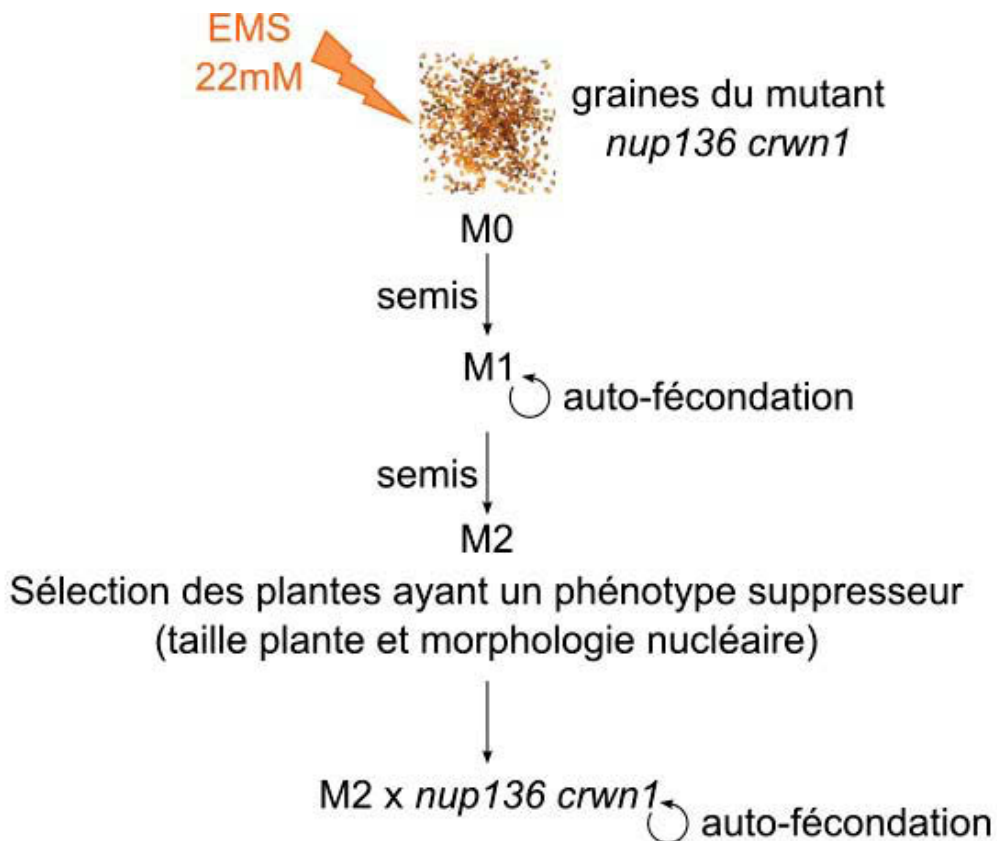
noyaux plus sphériques que dans une lignée sauvage (voir par exemple **Figure 5** de notre article soumis à *The Plant Cell* et **Figure 2.14**). Cette analyse, révèle que dans chacun des doubles mutants, c'est le mutant du nucléosquelette, *CRWN1* ou *KAKU4*, qui a l'effet le plus marqué et que la mutation de *NUP136* n'a qu'un faible impact sur la morphologie nucléaire. Toutefois, il reste compliqué d'évaluer l'effet relatif du NPC vis-à-vis de celui du nucléosquelette. Il semble que beaucoup de mutants affectent la morphologie nucléaire, et il serait judicieux d'effectuer une analyse plus large afin de déterminer la contribution de ses deux composants structuraux du noyau.

Lors de mon étude phénotypique et transcriptomique des multi-mutants de la périphérie nucléaire en utilisant les mutants pour les gènes *NUP136*, *KAKU4* et *CRWN1*, nous avons observé que le triple mutant *nup136 kaku4 crwn1* avait un très fort impact à la fois sur le développement de la plante (petite plante) et sur la dérégulation des gènes notamment ceux de la voie de réponse au stress biotique. Cependant, le double mutant *nup136 crwn1* bien qu'ayant un développement très affecté, ne dérégule qu'un faible nombre de gènes comparativement au mutant *nup136 kaku4*, ce dernier n'ayant qu'un assez faible impact sur le développement des plantes alors qu'il dérégule de nombreux gènes (triple mutant : 1699, *nup136 crwn1* : 251 et *nup136 kaku4* : 1027 ; **Figure 2.24**). Les anomalies de développement de nos différents mutants ne seraient donc pas strictement corrélées à la dérégulation de l'expression des gènes., ce qui est surprenant. Ces résultats pourraient souligner des interactions génétiques et fonctionnelles complexes entre les mutants du nucléosquelette et ceux du NPC mais il faut toutefois émettre certaines réserves.

En premier lieu, nous avons observé que nos mutants semblent très sensibles à leur condition de culture. Ainsi, la reproductibilité entre manips de RNA-Seq comme par exemple pour *kaku4* (manip RNA-Seq1, 2 et 3) ou entre allèles de *kaku4* (manip RNA-Seq2), semble limitée et nous a posé des difficultés d'interprétations de nos expériences. Pourtant, nous avons pris beaucoup de précautions avec l'ensemble des graines produites en même temps, l'utilisation d'un témoin Col-0 issu du même lot de graine dans chaque expérience, le même workflow d'analyse RNA-Seq pour toutes nos analyses, *etc.* Afin d'améliorer encore cette reproductibilité, plusieurs pistes pourraient être explorées par exemple l'utilisation de spike-in (ajout d'une quantité équivalente d'ADN témoin dans tous nos échantillons afin d'effectuer une normalisation) (Jean-Baptiste *et al.*, 2019), l'élimination de gènes *high inter-individual variability* de nos listes de gènes dérégulés (Cortijo *et al.*, 2019), l'utilisation d'un seul type cellulaire ou des cellules de ploïdie équivalente (Hu *et al.*, 2019), *etc.* ; il faut bien dire que les

pistes ne manquent pas. Etant donné la précocité de certains mutants impliquant *NUP136*, *KAKU4* et *CRWN1* et le fort impact que cela peut avoir sur la taille de la rosette, j'ai phénotypé toutes mes plantes lors de l'ouverture de la première fleur pour qu'elles soient toutes au même stade de développement et ainsi être comparées. Certaines de mes analyses sont réalisées à J21, si nous voulons utiliser ces résultats, il faudrait les refaire afin de prendre des plantes qui sont toutes au même stade de développement. De plus, nous avons remarqué que notre témoin Col-0 pouvait parfois être plus impacté que certains mutants dans des stades précoces (vers J+14) puis rattraper son retard de croissance à des stades plus tardifs. Il faut d'ailleurs remarquer que l'équipe d'E. Richards a réalisée toutes ses analyses du transcriptome à un stade plus tardif (J+28) que le nôtre (J+13) car il semblerait que l'impact des mutants *crwn* s'aggrave avec l'âge des plantes. J'ai d'ailleurs pu montrer cet effet en suivant le gène *PRI* en RT-PCR (**Figure 2.28**). Nous pourrions aussi effectuer une analyse sur les tissus reproducteurs, là où l'effet de nos mutants *kaku4* et *nup136* semblent les plus forts.

Au cours de cette réflexion sur la stratégie d'analyse des données de RNA-Seq, nous avons entrevu la possibilité de dissocier/décomplexifier les phénotypes de nos mutants. En effet, de précédentes études ont montré que le mutant *nup136* ainsi que le double mutant *crwn1 crwn2* entraînent l'activation constitutive de la voie du SA (Guo *et al.*, 2017; Tamura *et al.*, 2017; Choi, Strickler and Richards, 2019). Cette activation constitutive induit une dérégulation importante du transcriptome, une meilleure résistance aux pathogènes tout en provoquant le phénotype de petite taille des plantes. Ce même phénotype avait déjà été décrit pour le mutant *scn1* (Zhang and Li, 2005). Ce phénotype « petite plante » se retrouvant chez nos mutants *nup136 crwn1* et *nup136 kaku4 crwn1* et la voie de réponse au SA faisant également partie des voies métaboliques dérégulées dans nos mutants, nous avons regardé si nous pourrions nous affranchir du phénotype petite plante en utilisant des mutations affectant la voie de réponse au SA. Pour cela, nous nous sommes appuyés sur l'observation que la mutation perte de fonction de *SID2* (affecte la biosynthèse du SA) ou *NPR1* (facteur de transcription affectant les gènes cibles de la voie de réponse au SA comme *PRI*) suppriment le phénotype petite plante du mutant *crwn1 crwn2* au moins pour la taille de la rosette mais pas pour la taille des hampes pour *sid2* (Guo *et al.*, 2017; Choi, Strickler and Richards, 2019). Nous avons donc introduit les mutations de *SID2* ou *NPR1* dans notre mutant *nup136 crwn1*. Une nouvelle surprise nous attendait puisque les triples mutants *nup136 crwn1 sid2* et *nup136 crwn1 npr1* ne se comportent pas de la même façon. Si comme attendu, la mutation de *SID2* entraîne une restauration partielle du phénotype de taille de la plante, celle de *NPR1* au contraire accentue ce phénotype. Ce qui



- séquençage : pool de plantes ayant le phénotype suppresseur (M2)
pool de plantes ayant le phénotype *nup136 crwn1*
- identification des mutations responsables du phénotype suppresseur

Figure 3.3 : Etapes du crible EMS en cours sur le mutant *nup136 crwn1*

Réalisation d'une mutagenèse avec du méthanesulfonate d'éthyle (EMS) sur des graines du mutant *nup136 crwn1* (génération M0), ces graines ont été semées (génération M1) ; si présence d'une mutation dominante, son phénotype de suppresseur sera visible dans cette génération M1 ; pour les autres mutations, elles seront visibles seulement dans la génération M2 issue de l'autofécondation des M1. Lors de ce crible nous avons obtenu 1 plante dominante en M1 et une vingtaine de plantes en M2 suppresseur de la taille de la plante. Nous nous sommes arrêtés à cette étape. Les plantes ayant un phénotype suppresseur seront croisées avec le double mutant *nup136 crwn1* puis autofécondées avant d'être séquencées en deux pool (l'un correspondant au phénotype suppresseur et l'autre à celui du mutant *nup136 crwn1*) pour obtenir les mutations spécifiques des suppresseurs.

suggère que le phénotype des plantes est dû en partie à l'activation de la voie du SA mais est indépendante de NPR1. Il reste à étudier l'impact de ces mutants sur la morphologie nucléaire mais étant donné qu'aucune restauration du phénotype n'a été observé dans le mutant *crwn1 crwn2 sid2* nous ne nous attendons pas non plus à une restauration dans nos mutants *nup136 crwn1 sid2* et *nup136 crwn1 npr1* suggérant que la morphologie nucléaire ne serait pas directement impactée par le SA. Les mécanismes qui pourraient être impliqués dans les phénotypes du mutant *nup136 crwn1*, sont encore obscurs mais une de nos hypothèses est que NUP136 et/ou CRWN1 seraient impliqués dans les échanges nucléocytoplasmiques de la réponse au SA avec un rôle soit dans l'import de certains facteurs de transcription, soit dans l'export de certains ARNm ou bien seraient nécessaires à la formation d'un complexe impliqué dans la régulation de certains gènes.

Pour avancer sur les mécanismes expliquant nos différents phénotypes, nous avons initié un crible EMS dans le mutant *nup136 crwn1* pour rechercher des modificateurs du type suppresseur en recherchant des plantes où il y aurait une restauration du phénotype petite plante. Ce type de crible par exemple réalisé avec le mutant *snc1* a permis de relier les mécanismes de défense des plantes médié par le SA avec le transport nucléo-cytoplasmique (Zhang and Li, 2005; Cheng *et al.*, 2009; Wiermer *et al.*, 2012). Notons qu'à l'inverse, la recherche d'enhanceur augmentant l'intensité des phénotypes mutants est plus difficile à envisager du fait du phénotype développemental très sévère du mutant *nup136 crwn1*. Ainsi nous avons recherché des phénotypes développementaux atténué par rapport au double mutant initial, en particulier sur les critères de taille et de fertilité. Nous avons pour le moment réalisé un criblage pilote sur plusieurs milliers de graines M2 et isolé une vingtaine de suppresseurs potentiels avec des morphologies se rapprochant de la lignée sauvage (**Figure 3.3**). Un second critère de sélection est la restauration de la morphologie nucléaire, mais sur l'ensemble des révertants potentiels isolés à ce jour, aucun rétablissement d'une morphologie nucléaire sauvage n'a été observée. Ceci confirme pour l'instant que les phénotypes développementaux et ceux d'altération de la morphologie nucléaire ne semble pas reliés. Ce résultat n'est pas une surprise, nos résultats de croisements, montrent que des mutations dans des gènes impliqués dans la voie du SA comme *SID2* supprime le phénotype petite plante de notre mutant *nup136 crwn1* mais pas la morphologie nucléaire (cf Résultats [II.3.b](#)); **Figure2.28** pour la taille des plantes et données non montrées pour les noyaux). Ce crible EMS est prometteur, mais nécessite d'être amélioré en plusieurs points, un criblage initial basé sur la taille des plantes est assez simple

mais l'analyse de la morphologie pourrait être améliorée grâce à des lignées mutants portant des constructions fluorescentes additionnelles *via* la technique de clonage GoldenBraid. Nous développons notamment des T-DNA permettant le marquage simultané de la périphérie nucléaire (ou du NPC), de la chromatine et du nucléole qui nous permettrait un criblage plus rapide et la possibilité d'utiliser l'architecture du noyau comme critère primaire. Notre crible pourrait donc à ce moment être soit effectué pour rechercher une restauration de la morphologie nucléaire soit pour une restauration de la stature de la plante.

Pour conclure, mon travail a permis de mettre en évidence la connexion entre le pore nucléaire *via* le nuclear basket et le nucléosquelette grâce à des motifs conservés présents dans ces protéines et nécessaires pour leur interaction ainsi que pour leur localisation. Cependant, il reste encore de nombreuses expériences à faire pour comprendre le rôle de cette connexion dans l'organisation de la chromatine. Il serait intéressant d'étudier l'organisation de la chromatine dans les notre collection de multi-mutants y compris ceux combinant des mutants de la voie du SA. Cette étude pourrait se faire par Hi-C et/ou DNA-FISH. L'analyse par DNA-FISH a d'ailleurs été faite pour le triple mutant *kaku4 crwn1 crwn4* (Dubos et al 2020) et celles des mutants *nup136 crwn1* et *nup136 kaku4* a été réalisée, les résultats sont en cours d'analyses.

Références

Adam, S.A. and Goldman, R.D. (2012) ‘Insights into the differences between the A- and B-type nuclear lamins’, *Advances in Biological Regulation*, 52(1), pp. 108–113. Available at: <https://doi.org/10.1016/j.advenzreg.2011.11.001>.

Arpòn, J. *et al.* (2021) ‘Spatial modeling of biological patterns shows multiscale organization of *Arabidopsis thaliana* heterochromatin’, *Scientific Reports*, 11(1), p. 323. Available at: <https://doi.org/10.1038/s41598-020-79158-5>.

Banigan, E.J. and Mirny, L.A. (2020) ‘Loop extrusion: theory meets single-molecule experiments’, *Current Opinion in Cell Biology*, 64, pp. 124–138. Available at: <https://doi.org/10.1016/j.ceb.2020.04.011>.

Barrett, T. *et al.* (2013) ‘NCBI GEO: archive for functional genomics data sets--update’, *Nucleic Acids Research*, 41(Database issue), pp. D991-995. Available at: <https://doi.org/10.1093/nar/gks1193>.

Barton, L.J., Soshnev, A.A. and Geyer, P.K. (2015) ‘Networking in the nucleus: a spotlight on LEM-domain proteins’, *Current Opinion in Cell Biology*, 34, pp. 1–8. Available at: <https://doi.org/10.1016/j.ceb.2015.03.005>.

Baum, D.A. (2015) ‘A comparison of autogenous theories for the origin of eukaryotic cells’, *American Journal of Botany*, 102(12), pp. 1954–1965. Available at: <https://doi.org/10.3732/ajb.1500196>.

Beagan, J.A. and Phillips-Cremins, J.E. (2020) ‘On the existence and functionality of topologically associating domains’, *Nature Genetics*, 52(1), pp. 8–16. Available at: <https://doi.org/10.1038/s41588-019-0561-1>.

Bi, X. *et al.* (2017) ‘Nonrandom domain organization of the *Arabidopsis* genome at the nuclear periphery’, *Genome Research*, 27(7), pp. 1162–1173. Available at: <https://doi.org/10.1101/gr.215186.116>.

Blobel, G. (1985) ‘Gene gating: a hypothesis.’, *Proceedings of the National Academy of Sciences*, 82(24), pp. 8527–8529. Available at: <https://doi.org/10.1073/pnas.82.24.8527>.

Blunt, E.L. *et al.* (2020) ‘Coordination of NMCP1- and NMCP2-class proteins within the plant nucleoskeleton’, *Molecular Biology of the Cell*, 31(26), pp. 2948–2958. Available at: <https://doi.org/10.1091/mbc.E19-08-0464>.

Borah, S. *et al.* (2021) ‘Heh2/Man1 may be an evolutionarily conserved sensor of NPC assembly state’, *Molecular Biology of the Cell*. Edited by K. Weis, 32(15), pp. 1359–1373. Available at: <https://doi.org/10.1091/mbc.E20-09-0584>.

Brachner, A. and Foisner, R. (2011) ‘Evolution of LEM proteins as chromatin tethers at the nuclear periphery’, *Biochemical Society Transactions*, 39(6), pp. 1735–1741. Available at: <https://doi.org/10.1042/BST20110724>.

- Brickner, D.G. *et al.* (2019) ‘The Role of Transcription Factors and Nuclear Pore Proteins in Controlling the Spatial Organization of the Yeast Genome’, *Developmental Cell*, 49(6), pp. 936–947.e4. Available at: <https://doi.org/10.1016/j.devcel.2019.05.023>.
- Capelson, M. *et al.* (2010) ‘Chromatin-Bound Nuclear Pore Components Regulate Gene Expression in Higher Eukaryotes’, *Cell*, 140(3), pp. 372–383. Available at: <https://doi.org/10.1016/j.cell.2009.12.054>.
- Caputo, S. *et al.* (2006) ‘The Carboxyl-terminal Nucleoplasmic Region of MAN1 Exhibits a DNA Binding Winged Helix Domain*’, *Journal of Biological Chemistry*, 281(26), pp. 18208–18215. Available at: <https://doi.org/10.1074/jbc.M601980200>.
- Cavalier-Smith, T. (2002) ‘The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa.’, *International Journal of Systematic and Evolutionary Microbiology*, 52(2), pp. 297–354. Available at: <https://doi.org/10.1099/00207713-52-2-297>.
- Chen, Z. *et al.* (2009) ‘Biosynthesis of salicylic acid in plants’, *Plant Signaling & Behavior*, 4(6), pp. 493–496.
- Cheng, Y.T. *et al.* (2009) ‘Nuclear Pore Complex Component MOS7/Nup88 Is Required for Innate Immunity and Nuclear Accumulation of Defense Regulators in Arabidopsis’, *The Plant Cell*, 21(8), pp. 2503–2516. Available at: <https://doi.org/10.1105/tpc.108.064519>.
- Cheng, Z.J. *et al.* (2014) ‘Abscisic Acid Regulates Early Seed Development in Arabidopsis by ABI5-Mediated Transcription of SHORT HYPOCOTYL UNDER BLUE1’, *The Plant Cell*, 26(3), pp. 1053–1068. Available at: <https://doi.org/10.1105/tpc.113.121566>.
- Choi, J. and Richards, E.J. (2019) ‘The role of CRWN nuclear proteins in chromatin-based regulation of stress response genes’, *Plant Signaling & Behavior*, 15(1). Available at: <https://doi.org/10.1080/15592324.2019.1694224>.
- Choi, J., Strickler, S.R. and Richards, E.J. (2019) ‘Loss of CRWN Nuclear Proteins Induces Cell Death and Salicylic Acid Defense Signaling’, *Plant Physiology*, 179(4), pp. 1315–1329. Available at: <https://doi.org/10.1104/pp.18.01020>.
- Ciska, M. *et al.* (2019) ‘Evolutionary history and structure of nuclear matrix constituent proteins, the plant analogues of lamins’, *Journal of Experimental Botany*, 70(10), pp. 2651–2664. Available at: <https://doi.org/10.1093/jxb/erz102>.
- Cohen, T.V., Kosti, O. and Stewart, C.L. (2007) ‘The nuclear envelope protein MAN1 regulates TGF β signaling and vasculogenesis in the embryonic yolk sac’, *Development*, 134(7), pp. 1385–1395. Available at: <https://doi.org/10.1242/dev.02816>.
- Cortijo, S. *et al.* (2019) ‘Widespread inter-individual gene expression variability in *Arabidopsis thaliana*’, *Molecular Systems Biology*, 15(1), p. e8591. Available at: <https://doi.org/10.15252/msb.20188591>.
- Crevillén, P. *et al.* (2013) ‘A gene loop containing the floral repressor FLC is disrupted in the early phase of vernalization’, *The EMBO journal*, 32(1), pp. 140–148. Available at: <https://doi.org/10.1038/emboj.2012.324>.

- Davey, N.E. *et al.* (2011) ‘Attributes of short linear motifs’, *Molecular BioSystems*, 8(1), pp. 268–281. Available at: <https://doi.org/10.1039/C1MB05231D>.
- Davey, N.E. (2019) ‘The functional importance of structure in unstructured protein regions’, *Current Opinion in Structural Biology*, 56, pp. 155–163. Available at: <https://doi.org/10.1016/j.sbi.2019.03.009>.
- Dechat, T. *et al.* (2010) ‘Nuclear lamins’, *Cold Spring Harbor Perspectives in Biology*, 2(11), p. a000547. Available at: <https://doi.org/10.1101/cshperspect.a000547>.
- Devos, D.P., Gräf, R. and Field, M.C. (2014) ‘Evolution of the nucleus’, *Current Opinion in Cell Biology*, 28, pp. 8–15. Available at: <https://doi.org/10.1016/j.ceb.2014.01.004>.
- Dittmer, T.A. *et al.* (2007) ‘LITTLE NUCLEI genes affecting nuclear morphology in *Arabidopsis thaliana*’, *The Plant Cell*, 19(9), pp. 2793–2803. Available at: <https://doi.org/10.1105/tpc.107.053231>.
- Dittmer, T.A. and Misteli, T. (2011) ‘The lamin protein family’, *Genome Biology*, 12(5), p. 222. Available at: <https://doi.org/10.1186/gb-2011-12-5-222>.
- Dong, C.-H., Hu, X., *et al.* (2006) ‘A Putative *Arabidopsis* Nucleoporin, AtNUP160, Is Critical for RNA Export and Required for Plant Tolerance to Cold Stress’, *Molecular and Cellular Biology*, 26(24), pp. 9533–9543. Available at: <https://doi.org/10.1128/MCB.01063-06>.
- Dong, C.-H., Agarwal, M., *et al.* (2006) ‘The negative regulator of plant cold responses, HOS1, is a RING E3 ligase that mediates the ubiquitination and degradation of ICE1’, *Proceedings of the National Academy of Sciences of the United States of America*, 103(21), pp. 8281–8286. Available at: <https://doi.org/10.1073/pnas.0602874103>.
- Dong, P. *et al.* (2017) ‘3D Chromatin Architecture of Large Plant Genomes Determined by Local A/B Compartments’, *Molecular Plant*, 10(12), pp. 1497–1509. Available at: <https://doi.org/10.1016/j.molp.2017.11.005>.
- Dong, P. *et al.* (2020) ‘Plant and animal chromatin three-dimensional organization: similar structures but different functions’, *Journal of Experimental Botany*, 71(17), pp. 5119–5128. Available at: <https://doi.org/10.1093/jxb/eraa220>.
- Dubos, T. *et al.* (2020) ‘Automated 3D bio-imaging analysis of nuclear organization by NucleusJ 2.0’, *Nucleus*, 11(1), pp. 315–329. Available at: <https://doi.org/10.1080/19491034.2020.1845012>.
- Dubos, T. *et al.* (2022) ‘NODEJ: an ImageJ plugin for 3D segmentation of nuclear objects’, *BMC Bioinformatics*, 23(1), p. 216. Available at: <https://doi.org/10.1186/s12859-022-04743-6>.
- Dundr, M. and Misteli, T. (2010) ‘Biogenesis of Nuclear Bodies’, *Cold Spring Harbor Perspectives in Biology*, 2(12), p. a000711. Available at: <https://doi.org/10.1101/cshperspect.a000711>.
- Evans, D.E., Mermet, S. and Tatout, C. (2020) ‘Advancing knowledge of the plant nuclear periphery and its application for crop science’, *Nucleus*, 11(1), pp. 347–363. Available at: <https://doi.org/10.1080/19491034.2020.1838697>.

- Fang, Y. and Gu, Y. (2021) ‘Regulation of Plant Immunity by Nuclear Membrane-Associated Mechanisms’, *Frontiers in Immunology*, 12, p. 771065. Available at: <https://doi.org/10.3389/fimmu.2021.771065>.
- Feng, C.-M. *et al.* (2014) ‘Light-regulated gene repositioning in Arabidopsis’, *Nature Communications*, 5(1). Available at: <https://doi.org/10.1038/ncomms4027>.
- Feng, S. *et al.* (2014) ‘Genome-wide Hi-C Analyses in Wild-Type and Mutants Reveal High-Resolution Chromatin Interactions in Arabidopsis’, *Molecular Cell*, 55(5), pp. 694–707. Available at: <https://doi.org/10.1016/j.molcel.2014.07.008>.
- Fiserova, J., Kiseleva, E. and Goldberg, M.W. (2009) ‘Nuclear envelope and nuclear pore complex structure and organization in tobacco BY-2 cells’, *The Plant Journal: For Cell and Molecular Biology*, 59(2), pp. 243–255. Available at: <https://doi.org/10.1111/j.1365-313X.2009.03865.x>.
- Fontoura, B.M.A., Blobel, G. and Matunis, M.J. (1999) ‘A Conserved Biogenesis Pathway for Nucleoporins: Proteolytic Processing of a 186-Kilodalton Precursor Generates Nup98 and the Novel Nucleoporin, Nup96’, *The Journal of Cell Biology*, 144(6), pp. 1097–1112.
- Fransz, P. *et al.* (2002) ‘Interphase chromosomes in *Arabidopsis* are organized as well defined chromocenters from which euchromatin loops emanate’, *Proceedings of the National Academy of Sciences*, 99(22), pp. 14584–14589. Available at: <https://doi.org/10.1073/pnas.212325299>.
- Fransz, P. and de Jong, H. (2011) ‘From nucleosome to chromosome: a dynamic organization of genetic information’, *The Plant Journal*, 66(1), pp. 4–17. Available at: <https://doi.org/10.1111/j.1365-313X.2011.04526.x>.
- Gangappa, S.N., Berriri, S. and Kumar, S.V. (2017) ‘PIF4 Coordinates Thermosensory Growth and Immunity in Arabidopsis’, *Current Biology*, 27(2), pp. 243–249. Available at: <https://doi.org/10.1016/j.cub.2016.11.012>.
- García-Oliver, E., García-Molinero, V. and Rodríguez-Navarro, S. (2012) ‘mRNA export and gene expression: The SAGA–TREX-2 connection’, *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1819(6), pp. 555–565. Available at: <https://doi.org/10.1016/j.bbagr.2011.11.011>.
- Gomez-Cavazos, J.S. and Hetzer, M.W. (2012) ‘Outfits for different occasions: tissue-specific roles of Nuclear Envelope proteins’, *Current opinion in cell biology*, 24(6), pp. 775–783. Available at: <https://doi.org/10.1016/j.ceb.2012.08.008>.
- Görlich, D. and Kutay, U. (1999) ‘Transport Between the Cell Nucleus and the Cytoplasm’, *Annual Review of Cell and Developmental Biology*, 15(1), pp. 607–660. Available at: <https://doi.org/10.1146/annurev.cellbio.15.1.607>.
- Goto, C. *et al.* (2014) ‘The Novel Nuclear Envelope Protein KAKU4 Modulates Nuclear Morphology in Arabidopsis’, *The Plant Cell*, 26(5), pp. 2143–2155. Available at: <https://doi.org/10.1105/tpc.113.122168>.
- Goto, C. *et al.* (2020) ‘The nuclear envelope protein KAKU4 determines the migration order of the vegetative nucleus and sperm cells in pollen tubes’, *Journal of Experimental Botany*. Edited by Z. Wilson, 71(20), pp. 6273–6281. Available at: <https://doi.org/10.1093/jxb/eraa367>.

Graumann, K. *et al.* (2014a) ‘Characterization of two distinct subfamilies of SUN-domain proteins in Arabidopsis and their interactions with the novel KASH-domain protein AtTIK’, *Journal of Experimental Botany*, 65(22), pp. 6499–6512. Available at: <https://doi.org/10.1093/jxb/eru368>.

Graumann, K. *et al.* (2014b) ‘Characterization of two distinct subfamilies of SUN-domain proteins in Arabidopsis and their interactions with the novel KASH-domain protein AtTIK’, *Journal of Experimental Botany*, 65(22), pp. 6499–6512. Available at: <https://doi.org/10.1093/jxb/eru368>.

Graumann, K. (2014) ‘Evidence for LINC1-SUN Associations at the Plant Nuclear Periphery’, *PLoS ONE*, 9(3), p. e93406. Available at: <https://doi.org/10.1371/journal.pone.0093406>.

Graumann, K., Runions, J. and Evans, D.E. (2010) ‘Characterization of SUN-domain proteins at the higher plant nuclear envelope’, *The Plant Journal*, 61(1), pp. 134–144. Available at: <https://doi.org/10.1111/j.1365-313X.2009.04038.x>.

Grob, S. (2020) ‘Three-dimensional chromosome organization in flowering plants’, *Briefings in Functional Genomics*, 19(2), pp. 83–91. Available at: <https://doi.org/10.1093/bfgp/elz024>.

Grob, S. and Grossniklaus, U. (2019) ‘Invasive DNA elements modify the nuclear architecture of their insertion site by KNOT-linked silencing in Arabidopsis thaliana’, *Genome Biology*, 20, p. 120. Available at: <https://doi.org/10.1186/s13059-019-1722-3>.

Grob, S., Schmid, M.W. and Grossniklaus, U. (2014) ‘Hi-C Analysis in Arabidopsis Identifies the KNOT, a Structure with Similarities to the flamenco Locus of Drosophila’, *Molecular Cell*, 55(5), pp. 678–693. Available at: <https://doi.org/10.1016/j.molcel.2014.07.009>.

Grossman, E., Medalia, O. and Zwerger, M. (2012) ‘Functional Architecture of the Nuclear Pore Complex’, *Annual Review of Biophysics*, 41(1), pp. 557–584. Available at: <https://doi.org/10.1146/annurev-biophys-050511-102328>.

Gu, Y. *et al.* (2016) ‘Nuclear Pore Permeabilization Is a Convergent Signaling Event in Effector-Triggered Immunity’, *Cell*, 166(6), pp. 1526–1538.e11. Available at: <https://doi.org/10.1016/j.cell.2016.07.042>.

Guelen, L. *et al.* (2008a) ‘Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions’, *Nature*, 453(7197), pp. 948–951. Available at: <https://doi.org/10.1038/nature06947>.

Guelen, L. *et al.* (2008b) ‘Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions’, *Nature*, 453(7197), pp. 948–951. Available at: <https://doi.org/10.1038/nature06947>.

Gumber, H.K. *et al.* (2019) ‘Identification and characterization of genes encoding the nuclear envelope LINC complex in the monocot species Zea mays’, *Journal of Cell Science*, 132(3). Available at: <https://doi.org/10.1242/jcs.221390>.

Guo, T. *et al.* (2017) ‘Lamin-like Proteins Negatively Regulate Plant Immunity through NAC WITH TRANSMEMBRANE MOTIF1-LIKE9 and NONEXPRESSOR OF PR GENES1 in Arabidopsis thaliana’, *Molecular Plant*, 10(10), pp. 1334–1348. Available at: <https://doi.org/10.1016/j.molp.2017.09.008>.

- Herridge, R.P. *et al.* (2011) ‘Rapid analysis of seed size in *Arabidopsis* for mutant and QTL discovery’, *Plant Methods*, 7(1), p. 3. Available at: <https://doi.org/10.1186/1746-4811-7-3>.
- Hu, B. *et al.* (2019) ‘Plant lamin-like proteins mediate chromatin tethering at the nuclear periphery’, *Genome Biology*, 20(1). Available at: <https://doi.org/10.1186/s13059-019-1694-3>.
- Huang, S. *et al.* (2021) ‘A phase-separated nuclear GBPL circuit controls immunity in plants’, *Nature*, 594(7863), pp. 424–429. Available at: <https://doi.org/10.1038/s41586-021-03572-6>.
- Huang, Y. *et al.* (2021) ‘Polycomb-dependent differential chromatin compartmentalization determines gene coregulation in *Arabidopsis*’, *Genome Research* [Preprint]. Available at: <https://doi.org/10.1101/gr.273771.120>.
- Hüther, P. *et al.* (2020) ‘ARADEEPOPSIS, an Automated Workflow for Top-View Plant Phenomics using Semantic Segmentation of Leaf States’, *The Plant Cell*, 32(12), pp. 3674–3688. Available at: <https://doi.org/10.1105/tpc.20.00318>.
- Incarbone, M. and Dunoyer, P. (2013) ‘RNA silencing and its suppression: novel insights from in planta analyses’, *Trends in Plant Science*, 18(7), pp. 382–392. Available at: <https://doi.org/10.1016/j.tplants.2013.04.001>.
- Ishimura, A. *et al.* (2006) ‘Man1, an inner nuclear membrane protein, regulates vascular remodeling by modulating transforming growth factor β signaling’, *Development*, 133(19), pp. 3919–3928. Available at: <https://doi.org/10.1242/dev.02538>.
- Jarad, M. *et al.* (2020) ‘The Lamin-Like LITTLE NUCLEI 1 (LINC1) Regulates Pattern-Triggered Immunity and Jasmonic Acid Signaling’, *Frontiers in Plant Science*, 10. Available at: <https://doi.org/10.3389/fpls.2019.01639>.
- Jean-Baptiste, K. *et al.* (2019) ‘Dynamics of Gene Expression in Single Root Cells of *Arabidopsis thaliana*’, *The Plant Cell*, 31(5), pp. 993–1011. Available at: <https://doi.org/10.1105/tpc.18.00785>.
- Jin, Q. *et al.* (1998) ‘Yeast Nuclei Display Prominent Centromere Clustering That Is Reduced in Nondividing Cells and in Meiotic Prophase’, *The Journal of Cell Biology*, 141(1), pp. 21–29.
- Jupe, F. *et al.* (2019) ‘The complex architecture and epigenomic impact of plant T-DNA insertions’, *PLoS Genetics*, 15(1). Available at: <https://doi.org/10.1371/journal.pgen.1007819>.
- Kalverda, B. *et al.* (2010) ‘Nucleoporins Directly Stimulate Expression of Developmental and Cell-Cycle Genes Inside the Nucleoplasm’, *Cell*, 140(3), pp. 360–371. Available at: <https://doi.org/10.1016/j.cell.2010.01.011>.
- Kim, D.I. *et al.* (2016) ‘An improved smaller biotin ligase for BioID proximity labeling’, *Molecular Biology of the Cell*, 27(8), pp. 1188–1196. Available at: <https://doi.org/10.1091/mbc.E15-12-0844>.
- Kimura, Y. *et al.* (2014) ‘Localization of *Daucus carota* NMCP1 to the nuclear periphery: the role of the N-terminal region and an NLS-linked sequence motif, RYNLRR, in the tail domain’, *Frontiers in Plant Science*, 5. Available at: <https://doi.org/10.3389/fpls.2014.00062>.

