OXFORD BROOKES UNIVERSITY 2019
UNIVERSITE CLERMONT AUVERGNE

Ecole Doctorale
Sciences de la Vie, Santé, Agronomie, Environnement

Thèse

Présentée à l'Université Blaise Pascal pour l'obtention du grade de
A thesis submitted in partial fulfilment of the requirements of Oxford Brookes University for the degree of

DOCTOR OF PHILOSOPHY
DOCTEUR D'UNIVERSITE
Spécialité : Physiologie et Génétique Moléculaires

Characterisation of Nuclear Envelope-Associated Proteins (NEAPs) in *Arabidopsis thaliana*

Gwénaëlle DETOURNE

Soutenue le 29 Mai 2019

President of the Jury:
Dr Paul Fransz
Jury members:
Dr Verena Kriechbaumer
Dr Monica Pradillo
Director of Studies:
Professor David Evans
Professor Christophe Tatout

Génétique, Reproduction et Développement (GReD) – UMR CNRS 6293, INSERM U1103
Université Clermont Auvergne
28, Place Henri Dunant
63001 Clermont-Ferrand

Oxford Brookes University
Department of Biological and Medical Sciences
Headington Campus
OX3 0BP
Oxford UK

RESUME

Au cours de l'évolution, les cellules eucaryotes ont acquis une enveloppe nucléaire (NE) renfermant et protégeant le génome organisé en chromatine, une structure où l'ADN s'enroule autour de protéines histones. La NE est composé de deux membranes: du côté nucléoplasmique, la membrane nucléaire interne (INM) et du côté cytoplasmique, la membrane nucléaire externe. La NE permet la communication entre les deux compartiments par le biais des complexes de pores nucléaires et relie le cytosquelette au nucléosquelette via le complexe LINC (LInker of Nucleoskeleton to Cytoskeleton). Ainsi, le nucléosquelette associé à l'INM est nécessaire pour transmettre des signaux au noyau et induire des changements dans l'organisation de la chromatine et finalement dans l'expression des gènes.

Une nouvelle famille de protéines associées à l'enveloppe nucléaire (NEAP), proposées comme nouveaux composants du nucléosquelette de la plante, a récemment été mise en évidence dans la plante modèle *Arabidopsis thaliana*. Ces protéines sont codées par une famille de trois gènes et sont ciblées vers le noyau via un NLS où elles sont ancrées à l'INM via leur domaine transmembranaire C-terminal. Les protéines AtNEAPs possèdent également plusieurs longs domaines en spirale (coiled-coil) rappelant la structure des lamines chez les animaux. Cette thèse visait à réaliser une analyse fonctionnelle des AtNEAPs à l'aide de lignées mutantes T-DNA et CRISPR/Cas9. L'interactome AtNEAP a été étudié par des approches moléculaires (Yeast Two Hybrid), indiquant des interactions entre AtNEAPs pouvant former des homo- ou hétéro-dimères; ainsi que la localisation et la co-localisation *in vivo* couplées à de l'imagerie (apFRET), qui ont confirmé les interactions avec le facteur de transcription (TF) AtbZIP18. Les anticorps spécifiques à AtNEAP générés au cours de cette étude ont été utilisés pour confirmer l'expression *in vivo*. En outre, les résultats ont indiqué que les AtNEAPs font partie du nucléosquelette et jouent un rôle dans l'ancrage des TF à l'INM afin de maintenir la morphologie nucléaire et l'organisation de la chromatine.

Mots-clés : Périphérie nucléaire, NEAP, nucléosquelette, organisation de la chromatine, CRISPR/Cas9, bZIP18.

ABSTRACT

During evolution, eukaryotic cells have acquired a nuclear envelope (NE) enclosing and protecting the genome, which is organized in chromatin, a structure wrapping DNA around histone proteins. The NE is composed of two membranes: on the nucleoplasmic side, the Inner Nuclear Membrane (INM) and on the cytoplasmic side, the Outer Nuclear Membrane. The NE allows communication between both compartments through Nuclear Pore Complexes and bridges the cytoskeleton to the nucleoskeleton through the LInker of Nucleoskeleton to Cytoskeleton complex. Thus, the nucleoskeleton associated with the INM is needed to transmit signals to the nucleus and induce changes in chromatin organisation and ultimately gene expression.

A novel family of NUCLEAR ENVELOPE ASSOCIATED PROTEINS (NEAPs) proposed to be new components of the plant nucleoskeleton has been recently evidenced in the model plant *Arabidopsis thaliana*. AtNEAP proteins are encoded by a small gene family composed of three genes and are targeted through a nuclear localisation signal to the nucleus where they are anchored at the INM through their C-terminal transmembrane domain. AtNEAPs also possess several long coiled-coil domains reminiscent of the lamin structure in animals. This thesis aimed at performing a functional analysis of AtNEAPs using T-DNA insertion and CRISPR/Cas9 mutant lines. The AtNEAP interactome was investigated by molecular approaches (Yeast Two Hybrid), which indicated AtNEAP interactions with each other to form homo or hetero-dimers; as well as *in vivo* localisation and co-localisation coupled to image analyses (apFRET, acceptor photobleaching Fluorescence Resonance Energy Transfer), which confirmed interactions with the transcription factor (TF) AtbZIP18. AtNEAP specific antibodies generated during this study were used to confirm expression *in vivo*. Altogether, results indicated that AtNEAPs are part of the nucleoskeleton, with a role in anchoring TFs at the INM to maintain nuclear morphology and chromatin organisation.

Key words: Nuclear periphery, NEAP, nucleoskeleton, chromatin organisation, CRISPR/Cas9, bZIP18.

ACKNOWLEDGMENTS

First of all, I would like to thank Professor David Evans and Professor Christophe Tatout for having designed this PhD project in 2015 and I would like to thank the Postgraduate Research tutors who gave me the opportunity to work on this project, Dr Susan Brooks and regretted Professor Chris Hawes. I would like to thank all the members of the jury, Dr Verena Kriechbaumer, Dr Paul Fransz and Dr Monica Pradillo for having accepted to be reporters and for paying attention to the whole work achieved during this PhD.

A special thank you to all my supervisors, David, Christophe, Dr Emmanuel Vanrobays and Dr Katja Graumann, without forgetting Dr Aline Probst for everything you have done for me during this PhD. Thank you Christophe and Manu for having raised in me a huge interest in genetics and molecular biology and for the knowledge transmission through your teaching when I was a young student. Many thanks, David and Katja, for all the scientific approach but also for your warmful welcome in Oxford and for your strong personal support when I needed it the most…

Thank's to all the colleagues I worked with through all those years ; in Clermont-Ferrand, Sylvie, Sylviane, Sam, Sophie, Maxime, Lauriane, Cyril, Olivier, Simon, Aurélia and all the plant GReDins without forgetting JC, Claude, Ayhan, Cyril and Mounir for sport support, Fabrice, Rachel, Isa, Joëlle who first let me step into GReD life ; but also in Oxford, Chris, Joe, Francès, Verena, Vidya, Vanessa, John, Jirke, Sue and all the plant scientists, without forgetting Margaret, Dan and Iris. Your patience with me, kindness, support and fruitful exchanges were very helpful.

All this scientific work would not have been possible without the technical help and support of l'équipe administrative du GReD, Marie-Jo, Maryse, Nadia, Marie-Diane and the Research Degrees Team, Catherine Joyejob and Jill Organ, and his previous Administrator, Philip Voysey. So many thanks to all of them.

Un grand merci à ma famille et mes amis pour leur soutien tout au long de la thèse. Mes parents, premièrement, pour leur soutien sans faille quoiqu'il advint et leur aide dans mes nombreux déménagements entre autres (que de voyages !). Mon frère, Florian, ma grand-mère Marie-Louise qui me demandait régulièrement comment allaient mes petites plantes, mes grands-parents Denise et Jean-Claude et tous mes oncles, tantes et cousin(e)s ; ainsi que mes amis, Hélène, Carine, David, Thibaut, Aude, Léo, Marine, Géraldine et TimTim sans oublier Martin. Un merci tout particulier à Fred pour m'avoir soutenue dans les moments les plus difficiles et m'avoir ainsi permis d'aller jusqu'au bout…

A mon Papa, et mes Grands-Pères

CONTENT

ABBREVIATIONS

aa: amino acids

Ade: Adenine

apFRET: acceptor photobleaching Fluorescence Resonance Energy Transfer

APS: Ammonium Persulfate

At: *Arabidopsis thaliana*

BAF: Barrier to Autointegration Factor

BRLZ: Basic DB region leucine zipper

BSA: Bovine Serum Albumine

bZIP: basic-leucine zipper

CC: Coiled-coil

CFP: Cyan Fluorescent Protein

CHD: Chromodomain-Helicase-DNA-binding protein

ChIP: Chromatine Immuno-Precipitation

CRWN: Crowded Nuclei

C-ter: C-terminus

DBD: DNA-binding domain

DMSO: Dimethyl sulfoxide

dNTPs: desoxyribonucleotides tri-phosphates

DSB: DNA Double Strand Break

EAR: Ethylen-responsive element binding factor-associated Amphiphilic Repression

EDMD: Emery-Dreifuss Muscular Dystrophy

EF: FRET efficiency

EP: Electroporation

ER: Endoplasmic Reticulum

FAST: Fast Arabidopsis-seedlings mediated Transformation

FG: phenylalanine-glycine

F*ISH*: Fluorescence *In Situ* Hybridization

FRAP: Fluorescence Recovery After Photobleaching

GAR: Glycine-ARginine

GDP: Guanosine TriPhosphate

GFP: Green Fluorescent Protein

GTP: Guanosine DiPhosphate

HDAC: Histone DeACetylase

HGPS: Hutchinson-Gilford Progeria Syndrome

His: Histidine

HS: Heat Shock

IB: Infiltration Buffer

IF: Intermediate Filament

IP: Immunoprecipitation

INM: Inner Nuclear Membrane

Ino80: Inositol-requiring 80

ISWI: Imitation Switch

KASH: Klarsicht, ANC-1, Syne Homology

KD: Knock-Down

KO: Knock-Out

LADs: Lamin Associated Domains

LB: Luria Broth

LBR: Lamin B Receptor

LEM: LAP2, Emerin, MAN1

Leu: Leucine

LINC: LInker of Nucleoskeleton to Cytoskeleton

LINC: Little Nuclei

MARs: Matrix Attachment Regions

MS: Murashige and Skoog

MYTH: Membrane Yeast Two Hybrid

NADs: Nucleolus Associated Domains

NEAP: Nuclear Envelope-Associated Protein

NE: Nuclear Envelope

NEBD: NE BreakDown

NES: Nuclear Export Signal

NHEJ: Non-Homologous End Joining

NLS: Nuclear Localisation Signal

NMCP: Nuclear Matrix Constituent Protein

NPC: Nuclear Pore Complex

N-ter: N-terminus

NTF2: Nuclear Transport Factor 2

Nups: Nucleoporins

OD: Optical Density

ONM: Outer Nuclear Membrane

PBST: Phosphate Buffer Saline - Tween

PEG: Polyethylene Glycol

PI: Pre-Immune

PIC: Protease Inhibitor Cocktail

PTM: Post-Translational Modification

q-RT-PCR: quantitative Reverse Transcription - Polymerase Chain Reaction

RanGAP: RanGTPase Activating Protein

RanGEF: Ran Guanine nucleotide Exchange Factor

ROI: Region Of Interest

Rpm: rotations per minute

RT: Room Temperature

AtSAP18: SIN3 ASSOCIATED POLYPEPTIDE P18

SD: Synthetically Defined

SDS – PAGE: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

SINE: SUN-Interacting Nuclear Envelope

SIOX: Simple Interactive Object Extraction

SOC: Super Optimal broth with Catabolite repression

SUN: Sad1p/UNC-84

SW12/SNF2: SWItch/Sucrose NonFermentable

TCA: Trichloro-acetic acid

TE/LiAc: Tris EDTA/LithiumAcetate

TEMED: N,N,N,N-Tétraméthyléthylènediamine

TIK: Toll-Interleukin-Resistance KASH

TM: TransMembrane

TPL: TOPLESS

Trp: Tryptophan

TSI: Transcriptionally Silent Information

Y2H: Yeast Two Hybrid

YEB: Yeast Extract Broth

YNB: Yeast Nitrogen Base

YFP: Yellow Fluorescent Protein

YPD: Yeast extract Peptone Dextrose/D-glucose

WIP: WPP domain-Interacting Protein

WIT: WPP domain-Interacting Tail-anchored protein

WPP: tryptophan-proline-proline

LIST of FIGURES

LIST of TABLES

CHAPTER 1
INTRODUCTION

INTRODUCTION

One of the main differences between prokaryotes and eukaryotes is the presence of a nucleus, which allows the packaging of DNA into a specialized compartment and isolates it from the cytoplasm. Separation is made by two phospholipid bilayers forming an Inner Nuclear Membrane (INM) and an Outer Nuclear Membrane (ONM), which are the continuity of the endoplasmic reticulum (ER), an essential compartment for protein maturation.

The nucleus is mobile and its migration occurs principally through nucleo- and cytoskeleton interactions, thanks to properties of the Nuclear Envelope (NE), (Tamura and Hara-Nishimura, 2013; Zhou and Meier, 2014). During cell division, the NE is necessarily disrupted by a mechanism called NE Break Down (NEBD), a process in which NE and nucleoskeleton components are implicated (Murphy *et al*., 2010; Smoyer and Jaspersen, 2014). The nucleus structure plays fundamental roles for the cell, such as in stress responses, cell development or even reproduction and has to be dynamic in order to adopt various shapes and to regulate gene expression (Ungricht and Kutay, 2017; Yang *et al*., 2017; Zhou and Meier, 2014). This involves the interaction of structural components of the nucleus including the envelope and nucleoskeleton with specialised genomic regions (Pombo and Dillon, 2015). Thus, the main characteristics of nuclei, depending on cell type, are due to envelope and periphery components interacting with special genomic regions.

The work in this thesis was carried out to further characterize proteins localized at the nuclear periphery in *Arabidopsis thaliana* (*A. thaliana*), specifically the NUCLEAR ENVELOPE ASSOCIATED PROTEINS (AtNEAP), a protein family with suggested roles in organising nuclear shape and chromatin, being part of the nuclear periphery protein network, (Pawar *et al*., 2016). Before introducing this family in detail by describing previous work on AtNEAPs, the genomic organization of the nucleus and the properties of the nuclear envelope

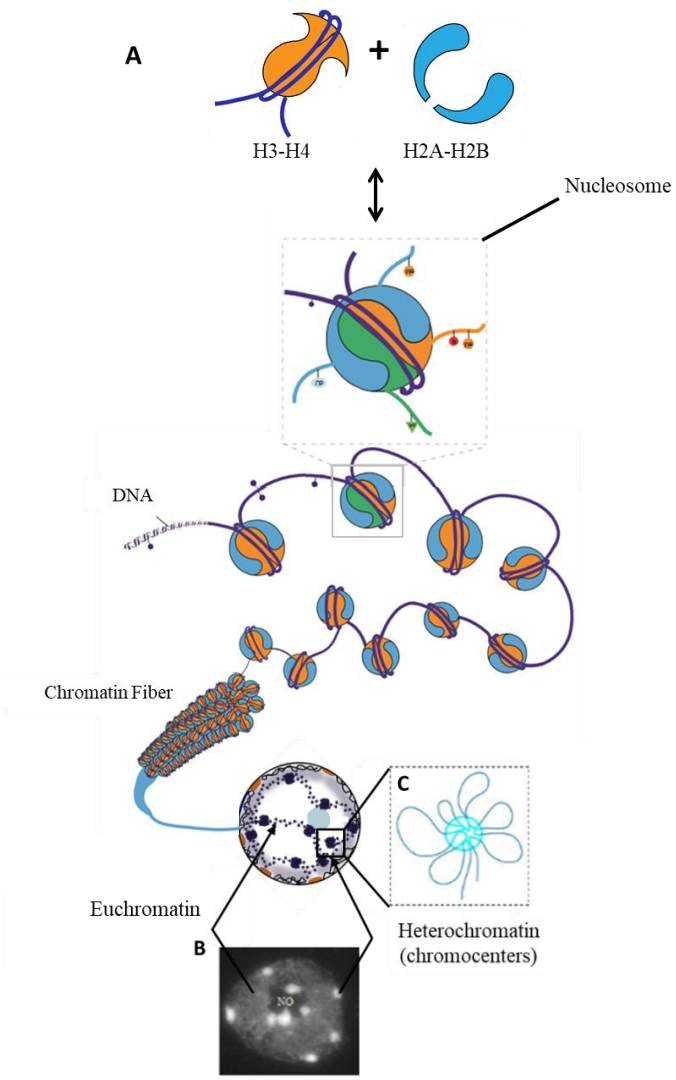


Figure 1-1: Chromatin formation and organisation (adapted from Probst et al., 2009)**. A.** The basic unit of chromatin is the nucleosome composed of histones H2A, H2B, H3, and H4 and 146bp of DNA. The beads of the string organization of nucleosomes folds into higher order chromatin structures. NO: nucleolus. **B.** Picture of an *Arabidopsis thaliana* nucleus stained with DAPI. Heterochromatin which is visible in microscopy in *A. thaliana* as bright foci called chromocentres is the most condensed chromatin state while euchromatin observed as a light grey background is the more relaxed chromatin state. **C**. Representation of the rosette organisation of chromatin. Heterochormatic sequences (in blue green) cluster together in chromocentre structures, while euchromatic sequences form chromatin loops anchored at the chromocentre, (Fransz et al., 2002, 2003, 2006; Grob et al., 2014, Feng et al., 2014).

and nuclear periphery components will be considered. Then, the aims of this PhD will be presented.

I – Chromatin formation and structure

In eukaryotes, the genomic DNA contained in the nucleus is organized into chromatin by association with histone proteins. A 146 base pair DNA sequence is wrapped around an octamer of histones forming the nucleosome, (Luger *et al*., 1997). This structure represents the basic unit of chromatin, ***Figure 1-1***.

There are five types of histones: H2A, H2B, H3, H4 and the histone linker H1. Every nucleosome is composed of a $(H3\text{-}H4)_2$ tetramer and two H2A-H2B dimers. Histone H1 allows the linkage of nucleosomes with each other for a higher level of compaction (Bharath *et al*., 2002; Rutowicz *et al*., 2018). Every type of histone belongs to a multigene family and, except for histone H4, there are one or several variants, which are non-allelic isoforms of canonical histones (Talbert *et al*., 2012). Incorporation of these variants influences nucleosome stability, DNA accessibility and, thus, gene expression. Indeed, particular variants contribute to specialized functions like centromere organization, silencing of transposable elements and repetitive sequences, X chromosome inactivation, specific gene activation or genomic stability, (Henikoff, 2008; Okada *et al*., 2005).

Post-translational modifications (PTM), mostly at the N-terminus (N-ter) of histones but also in their core domain are involved in *fine-tuning* gene expression. Histone PTMs can either directly affect the stability of the nucleosome or histone-DNA interaction or are interpreted by histone reader proteins that bind with specificity to certain histone PTMs translating this information into a more transcriptionally repressive or permissive chromatin environment.

The best-described PTMs are acetylation, methylation, phosphorylation, ubiquitinylation and sumoylation, (Bradbury, 1992; Desrosiers and Tanguay, 1986; Imhof and Becker, 2001; Maison *et al*., 2011; Probst *et al*., 2009; Strahl and Allis, 2000; Wu *et al*., 1986). The different combinations of PTMs together with the different canonical or variant histones within a nucleosome specify the epigenetic information (Jenuwein and Allis, 2001). Therefore, histone composition of nucleosomes and PTMs carried by histones modulate DNA accessibility and, in a higher scale, chromatin organization inside the nucleus.

During replication or transcription, the transcriptional machinery requires access to the underlying DNA. For that purpose, nucleosomes can slide along DNA or be fully or partially removed due to the action of chromatin remodellers. To date, four protein families are implicated in chromatin remodelling: SWITCH/SUCROSE NON FERMENTABLE (SWI/SNF), IMITATION SWITCH (ISWI), CHROMODOMAIN-HELICASE-DNA-BINDING PROTEIN (CHD), AND INOSITOL-REQUIRING 80 (INO80) (Clapier and Cairns, 2009; Henikoff, 2008; Petesch and Lis, 2012).

Classically, two different states of chromatin are distinguished according to their compaction levels (Heitz, 1928). The most condensed state is heterochromatin that shows high nucleosomal occupancy and enrichment in epigenetic marks repressive for transcription (Chodavarapu *et al*., 2010; Ricci *et al*., 2015) restricting access to DNA and limiting gene expression. The more decondensed chromatin, with lower nucleosomal occupancy, is euchromatin. It is more favourable for gene expression. However, these two principal chromatin states identified by cytological (Heitz, 1928), and molecular (Elgin and Grewal, 2003) analysis, have been further refined. Up to four main different states that result from combinations of several histone and DNA post-translational modifications, histone variants, transposable element composition and gene expression level were defined (Roudier *et al*., 2009, 2011; Sequeira-Mendes *et al*., 2014). In addition, it has been evidenced that NE and

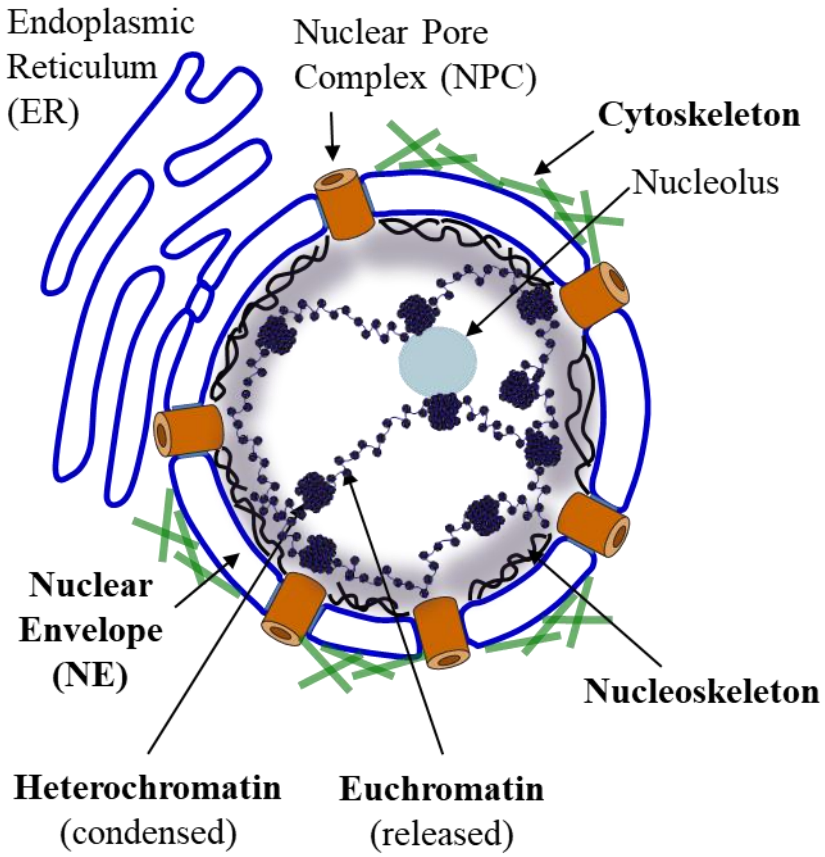


Figure 1-2: Nucleus organisation and structure. In eukaryotes, chromatin is separated from the cytoplasm by the nuclear envelope (NE) which is in continuity with the endoplasmic reticulum (ER). Throughout this membrane there are thousands of nuclear pore complexes (NPC), which allow communication between cytoplasm (outside) and nucleoplasm (inside). Also, on each side of the NE, cytoskeleton (green bars) and nucleoskeleton (black helix) made of protein complexes are responsible for keeping nucleus integrity and chromatin organisation. The two main states of chromatin (eu- and heterochromatin) are illustrated as arrays of relaxed or condensed nucleosomes (blacks beads).

nuclear periphery components have a role in structure and positioning of specific chromatin states within the nucleus, (Mattout *et al*., 2006; Murphy *et al*., 2010; Starr, 2009).

II – Nuclear envelope components

As presented previously, the NE is composed of two membranes: one on the nucleoplasmic side, INM, and another on the cytoplasmic side, ONM. The NE delimits the cytoplasm from the nucleoplasm and has many other functions, which have been well studied in animals. It allows communication between both compartments through Nuclear Pore Complexes (NPC) and bridges the cytoskeleton and the nucleoskeleton through the LInker of Nucleoskeleton to Cytoskeleton (LINC) complex, which permits nuclear migration and participates in maintaining nuclear shape and structure (Burke, 2012; Meier, 2001; Méjat and Misteli, 2010; Rose *et al*., 2004), ***Figure 1-2***.

II.1 – The Nuclear Pore Complex (NPC)

NPCs are embedded at sites of fusion between the ONM and INM and allow the NE to be permeable to a variety of macromolecules and signals. NPCs are ring-shape channels principally responsible for nucleo-cytoplasmic transport. The structure is octagonally symmetrical to its cylindrical axis and is composed of five different protein classes. The transmembrane ring is in contact with the NE, the core scaffold is made of an outer ring, an inner ring and a linker; cytoplasmic filaments, nuclear basket and central FG (hydrophobic core phenylalanine-glycine rich) are completing the complex, ***Figure 1-3*** (Alber *et al*., 2007; Brohawn *et al*., 2009; Grossman *et al*., 2012; Tamura and Hara-Nishimura, 2011, 2013). Hence, NPCs are formed by a large complex of proteins, with around 30 nucleoporins (Nups), which are highly conserved between vertebrates, yeast and plants (DeGrasse *et al*., 2009; Tamura *et al*., 2010). Despite a conserved structure, NPCs show some differences depending on the kingdom. Plant NPCs are 100 MDa in size, while vertebrates are 120 MDa and yeast

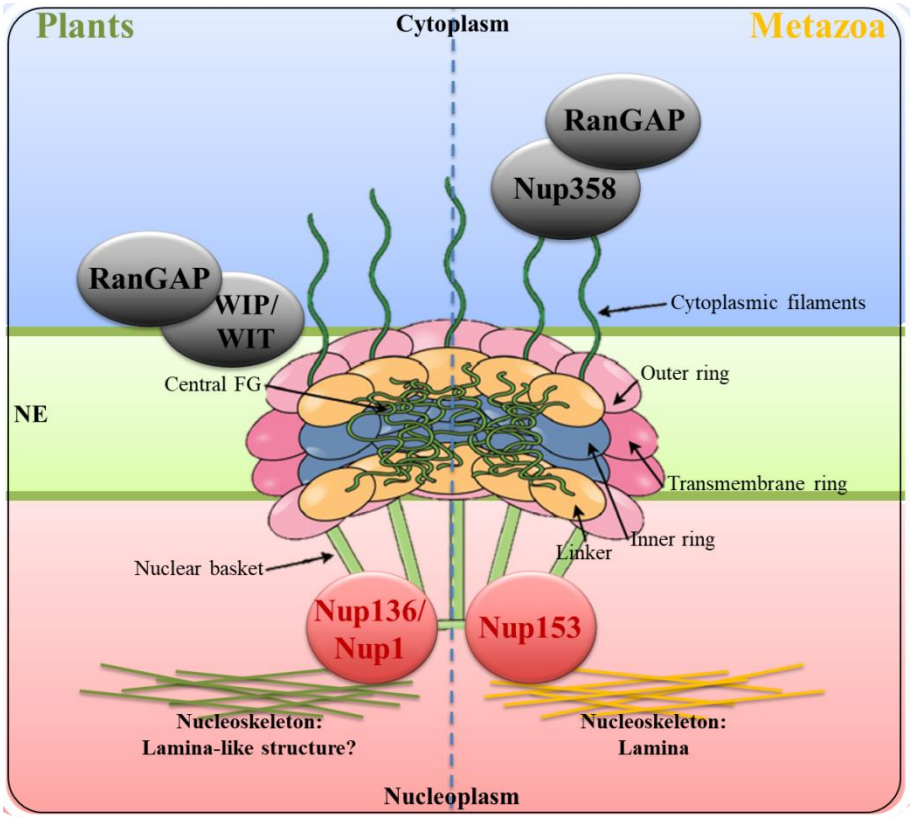


Figure 1-3: Schematic representation of the Nuclear Pore Complex (NPC) structure and main differences between plant and metazoan NPC components, (adapted from Tamura and Hara-Nishimura, 2013). The large complex of around 30 nucleoporins composing the NPC is well conserved through kingdoms. Major differences between plants and metazoan is the substitution of Nup358 by a WIT/WIP complex to anchor RanGAP close to the NPC for nucleo-cytoplasmic transport. Also, anchoring of the NPC to the nucleoskeleton here indicated as lamina-like structure / lamina by Nup153 in metazoa is replaced by Nup136/Nup1 in plants, (Zhou et al., 2012; Zhou & Meier, 2014; Tamura et al., 2010; Tamura and Hara-Nishimura, 2011).

only 50 MDa. Plants lack homologues to seven vertebrate proteins, of which one, Nup358 is important for anchoring RanGAP (RanGTPase Activating Protein, important for the nucleo-cytoplasmic transport (Ran cycle)) to the NE, (Hutten *et al*., 2008; Xu *et al*., 2007), and is substituted by a WIT/WIP complex in plants, (see ***II.2-b***), (Zhou and Meier, 2014; Zhou *et al*., 2012). Another is Nup153 anchoring the NPC to the nucleoskeleton in vertebrates, suggested to be replaced by Nup136/Nup1 in plants, which interacts dynamically with NPC at the NE, (Tamura and Hara-Nishimura, 2011; Tamura *et al*., 2010), ***Figure 1-3***.

II.1-a – Nucleo-cytoplasmic transport

Nucleo-cytoplasmic transport can be achieved by passive diffusion for small molecules like ions or small proteins below 50 kDa (Macara, 2001), but NPCs control the transit of macromolecules such as bigger proteins, ribosomes, RNA or RNA polymerases, by active transport, thanks to specific chaperones, named karyopherins, (Macara, 2001). Those chaperones are called importins or exportins depending on the direction of transport and have antagonist functions in the Ran cycle. Importins recognise a NLS (Nuclear Localization Signal) tag present on cargo proteins for transport into the nucleus and exportins recognise a NES (Nuclear Export Signal) for export from the nucleus, (Tran *et al*., 2014). The complex formed binds FG rich nucleoporins and is translocated between compartments. In the nucleoplasm, the importin-cargo complex is recognised by RanGTP, which changes importin conformation and releases cargo protein from importin. Importin-RanGTP is then translocated into the cytoplasm where RanGAP hydrolyses RanGTP in RanGDP and releases importin. To avoid any depletion of RanGTP inside the nucleus, RanGDP is brought back into the nucleus by its own import carrier, Nuclear Transport Factor 2 (NTF2), found in mammals, yeast and plants, (Macara, 2001; Zhao *et al*., 2006). RanGDP inside the nucleus is recycled to RanGTP by RanGEF (Ran Guanine nucleotide Exchange Factor) ready to disrupt new importin-cargo complexes. In this way, RanGAP and RanGDP are found only in the cytoplasm and RanGEF

and RanGTP only in the nucleoplasm, (Tran *et al*., 2014). The export mechanism is similar to import but this time RanGTP promotes and stabilizes exportin-cargo complexes. Once in the cytoplasm, RanGTP is hydrolysed to RanGDP by RanGAP and cargo protein is released from exportin. RanGDP and exportin are then brought back inside the nucleus by NTF2, (Macara, 2001; Zhao *et al*., 2006).

II.1-b – Association with INM and chromatin

Plant NPCs are non-randomly distributed over the NE. In vertebrates, NPC anchorage and position correlates with proteins of the inner nuclear periphery called lamins (see ***III.1***). This corresponds to observations in tobacco where NPC are anchored by a filamentous structure at the INM (Fiserova *et al*., 2009). Indeed, NPCs, more than acting as simple transport channels, have a role in pathogen response, (Gu *et al*., 2016), and many other functions. Among those, it has been shown that transcriptionally active and repressed genes associate with the NPC and the NE, (Dieppois and Stutz, 2010; Smith *et al*., 2015; Tran *et al*., 2014). Control of gene expression can be achieved by regulation of transcription factors or by modifying chromatin structure and accessibility to the transcriptional machinery, (see ***I***) (Capelson and Hetzer, 2009). Thus, NPCs have a transport-dependent role, for example, by importing transcription factors into the nucleus for specific gene activation, (Capelson and Hetzer, 2009), and also a transport-independent role in gene regulation, by tethering active regions of chromatin or recruiting actors of the transcription and mRNA export machineries, (Burns and Wente, 2014; Capelson and Hetzer, 2009; Tran *et al*., 2014).

Hence, NPCs are important to maintain nuclear shape, chromatin organisation, gene regulation, and together with the LINC complex (see ***II.2***), physically link the cytoskeleton to the nucleoskeleton and allow them to communicate.

II.2 – The LINC complex

The LINC complex spans the NE and bridges the cytoskeleton to the nucleoskeleton. In metazoa, this complex is composed of Sad1p/UNC-84 (SUN) domain proteins, which are localized in the INM, and of Klarsicht, ANC-1, Syne Homology (KASH) domain proteins, which are localized in the ONM. These two types of proteins interact in the periplasm via their C-terminal domains. On the nucleoplasmic side, SUN domain proteins interact with chromatin, lamins and other INM associated proteins via their N-terminal domain. On the cytoplasmic side, KASH domain proteins interact via their N-terminal domain with actin and different components of the cytoskeleton. In this way, communication between cytoplasm and chromatin is possible, which is important for many intra- and extracellular processes, (Starr, 2009). These include nuclear and chromosome positioning, cell division, and maintenance of nuclear shape, (Link *et al*., 2014). Indeed, defects in the LINC complex leads to diseases such as muscular dystrophy and progeria in humans (Burke, 2012; Méjat and Misteli, 2010; Tzur *et al*., 2006; Zhou and Meier, 2013).

The LINC complex is functionally conserved in eukaryotes, (Crisp *et al*., 2006; Graumann, 2014; Graumann *et al*., 2010; Murphy *et al*., 2010; Poulet *et al*., 2017). However, while SUN domain proteins are highly conserved across opisthokonts (a clade grouping metazoan and fungi), KASH domain proteins are more diverse in structure and function (Evans *et al*., 2014). The bridging complex has been described in humans, fly, worm, yeast and most recently in plants, (Meier, 2016; Tatout *et al*., 2014). In plants, it has been most studied in *Zea mays* (Murphy *et al*., 2010) and in *A. thaliana* (Graumann, 2014; Graumann *et al*., 2010; Meier, 2016). Further details will be presented below, concerning characteristics of SUN and KASH domain protein families, ***Figure 1-4***.

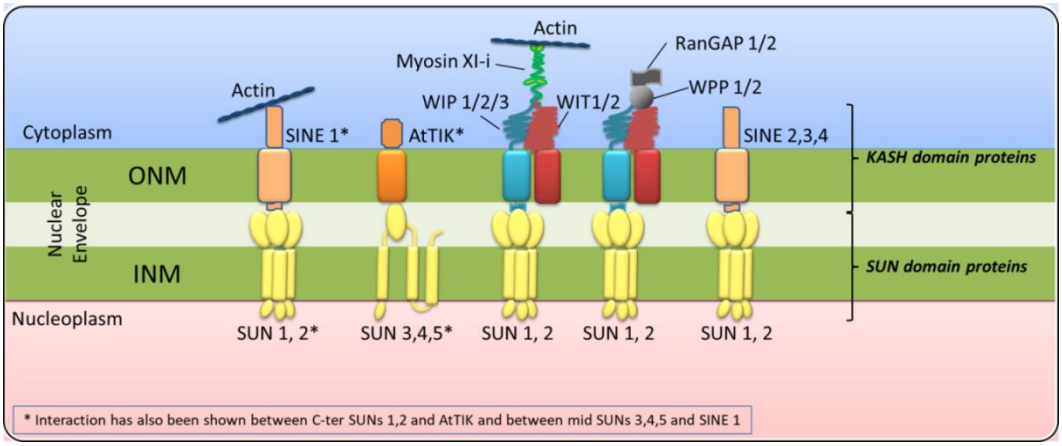


Figure 1-4: Different types of LInker of Nucleoskeleton and Cytoskeleton (LINC) complexes in *A. thaliana.* Due to a variety of Klarsicht-Anc1-Syne1 Homology (KASH) domain proteins in the Outer Nuclear Membrane (ONM) and the different Sad1-Unc84 (SUN) domain proteins in the Inner Nuclear Membrane (INM), several combinations of LINC complexes are possible and would be specific of cell types and developmental stages, (Meier, 2016).

*Interaction has also been shown between C-ter SUNs1-2 and AtTIK and between mid-SUNs3-5 and SINE1.

II.2-a – SUN domain protein family

Mammalian SUN domain proteins were found using bioinformatics analysis by comparison with SUN domain proteins found in the other kingdoms. The name comes from two proteins, which contain a C-terminal SUN domain: Sad1, a spindle pole body component described in yeast *Schizosaccharomyces pombe*, (Hagan and Yanagida, 1995) and UNC84, described in worm embryo, *Caenorhabditis elegans*, (Malone *et al.*, 1999). The SUN domain proteins are highly conserved in the different kingdoms and play a crucial role in cell survival. They are involved in nuclear migration, organization and shape determination, (Oda and Fukuda, 2011), chromosome and telomere positioning, cell cycle-dependent NEBD and NE reformation and in apoptosis, (Evans *et al*., 2014).

Plant SUN domain proteins were first described in *Z. mays* by (Murphy *et al*., 2010) and in *A. thaliana* by (Graumann *et al*., 2010). Computational methods indicated that the SUN domain protein family in *A. thaliana* is composed of five different proteins (from AtSUN1 to AtSUN5) and two subfamilies have been distinguished, (Graumann *et al*., 2010; Tatout *et al*., 2014). AtSUN1 and AtSUN2 have the SUN domain at the C-terminus while AtSUN3, AtSUN4 and AtSUN5 have their SUN domain in a central position, (Graumann, 2014). Thus, there are Cter-SUN proteins (AtSUN1 and AtSUN2) and Mid-SUN proteins (AtSUN3, AtSUN4 and AtSUN5), which have been demonstrated to be conserved from yeast to plant as a monophyletic group (Graumann, 2014; Murphy *et al*., 2010; Poulet *et al*., 2017). In addition to a different SUN domain position, these two subfamilies have a different number of transmembrane domains (TMD) and coiled-coil domains (CCD). CCDs are necessary to form oligomers, as previously shown with AtSUN1 and AtSUN2, which form homo- and heterodimers, (Graumann *et al*., 2010). Also, the two Arabidopsis subfamilies localise at the NE and at the ER but the Cter-SUN proteins are enriched at the nuclear envelope compared to the ER (Graumann, 2014). Moreover, both subfamilies interact with the same KASH domain

proteins, (introduced in detail in ***II.2-b*** below), such as AtWIP1, and AtTIK, (Graumann, 2014; Zhou *et al*., 2012, 2015), but only Cter-SUN proteins are currently known to interact with SINE proteins, (Evans *et al*., 2014; Zhou *et al*., 2014).

II.2-b – KASH domain protein family

Though no sequence homologues for opisthokont KASH domain proteins exist in plants, different proteins have been identified, containing KASH domains and also having conserved binding properties for plant SUNs (Meier, 2016; Zhou and Meier, 2013, 2014; Zhou *et al*., 2015). In Arabidopsis, so far, four protein families have been evidenced, the WPP (tryptophan-proline-proline)- DOMAIN-INTERACTING PROTEINS (AtWIP1-3) and their binding partners the WPP DOMAIN-INTERACTING TAIL-ANCHORED PROTEINS (AtWITs), (Zhou and Meier, 2014; Zhou *et al*., 2012), the SUN-INTERACTING NUCLEAR ENVELOPE (AtSINE1-5, (Zhou *et al*., 2014)) and the TOLL-INTERLEUKIN-RESISTANCE KASH (AtTIK) domains proteins, (Graumann, 2014; Meier, 2016). The KASH domain is composed of a TM domain followed by a XXPT motif in C-ter position important for interacting with SUN domain proteins. On the cytoplasmic side, through AtWIT interaction, AtWIP proteins interact with RanGAP, a factor implicated in nucleo-cytoplasmic transport through NPCs, (see ***II.1***) and AtSINE proteins interact with F-actin fibers, (Zhou *et al*., 2014). AtTIK protein is the least described at the moment.

Thus, as in metazoan, (Rothballer and Kutay, 2013), in plants different types of LINC complexes are anchored at the NE combining different KASH domain proteins, AtWIP, AtSINE and AtTIK and the two-subfamilies of SUN domain proteins, AtSUN1-2 and AtSUN3-5. Some types of LINC complexes could be specific to particular cell types or certain development stages, as suggested by expression profiles like AtSUN5 that is specifically expressed in pollen, (Meier, 2016), ***Figure 1-4***.

III – Nuclear periphery components

At the nuclear periphery in metazoans, a meshwork of proteins is present on the nucleoplasmic side of the INM. This meshwork is named the lamina and is composed of type V intermediate filament (IF) proteins, lamins, and also lamin binding proteins, (Burke and Stewart, 2013).

III.1 – Lamins and their associated proteins and domains

Metazoan lamins belong to the IF family and comprise the fifth type of IFs. These proteins are separated into two classes: A-type and B-type lamins. A-type lamins are mainly composed of lamin A and lamin C, encoded by a single gene, called *LMNA* in human. B-type lamins are mainly composed of lamin B1 and lamin B2, encoded by *LMNB1* and *LMNB2* respectively (Burke and Stewart, 2013). As an IF protein, the lamin monomer is formed by a central CCD also known as an α-helical rod domain (~50nm and 350aa), a short head domain of ~30aa at N-ter and a long C-ter tail domain of ~200aa containing an Ig-fold domain of ~3.5nm in diameter (Turgay *et al*., 2017). Lamin A and lamins B1 and B2 display also a C-ter CaaX motif where “C” is cysteine, “aa” aliphatic residues and “X” any amino-acid, usually a methionine. This CaaX motif is important for PTM processes like farnesylation to target newly synthesized lamins to the NE and carboxy-methylation to realize the correct cleavage of the protein, (Burke and Stewart, 2013). A lamin dimer is formed by two parallel monomers with both Ig-fold domains on the same side and then dimers associate head-to-tail to form polymers leading to a rod-shape fibre, 3.5nm thick (Turgay *et al*., 2017). This fibre displays typical paired globular domains, distant of approximately 20nm from each other on the rod, relative to tetrameric and hexameric regions containing lamin Ig-fold domains. Globular domains can also be associated with lamin binding partners (Turgay *et al*., 2017). A- and B-

type lamins are major components of lamin filaments and form a dense meshwork underneath the INM, representing 12.5% of the ~14nm-thick lamina (Turgay *et al*., 2017).

Lamins interact with the LINC complex by binding SUN proteins (Mattout *et al*., 2006), and, with other components of the lamina, anchor chromatin domains and regulatory molecules such as transcription factors, (Burke and Stewart, 2013). Thus, this particular structure forms a scaffold at the nuclear periphery, tethers peripheral actors to the lamina and plays a crucial role for cell survival by regulating gene expression, chromatin organization, or even nuclear shape and movements, (Mattout *et al*., 2006). Mutations in lamins and also in the other components of the lamina are deleterious for the organism and generate laminopathies, such as the Hutchinson-Gilford Progeria Syndrome (HGPS) in human, (Méjat and Misteli, 2010), or Emery-Dreifuss Muscular Dystrophy (EDMD), (Ho and Hegele, 2018; Mattout *et al*., 2006; Mounkes and Stewart, 2004; Samson *et al*., 2018).

Indeed, lamins, even if major components of the lamina, are not lone actors. Some lamin-binding partners are anchored into the INM with one, two or more TM domains, such as the lamin B receptor (LBR) and the LEM domain (LAP2, Emerin, MAN1 domain) protein family, tethering lamina to the NE (Dilsaver *et al*., 2018); some others mediate interactions with chromatin, such as Barrier to Autointegration Factor (BAF) linking chromatin to the lamina, (Wilson and Foisner, 2010), ***Figure 1-5A***. LEM domain proteins play many roles in cell signalling, (Huber *et al*., 2009), and through interactions with A- and B-type lamins, transmit signals to gene expression networks and also to translational machinery in the cytoplasm, (Ma and Blenis, 2009). LEM domain proteins are also able to bind DNA directly due to specific domains or chromatin proteins, (Wilson and Foisner, 2010).

Particular chromatin domains named Lamin Associated Domains (LADs) have been evidenced using a technique termed DamID. This technique reveals by adenine methylation, the DNA regions that enter into contact with a bait lamina protein (usually Lamin B1) fused

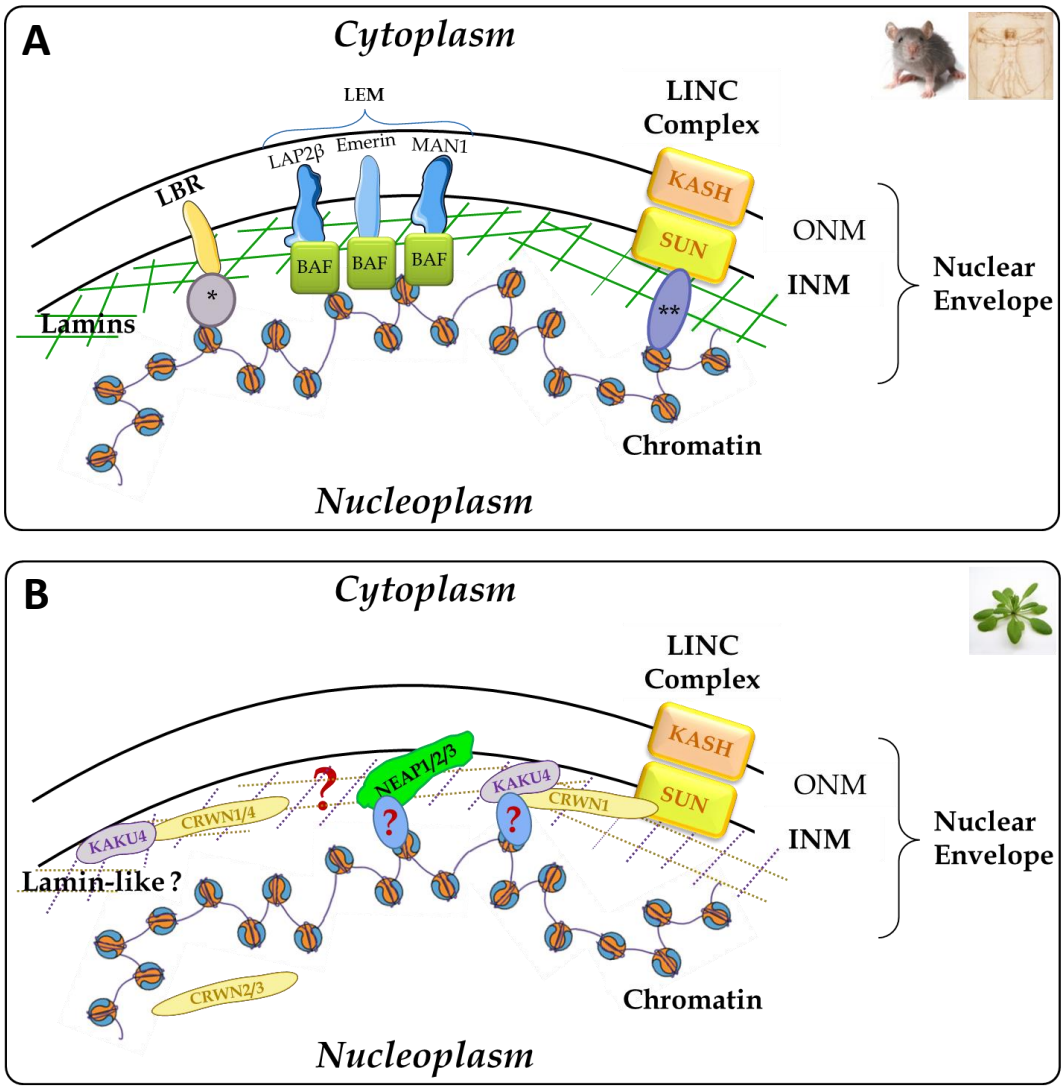


Figure 1-5: Comparison of the nuclear periphery structure and organisation in metazoan versus plants. **A.** In metazoa, the lamina is mainly composed of lamin filaments (green), which interact with proteins anchored at the INM, such as Lamin B Receptor (LBR) and LAP2β-Emerin-Man1 (LEM) proteins, and also with the LINC complex via SUN domain proteins. Thanks to intermediate proteins like Barrier of Autointegration Factor (BAF), chromatin domains are tethered at the lamina and thus at the nuclear periphery. Different proteins act as intermediate proteins and link chromatin to the lamina such as *Heterochromatic Protein 1 (HP1), **Bouquet1-2 (Bqt1-2), Silent Information Regulator 4, (Sir4) or Non-Disjunction 1 (Ndj1).
B. In plants, there is no homologue of lamins but KAKU4 and CRoWded Nuclei (CRWN) proteins are suggested as having the same role, (see *Chp1-III.2)*. Nuclear Associated Proteins (NEAP) are the most recent identified proteins and their function is investigated in this study.

to the bacterial *Dam* methylase, (van Steensel and Belmont, 2017; van Steensel and Henikoff, 2000). In differentiated human cells, LADs represent up to 30-35% of the human genome, usually harbouring repressive chromatin features, (Guelen *et al*., 2008). Two classes of LADs can be found: constitutive LADs (cLADs), that are present in all cell types, and facultative LADs (fLADs) that are LADs in a cell-type specific manner, (van Steensel and Belmont, 2017). Several studies suggest that this organization directly participates in silencing maintenance of particular chromatin regions. Indeed, localisation shift of those regions to the nuclear interior is usually related to a transcriptional gene reactivation of these regions, (Kind *et al*., 2013; Reddy *et al*., 2008). In 2010, two studies revealed the existence of other genomic regions with repressive chromatin features associated to the nucleolus, (van Koningsbruggen *et al*., 2010; Németh *et al*., 2010). These genomic regions were named NADs for Nucleolus Associated chromatin Domains. Interestingly, some NADs were already described as LADs, indicating a potential redundancy of these regions. Further studies revealed that after mitosis, some LADs are found in close association with nucleoli, demonstrating the existence of a stochastic reshuffling of a portion of LADs (possibly fLADs), (Kind *et al*., 2013).

In *A. thaliana*, NADs were identified by (Pontvianne *et al*., 2016) and genomic regions associated with the nuclear periphery were described by (Bi *et al*., 2017) using a derived Chip-seq approach (Re-ChIP-seq) with the nuclear pore protein NUP1 as bait. As in human cells, plant NADs and LADs-like are composed of genomic regions displaying heterochromatic features as the repressive histone modifications H3K9me2 and H3K9me3 (Bi *et al*., 2017; Pontvianne *et al*., 2016; van Steensel and Belmont, 2017). Importantly, genes present in these regions tend to be low expressed-genes, indicating a potential link between gene nuclear positioning and gene expression. However, unlike in human cells, plant LADs-like and NADs correspond to two mutually exclusive regions in *A. thaliana*, (Picart-Picolo *et al*., 2019).

Much less is known about the plant equivalents of the nucleoskeleton organisation and of proteins with such functions described in above paragraphs. Thus, this project aims to characterize new components of the NE and how the protein network at the nuclear periphery can possibly interact with chromatin.

In yeast, there are no lamin sequence homologues currently known, and the plant lamina or « plamina », (Fiserova *et al*., 2009), contains no lamins, LEM domain proteins or LBR homologues, (Rose *et al*., 2004). Neverthless, a meshwork of proteins is visible (Fiserova and Goldberg, 2010; Fiserova *et al*., 2009). A few plant proteins are suggested to have the same functions as mammalian lamins in controlling nuclear shape and chromatin organization, like CROWDED NUCLEI (CRWN) and KAKU4 proteins, detailed below (***III.2***), and more recently, AtNEAPs, which have the same characteristics as LEM proteins, suggesting similar functions. In addition, like LEM domain proteins, AtNEAPs may interact with chromatin. As this is the subject of the thesis, it will be described in detail in subsequent chapters, ***Figure 1-5B***.

III.2 – The plant lamina-like structure

Thus, no sequence homologues in plants have been found for metazoan lamins or lamin associated proteins. It seems likely that plants have developed their own proteins for supporting the NE and some recently identified proteins appear to have similar functions as lamins and are believed to be part of the plant nucleoskeleton.

III.2-a – Lamin-like proteins: CRWN family

(Moreno Díaz de la Espina *et al*., 1991) were the first to describe a fibrillar nuclear matrix underneath the INM in onion root cells similar to the one observed in metazoan. This matrix was isolated using specific extraction protocols from Laemmli and colleagues, (Laemmli, 1970; Lewis *et al*., 1984). Also, Matrix Attachment Regions (MARs) or Scaffold

Attachment Regions were described as DNA-nuclear matrix contact regions, (Breyne *et al*., 1994) and could be related to the current LADs and NADs, see ***III.1*** above.

Later, the first protein family to be identified was the NUCLEAR MATRIX CONSTITUENT PROTEIN (NMCP) family, which is highly conserved in plants but absent from metazoans and fungi, (Ciska *et al*., 2013). NMCP1 was the first protein of the family, identified using an antibody raised against the nuclear matrix purified from *Daucus carota* (Masuda *et al*., 1997). Although larger with a molecular mass of 134kDa, DcNMCP1 has a similar structure to metazoan lamins with a large central CCD and a putative NLS in its tail domain (Masuda *et al*., 1997). In Arabidopsis, four homologues have been identified in a genome-wide search for CCD proteins, (Rose *et al*., 2004). According to their mutant phenotypes that show "LITTLE NUCLEI", they were firstly named, AtLINC1-4 (Dittmer *et al*., 2007), but as it was confusing with the use of LINC complex for Linker of Nucleoskeleton and Cytoskeleton, they have been finally named AtCRWN1-4, (Wang *et al*., 2013), for "CROWDED NUCLEI". In the literature, all these names can be found, CRWN, LINC and NMCP, designing the same protein family in different plant species.

A mass spectrometry analysis of the Arabidopsis matrix at the nuclear periphery, (Sakamoto and Takagi, 2013) isolated about 1,600 proteins including AtCRWN1 and AtCRWN4 but not AtCRWN2 and AtCRWN3. AtCRWN1, AtCRWN2 and AtCRWN3 evolved at the same time and belong to the same monophyletic group as type 1 NMCP, and AtCRWN4 belongs to type 2 NMCP, (Wang *et al*., 2013). CRWN1 and CRWN4 are localized at the nuclear periphery, while CRWN2 and CRWN3 are preferentially localised at the nucleoplasm, (Dittmer *et al*., 2007; Sakamoto and Takagi, 2013).

It has also been shown that, like metazoan lamins, CRWN1 and CRWN4 are involved in maintenance of nuclear morphology (Dittmer *et al*., 2007; Sakamoto and Takagi, 2013; Wang *et al*., 2013). They are also involved in chromosome pairing, especially at

pericentromeric regions, evidenced by Hi-C in *crwn4* mutant plants, (Grob *et al*., 2014), and a chromocentre fusion has been observed in *crwn1crwn2* mutants that can be explained by the increased pairing at heterochromatic regions (Poulet *et al*., 2017; Wang *et al*., 2013).

The structure of the CRWN protein family, with a central CCD, similar to, but larger than the metazoan lamin CCD, is predicted to confer the ability to form filaments like lamins, (Ciska *et al*., 2013; Dittmer *et al*., 2007; Sakamoto and Takagi, 2013). Although no evidence were yet established concerning a real CRWN polymerisation, due to their similarities with metazoan lamins, CRWN proteins are not only suspected to be part of the plant lamina but also to be functional homologues of lamins, (Ciska and Moreno Díaz de la Espina, 2014; Ciska *et al*., 2013; Sakamoto and Takagi, 2013). Also, in mammals, SUN domain proteins interact with lamins; in plants, AtSUN1 and AtSUN2 are able to interact with AtCRWN1, (Graumann, 2014).

III.2-b – A CRWN-binding protein: KAKU4

AtCRWN1 and AtCRWN4 are able to interact with another nuclear periphery protein, named KAKU4, (Goto *et al*., 2014). Indeed, in a screen searching for mutants with an altered nuclear morphology, (Tamura *et al*., 2013) identified three proteins named KAKU for "nucleus" in Japanese. AtKAKU1 (see II-2-b) is a myosin XI-I known to interact with KASH domain proteins AtWIT1 and AtWIT2; AtKAKU2 is allelic to AtCRWN1, confirming the role of AtCRWN1 in nuclear morphology, (Dittmer *et al*., 2007); and AtKAKU4, a protein containing a NLS and a GAR domain with repeated glycine-arginine in C-ter position, and implicated in maintaining nuclear shape, (Goto *et al*., 2014).

KAKU4 is a CRWN-binding protein but neither CRWN nor KAKU4 seem to be responsible for the localisation at the nuclear periphery of each other, (Goto *et al*., 2014). Indeed AtKAKU4 has no paralogue and in an *Atkaku4* mutant AtCRWNs are well localised. On the contrary, the possible dependency on AtCRWN proteins for AtKAKU4 localisation

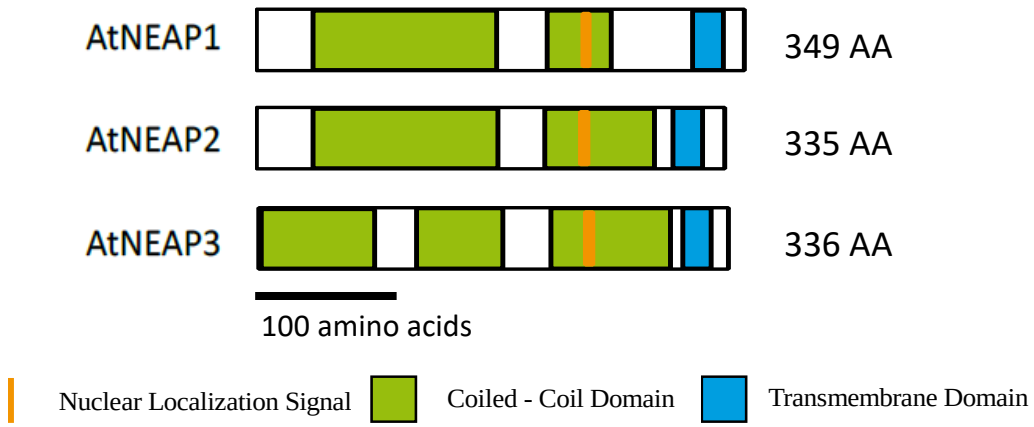


Figure 1-6: *Arabidopsis thaliana* Nuclear Envelope-Associated Protein (NEAP) family is composed of three proteins of about 350 amino acids (AA) containing coiled-coil (CC) domains, nuclear localisation signal (NLS) and a transmembrane (TM) domain at the C-terminus.

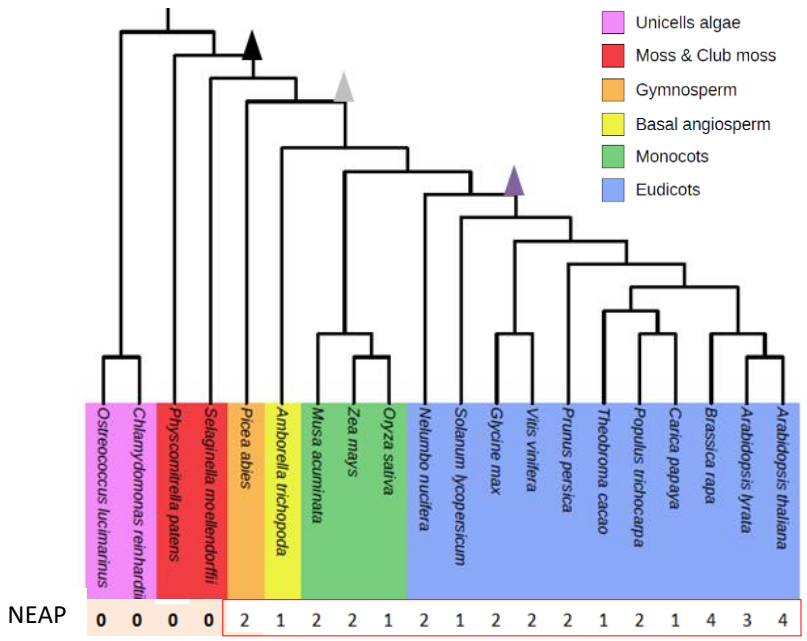


Figure 1-7: NEAP proteins during Plant Kingdom evolution. NEAP proteins are present in gymnosperms, basal angiosperms, monocots and eudicots but are absent from the most primitive species like moss and unicellular algae. Number of NEAP proteins in each studied species is indicated in the line above the phylogenic tree. Triangles indicate whole genome duplication events; adapted from Poulet et al., 2016.

has to be confirmed as the experiment was carried out only with an *Atcrwn1* single mutant and AtCRWN2, AtCRWN3 or AtCRWN4 could complement *Atcrwn1* loss-of-function. Also, AtKAKU4 could be localized at the INM even if no TM domain is present on the protein, (Goto *et al*., 2014). Interestingly, when AtKAKU4 as well as AtCRWN1 were over-expressed in plants, growth and deformation of the NE as well as intra-nuclear vesicle like NE invaginations have been observed in a dose-dependent manner, (Goto *et al*., 2014). This may indicate a function related to nuclear envelope assembly.

Thus, KAKU4, with CRWN1 and CRWN4, and to a lesser extent with CRWN2 and CRWN3 would form a protein network "lamina-like" at the nuclear periphery in plants. Their interaction with chromatin and the existence of LADs in plants is still a matter of intense investigations.

IV – The NUCLEAR ENVELOPE-ASSOCIATED PROTEINS

IV.1 – A new family of Inner nuclear membrane associated proteins

Based on a bioinformatic screen for new nuclear membrane proteins with KASH-like characteristics, and containing CCD and NLS, AtNEAP proteins were found and shown to possess a TM domain, (Pawar *et al*., 2016), ***Figure 1-6***. Phylogenetic analysis and sequence alignment tools (Poulet *et al*., 2016) revealed that the *NEAP* gene family first appeared in gymnosperms and are absent from archaic species as unicellular algae, ***Figure 1-7***. The monocots and the eudicots form monophyletic groups with specific-to-species gene duplication. Thus, during speciation of the *Brassicaceae* of which *A. thaliana* is a member, a duplication event resulted in three genes, *NEAP1*, *NEAP2*, *NEAP3*, (Poulet *et al*., 2016). As *AtNEAP4* is specific to Arabidopsis, truncated and transcribed at very low level, it has been considered as a pseudogene, (Poulet *et al*., 2016).

The first investigation to characterize AtNEAPs explored their localization. For that purpose, transient infiltration in *Nicotiana benthamiana* plants was used. These studies used constructs of transiently over-expressed AtNEAP proteins fused to a fluorescent label. Localization was assessed by confocal microscopy. All three proteins localized to the nuclear periphery and fluorescence recovery after photobleaching (FRAP) experiments showed that AtNEAP1 and AtNEAP2 are more tightly bound to the nuclear periphery than AtNEAP3. Moreover, mobile fractions of AtNEAP1 and AtNEAP2 are comparable to other NE or NE-associated proteins like AtSUN proteins. Studies of the different protein domains of AtNEAP3 indicated that the first CCD and TM domain are required for localization at the nuclear periphery rather than in the nucleoplasm. The NLS was shown to be required to target the protein to the nucleus instead of cytoplasm, (Pawar *et al*., 2016).

A second investigation explored whether AtNEAPs were able to form homo- or heterodimers by carrying out an acceptor photobleaching Fluorescence Resonance Energy Transfer (apFRET) experiment after co-infiltrating two AtNEAP proteins fused either with yellow or cyan fluorescent protein (YFP/CFP) in *Nicotiana benthamiana* leaves (see Methods section III-E). Results obtained indicated that AtNEAP1, AtNEAP2 and AtNEAP3 can interact with each other and with themselves. As these studies were performed with transiently over-expressed proteins, complementary studies were also carried out using the Membrane Yeast Two Hybrid (MYTH) system. These confirmed AtNEAP1-AtNEAP1, AtNEAP1-AtNEAP2 and AtNEAP1-AtNEAP3 interactions although these interactions were weak in MYTH, (Pawar *et al*., 2016). Other protein partners for AtNEAPs have been identified by Pawar *et al*, 2016. Indeed, apFRET and MYTH experiments showed that all three AtNEAPs interact with LINC complex components AtSUN1 and AtSUN2, (Pawar *et al*., 2016).

IV.2 – AtNEAP-interacting partner: AtbZIP18, a link with chromatin?

A MYTH screening using AtNEAP1 as bait, also revealed among others a basic-LEUCINE ZIPPER (AtbZIP18) protein as an interacting partner, which is a transcription factor (TF), (Gibalová *et al*., 2017; Pawar *et al*., 2016). Localization and co-localization studies show AtbZIP18 localized to the nucleus and the cytoplasm. When AtNEAP1 is co-expressed with AtbZIP18, AtNEAP1 fails to localize at the nuclear periphery and co-localizes with AtbZIP18 in the nucleoplasm (Pawar *et al*., 2016), indicating a potential *in-vivo* interaction for AtbZIP18 and AtNEAP1.

Thus, the transcription factor AtbZIP18 (Pawar *et al*., 2016) could be an interactor of AtNEAP1. AtbZIP18 belongs to a large family of transcription factors, named bZIP, implicated in a broad range of mechanisms (Dröge-Laser *et al*., 2018). In *A. thaliana* the protein family is composed of 78 members divided in 13 groups (A-M), (Dröge-Laser *et al*., 2018). The main characteristic is the presence of a BRLZ domain for a basic DNA-binding region followed by a leucine zipper allowing bZIP dimerization. AtbZIP18 belongs to group I, which is related to stress response, cell cycle regulation and various developmental aspects, (Dröge-Laser *et al*., 2018). It is implicated in pollen development and has been further characterized by (Gibalová *et al*., 2017). AtbZIP18 localizes to the ER and also in the nucleus, but is excluded from the nucleolus, (Gibalová *et al*., 2017). It has a rather ubiquitous expression pattern with higher levels of expression in mature pollen grains, embryo nuclei and roots. AtbZIP18 is thought to be redundant with AtbZIP34, one of its binding partners, and they both have a role in the male gametophyte. AtbZIP18 and AtbZIP34 could be repressors as in each single mutant there are more upregulated genes than downregulated. Moreover, AtbZIP18 has an Ethylene-responsive element binding factor-associated Amphiphilic Repression (EAR) motif, which is implicated in transcriptional inhibition through chromatin modification, (Gibalová *et al*., 2017; Kagale and Rozwadowski, 2010). Indeed, the EAR motif

is a common active transcriptional repression motif recruiting co-repressors such as AtSIN3, AtSAP18 (SIN3 ASSOCIATED POLYPEPTIDE P18) or TOPLESS (TPL), which interact with AtHDA19 to proceed to histone deacetylation and thus gene repression, (Kagale and Rozwadowski, 2011).

Through yeast two-hybrid (Y2H) experiments it has been shown that AtbZIP18 can also interact with AtbZIP61 and AtbZIP52, which also possesses an EAR motif; increasing the number of possible heterodimers, (Gibalová *et al*., 2017) and thereby different targeted genes.

Thus, it is of particular interest to investigate the interaction of AtbZIP18 with AtNEAP protein family in order to establish a link between chromatin and the nuclear periphery. Also, as AtbZIP18 would be a negative regulator of gene expression, this could help to better picture why and how heterochromatin is recruited at the nuclear periphery.

IV.3 – Aims of the research project

To be able to fully characterize AtNEAP proteins, reverse genetics is needed and different single mutants for all three genes have been selected and then crossed to obtain double and triple mutants, (Pawar-Menon, PhD thesis, 2015). The first set of mutants available included a single *Atneap1* KO, a single *Atneap3* KO, a double *Atneap1Atneap3* KO (Pawar *et al*., 2016) and a single *Atneap2* "leaky" probable knock-down (KD), see ***Results Chapter 3-I***.

The aim of the work presented in this thesis was to characterise the role, function and interactions of the AtNEAP protein family, building on previous work (Pawar-Menon, PhD thesis, 2015; Pawar *et al*., 2016), which indicated a location at the nuclear periphery and suggested a role in the interaction of chromatin with the nuclear envelope and nucleoskeleton through the LINC complex.

During this PhD, attempts were made to generate a triple *Atneap* mutant and this was achieved using CRISPR/Cas9 technology to generate a new single *Atneap2* KO mutant, which was crossed with the double *Atneap1Atneap3* KO mutant already available. Finally, the triple KO *Atneap1Atneap2Atneap3* and all the combinations of single and double mutants from this crossing were identified and selected. Study and characterisation of the mutants is described in section ***Results Chapter 3***.

New interaction partners of AtNEAP were explored using classical Y2H and confirmation of location and interaction with AtbZIP18 using high-resolution confocal microscopy (apFRET) was obtained. In addition, localisation *in-vivo* in *A. thaliana* was confirmed using a transient expression method called the FAST technique. Finally a number of tools were developed for future work, including antibodies specific to AtNEAPs, a complementation vector for mutant lines and protein extraction protocols for immunoprecipitation.

The results of the work will be discussed in the context of current knowledge of plant nuclear structure and constituents of the NE, nucleoskeleton and chromatin. Finally, possible future work will be discussed in order to more fully determine the role of this protein family in the nuclear periphery protein network; and its implication in the regulation of gene expression through interaction with the transcription factor AtbZIP18.

CHAPTER 2

MATERIALS and METHODS

Table 2.1: Amino-acid concentrations for Yeast media

Amino Acid	Abbreviation	Concentration (mg/ml)
Methionine	MET	20
Lysine	LYS	30
Uracile	URA	20
Histidine	HIS	20
Adenine	*ADE	20
Leucine	LEU	60
Tryptophane	TRP	20

*Add 4ml NaOH 5M to dissolve

MATERIALS and METHODS

I – Yeast

I.1 – Yeast strains

In this study, two *Saccharomyces cerevisiae* strains were used; **Y187**, *MATα, ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4Δ, gal80Δ, met–, URA3::GAL1$_{UAS}$-GAL1$_{TATA}$-LacZ, MEL1* for transformations with prey plasmids and **AH109**, *MATa, ura3-52, his3-200, trp1-901, leu2-3, 112, gal4Δ, gal80Δ, GAL2-ADE2, LYS2::GAL1-HIS3, URA3::MEL1-LacZ* for transformation with bait plasmids.

I.2 – Yeast growth and media

Yeasts were grown at 30°C for at least 3 days in different media, either liquid or solidified by adding agar 2%.

In order to start a fresh culture, a rich medium, Yeast extract Peptone Dextrose/D-glucose (YPD, 1% yeast extract, 2% peptone, 2% glucose/dextrose) was used.

In order to select yeasts transformants, a synthetically defined (SD, 2% glucose, 6.9g/l Yeast Nitrogen Base (YNB) medium) was used. SD medium was a minimal medium supplemented with essential amino acids (aa), some of which were omitted for selection of some plasmids and/or activated reporters. See ***Table 2.1***.

I.3 – Yeast transformation

Cells of either Y187 or AH109 strains grown on YPD were collected with a pipette tip and diluted in 1mL of 1X TE/LiAc and centrifuged at 13000g for 1min. This washing step was repeated with 1mL of 1X TE/LiAc and the pellet was resuspended in 500µL of 1X

TE/LiAc. For one transformation, 40µL of yeast was needed. DNA carrier was denatured at 95°C for 5min and put on ice.

Transformation was carried out with 300µL TE/LiAc/PEG 40%, 5µL DNA carrier, 200-500ng of plasmid (see ***IV.8***) and 40µL yeast prepared as above. The mix was incubated at 30°C for 30min, then 12µL DMSO was added and tubes were incubated at 42°C in a water bath for 15min in order to apply a heat shock. Finally, cells were plated on SD medium depleted for the selective amino-acids and incubated at 30°C for 3 to 5 days.

For screening a high number of proteins, cells of Y187 strain were transformed with a cDNA library of prey plasmids (Clontech, "Mate & Plate™ Library - Universal Arabidopsis (Normalized)") made from mRNA isolated from 11 Arabidopsis tissues.

I.4 – Yeast-Two-Hybrid (Y2H) screening

SD supplemented with all aa required but depleted in leucine and tryptophan (SD –Leu–Trp) was used to select diploid yeast containing bait and prey plasmids encoding tryptophan and leucine biosynthesis genes, respectively, that are otherwise absent from the cell. Thus, SD –Leu is used to select yeast containing prey plasmids and SD –Trp to select yeast containing bait plasmids.

When interaction occurs between bait and prey, the Gal4 transcription factor is reformed and activates the responsive *HIS3* and *ADE2* genes. Thus, diploid yeast containing bait and prey plasmids with interacting bait and prey proteins is able to grow on a SD medium depleted in leucine, tryptophan, histidine and adenine (SD/–Ade/–His/–Leu/–Trp).

Table 2.2: Antibiotic concentration and organism used for bacterial and plant selection

Antibiotic	Concentration (µg/ml)	Organism used on
Ampicillin	50	*E. coli / A. tumefaciens*
Gentamicin	25	*A. tumefaciens*
Kanamycin	50	*E. coli / A. tumefaciens*
Rifampicin	50	*A. tumefaciens*
Spectinomycin	50	*E. coli / A. tumefaciens*
Zeomycin	25	*E. coli*
Basta	10	*Arabidopsis thaliana*

II – Bacteria

II.1 – Bacterial strains

All cloning was performed in *Escherichia coli* (*E. coli*) using a chemically competent high efficiency DH5α strain. The DB3.1 strain was used to amplify empty gateway vectors containing the lethal *ccdB* gene (Bernard and Couturier, 1992).

For sub-cloning of binary vectors, the chemically competent *Agrobacterium tumefaciens* (*A. tumefasciens*) GV3101 strain was used, followed by plant transformation. GV3101 strain (C58 background) contains a rifampicin resistant gene (RIF) for selection during transformation, a nopaline type Ti plasmid pMP90 without its transport function (pTiC58DT-DNA) and a plasmid containing *vir* gene. The VIR T-DNA gene was inserted into the plant genome with essential elements (pTiC58DT-DNA, pMP90). Its T-DNA transfer function is damaged but can be transferred to the binary vector T-DNA to help smooth transfer. PMP90 (pTiC58DT-DNA) Ti plasmid contains screening tags: streptomycin and gentamycin for selection upon plant transformation such as Arabidopsis, tobacco, maize and potatoes.

II.2 – Bacterial growth and media

E. coli were grown O/N at 37°C in Luria Broth (LB) medium, either solidified by adding 1% agar or liquid, containing specific antibiotics depending on the vector transformed in the bacteria, see ***Table 2.2***.

A. tumefaciens were grown at 28°C in LB medium or Yeast Extract Broth (YEB), either solidified by adding 1% agar or liquid. Overnight shaking cultures (150rpm) were used for infiltration; otherwise cells were cultivated on plates for 48 hours, always at 28°C.

Rifampicin and gentamicin for helper plasmid T1 were required to select the GV3101 strain, plus additional antibiotic for selection of the plasmid of interest, see ***Table 2.2***.

II.3 – Transformation of E. coli

For electroporation (EP), the chemically competent *E. coli* DH5α strain was used. Cuvettes were cooled on ice and 50µL aliquots of cells were then thawed also on ice. Then 50-200ng of plasmid (see *IV.8*) were added to the cells and transferred to the cooled cuvettes. Samples were incubated for 30sec before electroporation at 1.8kV. Immediately, 450µL of Super Optimal broth with Catabolite repression (SOC, LB medium supplemented with 20mM glucose) was added and transferred to 2mL tubes for incubation at 37°C with shaking for 1h. Finally, 50µL were plated on LB Agar with specific antibiotics (see ***II.2***) and incubated at 37°C overnight.

For heat shock (HS) transformation, the following kits and cells were used; Agilent, "Ultracompetent cells XL-10" and NEB, "HiFi competent cells" and "5-alpha Competent *E. coli*". The respective reagents and protocols were used as indicated by the manufacturers. Refer to *IV.1* – List of primers for plasmid isolation procedure.

II.4 – Transformation of A. tumefaciens

Chemically competent *A. tumefaciens* GV3101 strain was used for transformations. Two different but similar techniques were used depending on the laboratory.

For EP, cuvettes were cooled on ice and 40µL aliquots of cells were thawed on ice. Then 50-200ng of plasmid were added to the cells and transferred to cooled tubes. Samples were incubated for 30sec before applying an electroporation of 1.8kV. Immediately, 1mL SOC was added and transferred to 2mL tubes for incubation at 28°C for 3h. Finally, 5µL were plated on LB Agar with specific antibiotics (see ***II.2*** and ***Table 2.2***) and incubated at 28°C for 48h.

For HS, 100µL of cells were thawed on ice and 0.5-1µg of plasmid was added to the cells and incubated on ice for 5min, then 5min in liquid nitrogen and 5min at 37°C. After heat

Table 2.3: Details of Arabidopsis lines used in this study

Arabidopsis lines used in this study							
Usual name		***First neap2***	***neap1***	***neap3***	***neap2 CRISPR***		
Official name		*Atneap2*	*Atneap1.1*	*Atneap3.1*	*Atneap2.1*		
Mutations	**Type**	T-DNA	T-DNA	T-DNA	CRISPR		
	Reference	GABI_589B02	Sail_846B07	WiscDsLoxHs086_02C	N/A		
Gene accession number		At5g26770	At3g05830	At1g09470	At5g26770		
Antibiotic resistance		Sulphadiazine	Basta	Kanamycine	N/A		
		X	X	X		***First triple neap***	**Combinations of mutants**
			X	X		***neap1neap3***	
			X		X	***neap1neap2***	
				X	X	***neap2neap3***	
			X	X	X	***triple neap***	

shock, 1mL YEB or LB was added and cells shaken at 28°C for 3h. Finally, 200µL were plated on YEB Agar with specific antibiotics (see ***II.2*** and ***Table 2.2***) and incubated at 28°C for 48h.

III – Plant

III.1 – Seed Stock

The wild type *Arabidopsis thaliana* (*A. thaliana*) *Columbia* (Col-0) and different mutant lines used in this study are listed in ***Tables 2.3 and 2.4***.

Wild type *Nicotiana benthamiana* (*N. benthamiana*) plants were used for *A. tumefaciens* infiltrations and transient protein expression (see ***III.5***).

III.2 – Seed germination and plant growth

For all plants, temperature and light conditions were 23°C and 8h dark/16h light. They were grown either in sterile soil or *in vitro* on plates with MS medium (Murashing and Skoog (Caisson Laboratories), 0.8% agar, 1% sucrose, pH5.7). In order to synchronise seed germination, all the seeds sown were left at 4°C for 48h before being transferred into growth chambers (Aralab).

For *in vitro A. thaliana* cultures, seeds were previously sterilized with 70% ethanol and 0.05% SDS and then rinsed with 95% ethanol and left to dry before sowing.