- Kolářová, K. *et al.* (2021) ‘Disruption of NAP1 genes in *Arabidopsis thaliana* suppresses the *fas1* mutant phenotype, enhances genome stability and changes chromatin compaction’, *The Plant Journal*, 106(1), pp. 56–73. Available at: <https://doi.org/10.1111/tpj.15145>.
- Kumar, M. *et al.* (2022) ‘The Eukaryotic Linear Motif resource: 2022 release’, *Nucleic Acids Research*, 50(D1), pp. D497–D508. Available at: <https://doi.org/10.1093/nar/gkab975>.
- Kume, K. *et al.* (2019) ‘Nuclear membrane protein Lem2 regulates nuclear size through membrane flow’, *Nature Communications*, 10, p. 1871. Available at: <https://doi.org/10.1038/s41467-019-09623-x>.
- Lefevre, H., Bauters, L. and Gheysen, G. (2020) ‘Salicylic Acid Biosynthesis in Plants’, *Frontiers in Plant Science*, 11. Available at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.00338> (Accessed: 11 October 2022).
- Li, E. *et al.* (2019) ‘Long-range interactions between proximal and distal regulatory regions in maize’, *Nature Communications*, 10(1), p. 2633. Available at: <https://doi.org/10.1038/s41467-019-10603-4>.
- Li, E.W., McKee-Muir, O.C. and Gilbert, P.M. (2018) ‘Cellular Biomechanics in Skeletal Muscle Regeneration’, in *Current Topics in Developmental Biology*. Elsevier, pp. 125–176. Available at: <https://doi.org/10.1016/bs.ctdb.2017.08.007>.
- Lichter, P. *et al.* (1988) ‘Rapid detection of human chromosome 21 aberrations by in situ hybridization.’, *Proceedings of the National Academy of Sciences*, 85(24), pp. 9664–9668. Available at: <https://doi.org/10.1073/pnas.85.24.9664>.
- Lieberman-Aiden, E. *et al.* (2009) ‘Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome’, *Science*, 326(5950), pp. 289–293. Available at: <https://doi.org/10.1126/science.1181369>.
- Liu, C. *et al.* (2017) ‘Prominent topologically associated domains differentiate global chromatin packing in rice from *Arabidopsis*’, *Nature Plants*, 3(9), pp. 742–748. Available at: <https://doi.org/10.1038/s41477-017-0005-9>.
- Liu, H. *et al.* (2018) ‘Distinct heat shock factors and chromatin modifications mediate the organ-autonomous transcriptional memory of heat stress’, *The Plant Journal*, 95(3), pp. 401–413. Available at: <https://doi.org/10.1111/tpj.13958>.
- Liu, Q., Shi, L. and Fang, Y. (2012) ‘Dicing Bodies’, *Plant Physiology*, 158(1), pp. 61–66. Available at: <https://doi.org/10.1104/pp.111.186734>.
- Liu, Y. *et al.* (2014) ‘Nucleolus-tethering system (NoTS) reveals that assembly of photobodies follows a self-organization model’, *Molecular Biology of the Cell*, 25(8), pp. 1366–1373. Available at: <https://doi.org/10.1091/mbc.E13-09-0527>.
- Liu, Z. *et al.* (2022) ‘A Nucleoporin NUP58 modulates responses to drought and salt stress in maize (*Zea mays* L.)’, *Plant Science*, 320, p. 111296. Available at: <https://doi.org/10.1016/j.plantsci.2022.111296>.

López-García, P. and Moreira, D. (2015) 'Open Questions on the Origin of Eukaryotes', *Trends in Ecology & Evolution*, 30(11), pp. 697–708. Available at: <https://doi.org/10.1016/j.tree.2015.09.005>.

Louwers, M. *et al.* (2009) 'Tissue- and Expression Level-Specific Chromatin Looping at Maize *b1* Epialleles', *The Plant Cell*, 21(3), pp. 832–842. Available at: <https://doi.org/10.1105/tpc.108.064329>.

Love, A.J. *et al.* (2016) 'Cajal bodies and their role in plant stress and disease responses', *RNA Biology*, 14(6), pp. 779–790. Available at: <https://doi.org/10.1080/15476286.2016.1243650>.

Lu, Q. *et al.* (2009) 'Arabidopsis homolog of the yeast TREX-2 mRNA export complex: components and anchoring nucleoporin: TREX-2 mRNA export complex', *The Plant Journal*, 61(2), pp. 259–270. Available at: <https://doi.org/10.1111/j.1365-313X.2009.04048.x>.

Lüdke, D. *et al.* (2022) 'NLR we there yet? Nucleocytoplasmic coordination of NLR-mediated immunity', *New Phytologist*, n/a(n/a). Available at: <https://doi.org/10.1111/nph.18359>.

Mao, Y.S., Zhang, B. and Spector, D.L. (2011) 'Biogenesis and Function of Nuclear Bodies', *Trends in genetics: TIG*, 27(8), pp. 295–306. Available at: <https://doi.org/10.1016/j.tig.2011.05.006>.

Martin, W.F., Garg, S. and Zimorski, V. (2015) 'Endosymbiotic theories for eukaryote origin', *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 370(1678), p. 20140330. Available at: <https://doi.org/10.1098/rstb.2014.0330>.

Masuda, K. *et al.* (1997) 'Peripheral Framework of Carrot Cell Nucleus Contains a Novel Protein Predicted to Exhibit a Long α -Helical Domain', *Experimental Cell Research*, 232(1), pp. 173–181. Available at: <https://doi.org/10.1006/excr.1997.3531>.

Masuda, K., Hikida, R. and Fujino, K. (2021) 'The plant nuclear lamina proteins NMCP1 and NMCP2 form a filamentous network with lateral filament associations', *Journal of Experimental Botany*. Edited by A. Dobritsa, 72(18), pp. 6190–6204. Available at: <https://doi.org/10.1093/jxb/erab243>.

Maul, G.G. *et al.* (1980) 'Investigation of the determinants of nuclear pore number', *Cytogenetic and Genome Research*, 26(2–4), pp. 175–190. Available at: <https://doi.org/10.1159/000131439>.

McBride, A.E. *et al.* (2009) 'Specific sequences within arginine-glycine-rich domains affect mRNA-binding protein function', *Nucleic Acids Research*, 37(13), pp. 4322–4330. Available at: <https://doi.org/10.1093/nar/gkp349>.

McKenna, J.F. *et al.* (2021) 'Maize (*Zea mays* L.) Nucleoskeletal Proteins Regulate Nuclear Envelope Remodeling and Function in Stomatal Complex Development and Pollen Viability', *Frontiers in Plant Science*, 12, p. 645218. Available at: <https://doi.org/10.3389/fpls.2021.645218>.

Mermet, S. *et al.* (submitted) 'Evolutionary conserved protein motifs drive interactions between the plant nucleoskeleton and nuclear pores'.

- Mikulski, P. *et al.* (2019) ‘The Chromatin-Associated Protein PWO1 Interacts with Plant Nuclear Lamin-like Components to Regulate Nuclear Size’, *The Plant Cell*, 31(5), pp. 1141–1154. Available at: <https://doi.org/10.1105/tpc.18.00663>.
- Morgan, D. (2007) *The cell cycle: principles of control*. Corby: Oxford University Press (Primers in biology).
- Mosalaganti, S. *et al.* (2018) ‘In situ architecture of the algal nuclear pore complex’, *Nature Communications*, 9(1), p. 2361. Available at: <https://doi.org/10.1038/s41467-018-04739-y>.
- Ning, J. *et al.* (2010) ‘A Raf-Like MAPKKK Gene DSM1 Mediates Drought Resistance through Reactive Oxygen Species Scavenging in Rice’, *Plant Physiology*, 152(2), pp. 876–890. Available at: <https://doi.org/10.1104/pp.109.149856>.
- Pan, C. *et al.* (2021) ‘PIF4 negatively modulates cold tolerance in tomato anthers via temperature-dependent regulation of tapetal cell death’, *The Plant Cell*, 33(7), pp. 2320–2339. Available at: <https://doi.org/10.1093/plcell/koab120>.
- Parry, G. (2015) ‘The plant nuclear envelope and regulation of gene expression’, *Journal of Experimental Botany*, 66(6), pp. 1673–1685. Available at: <https://doi.org/10.1093/jxb/erv023>.
- Pawar, V. *et al.* (2016) ‘A novel family of plant nuclear envelope-associated proteins’, *Journal of Experimental Botany*, 67(19), pp. 5699–5710. Available at: <https://doi.org/10.1093/jxb/erw332>.
- Pečenková, T. *et al.* (2022) ‘Immunity functions of Arabidopsis pathogenesis-related 1 are coupled but not confined to its C-terminus processing and trafficking’, *Molecular Plant Pathology*, 23(5), pp. 664–678. Available at: <https://doi.org/10.1111/mpp.13187>.
- Pecinka, A. *et al.* (2004) ‘Chromosome territory arrangement and homologous pairing in nuclei of Arabidopsis thaliana are predominantly random except for NOR-bearing chromosomes’, *Chromosoma*, 113(5), pp. 258–269. Available at: <https://doi.org/10.1007/s00412-004-0316-2>.
- Pei, L. *et al.* (2021) ‘Plant 3D genomics: the exploration and application of chromatin organization’, *New Phytologist*, 230(5), pp. 1772–1786. Available at: <https://doi.org/10.1111/nph.17262>.
- Pickersgill, H. *et al.* (2006) ‘Characterization of the Drosophila melanogaster genome at the nuclear lamina’, *Nature Genetics*, 38(9), pp. 1005–1014. Available at: <https://doi.org/10.1038/ng1852>.
- Pinkel, D. *et al.* (1988) ‘Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4.’, *Proceedings of the National Academy of Sciences*, 85(23), pp. 9138–9142. Available at: <https://doi.org/10.1073/pnas.85.23.9138>.
- Pontvianne, F. *et al.* (2016) ‘Identification of Nucleolus-Associated Chromatin Domains Reveals a Role for the Nucleolus in 3D Organization of the A. thaliana Genome’, *Cell Reports*, 16(6), pp. 1574–1587. Available at: <https://doi.org/10.1016/j.celrep.2016.07.016>.

- Pontvianne, F. and Liu, C. (2020) ‘Chromatin domains in space and their functional implications’, *Current Opinion in Plant Biology*, 54, pp. 1–10. Available at: <https://doi.org/10.1016/j.pbi.2019.11.005>.
- Poulet, A., Probst, A.V., *et al.* (2017) ‘Exploring the evolution of the proteins of the plant nuclear envelope’, *Nucleus*, 8(1), pp. 46–59. Available at: <https://doi.org/10.1080/19491034.2016.1236166>.
- Poulet, A., Duc, C., *et al.* (2017a) ‘The LINC complex contributes to heterochromatin organisation and transcriptional gene silencing in plants’, *Journal of Cell Science*, 130(3), pp. 590–601. Available at: <https://doi.org/10.1242/jcs.194712>.
- Poulet, A., Duc, C., *et al.* (2017b) ‘The LINC complex contributes to heterochromatin organisation and transcriptional gene silencing in plants’, *Journal of Cell Science*, 130(3), pp. 590–601. Available at: <https://doi.org/10.1242/jcs.194712>.
- Probst, A.V., Desvoyes, B. and Gutierrez, C. (2020) ‘Similar yet critically different: the distribution, dynamics and function of histone variants’, *Journal of Experimental Botany*. Edited by G. Parry, 71(17), pp. 5191–5204. Available at: <https://doi.org/10.1093/jxb/eraa230>.
- Rabl C., C. (1885) ‘Über Zelltheilung.’, *Morphologisches Jahrbuch*, p. 10:214-330.
- Rao, S.S.P. *et al.* (2014) ‘A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping’, *Cell*, 159(7), pp. 1665–1680. Available at: <https://doi.org/10.1016/j.cell.2014.11.021>.
- Rebollo, A., Savier, E. and Tuffery, P. (2021) ‘Pepscan Approach for the Identification of Protein–Protein Interfaces: Lessons from Experiment’, p. 14.
- Robinson, J.T. *et al.* (2011) ‘Integrative genomics viewer’, *Nature Biotechnology*, 29(1), pp. 24–26. Available at: <https://doi.org/10.1038/nbt.1754>.
- Rodríguez-Navarro, S. *et al.* (2004) ‘Sus1, a Functional Component of the SAGA Histone Acetylase Complex and the Nuclear Pore-Associated mRNA Export Machinery’, *Cell*, 116(1), pp. 75–86. Available at: [https://doi.org/10.1016/S0092-8674\(03\)01025-0](https://doi.org/10.1016/S0092-8674(03)01025-0).
- Rowley, M.J. and Corces, V.G. (2016) ‘The three-dimensional genome: principles and roles of long-distance interactions’, *Current Opinion in Cell Biology*, 40, pp. 8–14. Available at: <https://doi.org/10.1016/j.ceb.2016.01.009>.
- Sakamoto, T. *et al.* (2022) ‘Two-step regulation of centromere distribution by condensin II and the nuclear envelope proteins’, *Nature Plants* [Preprint]. Available at: <https://doi.org/10.1038/s41477-022-01200-3>.
- Sakamoto, Y. *et al.* (2020) ‘Subnuclear gene positioning through lamina association affects copper tolerance’, *Nature Communications*, 11. Available at: <https://doi.org/10.1038/s41467-020-19621-z>.
- Sakamoto, Y. and Takagi, S. (2013) ‘LITTLE NUCLEI 1 and 4 regulate nuclear morphology in *Arabidopsis thaliana*’, *Plant & Cell Physiology*, 54(4), pp. 622–633. Available at: <https://doi.org/10.1093/pcp/pct031>.

Santos, A.P. and Shaw, P. (2004) 'Interphase chromosomes and the Rab1 configuration: does genome size matter?', *Journal of Microscopy*, 214(2), pp. 201–206. Available at: <https://doi.org/10.1111/j.0022-2720.2004.01324.x>.

Schmid, M. *et al.* (2006) 'Nup-PI: The Nucleopore-Promoter Interaction of Genes in Yeast', *Molecular Cell*, 21(3), pp. 379–391. Available at: <https://doi.org/10.1016/j.molcel.2005.12.012>.

Schwartz, T. (2013) 'Functional Insights from Studies on the Structure of the Nuclear Pore and Coat Protein Complexes', *Cold Spring Harbor Perspectives in Biology*, 5(7), pp. a013375–a013375. Available at: <https://doi.org/10.1101/cshperspect.a013375>.

Sequeira-Mendes, J. *et al.* (2014) 'The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States[C][W]', *The Plant Cell*, 26(6), pp. 2351–2366. Available at: <https://doi.org/10.1105/tpc.114.124578>.

Shaw, P.J. and Brown, J.W. (2004) 'Plant nuclear bodies', *Current Opinion in Plant Biology*, 7(6), pp. 614–620. Available at: <https://doi.org/10.1016/j.pbi.2004.09.011>.

Shen, C. *et al.* (2021) 'Phase separation drives RNA virus-induced activation of the NLRP6 inflammasome', *Cell*, 184(23), pp. 5759–5774.e20. Available at: <https://doi.org/10.1016/j.cell.2021.09.032>.

Shin, Y. *et al.* (2017) 'Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets', *Cell*, 168(1–2), pp. 159–171.e14. Available at: <https://doi.org/10.1016/j.cell.2016.11.054>.

Sikorska, N. and Sexton, T. (2019) 'Defining Functionally Relevant Spatial Chromatin Domains: It is a TAD Complicated', *Journal of Molecular Biology*, 432. Available at: <https://doi.org/10.1016/j.jmb.2019.12.006>.

Sleeman, J.E. and Trinkle-Mulcahy, L. (2014) 'Nuclear bodies: new insights into assembly/dynamics and disease relevance', *Current Opinion in Cell Biology*, 28, pp. 76–83. Available at: <https://doi.org/10.1016/j.ceb.2014.03.004>.

Smith, S. *et al.* (2015) 'Marker gene tethering by nucleoporins affects gene expression in plants', *Nucleus*, 6(6), pp. 471–478. Available at: <https://doi.org/10.1080/19491034.2015.1126028>.

Smoyer, C.J. *et al.* (2016) 'Analysis of membrane proteins localizing to the inner nuclear envelope in living cells', *The Journal of Cell Biology*, 215(4), pp. 575–590. Available at: <https://doi.org/10.1083/jcb.201607043>.

Smythe, C., Jenkins, H.E. and Hutchison, C.J. (2000) 'Incorporation of the nuclear pore basket protein nup153 into nuclear pore structures is dependent upon lamina assembly: evidence from cell-free extracts of *Xenopus* eggs', *The EMBO journal*, 19(15), pp. 3918–3931. Available at: <https://doi.org/10.1093/emboj/19.15.3918>.

Snider, J. *et al.* (2010) 'Split-Ubiquitin Based Membrane Yeast Two-Hybrid (MYTH) System: A Powerful Tool For Identifying Protein-Protein Interactions', *Journal of Visualized Experiments : JoVE*, (36), p. 1698. Available at: <https://doi.org/10.3791/1698>.

Sosa, B.A. *et al.* (2013) ‘LINC Complexes Form by Binding of Three KASH Peptides to the Interfaces of Trimeric SUN proteins’, p. 26.

Sotelo-Silveira, M. *et al.* (2018) ‘Entering the Next Dimension: Plant Genomes in 3D’, *Trends in Plant Science*, 23(7), pp. 598–612. Available at: <https://doi.org/10.1016/j.tplants.2018.03.014>.

Stam, M., Tark-Dame, M. and Fransz, P. (2019) ‘3D genome organization: a role for phase separation and loop extrusion?’, *Current Opinion in Plant Biology*, 48, pp. 36–46. Available at: <https://doi.org/10.1016/j.pbi.2019.03.008>.

van Steensel, B. and Belmont, A.S. (2017) ‘Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression’, *Cell*, 169(5), pp. 780–791. Available at: <https://doi.org/10.1016/j.cell.2017.04.022>.

Supek, F. *et al.* (2011) ‘REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms’, *PLOS ONE*, 6(7), p. e21800. Available at: <https://doi.org/10.1371/journal.pone.0021800>.

Tamura, K. *et al.* (2010) ‘Identification and characterization of nuclear pore complex components in *Arabidopsis thaliana*’, *The Plant Cell*, 22(12), pp. 4084–4097. Available at: <https://doi.org/10.1105/tpc.110.079947>.

Tamura, K. *et al.* (2017) ‘Nup82 functions redundantly with Nup136 in a salicylic acid-dependent defense response of *Arabidopsis thaliana*’, *Nucleus*, 8(3), pp. 301–311. Available at: <https://doi.org/10.1080/19491034.2017.1279774>.

Tamura, K. (2020) ‘Nuclear pore complex-mediated gene expression in *Arabidopsis thaliana*’, *Journal of Plant Research* [Preprint]. Available at: <https://doi.org/10.1007/s10265-020-01177-0>.

Tamura, K. and Hara-Nishimura, I. (2011) ‘Involvement of the nuclear pore complex in morphology of the plant nucleus’, *Nucleus (Austin, Tex.)*, 2(3), pp. 168–172. Available at: <https://doi.org/10.4161/nucl.2.3.16175>.

Tamura, K. and Hara-Nishimura, I. (2013) ‘The molecular architecture of the plant nuclear pore complex’, *Journal of Experimental Botany*, 64(4), pp. 823–832. Available at: <https://doi.org/10.1093/jxb/ers258>.

Tang, Y. *et al.* (2021) ‘PNET2 is a component of the plant nuclear lamina and is required for proper genome organization and activity’, *Developmental Cell*, 57(1), pp. 19–31.e6. Available at: <https://doi.org/10.1016/j.devcel.2021.11.002>.

Tang, Y., Huang, A. and Gu, Y. (2020) ‘Global profiling of plant nuclear membrane proteome in *Arabidopsis*’, *Nature Plants*, 6(7), pp. 838–847. Available at: <https://doi.org/10.1038/s41477-020-0700-9>.

Tan-Wong, S.M., Wijayatilake, H.D. and Proudfoot, N.J. (2009) ‘Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex’, *Genes & Development*, 23(22), pp. 2610–2624. Available at: <https://doi.org/10.1101/gad.1823209>.

- Tatout, C. *et al.* (2014) 'The plant LINC complex at the nuclear envelope', *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology*, 22(2), pp. 241–252. Available at: <https://doi.org/10.1007/s10577-014-9419-7>.
- Uemura, E. *et al.* (2018) 'Large-scale aggregation analysis of eukaryotic proteins reveals an involvement of intrinsically disordered regions in protein folding', *Scientific Reports*, 8(1), p. 678. Available at: <https://doi.org/10.1038/s41598-017-18977-5>.
- Van Damme, D. *et al.* (2004) 'Molecular dissection of plant cytokinesis and phragmoplast structure: a survey of GFP-tagged proteins', *The Plant Journal*, 40(3), pp. 386–398. Available at: <https://doi.org/10.1111/j.1365-313X.2004.02222.x>.
- Van de Vosse, D.W. *et al.* (2013) 'A role for the nucleoporin Nup170p in chromatin structure and gene silencing', *Cell*, 152(5), pp. 969–983. Available at: <https://doi.org/10.1016/j.cell.2013.01.049>.
- Van Roey, K. *et al.* (2014) 'Short Linear Motifs: Ubiquitous and Functionally Diverse Protein Interaction Modules Directing Cell Regulation', *Chemical Reviews*, 114(13), pp. 6733–6778. Available at: <https://doi.org/10.1021/cr400585q>.
- Verma, V., Ravindran, P. and Kumar, P.P. (2016) 'Plant hormone-mediated regulation of stress responses', *BMC Plant Biology*, 16(1), p. 86. Available at: <https://doi.org/10.1186/s12870-016-0771-y>.
- Vertii, A. *et al.* (2019) 'Two contrasting classes of nucleolus-associated domains in mouse fibroblast heterochromatin', *Genome Research*, 29(8), pp. 1235–1249. Available at: <https://doi.org/10.1101/gr.247072.118>.
- Voisin, M., Vanrobays, E. and Tatout, C. (2018) 'Investigation of Nuclear Periphery Protein Interactions in Plants Using the Membrane Yeast Two-Hybrid (MbY2H) System', *Methods in Molecular Biology (Clifton, N.J.)*, 1840, pp. 221–235. Available at: https://doi.org/10.1007/978-1-4939-8691-0_16.
- Wang, F. *et al.* (2020) 'Crosstalk of PIF4 and DELLA modulates CBF transcript and hormone homeostasis in cold response in tomato', *Plant Biotechnology Journal*, 18(4), pp. 1041–1055. Available at: <https://doi.org/10.1111/pbi.13272>.
- Wang, H., Dittmer, T.A. and Richards, E.J. (2013) 'Arabidopsis CROWDED NUCLEI (CRWN) proteins are required for nuclear size control and heterochromatin organization', *BMC plant biology*, 13, p. 200. Available at: <https://doi.org/10.1186/1471-2229-13-200>.
- Wang, N. *et al.* (2021) 'Characterization of a Plant Nuclear Matrix Constituent Protein in Liverwort', *Frontiers in Plant Science*, 12. Available at: <https://doi.org/10.3389/fpls.2021.670306>.
- Wang, Q. *et al.* (2019) 'Roles of CRWN-family proteins in protecting genomic DNA against oxidative damage', *Journal of Plant Physiology*, 233, pp. 20–30. Available at: <https://doi.org/10.1016/j.jplph.2018.12.005>.

- Wang, S. *et al.* (2014) ‘A Noncanonical Role for the CKI-RB-E2F Cell Cycle Signaling Pathway in Plant Effector-Triggered Immunity’, *Cell host & microbe*, 16(6), pp. 787–794. Available at: <https://doi.org/10.1016/j.chom.2014.10.005>.
- Wiermer, M. *et al.* (2012) ‘Putative members of the Arabidopsis Nup107-160 nuclear pore sub-complex contribute to pathogen defense’, *The Plant Journal*, 70(5), pp. 796–808. Available at: <https://doi.org/10.1111/j.1365-313X.2012.04928.x>.
- Wilson, K.L. and Foisner, R. (2010) ‘Lamin-binding Proteins’, *Cold Spring Harbor Perspectives in Biology*, 2(4), p. a000554. Available at: <https://doi.org/10.1101/cshperspect.a000554>.
- Winey, M. *et al.* (1997) ‘Nuclear Pore Complex Number and Distribution throughout the Saccharomyces cerevisiae Cell Cycle by Three-Dimensional Reconstruction from Electron Micrographs of Nuclear Envelopes’, *Molecular Biology of the Cell*, 8, p. 14.
- Xing, H.-L. *et al.* (2014) ‘A CRISPR/Cas9 toolkit for multiplex genome editing in plants’, *BMC plant biology*, 14, p. 327. Available at: <https://doi.org/10.1186/s12870-014-0327-y>.
- Xu, X.M. *et al.* (2007) ‘NUCLEAR PORE ANCHOR, the Arabidopsis Homolog of Tpr/Mlp1/Mlp2/Megator, Is Involved in mRNA Export and SUMO Homeostasis and Affects Diverse Aspects of Plant Development’, *The Plant Cell*, 19(5), pp. 1537–1548. Available at: <https://doi.org/10.1105/tpc.106.049239>.
- Xu, Y.Z. *et al.* (2013) *Chromatin Remodelling During Host-Bacterial Pathogen Interaction, Chromatin Remodelling*. IntechOpen. Available at: <https://doi.org/10.5772/55977>.
- Yang, J. *et al.* (2020) ‘A lamin-like protein OsNMCP1 regulates drought resistance and root growth through chromatin accessibility modulation by interacting with a chromatin remodeller OsSWI3C in rice’, *New Phytologist*, 227(1), pp. 65–83. Available at: <https://doi.org/10.1111/nph.16518>.
- Zhang, A. *et al.* (2020) ‘Nuclear pore complex components have temperature-influenced roles in plant growth and immunity’, *Plant, Cell & Environment* [Preprint]. Available at: <https://doi.org/10.1111/pce.13741>.
- Zhang, X. *et al.* (2006) ‘Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis’, *Cell*, 126(6), pp. 1189–1201. Available at: <https://doi.org/10.1016/j.cell.2006.08.003>.
- Zhang, Y. and Li, X. (2005) ‘A putative nucleoporin 96 Is required for both basal defense and constitutive resistance responses mediated by suppressor of npr1-1, constitutive 1’, *The Plant Cell*, 17(4), pp. 1306–1316. Available at: <https://doi.org/10.1105/tpc.104.029926>.
- Zhao, W. *et al.* (2016) ‘The Arabidopsis CROWDED NUCLEI genes regulate seed germination by modulating degradation of ABI5 protein: CRWNs regulate seed germination’, *Journal of Integrative Plant Biology*, 58(7), pp. 669–678. Available at: <https://doi.org/10.1111/jipb.12448>.
- Zhou, M. *et al.* (2015) ‘Redox rhythm reinforces the circadian clock to gate immune response’, *Nature*, 523(7561), pp. 472–476. Available at: <https://doi.org/10.1038/nature14449>.

Zhou, X. *et al.* (2012a) ‘Novel plant SUN–KASH bridges are involved in RanGAP anchoring and nuclear shape determination’, *The Journal of Cell Biology*, 196(2), pp. 203–211. Available at: <https://doi.org/10.1083/jcb.201108098>.

Zhou, X. *et al.* (2012b) ‘Novel plant SUN–KASH bridges are involved in RanGAP anchoring and nuclear shape determination’, *The Journal of Cell Biology*, 196(2), pp. 203–211. Available at: <https://doi.org/10.1083/jcb.201108098>.

Zhou, X. *et al.* (2014) ‘Identification of unique SUN-interacting nuclear envelope proteins with diverse functions in plants’, *The Journal of Cell Biology*, 205(5), pp. 677–692. Available at: <https://doi.org/10.1083/jcb.201401138>.

Zhou, X., Groves, N.R. and Meier, I. (2015) ‘Plant nuclear shape is independently determined by the SUN-WIP-WIT2-myosin XI-i complex and CRWN1’, *Nucleus (Austin, Tex.)*, 6(2), pp. 144–153. Available at: <https://doi.org/10.1080/19491034.2014.1003512>.

Zhou, X. and Meier, I. (2014) ‘Efficient plant male fertility depends on vegetative nuclear movement mediated by two families of plant outer nuclear membrane proteins’, *Proceedings of the National Academy of Sciences*, 111(32), pp. 11900–11905. Available at: <https://doi.org/10.1073/pnas.1323104111>.

Annexes

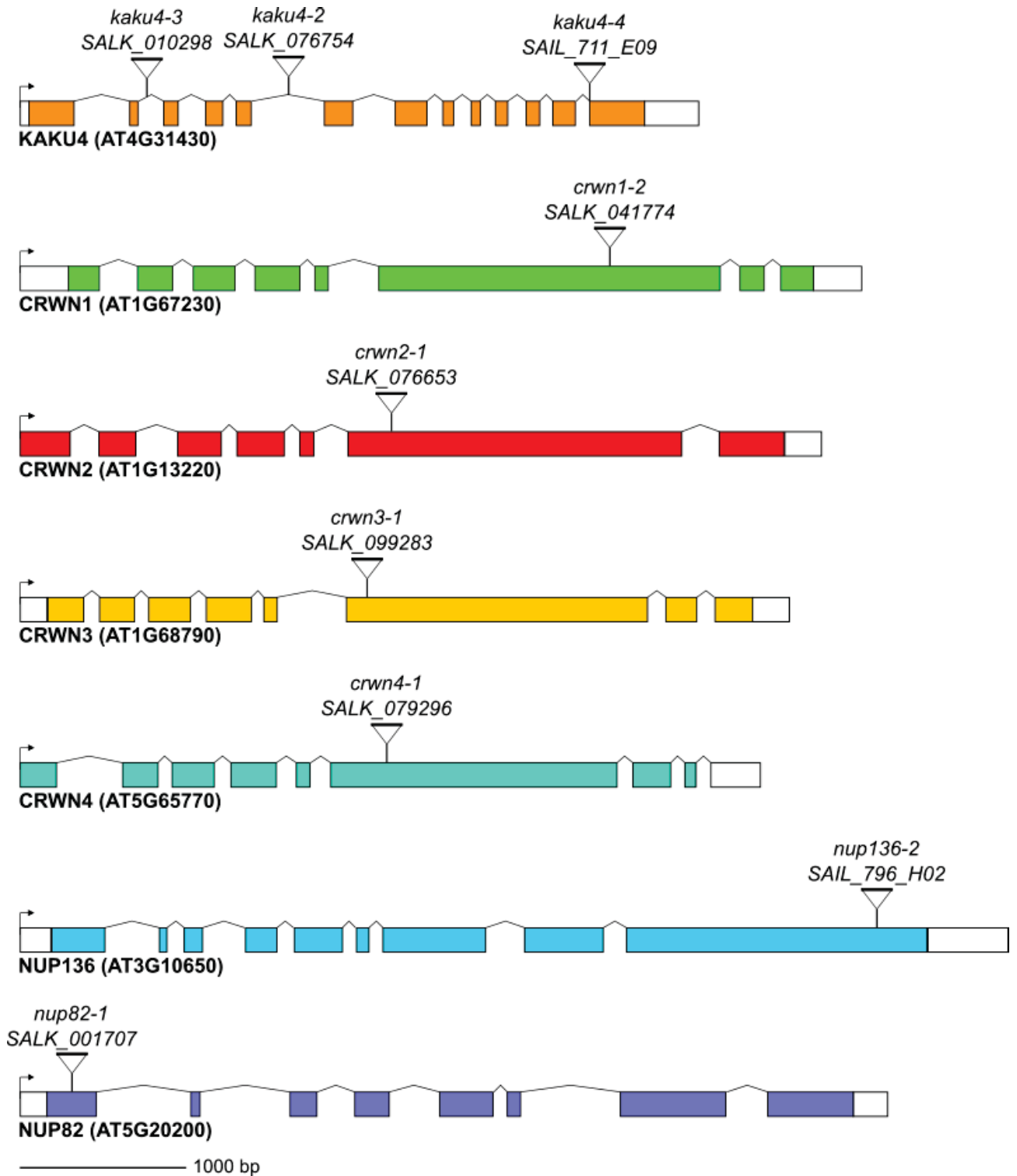
Annexe 1 : Différentes lignées mutantes d'insertion et leur vérification par RT-PCR

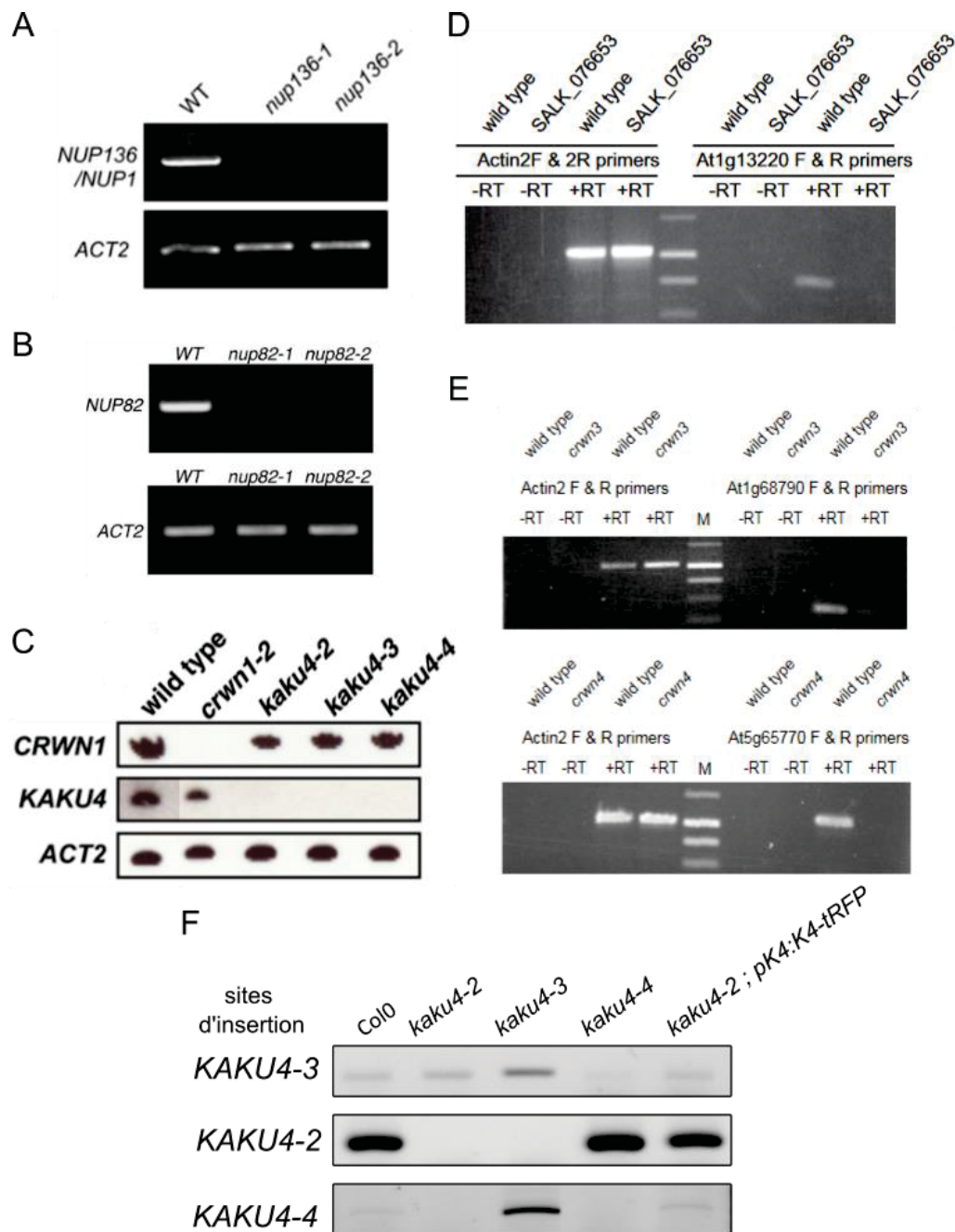
Annexe 2 : Revue “Advancing knowledge of the plant nuclear periphery and its application for crop science”

Annexe 3 : Données supplémentaires article “Evolutionary conserved protein motifs drive interactions between the plant nucleoskeleton and nuclear pores”

Annexe 4 : Article “Image analysis workflows to reveal the spatial organization of chromatin and chromosomes”

Annexe 1 : Différentes lignées mutantes d'insertion et leur vérification par RT-PCR





Vérification par RT-PCR des différents mutants utilisés pour créer nos multi-mutants.