III.3 – Crossing lines

In order to obtain multiple T-DNA insertion mutants, homozygous single or double T-DNA insertion mutants were crossed by directed pollination. *A. thaliana* plants to be crossed were grown as detailed above in ***III.2***. The mother plant was grown to a stage where it had a single floral stalk and few young flower buds. All open flowers, buds with white tips, immature budding meristems and mature siliques were removed with a pair of forceps so that

Table 2.4: List of constructs stably transformed into Arabidopsis lines

CDS transformed	Plant Background	Antibiotic resistance
YFP_NEAP1	Wild type Col-0	Basta
YFP_NEAP2		Basta
YFP_NEAP3		Basta
CFP_NEAP1		Basta
CFP_NEAP2		Basta
CFP_NEAP3		Basta
CRISPR_E1N2		Basta
CRISPR_E3N2		Basta

three to four buds remained for emasculation. The remaining flower buds were emasculated by first splitting the petals and sepals and then picking them out carefully along with all anthers. The mother plant was then pollinated using a mature flower from the father plant, by tapping the anthers on the style with the pollen visibly covering the stigma. This was repeated for all floral buds. Crossed buds were marked by tying coloured threads around them for identification. Once mature, a small paper bag was tied around individual siliques and they were allowed to dry inside the bag before collection. Hybrid seeds, expected to be heterozygous for both mutations from their homozygous parents were allowed to self-pollinate and their progeny was genotyped for identification of the homozygous double or triple mutant lines.

III.4 – Phenotype analysis

Root Assay was performed on wild type Col-0 and mutant lines sown on MS medium according to growth conditions, see ***III.2***. Square plates were used for this experiment, and kept vertical to let the primary root grow vertically at the medium surface. Seedlings were grown for 10 days after germination (dag) and primary root growth was assessed at different time points, 3, 5, 7 and 10 dag. Measures of root length were carried out with *ImageJ* software (see ***VIII.1***) on 45 independent seedlings per genotype and statistical analysis conducted as described in ***VII***.

Silique size was measured on 36 siliques at the same stage from 12 plants per genotype and statistical analysis carried out as described in ***VII***.

Seed number per half silique was scored from 36 siliques measured previously from 12 plants per genotype and statistical analysis carried out as described in ***VII***.

Staining of live cotyledon nuclei and fixation. Experiments were done in darkness. In this study, PicoGreen® solution (Molecular Probes) an ultra-sensitive fluorescent nucleic acid stain was used instead of DAPI as a DNA intercalating agent. From aliquots of Picogreen,

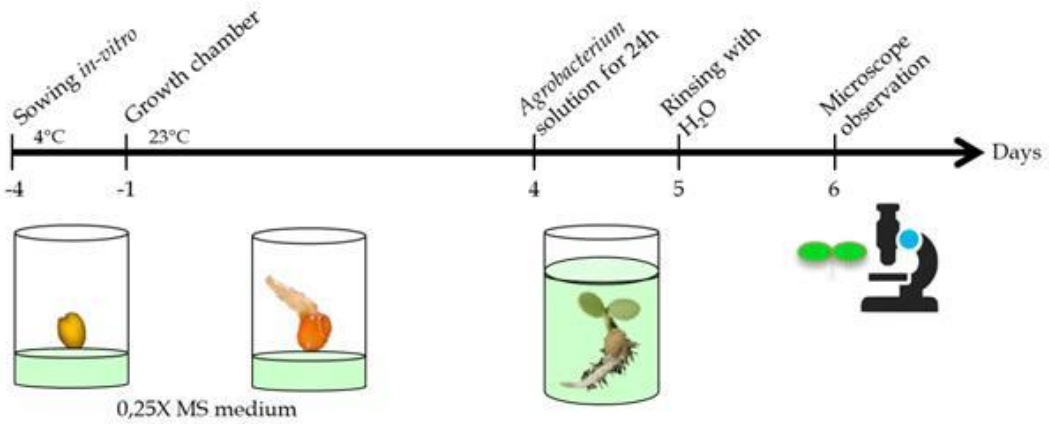


Figure 2-1: Main steps of adapted FAST protocol from Li and Nebenfuhr, 2010.

diluted in DMSO (1:10) and stored at -20°C, a solution of Picogreen diluted 1 in 200 with PBS 1X 0.01% Triton X-100 was prepared. Three to four 10 or 14 dag seedlings per line were taken and placed in water until all seedlings were harvested. Then, samples were placed in Picogreen solution and incubated for 15min at RT, 5min under vacuum and 30min more at RT. Next, samples were fixed with formaldehyde (formaldehyde 1%, PBS 1X, DMSO, H_2O) for 25min at RT and 5min under vacuum. Afterwards, fixative solution was replaced by methanol for a quick rinse followed by three washes of 5-10min with methanol and then for three washes of 5-10min with ethanol until the green colour disappeared. Samples were rehydrated with PBS 1X and then with 3 times 10min into PBS:glycerol 20:80 and left in this solution until slide mounting.

III.5 – Transient transformation

Transient transformation of A. thaliana was performed with an adapted protocol from (Li and Nebenführ, 2010) named Fast Agro-mediated Seedling Transformation (FAST) ***Figure 2-1***. In an ELISA plate (96 well), 110µL 0.25X MS medium (see ***III.2***) was poured into every well except those at the periphery, where 200µL H_2O was poured into the wells to keep the plate moisturized. Two seeds per well were sown for a total of 60 wells per plate and left for 48h at 4°C before placing plates in growth cabinets. On the day of seedling transformation (4dag), the transformed *A. tumefaciens* of interest (see ***II.4***) was resuspended from the plate on which it was growing to washing solution (10mM $MgCl_2$, 100µM acetosyringone) and OD measured. Seedlings were submitted to 100µL of co-cultivation medium (1.13g/l MS, 1% sucrose, 100µM acetosyringone, pH6.0 and just before use 0.005% (v/v; 50µl/l) Silwet L-77) in the presence of the transformed *A. tumefaciens* of interest at the OD required. Plates were wrapped in aluminium foil to keep them dark and then put back into the growth cabinet for 24h shaking at 75rpm. At 5dag, wells were washed with sterile H_2O,

the first time quickly and then three more times for 10min, keeping the last wash until microscope observation, with a maximum of 48h.

Transient transformation of N. benthamiana was performed using four to six week-old plants by infiltration with *A. tumefaciens*. Bacterial cultures (see ***II.4***) were removed from the incubator and 1mL from each sample was centrifuged at 8000rpm for 3min. The supernatant was removed and the pellets were gently resuspended in 1mL of Infiltration Buffer (IB, 0.5% D-Glucose, 0.5M MES, 0.02M $Na_3PO_4.12H_2O$, 0.0001% acetosyringone, H_2O) and centrifuged at 8000rpm for 3min. Supernatant was removed and 1mL IB was added to resuspend the pellet. Optical Density (OD) was measured by a Nanodrop spectrophotometer and was adjusted to 0.05 for p19 (a RNAi silencing suppressor) and 0.1 for other constructs. A 1mL syringe was used to push the resuspended bacteria into the leaf through a small hole pierced previously (Omarov *et al*., 2006). The plants were incubated at least for 2 days before checking protein expression by confocal imaging (see ***VI.3***) or harvesting plant tissue for protein extraction (see ***V.1***).

III.6 – Stable transformation – Deep floral transformation

A first overnight pre-culture was carried out with 200mL of the transformed *A. tumefaciens* cells in 15mL of LB containing the three antibiotics necessary for these cells. The day after, an overnight culture was prepared with 200mL of LB, 10mM $MgSO_4$ and with gentamycin and antibiotic corresponding to the vector of interest, but not rifampicin. The following morning, the OD_{600} was measured and the solution diluted until OD_{600} = 0.2. Incubation then was continued until it achieved an OD_{600} = 0.6-0.8. The solution was then centrifuged at 4000g for 10min, then the pellet was resuspended in 100mL of IB (5% sucrose, 0.1M $MgCl_2$, 0.02% Silwett). Then, all floral buds of a plant were immersed for 20s in this

solution. Afterwards, trays of plant pots were covered for 48h and finally transferred to the green house to be dried out.

IV – Nucleic Acids

IV.1 – List of primers

A list of primers is provided as supplemental tables at the end of this thesis (***Appendix I*** and ***Appendix II***).

IV.2 – Extraction of genomic DNA

Small pieces of young leaves were incubated in 8 tube PCR strips with DNA extraction buffer (0.2M Tris-HCl pH7.5, 0.25M NaCl, 0.025M EDTA pH8, 0.5% SDS, H_2O) for 10min at 99°C and spun down at 13500rpm for 8min. Then, supernatants were mixed with equal volumes of isopropanol, incubated for 15min at RT and spun down at 13500rpm for 15min. DNA pellets were washed with 70% ethanol, dried at 55°C and resuspended in H_2O.

IV.3 – RNA extraction

For RNA extraction, 8 to 12 seedlings of 14dag per sample were ground into powder with liquid nitrogen and were submitted to a number of different extraction protocols. Different kits used were “NucleoSpin RNA plant” from Macherey-Nagel and the “RNAeasy Plant Minikit” from Quiagen, both following the manufacturers’ protocol. A protocol using Trizol, Phenol, Chlororoform, and Isoamylic Alcohol was also used. This technique will be described in detail.

After a 5min-incubation at RT in 1mL Trizol (Tri-Reagent, Euromedex) and centrifugation at 4°C, 5min, 13200rpm, 200µl chloroform were added to the supernatants and samples incubated for 5min at RT and then centrifuged at 4°C, 15min, 13200rpm. The aqueous phase was taken and RNA precipitated with 600µL isopropanol. Samples were

incubated at least 30min at -20°C, then centrifuged at 4°C, 10min, 13200rpm and the final RNA pellet was washed with 1mL 70% ethanol. Pellets were resuspended in 50µl RNase-free H_2O.

Residual DNA present in the samples was removed by treatment with 4U DNaseI (1U/µL, RQ1 DNaseI, Promega) and 1X DnaseI buffer. DNase was then removed by Phenol-Chloroform-Isoamylic Alcohol Extraction (25:24:1) and RNAs were washed with Chlororoform-Isoamylic Alcohol (24:1). Finally, RNAs were precipitated overnight at -20°C, in 63% ethanol and 0.1M NaOAc pH5.2, and then washed with 1mL of 70% ethanol. The RNA pellet was dried out and resuspended in 32µl of sterile water and quantified using a NanoDrop-1000 Spectrophotometer (Thermo Fischer Scientific).

IV.4 – cDNA synthesis

Reverse transcription for semi-quantitative and quantitative PCR (qPCR) was carried out with 1.5µg of RNA heated for 5min at 70°C with 0.04µg/mL of oligodT. The reverse transcription (RT+) was then carried out by adding 200units of M-MLV Reverse Transcriptase (Promega, 200U/µL) for 1h at 42°C, 1X M-MLV RT buffer, RNasin buffer (40U/µL) and 0.5mM dNTPs. A negative control of reverse transcription (RT-) was performed under the same conditions without the enzyme. Finally, the cDNAs were diluted in sterile water (1:3) and 4µl used as a template for the semi-quantitative PCR (see ***Appendix I*** for primers).

IV.5 – PCR

Genotyping was carried out by PCR amplification using the GoTaq G2 Flexi polymerase (Promega) and 1 to 2µl of DNA template with the specific primers of the gene of interest or of the T-DNA insertion for the mutant lines. The list of primers used is given in ***Appendix I***. General PCR conditions were a first step of DNA denaturation at 95°C for 3min,

then a cycle repeated 36 times composed of one step of DNA denaturation at 95°C for 30sec, one step of annealing for primers at 52°C for 30sec and one step of DNA elongation at 72°C with 1min per Kb depending on the fragment size amplified. Afterwards, a final elongation at 72°C for 10min and a cooling until 20°C were applied.

IV.6 – Agarose gel electrophoresis

PCR products, gDNA and linearised plasmids were separated on 1.5 to 2% agarose gels prepared in 1X Tris Acetate EDTA buffer (TAE: 40mM Tris, 20mM acetic acid and 1mM EDTA). The agarose solution was heated in a microwave until it polymerised and was allowed to cool to 50°C, before addition of 0.625µg/mL of ethidium bromide (Thermoscientific) and pouring into a gel cast. If use of the GoTaq buffer without dye, 5 to 10µl of PCR product was diluted in 6X gel loading dye (NEB) and loaded into the agarose gel wells submerged in 1X TAE buffer. Alongside DNA samples, 6µL of Quickload® 100bp or 1kb+ DNA ladder (NEB) was also loaded. Gels were run at 100V until the dye front reached the end of the gels. DNA bands were imaged using a UV transilluminator (Ultra-Violet Products Ltd., Cambridge, UK) and Uvisave gel documentation camera (UVItec Ltd., Cambridge, UK).

IV.7 – List of vectors

A list of vectors is provided as supplemental tables at the end of this thesis, (***Appendix III, Appendix IV*** and ***Appendix V***).

IV.8 – Gateway cloning

Most of the constructs in this study were generated using Gateway technology (Marsischky and LaBaer, 2004). The first reaction of the Gateway system involves a plasmid pDONR which contains attB1 and attB2 recombination sites, an insertion zone for the DNA

fragment to be cloned, as well as antibiotic resistance allowing the selection of the transformed bacteria. All vectors used for this study are detailed in section ***IV.7*** and ***Appendix III.***

Fragments of interest were amplified by PCR under standard conditions using specific primers containing the Gateway attB1 and attB2 sequences at 5’ and 3' ends of the fragments (see ***Appendix II*** for primer sequences). Sequences of interest were then integrated into the pDONR-vector by a BP reaction. For this, 200ng of PCR products flanked at each end by the attB regions were mixed with 1µl BP clonase ™ II enzyme mix and 150ng of the pDONR-vector in a final volume of 5µl. The reaction mixture was incubated at 25°C for at least 1h and then 1µL of proteinase K (2µg/µL) was added to the mixture for 10min at 37°C to remove the recombinase. This first step yielded the pDONR plasmid containing the gene-of-interest associated with the specific antibiotic resistance cassette. Plasmids were then transformed into DH5α *E. coli* (see ***II.3***) grown on specific plates (see ***II.2***).

pDONR plasmids amplified by well-transformed DH5α *E. coli* were extracted as described below (see ***IV.9***), and submitted to a LR reaction after having been validated by sequencing. For that, 200ng of pDONR clones were mixed with 1µL LR clonase enzyme and 150ng of the destination vector, pDEST (see ***IV.8*** and ***Appendix IV*****)** in a final volume of 5µL. The reaction mixture was treated and inactivated as described above and pDEST plasmids transformed into DH5α *E. coli* (see ***II.3***) grown on specific plates (see ***II.2***).

pDEST plasmids amplified by well-transformed DH5α *E. coli* were extracted as described below (see ***IV.9***) and quality controlled by sequencing. When validated, they were either transformed into *A. tumefaciens* (see ***II.4***) for plant transformation (see ***III.5*** and ***III.6***) or into *S. pombe* (see ***I.3***) for Y2H (see ***I.4***).

IV.9 – Plasmid DNA extraction

"NucleoSpin Plasmid – Plasmid DNA Purification" from Macherey-Nagel and "Monarch Plasmid Miniprep Kit" from NEB were used for plasmid DNA extraction according to the manufacturer's protocol.

V – Protein

V.1 – Protein extraction

For controls in Y2H experiments (see ***I***), but also to test antibodies (see ***V.4*** and ***Chapter 4-III***), protein extracts were obtained from yeast colonies by resuspending them in 300µL of TCA Buffer (10mM Tris-HCl pH8, 20% TCA, 25mM NH_4OAc and 1mM EDTA) followed by an addition of about 100µL of glass microbeads and an incubation of 5min at -80°C. Solution was then vortexed for 1min and incubated on ice for 3min. This step was repeated three times. Lysate was transferred in a new tube (without beads) and washed 2 times with 100µL of TCA Buffer in order to isolate a maximum of proteins. Lysate was centrifuged for 10min at 4°C at 16000g and pellet resuspended in 150µL of Resuspension Buffer (100mM Tris-HCl pH11, 3% SDS). Finally, 150µL of 2X Laemmli buffer was added and samples denatured at 65°C for 10min before loading supernatants on a gel.

Total protein was extracted from non-infiltrated and transiently expressing *N. benthamiana* leaves for western blot analysis (see ***III.5***). Weighed empty 15mL tubes were placed in liquid nitrogen together with a mortar, pestle and spatula to precool. Two to four leaves per sample were cut and veins and midrib were removed. Leaves were ground to a fine powder in liquid nitrogen; the powder was then transferred into the cold 15mL tubes.

A first extraction buffer, developed to extract highly insoluble proteins, was used (0.1M Tris-HCl (pH 6.8), 4.5M Urea, 1M Thiourea, 2% CHAPS, 0.5% Triton X-100, 0.01M DTT, 1% Protein Inhibitor Cocktail (PIC, Sigma P9599-IML), Benzonase, 1.10^{-6}M PMSF).

For each 1g of plant material, 1mL extraction buffer was added. Samples were incubated at 4°C for 1h on a rotating mixer, then centrifuged at 13300rpm, 4°C for 10min. Supernatants were stored at -80°C before proceeding to protein precipitation with 10% of protein extract, 80% ice cold acetone and 10% TCA. Samples were left overnight at -20°C, then centrifuged at 4°C, 13000rpm for 15min. Pellets were washed twice with ice cold acetone and the centrifugation repeated (1mL acetone then 0.5mL in 2mL tubes). Afterwards, all acetone was removed and pellets were resuspended in SDS gel loading buffer (0.0625M Tris-HCl (pH 6.8), 25% Glycerol, 2% SDS, 0.05% Bromophenol blue, 8M Urea, 0.35M DTT; with the DTT stored at -20°C and added just before use). Samples were stored at -20°C.

A second extraction buffer was used in order to enrich protein extracts in nuclei. As previously, proteins were extracted from non-infiltrated and transiently expressing *N. benthamiana* leaves and also from three-week-old *A. thaliana* seedlings grown on plates (see *III-B.*). This buffer was adapted from Xia *et al*, 1997, named "Honda buffer" and composed of 2.5% Ficoll 400, 5% Dextran T40, 0.4M Sucrose, 0.035M Tris-HCl pH7.4, 10mM $MgCl_2$, 2.5mM DTT, 0.5mM PMSF, 0.1% Protease Inhibitor Cocktail (PIC). Once tissue was ground into powder, it was resuspended with 10mL of ice-cold Honda buffer per 1g of plant material. The mixture obtained was filtered through two layers of Miracloth (Calbiochem) and 0.5% (final concentration) Triton X-100 was added to the filtrate. Samples were incubated for 15min at 4°C on a rotating wheel (with head to tail rotation) and then centrifuged at 1500g at 4°C for 5min. Pellets were washed in 1mL Honda buffer to which Triton X-100 had been added to a final concentration of 0.1%, then centrifuged at 1500g, 4°C for 5min and finally washed with 1mL Honda buffer without Triton X-100 before the centrifugation step was repeated. Nuclei pellets were resuspended with 300µL SDS loading buffer or Laemmli 2X (Laemmli buffer (BIORAD) + β-mercaptoethanol 0.7M) and then stored at -20°C.

In order to prepare a protein extract enriched in nuclei and ready for an immunoprecipitation (IP) protocol, an adapted Honda protocol and buffer were used. Honda buffer 2 was made up with 0.02M Tris-HCl pH7.4, 0.44M Sucrose, 1.25% Ficoll 400, 2.5% Dextran T40, 0.01M $MgCl_2$, 0.5% Triton X-100, 0.005M DTT, 0.1% PIC (Roche, cOmplete and cOmplete Mini). Once tissue was ground into powder, resuspension was made with 30mL of ice-cold Honda buffer 2 per 4g of plant material. The mixture obtained was filtered through two layers of Miracloth (Calbiochem) and centrifuged at 2000g, 4°C for 15min. Nuclei pellets were transferred into 1.5mL tubes and washed from one to three times with 1mL Honda buffer 2 and centrifuged at 13000rpm, 4°C for 10min and resuspended in IP buffer (Law *et al*., 2010).

V.2 – Sodium Dodecyl Sulfate Polyacrylamide gel electrophoresis (SDS-PAGE)

Fresh 8% and 10% SDS gels were prepared (resolving gel: H_2O, 8% or 10% Acrylamide, 0.375M Tris-HCl (pH 8.8), 0.1% SDS, 0.05% APS, TEMED; stacking gel: H_2O, 4% Acrylamide, 0.125M Tris-HCl (pH 6.8), 0.1% SDS, 0.05% APS, TEMED). Precast gels from BIO-RAD were also used (8-16% and 10% Mini-PROTEAN TGX Precast Protein Gels).

For PAGE and blotting (see ***V.3***), the BIO-RAD miniprotean system was used. The gels were submerged in 1X Tris-Glycine electrophoresis buffer (5X (1L): 15.1g Tris Base, 94g Glycine, 10% SDS, H2O). Before loading on a gel, samples were boiled at 90°C for 5min or incubated at 37°C for 30min and centrifuged at 16000g for 5min. Gels were run at 150V for 15min and then at 100V for approximately 90min or until the samples had migrated through the gel. The “Precision Plus Protein™ Dual Color Standards” molecular mass marker ladder from Bio-Rad was used for every gel.

Table 2.5: Dilutions of antibodies used in this study

Antibodies & Fluorescent dyes		WB immunostaining
Primary Antibodies	Anti-GFP	1 in 3000 in 3% skimmed milk PBST
	Anti-NEAP1/2	1 in 250 in 3% skimmed milk PBST
	Anti-NEAP wobble	1 in 50 in 3% skimmed milk PBST
Secondary Antibodies	Goat anti-rabbit Cy5	1 in 400 in 3% skimmed milk PBST
	Goat anti-rabbit HRP	1 in 1000 in 3% skimmed milk PBST

V.3 – Western blotting and immunostaining

Three types of membrane were used for blotting including nitrile, PVDF or nitrocellulose membranes. Gels were transferred for 1h at 100V at RT in a BioRad Mini-protean II gel tank with an ice pack and 1X transfer buffer (10% 10X stock [1L: 144g Glycine, 30g Tris-Base, H_2O], 20% methanol, 70% H_2O ice-cold). Some gel transfers were also been carried out with the Trans-Blot Turbo^{TM} Transfer System from BIO-RAD, using BIO-RAD reagents for western blotting.

After transfer, membranes were placed for a few minutes in Ponceau S Staining Solution (0.1% (w/v) Ponceau S in 5% (v/v) acetic acid) to stain total proteins, then rinsed in Phosphate Buffer Saline (PBS). Membranes were blocked with 5% skimmed milk in PBST (0.5% Tween-20 in 1X PBS) for 1h on rotator. Then, blocking solution was replaced by primary antibody (IAb) solution (diluted in PBST-3% skimmed milk, see ***Table 2.5***), and left to incubate O/N rotating at 4°C and 10min at RT. Next, IAb was removed (and stored at -20°C for further use) and the membrane was washed 3 times quickly and 3 times for 10min with PBST. A secondary antibody (IIAb, diluted in PBST-3% skimmed milk, see ***Table 2.5***) was incubated for 1h in the dark at RT. Lastly, membranes were washed 3 times quickly and 3 times for 10min with PBST and stored in PBS at 4°C until imaging (in the dark).

V.4 – Antibody design

Peptides were designed to be highly specific to AtNEAP protein sequences (Eurogentec, Ltd, Southampton, UK). The first peptide (-QLDDKTRSLRE-) was specific to AtNEAP1 and AtNEAP2, including splicing variants, and antibodies generated for this study were named “anti-NEAP1/2”. The other peptide, (-H-DL-D/G-E/H-KK-E/H-SFRRNVVS-C-NH_2-), was specific to all three AtNEAPs but only for a small part of the peptide, the other part being a “wobble” version of the three sequences and antibodies generated with this

sequence were named anti-NEAP wobble. (See ***Figure 4-8***). Before rabbit immunization, pre-immune sera were tested and the results are detailed in ***Chapter 4-III*** and ***Figure 4-9***. Two rabbits were selected and immunized by Eurogentec with the two different peptides according to their "p28 day speedy protocol". At the end of the immunization program, Eurogentec evaluated the antibody titre by ELISA before affinity purification on a peptide column. Final volume before purification was about 30mL and quantity received for purified antibodies from one of the two rabbits were 3.4mL for anti-NEAP1/ 2 (2265µg/mL) and 3.2mL for anti-wobble (406µg/mL). Once the purified polyclonal antibodies received from the company, experiments were carried out to further test and optimize their use.

VI – Microscopy

VI.1 – Wide field microscopy

A MAAF DM 6000 (Leica-microsystems) wide field microscope was used with an Optigrid module for structured illumination and a sCMOS camera (ORCA FLASH 4.0 Hamamatsu) to capture large images of 6.5µm² with 2048x2048pixels containing up to 20-100 nuclei (XY= 0.103µm, Z= 0.2µm). The Optigrid module allowed an automatic deconvolution of the image.

VI.2 – Nuclear and chromatin organization measurements

From 3D image stacks acquired with the wide field microscope (see ***VI.1***), nuclear organization (nuclear morphology and chromatin organization) was quantified using an *ImageJ* plugin named *NucleusJ* (Desset *et al*., 2018; Poulet *et al*., 2014). It included all the necessary steps to process images of nuclei, to perform various analyses and to provide several quantitative parameters to describe the original image. *NucleusJ* then provided a set of

parameters including shape and size of nuclei, size and number of chromocentres as well as their position inside the nucleus relative to the nuclear periphery.

VI.3 – Confocal Imaging

At least two days post-infiltration of *N.benthamiana* leaves (see ***III.5***), transient protein expression was visualised and assessed by a Zeiss LSM 880 confocal microscope and with associated software (ZenLite 2012).

For ***live cell imaging***, an approximate 0.5cm^2 piece of infiltrated leaf was cut out and mounted in water on a microscope slide. A cover slip and a drop of oil were added. The x63 oil immersion lens with x2 zoom factor and CFP laser (10%) and YFP laser (1%) were used to excite, respectively, CFP at 458nm and YFP at 514nm. Emission of CFP was captured by a channel between 463-494nm and YFP emission was captured by second channel between 520-568nm. The pin hole was set at 0.9µm section to avoid cross-channel bleed (1.35 Airy Unit).

VI.4 – apFRET

Acceptor photobleaching Fluorescence Resonance Energy Transfer (apFRET) is based on CFP and YFP fluorophores. As the spectrum of CFP emission and YFP excitation overlap, CFP has the potential to transfer its emission energy to excite YFP. CFP is thus the “Donor” and YFP the “Acceptor”. Energy absorption by YFP is possible only if the two fluorophores are close enough to each other (<100Å). By bleaching YFP, there is no further energy absorption of CFP emission and a resulting rise in CFP emission. FRET efficiency (EF) is then calculated, defining the percentage of energy transferred between the two fluorophores. This is the percentage of rise of CFP emission between pre-bleach and post-bleach images.

To adapt this system for studying protein-protein interactions, one protein is fused to CFP and the other to YFP. If a rise in CFP fluorescence is observed, it indicates that CFP and

YFP were close enough to each other for energy transfer and suggests binding interactions for proteins of interest, (Karpova and McNally, 2006; Karpova *et al*., 2003). All apFRET experiments were performed by adapting the methods described by Graumann *et al*., 2010 using the Zeiss LSM 880 confocal microscope and settings as described above (see ***VI.4***). YFP emission was also measured as a control of YFP bleaching.

A rectangular region of interest (ROI) of 177µm^2 was drawn over a focussed region of NE that was then bleached with 40 iterations of the 514 nm laser at 100%. A total of 50 images were taken at the scan speed of 1 scan per second, 5 before bleaching and 45 post-bleach for a total of 100 constant sized ROI using independent nuclei for every combination of constructs tested.

EF was calculated by subtracting the first post-bleach value from the last pre-bleach value. For control CFP fluorescence values, each pre-bleach fluorescence was subtracted from the previous pre-bleach fluorescence value. The mean of these differences was set as the control EF value and reflected the normal change in CFP fluorescence during imaging.

EF was calculated using the formula $EF = I_t/I_{pre} \times 100$, where I_t is the difference of CFP fluorescence between post- and pre-bleach images, and I_{pre} is the average of pre-bleach CFP fluorescence. EF was expressed as its mean + standard error of the mean (SEM) and was compared to a non-bleached control (control CFP fluorescence before bleaching).

VII – Statistical analysis

The first step of the statistical analysis, carried out in Excel, was to test the equality of variances with an F-test. When the p-value obtained was <0.05, a t-test for unequal variances was carried out. When a p-value of >0.05 was obtained, a t-test for equal variances was carried out. For both t-test, when the p-value was <0.05, the difference between variances was

considered as significant and when the p-value was >0.05, it was determined to be not significant.

VIII – Bioinformatics

VIII.1 – ImageJ

Two main plugins of *ImageJ* (https://imagej.nih.gov/ij/) were used: *NucleusJ* to quantify Nuclear morphology (see ***VI.2***) and SIOX (Simple Interactive Object Extraction) to quantify leaf surfaces during phenotypic analysis of plant material. To measure root growth, the « straight line » tool in *ImageJ* was used.

VIII.2 – Software/Websites

Software used during this study was principally the Microsoft Office package and Adobe Illustrator. For designing primers, cloning *in silico* and sequencing analysis, the Primer3 website (https://primer3.org/), GenePalette (http://www.genepalette.org/), Serial Cloner (http://serialbasics.free.fr/Serial_Cloner-Download.html) and CodonCodeAligner (https://www.codoncode.com/aligner/download.htm) software were used. For image analysis, *ImageJ* (see above paragraph *VIII.B*), and ZenLite 2012 from Zeiss were used (https://www.zeiss.com/microscopy/int/products/microscope-software/zen-lite.html. Finally, for graphs, RStudio (https://www.rstudio.com) and StatGraph (www.statgraphics.com) were used.

RESULTS

CHAPTER 3

Characterization of the AtNEAP protein family

CHAPTER 3
Characterization of the AtNEAP protein family

The function of the AtNEAP protein family at the nuclear periphery remains to be elucidated. In order to investigate the potential role of these proteins in *A. thaliana*, a range of phenotypic analyses have been performed on different combinations of *Atneap* single, double and triple mutants.

In the first instance, only the single mutants *Atneap1* and *Atneap3* and the double *Atneap1Atneap3* mutant were available, with the *Atneap1Atneap3* knock-out (KO) mutant showing reduced primary root growth, (Pawar *et al*., 2016). Before obtaining a triple *neap* KO mutant, an *Atneap2* KO mutant was created using the CRIPSR/Cas9 technique and subsequently crossed with an *Atneap1Atneap3* double mutant. Phenotypic analysis and assessment by *NucleusJ* software, (Poulet *et al*., 2014), of nuclear morphology and chromatin organization changes have been undertaken on two types of leaf epidermal cell populations, pavement and guard cells.

I – Characterization of the triple neap mutant obtained from T-DNA insertion alleles

T-DNA insertion alleles for the three *AtNEAP* genes have been obtained from NASC and crossed to obtain double and triple mutants (Pawar-Menon, PhD thesis, 2015). In order to define if these mutants were loss-of-function, *AtNEAP* transcript levels were assessed by designing specific primer pairs, surrounding the T-DNA insertion sites (pair n°1) or located downstream the T-DNA insertion sites (pair n°2), (Pawar *et al*., 2016).

Plants used were three week-old *Atneap1Atneap2Atneap3* triple mutant (called *first triple*, see ***Methods III.1***) and wild type Col-0 as a control. Total transcript level was assessed

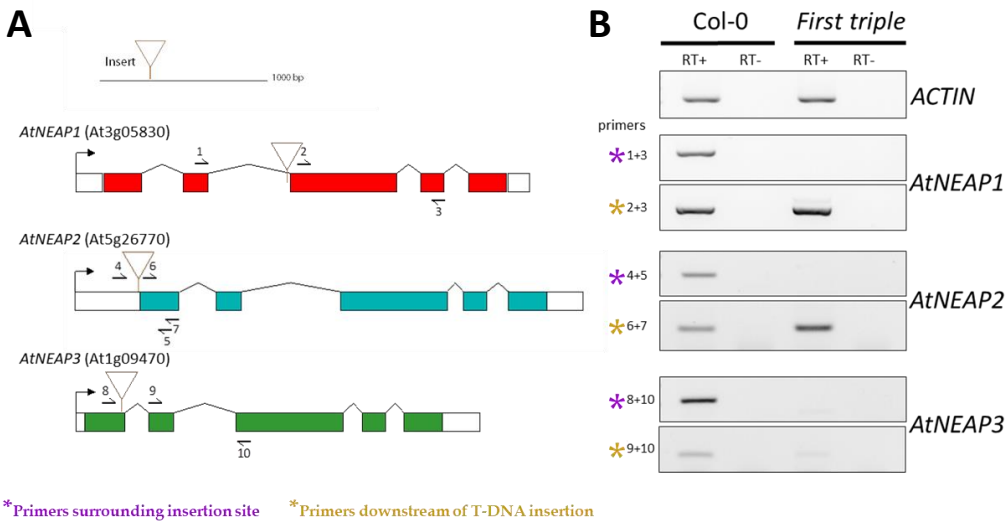


Figure 3-1: Transcript analysis of the *first triple* mutant. **A.** Representation of DNA structures of every *AtNEAP* gene, T-DNA insertions and primer positioning (arrow + number). Exons are indicated as coloured boxes and introns with lines. **B**. Expression of *AtNEAP* genes was assessed by RT-PCR using primers described in A (asterix + primer combination on the left) in wild type (Col-0) and *first triple neap* mutant. Primer pair n°1 surrounding the insertion site (purple), primer pair n°2 downstream of the insertion site (orange). *ACTIN* gene (*ACT2*, At3g18780) was used as a control. RT+ and RT-: with or without reverse transcriptase.

by comparison with *ACTIN* as a reference gene (***Figure 3-1***) using semi-quantitative RT-PCR.

First, the results from the RT-PCR confirmed the T-DNA insertions sites as no bands were detected in the *first triple* mutant with primer pair n°1 for all three *AtNEAP* genes (***Figure 3-1B***). For *AtNEAP1*, a transcript 3' of the T-DNA insertion is visible with primer pair n°2 but given the results from primer pair n°1, this is only a partial transcript potentially initiating in the T-DNA. For *AtNEAP3*, only a phantom band corresponding to a transcript in 3' of the T-DNA insertion was visible with primer pair n°2. It is quite common with T-DNA insertion in plants to observe partial transcripts initiated 3' of the T-DNA insertion although these transcripts are usually not functional. Hence, the *Atneap1* and *Atneap3* alleles are likely loss-of-function, *i.e.* Knock-Out (KO) alleles or at least Knock-Down (KD), (***Appendix VII***, Fig.7H-I in Pawar *et al.*, 2016). For *AtNEAP2*, a transcript 3' to the T-DNA insertion is also visible with primer pair n°2 but this time considering the position of these primers next to the beginning of exon 1, the transcript could be similar to the full length and functional. So, *Atneap2* is likely a functional allele of *AtNEAP2*.

Thus, these results indicated that the *first triple* mutant although containing three T-DNA insertions in each of the *AtNEAP* genes cannot be considered as a complete *AtNEAP* loss-of-function. As no specific antibody was available at that stage of the work to confirm the absence of AtNEAP proteins, it was decided to create a new *Atneap2* KO mutant using the CRISPR/Cas9 technique.

II – Generation of an Atneap2 KO mutant using CRISPR/Cas9 technology

The Cas9 enzyme is able to make a DNA double strand break at specific sites of the genome recognised by a small guide RNA (sgRNA), which recruits the Cas9 enzyme. This system can be adapted in order to choose a specific site where a break is wanted. Due to repair

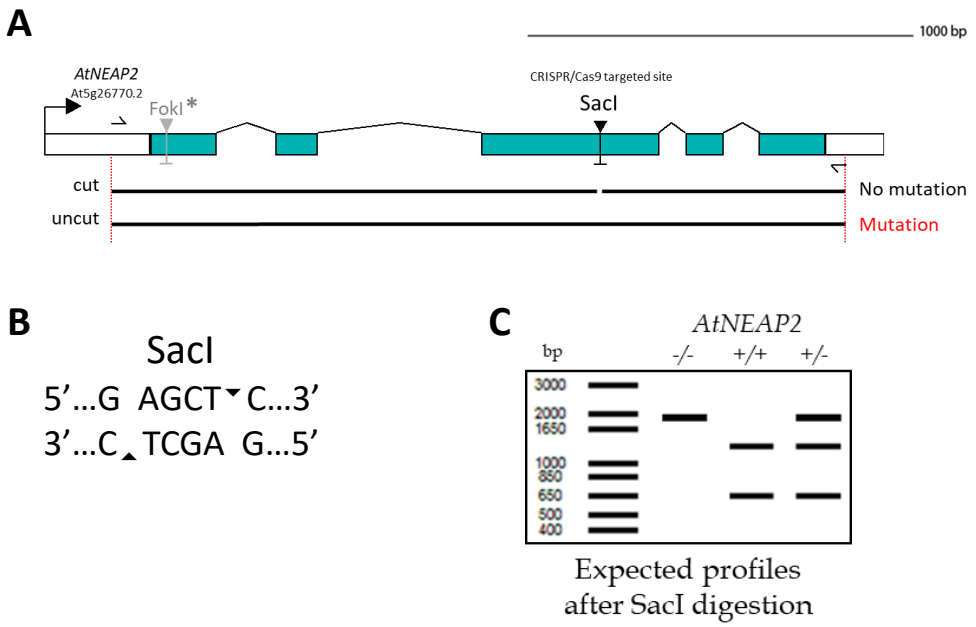


Figure 3-2: Details of *AtNEAP2* targeted site for CRISPR/Cas9. **A.** *AtNEAP2* gene sequence with the targeted site chosen in exon 3 with SacI enzyme restriction site. Primer position used for PCR amplification and expected PCR products are indicated below the gene structure. **B.** Target motif of SacI restriction enzyme. **C.** Expected profile of a mutant for exon 3 targeted site. *FokI restriction site was another CRISPR/Cas9 target site but no mutants were obtained.

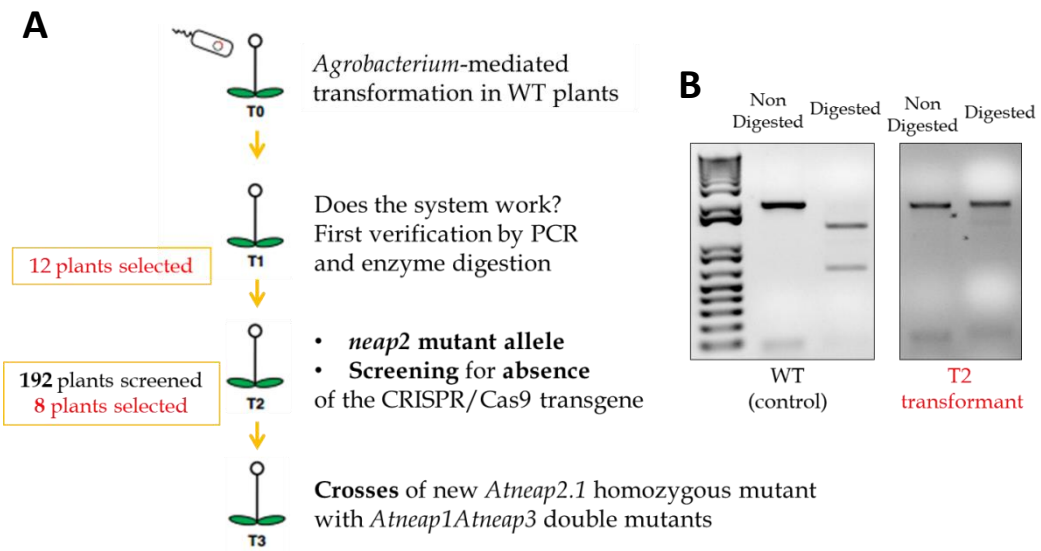


Figure 3-3: Detailed procedure for selection of CRISPR/Cas9 mutants. (adapted from Fauser et al., 2014). **A.** Steps towards the mutant selection. **B.** Typical result after SacI digestion of a wild type (left) and mutant plant (right).

of this DNA double strand break (DSB) by the cellular machinery, mainly through non-homologous-end-joining (NHEJ) in plants (Schiml *et al*., 2017). Different mutations will be created in the gene by addition or deletion of one or more bases.

For the *AtNEAP2* gene, two different target sites were chosen (***Figure 3-2***), and two special Destination vectors were created using Gateway technology, containing the sgRNA sequence specific to one or the other site chosen and also containing the Cas9 gene sequence, (Schiml and Puchta, 2016). Then, plants were transformed via *Agrobacterium tumefaciens* bacteria previously transformed with the cloned vectors.