A. Absence de transcrit pleine longueur de *NUP136* dans les mutants *nup136* (Tamura *et al.*, 2010). **B.** Absence de transcrit pleine longueur de *NUP82* dans les mutants *nup82* (Tamura *et al.*, 2017). **C.** Absence de transcrit pleine longueur de *KAKU4* et *CRWN1* des mutants *kaku4* et *crwn1-2* (Goto *et al.*, 2014). **D.** Absence de transcrit après le site d'insertion du T-DNA de *CRWN2* dans le mutant *crwn2*. **E.** Le mutant *crwn3* possède un faible transcrit après l'insertion du T-DNA dans les mutants *crwn3* (en bas) et absence de transcrit après le site d'insertion du T-DNA *CRWN4* et *crwn4*. (Wang, Dittmer and Richards, 2013). **F.** Vérification par RT-PCR de l'absence de transcrit autour du site d'insertion du T-DNA au niveau de *KAKU4* (*KAKU4-2*, *KAKU4-3*, *KAKU4-4* : PCR autour des sites d'insertion impliqués dans chacun des mutants *kaku4* ; cf ci-dessus : carte de *KAKU4*) dans les mutants *kaku4-2*, *kaku4-3*, *kaku4-4* et *kaku4-2 ; pK4:K4-tRFP*. Le mutant *kaku4-2* possède un transcrit en début du gène mais pas après l'insertion ; le mutant *kaku4-3* possède un transcrit tout le long du gène comme le mutant *kaku4-4* sauf que ce dernier n'a pas de transcrit sur la fin du gène.

Annexe 2 : Revue “Advancing knowledge of the plant nuclear periphery and its application for crop science”

Advancing knowledge of the plant nuclear periphery and its application for crop science

David E. Evans ^a, Sarah Mermet^b, and Christophe Tatout ^b

^aDepartment of Biological and Medical Sciences, Oxford Brookes University, Oxford, UK; ^bGReD, CNRS, INSERM, Université Clermont Auvergne, Clermont-Ferrand, France

ABSTRACT

In this review, we explore recent advances in knowledge of the structure and dynamics of the plant nuclear envelope. As a paradigm, we focused our attention on the Linker of Nucleoskeleton and Cytoskeleton (LINC) complex, a structurally conserved bridging complex comprising SUN domain proteins in the inner nuclear membrane and KASH domain proteins in the outer nuclear membrane. Studies have revealed that this bridging complex has multiple functions with structural roles in positioning the nucleus within the cell, conveying signals across the membrane and organizing chromatin in the 3D nuclear space with impact on gene transcription. We also provide an up-to-date survey in nuclear dynamics research achieved so far in the model plant *Arabidopsis thaliana* that highlights its potential impact on several key plant functions such as growth, seed maturation and germination, reproduction and response to biotic and abiotic stress. Finally, we bring evidences that most of the constituents of the LINC Complex and associated components are, with some specificities, conserved in monocot and dicot crop species and are displaying very similar functions to those described for *Arabidopsis*. This leads us to suggest that a better knowledge of this system and a better account of its potential applications will in the future enhance the resilience and productivity of crop plants.

ARTICLE HISTORY

Received 10 July 2020
Revised 9 October 2020
Accepted 14 October 2020

KEYWORDS

LINC Complex; sun domain; KASH domain; nucleoskeleton; crop improvement; rice; maize; Medicago

Introduction to the plant nuclear periphery

The aim of this review is to explore the application of recent advances in knowledge of the structure and dynamics of the plant nucleus and its future application in enhancing the resilience and productivity of crop plants. The nucleus is a highly complex compartment and its functions depend on its ordered and dynamic structure. In a living cell, the nucleus is capable of movement, changes in shape and volume, and reorganizing chromatin to alter the position of genetic material. It responds to physical and environmental stimuli. It is organized spatially by the presence of proteins that form a scaffold-like structure supporting, moving and linking the nucleus, its membranes and contents.

While the principles of the structural organization of the nucleus are universal to all eukaryotes studied, studies have revealed that plants have a unique combination of proteins achieving these functions [1–3]. These components

include functional homologues of the nucleoskeleton and of proteins that are intrinsic to the inner and outer nuclear membrane (INM and ONM) and that link to the cytoskeleton. While showing structural conservation, the nuclear pore complex (NPC) also has unique features [4] although it will not be the focus of this review. Here we will review knowledge of the plant nuclear periphery focusing on three main components. First, the proteins of the ONM linking with the cytoskeleton, organelles and the plasma membrane to provide a physical signaling pathway and to position and move the nucleus. Second the proteins of the INM, acting as an intermediary in the transmission of signals from outside the nucleus into the nucleus and chromatin through their connection with the nucleoskeleton. Finally, the nucleoskeleton, connected to chromatin to achieve, inter alia, chromosome positioning and 3D structure of the genome allowing changes in expression of the genome.

CONTACT David E. Evans  deevans@brookes.ac.uk  Department of Biological and Medical Sciences, Oxford Brookes University, Oxford OX3 0BP, UK

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The plant Linker of Nucleoskeleton and Cytoskeleton complex

At the center of this interlinked chain of proteins is the Linker of Nucleoskeleton and Cytoskeleton (LINC) complex. It is structurally conserved in all eukaryotes, made up of a protein bridge spanning the INM and ONM across the nuclear periplasm with anchors to link to the structural proteins of the nucleoplasm and cytoplasm. A small family of INM proteins, Sad1/UNC84 homology (SUN) domain proteins are highly conserved; two, SUN1 and SUN2 are present in plants. These have a C-terminal SUN domain (Cter-SUN) in the periplasm which interacts with a second family of proteins located in the ONM, the Klarsicht/ANC-1/Syne Homology (KASH) domain proteins. In addition, between one and three mid-SUN proteins are present, having the SUN domain in a central position. While the Cter-SUN are highly enriched in the INM, a substantial fraction of mid-SUNs is also ER localized [5,6].

KASH domain proteins are structurally conserved but difficult to identify in plants due to absence of sequence conservation. The SUN-binding KASH domain is located in the nuclear periplasm and together with the SUN domain they form the bridging complex. KASH domain proteins are very variable in structure and function [7]. The first plant KASH proteins identified were WPP-domain interacting proteins (WIPs) that anchor Ran GTPase activating protein (RanGAP) to the NE [7], a key mechanism required for nucleocytoplasmic transport through the NPC. WIPs have a typical KASH family structure, including, a transmembrane domain and a short SUN-domain interacting sequence of the three amino acids VPT [2,8]. They also have a cytoplasmic coiled-coil domain. Anchorage of RanGAP to the envelope involves SUN domain proteins, two KASH domain proteins, WIPs (WIP1 and WIP2) and further associated proteins, the WPP domain-interacting tail anchored proteins, WIT1 and WIT2. It may also include further WPP proteins [7].

The second family of plant KASH domain proteins described were the SINEs (SUN-interacting NE proteins) [1]. *SINE1* and *SINE2* of *Arabidopsis thaliana* (*A. thaliana*), are both expressed in roots, they

show differential expression in leaves [9], with *SINE1* expressed in guard cells and their progenitors, while *SINE2* is expressed in trichomes, epidermal and mesophyll cells, and only weakly in mature guard cells [9]. *SINE1* and *SINE2* have an N-terminal with homology to armadillo (ARM) repeat domains, which are known to bind actin and act as protein protein interaction domains [10]. Co-localization with actin has been verified for *SINE1* but not for *SINE2*. While *SINE1* has a function in guard cell movement [9,11], *SINE2* has been shown to be involved in immunity to a plant pathogen [9]. *SINEs* 3–4 have short, unique cytoplasmic domains and their function is yet to be fully described. In addition, a further dicot KASH domain protein was described. The Toll Interleukin Receptor (TIR) domain KASH protein (TIK) was identified during studies of mid-SUN interactors [6]. AtTIK has a TIR domain and a C-terminal TM domain upstream of a PPPS motif, a characteristic signature of the KASH domain and appears to be present only in a few Brassicaceae species (*A. thaliana*, *A. lyrata*, ...) [6]. Most recently, a novel family of KASH domain proteins has been described in the gramineae. Named Maize LINC KASH Grass-specific (MLKG) [12], they are described in detail for maize (below).

The higher plant nuclear lamina

The protein composition of the plant nuclear lamina is actively being researched. The lamina in animals, classically described as a layer of filamentous material known as the nucleoskeleton, is structured by type V intermediate filaments, which are mainly made up of lamins A and B [13,14]. Lamins play a wide range of roles contributing to nuclear structure and chromatin organization [15] and mutations in lamins especially in lamin A are associated with a large number of human diseases reminiscent of those found associated with the LINC complex [16]. Putative *Arabidopsis* homologues of intermediate filaments have been investigated by electron microscope studies [17] which revealed a meshwork of fibers underlying and connected to the NE and Nuclear Pore Complexes (NPCs), and by bioinformatics [18], genetics [19,20] and characterization of a lamina-like

proteome [21]. The best candidates for lamin-like proteins in plants are the NUCLEAR MATRIX CONSTITUENT PROTEIN (NMCP) discovered in carrot and its *Arabidopsis* orthologues CRoWded Nuclei (CRWN) [19,20]. The four *Arabidopsis* CRWNs have long coiled-coil domains. Two of them (CRWN1 and CRWN4) are located at the nuclear periphery, while CRWN2 and CRWN3 are localized to the nucleoplasm [19,21]. Interaction of CRWNs with SUN1 and SUN2 has been demonstrated [22]. Like animal lamin mutants which show small, deformed nuclei, *crwn1* and *crwn4* mutants have small, more rounded nuclei [21,23], though this is not seen for *crwn2* or *crwn3*. Plants with a *crwn1* mutation combined with *crwn2*, *crwn3* or *crwn4* have even smaller nuclei [23].

A further component of the plant lamina is KAKU4, discovered in a mutant screen, where the mutants have smaller, spherical nuclei [24]. In *Arabidopsis*, KAKU4 interact with CRWN1 and CRWN4 and its overexpression results in NE overgrowth. Other novel components which are likely to play a role in the plant nucleoskeleton are the Nuclear Envelope Associated Protein (NEAP) family [25]. NEAPs comprise coiled-coil domains, a nuclear localization signal and a predicted transmembrane domain near the C-terminus. AtNEAPs interact with themselves and with both Cter- and mid-SUNs, and over expression relocates CRWN1 from the nuclear periphery to the nucleoplasm [25]. NEAPs and KAKU4 function in the lamina have yet to be fully explored.

Impact of the plant nuclear periphery on nuclear features

Nuclear shape and size

Nuclear morphology includes nuclear volume/size and nuclear shape and displays a large range of variation in plants. It is a complex trait as nuclear, cell and organ sizes are tightly regulated in a coordinated manner [26]. In animals, the nucleus is most often depicted as round or oval while its shape is variable [27], can be altered in some diseases [28] and during aging [26]. Plant nuclei are also polymorphic in size and shape as

observed in different tissues [29,30]. Nuclear shape is thought to be stabilized by the connection of INM and ONM to one another at nuclear pores and to interactions at the INM between INM proteins, the nuclear lamina and chromatin and at the ONM by forces from the cytoskeleton all of which are likely mediated through the LINC complex. Many genes altering nuclear size and shape have been identified and most of them encode components of the microtubules [31], ONM [32,33], INM [6], NPC [34] and the lamina-like CRWN family referred to previously [23,24] and KAKU4 [24]. Surprisingly, most if not all these mutations induce smaller and rounded nuclei with the exception of mutants of GIP (γ -TuC Protein 3 (GCP3)-Interacting Proteins, involved in nucleating microtubules at the nuclear periphery [35,36] while in human cells, ghost shaped nuclei are observed [37].

Poulet et al [30] explored the role of components of the LINC complex, including KASH (*wifi: wip1 wip2 wip3 wit1 wit2* quintuple mutant) and SUN (*sun1 sun4 sun5* triple mutant) domain proteins in determining nuclear shape, size, heterochromatin organization, chromocentre position and activity of transcriptionally silent repetitive sequences. They observed altered nuclear shapes in three cell types investigated for SUN or KASH proteins and a nuclear lamina component (*crwn1 crwn2* double mutant). Mikulski et al [38] observed that the chromatin- and polycomb-associated component, PROLINE-TRYPTOPHANE-TRYPTOPHANEPROLINE INTERACTOR OF POLYCOMBS1 (PWO1), associates with CRWN1 in foci at the nuclear periphery and work together to maintain nuclear morphology as well as controlling expression of a group of target genes.

How might changes in the LINC complex and associated proteins as observed by Poulet et al [30] affect nuclear function? Two main hypotheses exist. The first hypothesis posits that changes in nuclear shape alter the rigidity of the nucleus; this could be beneficial for cells that need to squeeze through tight spaces, but deleterious to cells that are under mechanical duress. The second hypothesis proposes that changes in nuclear shape results in chromatin reorganization and thereby affects gene expression. It is important to note that

these two hypotheses are not mutually exclusive [27].

Nuclear movement

Plant nuclei are able to move in numerous cell types and circumstances, both as part of developmental processes or as a result of both biotic and abiotic stimulation [39]. Several cell models have been used to investigate nuclear movement in *Arabidopsis* such as root hair in which the nucleus migrates toward the tip as the root hair is growing [40]. KAKU1 a myosin XI-i homologue is involved in nuclear movement in root hairs. Such movements require the cytoskeleton especially actin filaments [32]. *wit1 wit2* mutant seedlings also impairs nuclear movement in root hair and Myosin XI-i was shown to interact with WIT1 and WIT2. Finally, WIT proteins are interacting with the KASH proteins WIP proteins anchored at the ONM [7,32,41]. This cascade of interaction nicely illustrates the connection between the nuclear envelope with the cytoskeleton and one of the key function of the LINC complex.

Another cell model is the pollen tube. In this case, during fertilization, the two sperm cell nuclei (SCN) and the vegetative nucleus (VN) migrate in the pollen tube to the unfertilized ovule. Zhou and Meier [42] demonstrated that WIT and WIP proteins are localized at the nuclear envelope of the VN and that they are essential for VN movement. Loss-of-function mutations in *WIT* and/or *WIP* gene families resulted in impaired VN movement that was no longer coupled with pollen tube tip growth. Zhou et al [33] went on to prove that the entire LINC complex was involved. Goto et al [43] recently show a role for the putative nuclear lamina protein KAKU4 in this process. KAKU4 is highly abundant in VNs and less so in SCNs. *kaku4* mutants show reversal in the order of migration of these two types of nuclei in elongating pollen tubes and reduced virility of the pollen.

Chromatin structure and gene expression

Certain proteins present at the nuclear periphery participate in the regulation of gene activity by modifying the organization of chromatin. Two different aspects can be considered depending on

whether broader chromatin domains or single gene transcription is considered.

Several studies have investigated the *A. thaliana* chromocentres which are heterochromatin domains enriched in silenced repeated sequences [44]. Chromocentres are of interest for this review as they are preferentially located at the nuclear periphery [30,45]. One of the most obvious impacts of the nuclear periphery on chromocentre organisation was observed with CRWN proteins. Fluorescent *in situ* Hybridization (FiSH) experiments clearly indicated that CRWNs are involved in cohesive forces within and between chromocentres as the double *crwn1 crwn2* mutant induces chromocentre fusion while more diffuse and decondensed chromocentres are observed in *crwn4* [23]. *crwn1* and *crwn4* single mutants were also studied by chromosome conformation capture (3C) and its whole genome application (Hi-C) [46–49]. *crwn1* and *crwn4* show an increased chromatin interactions between chromatin regions usually organised in different compartments and this is associated with an increased inter-chromosomal interactions. Finally, in a more general screen using 3D imaging investigating the effect of *crwn1*, *crwn4*, *kaku4*, *neap1* and *neap3* single mutants in *Arabidopsis*, Hu et al [49] suggested CRWN1 to be one of the key components maintaining specific chromatin domains at the nuclear periphery. Some effects were also recorded using the triple SUN mutant *sun1 sun4 sun5* in which chromocentres become more internal, being partially decondensed and leading to the derepression of some heterochromatic sequences [30].

As previously mentioned, CRWN1 interacts with PWO1, a component of the Polycomb-Group (PcG) associated factor. The PcG complex is a repressive complex and targeted genes are enriched in histone repressive marks such as H3K27me3. Interaction of PWO1 and CRWN1 in a transient expression assay revealed that this occurs to some extent near the nuclear periphery. Both proteins control expression of a similar set of target genes including those enriched in H3K27me3 [38,50]. Further studies show potential interaction of CRWN family members with transcription factors. CRWN1 interacts with NAC transcription factor NTL9 [51]. CRWN1 may also have the potential to directly interact with

chromatin as supported by the recent ChIP-Seq analysis using a CRWN1:2 HA line performed by Hu et al [49]. These sequences were named Plant Lamina-Associated Domains (PLADs), and nicely overlap with chromatin domains linked (or next or associated) to the nuclear periphery previously identified by Restriction Enzyme (RE)-ChIP [52]. Target sequences are mostly silent chromatin domains enriched in repressive chromatin marks, with low transcription and include some transposable elements which are hypothesized to be anchor points for CRWN1 chromatin binding at the nuclear periphery.

As described above, AtNEAP, another component of the nucleoskeleton, interacts with AtbZIP18, a transcription factor containing a DNA-binding BRLZ domain, a leucine zipper allowing bZIP dimerization and an Ethylene-responsive element binding factor-associated Amphiphilic Repression (EAR) motif which is implicated in transcriptional inhibition through chromatin modification [53,54]. The EAR motif has been shown to recruit histone deacetylase 19 (AtHDA19) leading to gene repression [54]. AtbZIP18 contributes to pollen development in which it could act as a repressor, as single mutant *Atbzip18* mostly induces upregulation of pollen expressed genes [53].

All the interactions described above suggest some interactions between the nuclear periphery and chromatin. The pioneer experiments performed by the group of Chang Liu with CRWN1 strongly support the existence of specific chromatin organization at the nuclear periphery in plants [49,52] reminiscent of the Lamin Associated Domains (LADs) well-described in animal genome.

Impact of the plant nuclear periphery on plant physiology

Seed maturation and germination

Maturation of seed is an essential step for the survival of many plant species. This developmental period is critical to prepare the seed to survive to desiccation but to produce the young seedling for the next-generation. Together with various external signals such as light and temperature,

phytohormones play a pivotal role in the regulation of seed germination. This regulation relies on the phytohormones abscisic acid (ABA) which promotes seed germination and gibberellin which inhibits seed germination by promoting seed maturation and dormancy. First evidence of a possible connection between the nucleoskeleton and the ABA pathway came from the investigation of the impact of nuclear size and chromatin organization on seed maturation and germination. ABA Insensitive 3 (ABI3) was shown to be required for the nuclear size reduction during seed maturation while CRWN1 and CRWN2 were required for the size increase during seed germination [55]. It is here interesting to note that one possible explanation of nuclear size regulation by CRWNs could be linked to their ability to induce membrane envelope overgrowth a phenomenon observed when CRWN are over-expressed in transient expression [24]. Membrane overgrowth is also observed when KAKU4 is over-expressed and this phenotype is even more pronounced when CRWN and KAKU4 are over-expressed together [24].

Further evidence of a connection between nuclear size and ABA pathway came from the investigation of the regulation of ABI5 degradation through the 26S proteasome pathway. ABI5 which is a key positive regulator in the ABA pathway is not only regulated by ABI3 but also by CRWN3 which is proposed to recruit ABI5 to specific nuclear body where it is degraded. This process mediated by the 26S proteasome pathway is dependent on the C-terminal of CRWN3 [56].

Reproduction and meiosis

Meiosis is a specialized cell division required for plant sexual reproduction. Through a single S phase and two rounds of cell division, gametes are produced with half the chromosome number. During this, the ordered movement of chromosome is essential to achieve the pairing of homologous chromosomes. Chromosome movement is driven by forces generated in the cytoplasm conveyed by NE proteins to the ends of chromosomes (*i.e.* at telomeres). Chromosome attachment to the NE requires tubulin and actin [57] and involves the LINC complex components SUN1 and SUN2

[58,59]. Studying pollen mother cells (PMCs) in prophase I demonstrated that AtSUN1 and AtSUN2 localized to the NE [58], in a pattern resembling telomere location and connected by a thread structure like those formed in yeast meiosis [60]. An *Atsun1 Atsun2* double mutant shows impaired meiosis. Polarized telomere location to the NE in leptoneuma does not take place, prophase I is delayed, there is incomplete synapsis and unresolved interlocks, univalents are observed at metaphase I and missegregations at anaphase I that lead to the formation of aneuploid gametes [58]. One interesting hypothesis is that the bouquet formation (i.e. clustering of telomeres at the INM) with telomere pairing may promote homologous chromosome synapsis, an important step toward recombination events to generate crossovers. Finally, components of the nuclear periphery not only affect meiosis but also fertilization as they are key components in nuclear migration during pollen tube growth as already mentioned in the nuclear movement section [42,43].

Organ development

The LINC complex is central to the processes of meiosis and mitosis which result in production of seeds and vegetative organs and therefore to the processes of cell and organ growth. The topic has been reviewed in detail recently [61]. Briefly, the behavior of the Cter-SUNs, SUN1 and SUN2 has been studied in mitotic division in synchronized tobacco cells [62]; in yeast and animals, they are known to be essential for centromere and centrosome anchorage and chromatin decondensation at the end of mitosis [63]. In *A. thaliana*, SUN1-YFP and SUN2-YFP fluorescence decreases and then increases in intensity at opposite sides of the NE as chromosomes condense. The NE then breaks down and is penetrated by spindle microtubules. SUN1 is then present in ER membranes around the mitotic spindle. As cells traverse anaphase, SUN1-YFP is present in tubule like structures around the segregated chromosomes. Both SUN1-YFP and SUN2-YFP are then located on decondensing chromatin facing the spindle poles before the membrane – containing the Cter-SUNs surround the chromatin. SUN1 and SUN2 aggregate first at the surface of chromatin facing the spindle

pole, then at the periphery of the spindle, and finally facing the cell plate, a spatial organization also observed for NMCP1 and NMCP2 in *Apium* [64]. The new cell wall is now formed by expansion of a membrane structure known as the phragmoplast, within which the cell plate grows centripetally. Other NE associated components also accumulate at the phragmoplast, including the Nucleoporin Rae1, Ran, RanGAP, WITs, and WIPs [65,66].

Response to stress

Plants are constantly facing changing environmental conditions such as biotic and abiotic stresses that can seriously affect their life cycle and productivity. In the past few years many reports have emerged on the contribution of the nuclear periphery to alter the plant immune system. As mentioned above, in *Arabidopsis* CRWN1 interacts with NAC WITH TRANSMEMBRANE MOTIF1-LIKE9 (NTL9). This interaction enhances NTL9 binding at PATHOGENESIS-RELATED1 (*PR1*) gene repressing *PR1* transcription, a key gene in plant defense mediated by the salicylic acid (SA) pathway. Therefore, CRWN1 is expected to be part of a repressive mechanism keeping silenced *PR1* [51]. This elegant analysis also delivered more mechanistic features of CRWN1 action. If CRWN1 transcription is induced by SA or upon bacterial infection, it is rapidly degraded by a proteasome-dependent pathway. Thus, reducing CRWN1 abundance contributes to *PR1* induction by transcription factors such as NONEXPRESSOR OF PR GENES1 (*NPR1*). *NPR1* is a major transcriptional activator of *PR1* which is imported through the NPC during biotic response. NPC and nucleoskeleton thus have complementary functions, CRWN maintaining the repression of *PR* genes through a possible alteration of chromatin marks [67] until signaling molecules such as *NPR1* transcription factor release this repression. Further analyses of transcription at a whole genome scale highlighted the induction of genes from the SA pathway in *crwn* mutants [51,67]. In an attempt to better understand the complex mutant phenotype observed in *crwn* mutants i.e. small plant phenotype, small nuclei, induction of SA, *crwn* mutants were combined with mutations affecting the SA pathway such as *NPR1* or *ISOCHORISMATE SYNTHASE1* (*ICS1*/

SID2). *crwn1 crwn2 npr1* [51] and *crwn1 crwn2 sid2* [67] compromised the small plant phenotype of *crwn1 crwn2* meaning that activation of SA response was indeed strongly contributing to the small plant phenotype of the *crwn* mutants. However, the small nuclei phenotype remained, suggesting that CRWN function regulation of nuclear morphology and SA-response can be disconnected [67]. Finally, recent evidence suggests that CRWN1 protein not only affect SA signaling but also jasmonic acid (JA) signaling in *Arabidopsis* [68] opening avenues to explore its impact on a more widespread pathogen response including not only bacteria but also fungal diseases.

Returning to the variation of nuclear size observed during seed maturation and germination, Van Zanten et al [55] suggested that the decreased nuclear size could represent a universal mechanism in acquisition of desiccation tolerance. However the possible dehydration tolerance of CRWN mutants which display a reduced nuclear size has not been tested. The authors measured nuclear size of leaves of the desiccation-tolerant resurrection plant *Craterostigma plantagineum* in normal and dehydrated conditions and found a reduced nuclear size in stressed condition suggesting that alteration of nuclear size can also occur in leaves. In *A. thaliana*, SUN3 might play an important role in stress tolerance by modulating UPR signaling, possibly via ABA-independent pathways. The wild-type and the mutant plants were found to be compromised in their tolerance to dehydration as indicated by increased transpirational water loss [69]. Finally, CRWN-family proteins were recently shown to protect against Reactive Oxygen Species (ROS) accumulation and against DNA damage [70].

All together, these reports suggest that the plant nuclear periphery acts as a sensor to external stresses especially in biotic stress response.

LINC complex and associated components in crop species

Conservation in monocot and dicot cultivated species

Decades of breeding have improved yield and the ability for plants to better respond to various stresses. Plant breeding is a constantly evolving strategy and from the green revolution in the 1960–70's to the current genomic selection era,

major progress has been made to enhance plant growth and genetic pools. Understanding plant physiology and its adaptation to changing environments is timely and knowledge transfer between model and cultivated plants has contributed and still offers good opportunities to identify new ways of improvement to reach this major challenge for the future. Because genetics by its own cannot explain all the phenotypic plasticity of plants, epigenetics is often cited for its 'strong potential for crop breeding' [71,72].

The nuclear periphery, as one layer of this complex regulation, described in the previous sections, actively participates in the transmission of external signals through the nuclear envelope to alter chromatin organization (*i.e.* epigenetics). However, investigation of plant LINC complexes and associated proteins in crop species remains sparse although many proteins can be identified by Blastp analysis or by exploring databases (Table 1, Figure 1).

Similarities and differences between *Arabidopsis* and crop species

Significant differences are observed among plant genomes regarding gene copy number which is not surprising considering the plasticity of plant genome size [73]. Plant genomes were subjected to several rounds of Whole Genome Duplications (WGD) such as the ρ - σ - τ WGD and the α - β - γ WGD series respectively in monocot and dicot lineages [74]. WGDs partially explain the high variability in paralogue numbers. For instance *Arabidopsis* would theoretically contain up to 3 (γ) x 2 (β) x 2 (α) (*i.e.* 12 expected paralogues) paralogues of the ancestral genes found in the basal angiosperm *Amborella trichopoda* if these had not been eliminated during evolution while species such as *Solanum lycopersicum* or *Vitis vinifera* will have been subjected only to the γ triplication (*i.e.* 3 expected paralogues) [75]. Thus, one can expect that the number of copies would be lower in these species as observed in Table 1 and Figure 1. Polyploid genomes such as wheat (allohexaploid species) contains not only paralogues but also three homeologues. This also explains the higher number of copies identified in wheat *versus* barley, rice and maize that are all diploid species. Finally it is also important to note that *Arabidopsis* with its



Table 1. LINC complex and associated components in monocot and dicot lineages.

Species	LINC complex										Nuclear periphery	
	KASH proteins (ONM)					SUN proteins (INM)					Nucleoskeleton	
	KASH TIK	KASH SINE	KASH WIP	KASH WIT	Other KASH	Cter-SUN	Mid-SUN	NEAP	CRWN/MMCP	KAKU4		
<i>A. thaliana</i> (TAIR+ Poulet et al., 2016)	ARTIK (AT5G44920)	ATSINET1 (AT1G54385)	ARWIP1 (AT4G26465)	ARWIT1 (AT5G11390)	MRKASH1 (Mt3g0990660)	MSUN1 (Mt8g043510)	MSUN3 (AT1G22882)	ARNEAP1 (AT3G05830)	ACRWN1 (AT1G67230)	AKAKU4 (AT4G31430)		
		ATSINET2 (AT3G03970)	ARWIP2 (AT5G56210)	ARWIT2 (ATTG68910)	MRKASH4 (Mt4g036225)		MSUN4 (AT1G71360)	ARNEAP2 (AT5G26770)	ACRWN2 (AT1G13220)			
		ATSINET3 (AT3G06600)	ARWIP3 (AT3G13360)		MRKASH5 (Mt6g016290)		MSUN5 (AT4G23950)	ARNEAP3 (AT1G09470)	ACRWN3 (AT1G66790)			
		ATSINET4 (AT4G24950)			MRKASH6 (Mt1g052620)			ARNEAP4 (AT1G09483)	ACRWN4 (AT1G66790)			
<i>M. truncatula</i> (Newman-Griffis et al., 2019 + PLAZA Dicots 4.5)		MSINET1 (Mt6g032885)	MRWIP1a (Mt6g070940)	MRWIT1 (Mt7g017690)	MRKASH1 (Mt3g0990660)	MSUN1 (Mt8g043510)	MSUN4-a (Mt3g006320)	MtNEAP2 (Mt1g484830)	MtCRWN1 (Mt7g018610)	MtKAKU4 (Mt3g108780)		
		MSINESa (Mt2g033900)	MRWIP1b (Mt3g009740)	MRWIT2 (Mt3g086660)	MRKASH4 (Mt4g036225)		MSUN4-b (Mt1g069420)	444/1e-155	859/0.0	1977/e-53		
		MSINESb (Mt5g054260)			MRKASH5 (Mt6g016290)		MSUN4-c (Mt4g030960)		MtCRWN2 (Mt6g015285)			
					MRKASH6 (Mt1g052620)		MSUN4-d (Mt4g030960)		588/0.0			
<i>C. arvensis</i> (PLAZA Dicots 4.5)		CaSINET1-a (Ca_13211.g)	CaWIPI1-a (Ca_01362.g)	CaWIT1 (Ca_16511.g)	CaWIPI1-a (Ca_01362.g)	CaCter-SUN (Ca_15702.g)	CaMid-SUN3 (Ca_21229.g)	CaNEAP2 (Ca_03601.g)	CaCRWN2 (Ca_10456.g)	CaKAKU4 (Ca_01833.g)		
		534/0.0	87.4/5e-18	369/4e-117	CaWIPI2 (Ca_15964.g)	441/1e-150	421/3e-140	373/4e-128	252/2e-67	182/2e-48		
		CaSINET1-b (Ca_15536.g)	CaWIPI1-b (Ca_20564.g)	CaWIT2 (Ca_15964.g)	CaWIPI2 (Ca_15964.g)		427/3e-142		644/0.0			
		517/2e-177	85.1/4e-17	421/1e-139					CaCRWN4 (Ca_09369.g)			
<i>V. vinifera</i> (Poulet et al., 2016 + PLAZA Dicots 4.5)		CaSINET2 (Ca_15305.g)	CaWIPI1-c (Ca_17833.g)	VvWIT1 (GSVVG01036512001)	VvWIT1 (GSVVG01036512001)	VvCter-SUN (GSVVG01001935001)	VvSUN3/4 (GSVVG01037194001)	VvNEAP2 (GSVVG01031318001)	VvCRWN1 (GSVVG01031076001)	VvKAKU4 (GSVVG01035528001)		
		62.4/2e-08	83.6/4e-17	517/4e-174	VvWIT2 (GSVVG01010127001)	441/1e-150	VvSUN5 (GSVVG01019538001)	VvNEAP3 (GSVVG01028212001)	VvCRWN3 (GSVVG0101972001)			
			CaWIPI2-b (Ca_20519.g)	479/4e-160								
<i>S. lycopersicum</i> (Poulet et al., 2016 + PLAZA Dicots 4.5)		SiSINET1/2 (Solyco3g062700.2)	SiWIP (Solyco2g066170.2)	SiWIT1 (Solyco2g065660.2)	SiWIP (Solyco2g066170.2)	SiCter-SUN (Solyco01g096780.2)	SiSUN5 (Solyco08g082540.2)	SiNEAP (Solyco01g101130.2)	SiCRWN1-a (Solyco03g045050.2)	SiKAKU4 (Solyco08g006480.2)		
				383/1e-122	SiWIT2 (Solyco5g010590.2)							
					SiWIT2 (Solyco5g010590.2)							
				395/6e-128								

(Continued)

Table 1. (Continued).

Species	Nuclear Envelope										Nuclear periphery		
	LINC complex										Nucleoskeleton		
	KASH proteins (ONM)					Other KASH					INM		SUN proteins (INM)
<i>Z. mays</i> (Gumber et al., 2019)	LINC KASH grass-specific	KASH TIK	KASH SINE	KASH WIP	KASH WIT	Other KASH	Cter-SUN	Mid-SUN	NEAP	CRWN/NNMCP	KAKU4		
	MLKG1 (Zm00001d002723)	MLKS1 (Zm00001d031134)	MLKP1 (Zm00001d003334)	MLKT1 (Zm00001d039272)	ZmSUN1 (Zm00001d015450)	ZmSUN3 (Zm00001d042643)	MNEAP1 (Zm00001d010918)	NCH1 (Zm00001d051600)	IMKAKU41 (Zm00001d026487)				
	MLKG2 (Zm00001d038539)	MLKS2 (Zm00001d052955)	MLKP2 (Zm00001d025667)	MLKT2 (Zm00001d010047)	ZmSUN2 (Zm00001d040331)	ZmSUN4 (Zm00001d012240)	MNEAP2 (Zm00001d024756)	NCH2 (Zm00001d043335)	IMKAKU2 (Zm00001d002012)				
		MLKP3 (Zm00001d005997)	MLKP4 (Zm00001d020879)			ZmSUN5 (Zm00001d011277)	MNEAP3 (Zm00001d002875)						
<i>S. bicolor</i> (PLAZA Monocots 4.5)	SbMLKG1 (Sb06g024760/ Sobic.006G172400)	SbSINE1-a (Sb05g022560/ Sobic.005G164900)	SbWIP1-a (Sb02g027920/ Sobic.002G242200)	SbWIT1-a (Sb09g08110/ Sobic.009G088000)	SbSUN1 (Sb04g005160/ Sobic.004G061900)	SbSUN4-a (Sb03g041510/ Sobic.003G374700)	SbNEAP2-a (Sb09g002860/ Sobic.009G033800)	Sb CRWN1 (Sb04g030240/ Sobic.004G264300)	SbKAKU4 (Sb06g031220/ Sobic.006G246900)				
	SbMLKG2-a (Sb09g023990/ Sobic.009G179500)	SbSINE1-b (Sb08g022270/ Sobic.008G179600)	SbWIP1-b (Sb07g028540/ Sobic.007G218900)	SbWIT1-b (Sb03g000580/ Sobic.003G004600)	SbSUN2 (Sb03g010590/ Sobic.003G125332)	SbSUN4-b (Sb03g026980/ Sobic.003G208900)	SbNEAP2-b (Sb03g011900/ Sobic.003G139700)	Sb CRWN4 (446/3e-136 Sobic.003G308200)	SbKAKU4 (65.9/1e-09)				
	SbMLKG2-b (282/2e-95 Sobic.004G352400)	SbWIP3 (77.4/2e-13 Sb06g019690/ Sobic.006G115466)											
<i>O. sativa</i> (Poulet et al., 2016 + PLAZA Monocots 4.5)	OsMLKG1 (Os04g0554300/ LOC_Os04g46790)	OsSINE1 (Os11g058000/ LOC_Os11g37100)	OsWIP1-a (Os04g0471300/ LOC_Os04g39540)	OsWIT1 (Os05g0241000/ LOC_Os05g15140)	OsSUN1/SAD1 (Os05g027020/ LOC_Os05g18770)	OsSUN3 (Os01g0876400/ LOC_Os01g65520)	OsNEAP1 (Os05g0135900/ LOC_Os05g04530)	OsCRWN1 (Os02g0710102/ LOC_Os02g48010)	OsKAKU4 (Os04g0655600/ LOC_Os04g56140)				
	OsMLKG2 (Os02g0824600/ LOC_Os02g57850)	OsSINE2 (Os12g0624800/ LOC_Os12g42960)	OsWIP1-b (Os08g0497900/ LOC_Os08g38890)	OsWIT1-b (314/1e-95)	OsSUN2 (Os01g0267600/ LOC_Os01g16220)	OsSUN5 (Os01g0599900/ LOC_Os01g41600)	OsNEAP2 (Os01g0292700/ LOC_Os01g18840)	OsCRWN4 (Os01g0767000/ LOC_Os01g56140)					
	HvMLKG1 (HORVU2H1G097210)	HvSINE1-a (HORVU6H1G028190)	HvWIP2-a (HORVU5H1G071910)	HvWIT1-b (HORVU7H1G007940)	HvSUN1-a (HORVU1H1G019590)	HvSUN3 (HORVU3H1G081880)	HvNEAP1 (HORVU6H1G069560)	HvCRWN1 (HORVU6H1G069560)					
	HvMLKG2 (HORVU2H1G110520)	HvSINE1-b (HORVU5H1G006220)	HvWIP2-b (HORVU2H1G081800)	HvWIT1-a (HORVU3H1G057330)	HvSUN1-b (HORVU3H1G037960)	HvSUN5 (HORVU3H1G052360)	HvNEAP2 (HORVU3H1G074480)	HvCRWN4 (HORVU3H1G074480)					
		HvSINE1-c (HORVU5H1G006040)	HvWIP3 (HORVU7H1G048940)										

(Continued)

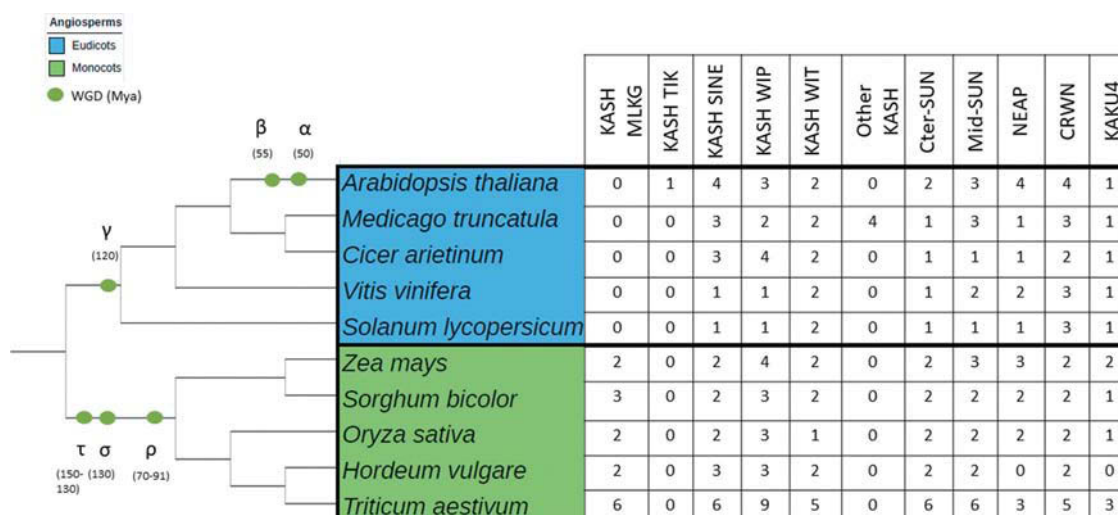


Figure 1. Number of Paralogues of LINC complex and associated components in monocot and dicot lineages.

rather small genome of about 150Mb contains a limited number of transposable elements (TEs, ~10% of the genome) mainly located at pericentromeric regions, while maize (3Gb genome) and wheat (~16Gb genome) have high number of TEs (up to 80% of their genome) dispersed along their chromosomes [73]. Regarding these genome features, *Arabidopsis* may then appear as an outlier among plant species. But to go further in these differences before going back to the observed functional similarities, genome size is not the only difference between *Arabidopsis* and crop genomes as chromosomes may adopt distinct organization in the nucleus. Some like yeast or wheat display a Rabl-like organization with telomeres and centromeres at opposite sides of the nucleus while others such as rice, maize, Sorghum and *Arabidopsis* do not [76--76-80]. Telomeres are also at various positions, being close to heterochromatin regions in tomato, dispersed around the nuclear periphery in wheat and rice and tightly associated with the nucleolus in *Arabidopsis* [79].

Crop species as model to study the impact of LINC complex and associated components in plant traits

Despite the clear differences that should also be kept in mind, *Arabidopsis*, maize and rice chromosome telomeres all form a bouquet structure and interact with the nuclear periphery during meiotic prophase 1 [58,81,82]. In all three cases, the SUN family is

involved. Following the simple concept of looking for functional similarities, reports about LINC complex and associated components have been collected from a small number of cereals and legume species as referred in Table 1 and Figure 1.

Cereals species

Rice (*Oriza sativa*) was the first plant species in which a SUN (SAD) protein was described, localized to NE and ER [83]. Recently, Shah et al [84] have studied the rice SUN domain proteins in depth and observe that all the rice SUN genes respond to abiotic stress (drought, salinity, temperature). Zhang et al [82,85] further studied the behavior of rice SUN domain proteins in meiosis. Using an *Ossun1*, *Ossun2* double mutant, Zhang et al [85] observed severe defects in telomere clustering, homologous pairing and crossover formation, indicating an essential role for both Cter-SUNs. Further analysis of the mutants, in a line in which topoisomerase initiated homologous recombination was disrupted, suggested that OsSUN1 and OsSUN2 are not completely redundant in function as the *Ossun1* single mutant had a normal phenotype but meiosis was disrupted in the *Ossun2* mutant. Finally, Yang et al [86] have recently described the interaction between OsNMCP1 and *Oriza sativa* SWITCH/SUCROSE NONFERMENTING (SWI/SNF) silencing complex. This interaction occurs between OsNMCP1 and OsSWI3C under drought stress conditions leads to the release of OsSWI3C from the SWI/SNF gene silencing complex, thus changing chromatin

accessibility in the genes related to root growth and drought resistance.

Maize (*Zea mays*) SUN domain proteins were classified in CCSD (canonical C-terminal SUN domain) and PM3-type (plant-prevalent Mid-SUN 3) [5]. Five genes for SUN domain proteins have been identified [5]. ZmSUN1 and ZmSUN2 are Cter-SUNs, while ZmSUNs 3–5 are mid-SUNs (Table 1). Murphy and Bass [59] studied the role of SUN domain proteins in maize meiosis. Maize forms a bouquet in zygotene. SUN2 is concentrated into a belt-like structure at the nuclear periphery which reorganizes to a half-belt in zygotene and returns to a full belt in pachytene. The half belt is associated with the cluster of telomeres in the bouquet and coincides with synapsis of homologous chromosomes. The half belt location of the SUN domain proteins is disrupted in the meiosis-specific mutants, *desynaptic* (*dy1*), *asynaptic1* (*as1*), and *divergent spindle1* (*dv1*). Based on the maize model, Bass proposed that first telomeres move to the NE, cluster together during the pre-zygotene-zygotene phase and connect to the microtubule system to move the nucleus within the cell and allow proper meiotic recombination. For these reasons, the LINC complex, which applies forces needed to tether telomeres to the nuclear envelope, has a major contribution in the meiotic process [81].

Further evidence for the functional components of the LINC complex, CRWN, KAKU4 and NEAPs in maize have been obtained in studies led by the Bass laboratory [12]. Using a combination of bioinformatics and cell biology, 22 genes have been identified and characterized that interact with SUN2 including the novel graminaceous KASH family (MLKG 1–2) (Table 1). MLKG1 interacts with SUN and is localized to the NE and ER actin network. Interaction with ZmSUN2 *in vivo* has been shown for MLKP2 and MLKG1; however, unlike most KASH domain proteins where the KASH domain is essential for NE localization, KASH deletion constructs remain associated with the nuclear periphery, suggesting other interactions that associate them with NE. More functional validation are now expected in maize thanks to this pioneer study which will strongly benefit the previous *Arabidopsis* studies. To our knowledge, there were no reports in other important species such as barley, wheat and sorghum.

Legume species

Medicago truncatula is a model legume used to explore the processes of nodulation. Zhou et al [9] and Newman-Griffis et al [87] explored its LINC complex components using a combination of bioinformatics, protein protein interaction and mutant analysis. They identified one *Medicago* SUN domain protein (MtSUN), and putative homologs of the *Arabidopsis* KASH proteins WIP1 (MtWIP1a, MtWIP1b) SINE1 (MtSINE1) and a further SINE family member, named SINE5 by Zhou et al (2014), MtSINE5a and MtSINE5b (Table 1). They also described the WIT1 and WIT2 homologues MtWIT1 and MtWIT2 and described a role for LINC complex components in root hair responses to environment that are involved in initiation of the symbiotic formation of the nodule. *Medicago* also contains a number of other KASH genes not attributable to known families designated MtKASH1, MtKASH4, MtKASH5 and MtKASH6 (Table 1, Figure 1). The additional complexity of the legume LINC family may be associated with its role in nodulation; however, homologues of these additional KASH genes were not detected in *Cicer arietinum*, the Chickpea, the second main cultivated grain legume (Table 1, Figure 1). SUN proteins were also investigated in Chickpea. Only one Cter-SUN (CaCterSUN) and one mid-SUN (CaSUN1) was detected in Chickpea while one homologue of CRWN1-3 (CaCRWN2) and one homologue of CRWN4 (CaCRWN4) was found (Table 1, Figure 1). One gene for KAKU4 (CaKAKU4) is also present. CaSUN1 is similar to AtSUN3 and was identified by mass spectrometry as one out of the 280 genes downregulated under dehydration [69].

Concluding remarks

This review focuses on advances in knowledge of the physical structure of the nucleus and the mechanisms by which nuclear dynamics might in the future be manipulated to enhance productivity and response to environmental factors. Increasingly, crop scientists are looking beyond classical genetics and genomics to consider epigenetic effects- modification to gene expression without changing the genetic code. Recent advances show that some of these epigenetic determinants rely on the structure

of the nucleus, the position and movement of genes as well as of the nucleus itself which are all-important in controlling the activity of the genome. As described in the first part of this review, the LINC complex and associated components regulate key functions in various aspects of the plant cycle (Figure 2). Most of them are conserved in a wide panel of plant species while some are Brassicaceae (KASH TIK) or monocot (KASH MLKG) specific (Figure 3). We have no doubt that more reports from cultivated species will come out in the coming years. However, application in plant improvement will have to face the pleiotropic effects of these genes and their impact on plant growth. Yield can be severely impaired in some of the mutants as observed in the constitutive activation of SA signaling pathway in *crwn* mutants inducing spontaneous expression of chlorotic lesions in the absence of pathogens due to hypersensitive response (HR) [67]. This class of mutation was well described by

maize geneticists almost 40 years ago as lesion-mimic mutants [88] and was considered as a model system to understand plant cell death in plants. Will this inconvenience prevent any application in agronomy? One can also expect that during evolution, each paralogue even each homeologue will diverged from its original ancestral homologue and acquired specific expression profiles and functions. As some of these genes are key players in nuclear/cell/organ size, meiosis, seed germination and stress responses, we propose there should be targeted research aimed at their functional characterization in crop species. Crops benefit from many genetic resources including TILLING populations, germplasm collections genome editing or advanced crossing populations [89] that will strongly benefit to the functional study of the nuclear periphery components.

A simple phylogenetic tree according to the Tree of Life [90] using 10 plant species is proposed on the left side of the figure. The proposed tree starts with

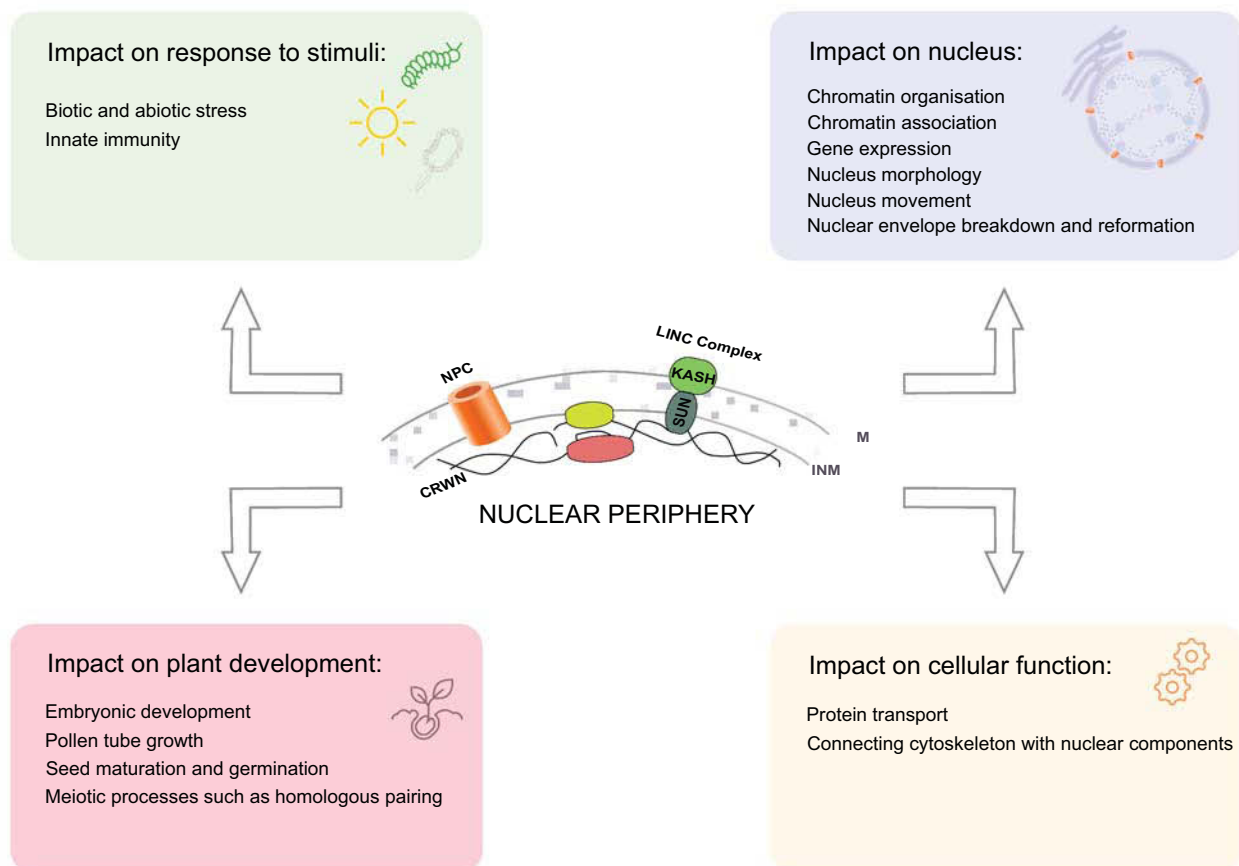


Figure 2. Schematic representation summarizing key functions of nuclear periphery components in plants.

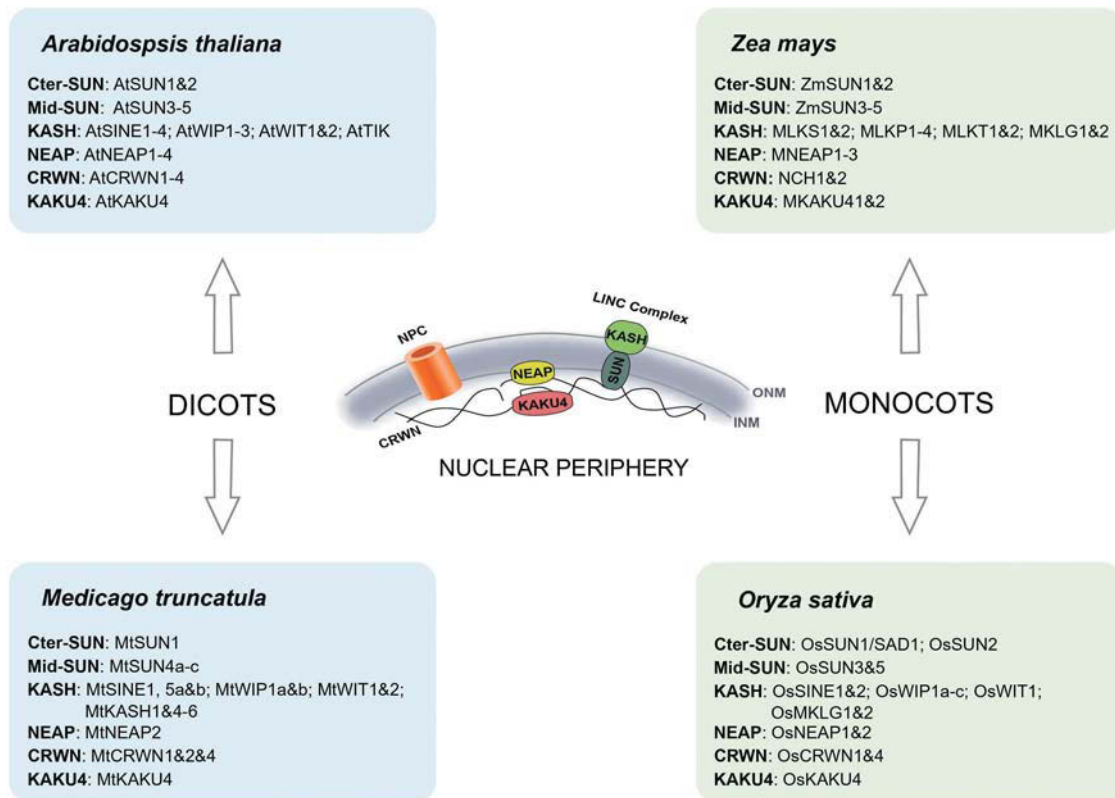


Figure 3. Schematic comparison of nuclear periphery components for monocot and dicot crops and the model species *Arabidopsis thaliana*.

angiosperms divergence (~170 Mya) and include the main whole Genome Duplications (WGD) relative to the dicot (blue) and monocot (green) lineages indicated as green circles on tree branches. Timing estimation of WGD events expressed in Million years (Mya) are indicated in between brackets [74,75,91]. τ is anterior to ρ and close to the origin of monocots is not completely resolved [91]. Number of paralogues is given on the right for each species for 11 gene families.

Species used in this table: *A. thaliana*: *Arabidopsis thaliana*, *M. truncatula*: *Medicago truncatula*, *C. arietinum*: *Cicer arietinum*, *V. vinifera*: *Vitis vinifera*, *S. lycopersicum*: *Solanum lycopersicum*, *Z. mays*: *Zea mays*, *S. bicolor*: *Sorghum bicolor*, *O. sativa*: *Oryza sativa ssp. Japonica*, *H. vulgare*: *Hordeum vulgare*, *T. aestivum*: *Triticum aestivum*

This table was produced from data found in the literature as well as in the databases available on PLAZA [92] using the Monocots PLAZA 4.5 for *T. aestivum*, *H. vulgare*, *O. sativa* and *S. bicolor* and Dicots PLAZA 4.5 for *S. lycopersicum*, *V. vinifera*, *C. arietinum* and *M. truncatula*. The protein sequences of *A. thaliana* or *Z. mays* (for MLKG proteins) were used as query sequence to perform

Blastp on PLAZA with a cut off of $1e-05$. The score in bits and the e-value of results of Blastp are written after the name of corresponding gene. The proteins noted in color are those found in the literature while the others are those found from databases. The different colors represent the different publications used in the table: blue: Poulet et al., 2016 [1], orange: Newman-Griffis et al., 2019 [87] and green: Gumber et al., 2019 [12].

Disclosure statement

No potential conflicts of interest were disclosed.

Funding

The work was supported by: CNRS, INSERM, Clermont Auvergne and Oxford Brookes Universities, 16-IDEX-0001 CAP 20-25 (WoW), CNRS delegation (CTa), COST-Action INDEPTH (CA16212) European Cooperation in Science and Technology [CA16212];

ORCID

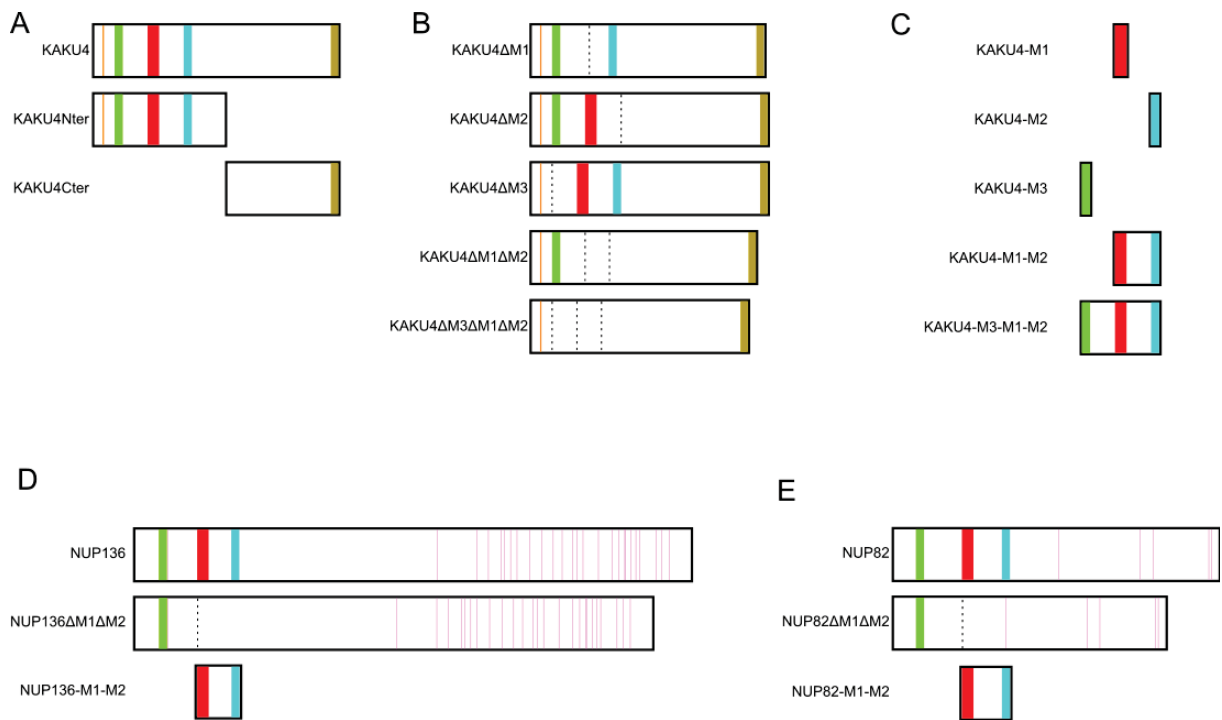
David E. Evans  <http://orcid.org/0000-0001-6248-1899>

References

- [1] Poulet A, Probst AV, Graumann K, et al. Exploring the evolution of the proteins of the plant nuclear envelope. *Nucleus*. 2017;19:1–14.
- [2] Meier I, Richards EJ, Evans DE. Cell biology of the plant nucleus. *Annu Rev Plant Biol*. 2017;68:139–172.
- [3] Tang Y, Huang A, Gu Y. Global profiling of plant nuclear membrane proteome in *Arabidopsis*. *Nat Plants*. 2020;6:838–847.
- [4] Tamura K. Nuclear pore complex-mediated gene expression in *Arabidopsis thaliana*. *J Plant Res* [Internet]. 2020 [cited 2020 Mar 18];133:449–455. Available from <http://link.springer.com/10.1007/s10265-020-01177-0>
- [5] Murphy SP, Simmons CR, Bass HW. Structure and expression of the maize (*Zea mays* L.) SUN-domain protein gene family: evidence for the existence of two divergent classes of SUN proteins in plants. *BMC Plant Biol*. 2010;10:269.
- [6] Graumann K, Vanrobays E, Tutois S, et al. Characterization of two distinct subfamilies of SUN-domain proteins in *Arabidopsis* and their interactions with the novel KASH-domain protein AtTIK. *J Exp Bot*. 2014;65:6499–6512.
- [7] Zhou X, Graumann K, Evans DE, et al. Novel plant SUN–KASH bridges are involved in RanGAP anchoring and nuclear shape determination. *J Cell Biol*. 2012;196:203–211.
- [8] Zhou X, Meier I. How plants LINC the SUN to KASH. *Nucleus*. 2013;4:206–215.
- [9] Zhou X, Graumann K, Wirthmueller L, et al. Identification of unique SUN-interacting nuclear envelope proteins with diverse functions in plants. *J Cell Biol*. 2014;205:677–692.
- [10] Coates JC. Armadillo repeat proteins: beyond the animal kingdom. *Trends Cell Biol*. 2003;13:463–471.
- [11] Biel A, Moser M, Meier IA. Role for plant KASH proteins in regulating stomatal dynamics. *Plant Physiol*. 2020;182:1100–1113.
- [12] Gumber HK, McKenna JF, Estrada AL, et al. Identification and characterization of genes encoding the nuclear envelope LINC complex in the monocot species *Zea mays*. *J Cell Sci*. 2019;132:1–16.
- [13] Aebi U, Cohn J, Buhle L, et al. The nuclear lamina is a meshwork of intermediate-type filaments. *Nature*. 1986;323:560–564.
- [14] Herrmann H. Intermediate Filaments: AU. Structure and Assembly. *Cold Spring Harb Perspect Biol*. 2016;8:a018242.
- [15] Dechat T, Pflieger K, Sengupta K, et al. Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. *Genes Dev*. 2008;22:832–853.
- [16] Méjat A, Misteli T. LINC complexes in health and disease. *Nucleus*. 2010;1:40–52.
- [17] Fiserova J, Kiseleva E, Goldberg MW. Nuclear envelope and nuclear pore complex structure and organization in tobacco BY-2 cells. *Plant J*. 2009;59:243–255.
- [18] Gardiner J, Overall R, Marc J. Putative *Arabidopsis* homologues of metazoan coiled-coil cytoskeletal proteins. *Cell Biol Int*. 2011;35:767–774.
- [19] Dittmer TA, Stacey NJ, Sugimoto-Shirasu K, et al. LITTLE NUCLEI genes affecting nuclear morphology in *Arabidopsis thaliana*. *Plant Cell*. 2007;19:2793–2803.
- [20] Masuda K, Xu ZJ, Takahashi S, et al. Peripheral framework of carrot cell nucleus contains a novel protein predicted to exhibit a long alpha-helical domain. *Exp Cell Res*. 1997;232:173–181.
- [21] Sakamoto Y, Takagi S. LITTLE NUCLEI 1 and 4 regulate nuclear morphology in *Arabidopsis thaliana*. *Plant Cell Physiol*. 2013;54:622–633.
- [22] Graumann K, Bassham D. Evidence for LINC1-SUN associations at the plant nuclear periphery. *PLoS ONE*. 2014;9:e93406.
- [23] Wang H, Dittmer TA, Richards EJ. *Arabidopsis* CROWDED NUCLEI (CRWN) proteins are required for nuclear size control and heterochromatin organization. *BMC Plant Biol*. 2013;13:200.
- [24] Goto C, Tamura K, Fukao Y, et al. Nuclear envelope protein KAKU4 modulates nuclear morphology in *Arabidopsis*. *Plant Cell*. 2014;26:2143–2155.
- [25] Pawar V, Poulet A, Détourné G, et al. A novel family of plant nuclear envelope-associated proteins. *J Exp Bot*. 2016;67:5699–5710.
- [26] Webster M, Witkin KL, Cohen-Fix O. Sizing up the nucleus: nuclear shape, size and nuclear-envelope assembly. *J Cell Sci*. 2009;122:1477–1486.
- [27] Skinner BM, Johnson EEP. Nuclear morphologies: their diversity and functional relevance. *Chromosoma*. 2017;126:195–212.
- [28] Zink D, Fischer AH, Nickerson JA. Nuclear structure in cancer cells. *Nat Rev Cancer*. 2004;4:677–687.
- [29] Chytilova E, Macas J, Sliwinska E, et al. Nuclear dynamics in *Arabidopsis thaliana*. *Mol Biol Cell*. 2000;11:2733–2741.
- [30] Poulet A, Duc C, Voisin M, et al. LINC complex contributes to heterochromatin organisation and transcriptional gene silencing in plants. *J Cell Sci*. 2017;130:590–601.
- [31] Masoud K, Herzog E, M-E C, et al. Microtubule nucleation and establishment of the mitotic spindle in vascular plant cells. *Plant J*. 2013;75:245–257.
- [32] Tamura K, Iwabuchi K, Fukao Y, et al. XI-i links the nuclear membrane to the cytoskeleton to control nuclear movement and shape in *Arabidopsis*. *Curr Biol*. 2013;23:1776–1781.
- [33] Zhou X, Groves NR, Meier I. Plant nuclear shape is independently determined by the SUN-WIP-WIT2-myosin XI-i complex and CRWN1. *Nucleus*. 2015;6:144–153.