Target sites chosen were in the first and in the third exons (or second and fourth exons depending on the splicing variant) due to the presence of a restriction site to facilitate screening of mutant plants. When a mutation is present at the target site, the restriction site is destroyed, giving loss-of ability for the restriction enzyme to recognize its site and to cut. In this way, it was possible to screen plants more easily, which have mutations in the *AtNEAP2* gene, (***Figure 3-2***).

Before obtaining a new homozygous mutant for *AtNEAP2*, selections over several generations were required. The first generation (T1) of transformed plants was treated with BASTA antibiotic to select plants, which had incorporated the T-DNA present in the cloned destination vector. Selected plants were left to grow until next generation, as mainly somatic mutations are expected since the Cas9 enzyme is expressed under the Ubiquitin promoter. Then, T2 plants were screened for two different parameters: presence or absence of the CRISPR/Cas9 transgene by PCR with specific transgene primers; and presence of a mutation in the *AtNEAP2* gene by PCR followed by an enzyme digestion of PCR products (***Figure 3-3***).

Plants, which had lost the CRISPR/Cas9 transgene and presenting a full-length band for the *AtNEAP2* gene after digestion, *i.e.* resistant DNA to digestion by restriction enzyme,

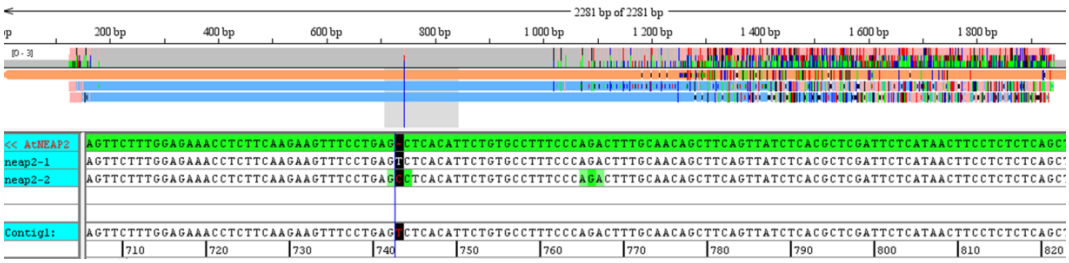


Figure 3-4: Alignment of sequences of *Atneap2-1* and *Atneap2-2* mutants after *AtNEAP2* full length amplification by PCR and sequencing. Software: CodonCodeAligner.

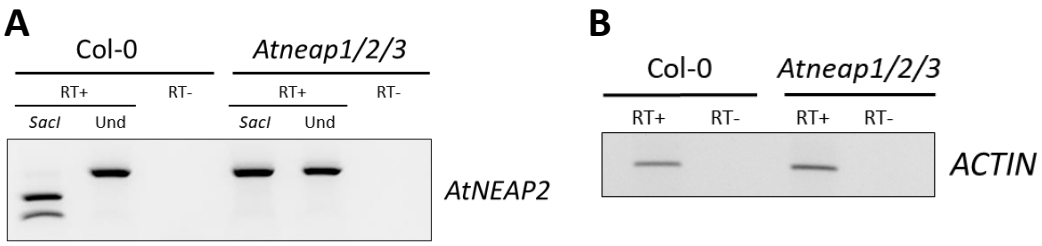


Figure 3-5: Transcript analysis of the *Atneap1/2/3* triple mutant.
A. Transcript accumulation of *AtNEAP2* in wild type (Col-0) and in *Atneap1/2/3*. A transcript is still produced in the *Atneap1/2/3* mutant but presence of the mutation is confirmed as the mutant transcript is resistant to SacI digestion. **B.** Transcript accumulation of the *ACTIN* gene (*ACT2*, At3g18780) is used as a control.

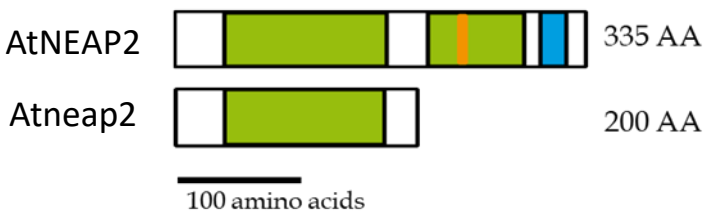


Figure 3-6: Schematic representation of wild type and putative mutant forms of AtNEAP2 proteins. Putative AtNEAP2 would be only 200 amino-acids long and miss one Coiled-Coil domain (green), the Nuclear Localisation Signal (orange) and the Transmembrane Domain (blue).

signifying mutation in the sequence, were grown to the next generation. At that step, only mutations in the fourth exon were obtained, so no further experiments have been carried out to obtain a mutant allele on the second exon as this did not work at the first attempt. T3 plants were then subjected to the same form of selection as for T2, potential homozygous mutants being expected at this stage. Finally, selection was carried out on the next generation (T4) and two different mutant alleles on the fourth exon were obtained for *AtNEAP2*, respectively named *Atneap2-1* and *Atneap2-2*. These two mutants have a single nucleotide insertion in the central exon, a T in position 743 for *Atneap2-1*, and *Atneap2-2* was a probable transheterozygous, e.g two different versions of the insertion at the same position, (***Figure 3-4***).

Each *Atneap2* mutant was then crossed with the *Atneap1Atneap3* double mutant and T2 generations were genotyped in order to find the triple *Atneap* mutant. The probability was only 1/64 to find it in this generation, so plants homozygous mutant for two genes and heterozygous for the third gene were selected in order to obtain the triple mutant easily in the next generation (1/4), and also to be able to see if the triple *neap* mutant was lethal.

Three triple *neap* mutant plants were finally obtained with the *Atneap2.1* allele and the further studies carried out were with the single *Atneap2.1* and triple *Atneap1Atneap2.1Atneap3* (named *Atneap1/2/3*) mutants.

In the meantime, further characterization of the *Atneap2.1* mutant was carried out with RT-PCR analysis and revealed that a transcript was still produced but contained the mutated SacI site, which is easy to follow by using this specific restriction enzyme. (***Figure 3-5***). *In-silico* studies indicated that an early stop codon appears after the insertion site, potentially leading to a truncated protein missing its NLS, one CC domain and the TM domain, (***Figure 3-6***, ***Appendix V***). *In vivo* studies with transient expression in *N. benthamiana* plants of a similar AtNEAP2 truncated protein (missing NLS, one CC and TM domains) fused to a GFP

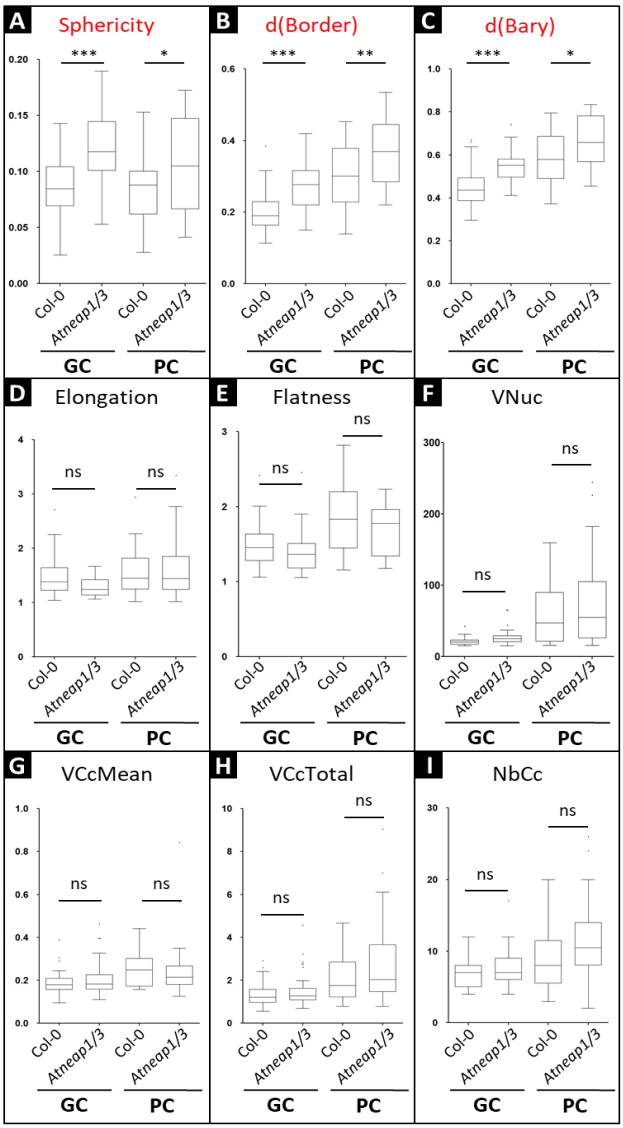


Figure 3-7: Nuclear morphology parameters in *Atneap1Atneap3* double mutant (*Atneap1/3*) and a wild type Col-0 as a control. Plantlets of 14 dag were grown on MS medium. For the guard cell population (GC), a total of 54 nuclei for Col-0 and 63 nuclei for *Atneap1/3*, and for the pavement cell population, a total of 21 nuclei for Col-0 and 36 nuclei for *Atneap1/3* were assessed by *NucleusJ* software. ns: non significant; Nine parameters are shown: sphericity (**A**), distance from chromocentre border (**B**, d(Border)) or barycentre (**C**, d(Bary)) to the NE, elongation **(D),** flatness **(E),** volume of the nucleus (**F**, VNuc), mean of the chromocentre volume (**G**, VCcMean), total volume of chromocentres (**H**, VCcTotal) and number of chromocentres (**I**, NbCc).
ns: non significant; *p< 0.05 ; ** p<0.01 ; *** p<0.001.

tag indicated that the truncated protein is not localised in the nucleus and displays a weak and diffuse GFP staining of the tobacco cells suggesting an unstable protein with a rapid turn-over, (data not shown).

Thus, it seems that the truncated AtNEAP2 protein is no longer nuclear and that the CRISPR/Cas9 induced mutation created a new KO *Atneap2* allele named *Atneap2.1*.

III – Phenotyping and studying nuclear organisation of neap mutants

As mutants of NE proteins often lead to alteration of nuclear shape and/or size, as a first attempt, nuclear morphology of single and double *Atneap* mutants was investigated in root hair cells containing elongated nuclei. However, this assay did not show obvious differences to wild type plants (Pawar *et al*., 2016). In order to establish if there are changes in nuclear morphology or chromatin organisation in a more quantitative manner, calculations from 3D images were carried out, using *NucleuJ* plugin into *ImageJ* software, this time on epidermis cells.

The first experiment used three week old seedlings and was carried out in triplicate with the *Atneap1Atneap3* double mutant (*Atneap1/3*) and a wild type (WT) Col-0 as a control. Results showed that two parameters were significantly modified in the mutant compared to WT; nuclear sphericity and the distances of chromocentre border or barycentre to the nuclear periphery, which were increased in both cell populations studied, namely pavement and guard cells, (***Figure 3-7A-C***). Other parameters, elongation, flatness, nuclear volume, chromocentre volume and number of chromocentres remained constant in the mutant compared to WT, (***Figure 3-7D-I***).

Although no specific phenotype had been observed for single *Atneap1* and *Atneap3*, phenotypic analysis was carried out on the new *Atneap2.1* mutant obtained as well as on the triple *Atneap1/2/3* mutant once available, ***Figure 3-8***. Different growth phenotypes were

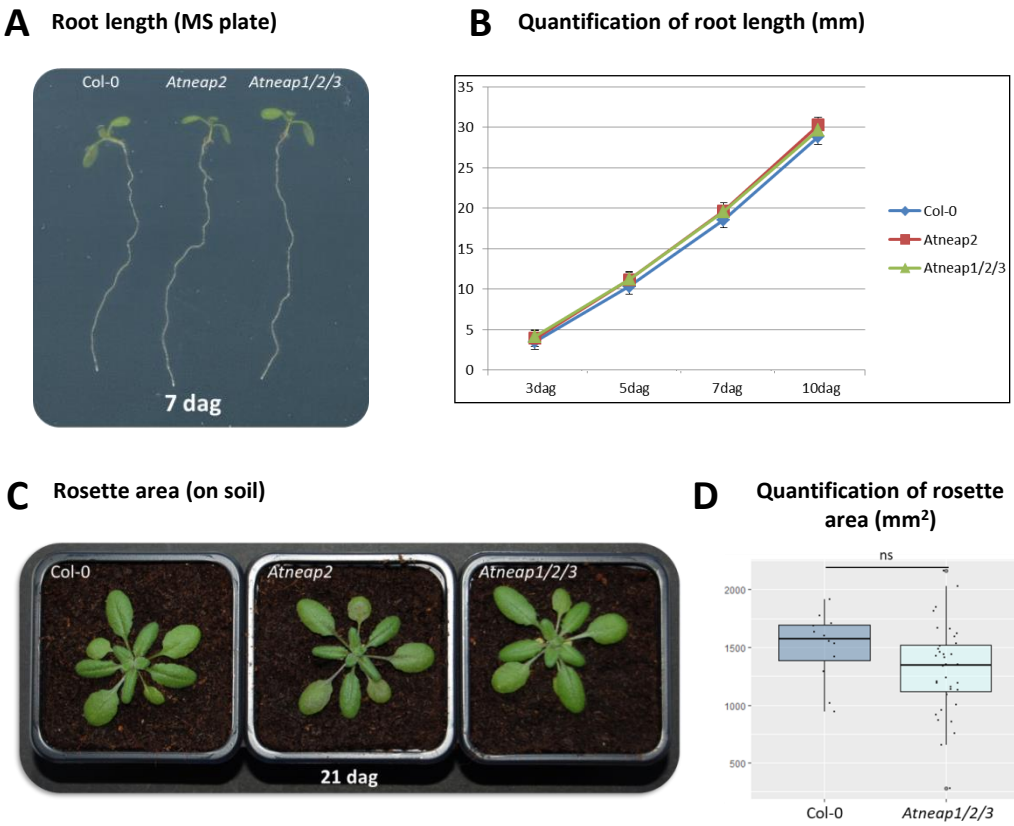


Figure 3-8: General growth phenotype of *Atneap2.1*, *Atneap1/2/3* compared to a wild type Col-0. A. Plantlets of 7dag grown on MS medium. **B.** Quantification of root length (mm) at four time points 3, 5, 7 and 10dag. **C.** Rosette of 21dag plants grown on soil. **D.** Quantification of rosette area (mm^2) of 22dag plants. ns: non significant; * $p< 0.05$; ** $p<0.01$; *** $p<0.001$.

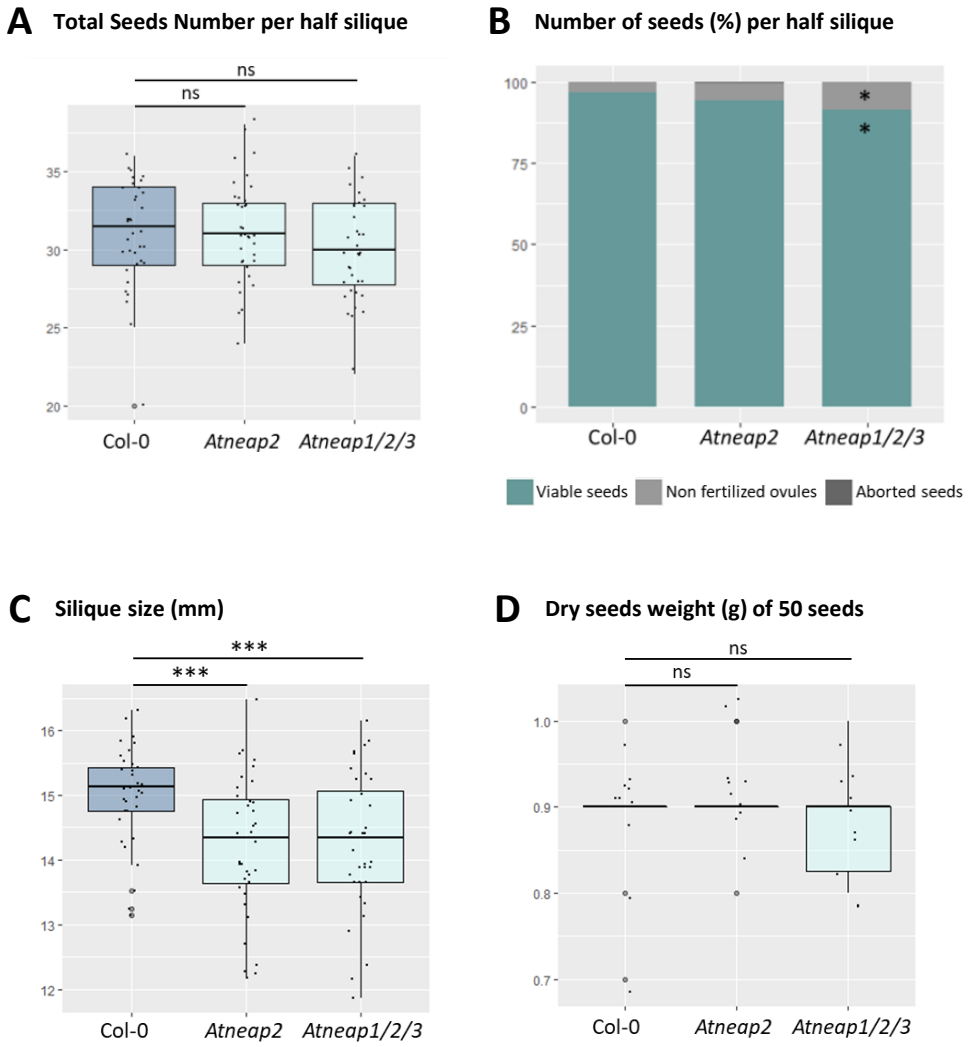


Figure 3-9: Phenotypic analysis of *Atneap2.1*, *Atneap1/2/3* compared to a wild type Col-0. A. Total seeds number per half silique, n=30. **B.** Number of seeds (%) per half silique, n=30. **C.** Silique size (mm), n=30. **A, B, C.** All the plants were grown on soil for a total of 12 plants per genotype. **D.** Dry seeds weight (g) of 50 seeds, n=500.

ns: non significant; * p< 0.05 ; ** p<0.01 ; *** p<0.001.

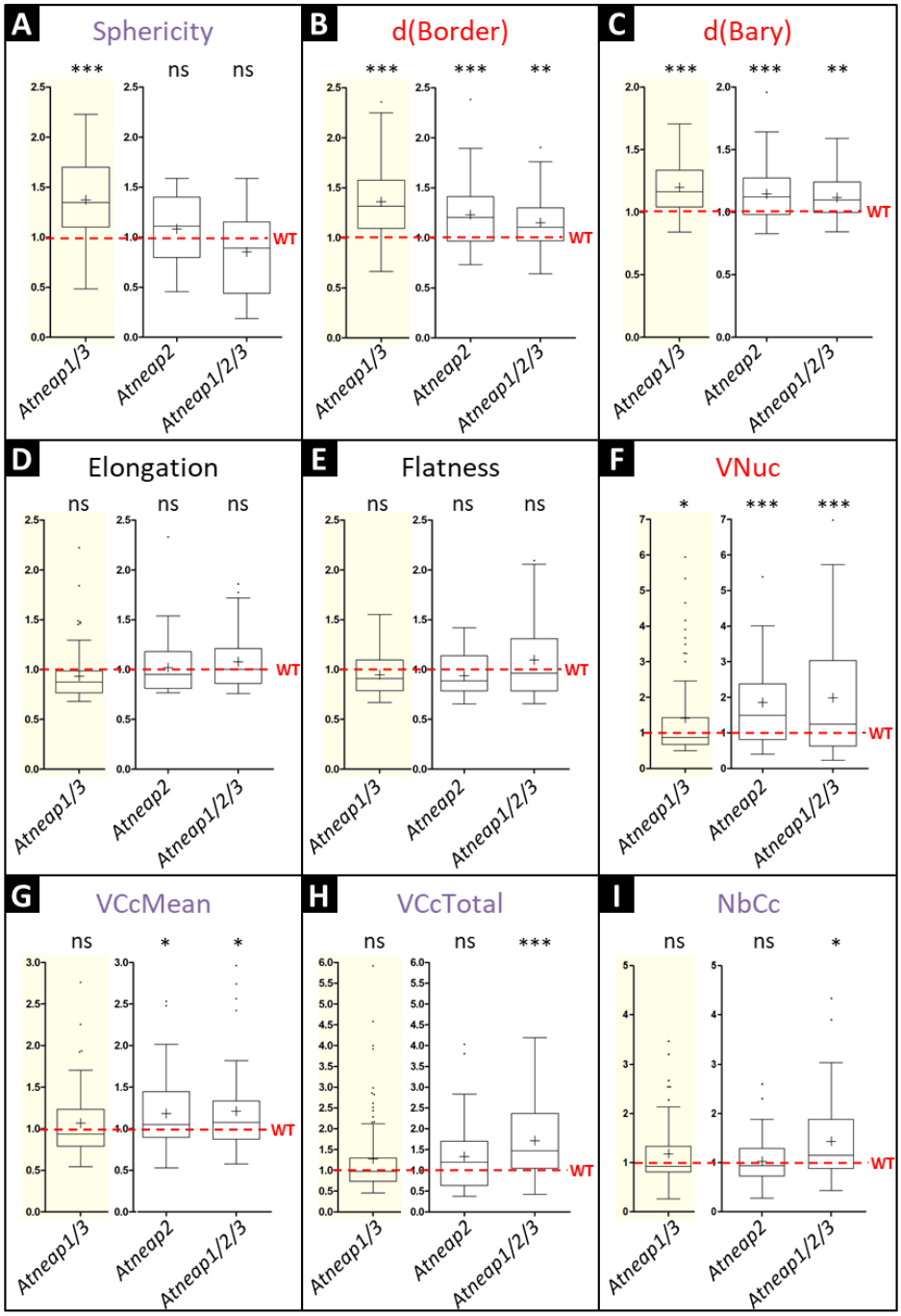


Figure 3-10: Nuclear morphology parameters in *Atneap1/3*, *Atneap2.1*, *Atneap1/2/3* normalized to a wild type Col-0. Plantlets used were 14 dag grown on MS medium. Nuclei were assessed by *NucleusJ* plugin in *ImageJ* software. Yellow parts represent data from the first experiment shown in detail in *Figure 3-7*. Data were normalized with specific WT (Col-0) from every independent experiment. The same nine parameters shown in *Figure 3-7* were studied.
ns: non significant; * p< 0.05 ; ** p<0.01 ; *** p<0.001.

assessed as the root growth kinetics until 10dag on a total of 48 independent seedlings (***Figure 3-8A, B***), leaf surface area at 21dag and the time point of the vegetative/reproductive switch on a total of 12 independent plants (***Figure 3-8C, D***). No difference was observed for those parameters, ***Figure 3-8A-D***.

Then, reproductive tissues were studied for silique size and number and weight of seeds, (***Figure 3-8E-H***). A total of 30 half-siliques from 12 independent plants (***Figure 3-8E, F, G***), and 500 dried seeds (***Figure 3-8H***) were assessed in these experiments. A weak but significantly reduced size of siliques was observed for single *Atneap2.1* and triple *Atneap1/2/3* mutants compared to WT. Also, the triple *Atneap1/2/3* mutant showed less viable seeds, more non-fertilized ovules, but a similar seed weight, indicating a potential defect in meiosis or embryo formation. This is currently being investigated in collaboration with Monica Pradillo, Complutense University of Madrid.

Then, nine nuclear morphology parameters were measured using *NucleusJ* as previously for the *first triple* mutant (***Figure 3-7***). Six independent cotyledon pairs from 21dag seedlings for each genotype, *Atneap2.1*, *Atneap1/2/3* and a WT Col-0 as a control were assessed for a total of 42 to 50 nuclei analysed, ***Figure 3-9***. Results are presented as ratio, normalized to the WT of reference for this experiment. On the left of each panel, in yellow, the results from the first experiment with the double *Atneap1/3* are shown normalized with the WT of reference for the respective experiment. As previously observed for *Atneap1/3*, the distance of the chromocentre border or barycentre, to the nuclear periphery is increased also in *Atneap2.1* and *Atnepa1/2/3* compared to the WT, ***Figure 3-9B, C***. Surprisingly, the increased sphericity observed in *Atneap1/3* was not found in *Atneap2.1* and *Atnepa1/2/3*, with no significant difference compared to the WT, and even a tendency to a lower sphericity for the *triple neap* mutant, ***Figure 3-9A***. Also, nuclear volume is increased in all mutants, ***Figure 3-9F***, and correlates with an increase of the mean of the chromocentre volume in *Atneap2.1*

and *Atnepa1/2/3*, ***Figure 3-9G***. Only the triple *neap* mutant showed an increased total chromocentre volume and number of chromocentres compared to WT, ***Figure 3-9H, I***. Elongation and flatness of the nuclei were not modified ***Figure 3-9D, E***, consistently with the previous experiment.

Thus, in absence of the AtNEAP proteins, nuclei and chromocentres are larger, and chromocentres are at a greater distance from the nuclear periphery, suggesting their tethering to the nuclear periphery is altered.

Conclusion

Using the CRISPR/Cas9 technology, generation of a triple KO mutant for the AtNEAP family was long but successful. Analysis of the general growth phenotype under optimal conditions did not reveal any effect of loss-of-function of *AtNEAP2* only, nor *AtNEAP1*, *AtNEAP2* and *AtNEAP3* simultaneously. Defects in the mutants were only observed when looking at reproductive tissues, being a reduced silique size and a reduced number of viable seeds. These results indicate firstly that major defects could arise during gamete synthesis and fecundation steps. That is why meiosis and embryo formation are being investigated in collaboration. Secondly, while no growth defects were observed under optimal growth conditions, it remains to be investigated how *Atneap* mutant plants respond to different stress conditions.

Results from these studies indicate that the entire removal of this protein family doesn’t have a large effect on general phenotype for plants growing in optimal conditions. Further study of nuclear architecture and chromatin organisation showed that chromocentre position was affected. To investigate whether these changes in chromocentre position impact transcriptional silencing of repetitive sequences organized in chromocentre, analysis is being carried out on the expression of 180bp satellite repeats, the Transcriptional Silent Information (TSI) as well as three euchromatic genes (*UBC28*, *UEV1C* and *HXK1*) in WT Col-0 and triple

mutant *Atneap1/2/3*. This will help us to decide the best strategy to apply before proceeeding to a complete transcriptome analysis at the genome level (RNA-seq) in order to investigate the role of AtNEAP proteins on gene transcription and chromatin organisation, to establish whether this occurs in a site-specific manner or not.

In parallel, a complementation vector, pAtNEAP1::4xc-Myc-AtNEAP1, for *Atneap* mutant lines, was designed, is being synthesized and planned to be transformed into Col-0 and triple *Atneap1/2/3* mutant lines via *A. tumefaciens* (see ***Methods III.6***). Nuclei measurements will be assessed following same procedures see ***Chapter 3-III*** above. The complemented lines could also be interesting tools for the future to explore possible interacting partners by IP using anti-Myc antibody followed by mass spectrometry sequencing.

CHAPTER 4

AtNEAP protein interactome

CHAPTER 4
AtNEAP protein interactome

A second research objective was to identify novel protein partners of AtNEAP proteins to get new insights into their possible function at the nuclear periphery. Three aspects were explored: Y2H screens, characterization of a known interactor of AtNEAP protein AtbZIP18 and the design of antibodies against AtNEAP proteins.

As detailed in (Meng *et al*., 2005), several methods exist in order to investigate protein-protein interactions using yeast in the Y2H and MYTH systems (MYTH tests interactions at a cellular membrane and is particularly suited for membrane proteins) or by using fluorescent imaging-based biophysical techniques such as apFRET, or BiFC. (Pawar *et al*., 2016) showed that AtNEAP proteins were able to form homo- and hetero-dimers through apFRET experiments. In this study, interactions were explored using the classical Y2H system that tests interaction of proteins in the nucleus (Fields and Song, 1989). Experiments were carried out between AtNEAP proteins and known proteins localised at the nuclear periphery, or between AtNEAP proteins and an *A. thaliana* cDNA library to look for novel interactors through new screens.

A previous MYTH screening, using AtNEAP1 as bait (Pawar *et al*., 2016), revealed one interacting protein, a basic-leucine zipper (AtbZIP18), which is a transcription factor, (Gibalová *et al*., 2017). This study showed a co-localisation for AtNEAP1 and AtbZIP18 in the nucleoplasm, (Pawar *et al*., 2016). Thus, apFRET experiments were carried out in order to confirm this suggested interaction as well as bZIP18 domain deletion constructs to characterize the specific interaction site with AtNEAPs by Y2H experiments.

Finally, in order to study the AtNEAP interactome, specific antibodies against AtNEAPs were required for pull-down assays followed by mass spectrometry analysis. Thus,

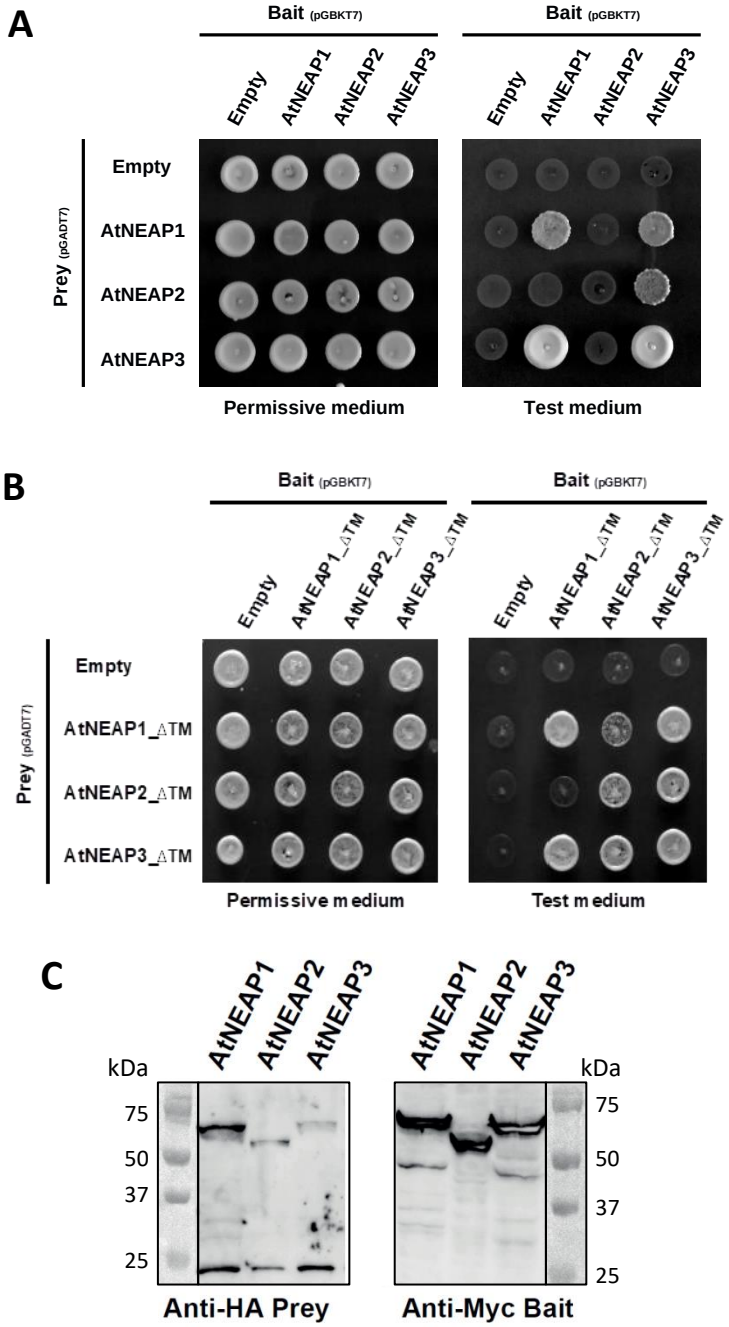


Figure 4-1: Interaction between AtNEAPs and AtNEAPs_ΔTM proteins with each other respectively. Yeast strains were tested on permissive medium (left) depleted in leucine and tryptophan to select diploids only and on test medium (right) depleted in leucine, tryptophan, histidine and adenine to select diploids with interacting proteins. **C.** WB to confirm expression of AtNEAP proteins as baits (c-Myc antibody) and preys (HA antibody).

antibodies were designed, produced in rabbits, verified and tested on different protein extracts.

I – Classic Yeast Two Hybrid (Y2H)

Only the classical Y2H system was used during this study. In the first instance, drop tests were undertaken with known baits and preys. Yeast strains containing prey or bait vectors were mated, and zygotes grown on permissive medium. Interaction with bait and prey was then tested on selective medium depleted in Trp, Leu, His and Ade. The first experiment, (***Figure 4-1A***), testing interaction of full length AtNEAP proteins with each other (baits and preys) confirmed homodimerization for AtNEAP1 and AtNEAP3 and heterodimerization for AtNEAP1-AtNEAP3 and AtNEAP2-AtNEAP3. As the TM may sequester the AtNEAP proteins at the membranes (nuclear, plasma or ER) and could impair the transcriptional activation of the reporter genes in Y2H, the same experiment was repeated with AtNEAPs depleted of the TM domain, (***Figure 4-1B***). The results showed the same interactions as previously observed ***Figure 4-1A***, *i.e.* AtNEAP1 and AtNEAP3 homodimerization and AtNEAP1-AtNEAP3, AtNEAP2-AtNEAP3 heterodimers. Results also showed AtNEAP2-AtNEAP1 heterodimer and AtNEAP2 homodimer. These results in Y2H are consistent with the ones observed by (Pawar *et al*., 2016) by apFRET and MYTH (Pawar-Menon, PhD thesis, 2015), with every AtNEAP homodimer and every combination of heterodimers. They indicate that although the TM domain does not block transcriptional activation of the reporter genes, some interactions were revealed only when the TM domain is deleted.

In the experiment shown in ***Figure 4-1C***, bait and prey fusion proteins were extracted from yeast (see ***Methods V.1***) in order to confirm the synthesis of those proteins in the yeast system. Prey constructs have a HA tag and bait constructs have a c-Myc tag allowing detection on a blot with specific tag antibodies.

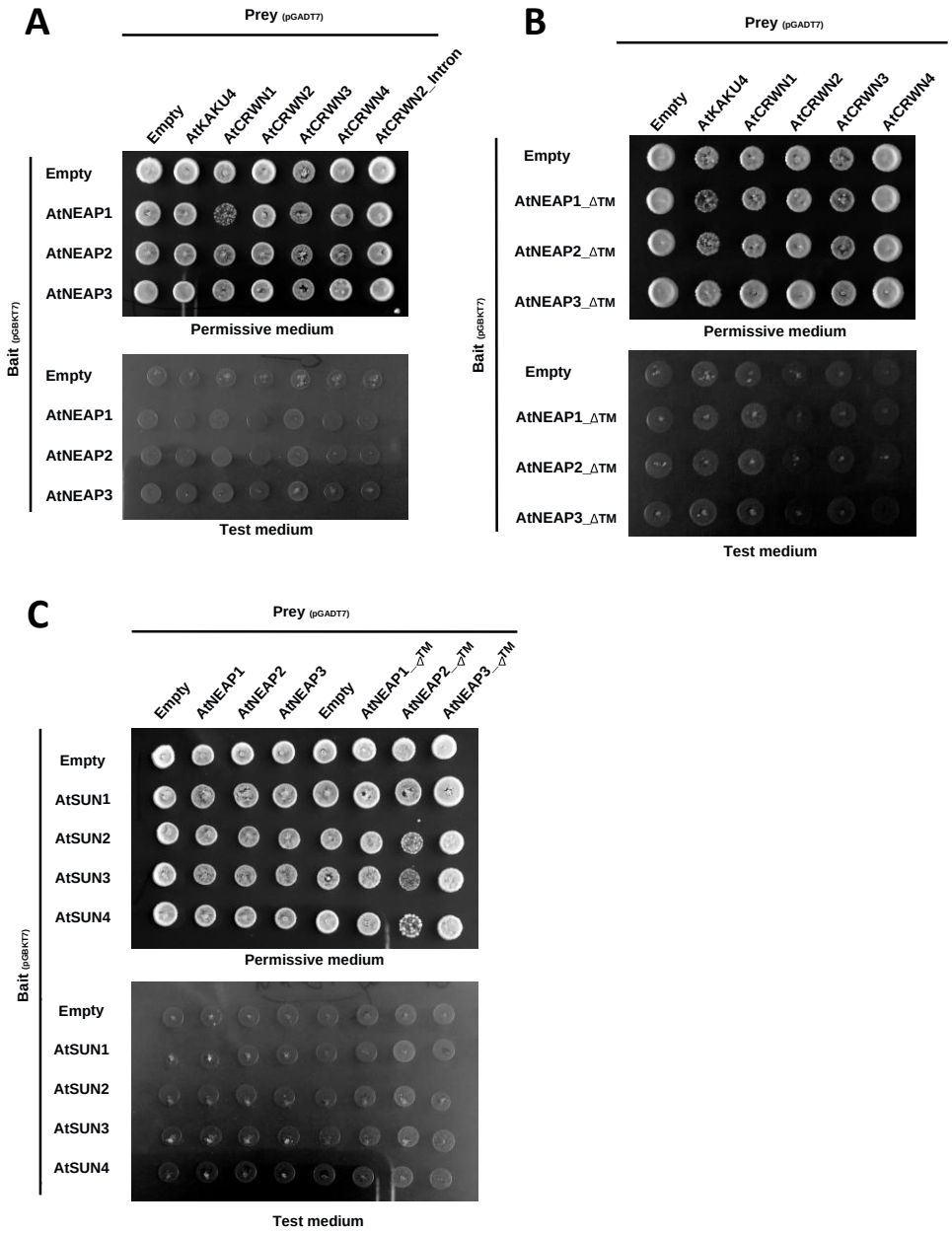


Figure 4-2: Interaction between AtNEAP proteins and known nuclear periphery and INM proteins. Yeast strains were tested on permissive medium (left) depleted in leucine and tryptophan to select diploids only and on test medium (right) depleted in leucine, tryptophan, histidine and adenine to select diploids with interacting proteins. Bait AtNEAPs (**A**) and AtNEAPs_ΔTM (**B**) vs prey AtCRWN1-4 and AtKAKU4. **C.** Bait AtSUN1-4 vs prey AtNEAPs and AtNEAPs_ΔTM.

Next, AtNEAP protein interactions were tested with other known proteins situated at the nuclear periphery (AtKAKU4 and AtCRWNs) or from the INM (AtSUN domain proteins). AtKAKU4 and AtCRWN1-4 as preys were tested against AtNEAPs as baits, either full length or without the TM domain ***Figure 4-2A, B***. The same experiment was carried out with AtSUN1-4 as baits and AtNEAPs as preys, ***Figure 4-2C***. For these drop tests, no interaction at all was detected on selective medium.

In order to look for new interactors without any *a priori* using Y2H, library screenings were carried out with each full length AtNEAP as bait. The prey library was composed of 2 million independent cDNA clones, number determined thanks to serial dilutions made from the mated culture and a count of the number of diploïds obtained on plates with permissive medium. The mating efficiency was 9.4% for the prey library crossed with bait AtNEAP1, 12.8% for the prey library crossed with bait AtNEAP2 and 18.9% for the prey library crossed with bait AtNEAP3. As the minimum for mating efficiency is 5%, this experiment was validated, but, unfortunately, on the selective medium, very few clones were recovered from these screens. Each of them was tested to avoid false positive or contamination. After those tests, only one clone for AtNEAP1 (AT1G45474), none for AtNEAP2 and two for AtNEAP3 (AT2G22360 and AT1G51510) corresponded to potential interacting proteins. As the cDNA library was constructed by cloning the cDNAs in three possible frames, sequencing of these clones was performed to define if the expressed proteins were in the correct frame. Finally, it appeared that none of the candidate clones were in the +1 frame. This whole Y2H screen experiment was repeated at another time with similar results and for that reason the Y2H investigation with the cDNA library was stopped at this stage.

In summary, AtNEAP proteins interact with each other as homo- or heterodimers but despite quite significant efforts to evaluate their interactions with known proteins of the INM

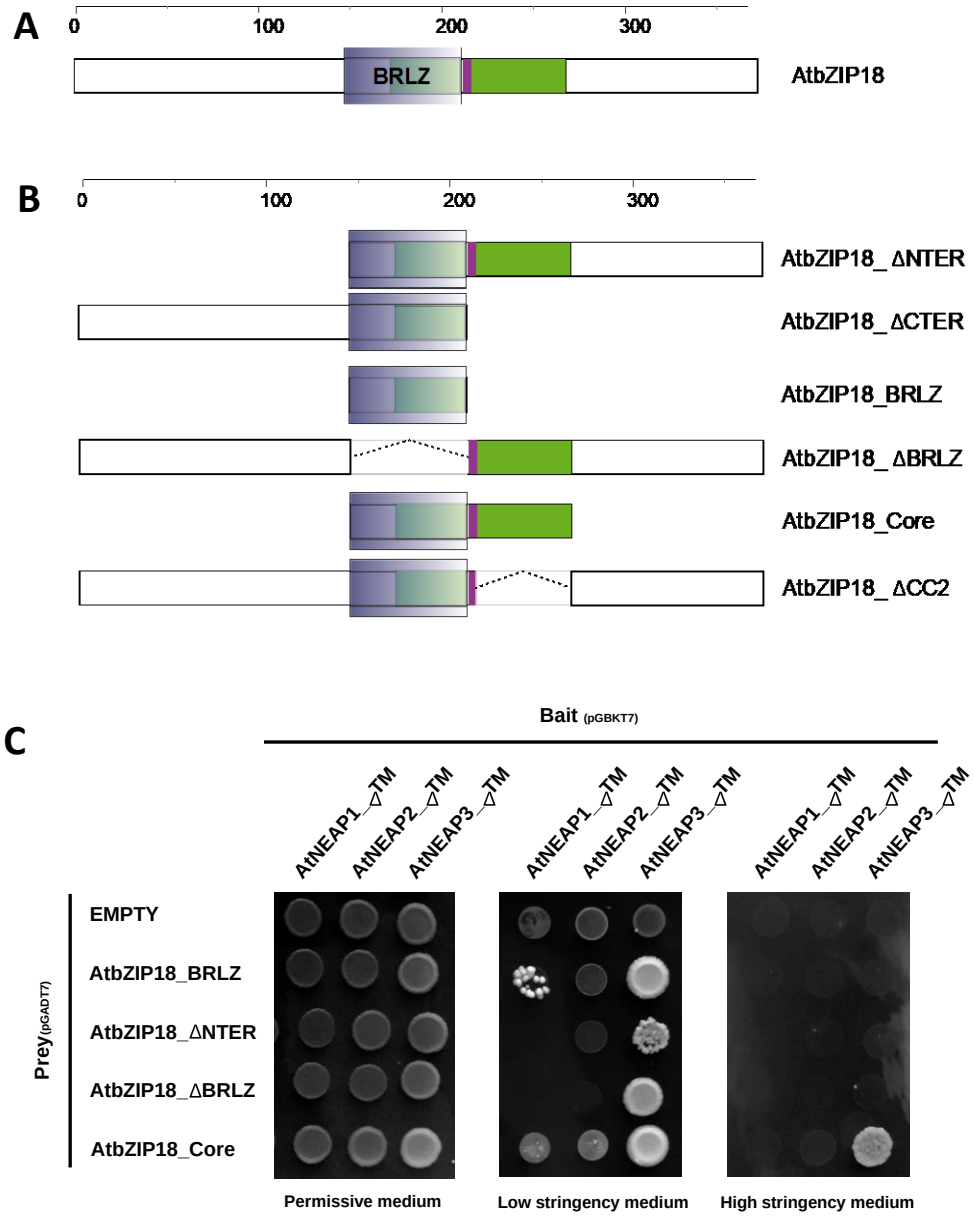


Figure 4-3: AtbZIP18 domain deletion constructs and Y2H experiments. A. Schematic representation of AtbZIP18 full length with BRLZ for basic DNA-binding domain leucine-zipper (blue), coiled-coil domains in green and EAR motif in purple. **B.** Domain deletion constructs. **C.** Y2H drop tests with 4 on 6 AtbZIP18 deletion constructs and AtNEAPs_ΔTM. Permissive medium is depleted in leucine and tryptophan; Low stringency medium is depleted in leucine, tryptophan and histidine; High stringency medium is depleted in leucine, tryptophan, histidine and adenine.

and the nuclear periphery (AtSUN domain proteins, AtKAKU4 and AtCRWNs) as well as the search of new interactors (cDNA library screenings) no new protein partners were identified.

Thus, focusing on the interaction partner previously evidenced by (Pawar *et al*., 2016), AtbZIP18 as bait was crossed with AtSUN1-4, AtKAKU4 and AtCRWN1-4 as preys to investigate their possible interaction with AtbZIP18. Unfortunately, first attempts of Y2H experiments revealed that AtbZIP18 as prey was really slow to grow and AtbZIP18 as bait was auto-activating, consistent with the observations of (Gibalová *et al*., 2017) who removed the auto-activation domain located in the N-ter of the protein, before the BRLZ domain, ***Figure 4-3A***. In the meantime, domain deletions of AtbZIP18 were designed and constructed, as shown in ***Figure 4-3B***, in order to better characterize the specific interaction domain with AtNEAPs. Some of the domain deletion constructs were missing the N-ter auto-activating domain, so normal Y2H experiments were carried out with them, ***Figure 4-3C***. The constructs AtbZIP18_BRLZ, AtbZIP18_ΔNTER, AtbZIP18_ΔBRLZ and AtbZIP18_Core as preys were crossed with AtNEAP1_ΔTM, AtNEAP2_ΔTM and AtNEAP3_ΔTM as baits. Diploïds were grown on permissive, low and high stringency media. Results showed that on high stringency medium, only AtNEAP3_ΔTM interacted with the core of AtbZIP18 containing the BRLZ domain, the EAR motif, and CC domains, ***Figure 4-3B, C***. On the low stringency medium, interactions were observed for AtNEAP3_ΔTM with every domain deletion construct. Also, the core of AtbZIP18 interacted with AtNEAP1_ΔTM and AtNEAP2_ΔTM; and the BRLZ domain of AtbZIP18 interacted with AtNEAP1_ΔTM. This full experiment was repeated with full length AtNEAPs but no interaction was observed (data not shown).

The constructs AtbZIP18_BRLZ, AtbZIP18_ΔNTER, AtbZIP18_ΔBRLZ and AtbZIP18_Core as full length AtbZIP18 were fused to GFP and transformed in *N. benthamiana* showing a nucleoplasmic localisation (data not shown).

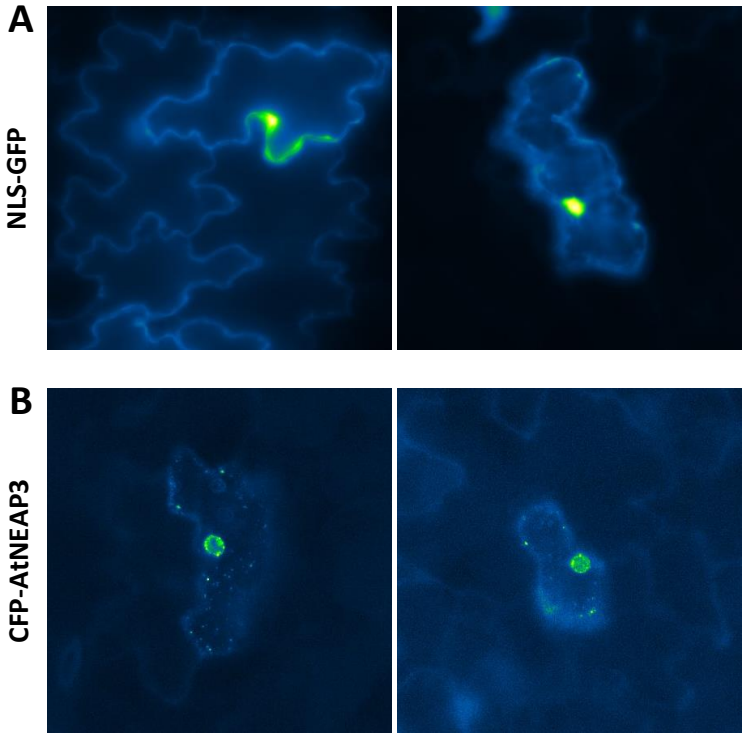


Figure 4-4: Transient transformation of *A. thaliana* cotyledon epidermal cells. Plantlets of 6 dag expressing (**A**) NLS-GFP construct or (**B**) CFP-AtNEAP3 construct. Observation were performed 48h after co-cultivation with *A. tumefaciens* at OD_{600} = 0.5 (**A)** and OD_{600} = 1 (**B**). MMAF microscope, objective X63.

In previous MYTH experiments performed by Maxime Voisin, a former PhD student of the GReD team, another transcription factor, AtMaMYB (At5G45420), identified as an AtSUN3 interactor also interacted with AtbZIP18, ***Appendix VII***. Furthermore, a preliminary MYTH experiment suggested that AtNEAP1 also interacts with AtMaMYB, work achieved by Maxime Voisin. These new interactions will have to be confirmed and extended to AtNEAP2 and AtNEAP3 but already suggest that AtNEAP1 may interact with at least two transcription factors at the nuclear periphery and that a potential network with transcription factors, which would be anchored at the nuclear periphery, is emerging, ***Appendix VII***.

II – Localisation, Co-localisation and apFRET

Transient expression in tobacco is a frequently used technique for Arabidopsis research projects as an alternative to transient expression in Arabidopsis cell culture lines which are not as well developed as an experimental system as cell culture in animals. Hence, the tobacco transient expression assay is a technique of choice when Arabidopsis is the model of study but it remains a heterologous system. In order to observe protein localisation *in vivo* but in Arabidopsis, the FAST technique of (Li and Nebenführ, 2010) was developed for Clermont-Ferrand laboratory conditions. The main aim was to study the localisation of the AtNEAP proteins in WT plants and also to visualise any change of localisation of AtNEAP proteins in the available collection of NE protein mutants when transiently expressed in Arabidopsis.

The protocol was first tested with a chimeric fusion protein, NLS-GFP, used as a positive control for nuclear localisation. As shown in ***Figure 4-4A***, NLS-GFP is located in nuclei and confirmed that the protocol was working. Then, AtNEAP proteins fused with CFP in the N-ter were transiently transformed into Arabidopsis seedlings. Positively transformed cells were observed only for the CFP-AtNEAP3 construct and confirmed the nuclear

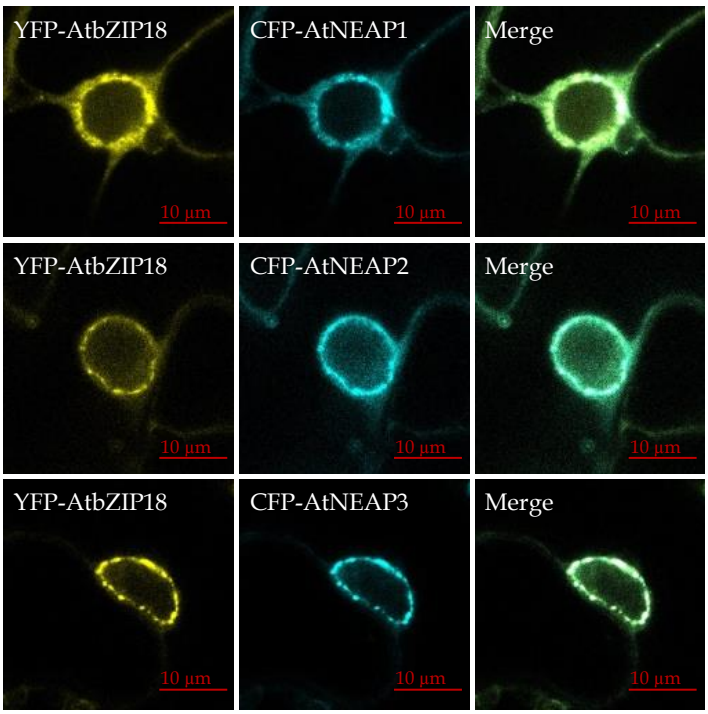


Figure 4-5: Transient expression in leaf epidermal cells of *N. benthamiana*. Three week old plants were co-infiltrated with *A. tumefaciens* containing p19 at DO_{600} = 0.5, YFP-AtbZIP18 construct at DO_{600} = 1 and depending on the condition, CFP-AtNEAP1, CFP-AtNEAP2 or CFP-AtNEAP3 construct at DO_{600} = 1. Images were obtained using a Zeiss LSM800 confocal microscope using a X63water-immersion objective. Scale bar = 10µm.

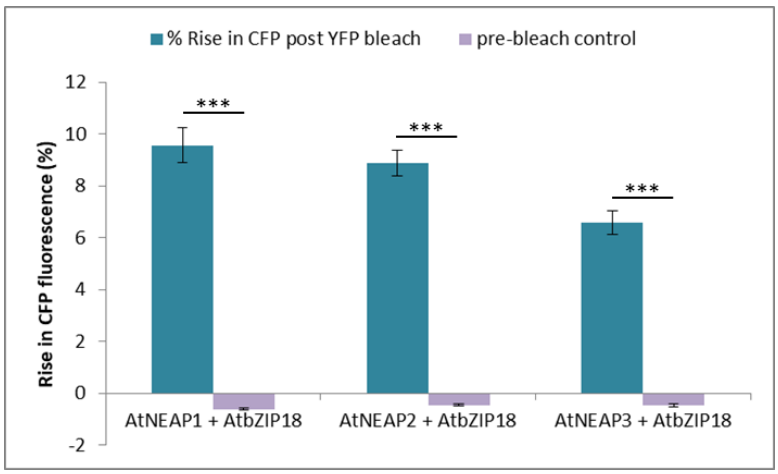


Figure 4-6: AtbZIP18 interact with AtNEAPs. apFRET was performed using 105 nuclei for AtNEAP1+AtbZIP18, 100 nuclei for AtNEAP2+AtbZIP18 and 97 nuclei for AtNEAP3+AtbZIP18. *** $p<0.001$

periphery localisation of AtNEAP3 *in vivo* in *A. thaliana*, ***Figure 4-4B***. Unfortunately, this protocol was not efficient: whilst if in most of the replicates 5 to 6 out of 8 seedlings were expressing the NLS-GFP construct, some of the replicates did not express AtNEAP and in the best experiments only 3 out of 6 seedlings expressed the CFP-AtNEAP3 construct. Given the difficulties to set up the protocol in WT, it was decided not to investigate mutant backgrounds or the co-infiltration of two constructs to observe co-localisation. Therefore, all the subsequent experiments of transient expression were performed in *N. benthamiana*.

Infiltrations of *N. benthamiana* leaves were carried out to study co-localisation and interactions with the apFRET technique. As shown in ***Figure 4-5***, in each case, both AtbZIP18 and one of the AtNEAPs co-localised at the nuclear periphery. Note that this is significantly different to previous data (Pawar *et al*., 2016) where AtbZIP18 was nucleoplasmic and not only restricted to the nuclear periphery when co-expressed with one of the AtNEAPs. Co-localisation indicates that two proteins localise in the same area but in order to prove interactions, *i.e.* a close proximity <100Å, apFRET experiments were carried out on YFP-AtbZIP18 and every CFP-AtNEAP fusion protein. Two independent replicates were carried out per combination with apFRET on 43 and 44 nuclei respectively for AtNEAP1+AtbZIP18 (N1 combination), 45 and 44 nuclei for AtNEAP2+AtbZIP18 (N2 combination), 45 and 52 nuclei for AtNEAP3+AtbZIP18 (N3 combination). An additional replicate was done for N1 and N2 combinations with 18 and 11 nuclei assessed respectively. The percentage of rise in CFP emission post YFP bleach for every condition was then calculated on a total population of 105 nuclei for N1, 100 nuclei for N2 and 97 nuclei for N3. Results shown in ***Figure 4-6*** confirmed interaction between AtbZIP18 and all three AtNEAP proteins with a p-value < 0.001 and suggest that the transcription factor AtbZIP18 can be tethered at the nuclear periphery maybe through its interaction with AtNEAP proteins. Further

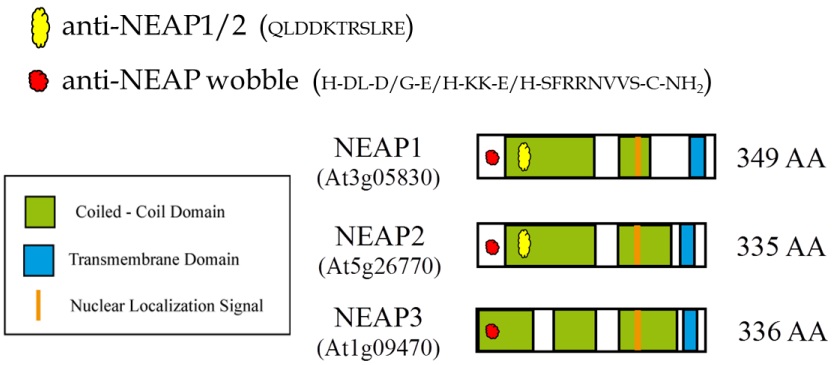


Figure 4-7: Schematic representation of target sites of specific antibodies.

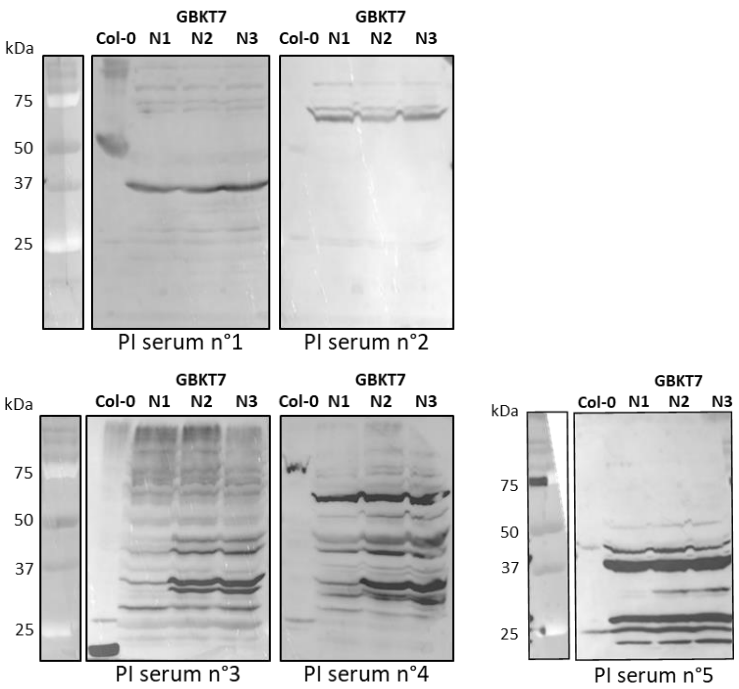


Figure 4-8: Test of Pre-immune (PI) sera on yeast extracts containing bait AtNEAP fusion proteins* and a WT Col-0 plant extract. Pre-immune sera were diluted at 1/1000 in 3% PBST-milk, and interaction revealed by a secondary goat anti-rabbit-HRP diluted at 1/50 000. *Bait AtNEAP fusion proteins are the ones used in Y2H experiments, *i.e.* pGBKT7-AtNEAPs construct and revealed previously with anti-Myc antibody (see *Figure 4-1C*).

experiments using deletion mutants and *in vivo* studies in mutant backgrounds will be needed to confirm this hypothesis.

III – Generation of AtNEAP antibodies

Specific AtNEAP antibodies were not available at the beginning of the project and were required for several reasons. Firstly to better characterize different *Atneap* mutants, via Western blotting. Secondly to investigate protein-protein interaction via immunoprecipitation (IP) and co-IP; thirdly for immunohistochemistry of *A. thaliana* tissues.

In order to produce antibodies, two different peptides were designed in collaboration with Eurogentec to be highly specific to AtNEAP protein sequences (See ***Methods V.4***). The first peptide of 11 aa is located in the N-terminal part of the AtNEAP proteins within the first CC domain and is specific to AtNEAP1 and AtNEAP2, splicing variants included (used to produce "*anti-NEAP1/2*" antibody). The second peptide is also of 11 aa and situated before the CC domain of AtNEAP1 and AtNEAP2 and at the beginning of the first CC domain of AtNEAP3. This peptide is designed to recognize all three AtNEAPs and includes a wobble version of the three sequences (used to produce "*anti-NEAP wobble*" antibody), (***Figure 4-7***). Before inoculating rabbits, pre-immune (PI) sera from five different rabbits were tested on yeast extracts containing AtNEAP fusion proteins and a total protein extract of a WT Col-0 plant, ***Figure 4-8***. The five pre-immune sera revealed a non-specific band around 25kDa for every sample. In Col-0 plant extract, a band around 50kDa was detected in PI n°1 and n°2; and PI n°4 revealed a band around 75kDa. Also, for yeast extract, PI n°5 did not detect any band above 50kDa, ***Figure 4-8***. Thus of five rabbits, only two were kept, n°3 and n°5, as they did not detect a band of similar size to the AtNEAP fusion proteins (60-70kDa) or similar to native proteins (30-40 kDa). The two selected rabbits were immunised by Eurogentec with the two different peptides according to a "p28 day speedy protocol" (immunization in 28 days), See ***Methods V.4***.

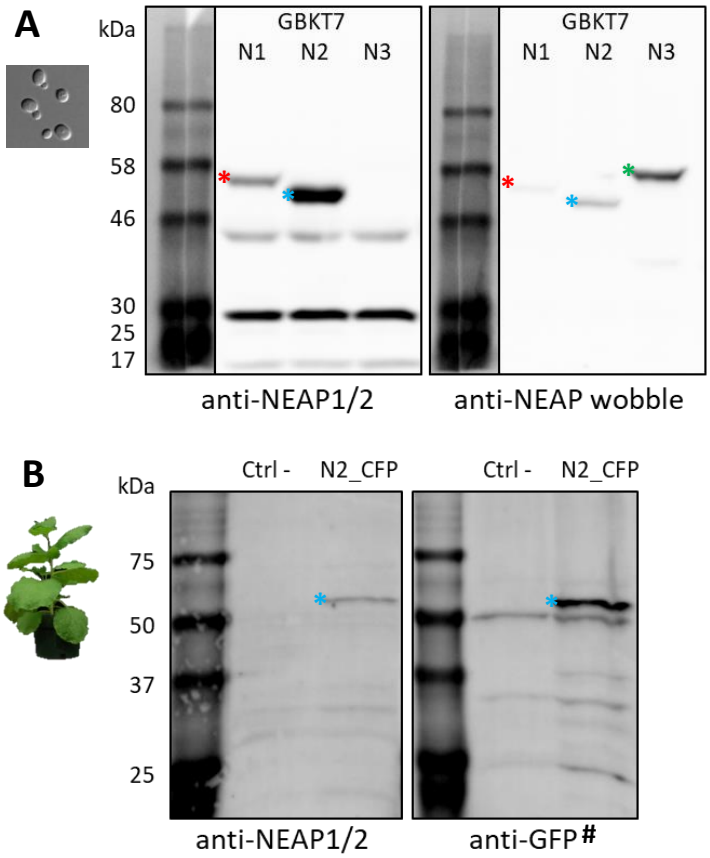


Figure 4-9: Test of anti-NEAP1/2 and anti-NEAP wobble antibodies on different protein extracts. A. Proteins were extracted from yeast (*S. cerevisiae*) containing over-expressed bait fusion AtNEAP proteins used in Y2H experiments. From left to right: pGBKT7-AtNEAP1, pGBKT7-AtNEAP2, pGBKT7-AtNEAP3. *expected bands. **B.** Proteins were extracted from infiltrated *N. benthamiana* plants with AtNEAP2-CFP. A non-infiltrated plant was used as a negative control (Ctrl -). Red, blue or green asterisks showing bands representing AtNEAP1, AtNEAP2 or AtNEAP3 respectively. Antibody dilutions were 1 in 250 for anti-NEAP1/2 and 1 in 50 for anti-NEAP wobble, and are available in ***Methods Table 2.5*** for more details. **#**. As CFP and GFP have similar protein sequences, anti-GFP antibody is able to recognize either GFP or CFP.

Once the purified polyclonal antibodies received, they were tested on several protein extracts from different organisms. Firstly, same protein extracts used in the tests of the PI sera were used, containing AtNEAP fusion proteins over-expressed in yeast (***Figure 4-9A***). Proteins synthesised from yeast were rather soluble and easy to run on SDS-PAGE. As observed in ***Figure 4-8*** with PI sera (n°3 and n°5), some bands were visible around 30 and 40kDa but at a different expected size compared to AtNEAP fusion proteins and for that reason are considered as non-specific bands. Indeed, according to *in-silico* predictions with Serial Cloner software, expected bands for pGBKT7-AtNEAP1 construct should be around 57kDa, pGBKT7-AtNEAP2 around 54kDa and pGBKT7-AtNEAP3 around 55kDa (as estimated molecular weight (MW) of Gal4 DNA-binding domain (GalDBD) from the pGBKT7 construct is 16kDa; AtNEAP1: 41kDa; AtNEAP2: 38kDa; AtNEAP3: 39kDa). In ***Figure 4-9A***, on both left and right panels, with anti-NEAP1/2 and anti-NEAP wobble respectively, bands of approximately 55kDa for AtNEAP1 and 50kDa for AtNEAP2 fusion proteins were detected. Only with anti-NEAP wobble a band was visible around 56kDa for AtNEAP3 fusion protein. Bands were approximately at the expected MW even though AtNEAP1 and AtNEAP2 fusion proteins appeared at a lower MW and AtNEAP3 fusion protein at a bigger MW. These results are consistent with the ones presented in ***Figure 4-1C*** and also in Pawar-Menon, PhD thesis, 2015, with observed bands at 61, 60 and 65 kDa for AtNEAP1, AtNEAP2 and AtNEAP3 respectively for NEAPs fused to YFP when revealed with an anti-GFP antibody (YFP being around 27kDa). Thus, using fusion proteins expressed in yeast, the anti-NEAP1/2 antibody detected AtNEAP1 and AtNEAP2 while the anti-NEAP wobble detected all the AtNEAPs, showing higher affinity for AtNEAP3. This result was not expected from the initial design of the peptides but offers the advantage of two antibodies with different specificities. Finally, the anti-NEAP wobble shows a lower efficiency and requires to be used at a lower dilution.

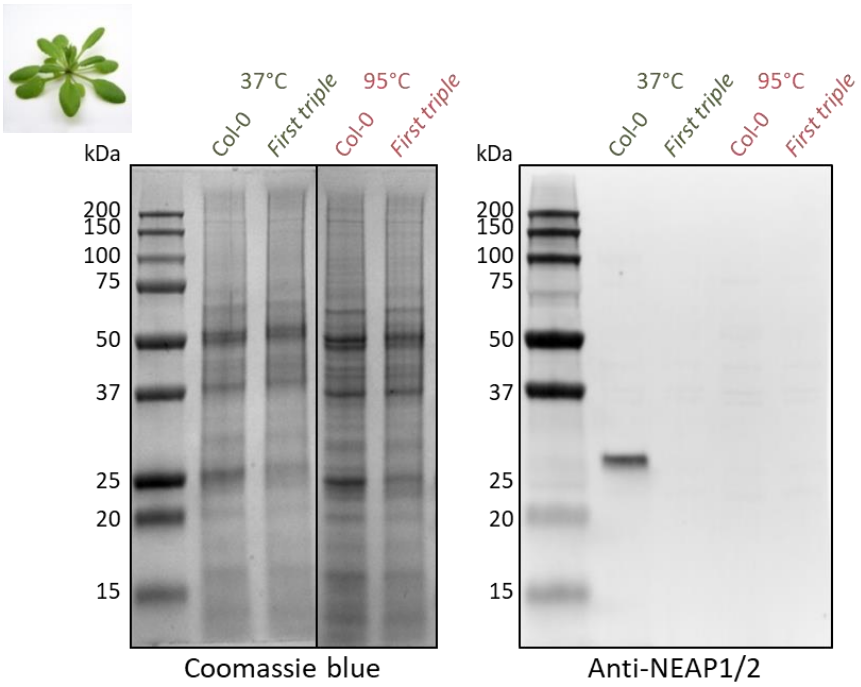


Figure 4-10: Test of anti-NEAP1/2 antibody on native AtNEAP proteins in *A. thaliana*. Proteins were extracted from 3 week-old WT Col-0, or *first triple neap* mutant plants. Same loaded gel was stained with Coomassie blue to attest the loading (left panel). Antibody dilution was 1 in 250 for anti-NEAP1/2 and more details are available in ***Methods Table 2.5***.

Then, tests were carried out on transiently over-expressed AtNEAP2-CFP fusion proteins in *N. benthamiana*, using an extraction protocol specific for membrane proteins optimized by (Pawar *et al*., 2016), ***Figure 4-9B***. Negative control was a non-infiltrated *N. benthamiana* plant. On the left panel of ***Figure 4-9B***, anti-NEAP1/2 antibody revealed a band around 56kDa for the plants infiltrated with AtNEAP2-CFP. A band with the same MW was also revealed with anti-GFP antibody for the same plant. No band around 56kDa was observed for the negative control plant for both antibodies. The predicted MW for AtNEAP2-CFP fusion protein was around 60kDa as CFP is 22kDa and AtNEAP2 is 38kDa. Then, MW observed was slightly lower than expected but this result was consistent with data previously observed by Pawar-Menon, PhD thesis, 2015. Attempts to use the anti-NEAP wobble on plant extracts failed and this antibody seems to be weaker and also needs to be used at a low dilution (a least 1 in 50). Thus, AtNEAP2-CFP was revealed by both anti-NEAP1/2 and anti-GFP antibodies and confirmed the ability of anti-NEAP2 antibody to recognize AtNEAP2, when fusion proteins are over-expressed in *N. benthamiana*.

Finally, antibodies were also tested on *A. thaliana* extracts, but additional protein extraction protocols were required for enriching extracts in nuclear proteins, (***Methods V.1***; ***Figure 4-10***). Some difficulties became apparent for extracting AtNEAP native proteins from *A. thaliana* plants, so adjustments in the protocol were made. One of those was to incubate protein lysate at 37°C for 30min instead of the classical incubation at 95°C for 5min before a loading on a SDS-gel. Avoiding the boiling step prevents those particular proteins with TM and CC domains from aggregation, limiting their migration into the SDS-PAGE. Three-week-old *A. thaliana* plants grown on MS medium, either wild-type (Col-0) or *first triple neap* mutant (*First triple*) were used. Results shown in ***Figure 4-10*** revealed only one clear band around 30kDa for Col-0 incubated at 37°C before loading. So, firstly, it seems that replacing the boiling step should be recommended for enriching the soluble lysate in AtNEAP proteins.

However, as previously observed, the detected band containing possibly AtNEAP1 and AtNEAP2 proteins had a MW of 30kDa, *i.e.* lower than the estimated MW of 41 and 38kDa respectively. After deduction of the estimated MW of the Gal4DBD and CFP tags, AtNEAP1 shows an apparent MW of 39kDa and AtNEAP2 of 34kDa on the previous blots. No band was detected for the *First triple* mutant.

Before drawing conclusions on the ability of the anti-NEAP1/2 antibody to recognize AtNEAP1 and AtNEAP2 native proteins specifically in Arabidopsis, this experiment should be repeated as this positive result has been obtained twice on two independent western blotting experiments but using the same protein extract. Also, no band was observed for the *first triple* mutant, but as the *Atneap2* mutant allele with T-DNA insertion was not a KO, a band for AtNEAP2 was expected. Then, either, something went wrong during the protein extraction, or these results indicate that no AtNEAP2 protein is synthesised in the *Atneap2* mutant allele with T-DNA insertion, or that expression levels are too low for detection, or the protein is more rapidly degraded.

Altogether, the results of antibody tests on different protein extracts demonstrated that the anti-NEAP1/2 antibody was able to recognize AtNEAP1 and AtNEAP2 expressed and extracted from yeast and tobacco. Also, this antibody would be able to recognize native proteins in Arabidopsis nuclei extracts. Thus, it would be the first time that *AtNEAP* expression has been shown in native tissues using an antibody. The anti-NEAP wobble antibody seemed to be weaker as it was tested on plant extracts with no success (data not shown) but it was able to recognize at least each AtNEAP in yeast extracts, with a higher affinity for AtNEAP3, which would be complementary to the other antibody.

Conclusion

After several attempts with Y2H looking for AtNEAP partners, no new relevant partner was identified and previous MYTH studies identifying AtbZIP18 and AtMaMYB

were not confirmed. However, this is likely to be a result of the differences in the properties of the Y2H and MYTH systems. The apFRET experiments carried out confirmed the physical interaction between AtbZIP18 and AtNEAPs, suggested by MYTH for AtbZIP18+AtNEAP1 in Pawar *et al*, 2016.

The results obtained also showed that AtbZIP18 is localised at the nuclear periphery in transient expression in *N. benthamiana* when co-infiltrated with AtNEAPs, while those proteins were suggested to be localised in the nucleoplasm in Pawar *et al*., 2016. In addition, Gibalová *et al*., 2017, showed AtbZIP18 alone to be in the nucleoplasm and in the perinuclear region. Thus, this observation could indicate that AtNEAPs and AtbZIP18 influence localisation of each other. A recent analysis of the putative domains contained in AtbZIP18 protein sequence revealed the presence of CC domains close to the BRLZ domain, which could be responsible for interaction with AtNEAPs. Further investigation is now required on specific function of AtbZIP18 with AtNEAPs at the nuclear periphery. Concerning AtMaMYB, only preliminary evidence from MYTH suggests the existence of a small network of interaction between AtSUN3, AtNEAP1, AtMaMYB and AtbZIP18 proteins, ***Appendix VII***. More experiments performed by Bisa Andov, PhD student at Oxford Brookes University, focused on AtMaMYB, are ongoing and will assess *in vivo* the relevance of this network.

In addition, specific anti-AtNEAP antibodies were obtained and proof of specificity has been demonstrated using over-expressed fusion proteins from yeast and tobacco (and possibly on native proteins extract from *A. thaliana*). To explore other molecular approaches to investigate the AtNEAPs interactome, preliminary experiments have been recently undertaken to explore the use of these new antibodies for IP and co-IP from *A. thaliana* extract. Some difficulties appeared at first steps in extracting native AtNEAP proteins and for enriching lysates in non-soluble NE proteins. Recently, a protocol was set-up by Frances Tolmie (Oxford Brookes University), using a method based on one from the group of Hank

Bass, Florida State University, who recently succeeded to immunoprecipitate SUN2 and who co-immunoprecipitated many nuclear envelope and nuclear periphery proteins (Gumber *et al*., 2019). Protein extraction protocols with enrichment in nuclear proteins have to be applied to AtNEAP proteins to produce plant materials for IP and mass spectrometry. Other approaches include (i) the construction of 6xHis-ATNEAPs and GST-AtNEAPs, respectively for nickel and GST pull-down experiments and (ii) the establishment of new transgenic lines expressing pAtNEAP1::4xc-Myc-AtNEAP1 (see ***Chapter 3 Conclusion***) in a *triple neap* mutant background for complementation and for IP using anti-Myc antibody followed by mass spectrometry sequencing, which will open new possibilities in the future.

DISCUSSION

DISCUSSION

Through different approaches adopted in this thesis, the AtNEAP protein function and interaction network at the nuclear periphery has been further characterised. Reverse genetics using CRISPR/Cas9 technology gave the opportunity to obtain a new triple KO *Atneap1 Atneap2 Atneap3* (*Atneap1/2/3*) mutant. In these plants lacking functional versions of the three AtNEAP paralogues, several phenotypic characteristics have been observed: at the whole plant level, phenotypic alterations were observed in reproductive tissues suggesting a functional role in meiosis or embryo formation. At the cellular level, changes in nuclear organisation compared to WT Col-0 plants were recorded and suggest a role for AtNEAP proteins in the localisation and possible anchoring of chromocentres at the nuclear periphery. In parallel, molecular approaches including Y2H and *in vivo* localisation and co-localisation associated with apFRET confirmed the ability of AtNEAP proteins to form homo- and heterodimers, and to interact with domains of the TF AtbZIP18. A biochemical approach including the generation of specific AtNEAP antibodies confirmed expression *in-vivo* and revealed the strong hydrophobic properties of AtNEAP resulting in difficulties in protein extraction procedures. Altogether, the results support the role of AtNEAP proteins in anchoring AtbZIP18 TF at the INM to maintain nuclear morphology and chromatin organisation. In this general discussion, firstly, short-term approaches will be suggested to demonstrate the functional role of AtNEAP proteins at the nuclear periphery. Secondly, a potential mechanism of action of AtNEAP proteins in tethering chromocentres at the nuclear periphery and a role through interaction with AtbZIP18 in gene repression will be discussed. Finally, future work (long-term approaches) and hypothesis will be suggested.

I – Role of AtNEAPs in tethering chromocentres to the nuclear periphery

In order to study AtNEAP protein function in *A. thaliana*, reverse genetics has been used. It was important to generate a triple KO mutant including a loss-of-function allele of *AtNEAP2*. This allele was not available at the beginning of this study, and two years were required to create a new *Atneap2* mutant via the CRISPR/Cas9 technique in wild type plants and to introgress the new mutant allele into the double *Atneap1Atneap3* mutant already available. No mutant plants could be recovered for the first CRISPR/Cas9 target site, ideally located at the beginning of *AtNEAP2* in the first exon. This could be explained by the fact that Cas9 mutations randomly happen so does the reparation mechanism in Arabidopsis which is preferentially NHEJ, (Schiml *et al.*, 2017). Nevertheless, a mutation in the second CRISPR/Cas9 target site was obtained located in the third exon before the NLS and TM domain and it was decided to continue the work using this mutant allele. This single nucleotide insertion introduces an early stop codon and would lead to a truncated protein with neither NLS nor TM domain. A transient expression experiment of the truncated AtNEAP2 protein confirmed that this mutant protein is not targeted to the nucleus and weakly expressed suggesting some instability (rapid turnover). Then, preliminary phenotyping screens were performed on the triple *Atneap1/2/3* mutant.

A first analysis revealed an impact on the siliques, which are significantly reduced in size, and also contained a reduced number of viable seeds correlated with a higher number of non-fertilized ovules, as the total number of seeds is not affected. This result seems consistent with the higher level of transcription of *AtNEAPs* in seeds, especially in embryo, (Pawar *et al.*, 2016). It also raised the question of a potential effect on meiosis or embryo formation in the *Atneap* mutants which is currently being investigated in collaboration with Monica Pradillo's group in Spain.