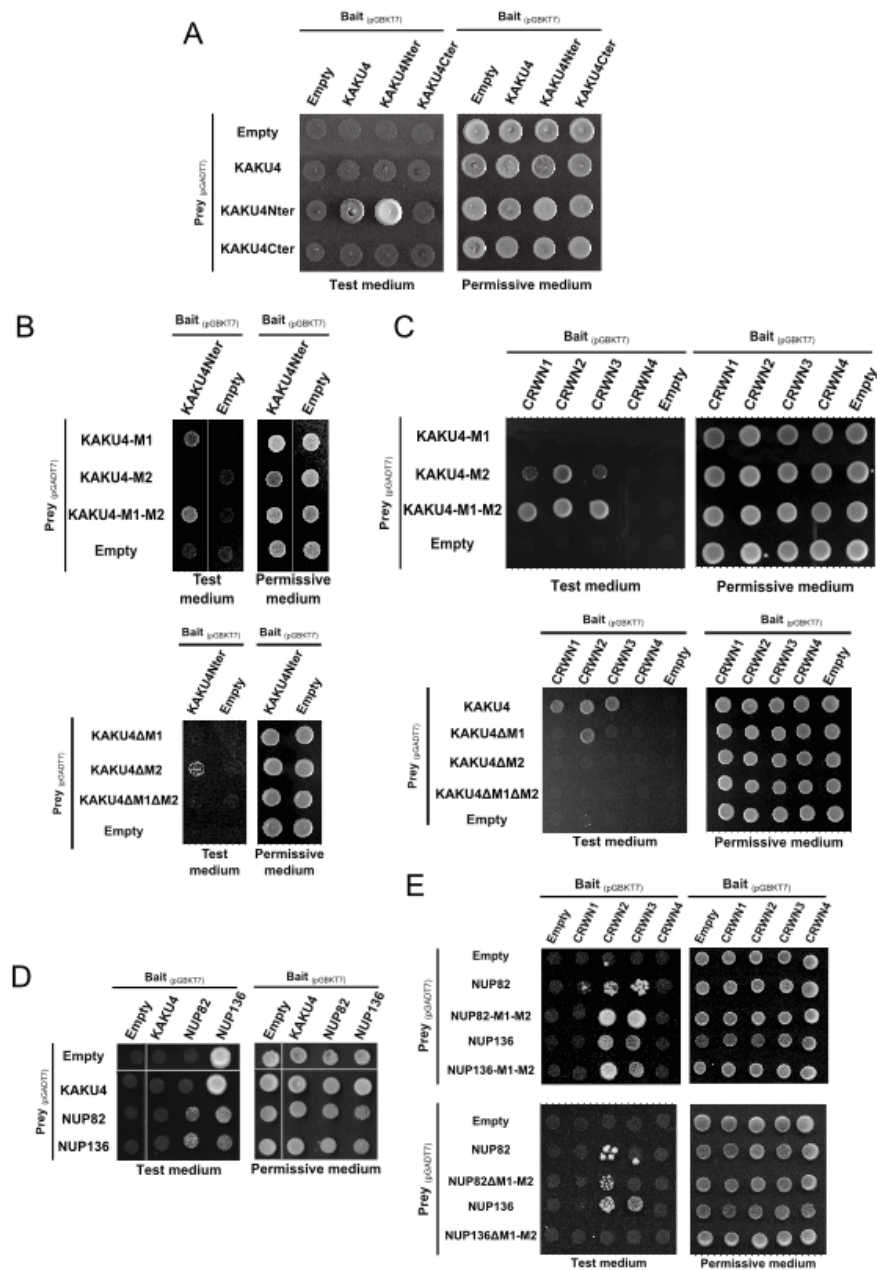
- [34] Tamura K, Hara-Nishimura I. Involvement of the nuclear pore complex in morphology of the plant nucleus. *Nucleus*. 2011;2:168–172.
- [35] Batzenschlager M, Lermontova I, Schubert V, et al. Arabidopsis MZT1 homologs GIP1 and GIP2 are essential for centromere architecture. *PNAS*. 2015;112:8656–8660.
- [36] Janski N, Masoud K, Batzenschlager M, et al. The GCP3-interacting proteins GIP1 and GIP2 are required for γ -tubulin complex protein localization, spindle integrity, and chromosomal stability. *Plant Cell*. 2012;24:1171–1187.
- [37] Melcer S, Gruenbaum Y. Nuclear morphology: when round kernels do the charleston. *Curr Biol*. 2006;16:R195–7.
- [38] Mikulski P, Hohenstatt ML, Farrona S, et al. Protein PWO1 interacts with plant nuclear lamin-like components to regulate nuclear size. *Plant Cell*. 2019;31:1141–1154.
- [39] Griffis AHN, Groves NR, Zhou X, et al. Nuclei in motion: movement and positioning of plant nuclei in development, signaling, symbiosis, and disease. *Front Plant Sci* [Internet]. 2014 [cited 2017 Mar 7];5. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3982112/>
- [40] Ketelaar T, Faivre-Moskalenko C, Esseling JJ, Ruijter NCA de, Grierson CS, Dogterom M, Emons AMC. Positioning of nuclei in arabidopsis root hairs an actin-regulated process of tip growth. *Plant Cell*. 2002;14:2941–2955.
- [41] Zhao Q, Brkljacic J, Meier I. Two distinct interacting classes of nuclear envelope-associated coiled-coil proteins are required for the tissue-specific nuclear envelope targeting of Arabidopsis RanGAP. *Plant Cell*. 2008;20:1639–1651.
- [42] Zhou X, Meier I. Efficient plant male fertility depends on vegetative nuclear movement mediated by two families of plant outer nuclear membrane proteins. *Proc National Acad Sci*. 2014; 111:11900–11905.
- [43] Goto C, Tamura K, Nishimaki S, et al. KAKU4-mediated deformation of the vegetative nucleus controls its precedent migration over sperm cells in pollen tubes. 2020;71:6273–6281.
- [44] Fransz P, De Jong JH, Lysak M, et al. Interphase chromosomes in Arabidopsis are organized as well defined chromocenters from which euchromatin loops emanate. *Proc Natl Acad Sci USA*. 2002;99:14584–14589.
- [45] Fang Y, Spector DL. Centromere positioning and dynamics in living Arabidopsis plants. *Mol Biol Cell*. 2005;16:5710–5718.
- [46] Feng S, Cokus SJ, Schubert V, et al. Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in Arabidopsis. *Mol Cell*. 2014;55:694–707.
- [47] Grob S, Schmid MW, Hi-C GU. Analysis in Arabidopsis identifies the KNOT, a structure with similarities to the flamenco locus of Drosophila. *Mol Cell*. 2014;55:678–693.
- [48] Liu C, Cheng Y-J, Wang J-W WD. Prominent topologically associated domains differentiate global chromatin packing in rice from Arabidopsis. *Nat Plants*. 2017;3:742–748.
- [49] Hu B, Wang N, Bi X, et al. Plant lamin-like proteins mediate chromatin tethering at the nuclear periphery. *Genome Biol*. 2019;20:87.
- [50] Choi J, Richards EJ. The role of CRWN nuclear proteins in chromatin-based regulation of stress response genes. *Plant Signal Behav*. 2020;15:1694224.
- [51] Guo T, Mao X, Zhang H, et al. Negatively regulate plant immunity through NAC WITH TRANSMEMBRANE MOTIF1-LIKE9 and NONEXPRESSOR OF PR GENES1 in Arabidopsis thaliana. *Mol Plant*. 2017;10:1334–1348.
- [52] Bi X, Cheng Y-J, Hu B, et al. Nonrandom domain organization of the Arabidopsis genome at the nuclear periphery. *Genome Res*. 2017;27:1–12.
- [53] Gibalová A, Steinbachová L, Hafidh S, et al. Characterization of pollen-expressed bZIP protein interactions and the role of ATbZIP18 in the male gametophyte. *Plant Reprod*. 2017;30:1–17.
- [54] Kagale S, Links MG, Genome-Wide RK. Analysis of ethylene-responsive element binding factor-associated amphiphilic repression motif-containing transcriptional regulators in Arabidopsis. *Plant Physiol*. 2010;152:1109–1134.
- [55] van Zanten M, Koini MA, Geyer R, et al. Seed maturation in Arabidopsis thaliana is characterized by nuclear size reduction and increased chromatin condensation. *Proc National Acad Sci* 2011; 108:20219–20224.
- [56] Zhao W, Guan C, Feng J, et al. The Arabidopsis CROWDED NUCLEI genes regulate seed germination by modulating degradation of ABI5 protein: cRWNs regulate seed germination. *J Integr Plant Biol*. 2016;58:669–678.
- [57] Sheehan MJ, Pawlowski WP. Live imaging of rapid chromosome movements in meiotic prophase I in maize. *PNAS*. 2009;106:20989–20994.
- [58] Varas J, Graumann K, Osman K, et al. Absence of SUN1 and SUN2 proteins in *Arabidopsis thaliana* leads to a delay in meiotic progression and defects in synapsis and recombination. *Plant J*. 2015;81:329–346.
- [59] Murphy SP, Gumber HK, Mao Y, et al. A dynamic meiotic SUN belt includes the zygotene-stage telomere bouquet and is disrupted in chromosome segregation mutants of maize (*Zea mays* L.). *Front Plant Sci* [Internet]. 2014 [cited 2020 Jul 8];5. Available from: <https://www.frontiersin.org/articles/10.3389/fpls.2014.00314/full>
- [60] Koszul R, Kleckner N. Dynamic chromosome movements during meiosis: a way to eliminate unwanted connections? *Trends Cell Biol*. 2009;19:716–724.
- [61] Fernández-Jiménez N, Pradillo M, Tatout C. The role of the nuclear envelope in the regulation of chromatin

- dynamics during cell division. *J Exp Bot* [Internet]. [cited 2020 Jul 8] Available from: <https://academic-oup-com.insb.bib.cnrs.fr/jxb/article/doi/10.1093/jxb/eraa299/5863386>
- [62] Graumann K, Evans DE. Nuclear envelope dynamics during plant cell division suggest common mechanisms between kingdoms. *Biochem J*. 2011;435:661–667.
- [63] Güttinger S, Laurell E, Kutay U. Orchestrating nuclear envelope disassembly and reassembly during mitosis. *Nat Rev Mol Cell Biol*. 2009;10:178–191.
- [64] Kimura Y, Kuroda C, Masuda K. Differential nuclear envelope assembly at the end of mitosis in suspension-cultured *Apium graveolens* cells. *Chromosoma*. 2010;119:195–204.
- [65] Patel S, Rose A, Meulia T, et al. WPP-domain proteins are developmentally associated with the nuclear envelope and promote cell division. *Plant Cell*. 2004;16:3260–3273.
- [66] Xu XM, Meulia T, Meier I. Anchorage of plant RanGAP to the nuclear envelope involves novel nuclear-pore-associated proteins. *Curr Biol*. 2007;17:1157–1163.
- [67] Choi J, Strickler SR, Richards EJ. Loss of CRWN nuclear proteins induces cell death and salicylic acid defense signaling. *Plant Physiol*. 2019;179:1315–1329.
- [68] Jarad M, Mariappan K, Almeida-Trapp M, et al. LITTLE NUCLEI 1 (LINC1) regulates pattern-triggered immunity and jasmonic acid signaling. *Front Plant Sci*. 2020;10:1639.
- [69] Jaiswal DK, Mishra P, Subba P, et al. Membrane-associated proteomics of chickpea identifies Sad1/UNC-84 protein (CaSUN1), a novel component of dehydration signaling. *Sci Rep*. 2014;4:4177.
- [70] Wang Q, Liu S, Lu C, et al. Roles of CRWN-family proteins in protecting genomic DNA against oxidative damage. *J Plant Physiol*. 2019;233:20–30.
- [71] Pecinka A, Chevalier C, Colas I, et al. Chromatin dynamics during interphase and cell division: similarities and differences between model and crop plants. *J Exp Bot*. 2019;71:5205–52022.
- [72] Tirnaz S, Epigenetics: BJ. Potentials and challenges in crop breeding. *Mol Plant*. 2019;12:1309–1311.
- [73] Gaut BS, Ross-Ibarra J. Selection on major components of angiosperm genomes. *Science*. 2008;320:484–486.
- [74] Zwaenepoel A, Van de Peer Y. Inference of ancient whole-genome duplications and the evolution of gene duplication and loss rates. *Mol Biol Evol*. 2019;36:1384–1404.
- [75] Jiao Y, Leebens-Mack J, Ayyampalayam S, et al. A genome triplication associated with early diversification of the core eudicots. *Genome Biol*. 2012;13:1–14.
- [76] Uber Zelltheilung. *RC, Jahrbuch M*. 1885;10:214–330
- [77] Santos AP, Shaw P. Interphase chromosomes and the Rab1 configuration: does genome size matter? *J Microsc*. 2004;214:201–206.
- [78] van Driel R, Fransz P. Nuclear architecture and genome functioning in plants and animals: what can we learn from both? *Exp Cell Res*. 2004;296:86–90.
- [79] Fransz P, de Jong H. From nucleosome to chromosome: a dynamic organization of genetic information. *Plant J*. 2011;66:4–17.
- [80] Armstrong SJ, Jones GH. Female meiosis in wild-type *Arabidopsis thaliana* and in two meiotic mutants. *Sex Plant Reprod*. 2001;13:177–183.
- [81] Bass HW, Marshall WF, Sedat JW, et al. Telomeres cluster de novo before the initiation of synapsis: a three-dimensional spatial analysis of telomere positions before and during meiotic prophase. *J Cell Biol*. 1997;137:5–18.
- [82] Zhang F, Tang D, Shen Y, et al. Protein ZYGO1 mediates bouquet formation to promote homologous pairing, synapsis, and recombination in rice meiosis. *Plant Cell*. 2017;29:2597–2609.
- [83] Moriguchi K, Suzuki T, Ito Y, et al. Functional isolation of novel nuclear proteins showing a variety of subnuclear localizations. *Plant Cell*. 2005;17:389–403.
- [84] Shah MNA, Arabia S, Islam T, et al. Molecular evolution of SUN-domain containing proteins in diverse plant species and their expression profiling in response to developmental and perturbation stimuli. *Phytochemistry*. 2019;157:28–42.
- [85] Zhang F, Ma L, Zhang C, et al. The SUN domain proteins OsSUN1 and OsSUN2 play critical but partially redundant roles in meiosis. *Plant Physiol*. [Internet]; 2020 [cited 2020 Jun 22]. Available from: <http://www.plantphysiol.org/content/early/2020/06/18/pp.20.00140>
- [86] Yang J, Chang Y, Qin Y, et al. A lamin-like protein OsNMCP1 regulates drought resistance and root growth through chromatin accessibility modulation by interacting with a chromatin remodeller OsSWI3C in rice. *New Phytol*. 2020;227:65–83.
- [87] Newman-Griffis AH, Cerro P del, Charpentier M, Meier I. Medicago LINC complexes function in nuclear morphology, nuclear movement, and root nodule symbiosis. *Plant Physiol*. 2019;179:491–506.
- [88] Walbot V, Hoisington DA, Neuffer MG. Disease lesion mimic mutations [Internet]. In: Kosuge T, Meredith CP, Hollaender A, et al. editors. Genetic engineering of plants: an agricultural perspective. Boston, MA: Springer US; 1983 cited 2020 Jul 8. p. page 431–42. DOI: 10.1007/978-1-4684-4544-2_29
- [89] Barabaschi D, Tondelli A, Desiderio F, et al. Next generation breeding. *Plant Sci*. 2016;242:3–13.
- [90] Forterre P. The universal tree of life: an update. *Front Microbiol*. 2015;6:717.
- [91] Tang H, Bowers JE, Wang X, et al. Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. *Proc Natl Acad Sci U S A*. 2010;107:472–477.
- [92] Van Bel M, Diels T, Vancaester E, et al. PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Res*. 2018;46:D1190–6.



Supplemental Figure S2: Schematic representation of all constructs used in this study.

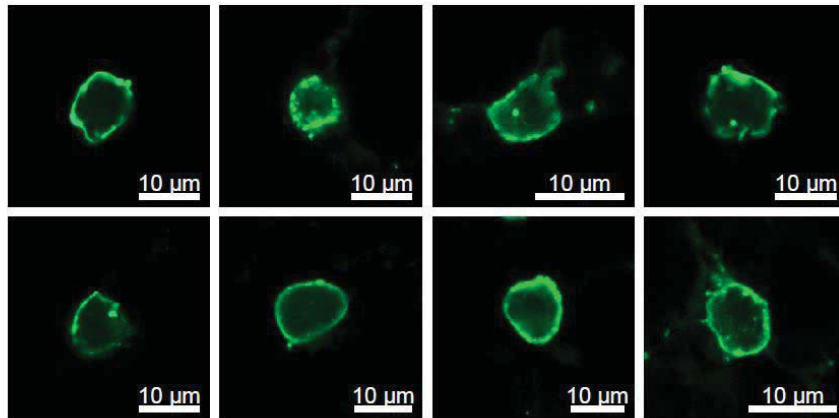
(A-C) Schematic representation of KAKU4 constructs with N-ter and C-ter regions **(A)**, deletions (dashed line) of motifs M1 (red), M2 (blue) and/or M3 (green) **(B)** or motifs M1, M2 and/or M3 only **(C)**. **(D-E)** Schematic representation of constructs: full length, deleted of motifs M1 and M2 or M1 and M2 only of NUP136 **(D)** or NUP82 **(E)**. FG repeats are indicated in purple, NLS in orange and GAR domain in light brown.



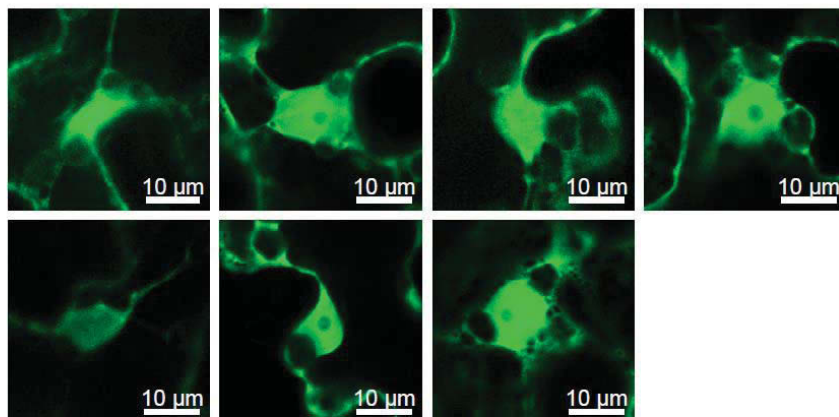
Supplemental Figure S3: Protein interactions between the nucleoskeleton and the NPC basket components.

(A-E) Yeast two-hybrid assay on test (SD/–Leu/–Trp/–Ade/–His) or permissive (SD/–Leu/–Trp) media. Empty vector pGBKT7 and pGADT7 are used respectively as bait and prey negative controls. (A) Interactions of KAKU4 with itself. (B) Role of KAKU4 M1 and M2 motifs of KAKU4 in its interaction with itself. (C) Role of M1 and M2 motifs in KAKU4 for the interactions of KAKU4 with CRWN proteins. (D) Interactions between KAKU4 and NUP82, NUP136. (E) Role of their M1 and M2 motifs in the interaction NUP82 and NUP136 with CRWN proteins.

A



B

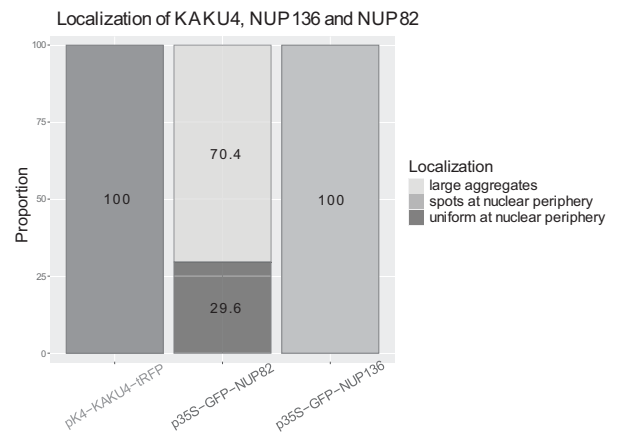
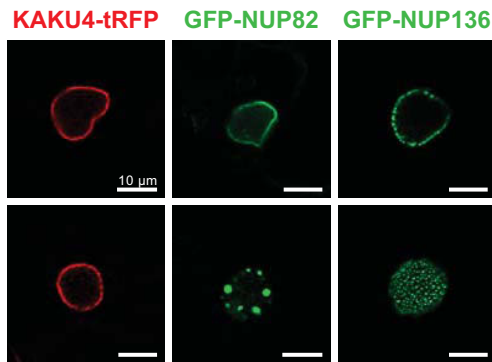


Supplemental Figure S4: Assay of nuclear localization of KAKU4 N and Cter regions.

A) Nter (1-310) and **B)** Cter (311-584) domains of KAKU4-GFP expressed transiently in *Nicotiana benthamiana*.

A

Gene name	Proteins	Tree length	dN/dS	P-value
<i>KAKU4</i>	Full-length	1.18	0.35	6.85E-30
<i>NUP82</i>	Full-length	1.34	0.53	2.93E-18
<i>NUP136</i>	Full-length	1.30	0.40	5.12E-61
<i>KAKU4</i>	N-terminal	1.21	0.35	4.54E-21
<i>NUP82</i>	N-terminal	1.33	0.54	7.81E-09
<i>NUP136</i>	N-terminal	1.14	0.29	6.80E-28

B

Supplemental Figure S5: PAML analysis and localization assay at the nuclear periphery of KAKU4, NUP82 and NUP136.

(A) Results of the test to determine whether the evolution of KAKU4, NUP82 and NUP136 is random (genetic drift) or under selection using Codeml NSsites from the PAML package. P-values are given for the full-length or N-terminal part of KAKU4, NUP82 and NUP136. Genetic drift hypothesis is rejected if $P\text{-value} < 0.5$. **(B)** Localization of pKAKU4-KAKU4-tRFP, p35S-GFP-NUP82 and p35S-GFP-NUP136 constructs in transient expression assay in *Nicotiana benthamiana* leaves. Two examples (top and bottom panels) are given. In a typical transient expression assay, KAKU4 localizes as a more uniform profile at the nuclear periphery ($n=43$; 100% uniform profiles) than NUP82 that displays either a uniform pattern ($n=8$; 29.6%) or large aggregates ($n=19$; 70.4%). NUP136 always displays a spotty pattern ($n=81$; 100%).

GeneID (IDs from NCBI ref_seq, PLAZA (Monocots and Dicots 4.5) or Congenie.org)	ProteinID (UniProt ID)	Proteins family	Species names	Species name abbreviations
NP_974653.1	Q949W6	KAKU4	<i>Arabidopsis thaliana</i>	A.th.KAKU4
XP_002869326.1	D7MB66	KAKU4	<i>Arabidopsis lyrata</i>	A.ly.KAKU4
XP_009126975.1	M4D467	KAKU4	<i>Brassica rapa</i>	B.ra.KAKU4
XP_021903194.1	UPI000B8C8B3C	KAKU4	<i>Carica papaya</i>	C.pa.KAKU4
PtQ48030.1	B9ILG6	KAKU4	<i>Populus trichocarpa</i>	P.tr.KAKU4
EOY29204.1	A0A061GQ30	KAKU4	<i>Theobroma cacao</i>	T.ca.KAKU4
XP_020410513.1	A0A251RDA9	KAKU4	<i>Prunus persica</i>	P.pe.KAKU4
XP_002280589.3	F6H3A9	KAKU4	<i>Vitis vinifera</i>	V.vi.KAKU4
XP_006578198.1	A0A0R0K9N2	KAKU4	<i>Glycine max</i>	G.ma.KAKU4
XP_004244492.1	UPI0002BCA6BB	KAKU4	<i>Solanum lycopersicum</i>	S.ly.KAKU4
XP_002447248.1	UPI0001A871E3	KAKU4	<i>Sorghum bicolor</i>	S.bi.KAKU4
NP_001361003.1		KAKU4	<i>Zea mays</i>	Z.ma.KAKU4
KAF2936273.1		KAKU4	<i>Oryza sativa</i>	O.sa.KAKU4
XP_020103774.1	UPI000981376B	KAKU4	<i>Ananas comosus</i>	A.co.KAKU4
XP_009399593.1	UPI00051210D5	KAKU4	<i>Musa acuminata</i>	M.ac.KAKU4
XP_020523650.1	UPI0009BE8A11	KAKU4	<i>Amborella trichopoda</i>	A.tr.KAKU4
NP_197520.1	F4K465	NUP82	<i>Arabidopsis thaliana</i>	A.th.NUP82
XP_020876167.1	UPI000A29CAF8	NUP82	<i>Arabidopsis lyrata</i>	A.ly.NUP82
RID73743.1	A0A398AES8	NUP82	<i>Brassica rapa</i>	B.ra.NUP82
PLAZA : Cpa.g.sc46.13		NUP82	<i>Carica papaya</i>	C.pa.NUP82
XP_0244446140.1	A0A2K1X065	NUP82	<i>Populus trichocarpa</i>	P.tr.NUP82
EOY31095.1	A0A061GMJ9	NUP82	<i>Theobroma cacao</i>	T.ca.NUP82
XP_007203780.1	M5VUD4	NUP82	<i>Prunus persica</i>	P.pe.NUP82
XP_010656434.1	D7TC29	NUP82	<i>Vitis vinifera</i>	V.vi.NUP82
XP_006578449.1	K7KK18	NUP82	<i>Glycine max</i>	G.ma.NUP82
XP_004242779.1	A0A3Q7H7U7	NUP82	<i>Solanum lycopersicum</i>	S.ly.NUP82
XP_002437257.1	C5Z5P6	NUP82	<i>Sorghum bicolor</i>	S.bi.NUP82
ACG45455.1	B6U7X2	NUP82	<i>Zea mays</i>	Z.ma.NUP82
XP_015640949.1	UPI0001C7C628	NUP82	<i>Oryza sativa</i>	O.sa.NUP82
XP_020088501.1	UPI00098189AB	NUP82	<i>Ananas comosus</i>	A.co.NUP82
PLAZA : Ma07_g10020		NUP82	<i>Musa acuminata</i>	M.ac.NUP82
ERN03727.1	W1P1M4	NUP82	<i>Amborella trichopoda</i>	A.tr.NUP82
NP_001327263.1	Q9CAF4	NUP136	<i>Arabidopsis thaliana</i>	A.th.NUP136
XP_002884804.1	D7L919	NUP136	<i>Arabidopsis lyrata</i>	A.ly.NUP136
RID71392.1	A0A398A0P7	NUP136	<i>Brassica rapa</i>	B.ra.NUP136

PLAZA : Cpa.g.sc33.123 + Cpa.g.sc33.124		NUP136	<i>Carica papaya</i>	C.pa.NUP136
XP_024462942.1	A0A2K1ZA20	NUP136	<i>Populus trichocarpa</i>	P.tr.NUP136
EOY19316.1	A0A061FQH4	NUP136	<i>Theobroma cacao</i>	T.ca.NUP136
XP_020411742.1	A0A251QG05	NUP136	<i>Prunus persica</i>	P.pe.NUP136
PLAZA : GSVIVG01034600001 + GSVIVG01034601001		NUP136	<i>Vitis vinifera</i>	V.vi.NUP136
XP_003556871.1	I1NED8	NUP136	<i>Glycine max</i>	G.ma.NUP136
XP_010312503.1	A0A3Q7IPC9	NUP136	<i>Solanum lycopersicum</i>	S.ly.NUP136
EER93329.2	C5WY53	NUP136	<i>Sorghum bicolor</i>	S.bi.NUP136
ONM10062.1	A0A1D6L6Z4	NUP136	<i>Zea mays</i>	Z.ma.NUP136
PLAZA LOC_Os03g57425 + LOC_Os03g57430		NUP136	<i>Oryza sativa</i>	O.sa.NUP136
XP_020091461.1		NUP136	<i>Ananas comosus</i>	A.co.NUP136
XP_009389436.1		NUP136	<i>Musa acuminata</i>	M.ac.NUP136
XP_006826428.1	W1NEY0	NUP136	<i>Amborella trichopoda</i>	A.tr.NUP136
Congenie.org : MA_101453p001		NUP136	<i>Picea abies</i>	P.ab.NUP136
Congenie.org : PITA_000024509		NUP136	<i>Pinus taeda</i>	P.ta.NUP136
XP_024382323.1	A0A2K1K630	NUP136	<i>Physcomitrella patens</i>	P.pa.NUP136
PTQ48030.1	A0A2R6XPI8	NUP136	<i>Marchantia polymorpha</i>	M.po.NUP136

Supplemental Table S1: Genes and proteins Sequences. Gene sequences were collected from NCBI ref_seq, PLAZA or Plantgenie.org and protein sequences from Uniprot database. Sequences were selected from 24 species from the green lineage that are listed as full names or abbreviations as used in fasta files or phylogenetic trees. Protein IDs are grouped according to whether they belong to the KAKU4, NUP82 and NUP136 protein family.

Protein Name	Uniprot ID	Gene Name	Class	Sub-Class 1	Sub-Class 2	Protein size	Nb AA order	Nb AA disorder	Order proportion	Disorder proportion
AAAS_ARATH	Q8GWR1	ALADIN	NUP	NA		447	416	31	0,931	0,0694
C3H16_ARATH	Q9FWS3	CG1	NUP	FG-NUP		435	138	297	0,3172	0,6828
CRWN1_ARATH	F4HRT5	CRWN1	NSK	NA		1132	254	878	0,2244	0,7756
CRWN2_ARATH	Q9SAF6	CRWN2	NSK	NA		1128	225	903	0,1995	0,8005
CRWN3_ARATH	Q9CA42	CRWN3	NSK	NA		1085	217	868	0,2000	0,8000
CRWN4_ARATH	Q9FLH0	CRWN4	NSK	NA		1010	165	845	0,1634	0,8366
F4K465_ARATH	F4K465	NUP82	NUP	NA	NPC basket	762	304	458	0,3990	0,6010
GLE1_ARATH	Q0WPZ7	GLE1	NUP	NA		611	392	219	0,6416	0,3584
GP210_ARATH	F4KHD8	GB210	NUP	NA		1923	1864	59	0,9693	0,0307
HOS1_ARATH	Q84JU6	HOS1	NUP	NA		927	726	201	0,7832	0,2168
KAKU4_ARATH	Q949W6	KAKU4	NSK	NA		485	205	280	0,4227	0,5773
NP214_ARATH	F4I1T7	NUP214	NUP	FG-NUP		1819	751	1068	0,4129	0,5871
NP93A_ARATH	O22224	NUP93A	NUP	NA		861	800	61	0,9292	0,0708
NU107_ARATH	Q8L748	NUP107	NUP	NA		1101	1032	69	0,9373	0,0627
NU133_ARATH	F4IGA5	NUP133	NUP	NA		1285	1152	133	0,8965	0,1035
NU155_ARATH	F4HXV6	NUP155	NUP	NA		1464	1340	124	0,9153	0,0847
NU160_ARATH	Q9C811	NUP160	NUP	NA		1495	1392	103	0,9311	0,0689
NU205_ARATH	F4KBW6	NUP205	NUP	NA		1838	1756	82	0,9554	0,0446
NU50A_ARATH	Q9C829	NUP50A	NUP	FG-NUP	NPC basket	440	146	294	0,3318	0,6682
NU50B_ARATH	Q9LW88	NUP50B	NUP	FG-NUP	NPC basket	465	109	356	0,2344	0,7656

NU93B_ARATH	F4J284	NUP93B	NUP	NA		860	772	88	0,8977	0,1023
NU98A_ARATH	Q8RY25	NUP98A	NUP	FG-NUP		1041	315	726	0,3026	0,6974
NU98B_ARATH	F4ID16	NUP98B	NUP	FG-NUP		997	246	751	0,2467	0,7533
NUA_ARATH	A4GSN8	NUA	NUP	NA	NPC basket	2093	585	1508	0,2795	0,7205
NUP1_ARATH	Q9CAF4	NUP136	NUP	FG-NUP	NPC basket	1309	174	1135	0,1329	0,8671
NUP35_ARATH	O04326	NUP35	NUP	NA		329	127	202	0,3860	0,6140
NUP43_ARATH	Q24JJ9	NUP43	NUP	NA		361	319	42	0,8837	0,1163
NUP54_ARATH	Q8GYF7	NUP54	NUP	FG-NUP		377	231	146	0,6127	0,3873
NUP58_ARATH	Q8RWH9	NUP58	NUP	FG-NUP		513	221	292	0,4308	0,5692
NUP62_ARATH	Q8L7F7	NUP62	NUP	FG-NUP		739	148	591	0,2003	0,7997
NUP85_ARATH	Q8RXH2	NUP85	NUP	NA		716	699	17	0,9763	0,0237
NUP88_ARATH	Q9FFK6	NUP88	NUP	NA		810	676	134	0,8346	0,1654
NUP96_ARATH	Q8LLD0	NUP96	NUP	NA		1046	923	123	0,8824	0,1176
RAE1_ARATH	Q38942	RAE1	NUP	NA		349	314	35	0,8997	0,1003
SC13B_ARATH	O64740	SEC13B	NUP	NA		302	293	9	0,9702	0,0298
SEH1_ARATH	Q93VR9	SEH1	NUP	NA		326	270	56	0,8282	0,1718
SUN1_ARATH	Q9FF75	SUN1	INM	NA		471	340	131	0,7219	0,2781
SUN2_ARATH	Q9SG79	SUN2	INM	NA		455	302	153	0,6637	0,3363
SUN3_ARATH	F4I316	SUN3	INM	NA		660	350	310	0,5303	0,4697
SUN4_ARATH	F4I8I0	SUN4	INM	NA		596	386	210	0,6477	0,3523
SUN5_ARATH	F4JPE9	SUN5	INM	NA		561	483	78	0,8610	0,1390

Supplemental Table S2: Intrinsically disorder prediction by VSL2 for NUPs, nucleoskeleton and SUN domain proteins. Protein and Gene names, Uniprot IDs is given and classified into NUPs, NSK (nucleoskeleton) and Inner Nuclear Membrane (INM) proteins. FG-NUPs and NUPs from the NPC basket are indicated. The number of ordered and disordered amino-acids (AA) were defined using precomputed predictions available at DescribeProt server (Materials and Methods) and then used to compute the proportion of disorder for each protein.

Tableau non incorporé car très grand

Supplemental Table S3: Intrinsically disorder prediction by VSL2, DEPICTER and DISOPRED3. Score for prediction of disordered residues ranging from 0 (order) to 1 (disorder) are predicted for each amino-acid of KAKU4, NUP82 and NUP136 using VSL2, DEPICTER and DISOPRED3. Profiles using full-length and N-terminal protein sequences are depicted respectively in Supplemental Figure S4 in Figure 6A.

	NUP136-nY	NUP82-nY	nY-KAKU4	nY-CRWN1	nY-CRWN2
NUP136-cY	-	-	-	-	-
NUP82-cY	+	-	+	-	+
cY-KAKU4	+/-	-	+	-	-
cY-CRWN1	-	-	-	-	NT
cY-CRWN2	-	-	-	NT	-

Supplemental Table S4: Bimolecular fluorescence complementation (BiFC) assay in *Nicotiana benthamiana* leaves. Fibrillarin-RFP was co-expressed as a marker for visualizing nuclei in the transformed cells. nYFP and cYFP are the N-terminal and C-terminal half of YFP respectively. “+” positive interaction, “+/-” weak interaction; “-” absence of recorded interaction, “NT” not tested.