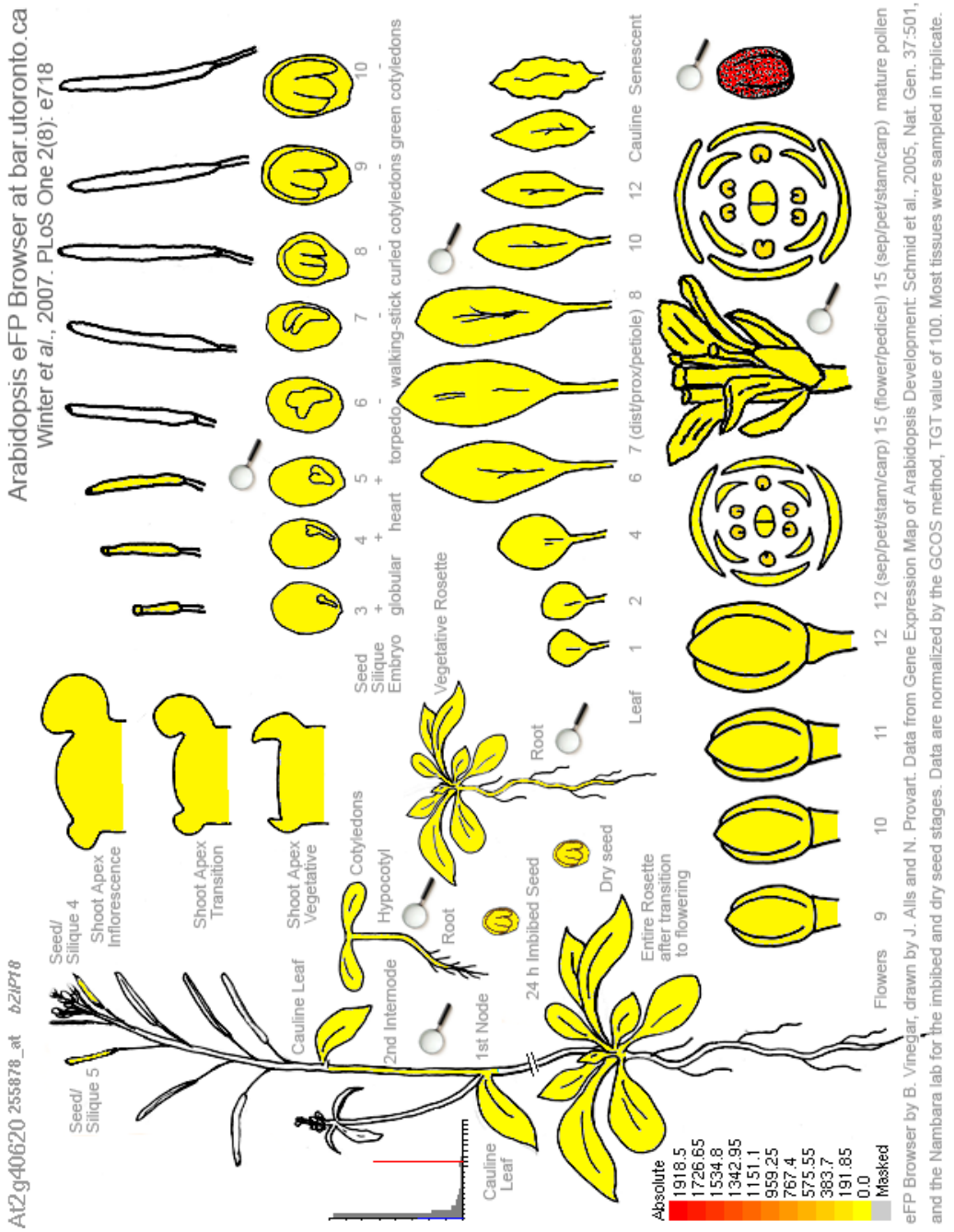


Figure 5-1: Transcription levels of AtbZIP18 mRNA in different tissues from *GeneVestigator*. AtbZIP18 is highly expressed in the mature pollen grain.

A finer analysis of nuclear morphology and chromatin organisation of mutant nuclei showed that chromocentres are more internal as the distance between chromocentres and the nuclear periphery is increased. Even if this experiment needs to be done other times with an increased number of nuclei in both guard and pavement cell populations, these preliminary results could indicate a defect in a putative physical tethering of the chromocentres at the nuclear periphery when AtNEAP proteins are absent. A similar result was obtained with the triple *Atsun1/4/5* mutant, which also shows a decompaction of chromocentres and a release of gene silencing at some repeated sequences (Poulet *et al*., 2016). However, no chromocentre decompaction was observed in the triple *Atneap1/2/3* mutant and preliminary RT-qPCR results using the same repeated sequences as for the triple *Atsun1/4/5* (180bp, TSI) failed to detect any defect in gene silencing. If AtNEAP proteins participate in the tethering of specific chromatin regions at the nuclear periphery, it does not seem to affect repeated sequences or chromocentre compaction. So far, how an alteration in chromocentre position affects genome expression or is linked to the phenotypical differences observed in *Atneap* mutant plants remains to be shown.

II – AtNEAP proteins interact with the transcription factor AtbZIP18

AtbZIP18 was identified in a MYTH screen using AtNEAP1 as bait (Pawar *et al*., 2016). AtbZIP18 is a transcription factor (TF) expressed everywhere in the plant but with a higher level of transcription in the mature pollen grain, (Gibalová *et al*., 2017), ***Figure 5-1***. Among *AtNEAP* genes, *AtNEAP1* and *AtNEAP2* are more highly transcribed compared to *AtNEAP3* in every tissue with a relative stronger transcription in seeds (***Appendix VII***, Pawar *et al*., 2016, Supplementary Fig4). In this work, interaction of AtNEAP proteins with AtbZIP18 has been confirmed *in vivo* by apFRET, ***Figure 4-6***. Note that during this experiment, AtbZIP18 localisation, ***Figure 4-5***, was significantly different from previous data

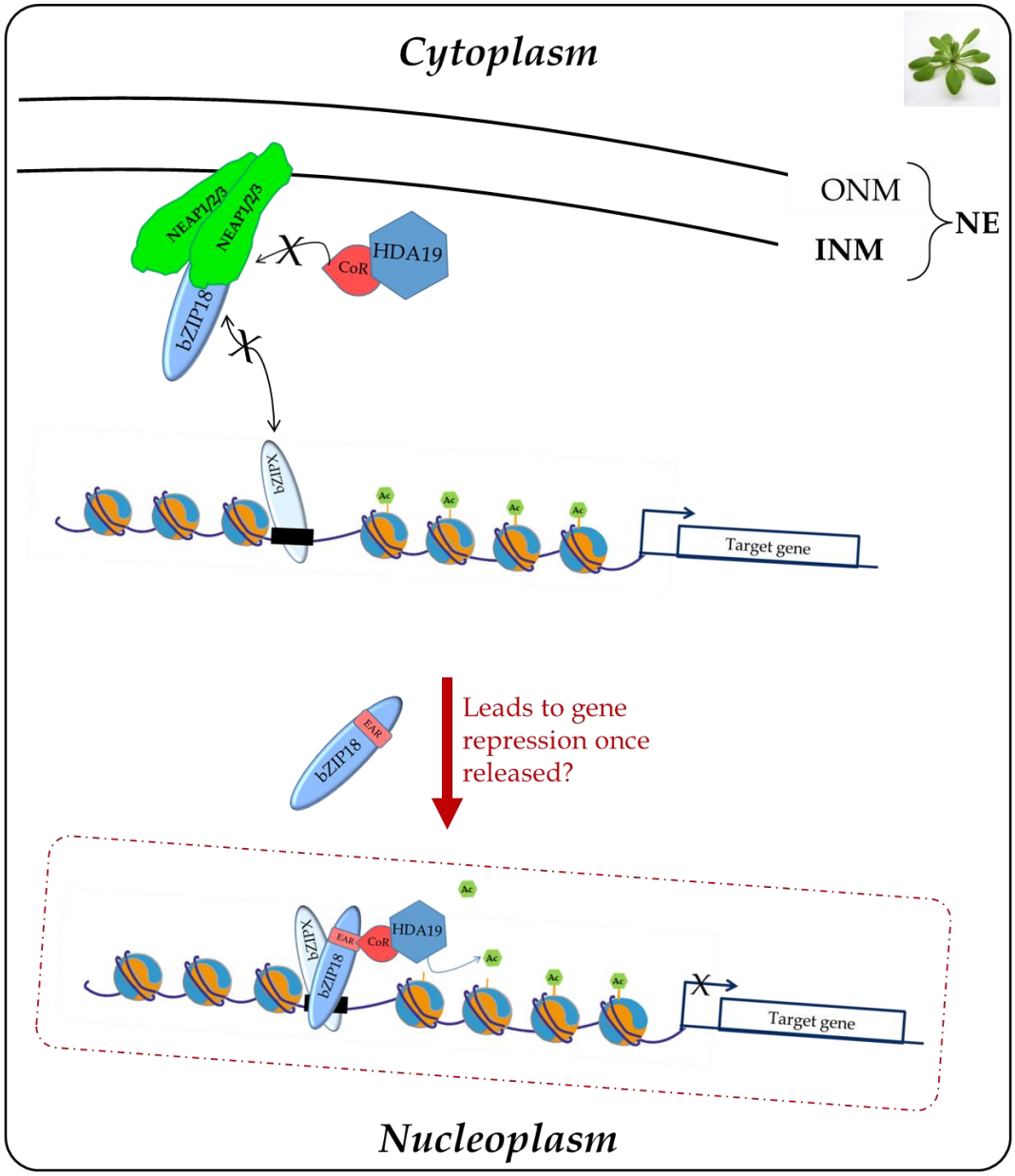


Figure 5-2: Proposed model of AtNEAP and AtbZIP18 function in gene repression in *A. thaliana*. AtbZIP18 would be sequestered by AtNEAP at the nuclear periphery, masking the EAR motif and blocking interactions with CoR and dimerisation with another bZIP. A loss of tethering by AtNEAP would lead to gene repression mediated by bZIP18 and other partners. bZIPX: AtbZIP34, 52 or 61. CoR: Co-repressors such as AtSIN3, AtSAP18 or TPL. Black rectangle represents the DNA motif for interaction with bZIP TFs.HDA19 is a histone deacetylase and leads to gene repression but the mechanism of the spreading remains to be elucidated.

(Pawar *et al*., 2016) where it was nucleoplasmic and not only restricted to the nuclear periphery. This could be due to different issues regarding the experiment in itself. Indeed, depending on the timing for observation, between two – to five days post-infiltration, as it is transient, the level of expression can drastically change and induce mis-localisation if proteins are too much over-expressed. In this case, it tends to leak into the nucleoplasm instead of being restricted to the nuclear periphery. This is probably the case in Pawar-Menon, PhD thesis, 2015, where AtNEAP1 on its own was sometimes peripheral, sometimes nucleoplasmic. Finally, it has been shown in Gibalova *et al.*, 2017, that AtbZIP18 on its own was enriched at the nuclear periphery. Therefore, it does make sense that the co-localisation and interaction of AtbZIP18/AtNEAP happen at the nuclear periphery as shown in this study, ***Figure 4-5 and Figure 4-6***, but if fusion proteins are too much over-expressed, this interaction may be seen into the nucleoplasm.

A recent analysis of the predicted domains of bZIP transcription factors including AtbZIP18 revealed the presence of CC domains overlapping the BRLZ domain implicated in its interaction with DNA and dimerization with other bZIP TFs (Dröge-Laser *et al*., 2018). CC domains are known to be important for Protein-Protein Interaction (PPI) and even if the CC domain is common to all bZIP TFs, (Dröge-Laser *et al*., 2018), in the case of AtbZIP18 it seems to be responsible for interaction with AtNEAP proteins according to Y2H experiments with AtbZIP18 domain deletions. This could raise the possibility that AtbZIP18 can either interact with AtNEAPs, and therefore be tethered at the NE, or at specific DNA target sites to regulate transcription, but would not be able to do both at the same time, ***Figure 5-2***. Indeed, by linking AtbZIP18 through its CC domain, AtNEAP could mask the Ethylene-responsive element binding factor-associated Amphiphilic Repression (EAR) motif (LxLxL) motif and prevent fixation of co-repressors on AtbZIP18, ***Figure 5-2***. Also, it is possible that AtNEAP/AtbZIP18 interaction inhibits AtbZIP18 dimerization and thus DNA binding. Then,

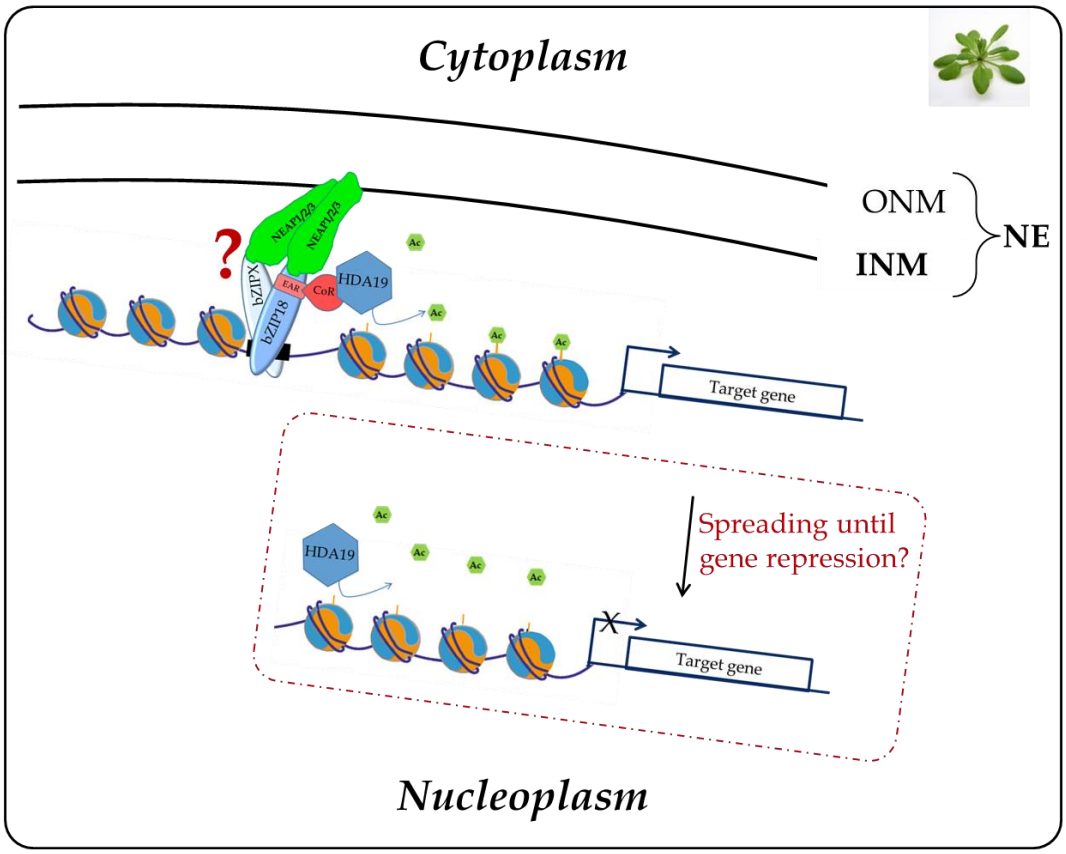


Figure 5-3: Alternative proposed model of AtNEAP and AtbZIP18 function in gene repression in *A. thaliana*. The TFs bZIP function as dimers and would be tethered by AtNEAP dimers, bZIPX: AtbZIP34, 52 or 61. CoR: Co-repressors such as AtSIN3, AtSAP18 or TPL. Black rectangle represents the DNA motif for interaction with bZIP TFs. HDA19 is a histone deacetylase and leads to gene repression but the mechanism of the spreading remains to be elucidated.

due to a stimuli or a stress from the environment, AtbZIP18 could be released from the NE tethering and would be able to trigger a gene repression pathway, ***Figure 5-2***. Therefore, it would be a matter of interest to investigate whether AtbZIP18 once bound to AtNEAP is still able to dimerize with another AtbZIP and to bind DNA or co-factors.

Alternatively, if the interaction of AtNEAP proteins with AtbZIP18 simultaneously bound to DNA is possible, it would suggest that AtNEAPs could be partially responsible for the tethering of chromatin domains at the nuclear periphery by linking TFs, ***Figure 5-3***. AtbZIP18 is a putative transcription repressor as its loss-of-function leads to the up-regulation of 117 genes out of 133 differentially expressed genes (Fold Change ≥2), (Gibalová *et al*., 2017). One possible explanation of this repressive activity is the presence of an EAR motif located at the 3' of the BRLZ domain within AtbZIP18. The EAR motif is a common motif found in TFs implicated in gene repression as this motif is known to be involved in transcriptional inhibition through chromatin modifications, (Gibalová *et al*., 2017; Kagale and Rozwadowski, 2010), ***Figure 5-3***. Indeed, the EAR motif is important for interaction of the TF with chromatin remodelling factors. These co-repressors are able to recruit AtHDA19, which is a histone deacetylase (HDAC), leading to gene repression, (Kagale and Rozwadowski, 2010, 2011).

AtbZIP18 was the only TF revealed by MYTH but recent experiments from the lab suggest that another transcription factor called AtMaMYB is also able to interact both with AtNEAP1 and AtbZIP18 in MYTH (Voisin and Vanrobays unpublished). Although these new interactions have to be confirmed *in vivo*, it suggests that other TFs interact with AtNEAPs.

III – Future work and perspectives

During the initial steps of this work, a Y2H screen was performed but failed to detect any new interactors. Also, neither AtbZIP18 nor AtMaMYB were identified in these screens. The failure to detect new partners could be explained by the fact that this system, compared to

the MYTH system, is not adapted for the specific requirements of TM proteins such as AtNEAPs, probably leading to a mis-localisation of AtNEAPs in yeast. Indeed, in Pawar-Menon, PhD thesis, 2015, a MYTH experiment revealed a weak interaction between AtSUN1/AtNEAP1 and AtSUN2/AtNEAP1 which was, nevertheless, confirmed by apFRET, (Pawar *et al*., 2016).

Thus, it would be interesting to perform these Y2H screens, this time with AtNEAPs having the TM domain deleted. The fact that results obtained so far to identify AtNEAP partners have resulted only in the identification of two TFs, AtbZIP18 and AtMaMYB, suggests that interaction between AtNEAP proteins and chromatin is indirect through TFs. If AtNEAP proteins interact with other components of the nuclear periphery, this will have to be identified by other strategies than Y2H.

One such strategy could be to apply immunoprecipitation (IP) of AtNEAP proteins followed by Mass-Spectrometry (MS) sequencing. The new generated antibodies successfully detected AtNEAP proteins on WB analysis, at least when over-expressed fusion proteins, ***Figure 4-9***, and it would be interesting to validate the absence of AtNEAP proteins in the triple *Atneap1/2/3*. In the future, IP protocols will have to be established, which have to be appropriate for the very hydrophobic AtNEAP proteins. This could first be tested on AtNEAP proteins expressed in yeast and then by using a line expressing c-Myc- or Flag-HA-tagged AtNEAP1 protein in a triple *Atneap1/2/3* background. Preliminary attempts to perform IP were tested but failed as AtNEAP proteins proved very challenging to extract, possibly remaining associated with the nuclear membrane and remaining in the insoluble fraction. Indeed, in the study presented in this thesis, a wide investigation of new AtNEAP interactants was initiated but was not successful due to failure at early steps of the protocol when preparing samples for IP due to the insolubility of the AtNEAPs.

It is striking to note that some early experiments to purify the putative plant lamina/nucleoskeleton by (Sakamoto and Takagi, 2013) also failed to detect AtNEAPs, AtbZIP18 or AtMaMYB. This could highlight the difficulties linked to the relative insolubility of AtNEAP proteins in standard buffers and explain their absence in the list of 660 proteins of the crude lamina (Sakamoto and Takagi, 2013). More recently, (Goto *et al*., 2019), performed a wide nuclear proteome investigation of *A. thaliana* with MS analysis. Of 1541 proteins identified, some nucleoskeleton proteins were found, such as AtCRWN1, AtCRWN4, AtKAKU4, AtSUN1, AtSUN2, but not AtNEAPs, AtbZIP18 or AtMaMYB. It seems that further protocol optimization is required in order to be able to do MS and define what the AtNEAP interactome *in vivo* exactly is.

The fact that the triple *Atneap1/2/3* mutant did not show an apparent phenotype either during the vegetative phase or in the shoot or root, suggests firstly that AtNEAP proteins may have a redundant function with other actors at the nuclear periphery. Alternatively, AtNEAP function could be related to stress response and therefore, mutants could show a stress-induced phenotype. Thus, this mutant, as well as all the single and double mutants could be challenged under different stress conditions like heat, cold, drought, salt or light duration and intensity. Indeed, plants have to adapt to changes of light and temperature between night and day and depending on weather. It has been shown that during photomorphogenesis in plants, heterochromatin organization is profoundly reorganized (Bourbousse *et al*., 2015) and a heat shock of 30h at 37°C induced heterochromatin decondensation, (Pecinka *et al*., 2010; Tittel-Elmer *et al*., 2010). Therefore, the plasticity of the genome is really important for plants during environment-dependent switches, which are particularly challenging. A defect in chromatin organization and positioning of the chromocentres could be deleterious at these critical steps in mutant plants. That is why a study of *Atneap* mutants under stress conditions could therefore reveal effects on growth conditions and nuclear morphology integrity. Also,

as proposed in models (***Figure 5-2 and 5-3***), AtNEAP proteins could be part of a signalling pathway; linking AtbZIP18 whose function is to repress specific target genes through the help of AtHDA19. Roles of this HDAC have been investigated and AtHDA19 was reported to control root cell elongation, modulate seed germination and to be implicated in salt- as well as abscisic acid stress-response, (Chen and Wu, 2010; Chen *et al*., 2015). Thus, the specific roles of HDA19 could be helpful in order to determine which kind of stresses could be applied to the triple *Atneap* mutant or the single *Atbzip18*.

So far, the triple *Atneap* mutant characterization has been mainly phenotypic. A deeper molecular characterization is required in order to get insight into the molecular roles of the AtNEAP protein family. Image analysis of the nuclear periphery in the triple *Atneap* mutant revealed mislocalisation of chromocentres as the distance from the nuclear periphery was increased. This was also visible in the triple *Atsun1/4/5* mutant, (Poulet *et al*., 2016) and in mammals presenting laminopathies, nuclear shape is also altered with lobulation of the NE, thickening of the nuclear lamina and loss of peripheral heterochromatin, (Mattout *et al*., 2006). An altered pattern of heterochromatin positioning leads to transcriptional alterations and for that purpose, RNAseq analysis might be relevant to reveal silencing release and sporadic transcription at the centromeric and pericentromeric regions or change in gene expression.

A rearrangement of heterochromatin would ultimately modify deeply the epigenome organization in the *Atneap* triple mutant. Therefore it could be interesting to perform whole genome bisulfite sequencing analysis as well as ChIP-seq on permissive and repressive histone marks such as, respectively, H3K4me3, H3K9Ac, and H3K9me2, H3K27me1, H3K27me3, to see which genes are affected by changes in this mutant. Also, to draw chromatin-chromatin interaction map at the whole genome level, Hi-C analyses could be carried out to look at potential modifications.

All those genome-wide analyses could also be carried out on stress-induced plants to exacerbate mutant effects on genome plasticity. In parallel, crossing the *Atneap1/2/3* mutant with the At*bZIP18* mutant would be worthwhile to investigate their interaction pattern and function in the protein network at the nuclear periphery. Indeed, *Atbzip18* single mutants, in addition to showing an increased number of aborted pollen grains and defects in living ones, shows a global gene up-regulation, (see ***paragraph II*** above). RNA-Seq experiments comparing the triple *Atneap1/2/3* with a quadruple *Atneap1/2/3, Atbzip18* could reveal which part of the transcriptome that might be altered in *Atneap1/2/3* mutants is mediated by AtbZIP18. In the same way, ChIP-Seq experiments could be designed to investigate these target sites if an AtbZIP18-GFP construct expressed under its own promoter is introduced into the *Atneap1/2/3* mutant line. A wide study using DAP-seq technology, (O'Malley *et al*., 2016), determined specific DNA-binding motives of multiple TFs including AtbZIP18 and some of its partners of TF group I, AtbZIP52 and 51. These motifs are -TGACAGCTGT- with a higher confidence for the core -CAGCT- and this information could be helpful for discovering AtbZIP18 target genes linked to the nuclear periphery. Indeed, from RNA-Seq results, the putative AtbZIP18 target genes can be identified and upstream regions can be screened for this common motif. Then, a co-expression analysis of these genes can be performed.

Altogether, these multiple approaches would be very promising to better define the impact of the nuclear periphery on gene expression and especially to further elucidate the role of AtNEAP proteins at the nuclear periphery, in anchoring chromatin and in the nucleoskeleton.

REFERENCES

REFERENCES

Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., *et al*. (2007). The molecular architecture of the nuclear pore complex. Nature *450*, 695–701.

Bernard, P., and Couturier, M. (1992). Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. J. Mol. Biol. *226*, 735–745.

Bharath, M.M.S., Chandra, N.R., and Rao, M.R.S. (2002). Prediction of an HMG-box fold in the C-terminal domain of histone H1: insights into its role in DNA condensation. Proteins *49*, 71–81.

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*, 1162–1173.

Bourbousse, C., Mestiri, I., Zabulon, G., Bourge, M., Formiggini, F., Koini, M.A., Brown, S.C., Fransz, P., Bowler, C., and Barneche, F. (2015). Light signaling controls nuclear architecture reorganization during seedling establishment. Proc. Natl. Acad. Sci. U.S.A. *112*, E2836-2844.

Bradbury, E.M. (1992). Reversible histone modifications and the chromosome cell cycle. Bioessays *14*, 9–16.

Breyne, P., Van Montagu, M., and Gheysen, G. (1994). The role of scaffold attachment regions in the structural and functional organization of plant chromatin. Transgenic Res. *3*, 195–202.

Brohawn, S.G., Partridge, J.R., Whittle, J.R.R., and Schwartz, T.U. (2009). The nuclear pore complex has entered the atomic age. Structure *17*, 1156–1168.

Burke, B. (2012). It takes KASH to hitch to the SUN. Cell *149*, 961–963.

Burke, B., and Stewart, C.L. (2013). The nuclear lamins: flexibility in function. Nat. Rev. Mol. Cell Biol. *14*, 13–24.

Burns, L.T., and Wente, S.R. (2014). From hypothesis to mechanism: uncovering nuclear pore complex links to gene expression. Mol. Cell. Biol. *34*, 2114–2120.

Capelson, M., and Hetzer, M.W. (2009). The role of nuclear pores in gene regulation, development and disease. EMBO Rep. *10*, 697–705.

Chen, L.-T., and Wu, K. (2010). Role of histone deacetylases HDA6 and HDA19 in ABA and abiotic stress response. Plant Signal Behav. *5*, 1318–1320.

Chen, C.-Y., Wu, K., and Schmidt, W. (2015). The histone deacetylase HDA19 controls root cell elongation and modulates a subset of phosphate starvation responses in Arabidopsis. Sci. Rep. *5*, 15708.

Chodavarapu, R.K., Feng, S., Bernatavichute, Y.V., Chen, P.-Y., Stroud, H., Yu, Y., Hetzel, J.A., Kuo, F., Kim, J., Cokus, S.J., *et al*. (2010). Relationship between nucleosome positioning and DNA methylation. Nature *466*, 388–392.

Ciska, M., and Moreno Díaz de la Espina, S. (2014). The intriguing plant nuclear lamina. Front Plant Sci. *5*, 166.

Ciska, M., Masuda, K., and Moreno Díaz de la Espina, S. (2013). Lamin-like analogues in plants: the characterization of NMCP1 in Allium cepa. J. Exp. Bot. *64*, 1553–1564.

Clapier, C.R., and Cairns, B.R. (2009). The Biology of Chromatin Remodeling Complexes. Annual Review of Biochemistry *78*, 273–304.

Crisp, M., Liu, Q., Roux, K., Rattner, J.B., Shanahan, C., Burke, B., Stahl, P.D., and Hodzic, D. (2006). Coupling of the nucleus and cytoplasm: role of the LINC complex. J. Cell Biol. *172*, 41–53.

DeGrasse, J.A., DuBois, K.N., Devos, D., Siegel, T.N., Sali, A., Field, M.C., Rout, M.P., and Chait, B.T. (2009). Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. Mol. Cell Proteomics *8*, 2119–2130.

Desrosiers, R., and Tanguay, R.M. (1986). Further characterization of the posttranslational modifications of core histones in response to heat and arsenite stress in Drosophila. Biochem. Cell Biol. *64*, 750–757.

Desset, S., Poulet, A., and Tatout, C. (2018). Quantitative 3D Analysis of Nuclear Morphology and Heterochromatin Organization from Whole-Mount Plant Tissue Using NucleusJ. Methods Mol. Biol. *1675*, 615–632.

Dieppois, G., and Stutz, F. (2010). Connecting the transcription site to the nuclear pore: a multi-tether process that regulates gene expression. J. Cell. Sci. *123*, 1989–1999.

Dilsaver, M.R., Chen, P., Thompson, T.A., Reusser, T., Mukherjee, R.N., Oakey, J., and Levy, D.L. (2018). Emerin induces nuclear breakage in Xenopus extract and early embryos. Mol. Biol. Cell mbcE18050277.

Dittmer, T.A., Stacey, N.J., Sugimoto-Shirasu, K., and Richards, E.J. (2007). LITTLE NUCLEI genes affecting nuclear morphology in Arabidopsis thaliana. Plant Cell *19*, 2793–2803.

Dröge-Laser, W., Snoek, B.L., Snel, B., and Weiste, C. (2018). The Arabidopsis bZIP transcription factor family-an update. Curr. Opin. Plant Biol. *45*, 36–49.

Elgin, S.C.R., and Grewal, S.I.S. (2003). Heterochromatin: silence is golden. Curr. Biol. *13*, R895-898.

Evans, D.E., Pawar, V., Smith, S.J., and Graumann, K. (2014). Protein interactions at the higher plant nuclear envelope: evidence for a linker of nucleoskeleton and cytoskeleton complex. Front Plant Sci. *5*, 183.

Fauser, F., Schiml, S., and Puchta, H. (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. Plant J. *79*, 348–359.

Feng, S., Cokus, S.J., Schubert, V., Zhai, J., Pellegrini, M., and Jacobsen, S.E. (2014). Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in Arabidopsis. Mol. Cell *55*, 694–707.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature *340*, 245–246.

Fiserova, J., and Goldberg, M.W. (2010). Relationships at the nuclear envelope: lamins and nuclear pore complexes in animals and plants. Biochem. Soc. Trans. *38*, 829–831.

Fiserova, J., Kiseleva, E., and Goldberg, M.W. (2009). Nuclear envelope and nuclear pore complex structure and organization in tobacco BY-2 cells. Plant J. *59*, 243–255.

Fransz, P., De Jong, J.H., Lysak, M., Castiglione, M.R., and Schubert, I. (2002). Interphase chromosomes in Arabidopsis are organized as well defined chromocenters from which euchromatin loops emanate. Proc. Natl. Acad. Sci. U.S.A. *99*, 14584–14589.

Fransz, P., Soppe, W., and Schubert, I. (2003). Heterochromatin in interphase nuclei of Arabidopsis thaliana. Chromosome Res. *11*, 227–240.

Fransz, P., ten Hoopen, R., and Tessadori, F. (2006). Composition and formation of heterochromatin in Arabidopsis thaliana. Chromosome Res. *14*, 71–82.

Gibalová, A., Steinbachová, L., Hafidh, S., Bláhová, V., Gadiou, Z., Michailidis, C., Műller, K., Pleskot, R., Dupľáková, N., and Honys, D. (2017). Characterization of pollen-expressed bZIP protein interactions and the role of ATbZIP18 in the male gametophyte. Plant Reprod. *30*, 1–17.

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., Hashizume, S., Fukao, Y., Hara-Nishimura, I., and Tamura, K. (2019). Comprehensive nuclear proteome of Arabidopsis obtained by sequential extraction. Nucleus.

Graumann, K. (2014). Evidence for LINC1-SUN associations at the plant nuclear periphery. PLoS ONE *9*, e93406.

Graumann, K., Runions, J., and Evans, D.E. (2010). Characterization of SUN-domain proteins at the higher plant nuclear envelope. Plant J. *61*, 134–144.

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. Mol. Cell *55*, 678–693.

Grossman, E., Medalia, O., and Zwerger, M. (2012). Functional architecture of the nuclear pore complex. Annu. Rev. Biophys. *41*, 557–584.

Gu, Y., Zebell, S.G., Liang, Z., Wang, S., Kang, B.-H., and Dong, X. (2016). Nuclear Pore Permeabilization Is a Convergent Signaling Event in Effector-Triggered Immunity. Cell *166*, 1526-1538.e11.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., *et al*. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948–951.

Gumber, H.K., McKenna, J.F., Estrada, A.L., Tolmie, A.F., Graumann, K., and Bass, H.W. (2019). Identification and characterization of genes encoding the nuclear envelope LINC complex in the monocot species Zea mays. J. Cell. Sci. *132*.

Hagan, I., and Yanagida, M. (1995). The product of the spindle formation gene sad1+ associates with the fission yeast spindle pole body and is essential for viability. J. Cell Biol. *129*, 1033–1047.

Heitz, E. (1928). Das Heterochromatin der Moose. I Jahrb. Wiss. Bot. 69: 762–818.

Henikoff, S. (2008). Nucleosome destabilization in the epigenetic regulation of gene expression. Nat. Rev. Genet. *9*, 15–26.

Ho, R., and Hegele, R.A. (2018). Complex effects of laminopathy mutations on nuclear structure and function. Clin. Genet.

Huber, M.D., Guan, T., and Gerace, L. (2009). Overlapping functions of nuclear envelope proteins NET25 (Lem2) and emerin in regulation of extracellular signal-regulated kinase signaling in myoblast differentiation. Mol. Cell. Biol. *29*, 5718–5728.

Hutten, S., Flotho, A., Melchior, F., and Kehlenbach, R.H. (2008). The Nup358-RanGAP complex is required for efficient importin alpha/beta-dependent nuclear import. Mol. Biol. Cell *19*, 2300–2310.

Imhof, A., and Becker, P.B. (2001). Modifications of the histone N-terminal domains. Evidence for an “epigenetic code”? Mol. Biotechnol. *17*, 1–13.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074–1080.

Kagale, S., and Rozwadowski, K. (2010). Small yet effective: the ethylene responsive element binding factor-associated amphiphilic repression (EAR) motif. Plant Signal. Behav. *5*, 691–694.

Kagale, S., and Rozwadowski, K. (2011). EAR motif-mediated transcriptional repression in plants. Epigenetics *6*, 141–146.

Karpova, T., and McNally, J.G. (2006). Detecting protein-protein interactions with CFP-YFP FRET by acceptor photobleaching. Curr Protoc Cytom *Chapter 12*, Unit12.7.

Karpova, T.S., Baumann, C.T., He, L., Wu, X., Grammer, A., Lipsky, P., Hager, G.L., and McNally, J.G. (2003). Fluorescence resonance energy transfer from cyan to yellow fluorescent protein detected by acceptor photobleaching using confocal microscopy and a single laser. J Microsc *209*, 56–70.

Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013). Single-cell dynamics of genome-nuclear lamina interactions. Cell *153*, 178–192.

van Koningsbruggen, S., Gierlinski, M., Schofield, P., Martin, D., Barton, G.J., Ariyurek, Y., den Dunnen, J.T., and Lamond, A.I. (2010). High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. Mol. Biol. Cell *21*, 3735–3748.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680–685.

Law, M.J., Lower, K.M., Voon, H.P.J., Hughes, J.R., Garrick, D., Viprakasit, V., Mitson, M., De Gobbi, M., Marra, M., Morris, A., *et al*. (2010). ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. Cell *143*, 367–378.

Lewis, C.D., Lebkowski, J.S., Daly, A.K., and Laemmli, U.K. (1984). Interphase nuclear matrix and metaphase scaffolding structures. J. Cell Sci. Suppl. *1*, 103–122.

Li, J.-F., and Nebenführ, A. (2010). FAST technique for Agrobacterium-mediated transient gene expression in seedlings of Arabidopsis and other plant species. Cold Spring Harb Protoc *2010*, pdb.prot5428.

Link, J., Leubner, M., Schmitt, J., Göb, E., Benavente, R., Jeang, K.-T., Xu, R., and Alsheimer, M. (2014). Analysis of meiosis in SUN1 deficient mice reveals a distinct role of SUN2 in mammalian meiotic LINC complex formation and function. PLoS Genet. *10*, e1004099.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251–260.

Ma, X.M., and Blenis, J. (2009). Molecular mechanisms of mTOR-mediated translational control. Nat. Rev. Mol. Cell Biol. *10*, 307–318.

Macara, I.G. (2001). Transport into and out of the nucleus. Microbiol. Mol. Biol. Rev. *65*, 570–594, table of contents.

Maison, C., Bailly, D., Roche, D., Montes de Oca, R., Probst, A.V., Vassias, I., Dingli, F., Lombard, B., Loew, D., Quivy, J.-P., *et al*. (2011). SUMOylation promotes de novo targeting of HP1α to pericentric heterochromatin. Nat. Genet. *43*, 220–227.

Malone, C.J., Fixsen, W.D., Horvitz, H.R., and Han, M. (1999). UNC-84 localizes to the nuclear envelope and is required for nuclear migration and anchoring during C. elegans development. Development *126*, 3171–3181.

Marsischky, G., and LaBaer, J. (2004). Many paths to many clones: a comparative look at high-throughput cloning methods. Genome Res. *14*, 2020–2028.

Masuda, K., Xu, Z.J., Takahashi, S., Ito, A., Ono, M., Nomura, K., and Inoue, M. (1997). Peripheral framework of carrot cell nucleus contains a novel protein predicted to exhibit a long alpha-helical domain. Exp. Cell Res. *232*, 173–181.

Mattout, A., Dechat, T., Adam, S.A., Goldman, R.D., and Gruenbaum, Y. (2006). Nuclear lamins, diseases and aging. Current Opinion in Cell Biology *18*, 335–341.

Meier, I. (2001). The plant nuclear envelope. Cell. Mol. Life Sci. *58*, 1774–1780.

Meier, I. (2016). LINCing the eukaryotic tree of life - towards a broad evolutionary comparison of nucleocytoplasmic bridging complexes. J. Cell. Sci. *129*, 3523–3531.

Méjat, A., and Misteli, T. (2010). LINC complexes in health and disease. Nucleus *1*, 40.

Meng, J.-J., Rojas, M., Bacon, W., Stickney, J.T., and Ip, W. (2005). Methods to study protein-protein interactions. Methods Mol. Biol. *289*, 341–358.

Moreno Díaz de la Espina, S., Barthellemy, I., and Cerezuela, M.A. (1991). Isolation and ultrastructural characterization of the residual nuclear matrix in a plant cell system. Chromosoma *100*, 110–117.

Mounkes, L.C., and Stewart, C.L. (2004). Aging and nuclear organization: lamins and progeria. Current Opinion in Cell Biology *16*, 322–327.

Murphy, S.P., Simmons, C.R., and Bass, H.W. (2010). 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. *10*, 269.

Németh, A., Conesa, A., Santoyo-Lopez, J., Medina, I., Montaner, D., Péterfia, B., Solovei, I., Cremer, T., Dopazo, J., and Längst, G. (2010). Initial genomics of the human nucleolus. PLoS Genet. *6*, e1000889.

Oda, Y., and Fukuda, H. (2011). Dynamics of Arabidopsis SUN proteins during mitosis and their involvement in nuclear shaping. Plant J. *66*, 629–641.

Okada, T., Endo, M., Singh, M.B., and Bhalla, P.L. (2005). Analysis of the histone H3 gene family in Arabidopsis and identification of the male-gamete-specific variant AtMGH3. Plant J. *44*, 557–568.

O'Malley, R.C., Huang, S.-S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R. (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. Cell *165*, 1280–1292.

Omarov, R., Sparks, K., Smith, L., Zindovic, J., and Scholthof, H.B. (2006). Biological relevance of a stable biochemical interaction between the tombusvirus-encoded P19 and short interfering RNAs. J. Virol. *80*, 3000–3008.

Pawar, V., Poulet, A., Détourné, G., Tatout, C., Vanrobays, E., Evans, D.E., and Graumann, K. (2016). A novel family of plant nuclear envelope-associated proteins. J. Exp. Bot. *67*, 5699–5710.

Pawar-Menon Vidya, 2015. PhD Thesis. Novel plant nuclear envelope-associated coiled-coil proteins. Oxford Brookes University.

Pecinka, A., Dinh, H.Q., Baubec, T., Rosa, M., Lettner, N., and Mittelsten Scheid, O. (2010). Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in Arabidopsis. Plant Cell *22*, 3118–3129.

Petesch, S.J., and Lis, J.T. (2012). Overcoming the nucleosome barrier during transcript elongation. Trends Genet. *28*, 285–294.

Picart-Picolo, A., Picault, N., and Pontvianne, F. (2019). Ribosomal RNA genes shape chromatin domains associating with the nucleolus. Nucleus 1–6.

Pombo, A., and Dillon, N. (2015). Three-dimensional genome architecture: players and mechanisms. Nat. Rev. Mol. Cell Biol. *16*, 245–257.

Pontvianne, F., Carpentier, M.-C., Durut, N., Pavlištová, V., Jaške, K., Schořová, Š., Parrinello, H., Rohmer, M., Pikaard, C.S., Fojtová, M., *et al*. (2016). Identification of Nucleolus-Associated Chromatin Domains Reveals a Role for the Nucleolus in 3D Organization of the A. thaliana Genome. Cell Rep. *16*, 1574–1587.

Poulet, A., Arganda-Carreras, I., Legland, D., Probst, A.V., Andrey, P., and Tatout, C. (2014). NucleusJ: an ImageJ plugin for quantifying 3D images of interphase nuclei. Bioinformatics.

Poulet, A., Probst, A.V., Graumann, K., Tatout, C., and Evans, D. (2016). Exploring the evolution of the proteins of the plant nuclear envelope. Nucleus *8*, 46–59.

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.

Probst, A.V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. Nat. Rev. Mol. Cell Biol. *10*, 192–206.

Reddy, K.L., Zullo, J.M., Bertolino, E., and Singh, H. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. Nature *452*, 243–247.

Ricci, M.A., Manzo, C., García-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. Cell *160*, 1145–1158.

Rose, A., Patel, S., and Meier, I. (2004). The plant nuclear envelope. Planta *218*, 327–336.

Rothballer, A., and Kutay, U. (2013). The diverse functional LINCs of the nuclear envelope to the cytoskeleton and chromatin. Chromosoma *122*, 415–429.

Roudier, F., Teixeira, F.K., and Colot, V. (2009). Chromatin indexing in Arabidopsis: an epigenomic tale of tails and more. Trends Genet. *25*, 511–517.

Roudier, F., Ahmed, I., Bérard, C., Sarazin, A., Mary-Huard, T., Cortijo, S., Bouyer, D., Caillieux, E., Duvernois-Berthet, E., Al-Shikhley, L., *et al*. (2011). Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. EMBO J. *30*, 1928–1938.

Rutowicz, K., Lirski, M., Mermaz, B., Schubert, J., Teano, G., Mestiri, I., Kroteń, M.A., Fabrice, T.N., Fritz, S., Grob, S., *et al*. (2018). Linker histones regulate fine-scale chromatin organization and modulate developmental decisions in Arabidopsis. BioRxiv 458364.

Sakamoto, Y., and Takagi, S. (2013). LITTLE NUCLEI 1 and 4 regulate nuclear morphology in Arabidopsis thaliana. Plant Cell Physiol. *54*, 622–633.

Samson, C., Petitalot, A., Celli, F., Herrada, I., Ropars, V., Le Du, M.-H., Nhiri, N., Jacquet, E., Arteni, A.-A., Buendia, B., *et al*. (2018). Structural analysis of the ternary complex between lamin A/C, BAF and emerin identifies an interface disrupted in autosomal recessive progeroid diseases. Nucleic Acids Res. *46*, 10460–10473.

Schiml, S., and Puchta, H. (2016). Revolutionizing plant biology: multiple ways of genome engineering by CRISPR/Cas. Plant Methods *12*, 8.

Schiml, S., Fauser, F., and Puchta, H. (2017). CRISPR/Cas-Mediated In Planta Gene Targeting. Methods Mol. Biol. *1610*, 3–11.

Sequeira-Mendes, J., Aragüez, I., Peiró, R., Mendez-Giraldez, R., Zhang, X., Jacobsen, S.E., Bastolla, U., and Gutierrez, C. (2014). The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. Plant Cell *26*, 2351–2366.

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.

Smoyer, C.J., and Jaspersen, S.L. (2014). Breaking down the wall: the nuclear envelope during mitosis. Curr. Opin. Cell Biol. *26*, 1–9.

Starr, D.A. (2009). A nuclear-envelope bridge positions nuclei and moves chromosomes. J. Cell. Sci. *122*, 577–586.

van Steensel, B., and Belmont, A.S. (2017). Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. Cell *169*, 780–791.

van Steensel, B., and Henikoff, S. (2000). Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. Nat. Biotechnol. *18*, 424–428.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45.

Talbert, P.B., Ahmad, K., Almouzni, G., Ausió, J., Berger, F., Bhalla, P.L., Bonner, W.M., Cande, W.Z., Chadwick, B.P., Chan, S.W.L., *et al*. (2012). A unified phylogeny-based nomenclature for histone variants. Epigenetics Chromatin *5*, 7.

Tamura, K., and Hara-Nishimura, I. (2011). Involvement of the nuclear pore complex in morphology of the plant nucleus. Nucleus *2*, 168–172.

Tamura, K., and Hara-Nishimura, I. (2013). The molecular architecture of the plant nuclear pore complex. J. Exp. Bot. *64*, 823–832.

Tamura, K., Fukao, Y., Iwamoto, M., Haraguchi, T., and Hara-Nishimura, I. (2010). Identification and characterization of nuclear pore complex components in Arabidopsis thaliana. Plant Cell *22*, 4084–4097.

Tamura, K., Iwabuchi, K., Fukao, Y., Kondo, M., Okamoto, K., Ueda, H., Nishimura, M., and Hara-Nishimura, I. (2013). Myosin XI-i links the nuclear membrane to the cytoskeleton to control nuclear movement and shape in Arabidopsis. Curr. Biol. *23*, 1776–1781.

Tatout, C., Evans, D.E., Vanrobays, E., Probst, A.V., and Graumann, K. (2014). The plant LINC complex at the nuclear envelope. Chromosome Res. *22*, 241–252.

Tittel-Elmer, M., Bucher, E., Broger, L., Mathieu, O., Paszkowski, J., and Vaillant, I. (2010). Stress-induced activation of heterochromatic transcription. PLoS Genet. *6*, e1001175.

Tran, E.J., King, M.C., and Corbett, A.H. (2014). Macromolecular transport between the nucleus and the cytoplasm: Advances in mechanism and emerging links to disease. Biochim. Biophys. Acta *1843*, 2784–2795.

Turgay, Y., Eibauer, M., Goldman, A.E., Shimi, T., Khayat, M., Ben-Harush, K., Dubrovsky-Gaupp, A., Sapra, K.T., Goldman, R.D., and Medalia, O. (2017). The molecular architecture of lamins in somatic cells. Nature *543*, 261–264.

Tzur, Y.B., Wilson, K.L., and Gruenbaum, Y. (2006). SUN-domain proteins: "Velcro" that links the nucleoskeleton to the cytoskeleton. Nat. Rev. Mol. Cell Biol. *7*, 782–788.

Ungricht, R., and Kutay, U. (2017). Mechanisms and functions of nuclear envelope remodelling. Nat. Rev. Mol. Cell Biol. *18*, 229–245.

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 Biol. *13*, 200.

Wilson, K.L., and Foisner, R. (2010). Lamin-binding Proteins. Cold Spring Harb Perspect Biol *2*, a000554.

Wu, R.S., Panusz, H.T., Hatch, C.L., and Bonner, W.M. (1986). Histones and their modifications. CRC Crit. Rev. Biochem. *20*, 201–263.

Xu, X.M., Meulia, T., and Meier, I. (2007). Anchorage of plant RanGAP to the nuclear envelope involves novel nuclear-pore-associated proteins. Curr. Biol. *17*, 1157–1163.

Yang, Y., Wang, W., Chu, Z., Zhu, J.-K., and Zhang, H. (2017). Roles of Nuclear Pores and Nucleo-cytoplasmic Trafficking in Plant Stress Responses. Front. Plant Sci. *8*, 574.

Zhao, Q., Leung, S., Corbett, A.H., and Meier, I. (2006). Identification and characterization of the Arabidopsis orthologs of nuclear transport factor 2, the nuclear import factor of ran. Plant Physiol. *140*, 869–878.

Zhou, X., and Meier, I. (2013). How plants LINC the SUN to KASH. Nucleus *4*, 206–215.

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. Proc. Natl. Acad. Sci. U.S.A.

Zhou, X., Graumann, K., Evans, D.E., and Meier, I. (2012). Novel plant SUN-KASH bridges are involved in RanGAP anchoring and nuclear shape determination. J. Cell Biol. *196*, 203–211.

Zhou, X., Graumann, K., Wirthmueller, L., Jones, J.D.G., and Meier, I. (2014). Identification of unique SUN-interacting nuclear envelope proteins with diverse functions in plants. J. Cell Biol. *205*, 677–692.

Zhou, X., Graumann, K., and Meier, I. (2015). The plant nuclear envelope as a multifunctional platform LINCed by SUN and KASH. J. Exp. Bot. *66*, 1649–1659.

APPENDIX

APPENDIX

Appendix I: Primer table for genotyping and transcript analysis

Primer Description	Primer name	Forward/ Reverse	Tm (°C)	Sequence
NEAP1_SAIL846_B07	CT487	F	49	CTCTGCAGCTTTCTTGTCTGG
	CT488	R	47	AGCTTGAAGCTTCTGCATCTG
SAIL_T-DNA	AP9	F	67	TAGCATCTGAATTTCATAACCAATCTCGATACAC
NEAP2_GABI_589B02: insertion:chr5 9409811	CT497	F	45	AAAGGGCCATTGATTACCAAG
	CT498	R	45	AGAAATTCGGAAGGGAAAGAC
GABI_9474_T-DNA	CT491	R	50	ATAATAACGCTGCGGACATCTACATTTT
neap2_surroundingCRISPRmutation	CT740	F	58	AGTGGACGATCTGAGATCACA
	CT741	R	59	TCTGCTCTTCTTTGTCTCTCCA
NEAP3_WiscDsLoxHs086_02C	CT492	F	59	TTCCTACCAAACCCAGAAACC
	CT493	R	57	TCAGCCAATTCCTTCACAAAC
WiscDsLoxHs_T-DNA	CT494	F	55	TGATCCATGTAGATTTCCCGGACATGAAG
NEAP1.1_RT-PCR_avalTDNA_450-961	CT569	F	60	GCAGAGGCAAGTGCTGATTC
NEAP2.3_RT-PCR_autourTDNA_161-345	CT570	F	59	AAGATTTCCCCACATGCATTGAA
	CT571	R	60	AACGTTCCGCCTGAAACTCT
NEAP2.3_RT-PCR_5'CDS	CT572	F	59	GTCGGATTCCGTCAAAACGA
	CT573	R	60	CACTTGCTTTAACTCAGCTGCC
NEAP3_RT-PCR_autourTDNA_123-322	CT574	F	60	ATGTGGTGTCTTTGGCCACT
NEAP3_RT-PCR_avalTDNA_215-322	CT576	F	57	GAGGCAGAAACAAGAGTTAAGAG
actin_qPCR_ACT2_At3g18780	CT649	F	65	GAGAGATTCAGATGCCCAGAAGTC
	CT650	R	57	TGGATTCCAGCAGCTTCCA
NEAP1.1_qRT-PCR	CT554	F	58	GGGAAAGAAGGGACATGGAG
	CT555	R	59	CTTGAACCCTGAATTGTCCCAG
NEAP2.3_qRT-PCR	CT556	F	58	ATCGGTTTCAGGGGAGAAAC
	CT557	R	60	TGCAACTACTGCATTTACACACC
NEAP3_qRT-PCR	CT558	F	60	TGGGAGAGCGAAACGATGTG
	CT559	R	60	TCAACGCCTAGAAAACGCAAC

Appendix II: Primer table used for vector constructs

Primer Description	Primer name	Forward/ Reverse	Tm (°C)	Sequence
AttB1_NEAP1_5'	CT495	F	53	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC ATGTCTTATTCTGAAAAAACGA
AttB2_NEAP1_3'	CT496	R	55	GGGG AC CAC TTT GTA CAA GAA AGC TGG GTG TCATCTCTTGGAGACTACC
AttB2_NEAP1ΔTM_3'_(Δ324-stop)	CT499	R	55	GGGG AC CAC TTT GTA CAA GAA AGC TGG GTG TTTATCCCAAGACTCATTGTTA
AttB1_NEAP2_5'	CT500	F	56	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC ATGTCGGATTCCGTCAAAAC
AttB2_NEAP2ΔTM_3'_(Δ311-stop)	CT501	R	55	GGGG AC CAC TTT GTA CAA GAA AGC TGG GTG TTGTTTCTCCCCTGAAACC
AttB1_NEAP3_5'	CT502	F	54	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC ATGCCAACTTCTGTTAGTCT
AttB2_NEAP3ΔTM_3'_(Δ314-stop)	CT503	R	55	GGGG AC CAC TTT GTA CAA GAA AGC TGG GTG CAAGTTTTGGTTGCTAGAATG
NEAP2_CRISPR_25_Exon3	CT577	F	58	ATTGAAGGCACAGAATGTGAGCTC
	CT578	R	58	AAACGAGCTCACATTCTGTGCCTT
NEAP2_CRISPR_56_Exon1	CT579	F	52	ATTGCGCTTTTGAAAGATTTGGAT
	CT580	R	52	AAACATCCAAATCTTTCAAAAGCG
CRISPR_M13_RP	AP198	R	18	CAGGAAACAGCTATGACC
CRISPR_ss42	AP619		22	TCCCAGGATTAGAATGATTAGG
CRISPR_ss43	AP620		22	CGACTAAGGGTTTCTTATATGC
CRISPR_ss61	AP621		24	GAGCTCCAGGCCTCCCAGCTTTCG

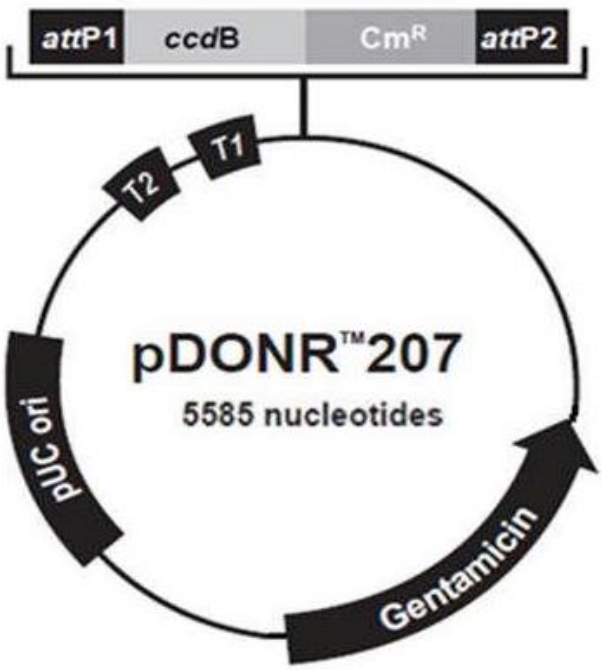
Appendix III: pDONR Gateway vectors constructed in this study

Table (**A**) and maps (**B - D**) of the different pDONR Gateway vectors.

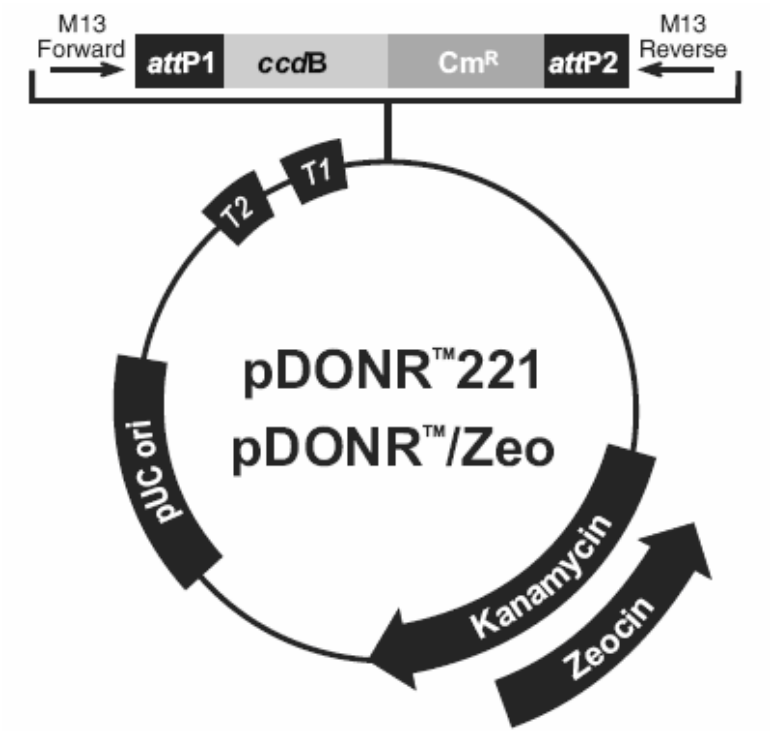
A

Type of vector	Construct description	Bacterial resistance	Promoter
pDONR_207	AtNEAP2	Gentamycin	**UC origin**
pDONR_Zeo	Atneap2_CRISPRmutation	Zeocyn	**UC origin**
	AtNEAP1		
	AtNEAP3		
	AtNEAP1_ΔTM		
	AtNEAP2_ΔTM		
	AtNEAP3_ΔTM		
pEn-Chimera	**Exon1_NEAP2_CRISPR**	**Ampicillin**	**AtU6-26**
	Exon3_NEAP2_CRISPR		

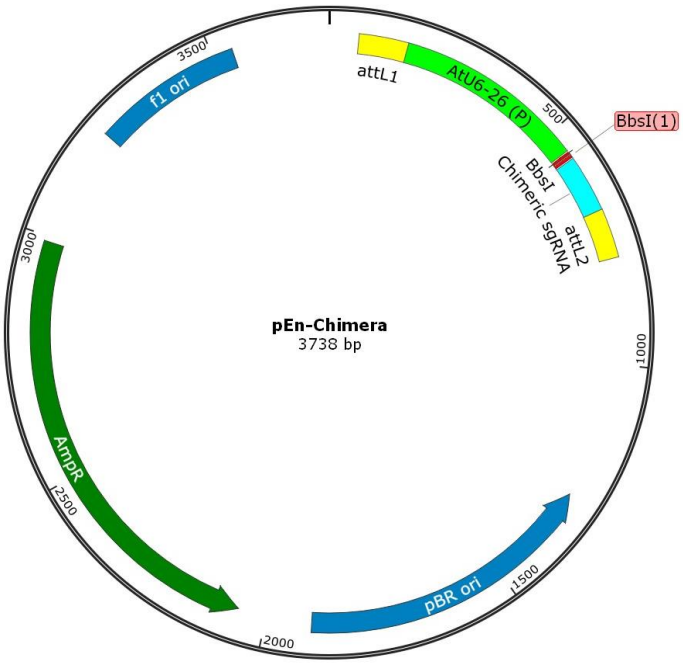
B



C


M13
Forward
attP1
ccdB
CmR
attP2
M13
Reverse
T1
T2
pDONR™221
pDONR™/Zeo
pUC ori
Kanamycin
Zeocin

D


3500
f1 ori
attL1
AtU6-26 (P)
500
BbsI(1)
BbsI
Chimeric sgRNA
attL2
3000
pEn-Chimera
3738 bp
1000
AmpR
2500
pBR ori
1500
2000

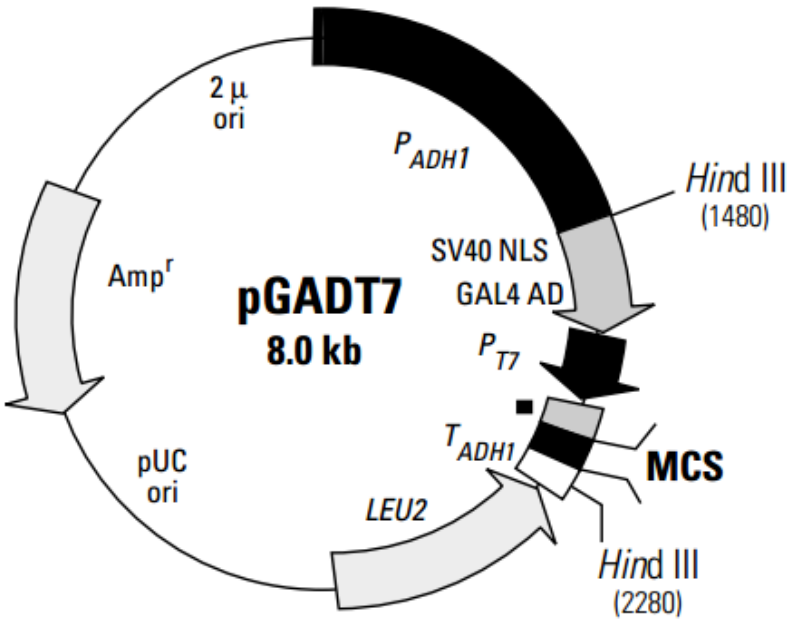
Appendix IV: pDEST Gateway vectors constructed in this study for Y2H

Table (**A**) and maps (**B** and **C**) of the different pDEST Gateway vectors.

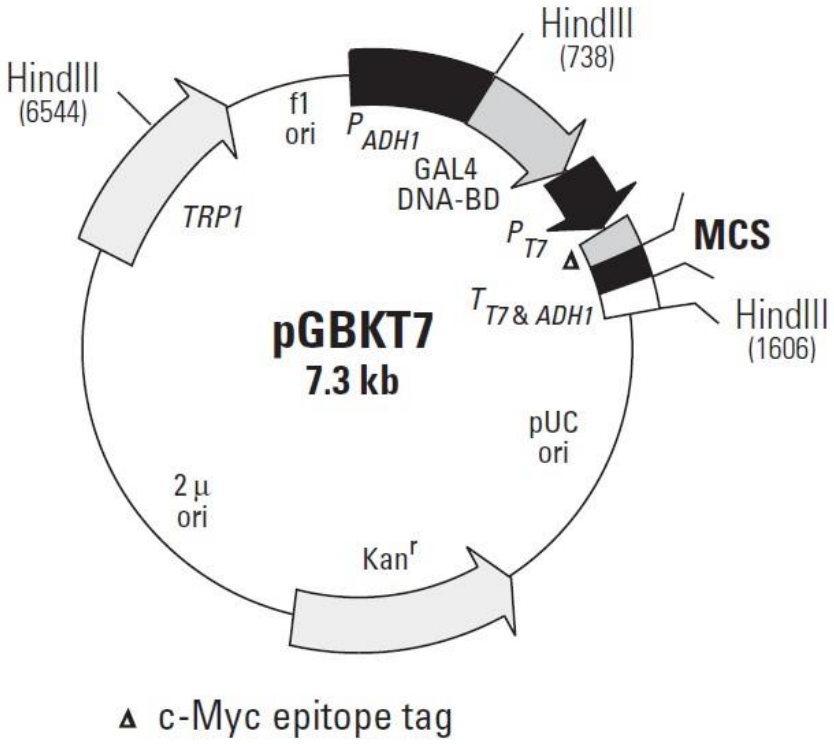
A

Type of vector	Construct description	Bacterial resistance	Promoter	Side of the fusion
pGADT7	AtNEAP1	Ampicillin	ADH1	N-ter
	AtNEAP2			
	AtNEAP3			
	AtNEAP1_ΔTM			
	AtNEAP2_ΔTM			
	AtNEAP3_ΔTM			
pGBKT7	AtNEAP1	Kanamycin		
	AtNEAP2			
	AtNEAP3			
	AtNEAP1_ΔTM			
	AtNEAP2_ΔTM			
	AtNEAP3_ΔTM			

B


2 µ ori
P_{ADH1}
Hind III (1480)
SV40 NLS
GAL4 AD
Amp^r
pGADT7
8.0 kb
P_{T7}
T_{ADH1}
MCS
pUC ori
LEU2
Hind III (2280)
■ HA epitope tag

C


HindIII (738)
HindIII (6544)
f1 ori
P_{ADH1}
GAL4 DNA-BD
TRP1
P_{T7}
MCS
$T_{T7\ \&\ ADH1}$
HindIII (1606)
pGBKT7
7.3 kb
pUC ori
2 µ ori
Kan^r
▲ c-Myc epitope tag

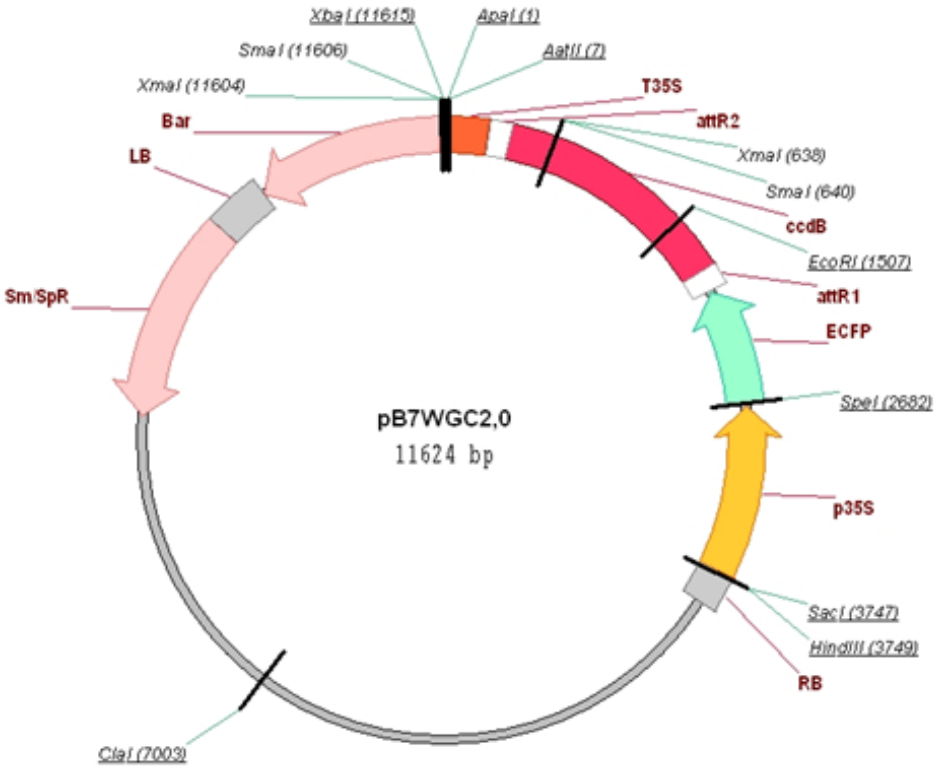
Appendix V: pDEST Gateway vectors constructed in this study for plant transformation

Table (**A**) and maps (**B** - **E**) of the different pDEST Gateway vectors.

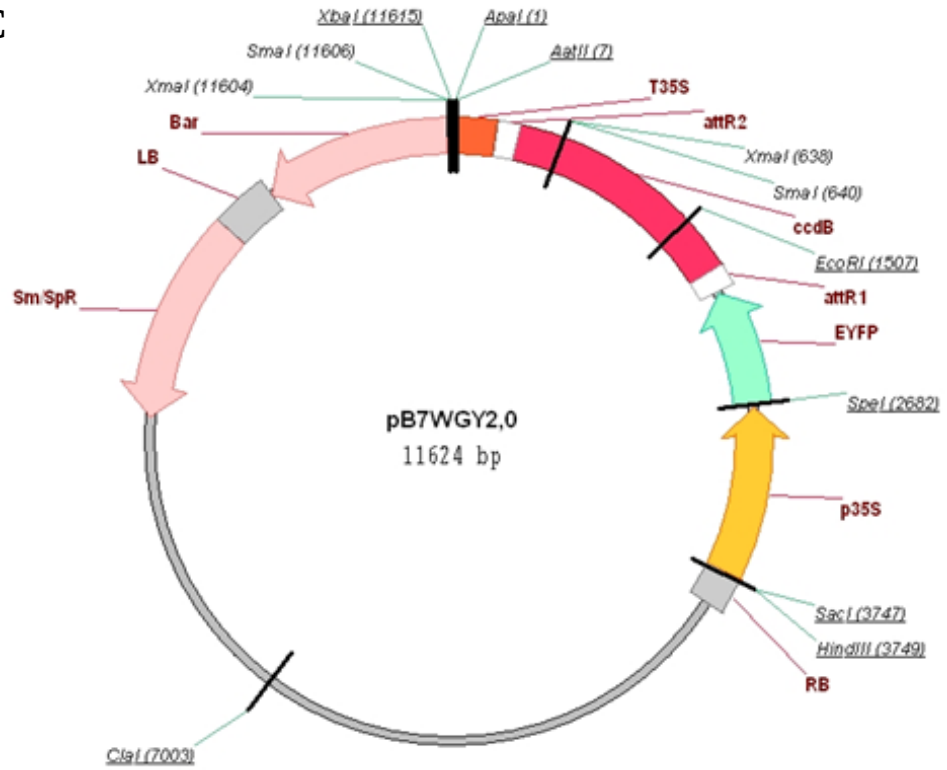
A

Type of vector	Construct description	Bacterial/Plant resistance	Promoter	Side of the fusion
pB7WGC2	CFP_AtNEAP1	Spectinomycin / Basta	35S	N-ter
	CFP_AtNEAP2			
	CFP_Atneap2_CRISPRmutation			
	CFP_AtNEAP3			
pB7WGY2	YFP_AtNEAP1			
	YFP_AtNEAP2			
	YFP_Atneap2_CRISPRmutation			
	YFP_AtNEAP3			
	YFP_AtNEAP1_ΔTM			
	YFP_AtNEAP2_ΔTM			
	YFP_AtNEAP3_ΔTM			
pGWB605	NLS_GFP			C-ter
pDe_Cas9	Exon1NEAP2_CRISPR		cUBI4-2	N/A
	Exon3NEAP2_CRISPR			

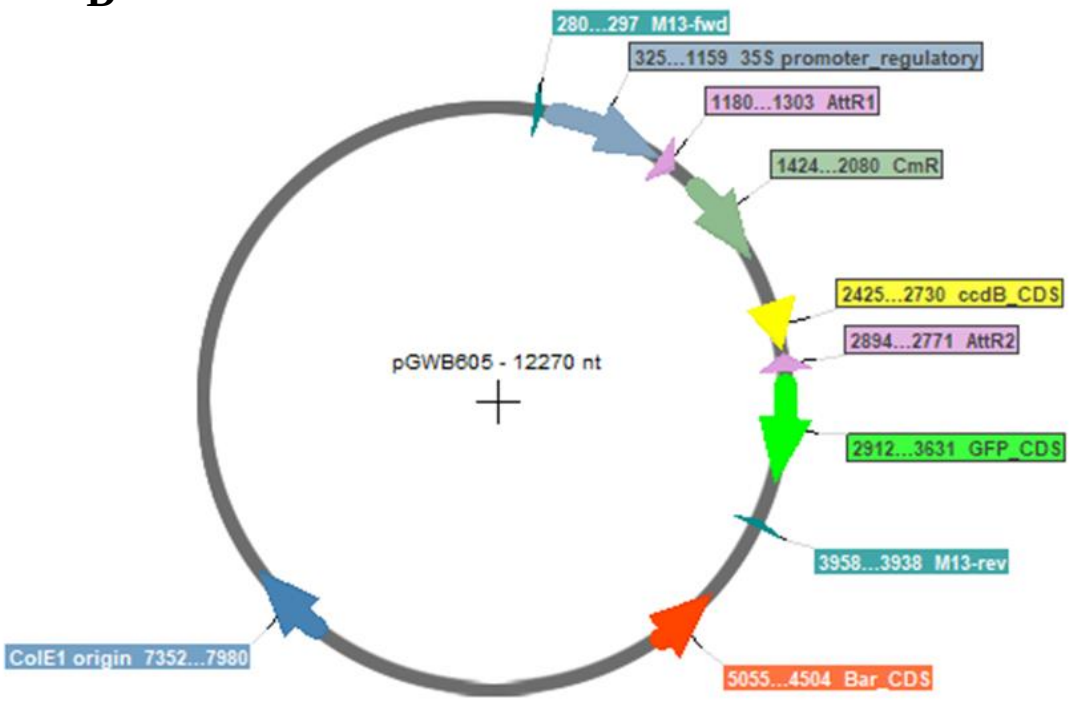
B


XbaI (11615)
ApaI (1)
SmaI (11606)
AatII (7)
XmaI (11604)
T35S
Bar
attR2
LB
XmaI (638)
SmaI (640)
ccdB
EcoRI (1507)
Sm/SpR
attR1
ECFP
SpeI (2682)
pB7WGC2,0
11624 bp
p35S
SacI (3747)
HindIII (3749)
RB
ClaI (7003)

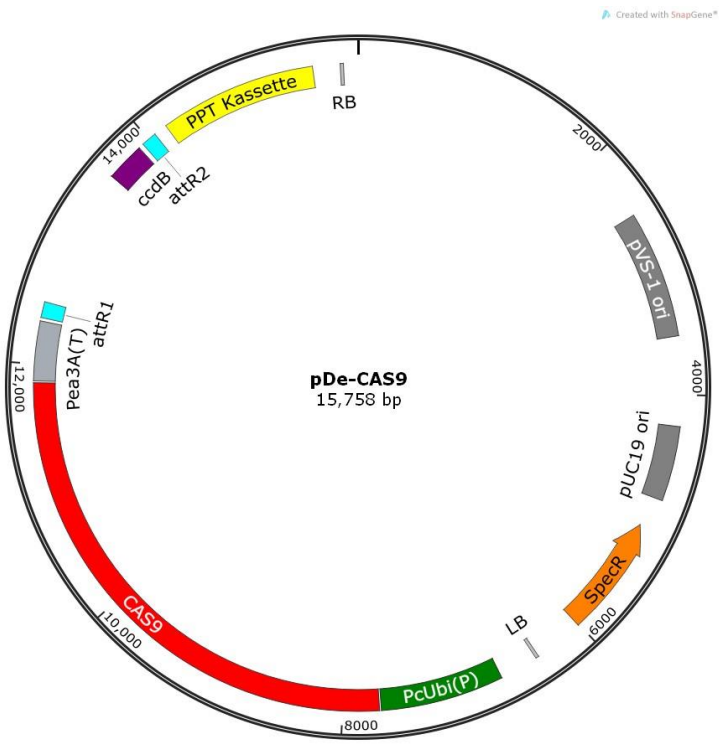
C


XbaI (11615)
ApaI (1)
SmaI (11606)
AatII (7)
XmaI (11604)
T35S
Bar
attR2
LB
XmaI (638)
SmaI (640)
ccdB
EcoRI (1507)
Sm/SpR
attR1
EYFP
SpeI (2682)
pB7WGY2,0
11624 bp
p35S
SacI (3747)
HindIII (3749)
RB
ClaI (7003)

D


280...297 M13-fwd
325...1159 35S promoter_regulatory
1180...1303 AttR1
1424...2080 CmR
2425...2730 ccdB_CDS
2894...2771 AttR2
2912...3631 GFP_CDS
3958...3938 M13-rev
5055...4504 Bar_CDS
ColE1 origin 7352...7980
pGWB605 - 12270 nt

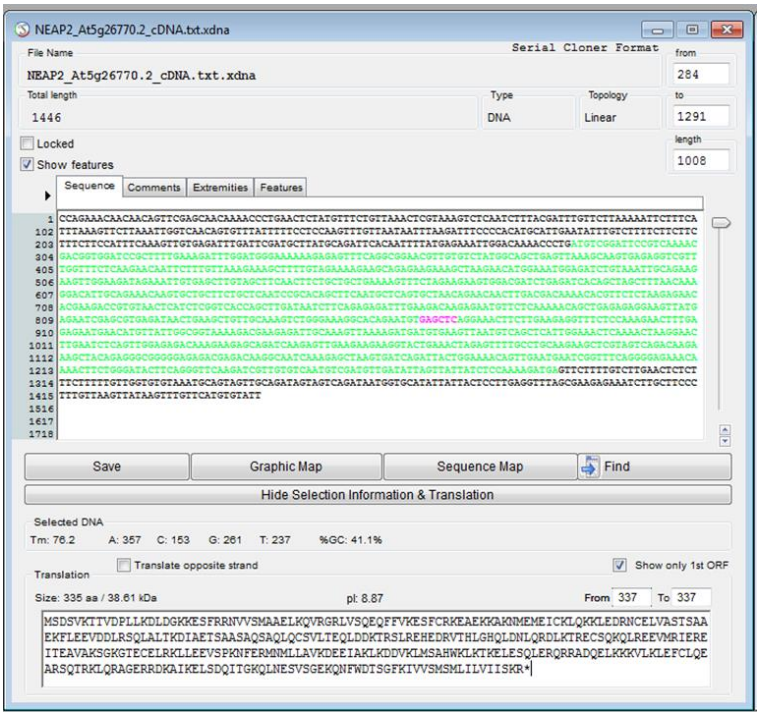
E


Created with SnapGene®
PPT Kassette
RB
ccdB
attR2
pVS-1 ori
pUC19 ori
SpecR
LB
PcUbi(P)
CAS9
Pea3A(T)
attR1
pDe-CAS9
15,758 bp
2000
4000
6000
8000
10,000
12,000
14,000

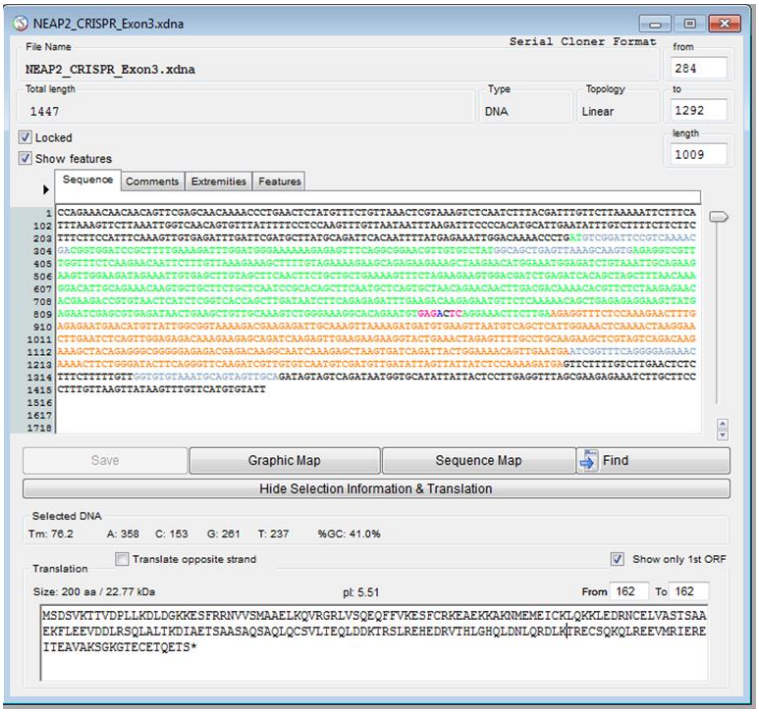
Appendix VI: In-silico WT and mutant transcripts and protein for AtNEAP2

A. *In-silico AtNEAP2* transcripts and prediction of AtNEAP2 protein. **B.** Potential *Atneap2* transcript and Atneap2 mutant protein. Wild type transcript (green), targeted site for CRIPSR (pink) and untranscribed region after an early stop codon (orange) are indicated. **Software:** Serial cloner.