Gene/Construct	Sequences (5' → 3')	Use
KAKU4	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGATTCCGTCTCCGGT GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATTTGGCCCGTCCTTT	Gateway Cloning
KAKU4Nter	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGATTCCGTCTCCGGT GGGGACCACTTTGTACAAGAAAGCTGGGTGTTAGAGCATATCAGTCGCCTAA	Gateway Cloning
KAKU4Cter	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAAAGGAAATGCTTCGTCT GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATTTGGCCCGTCCTTT	Gateway Cloning
KAKU4ΔGAR	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGATTCCGTCTCCGGT GGGGACCACTTTGTACAAGAAAGCTGGGTGTTAGTATGCCAAGACCTTGGT	Gateway Cloning
KAKU4-M1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATTGAGCAGCTTCTCTTG GGGGACCACTTTGTACAAGAAAGCTGGGTGTTAATCTACAACCCTAGCCTT	Gateway Cloning
KAKU4-M2	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGGATCGCCTGTGGACGTG GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATGGTGATCGAAAATCTAA	Gateway Cloning
KAKU4-M3	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCCTCAGCAAACCCCTAGT GGGGACCACTTTGTACAAGAAAGCTGGGTGTTAGACGACGGAAGAAATGAA	Gateway Cloning
KAKU4-M1-M2	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATTGAGCAGCTTCTCTTG GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATGGTGATCGAAAATCTAA	Gateway Cloning
KAKU4-M3-M1-M2	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCCTCAGCAAACCCCTAGT GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATGGTGATCGAAAATCTAA	site-directed mutagenesis
Deletion M1 of KAKU4	CAGTTCTAAGCGTGATACATCCAGCGTCCCT AGGGGACGCTGGGATGTACACGCTTAGAACTG	site-directed mutagenesis
Deletion M2 of KAKU4	AAGCAACTGAAGATGGTGCTGCTGCGAATAATTTAGATTTTCG CGAAAATCTAAATTTATTCGCAGCAGCACCATCTTCAGTTGCTT	site-directed mutagenesis
Deletion M3 of KAKU4	CCTCAGCAAACCCCTAGTTTCATTTCTTCCGTCTGTC GACGACGGAAGAAATGAACTAGGGTTTTGCTGAGG	Dual system mutagenesis
NUP82	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCCACTCAAGGAGAA CGGGGACCACTTTGTACAAGAAAGCTGGGTGCTTTTAAACAGCAGACACC	Gateway Cloning
NUP82-M1-M2	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCTTGAAAGGCTAATGGAA GGGGACCACTTTGTACAAGAAAGCTGGGTGTTAGGAACACTTGTATGATGTC	Gateway Cloning
Deletion M1-M2 of NUP82	ACGACGTTGCAATCTCGGAGCAAGGTTTTGTAGCAAGAAA TTTCTTGCTACAAAACCTTGCTCCGAGATTGCAACGTCGT	site-directed mutagenesis
NUP136	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCGAGCGCGGCACGG GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATTTCTTCCCTGGTGGA	Gateway Cloning
NUP136-M1-M2	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCTTGAGAAAATCTTACAGG GGGGACCACTTTGTACAAGAAAGCTGGGTGTTAAGGAGTAACCTTCTGAAGG	Gateway Cloning
Deletion M1-M2 of NUP136	CTAAGGATGGATTTACAGATTCTATGTTAGGGTTACGTGG CCACGTAACCCTAACATAGAATCTGTAAATCCATCCTTCG	site-directed mutagenesis
CRWN1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGTCCACGCCGTTGAAG GGGGACCACTTTGTACAAGAAAGCTGGGTGCTACGTCGTCAGAAAGT	Gateway Cloning
CRWN2	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGACGCCGAGAAGCGAG GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATGTAGTGAGAAAAGTCC	Gateway Cloning
CRWN3	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGTTCACTCCGCAAAGG	

	GGGGACCACTTTGTACAAGAAAGCTGGGTGTCATGTTGTGAAAAAGACC	Gateway Cloning
CRWN4	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCAACTTCTTCTCGTTC GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATGTGCATATATCTTCTG	Gateway Cloning
Fib2	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAGACCTC CTCTAACTGG GGGGACCACTTTGTACAAGAAAGCTGGGTGCTAAGCAGCAGTAGCAGCC	Gateway Cloning
<i>kaku4-2</i>	GGATGAGGAAAAGTTTCCAGC ACCAAGCATTTCAACGACTCAG	T-DNA Genotyping
<i>nup136-2</i>	CGAGACTTTGGAGAAGAAGGC CACCTTTTCCCTTTGGAGGATC	T-DNA Genotyping

Supplemental Table S5: Primer sequences used in this study.

Name	Genotype	Type
pDONR-ZEO	Zeo ^R	Gateway
pGWB606 (p35S::GFP)	Spec ^R , Basta ^R , p35S::GFP Nter	Gateway
pGWB555 (p35S::mRFP)	Spec ^R , Hygro ^R , p35S::mRFP Nter	Gateway
pGWB559 (pK4::tRFP)	Spec ^R , Hygro ^R , Cter tRFP	Gateway
pDEST-GADT7	Tag HA, Gal4-AD, Amp ^R , Leu+	Gateway
pDEST-GBKT7	Tag c-Myc, Gal4-DBD, Kan ^R , Trp+	Gateway
pB4GWnY	Spec ^R , Kan ^R , p35S::nYFP Cter	Gateway
pB4GWcY	Spec ^R , Kan ^R , p35S::cYFP Cter	Gateway
pB4nYGW	Spec ^R , Kan ^R , p35S::nYFP Nter	Gateway
pB4cYGW	Spec ^R , Kan ^R , p35S::cYFP Nter	Gateway

Supplemental Table S6: Plasmid backbones used in this study.

	Cloning		Localisation			Yeast-two-hybrid			BiFC	
	pDONR-ZEO	PGWB606 (p35S::GFP-X)	PGWB555 (p35S::mRFP-X)	PGWB559 (pK4::X- RFP)	pDEST-GADT7	pDEST-GBKT7	pB4GWnY (p35S::X- nYFP)	pB4GWcY (p35S::X- cYFP)	pB4nYGW (p35S::nYFP-X)	pB4cYGW (p35S::cYFP-X)
KAKU4	X	X	X	X	X			X	X	
KAKU4Nter	X	X		X	X					
KAKU4Cter	X	X		X	X					
KAKU4ΔM1	X	X		X						
KAKU4ΔM2	X	X		X						
KAKU4ΔM1ΔM2	X	X		X						
KAKU4ΔM3	X	X								
KAKU4ΔM3ΔM1ΔM2	X	X								
KAKU4-M1	X	X		X						
KAKU4-M2	X	X		X						
KAKU4-M1-M2	X	X		X						
KAKU4-M3	X	X								
KAKU4-M3-M1-M2	X	X								
NUP82	X	X		X	X	X	X			
NUP82ΔM1-M2	X	X		X						
NUP82-M1-M2	X	X		X						
NUP136	X	X		X	X	X	X			
NUP136ΔM1-M2	X	X		X						
NUP136-M1-M2	X	X		X						
CRWN1	X	X		X	X			X	X	
CRWN2	X	X		X	X*			X	X	
CRWN3	X	X		X	X					
CRWN4	X	X		X	X			X	X	
Fibrillarin	X		X							

Supplemental Table S7: Overview of constructs used in this study.

* The CRWN2 cDNA contain an intron after the the coding sequence for the CC domains

Annexe 4 : Article “Image analysis workflows to reveal the spatial organization of chromatin and chromosomes”

Image analysis workflows to reveal the spatial organization of chromatin and chromosomes

Ricardo S Randall^{1,*}, Claire Jourdain^{2,*}, Anna Nowicka^{3,*}, Kateřina Kaduchová^{3,*}, Michaela Kubová^{4,*}, Mohammad A. Ayoub^{5,*}, Veit Schubert^{5,*}, Christophe Tatout^{6,*}, Isabelle Colas^{7,*}, Kalyanikrishna², Sophie Desset⁶, Sarah Mermet⁶, Aurélia Stevens⁶, Ivona Kubalová³, Terezie Mandáková⁴, Stefan Heckmann⁵, Martin A. Lysak⁴, Martina Panatta⁸, Raffaella Santoro⁸, Daniel Schubert², Ales Pecinka³, Devin Routh⁹ and Célia Baroux¹

*, equal contribution

Affiliations

¹ Department of Plant and Microbial Biology, Zürich-Basel Plant Science Center, University of Zürich, Switzerland

² Institute of Biology, Freie Universität Berlin, Germany

³ Institute of Experimental Botany, v. v. i. (IEB), Centre of the Region Haná for Biotechnological and Agricultural Research (CRH), Olomouc, Czech Republic

⁴ Central European Institute of Technology (CEITEC), Department of Experimental Biology and National Centre for Biomolecular Research, Masaryk University, Brno 625 00, Czech Republic

⁵ Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, D-06466 Seeland, Germany

⁶ Institut Génétique, Reproduction et Développement (GRoD), Université Clermont Auvergne, CNRS, INSERM, 63001, Clermont-Ferrand, France

⁷ The James Hutton Institute, Errol Road, Invergowrie, DD2 5DA, Scotland, United Kingdom

⁸ Department of Molecular Mechanisms of Disease, DMMD, University of Zurich, 8057, Zurich, Switzerland.

⁹ Service and Support for Science IT (S3IT), Universität Zürich, Zürich, Switzerland [Correspondence](#)

cbaroux@botinst.uzh.ch

ORCID numbers

Author	ORCID Number
Célia Baroux	0000-0001-6307-2229
Ricardo S Randall	0000-0002-2778-519X
Ivona Kubalová	0000-0002-5673-9715
Veit Schubert	0000-0002-3072-0485
Isabelle Colas	0000-0001-6980-9906
Mohammad Ayoub	0000-0001-9371-3277
Stefan Heckmann	0000-0002-0189-8428
Michaela Kubová	0000-0003-3474-7324
Terezie Mandáková	0000-0001-6485-0563
Martin A. Lysak	0000-0003-0318-4194
Daniel Schubert	0000-0003-2390-0733

Author	ORCID Number
Kalyanikrishna	0000-0002-8026-246X
Claire Jourdain	0000-0002-3568-580X
Anna Nowicka	0000-0002-5762-3482
Ales Pecinka	0000-0001-9277-1766
Katerina Lahenova	0000-0001-9259-5942
Aurélia Stevens	0000-0003-1378-2569
Sarah Mermet	0000-0001-6839-3224
Sophie Desset	0000-0002-4897-4977
Christophe Tatout	0000-0001-5215-2338
Martina Panatta	0000-0002-8180-8380
Raffaella Santoro	0000-0001-9894-2896

Abstract

Nucleus, chromatin, and chromosome organisation studies heavily rely on fluorescence microscopy imaging to elucidate the distribution and abundance of structural and regulatory components. Three-dimensional (3D) image stacks are a source of quantitative data on signal intensity level and distribution and on the type and shape of distribution patterns in space. Their analysis can lead to novel insights that are otherwise missed in qualitative-only analyses. Quantitative image analysis requires specific software and workflows for image rendering, processing, segmentation, setting measurement points and reference frames and exporting target data before further numerical processing and plotting. These tasks often call for the development of customized computational scripts and require an expertise which is not broadly available in the community of experimental biologists. Yet, the increasing accessibility of high- and super-resolution imaging methods fuel the demand for user-friendly image analysis workflows. Here, we provide a compendium of strategies developed by participants of a training school from the COST action INDEPTH to analyse the spatial distribution of nuclear and chromosomal signals from 3D image stacks, acquired by diffraction-limited confocal microscopy and super-resolution microscopy methods (SIM and STED). While the examples make use of one specific commercial software package, the workflows can easily be adapted to concurrent commercial and open-source software. The aim is to encourage biologists lacking custom-script-based expertise to venture into quantitative image analysis and to better exploit the discovery potential of their images.

Abbreviations

3D FISH, three-dimensional fluorescence in situ hybridization	HEI10, HUMAN ENHANCER OF INVASION 10
3D, three-dimensional	NCO, Non-Crossover
ASY1, ASYNAPTIC 1	NE, Nuclear Envelope
CC, chromocenters	Oligo-FISH, oligonucleotide fluorescence in situ hybridization
CO, Crossover	RNPII, RNA Polymerase II
DAPI, 4',6-diamidino-2-phenylindole	SC, Synaptonemal Complex
DMC1, DNA MEIOTIC RECOMBINASE 1	SIM, structured illumination microscopy
DSB, Double Strand Break	ZMM (ZIP, MSH4, MSH5 and MER3 proteins)
FISH, fluorescence in situ hybridization	ZYP1, ZIPPER-LIKE PROTEIN 1
GFP, GREEN FLUORESCENT PROTEIN	

Keywords

Nucleus, chromatin, 3D organization, spatial distribution, image analysis, segmentation, quantification, mitosis, meiosis, chromosome, metaphase, pachytene, cross overs, nuclear speckles, nuclear bodies, RNA Pol II, transcription factories, oligo FISH, STED imaging, SIM

Funding details

This work was funded by grants from: the Swiss National Science Foundation [310030_185186 and IZCOZO_198171 to CB, 310030_201268 to RS], the Velux Foundation and the Ricola foundation to CB; the European Research Council Grant Shuffle (Project ID: 669182) and Biotechnology and Biological Sciences Research Council (BB/T008636/1) to IC; the Deutsche Forschungsgemeinschaft to VS (Schu 762/11-1); the UZH-FUB Global Strategy and Partnerships Funding Scheme (grant to K), the German Academic Exchange Service (Deutsche Akademische Austauschdienst Dienst; grant to K); the International Max Planck Research School for Biology and Computation (IMPRS-BAC, PhD grant to CJ); the Czech Science Foundation (project no. 18-20134S) and the Operational Programme Research, Development and Education – „Project Internal Grant Agency of Masaryk University" (No. CZ.02.2.69/0.0/0.0/19_073/0016943 to ML) and “Plants as a tool for sustainable global development”

(No. CZ.02.1.01/0.0/ 0.0/16_019/0000827 to KK and AP); the German Federal Ministry of Education and Research (BMBF) in frame of the grant HERBY (FKZ-031B0188) to SH; the Czech Science Foundation grant (21-02929S) and Purkyně fellowship to AP; the European Research Council [ERC-AdG-787074-NucleolusChromatin to R.S.]; CT, AS, SM and SD are supported by CNRS, INSERM, Université Clermont Auvergne (UCA), 16-IDEX-0001 CAP 20-25 CIR1 and CIR3

Disclosure statement

<https://authorservices.taylorandfrancis.com/editorial-policies/competing-interest/>

No potential competing interest was reported by the authors.

Data availability statement

Training image datasets are available on the INDEPTH-OMERO repository [1, 2].

Detailed instructions for each image-segmentation workflow and demo images are provided as supplemental material.

Introduction

Elucidating the spatial organization of eukaryotic genomes, their structural and compositional dynamics during cellular processes and functional relationship with the nucleus, is a keystone of three-dimensional (3D) genomics. 3D genomics aims to decipher the functional, (3D) organizing principles of the chromosomes, chromatin domains and nucleus that contribute to transcription, replication, repair, and recombination. Understanding the 3D genome requires multidisciplinary methods including high-throughput, sequencing-based, molecular profiling techniques, computational simulation-based biophysical and mathematical modeling, and microscopy imaging at high-to-super resolution and in three-dimensions (3D) [3, 4]. Microscopy followed by image analysis provides the opportunity to measure chromosome and chromatin structures down to the nanoscale, with a few kilobases resolution. This can inform on the genomic interactions *in situ* and the spatial organization of genomic domains in relation to the 3D nuclear space and its functional compartments [1, 5].

Venturing into these opportunities to probe for the spatial organization of the genome *in situ* requires dedicated imaging and image analysis procedures, recently captured by the concept of quantitative, data-driven microscopy [3]. Quantitative image analysis for nuclear and chromosomal studies can be implemented at different levels of complexity, depending on the research question/s and, often, the expertise available. For instance, a

simple level consists in the scoring of a structure or pattern on the image based on user-defined classification. This can be applied when the immunolabelled chromatin protein, or FISH-labelled genomic domain, shows a very distinct distribution pattern (eg. punctuate vs diffuse), varying between treatments or genotypes. In this case, quantifying the relative occurrence of pattern categories by scoring may be sufficient to address the original question. Manual scoring can also be used to quantify a moderate number of labeled regions (e.g., number of FISH signals or nuclear bodies). These categorical, quantitative approaches have the virtue to be accessible to all experimentalists, without sophisticated software. Such approaches can nevertheless reliably identify a relatively simple signal distribution, providing a limited number of samples, and double-blind scoring to avoid cognitive biases. Yet, for many images (e.g., from high-throughput imaging), images with multiple labels, showing complex spatial patterns of signal distribution, with continuous (rather than discrete) variation in signal abundance, or a combination of all, require computationally driven processing approaches for quantitative analyses. A core step required is image segmentation. This process partitions the image based on the signal distribution into digital objects identifying biologically relevant structures. Various image segmentation methods and algorithms exist. These perform differently depending on the signal distribution [6], with deep-learning approaches for automated segmentation tasks at a large-scale

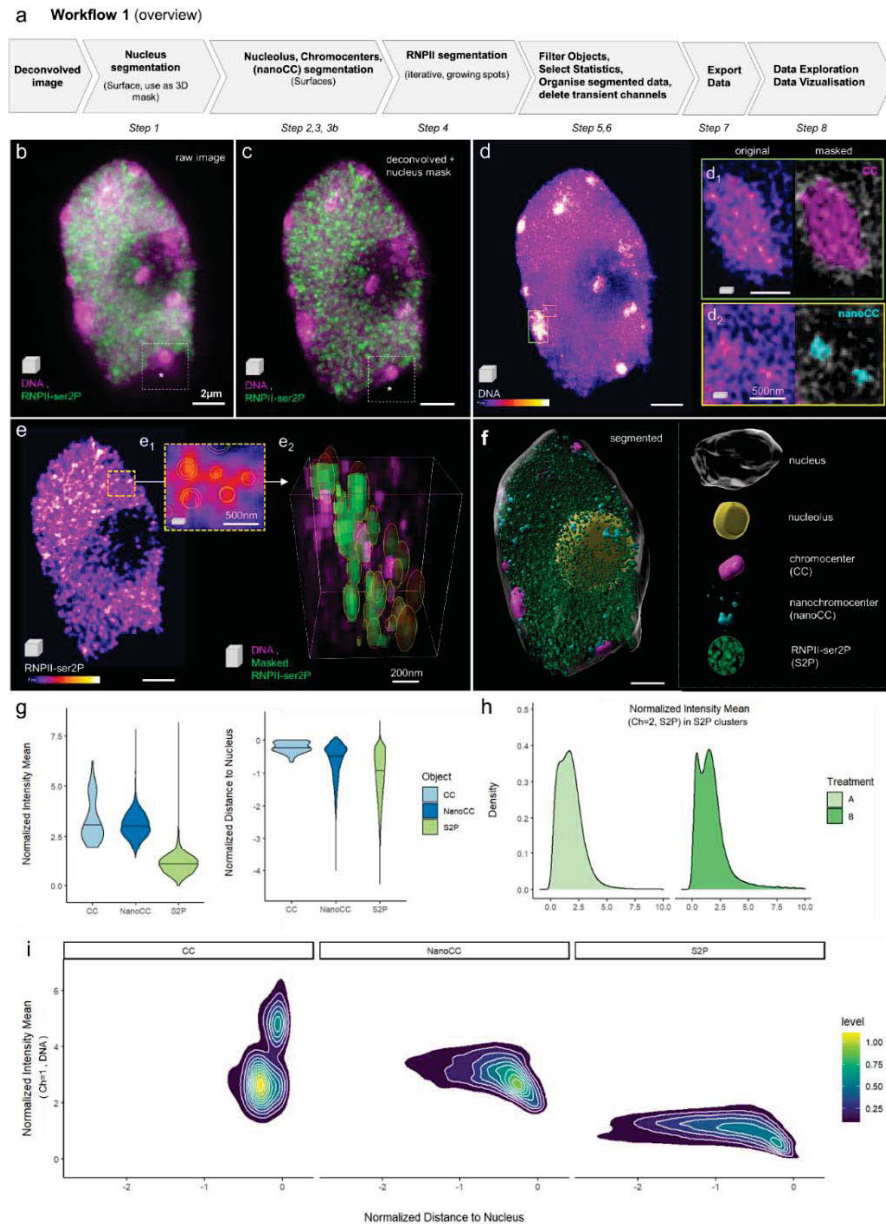


Figure1. Analysis of the spatial distribution of RNA Pol II clusters in intact nuclei

(a) Overview of the workflow illustrated in b-i; **(b)** 3D projection of a 3D-STED image reporting on immunolabelled RNA Pol II (isoform phosphorylated on SerP, green, RNP2-ser2P) and DNA (magenta, Hoechst 580CP, [17]), raw image; **(c)** Same image following deconvolution, nucleus contour segmentation and masking; **(d)** Intensity-coded coloring mode (Fire) of the DNA channel and frames magnified in the insets showing examples of chromocenters (CC, d1) and nanochromocenter (nanoCC, d2) in the original channel (left) and after segmentation and pseudo-coloring (right); **(e)** Intensity-coded coloring mode (Fire) of the RNPII-ser2P channel showing a dense distribution of clusters with identifiable intensity peaks, enabling segmentation as adaptive spots (e1, e2), e1: single plane showing the spot contours; e2, 3D segment of the image after segmentation, clusters pseudo-colored in green, DNA in magenta. **(f)** Fully segmented image containing surface (nucleus, nucleolus, CC and nanoCC) and spot objects (RNPII-ser2P, abbreviated S2P). **(g-i)** Data exploration using *DataViz* (github.com/barouxlabs/DataViz, Supplemental File 1- Video 2). **(g)**, Violin plots showing a similar DNA density distribution in CC and nanoCC but much lower density in S2P clusters (intensity mean, DNA channel, normalized per image) and a sharp peripheral location of CC as formerly described (Andrey et al., 2010; Franz et al., 2002), contrasting with the more dispersed distribution of nanoCC and S2P clusters (distance to nucleus surface (0) normalized using the nucleus center of mass as reference); **(h)** Example showing an application of the workflow, to compare the distribution of RNPII cluster intensities between two treatments: A and B. **(i)** Another example illustrating one of the many analyses enabled by the workflow and *DataViz*, with density scatter plots of DNA intensity means in RNPII clusters as a function of their distance to the nucleus surface. Scale bars: b-f, 2µm; insets, as indicated.

being continuously developed [7]. Once the image is segmented, multiple features can be extracted from the 3D digital objects, for instance: object number, size and shape; signal intensity and variance per object type, texture of the signal, channel and position in the image; distance relationships, and spatial distribution. Practically, these features are highly relevant to analyse the spatial organization of chromatin, chromosome and nuclear components *in situ*.

The field would greatly benefit from the broader deployment of image processing-based quantitative analyses [1, 4]. Yet a major hurdle for most biology-only oriented labs is the lack of computational expertise for customizing image processing scripts and large data handling, the lack of template workflows, or both. Key concepts, from image acquisition to quantitative data, have been framed in recent years, for applications in cell biology, but also to set good practice and standards in the field [3, 8]. Efforts are undertaken to promote education and support in image analysis for scientists dealing with biological images [9]. In this context, this resource paper aims at providing a compendium of image analysis workflows for nucleus and chromatin studies, taking seven case-studies as examples. The workflows are based on a user-friendly, commercial image processing software (Imaris, Bitplane, Switzerland) but are conceptually applicable to concurrent (commercial or open source) software. These workflows were developed by participants of the training school of the INDEPTH COST action [2].

The workflows associated with each case study are briefly described below and are illustrated in the related figures. Each workflow is associated with a supplemental file folder that includes: a step-by-step guideline (text); a table summarizing the main steps functions and parameters used on the training image; a training image and, for some workflows, a video tutorial.

[Analysing the spatial distribution of transcription clusters](#)

In mammals, a radial gradient model of transcription in the nucleus has been proposed [10, 11]. In plants, including the *Arabidopsis thaliana* (Arabidopsis) plant model, little is

known about the spatial (3D) distribution of transcription. Transcriptional activity in the nucleus can be visualized *in situ* by immunolabeling the active isoform of RNA Polymerase II (RNPII). In Arabidopsis, super-resolution imaging of RNPII has shown a reticulated pattern throughout the nucleoplasm along which distributed clusters of variable size and intensity exist [12, 13].

To resolve the spatial (3D) distribution of RNPII signals in Arabidopsis nuclei, we imaged RNPII and DNA using 3D-STED microscopy. To quantify RNPII foci distribution we designed an image analysis workflow ([Figure 1a](#) and [Supplemental File 1](#)). Sample preparation and imaging are described elsewhere [14]. Deconvolved STED images are segmented using the Imaris software (Bitplane, Switzerland) to create digital objects corresponding to the nucleus, the nucleolus, the chromocenters and the RNPII signals ([Figure 1b-d](#), [Supplemental File 1 - Video 1](#)). The surface object corresponding to the nucleus is also used to apply a 3D mask to separate the true image from the background signal (compare the insets in [Figure 1a-b](#)). While the nucleus and nucleolus are segmented based on smoothed, manual contours, heterochromatin is segmented using the supervised automated tool. Chromocenters (CC) are typically large, brightly stained regions. In Arabidopsis nuclei, these are discrete and relatively easy to segment ([Figure 1d](#), [inset d1](#)). Super-resolution imaging revealed additional heterochromatin regions we termed nanochromocenters (nanoCC), that can also be segmented ([Figure 1d](#), [inset d2](#)). RNPII signal shows a complex nuclear distribution in Arabidopsis nuclei: rather than being discrete, it spreads unevenly in a reticulated manner with, however, clearly identifiable local clusters (Schubert et al., 2015). Our aim was to segment the image to discretise RNPII signal and focus on the clusters, considering their variable size, to further analyse their variability in intensity, size and spatial distribution. We applied the growing spot function in an iterative manner and could capture 70-80% of the RNPII signal in spots of variable size ([Figure 1e](#), [inset e2](#)). This stepwise segmentation resulted in a digital image composed of objects capturing the nucleus, the nucleolus, the chromocenters and RNPII

a Workflow 2 (overview)

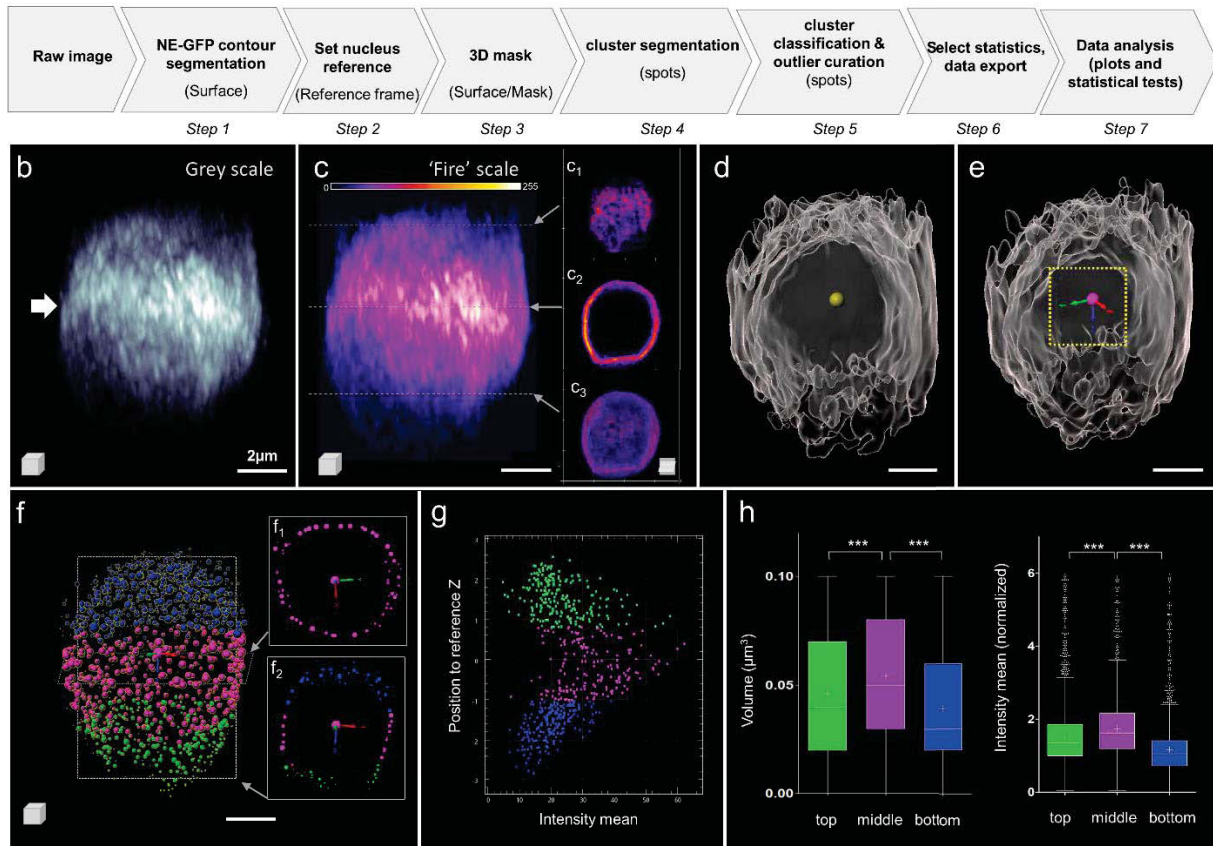


Figure 2: Analysis of the spatial distribution of a fluorescently tagged protein associated with the nuclear envelope.

(a) Overview of the Image analysis workflow. Details of the parameters are in supplements. **(b)** Raw image of NE-GFP (Nuclear Envelope - associated protein fused to GFP) signal in a root nucleus; 3D rendering in grey levels suggests an enrichment of the protein at the equatorial region of the nucleus (white arrow). **(c)** Same image (3D) as in (b) using a fire color scale for NE-GFP signal intensities display (0-255), c1-c3 insets: cross sections at selected top, middle and bottom planes, respectively. **(d)** Result of the segmentation of the NE-GFP signal domain as a surface (grey); a spot (yellow) is created at the surface' center-of-mass. **(e)** A new XYZ coordinate system (reference frame) is docked at the center-of-mass. **(f)** The NE-GFP signal is segmented as spots of adaptive size ('growing spots') using the channel masked by the surface; spots are classified according to their axial (z) position, the equatorial region is defined $\pm 2\mu\text{m}$ around the origin. Three spot classes are created located at the top, middle and bottom of the nucleus (blue, magenta, green, respectively). f1, f2 insets: XY and XZ sections. **(g)** The intensity mean of the spots is plotted as a function of their axial position (z) relative to the new reference frame for the image shown in (b-f). The colors indicate the 'top', 'middle' and 'bottom' classes, respectively. **(h)** The volume and normalized intensity mean of NE-GFP spots are plotted for each class, for n=8 nuclei images segmented following this workflow. Kruskal-Wallis and Dunn's multiple comparison tests with bottom vs middle and top vs middle indicate statistically significant differences with $P < 0.001$ (***) for both variables. Scale bars: 2 μm

clusters (Figure 1f). Variables of interest, such as signal intensity per channel, object size and shape and distance between objects (spot-to-spot, spot-to-surface) were exported for each object type and channel.

The high number of variables, object type, channels, image replicates and levels of comparison (such as genotypes and treatment) dramatically increases data complexity. To facilitate data exploration, we built a stand-alone data visualization interface named *DataViz*

(<https://github.com/barouxlabs/DataViz>) which allows one to interactively plot all, or a subset of, data. This also enables custom variable creation for the normalization of distances and intensity per image (Figure 1g-i, Supplemental File 1 -Video 2). Here, we provide a few examples of violin plots (Figure 1g), density distributions (Figure 1h) and scatter plots with density contours (Figure 1i). The mean intensity, normalized per nucleus, of the DNA signal shows that the average and range of chromatin compaction in CC and nanoCC is largely similar, while nanoCC occasionally show a higher compaction (upper tail of the violin plot). By contrast, and as expected for transcriptionally active regions, chromatin is, in average, 2-3x less condensed in the RNPII-S2P clusters. Also, plotting the shortest distance of each object to the nucleus surface (negative values indicate distance towards the nucleus interior), confirms the peripheral localization of CC as already described [15, 16], an apparent enrichment of the nanoCC towards the periphery (although less pronounced than CC), and spreading of the RNPII-S2P clusters from the periphery towards the nuclear interior (with an apparent decreasing occurrence linked to the presence of the nucleolus). Further, plotting the density distribution of RNPII-S2P signals in the clusters (normalized mean intensity) allows detecting different structures of the RNPII landscape between different treatment (A and B in the example provided Figure 1h). Finally, *DataViz* enables exploring the relationship between two continuous variables using scatter plots, with or without density contours. In the example provided Figure 1i, we interrogated the relationship between the distance to the nucleus boundary and the mean DNA intensity

for each of the nuclear domains segmented as CC, nanoCC and S2P clusters. The plots suggested (i) two categories of CC distinctive mostly by their intensity but slightly different with regards to their peripheral position, and that (ii) nanoCC and transcription clusters closer to the periphery are on average less compact than their counterpart located more towards the nuclear interior. These are only a few examples of the numerous possible plots that collectively contribute to data mining and discovery.

The segmentation process described in detail in the supplemental material corresponds to a user-supervised workflow. The input values (threshold, smoothing factor, or filtering values) are either software-defined values (and depend on image attributes) or customized by the user to best capture the biological objects. The parameters are then saved and re-applied for subsequent image replicate analyses. If the image quality is highly reproducible, it is further possible to apply automated batch-segmentation (following the software provider's instructions). For a trained user, the workflow takes *ca.* 45min per image or less. Finally, this workflow can be further applied for the analysis of other types of nuclear signals showing punctate distribution similar to that in our example.

[Analysis of the spatial distribution of proteins located at the nuclear periphery](#)

To date, the distribution of nuclear envelope (NE)-associated proteins is poorly documented in plants. 3D microscopy-based observations may provide new insights into the organization of chromatin domains at the nuclear periphery at the single-cell level.

In this example, we developed a workflow to quantify the spatial distribution pattern of an NE-associated protein, labelled with GFP (NE-GFP) (Figure 2a). We imaged Arabidopsis root nuclei from 1 week old seedlings using a confocal microscope equipped with an Airyscan module [18], (Supplemental File 3- Image2). The NE-GFP protein is located at the NE and form clusters of variable sizes with an apparent asymmetric distribution (arrow, Figure 2b).

a Workflow 3 (overview)

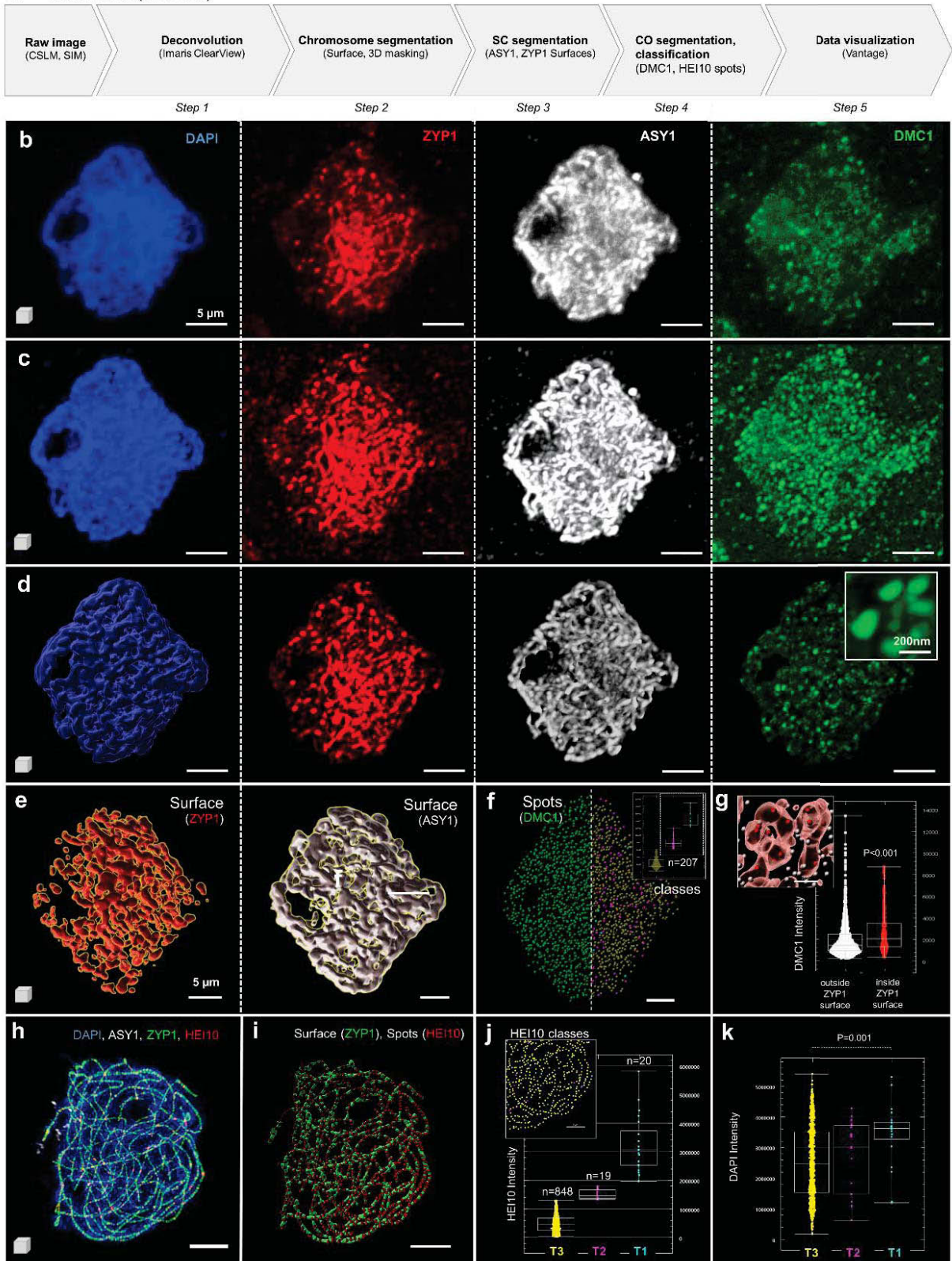


Figure 3 – Analysis of cross-over distribution in meiocytes

Using an intensity-based channel coloring mode confirmed the enrichment of NE-GFP at the equatorial plane of the nucleus, compared to the top and bottom poles (Figure 2c and insets). The first step in our procedure is to segment the global domain of NE-GFP signal using the 'Surface' function of Imaris (Figure 2d). In a second step, we created a spot at the center of mass of the segmented NE-GFP surface (yellow spot, Figure 2d) and used it to dock a 'Reference frame' object (Figure 2e). This new XYZ coordinate system at the nucleus' center is used later for spot classification. The third step is to use the surface to mask the residual NE-GFP signal outside the surface (see the detailed procedure in supplemental file 2). The fourth step consists in the segmentation of the NE-GFP signal which appears as clusters of varying sizes and intensities. To capture this variability, the 'spot' tool was used with the 'region growing' option to create spots of an adaptive size suitable to model the NE-GFP signal distribution. Next (step 5), the spots are classified in three categories (top:blue, middle:magenta, bottom:green) according to their axial position in the new coordinate system defined at step 2 (Figure 2f). The 'middle' class is defined by a region encompassing the origin of the new coordinate system from -2 to +2 μm along the z axis. The 'top' and 'bottom' classes capture the spots above and below this equatorial region, respectively. In addition, this step includes a curation of the segmentation result to (i) keep spots strictly located at the periphery (removing outliers located internally due to surface invaginations) and (ii) select spots of biologically relevant size (up to 600nm diameter, $\sim 1.1 \mu\text{m}^3$) (see the detailed procedure in supplemental file 2). A 2D-plot of the mean intensity of segmented NE-GFP spots according to their axial (x) position in the new coordinated system (step 2) reveals higher intensities among spots located at the equatorial plane ('middle' class, Figure 2g). The segmentation of multiple

images allowed verifying this finding following data export (step 6) and plotting (step 7, Figure 2h, n=8 nuclei). Importantly, as fluorescence intensities vary between images, the mean intensity of each spot is normalized by the mean intensity in the NE-GFP surface, per image (Figure 2h). This analysis revealed that both the volume and the mean intensity of NE-GFP clusters in the equatorial plane ('middle' class) are significantly different from that of the clusters located at the polar regions ('top' and 'bottom' classes), (Figure 2h, Kruskal-Wallis and Dunn's multiple comparison test with $P < 0.0001$ for all pairs). In conclusion, this image analysis workflows enables quantifying the spatial heterogeneity of proteins associated with the nuclear envelope. In combination with mutant genetics, this approach enables to assess the quantitative influence of candidate regulators and that of intrinsic (protein) domains on the spatial protein localization.

Analysis of protein distribution on meiotic chromosomes

Meiosis is a special type of cell division occurring during sexual reproduction and enabling genetic recombination. During the first stage of meiosis, prophase I, homologous chromosomes align along their entire length by a protein structure called the synaptonemal complex (SC). This process is essential for crossover (CO) formation in many eukaryotes. Prophase I is itself divided into five substages – Leptotene, Zygotene, Pachytene, Diplotene and Diakinesis. Each stage can be monitored by immunostaining of specific proteins involved in SC formation. The most common targets are ASY1 (ASYNAPTIC 1) and ZYP1 (ZIPPER-LIKE 1), [19]. During Prophase I, homologous recombination starts with the formation of SPO11-programmed DNA double-stranded breaks (DSB; [20]). These DSBs are subsequently processed and recombinases such as DMC1 (DNA MEIOTIC RECOMBINASE 1) mediate strand invasion, essential for CO formation [21]. In barley a large number of DSBs is formed [22]

Figure 3 – Analysis of cross-over distribution in meiocytes (continued)

(a) Overview of the image analysis workflow in 5 steps illustrated on two images marking the synaptonemal complex (SC) and crossing overs (CO) with different components (provided in [supplemental File 3](#)): **image 3a** (b-g) represents a barley meiocyte at zygotene stage immunostained for DMC1, ZYP1 and ASY1 and counterstained for DNA using DAPI. The image was acquired by confocal microscopy (ZEISS LSM 710) as described (Colas et al., 2019). **Image 3b** (h-k) represents a barley meiocyte at the late pachytene stage immunostained for ASY1 (Ch=2), ZYP1 (Ch=3) and HEI10 (Ch=4) and counterstained for DNA using DAPI (Ch=1). The image was acquired by 3D-SIM (Colas et al., 2019). **(b)** Original image acquired by confocal imaging, the different labeling are indicated. **(c)** image following deconvolution to resolve the synaptonemal complex (SC) and immunostained cross-over (CO). **(d)** segmentation of the chromosomes as surface and masking of the ZYP1, ASY1 and DMC1 channels to remove background signal. It allows resolving DMC1 foci at high resolution (inset). **(e)** SC segmentation using the ZYP1 and ASY1 masked channels **(f)** DMC1 foci segmentation (left) and classification according to their intensity (right and inset= intensity plot per category), **(g)** classification of DMC1 foci according to their distance relative to the ZYP1 surface. **(h)** Original 3d SIM image (image 3b), **(i)** same image following ZYP1 and HEI10 segmentation, **(j)** HEI10 spots were classified according to their intensity (T1, T2, T3 on graph and inset); 20 foci were scored (automatic) for the T1 class as described in earlier studies, **(k)** HEI10 classes differ by the DNA density. Scale bars: 5 μm except for the inset d, DMC1 channel (200nm), Plots (f, g, j, k): Imaris Vantage.

but only 13-22% (depending on the cultivar and scoring method) are repaired to crossover (CO) while the rest are repaired as non-crossovers (NCO) [23, 24]. What controls the fate of DSB (CO vs NCO) is poorly understood. A current hypothesis involves HEI10 (HUMAN ENHANCER OF INVASION 10), a ZMM class-of-protein in the CO repair pathway, as an early indicator [25].

To elucidate whether HEI10 also contributes to DSB fate designation in barley, one approach is to elucidate the dynamics of HEI10 foci along the prophase chromosomes at early, mid, and late stages, and in relation to DMC1 at early prophase. This approach requires 3D imaging of (immunostained) meiotic proteins on prophase chromosomes and the scoring of HEI10 vs DMC1 foci in relation to the prophase stage. We describe here a workflow to process 3D-SIM images to enable the scoring and classification of DMC1 and HEI10 foci depending on their size and intensity.

The images used to illustrate the workflow (Figure 3a) are that of nuclei of barley male meiotic cells at different stages labeled for components of the SC (ASY1, ZYP1), processed DSBs (DMC1) and recombination intermediates (HEI10). Image 3a shows a zygotene stage nucleus labeled with ASY1 (white), ZYP1 (red), DMC1 (green) and counterstained with DAPI (blue), acquired on a confocal microscope (Zeiss LSM 710) as per Colas *et al.* [24], (Figure 3b-g, Supplemental File 3- Image3a). Image 3b shows a pachytene stage nucleus labeled with ASY1 (white), ZYP1 (green), HEI10 (red) and counterstained with DAPI (blue) (Figure 3h-k, Supplemental File 3- Image3b), and acquired using 3D-SIM similar as for rye [26]. The aim for both image analyses was to segment HEI10 or DMC1 foci and to analyse their properties relative to the SC, their size and their intensity. For this, the chromosomes, SC complex and HEI10 or DMC1 foci are segmented separately. Different strategies can be used for filtering and classifying the HEI10 or DMC1 foci and are illustrated with a workflow for image 3a and another for image 3b (see detailed guidelines and parameters in supplemental file 3).

In brief, Image 3a (Figure 3b) was deconvoluted using Imaris module Clearview

9.5 to reconstruct the image at optical resolution (Figure 3c), which is essential for subsequently estimating the CO diameter l (Figure 3c, DMC1 channel inset). Next, the DAPI staining was used to generate a 3D surface of chromosomes serving as a mask to remove signal noise in the image (Figure 3d). Note that here, the aim was not to segment the chromosomes very precisely, as the masking would result in the exclusion of ASY1, ZYP1 and DMC1 signals that do not all colocalise with DNA at that meiotic stage. Hence, permissive criteria were preferred in this case. Next, the SC complex was segmented on both the ASY1 and ZYP1 channels, creating two distinct surfaces (Figure 3e). Finally, DMC1 foci were segmented as spot objects using an estimated seed size of 200 nm (Figure 3 f). The algorithm detects all possible foci with both low and high intensities. Classical studies have so far focused on high-intensity foci in a number of a few hundred (Colas *et al.*, 2019; Higgins *et al.*, 2012). By contrast, the workflow described here enables one to capture all foci, first, irrespective of their intensity, and additionally to classify them according to intensity, during the creation process (Figure 3f, right panel). Alternatively, spots can be classified based on normalized signal intensity in a third-party software application following data export (for instance, using DataViz, see Workflow1). In an intensity sum-based classification we scored 217 DMC1 spots with medium-to-high intensity as previously reported for a similar stage of meiosis [22, 24]. The remaining low-intensity spots may correspond to either immunolabeling noise or unbound proteins. Next, we asked whether DMC1 localisation correlated with the SC. Indeed, following the classification of DMC1 spots in two groups, inside or outside the ZYP1 surface, we found a significant enrichment of DMC1 signal (intensity mean) when foci colocalize with ZYP1 (Figure 3g). This is only one of the many examples of correlative analyses that can be carried-out in such segmented images.

Image 3b analysis (Figure 3i-k) followed a similar workflow, although chromosome segmentation and masking were omitted in this case, to focus on the SC (ZYP1) and co-localizing recombination intermediates (HEI10)

a Workflow 4 (overview)

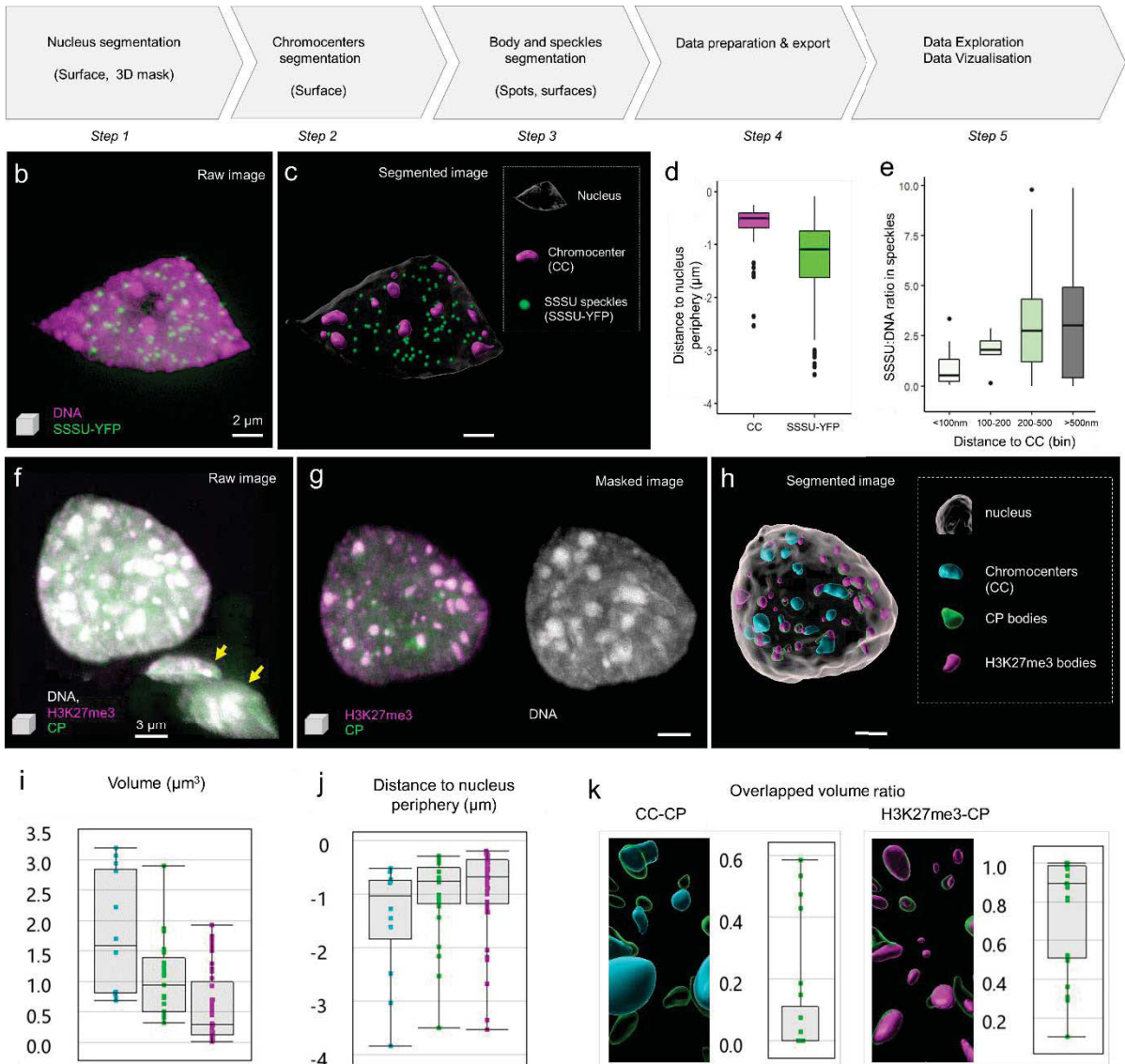


Figure 4. Analysis of the spatial distribution of nuclear speckles & bodies

(a) Overview of the image analysis workflow, details are provided in supplemental files 4. **(b)** A raw, STED image (3D projection) showing an isolated leaf nucleus stained for DNA (magenta, Hoechst 580CP, [17]) and immunostained for SSSU (green). **(c)** Segmentation result: the nucleus, chromocenters (CC) and the nuclear speckles (SSSU) were segmented as surface objects (legend, right panel). **(d)** The position of CC and SSSU speckles was plotted relative to the nucleus' periphery defined by the surface's boundary (0= at the boundary; negative values= towards the interior), $n=9$ nuclei analysed. **(e)** The relative enrichment of SSSU on chromatin was plotted as the SSSU:DNA mean signal intensity ratio for different classes of speckles defined by their distance to CC (in μm). Plots were generated using dataviz using data from $n=10$ segmented nuclei. **(f)** A confocal image (3D projection) of a nuclei of mouse naïve pluripotent embryonic stem cells Stained for DNA (grey, DAPI), immunostained for a chromatin protein (CP, green) and h3k27me3 (magenta) forming large nuclear bodies, the arrows show truncated nuclei in the field of view undesirable for downstream analyses. **(g)** Same image after 3D masking using the nucleus surface created at step 1. **(h)** Results of image segmentation: the nucleus, chromocenters (CC) and the nuclear bodies (CP and h3k27me3) were segmented as surface objects (legend, right panel). **(i-k)** Quantitative analysis of CC and nuclear bodies: volume: (i) distance to the nucleus periphery (j) and overlapping volume ratios (k, left: CC and CP overlap, right: CP and h3k27me3 overlap). Plots were generated on the demo image 4b using Imaris Vantage. Scale bar: (a-b), $2\mu\text{m}$; (f-h), $3\mu\text{m}$

segmentation. In this case, the ZYP1 surface was used to mask the HEI10 channel to specifically focus on foci associated with ZYP1 (Figure 3i). HEI10 spots objects were then classified according to their intensity, considering notably the first two 2% quantile versus the rest (Figure 3j) to create the T1, T2 and T3 classes, an approach formerly described to analyse meiosis in a fungal species [27]. Notably, the surprisingly high number of small HEI10 foci during late pachytene suggests the need for further investigation. To further describe the properties of HEI10 classes, we investigated different relationships and found that in most cells, typically high-intensity HEI10 foci (T1) localise, on average, on chromosome regions with higher DNA compaction (intensity) compared to low-intensity HEI10 foci (T3) (Figure 3k).

This image processing workflow facilitates the scoring of class I CO and NCO foci across multiple images, stages and genotypes, a task largely done manually until now. In addition, segmentation is near-exhaustive and includes low-intensity foci that were discarded from manual scoring in former studies. This raises the question of the dynamics of HEI10 and DMC1 foci formation, with possible intermediate stages represented by low-intensity foci. In addition, it opens the possibility to refine the analysis of CO/NCO spatial organisation and their fine-scale structure. For instance, the localisation of CO/NCO foci can be measured relative to the SC components as a function of their intensity, and as a function of local chromatin compaction.

Analysis of nuclear speckle distribution

A distinguishing feature of nuclear topography is the ability to accommodate a variety of subnuclear compartments including nuclear bodies. Nuclear bodies are distinct sites within a confined volume of the nucleus that spatially compartmentalize the nuclear environment to facilitate more efficient interaction between nuclear components. Like membrane-bound organelles, they also maintain an effective steady-state structure, probably by a mechanism different from that of the cytoplasmic organelles [28]. The first

identified and best-characterized plant nuclear bodies are the nucleolus and Cajal bodies. Several other smaller structures have however been identified, including speckles, paraspeckles, coiled bodies and photobodies [29-31]. Unmasking the mechanisms by which cells assemble, maintain, and regulate nuclear bodies & speckles, and the environmental and developmental factors contributing to the process will shed light on their biological functions. For instance, splicing regulator (SR) proteins in plants localize in speckles, whose size and shape are dependent on cell type, metabolic state, and transcriptional activity [30-32].

One way to elucidate nuclear body & speckles dynamics is through microscopy imaging and image analysis. This approach enables analysing their spatial distribution and their composition relative to other nuclear components and DNA (chromatin) density. In this example, we show-case a simple image analysis workflow to analyse the distribution of nuclear speckles and bodies. We used two images: [supplemental file 4 – image 4a](#) reports on the nuclear localisation of a plant chromatin remodeler, a SWI/SNF subunit (called SSSU here) forming nuclear speckles in leaf nuclei; [supplemental file 4 – image 4b](#) reports on the nuclear localisation of a mammalian chromatin protein (called CP here) and of H3K27me3 forming large nuclear bodies in nuclei of mouse naïve pluripotent embryonic stem cells [33].

SSSU was found to interact with PWO1 and CRWN1 (Kalyanikrishna, Jourdain, Schubert, unpublished), a set of proteins involved in epigenetic gene regulation and chromatin organization at the nuclear periphery [34]. CRWN1 (CROWDED NUCLEI 1) is a plant nuclear lamina candidate CRWN1 in Arabidopsis and interacts with PWO1 (PROLINE-TRYPTOPHANE-TRYPTOPHANE-PROLINE INTERACTOR OF POLYCOMBS1), a plant-specific protein associated with histones and PRC2 (POLYCOMB REPRESSIVE COMPLEX 2)[34]. Due to its possible interaction with CRWN1, we asked whether SSSU is also located preferentially at the nuclear periphery. To answer this question, we tagged SSSU with YFP (YELLOW FLUORESCENT PROTEIN) and imaged

a Workflow 5 (overview)

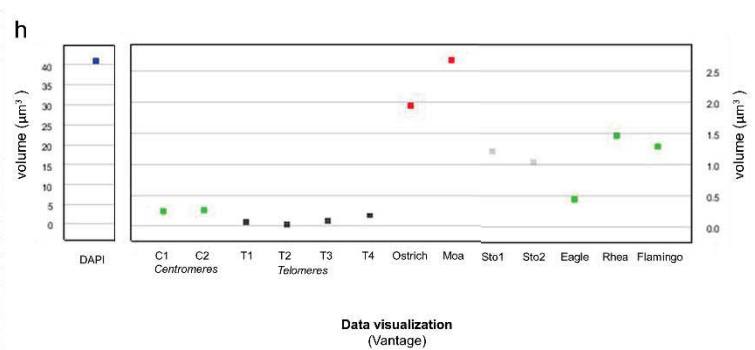
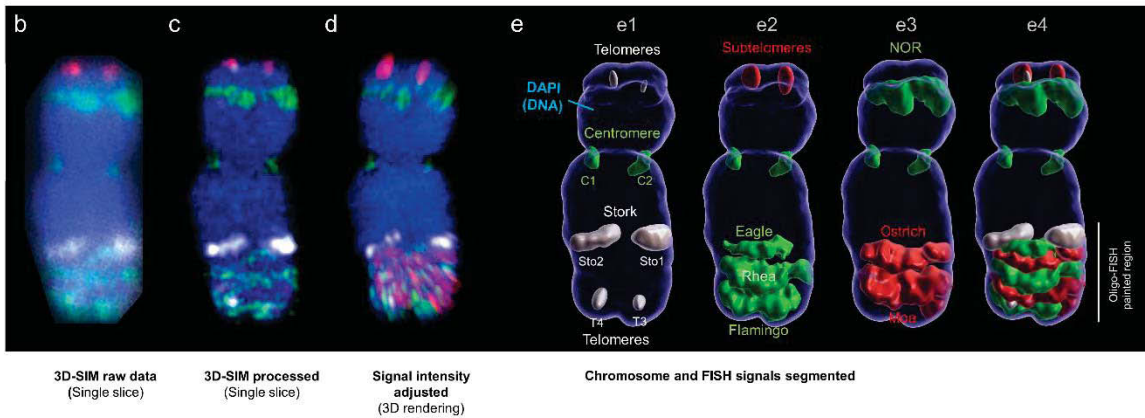


Figure 5. Analysis of metaphase chromosome ultrastructure using volume measurement of oligo-FISH labeled regions.

(a) Overview of the image analysis workflow. **(b)** 3D-SIM raw image slice from a stack containing 30 slices at widefield resolution. **(c)** 3D-SIM processed image slice showing increased resolution. **(d)** Display adjustment to optimize the visualization of signals with varying intensities. **(e)** Segmentation results: the chromosome is segmented using the DAPI channel and the generated 3D surface is used as a mask to specifically retain chromosomal FISH signals and exclude background. The segmentation is presented sequentially for different FISH probe groups (e1-e3), and the result is shown in the merge (e4). e1, telomere, centromere and Stork probes; e2, Subtelomeres, Eagle, Rhea and Flamingo probes; e3, 45SrDNA (Nucleolar Organising Region, NOR), Ostrich and Moa probes. The Oligo-FISH probes label the bottom part of chromosome 5HL. **(f)** Side view of a 3D movie generated via the ‘Animation’ tool (Supplemental File 5 – video 1). **(g)** Volume data are read in the ‘Statistics’ Tab for selected surfaces. **(h)** Data visualization using the ‘Vantage’ tool for individual objects (top).

nuclei expressing SSSU-YFP using STED microscopy. The image analysis workflow consists in only a few steps (Figure 4a): STED images reporting on the immunolabeled SSSU-YFP and DNA counterstaining (Figure 4b) are segmented for the nucleus, chromocenters (CC) and SSSU- speckles using the surface tool. For the segmentation of the nucleus, smooth, manual contours were used, while for segmenting CC automated, parameter-controlled settings were applied. SSSU speckles were segmented as spot objects of ca 200 nm diameter. Segmentation data of several images (n=9) were exported and plotted using DataViz (see workflow 1, [supplemental file 1-video2](#)). SSSU-YFP speckles showed a broad spatial distribution, with no clear preferential enrichment towards the periphery (Figure 4d), in contrast to chromocenters as previously shown [15, 16]. We found, however, that SSSU speckles are not uniform: they differ in their relative enrichment (SSSU:DNA ratio), which correlates with the proximity to CC (Figure 4e). Our analysis demonstrates that the nuclear speckles formed by SSSU are not preferentially enriched at the nuclear periphery as would have been expected from their biochemical interaction with CRWN1. The analysis suggests a differential enrichment depending on the proximity to other nuclear bodies, the CC, a relationship whose functional relevance remains to be investigated. This preliminary finding was unexpected and was revealed thanks to the possibility to explore multiple relationships between distance and intensity measurements in DataViz using segmentation data generated by this workflow.

In the second example, we were interested in the CP protein localization relative to the repressive nuclear compartments formed by heterochromatin (chromocenters) and H3K27me3 in nuclei of mouse naïve pluripotent embryonic stem cells. Nuclei stained for DNA and immunostained for CP and H3K27me3 were imaged at high resolution by confocal microscopy (Figure 4f). To analyse the distribution of CP bodies, we segmented the nucleus, the chromocenters (CC), CP and H3K27me3 nuclear bodies as surfaces of adaptive size (Figure 4g-h). Volume measurements show that CP bodies are smaller

than CC but larger than H3K27me3 bodies (Figure 4i, $P < 0.001$, Wilcoxon test) and similarly distributed towards the periphery compared to CC and H3K27me3 bodies (Figure 4j). The image shows an intricate relationship between CP bodies, CC and H3K27me3 bodies. Measuring the overlapped volume ratio is a useful approach to quantify the fraction of spatially colocalizing bodies (Figure 4k), revealing in our case a frequent overlap of 50% or more of CP bodies with H3K27me3 bodies, whereas the overlap with CC is less frequent and occurs to a lower extent (<20%). This image analysis workflows thus allows to quantify features of nuclear bodies distribution that are otherwise underappreciated with qualitative data only. Based on this simple workflow, further processing steps can be implemented that would contribute a refined analysis of the spatial pattern of the CP protein relative to chromatin density and H3K27me3 levels. This can include for instance the creation of intensity-based colocalization or ratio channels (not shown).

[Analysis of the higher-order chromatin organization in mitotic chromosomes](#)

Oligo-FISH combined with spatial super-resolution structured illumination microscopy (3D-SIM) was useful for discovering that the chromatids of barley metaphase chromosomes are organized at the higher-order level as a helix, the so-called chromonema. Additionally, volume measurements of the oligo-FISH painted regions allowed calculating the chromatin compaction density within the chromosomes. Based on the DNA quantity used to design the oligo-FISH probes and the measured volumes, the DNA density along chromosome arm 5HL was calculated. Thus, at interstitial arm regions a chromatin density ($34.1 \text{ Mb}/\mu\text{m}^3$) ~ 1.7 times higher than that at the region adjacent to the subtelomere ($19.5 \text{ Mb}/\mu\text{m}^3$) has been identified [35].

Centromeres, 45SrDNA (NOR, nucleolus organizing regions) telomeres, and subtelomeres of somatic metaphase chromosomes were labeled with specific FISH probes as described [35]. Besides, half- and full helical turns of the chromonema were painted by oligo-FISH at the long arm region of

a Workflow 6

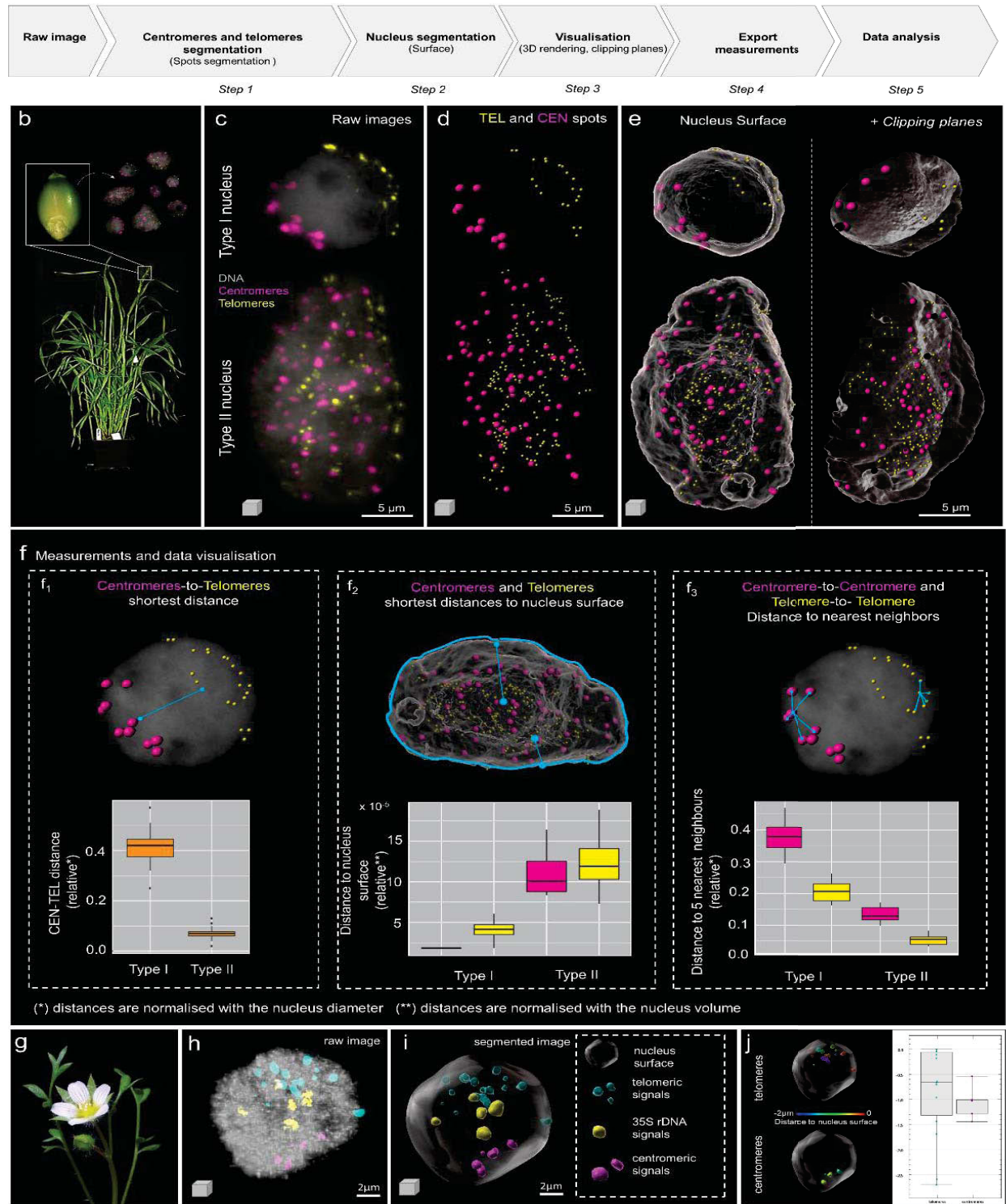


Figure 6. Analysis of centromere and telomere positioning in the interphase nucleus.

chromosome 5H (([Figure 5](#), probes were named according to birds: Stork, Eagle, Ostrich, Rhea, Moa, Flamingo), [35].

Workflow 5 describes the processing procedure to segment an individual chromosome and the different FISH signals to obtain quantitative measurements on the labeled regions ([Figure 5a](#); [Supplemental File 5](#)). In this example, 3D-SIM raw data image stacks ([Figure 5b](#)) were acquired using an Elyra PS.1 microscope system equipped with a 63×/1.4 Oil Plan-Apochromat objective and processed with the software ZENBlack (Carl Zeiss GmbH) (Weisshart et al., 2016) ([Figure 5c](#)). After file conversion (Imaris 9.7 converter, Bitplane), DAPI and FISH signal intensities were adjusted for improving visualization using the 'Display Adjustment' tool ([Figure 5d](#)). The DAPI-labeled chromosome was segmented using the surface tool and the surface was used to mask the image to remove background signal outside this region of interest. Additional surfaces of the other, differently colored FISH signals were generated ([Figure 5e-f](#), [Movie 5](#)). The surface volume data were established ([Figure 5g](#)), exported for further analysis by compiling several images, and used to calculate the volumetric density of the different FISH-labeled regions along the chromosome ([Figure 5h](#)) [35]. An example plot for one chromosome is shown in [Figure 5i](#) using Vantage.

[Analysis of centromeres and telomeres' positioning](#)

Arabidopsis and barley are eukaryotic models contrasting in their three-dimensional (3D) interphase chromosome organization. Arabidopsis has a small genome of about 157 Mbp per haploid DNA content (1C) packed into 5 chromosomes ($2n = 10$), whereas the barley genome is large, with around 5.1 Gbp/1C divided into 7 chromosomes ($2n = 14$) [36, 37]. In Arabidopsis, centromeres are distributed relatively equally around the nuclear periphery to which they are attached, while telomeres associate with nucleoli and each chromosome occupies a discrete territory within the nuclear space [38]. In barley, interphase chromosomes are organized in the so-called Rabl configuration with the centromeres and telomeres clustered at opposite nuclear poles [37]. While the Rabl configuration was long

thought to be prevalent among monocot species, recent studies show that it also occurs in dicot species and that variations exist within the same phylogenetic group [39]. In addition, this peculiar organisation can occur in a tissue-specific manner, as in rice[40]. To better characterise the occurrence of Rabl vs non-Rabl configurations and their possible intermediates, in different species and tissue types, there is the need to define an image analysis workflow quantifying telomere and centromere distribution in the nuclear space. We present such a workflow (overview [Figure 6a](#)) illustrated with two examples, corresponding to studies of chromosome organisation in a monocot species (barley, [Figure 6b-f](#)) and in a dicot species (*Limnanthes floccosa* subsp. *bellingiana*, [Figure 6g-j](#)). Details, parameters, and demo images are available in [Supplemental Files 6](#).

A first example is given for barley nuclei ([Figure 6b-f](#)). Nuclei extracted from seeds ([Figure 6b](#)) were flow-sorted as described [41] and labeled by FISH using fluorescently labeled oligoprobes (Cy3-labelled *CEREBA*-centromeric repeat; [42] and Cy5-labeled Arabidopsis-type telomeric repeats [43]. Z-stack images were acquired with an epifluorescence microscope connected with a spinning disk (Andor, Oxford Instruments, UK). Centromeric and telomeric FISH signals and the DNA counterstain (DAPI) were pseudo-colored in magenta, yellow, and grey respectively. Two types of seed nuclei are shown (Type I, Type II, [Figure 6c](#)). Images are first segmented on the channels reporting on FISH signals to create spot objects corresponding to centromeres (CEN) and telomeres (TEL), ([Figure 6d](#)) using the automated tool. The spots diameter is adjusted to the average size of FISH foci (ca 300nm). The nuclear surface is rendered using a low smoothing factor ([Figure 6e](#)). For better visualization of the spot distribution inside the nucleus, the surface is set to transparent or can be digitally sectioned using a clipping plane ([Figure 6e](#)). This segmentation and 3D visualization approach allowed us to realise that the seed nuclei population was composed of two categories of nuclei. In type I nuclei, centromeric and telomeric spots are grouped to opposite sides of the nucleus, reflecting Rabl-

Figure 6. Analysis of centromere and telomere positioning in the interphase nucleus (continued).

(a) Workflow overview showing the main steps to process the 3D image and identify centromeres and telomeres and their position in an interphase nucleus. The workflow is illustrated with seed nuclei from barley (a-f) and leaf nuclei from *Limnanthes* (g-i). **(b)** Barley plant, seeds and isolated nuclei stained by FISH for centromere and telomeric repeats (see main text for details). **(c)** Raw images (3D projections) of type I and type II nuclei showing centromeric (magenta) and telomeric (yellow) FISH probes signals, counterstained for DNA (DAPI, grey). **(d)** Telomeric (TEL) and centromeric (CEN) signals were segmented as spots. **(e)** 3D rendering together with nucleus surfaces (grey) following segmentation, whole nuclei (left) or clipped (right), exposing the CEN and TEL signals in the interior of the nucleus. **(f)** The distribution of telomeres and centromeres is described according to three measurements derived from spot-to-spot or spot-to-surface statistics: shortest distance between centromeres and telomeres (f1), shortest distance of centromeres to the nucleus surface and shortest distance of telomeres to the nucleus surface (f2), inter-centromere and inter-telomere distances computed as the average distance to the nearest 5 neighbor spots of the same category (f3). In blue, schematic representation of the measured distance. Distances were exported and normalized to the nucleus diameter (f1, f3) or nucleus volume (f2) and plotted using the ggplot GUI online tool (<https://shiny.gmw.rug.nl/ggplotgui/>). The lower and upper hinges of the boxplots correspond to the first and third quartiles of the data, respectively, the black lines within the boxes mark the median. Five to ten nuclei were used for each measurement. Black spots beyond the whiskers represent outliers. **(g-j)** Illustration of the workflow on a *Limnanthes* leaf nucleus, **(g)** *Limnanthes floccosa* subsp. *bellingneriana*, **(h)** Raw image (3D projection) of a nucleus stained for centromeric repeats (magenta), telomeric repeats (cyan) and rDNA repeats (yellow) by FISH, counterstained for DNA (DAPI, grey), imaged by confocal laser scanning microscopy, **(i)** 3D nucleus following segmentation of FISH signals and DNA as surfaces. **(j)** Distance of the different segmented groups relative to the nucleus surface were plotted in Imaris Vantage; images showing a distance-coded coloring are shown for centromeres and telomeres.

like features. In type II nuclei, they distribute in the whole 3D nuclear space, which corresponds to a non-Rabl configuration. To support this observation by quantitative measurements, we exported three types of distance measurements (Figure 6f): (i) shortest distance between centromeric and telomeric spots (Figure 6f₁), (ii) shortest distance of centromeres and telomeres to the surface, corresponding to the nucleus border (Figure 6f₂), and (iii) average distance to top 5 neighboring spots for each group (centromeres, telomeres, Figure 6f₃). Because distances depend on the nucleus size, we normalized them using the nucleus diameter (graphs Figure 6f express relative distances). The quantitative analysis shown Figure 6f based on ca. 20 nuclei supported a contrasted spatial distribution of telomeres and centromeres in the two categories, with notably clear segregation of telomere and centromere groups in Type I nuclei (f1), located closely to the nuclear surface (f2). We also noticed a shorter distance between telomeres and centromeres in Type II nuclei, which was unexpected. Because this type of nuclei is frequently highly endoreduplicated, this led us to investigate further the relationship between ploidy and chromosomal organization (Nowicka, Pecinka et al., submitted).

A second example is shown using nuclei from *L. floccosa* subsp. *bellingneriana* (Figure 6g-j). Nuclei extraction from different types of tissue, FISH protocol and imaging were previously described by [39, 44-46]. A similar procedure was applied to segment the nucleus based on DAPI staining (grey) and FISH signal reporting on the telomeres (cyan), centromeres (magenta) and 35S rDNA loci (yellow) (Figure 6h-i). In this example, centromeres and telomeres clearly showed clustering towards the nuclear periphery as shown with a median distance of the spots around 0.6µm (TEL) to 1µm (CEN) in a nucleus of ca. 12 µm diameter (Figure 6j). We used this workflow for the analysis of nuclear organization in Crucifer genomes [39], in seven diploid species with up to 26-fold variation in genome size. This allowed to unveil species-specific patterns in nuclear organization [39].

For a trained user, the workflow takes approximately 20 minutes or less per image. This workflow can be used to compare the spatial distribution of chromosomes at interphase using centromeres and telomeres as references. Distance measurements across image replicates offer the possibility to detect quantitative differences invisible by the eye, between tissue types and cell types and to characterize potential mutant phenotypes in genetic analyses.

Analysis of mitotic chromosome orientation during division

Mitosis is the process by which organisms increase the number of cells. In plants, the highest number of mitotically active cells can be found in the root and shoot apical meristems (RAM and SAM, respectively; [47]). We focused the analysis on chromosome organization and orientation in living barley roots. Both cell division orientation and cell elongation contribute to the oriented growth of the root. Changes in the mitotic division orientation affects root shape and anatomy [48]. In *Vicia faba*, chromosome positioning correlates with the cell division plane and ultimately cell shape [49]. Notably, it was speculated that cell size could be a limiting factor forcing the spindle axis to be tilted, deviating slightly from the main axis of cell and organ elongation. Analyzing the orientation of mitotic chromosomes during cell division is thus relevant to understand this intricate relationship.

We designed a 3D microscopic image analysis workflow described in Figure 7a and detailed in Supplemental Files 7 containing a protocol and troubleshooting tips. We used barley chromatin and centromere fluorescent marker lines (FMLs) expressing translational fusions of histone H2B with CYAN FLUORESCENT PROTEIN (CFP-H2B) and ̢-CENTROMERIC HISTONE H3 with RED FLUORESCENT PROTEIN (RFP-CENH3), respectively (Kaduchová, Pecinka et al., in preparation). Z-stack images were acquired using a Leica TCS SP8 STED3X confocal microscope equipped with a Leica Application Suite X (LAS-X) software version 3.5.5 with a Leica Lightning module (Leica, Buffalo Grove, IL, USA) (Leica Microsystems, Wetzlar, Germany).

a Workflow 7 (overview)

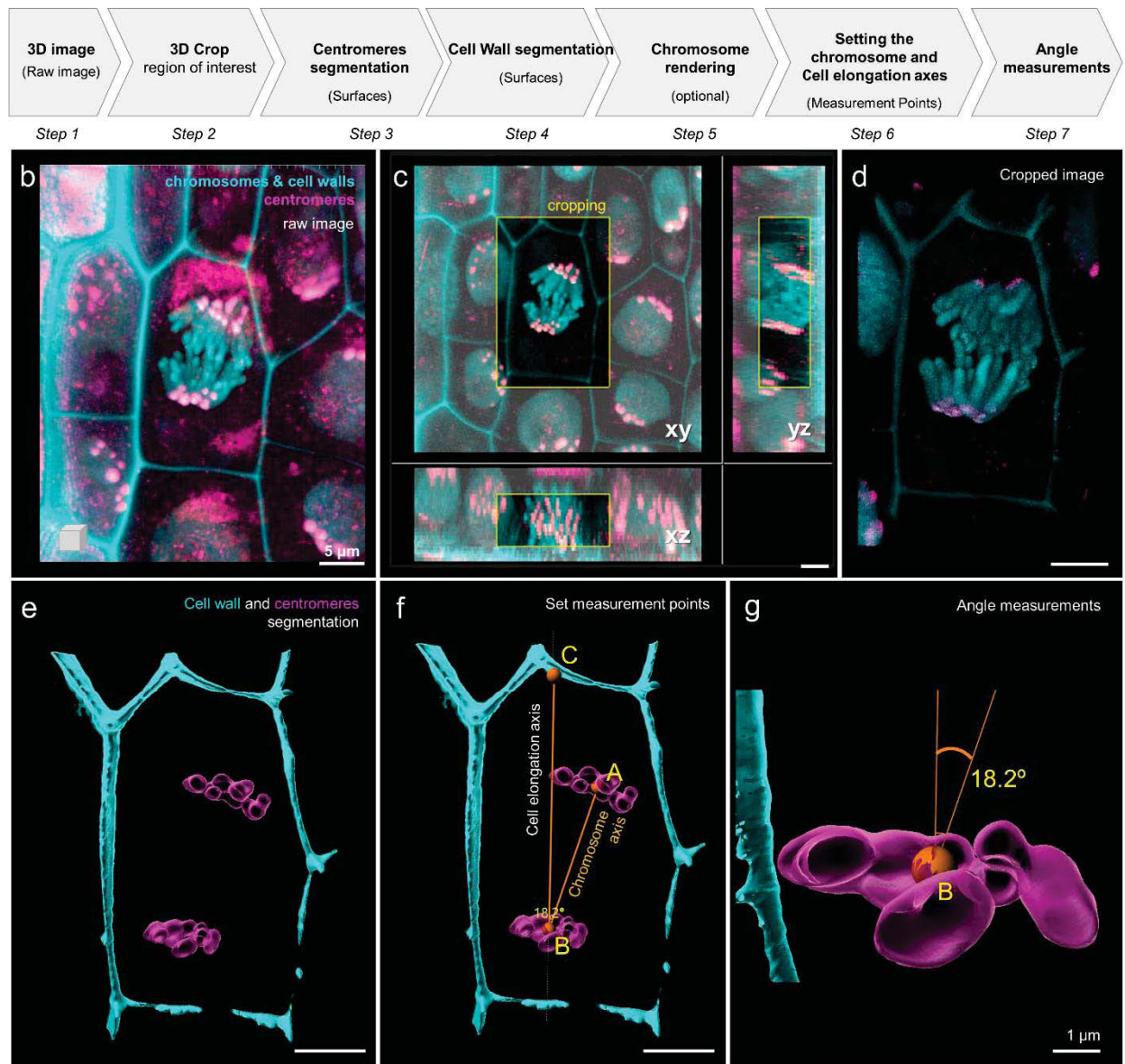


Figure 7. Division angle measurement using surface rendered cellular structures in living barley root cells.

(a) Workflow overview showing the sequence of tasks to process a raw picture up to the setting of measurement lines within the 3D cell space. (b) Confocal imaging of barley root tissue from a young seedling expressing CFP-H2B marking the chromosomes (cyan) and RFP-CenH3 marking the centromeres (magenta). In addition, cell wall autofluorescence upon UV excitation was used to mark the cell's boundaries (cyan). The image is a partial projection from a z-stack. (c) 3D cropping of the image to select a region of interest containing a dividing cell in mitotic anaphase (yellow frame). Orthogonal projections are shown in xy, yz and xz. (d) The cropped image in rendered in 3D using the 'blend' mode. (e) 'Surface' rendering of segmented centromeres (magenta) and the cell wall (cyan). (f) Setting of 'Measurement points' and their connective lines. AB defines the axis along which chromosomes are pulled (orthogonal to the chromosome plates), BC defines the cell's elongation axis. (g) Detailed visualization of the lower metaphase plate and angle formed between both axes defined by AB and BC measurement lines. The angle is measured in 3D by Imaris. Scale bars: b-f, 5 μm ; g, 1 μm

In addition, we took advantage of the fact that barley cell walls have an autofluorescence detectable in CFP emission spectra [50], allowing simultaneous visualization of chromosomes, centromeres, and cell walls. Centromere signals were pseudo-colored in magenta, chromatin with cell walls in cyan. The raw image (Figure 7b) presenting several cells in the root was cropped around one cell showing chromosomes at anaphase (Figure 7c). Centromeres were segmented with the 'spots' tool (Figure 7d). Chromosomes and cell walls were segmented using the 'surface' tool (Figure 7e). Using the tool 'Measurement point' we created spots (connected by a measurement line) at key positions, providing information on the cell elongation axis (A-B), on the pulling axis of the chromosomes (A'-B) and a reference axis (B-C) (Figure 7f, see supplemental File 7 for detailed explanation), which allowed for angle measurements (Figure 7g).

For a trained user, the workflow takes approximately 30 min per image. This workflow will allow one to measure the relationship between the orientation of the spindle axis during division and cell shape (elongation) and its variation between tissue and cell types. In addition, the possibility to measure this relationship opens the possibility to quantify the effect of genetic or environmental factors with large or subtle effects on the cell division axis.

Conclusive remarks

We present here a set of seven image analysis workflows enabling the quantitative study of the spatial organization of chromosomes and chromatin components. The workflows cover applications for studies at interphase (workflows 1, 2, 4, 6), mitosis (workflow 5, 7) or meiosis (workflow 3). Workflows 1,2 and 4 demonstrated the possibility to discover spatial distribution patterns, taking as examples transcription clusters, nuclear bodies and speckles and nuclear envelope-associated proteins. Such patterns were revealed thanks to the exploration of possible relationships between distance and intensity measurements among the different objects of the segmented images. Workflow 6, exploring genome organization at interphase, illustrates the quantitative power of

image segmentation to precisely measure the spatial positioning in the nuclear space, and the clustering of telomeres and centromeres. These features describe different types of 3D genome organization depending on cell type and species. The interest of performing image analysis for chromosome studies was further illustrated with workflows 3 and 5 focusing on condensed chromosomes at meiosis or mitosis, respectively. Workflow 1 demonstrates the usefulness of image segmentation for quantifying the number and distribution of cross-over components on meiotic chromosomes and revealing the possible enrichment in relation to the synaptonemal complex and that of chromosomal regions. Workflow 5 shows that volumetric measurements of FISH signals enable determining chromatin density (compaction) at different genomic regions. Finally, workflow 7 proposes an approach to measure the angles between chromosome and cell elongation axes, to investigate the relationship between cell division orientation and chromosomal positioning.

Image segmentation delivers a wealth of information related to signal intensity, distribution pattern (texture), shape, size and distance relationships between segmented objects [3]. Thus, images become associated with many variables and entry types, generating big data. Those can either be explored in a non-hypothesis-driven way using multidimensional data analysis (Bagheri et al., 2022) or in a hypothesis-driven manner following a careful choice of data for export. Even when exporting a selective number of image descriptors, the analysis of replicate image datasets, in different conditions (treatments or genotypes), labeled for multiple components, quickly generates a large numerical dataset. Versatile data visualization interfaces become handy at this stage. Here, we provided some examples among the numerous available solutions. We developed a customized shiny-based data visualisation interface, DataViz, for processing (normalising, filtering), exploring and plotting intensity, morphology and distance measurement data exported from segmented images. Normalisation of intensity or distance measurements per image is important for

considering variations that may arise between images during sample preparation, imaging or image acquisition [3, 8]. The examples provided here propose different strategies depending on the image analysis question. Versatile data visualisation greatly facilitates the explorative work, which in turn has the potential to seed discoveries, revealing unexpected patterns or relationships and driving further analyses or experiments.

Finally, though these workflows were developed to analyse nuclei and chromosome organization mostly in plant cells, these are conceptually applicable to nuclei of other species. An example is shown in workflow 4 with the analysis of nuclear bodies in mouse embryonic stem cells. Also, while based on a particular (commercial) software piece, they are in essence transferable to other concurrent, commercial or open-source software [51]. This compendium of workflows, providing illustrations, training images and detailed guidelines, aims at inspiring experimentalists in the field of life science with no or little expertise

in image processing. This effort responds to the rapid development of microscopy imaging techniques, and the need to render image analysis accessible to a wider community, to exploit image data to an unprecedented level of analysis.

Acknowledgments

We acknowledge the microscopy imaging facilities of the University of Zürich (Jana Doehner, Urs Ziegler; ZMB), of the CEITEC at the Masaryk University (CELLIM, supported by the MEYS CR project no. LM2018129), of the Auvergne Bioimaging platform (CLIC microscopy facility).

The authors acknowledge the 'INDEPTH' COST Action CA16212 for funding of the Training School during which the workflows were developed, Peter Majer and Michael Mahler (Bitplane AG) for suggestions during the training school.

References

1. Dumur T, Duncan S, Graumann K, Desset S, Randall RS, Scheid OM, et al. Probing the 3D architecture of the plant nucleus with microscopy approaches: challenges and solutions. *Nucleus* 2019; 10:181-212. doi: 10.1080/19491034.2019.1644592
2. Tatout C, Mougeot G, Parry G, Baroux C, Pradillo M, Evans D. The INDEPTH (Impact of Nuclear Domains On Gene Expression and Plant Traits) Academy - a community resource for plant science. *J Exp Bot* 2022. doi: 10.1093/jxb/erac005
3. Bagheri N, Carpenter AE, Lundberg E, Plant AL, Horwitz R. The new era of quantitative cell imaging-challenges and opportunities. *Mol Cell* 2022; 82:241-7. doi: 10.1016/j.molcel.2021.12.024
4. Jerkovic I, Cavalli G. Understanding 3D genome organization by multidisciplinary methods. *Nat Rev Mol Cell Biol* 2021; 22:511-28. doi: 10.1038/s41580-021-00362-w
5. Boettiger A, Murphy S. Advances in Chromatin Imaging at Kilobase-Scale Resolution. *Trends Genet* 2020; 36:273-87. doi: 10.1016/j.tig.2019.12.010
6. Dar AS, Padha D. Medical Image Segmentation: A Review of Recent Techniques, Advancements and a Comprehensive Comparison. *International Journal of Computer Sciences and Engineering* 2019; 7:114-24. doi: 10.26438/ijcse/v7i7.114124
7. Lucas AM, Ryder PV, Li B, Cimini BA, Eliceiri KW, Carpenter AE. Open-source deep-learning software for bioimage segmentation. *Mol Biol Cell* 2021; 32:823-9. doi: 10.1091/mbc.E20-10-0660
8. Caicedo JC, Cooper S, Heigwer F, Warchal S, Qiu P, Molnar C, et al. Data-analysis strategies for image-based cell profiling. *Nat Methods* 2017; 14:849-63. doi: 10.1038/nmeth.4397
9. Jamali N, Dobson ETA, Eliceiri KW, Carpenter AE, Cimini BA. 2020 BioImage Analysis Survey: Community experiences and needs for the future. *Biological Imaging* 2021; 1:e4. doi: 10.1017/S2633903X21000039

10. Cremer M, Schmid VJ, Kraus F, Markaki Y, Hellmann I, Maiser A, et al. Initial high-resolution microscopic mapping of active and inactive regulatory sequences proves non-random 3D arrangements in chromatin domain clusters. *Epigenetics Chromatin* 2017; 10:39. doi: 10.1186/s13072-017-0146-0
11. Girelli G, Custodio J, Kallas T, Agostini F, Wernersson E, Spanjaard B, et al. GPSeq reveals the radial organization of chromatin in the cell nucleus. *Nat Biotechnol* 2020; 38:1184-93. doi: 10.1038/s41587-020-0519-y
12. Schubert V. RNA polymerase II forms transcription networks in rye and Arabidopsis nuclei and its amount increases with endopolyploidy. *Cytogenet Genome Res* 2014; 143:69-77. doi: 10.1159/000365233
13. Schubert V, Weisshart K. Abundance and distribution of RNA polymerase II in Arabidopsis interphase nuclei. *J Exp Bot* 2015; 66:1687-98. doi: 10.1093/jxb/erv091
14. Ashenafi MS, Baroux C. Automated 3D Gene Position Analysis Using a Customized Imaris Plugin: XTFISHInsideNucleus. *Methods Mol Biol* 2018; 1675:591-613. doi: 10.1007/978-1-4939-7318-7_32
15. Andrey P, Kieu K, Kress C, Lehmann G, Tirichine L, Liu Z, et al. Statistical analysis of 3D images detects regular spatial distributions of centromeres and chromocenters in animal and plant nuclei. *PLoS Comput Biol* 2010; 6:e1000853. doi: 10.1371/journal.pcbi.1000853
16. Franz P, De Jong JH, Lysak M, Castiglione MR, Schubert I. Interphase chromosomes in Arabidopsis are organized as well defined chromocenters from which euchromatin loops emanate. *Proc Natl Acad Sci U S A* 2002; 99:14584-9. doi: 10.1073/pnas.212325299
17. Bucevicius J, Keller-Findeisen J, Gilat T, Hell SW, Lukinavicius G. Rhodamine-Hoechst positional isomers for highly efficient staining of heterochromatin. *Chem Sci* 2019; 10:1962-70. doi: 10.1039/c8sc05082a
18. Huff J. The Airyscan detector from ZEISS: confocal imaging with improved signal-to-noise ratio and super-resolution. *Nature Methods* 2015; 12:i-ii. doi: 10.1038/nmeth.f.388
19. Orr JN, Lewandowska D, Waugh R, Colas I. Chromosome Synapsis and Recombination in Cereals. *Annual Plant Reviews online*:1-26.
20. Keeney S. Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis. *Genome Dyn Stab* 2008; 2:81-123. doi: 10.1007/7050_2007_026
21. Bishop DK. RecA homologs Dmc1 and Rad51 interact to form multiple nuclear complexes prior to meiotic chromosome synapsis. *Cell* 1994; 79:1081-92. doi: 10.1016/0092-8674(94)90038-8
22. Higgins JD, Perry RM, Barakate A, Ramsay L, Waugh R, Halpin C, et al. Spatiotemporal asymmetry of the meiotic program underlies the predominantly distal distribution of meiotic crossovers in barley. *Plant Cell* 2012; 24:4096-109. doi: 10.1105/tpc.112.102483
23. Barakate A, Higgins JD, Vivera S, Stephens J, Perry RM, Ramsay L, et al. The synaptonemal complex protein ZYP1 is required for imposition of meiotic crossovers in barley. *Plant Cell* 2014; 26:729-40. doi: 10.1105/tpc.113.121269
24. Colas I, Barakate A, Macaulay M, Schreiber M, Stephens J, Vivera S, et al. desynaptic5 carries a spontaneous semi-dominant mutation affecting Disrupted Meiotic cDNA 1 in barley. *J Exp Bot* 2019; 70:2683-98. doi: 10.1093/jxb/erz080
25. Serra H, Lambing C, Griffin CH, Topp SD, Nageswaran DC, Underwood CJ, et al. Massive crossover elevation via combination of HEI10 and recq4a recq4b during Arabidopsis meiosis. *Proc Natl Acad Sci U S A* 2018; 115:2437-42. doi: 10.1073/pnas.1713071115
26. Hesse S, Zelkowski M, Mikhailova EI, Keijzer CJ, Houben A, Schubert V. Ultrastructure and Dynamics of Synaptonemal Complex Components During Meiotic Pairing and Synapsis of Standard (A) and Accessory (B) Rye Chromosomes. *Front Plant Sci* 2019; 10:773. doi: 10.3389/fpls.2019.00773
27. De Muyt A, Zhang L, Piolot T, Kleckner N, Espagne E, Zickler D. E3 ligase Hei10: a multifaceted structure-based signaling molecule with roles within and beyond meiosis. *Genes Dev* 2014; 28:1111-23. doi: 10.1101/gad.240408.114
28. Mao YS, Zhang B, Spector DL. Biogenesis and function of nuclear bodies. *Trends Genet* 2011; 27:295-306. doi: 10.1016/j.tig.2011.05.006
29. Ali GS, Golovkin M, Reddy AS. Nuclear localization and in vivo dynamics of a plant-specific serine/arginine-rich protein. *Plant J*

- 2003; 36:883-93. doi: 10.1046/j.1365-313x.2003.01932.x
30. Docquier S, Tillemans V, Deltour R, Motte P. Nuclear bodies and compartmentalization of pre-mRNA splicing factors in higher plants. *Chromosoma* 2004; 112:255-66. doi: 10.1007/s00412-003-0271-3
31. Reddy AS, Day IS, Gohring J, Barta A. Localization and dynamics of nuclear speckles in plants. *Plant Physiol* 2012; 158:67-77. doi: 10.1104/pp.111.186700
32. Lorkovic ZJ, Hilscher J, Barta A. Use of fluorescent protein tags to study nuclear organization of the spliceosomal machinery in transiently transformed living plant cells. *Mol Biol Cell* 2004; 15:3233-43. doi: 10.1091/mbc.e04-01-0055
33. Tosolini M, Brochard V, Adenot P, Chebrout M, Grillo G, Navia V, et al. Contrasting epigenetic states of heterochromatin in the different types of mouse pluripotent stem cells. *Sci Rep* 2018; 8:5776. doi: 10.1038/s41598-018-23822-4
34. Mikulski P, Hohenstatt ML, Farrona S, Smaczniak C, Stahl Y, Kalyanikrishna, et al. The Chromatin-Associated Protein PWO1 Interacts with Plant Nuclear Lamin-like Components to Regulate Nuclear Size. *Plant Cell* 2019; 31:1141-54. doi: 10.1105/tpc.18.00663
35. Kubalová I, Câmara AS, Cápál P, Beseda T, Rouillard J-M, Krause GM, et al. Helical metaphase chromatid coiling is conserved. *bioRxiv* 2021:2021.09.16.460607. doi: 10.1101/2021.09.16.460607
36. Bennett MD, Leitch IJ, Price HJ, Johnston JS. Comparisons with *Caenorhabditis* (approximately 100 Mb) and *Drosophila* (approximately 175 Mb) using flow cytometry show genome size in *Arabidopsis* to be approximately 157 Mb and thus approximately 25% larger than the *Arabidopsis* genome initiative estimate of approximately 125 Mb. *Ann Bot* 2003; 91:547-57. doi: 10.1093/aob/mcg057
37. Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, et al. A chromosome conformation capture ordered sequence of the barley genome. *Nature* 2017; 544:427-33. doi: 10.1038/nature22043
38. Pecinka A, Schubert V, Meister A, Kreth G, Klatte M, Lysak MA, et al. Chromosome territory arrangement and homologous pairing in nuclei of *Arabidopsis thaliana* are predominantly random except for NOR-bearing chromosomes. *Chromosoma* 2004; 113:258-69. doi: 10.1007/s00412-004-0316-2
39. Shan W, Kubova M, Mandakova T, Lysak MA. Nuclear organization in crucifer genomes: nucleolus-associated telomere clustering is not a universal interphase configuration in Brassicaceae. *Plant J* 2021; 108:528-40. doi: 10.1111/tpj.15459
40. Santos AP, Shaw P. Interphase chromosomes and the Rab1 configuration: does genome size matter? *J Microsc* 2004; 214:201-6. doi: 10.1111/j.0022-2720.2004.01324.x
41. Nowicka A, Kovacik M, Tokarz B, Vrana J, Zhang Y, Weigt D, et al. Dynamics of endoreduplication in developing barley seeds. *J Exp Bot* 2021; 72:268-82. doi: 10.1093/jxb/eraa453
42. Hudakova S, Michalek W, Presting GG, ten Hoopen R, dos Santos K, Jasencakova Z, et al. Sequence organization of barley centromeres. *Nucleic Acids Res* 2001; 29:5029-35. doi: 10.1093/nar/29.24.5029
43. Richards EJ, Ausubel FM. Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* 1988; 53:127-36. doi: 10.1016/0092-8674(88)90494-1
44. Hurel A, Phillips D, Vrielynck N, Mezard C, Grelon M, Christophorou N. A cytological approach to studying meiotic recombination and chromosome dynamics in *Arabidopsis thaliana* male meiocytes in three dimensions. *Plant J* 2018; 95:385-96. doi: 10.1111/tpj.13942
45. Mandakova T, Lysak MA. Chromosome Preparation for Cytogenetic Analyses in *Arabidopsis*. *Curr Protoc Plant Biol* 2016; 1:43-51. doi: 10.1002/cppb.20009
46. Mandakova T, Lysak MA. Painting of *Arabidopsis* Chromosomes with Chromosome-Specific BAC Clones. *Curr Protoc Plant Biol* 2016; 1:359-71. doi: 10.1002/cppb.20022
47. Kirschner GK, Stahl Y, Von Korff M, Simon R. Unique and Conserved Features of the Barley Root Meristem. *Front Plant Sci* 2017; 8:1240. doi: 10.3389/fpls.2017.01240
48. Van Norman JM. Asymmetry and cell polarity in root development. *Dev Biol* 2016; 419:165-74. doi: 10.1016/j.ydbio.2016.07.009
49. Oud JL, Nanninga N. Cell shape, chromosome orientation and the position of the plane of

division in *Vicia faba* root cortex cells. *Journal of Cell Science* 1992; 103:847-55. doi: 10.1242/jcs.103.3.847

50. Donaldson L. Autofluorescence in Plants. *Molecules* 2020; 25. doi: 10.3390/molecules25102393

51. Levet F, Carpenter A, Eliceiri K, Kreshuk A, Bankhead P, Haase R. Developing open-source software for bioimage analysis: opportunities and challenges [version 1; peer review: 2 approved]. *F1000Research* 2021; 10. doi: 10.12688/f1000research.52531.1

Résumé

Les eucaryotes organisent et protègent leur génome au sein du noyau, un organite délimité par une enveloppe nucléaire ponctuée par des pores nucléaires qui permettent les échanges nucléocytoplasmiques, et en dessous de laquelle se localise le nucléosquelette, une structure filamenteuse en contact avec le matériel génétique. Enveloppe nucléaire, pores nucléaires et nucléosquelette constituent non seulement des éléments structuraux qui participent à l'établissement de la morphologie du noyau mais sont également suggérés comme constituant la périphérie nucléaire, un domaine nucléaire qui régule l'expression du génome.

Dans ce travail, nous avons utilisé la plante modèle *Arabidopsis thaliana* pour étudier KAKU4 et CROWDED NUCLEI 1-4 (CRWN) qui sont les principales protéines entrant dans la composition du nucléosquelette des plantes. En cherchant à déterminer comment le nucléosquelette se localisait à la périphérie nucléaire, nous avons identifié de courts motifs protéiques dans la région N-terminale de KAKU4 qui sont également retrouvés dans NUP82 et NUP136, deux nucléoporines appartenant au pore nucléaire. Par des méthodes de double hybride, d'expression transitoire et de complémentation fonctionnelle, j'ai montré que ces motifs sont nécessaires à l'interaction physique entre le pore nucléaire et le nucléosquelette et permettent la localisation du nucléosquelette à la périphérie nucléaire. De plus, une analyse phylogénétique suggère que NUP82, NUP136 et KAKU4 dériveraient d'une nucléoporine ancestrale commune à partir de laquelle KAKU4 aurait divergé au cours de l'évolution puis acquis une nouvelle fonction associée au nucléosquelette. J'ai étendu ces liens évolutifs à la protéine CRWN1 qui possède le même motif protéique que KAKU4 que nous avons appelé *peripheral localisation signal* (PLS) et qui détermine pour ces deux protéines l'enrichissement à la périphérie nucléaire. Ces liens évolutifs étroits entre acteurs de la périphérie nucléaire a été renforcé par la création d'une collection de multi-mutants dont la caractérisation met en évidence l'existence d'interactions génétiques fortes entre le pore nucléaire et le nucléosquelette. Ainsi, les mutations *nup136* et *crwn1* induisent des effets synergiques qui se traduisent par des défauts sévères dans le développement de la plante au stade végétatif et reproductif et une dérégulation forte de gènes impliqués dans la réponse aux stress biotiques et abiotiques, soulignant le rôle essentiel de l'acide salicylique dans les anomalies de développement de nos mutants.

L'ensemble de ces travaux mettent en évidence une origine évolutive et des fonctions communes entre plusieurs acteurs de la périphérie nucléaire et leur rôle central dans l'intégration des signaux de réponse de la plante lors de stress biotiques ou abiotiques.