A

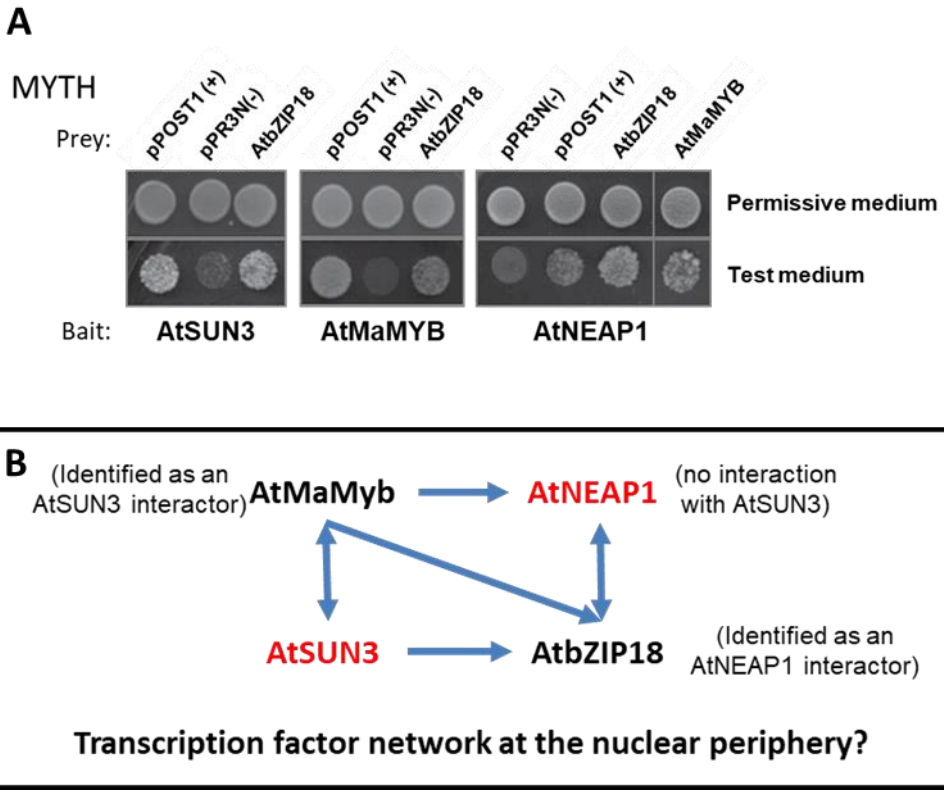


B



Appendix VII: Interaction of AtbZIP18 and AtMaMYB transcription factors with AtNEAP1

A. MYTH experiment carried out by Voisin, unpublished. Yeast strains were tested on permissive medium (upper panel) depleted in leucine and tryptophan to select diploïds only and on test medium (lower panel) depleted in leucine, tryptophan, histidine and adenine to select diploids with interacting proteins. Baits AtSUN3, AtMaMYB and AtNEAP1 were tested with pPOST and pPR3N respectively as positive and negative controls and with prey AtbZIP18. Bait AtNEAP1 was also tested with prey AtMaMYB. Data from Voisin, unpublished. **B.** Model of a potential network with transcription factors anchored at the nuclear periphery by AtSUN3 and AtNEAP1.



Appendix VIII: Scientific contribution during the PhD

Posters:

2016 – SEB meeting in Brighton, UK

2018 – JED Clermont-Ferrand, France

2019 – Post Graduate Symposium, Oxford, UK

Scientific papers:

Pawar, V., Poulet, A., **Détourné, G.**, Tatout, C., Vanrobays, E., Evans, D.E., and Graumann, K. (2016). A novel family of plant nuclear envelope-associated proteins. J. Exp. Bot. 67, 5699–5710

Duc, C., Benoit, M., **Détourné, G.**, Simon, L., Poulet, A., Jung, M., Veluchamy, A., Latrasse, D., Le Goff, S., Cotterel, S., Tatout, C., Benhamed, M., and Probst, AV. (2017). Arabidopsis ATRX Modulates H3.3 Occupancy and Fine-Tunes Gene Expression. Plant Cell 29, 1773–1793

Characterisation of a novel family of plant nuclear envelope associated proteins (NEAP) in *Arabidopsis thaliana*

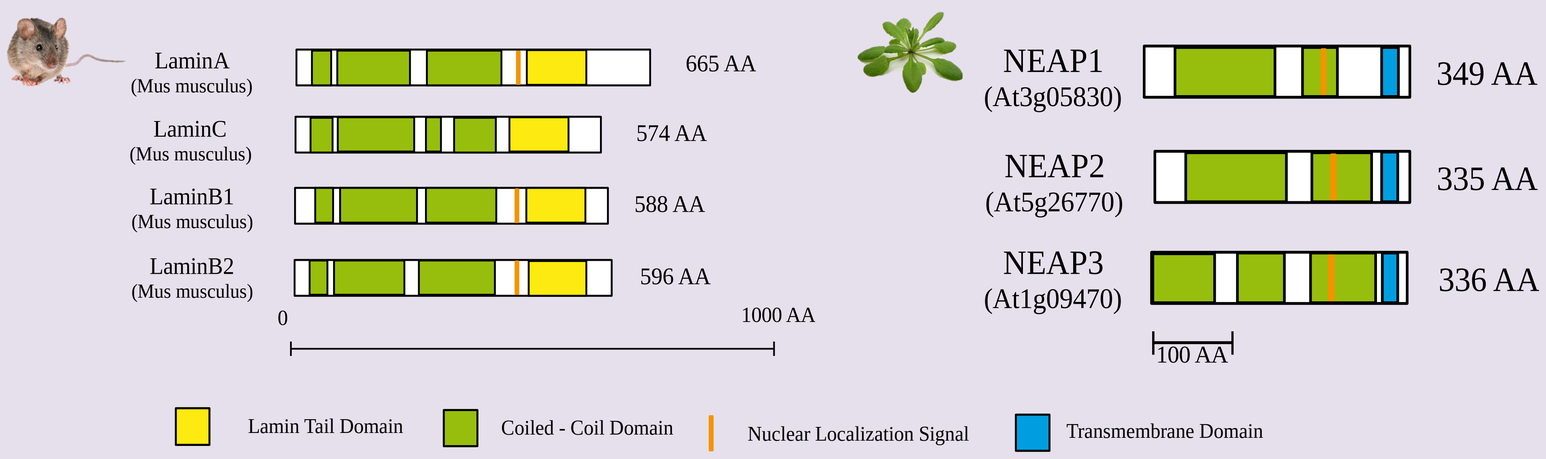
Gwénaëlle DETOURNE[1,2], Vidya PAWAR[2], Axel POULET[1,2], Emmanuel VANROBAYS[1], Katja GRAUMANN[2], Christophe TATOUT[1] and David EVANS[2]

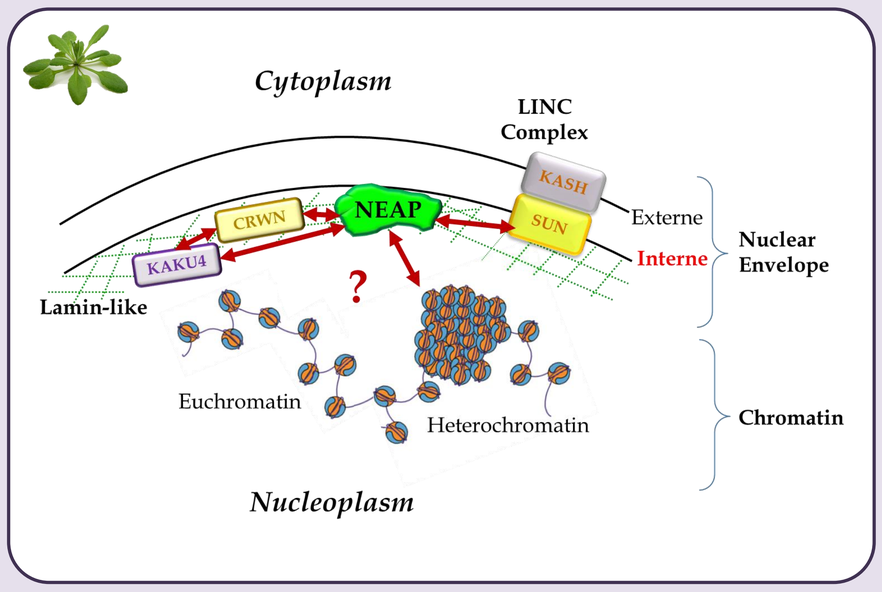
[1] CNRS UMR6293 INSERM U1103, Génétique Reproduction et Développement (GReD), Université Clermont Auvergne, Clermont-Ferrand, France
[2]Department of Biological and Medical Sciences, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, United-Kingdom

Animals and fungi contain a meshwork structure beneath the inner nuclear membrane (INM) which is called, in animals, the lamina, and composed of lamins, Lamin B Receptor and LEM (LAP2β, Emerin and Man1) protein families. The lamina interacts with chromatin and lamina mutations affect chromatin organisation, modify gene expression and are associated with severe human diseases. Much less is known about the plant lamina and this project aims to characterize new components of the INM and their possible interactions with chromatin.

Screening for new proteins with nuclear membrane properties

NEAPs contain coiled-coil domains , a transmembrane domain and a nuclear localisation signal

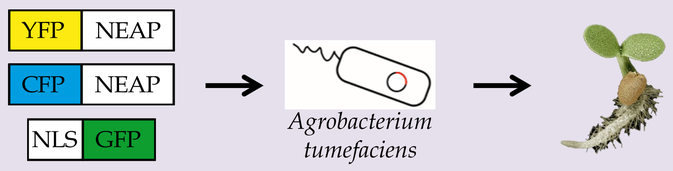


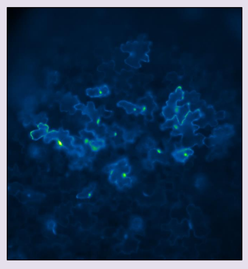


Investigate participation of NEAP in a protein network at the nuclear periphery and potential interactions with chromatin

NEAP *in vivo* localisation by transient expression in *A. thaliana*

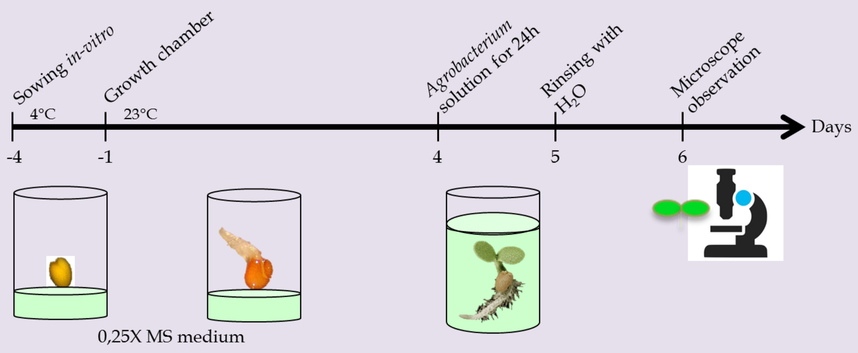
Transient expression with fluorescent fusion proteins

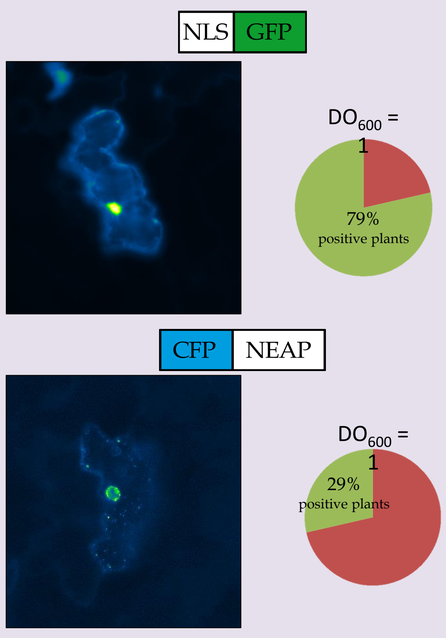




Patches of « expressing cells »

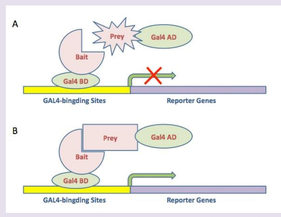
Protocol details:



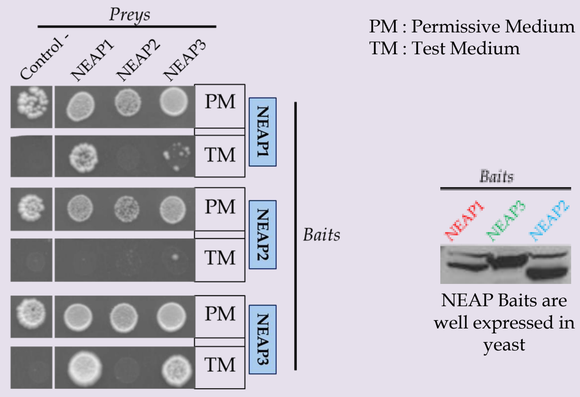


- **Investigate NEAP nuclear localisation in lamina mutants *(neap, crwn, kaku4, etc…)***
- **Investigate NEAP interaction *in planta* (FRET experiments)**

NEAPs form homo- and hetero- dimers



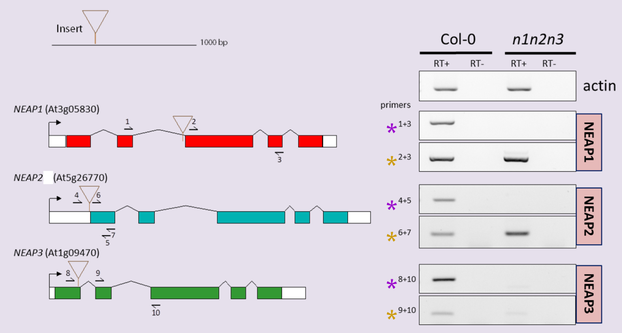
Yeast Two Hybrid system



NEAP Baits are well expressed in yeast

- **Screening of cDNA library (Clontech) using each NEAP as bait**
- **Investigate NEAP interaction with other proteins of the nuclear periphery**

neap triple mutant impacts nuclear morphology and chromatin organisation

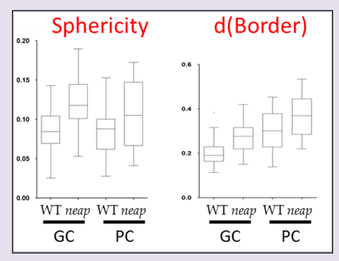


* Primers surrounding insertion site

* Primers downstream of T-DNA insertion

GC : Guard cells PC : Pavement cells

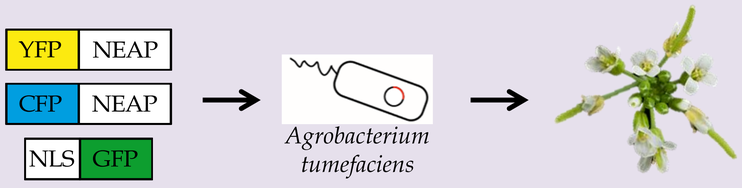
***NucleusJ* analysis**



- **Generate CRISPR-Cas9 mutant for NEAP2**
- **Characterisation of *neap* triple mutant**

Interaction of NEAPs with chromatin

Stable transformation of NEAP fluorescent fusion proteins



- **ChIP using GFP antibody: identification of sequences or chromatin domains interacting with NEAPs**
- **GFP-trap using GFP antibody: identification of new NEAP protein partners**

Conclusions and Perspectives

→ New family of nuclear envelope proteins (NEAPs) identified
- ✓ NEAPs interact with each other
- ✓ They are localised at the nuclear periphery

→ Developing tools and methods to identify:
- ❖ new NEAP-interacting proteins
- ❖ NEAP-associated chromatin domains

Gwénaëlle DETOURNE, first year PhD student
Co-tutelle between Oxford (UK) and Clermont-Ferrand (F)
Contact: gwenaelle_detourne@hotmail.com

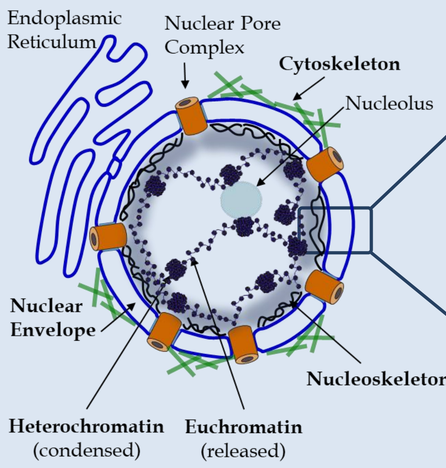
Characterisation of a novel family of plant nuclear envelope associated proteins (NEAP) in *Arabidopsis thaliana*

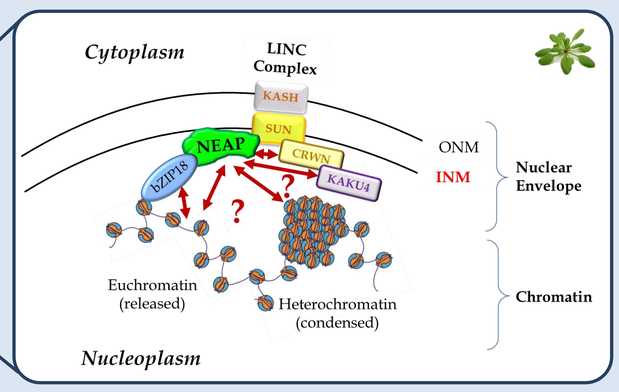
Gwénaëlle DETOURNE[1,2], Emmanuel VANROBAYS[1], Katja GRAUMANN[2], Aline V. Probst[1], Christophe TATOUT[1] and David EVANS[2]

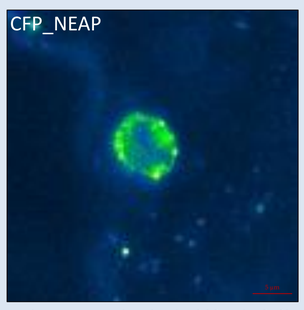
[1] CNRS UMR6293 INSERM U1103, Génétique Reproduction et Développement (GReD), Université Clermont Auvergne, Clermont-Ferrand, France
[2]Department of Biological and Medical Sciences, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, United-Kingdom

The nucleoskeleton underneath the nuclear envelope is needed to transmit signals to the nucleus and induce changes in chromatin organisation and ultimately gene expression. A novel family of Nuclear Envelope Associated Proteins (NEAPs) proposed to be new components of the plant nucleoskeleton has been recently evidenced in the model plant *Arabidopsis thaliana*. They are anchored at the inner nuclear membrane (INM) and results suggest they are part of a protein network responsible for maintaining nuclear morphology and chromatin organisation.

The Nuclear Periphery organisation and the NEAPs







Transient expression in *Arabidopsis thaliana*

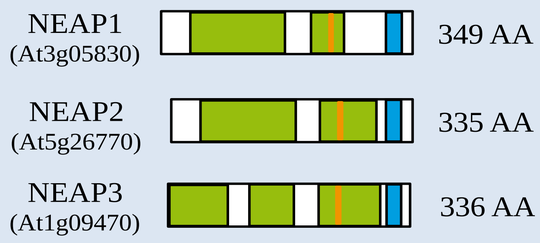
NEAPs are localised at the **Nuclear Periphery**

NEAP proteins contain

Coiled-Coil domains

a **TransMembrane Domain**

and a **Nuclear Localisation Signal**



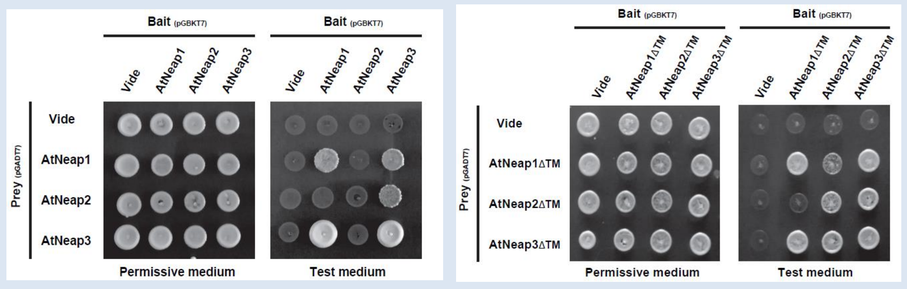
Investigate participation of NEAP in a protein network at the nuclear periphery and potential interactions with chromatin

Study of the NEAP interactome

Yeast Two Hybrid system

Full length NEAPs

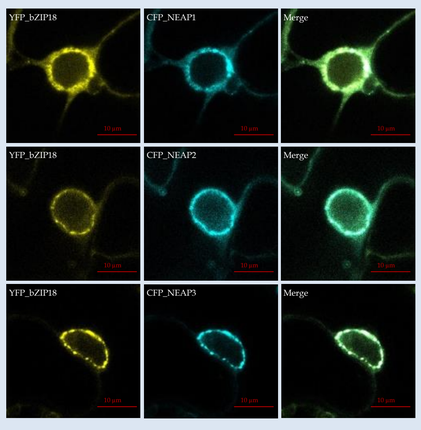
NEAPs without TM domain



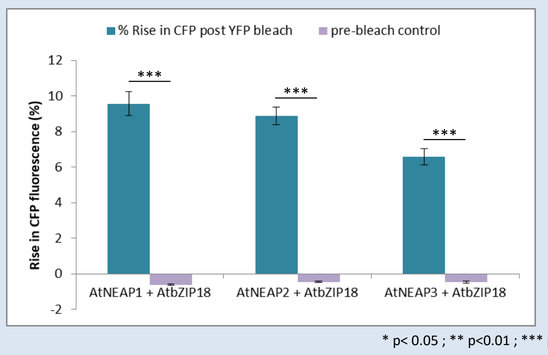
NEAP proteins are able to interact with each other
No interactions with other NE proteins found in Y2H nor with new interactors

But Interaction of Transcription Factor **bZIP18** with **NEAP1** found by Membrane Y2H (Pawar et al, 2016)

→Investigate bZIP18/NEAP interaction *in planta* by apFRET experiment
(acceptor photobleaching Fluorescence Resonance Energy Transfer)



Transient expression in *Nicotiana benthamiana* leaf epidermal cells

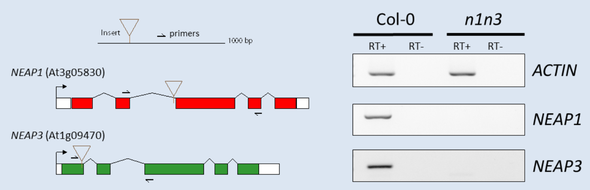


* p< 0.05 ; ** p<0.01 ; *** p<0.001

Interactions confirmed for **AtbZIP18** and **AtNEAPs**

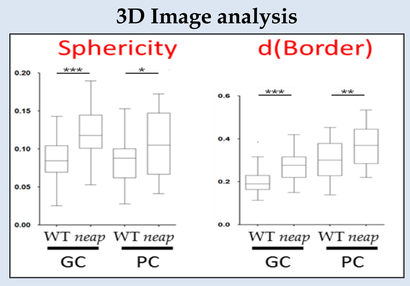
Functional analysis of *neap* mutants

A ***neap1neap3* double mutant**

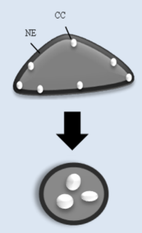


neap1neap3 is a double knock-out mutant

3D Image analysis

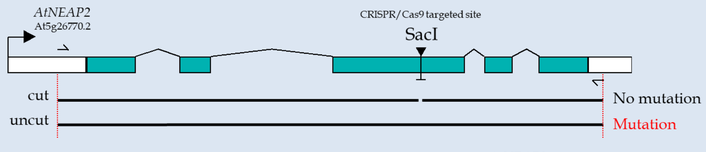


GC : Guard cells
PC : Pavement cells

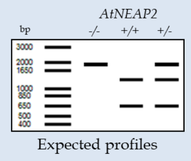


Double *neap* mutant alters **nucleus shape** and **chromocentre position**

B ***neap2* mutant created by CRISPR/Cas9**



SacI
5′…G AGCT C…3′
3′…C TCGA G…5′

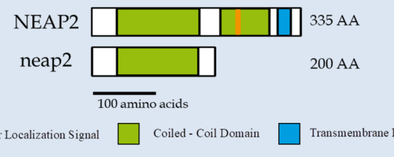


Expected profiles

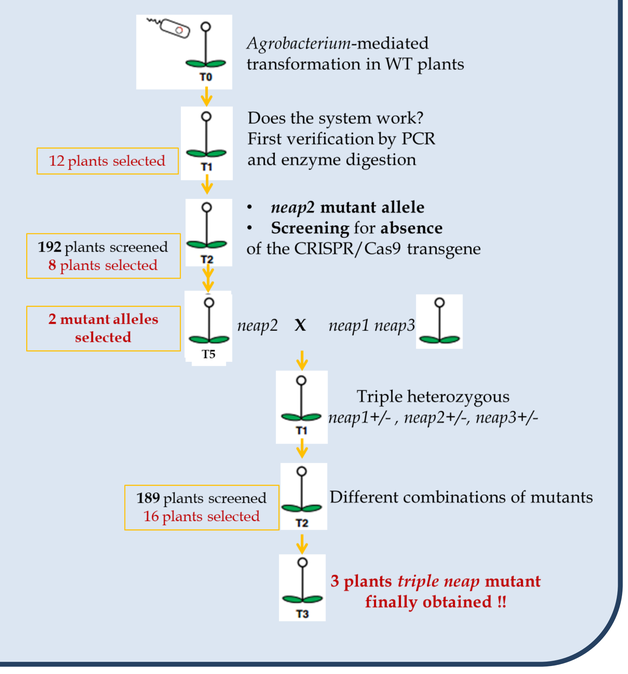
Single nucleotide insertion

AtNEAP2	TCCTGAG – CTCACAT
Atneap2-1	TCCTGAG**T**CTCACAT
Atneap2-2	TCCTGAG**C**CTCACAT

Early stop codon after insertion site leads to a truncated protein missing NLS, one CC and TM domains



C **Obtaining triple *neap* mutant**

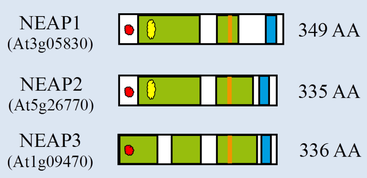


Designing AtNEAP antibodies

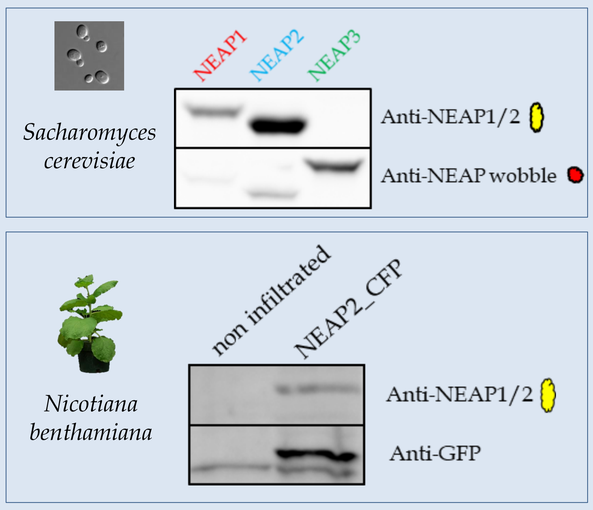
anti-NEAP1/2 (QLDDKTRSLRE)

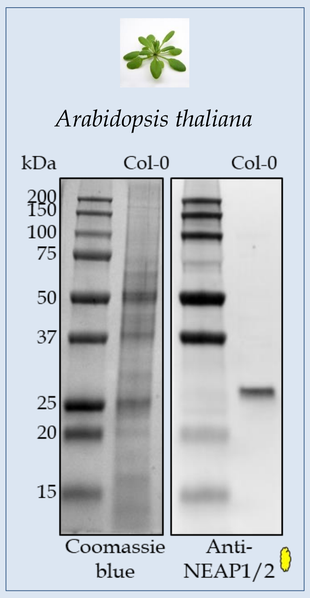
anti-NEAP wobble (H-DL-D/G-E/H-KK-E/H-SFRRNVVS-C-NH_2)

Targeting N-terminus of NEAPs



Antibodies work for Western blotting





→ **Characterize *neap* single, double and triple mutants concerning protein synthesis**

→ **Use them for ImmunoPrecipitation (and co-IP) in order to look for new NEAP interactors by Mass Spectrometry**

Conclusions and Perspectives

- Characterize single *neap2* and triple *neap* mutants
  - 3D image analysis for nuclear morphology
  - Identify NEAPs-regulated genes (RNAseq)
- Look for new NEAP partners by different techniques
  - IP/co-IP followed by Mass Spectrometry
  - apFRET analysis with known NE proteins

Gwénaëlle DETOURNE, third year PhD student
Co-tutelle between Oxford (UK) and Clermont-Ferrand (F)
Contact: 15059427@brookes.ac.uk

Nigel Groome Studentship

Characterisation of a novel family of plant nuclear envelope associated proteins (NEAP) in *Arabidopsis thaliana*

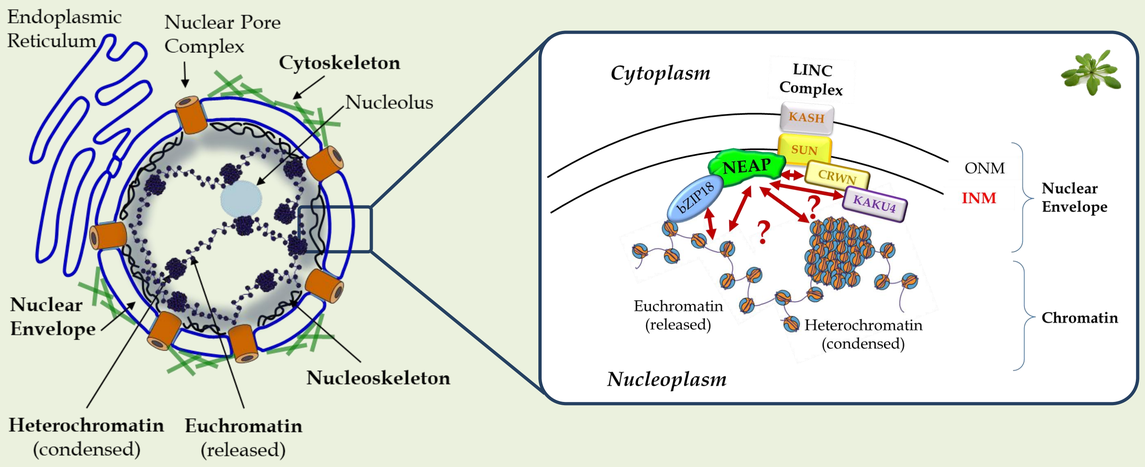
Gwénaëlle DETOURNE[1,2], Emmanuel VANROBAYS[1], Katja GRAUMANN[2], Aline V. Probst[1], Christophe TATOUT[1] and David EVANS[2]

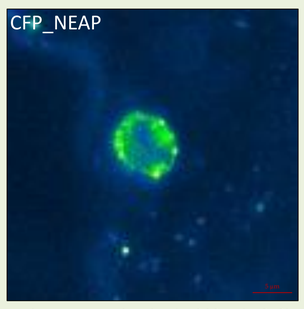
[1] CNRS UMR6293 INSERM U1103, Génétique Reproduction et Développement (GReD), Université Clermont Auvergne, Clermont-Ferrand, France

[2]Department of Biological and Medical Sciences, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, United-Kingdom

The nucleoskeleton underneath the nuclear envelope is needed to transmit signals to the nucleus and induce changes in chromatin organization and ultimately gene expression. A novel family of Nuclear Envelope Associated Proteins (NEAPs) proposed to be new components of the plant nucleoskeleton has been recently evidenced in the model plant *Arabidopsis thaliana*. They are anchored at the inner nuclear membrane (INM) and results suggest they are part of a protein network responsible for maintaining nuclear morphology and chromatin organization.

The Nuclear Periphery organisation and the NEAPs





Transient expression in *Arabidopsis thaliana*

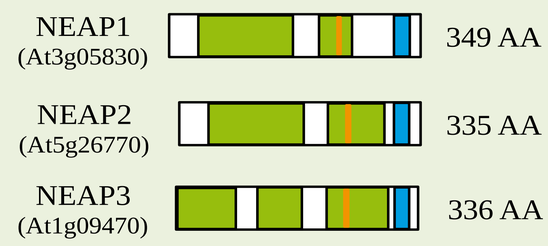
NEAPs are localised at the **Nuclear Periphery**

NEAP proteins contain

Coiled-Coil domains

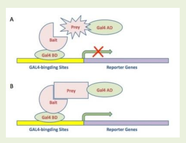
a **TransMembrane Domain**

and a **Nuclear Localisation Signal**



Investigate participation of NEAP in a protein network at the nuclear periphery and potential interactions with chromatin

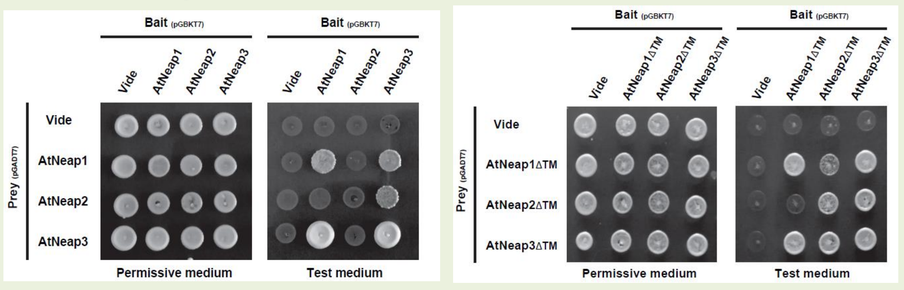
Study of the NEAP interactome



Yeast Two Hybrid system

Full length NEAPs

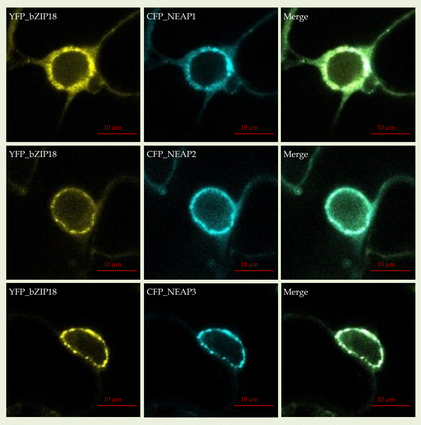
NEAPs without TM domain



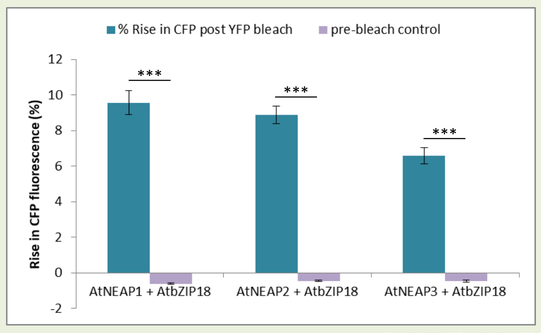
NEAP proteins are able to interact with each other
Other interactors of NEAP are difficult to identify by classical Y2H

Interaction of Transcription Factor **bZIP18** with **NEAP1** found by Membrane Y2H
(Pawar et al, 2016)

→Investigate bZIP18/NEAP interaction *in planta* by apFRET experiment
(acceptor photobleaching Fluorescence Resonance Energy Transfer)



Transient expression in *Nicotiana benthamiana* leaf epidermal cells



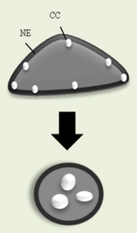
* p< 0.05 ; ** p<0.01 ; *** p<0.001

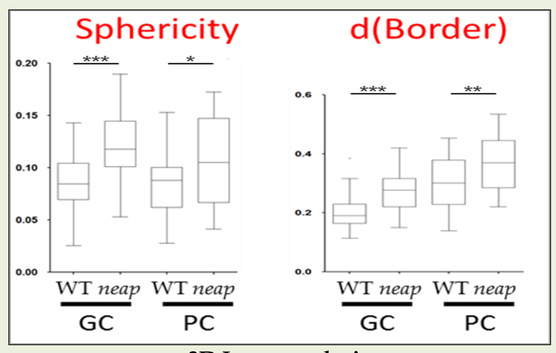
Interactions confirmed between **AtbZIP18** and **AtNEAPs**

Functional analysis of *neap* mutants

A ***neap1neap3* double mutant**

Nucleus shape and **chromocentre position** are **altered** in double *neap1neap3* mutant

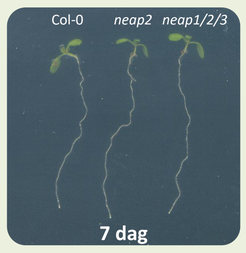


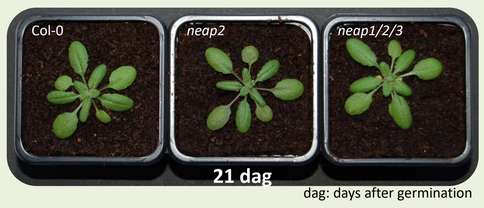


GC : Guard cells
PC : Pavement cells

3D Image analysis

B **New mutants obtained: *neap2* mutant (by CRISPR/Cas9) and triple *neap* mutant, « *neap1/2/3* »**

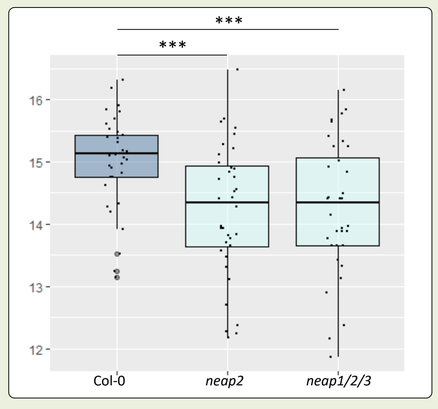




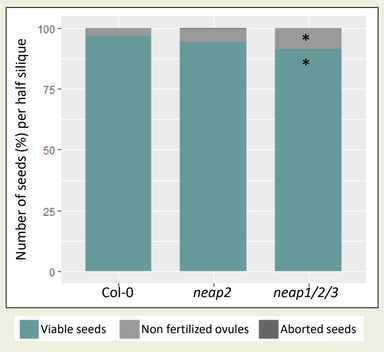
dag: days after germination

General growth is **not** altered

Silique size (mm) is **reduced**



Weak effects on seed production



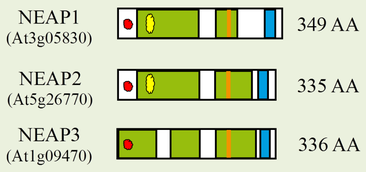
→ Look at nuclear morphology with 3D image analysis
→ Look at meiosis and embryo formation in mutants

Designing AtNEAP antibodies

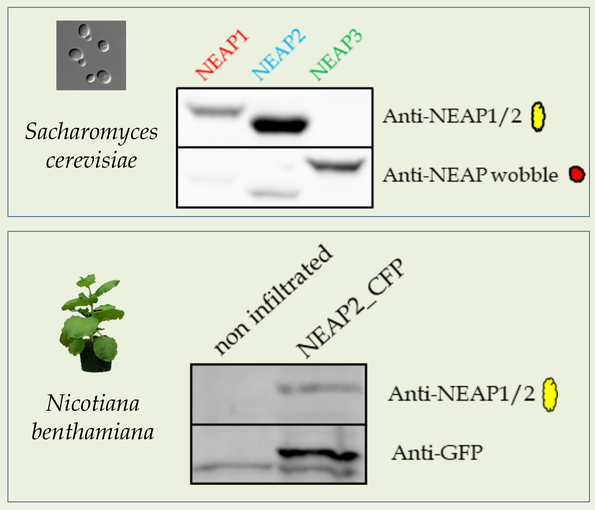
anti-NEAP1/2 (QLDDKTRSLRE)

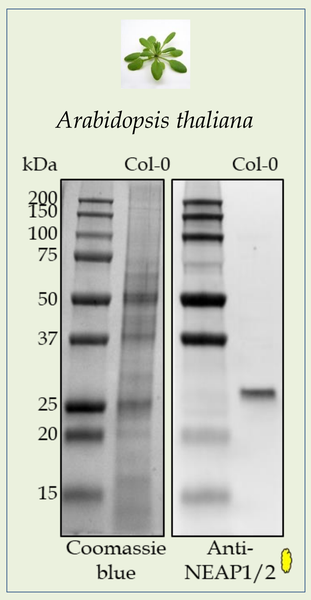
anti-NEAP wobble (H-DL-D/G-E/H-KK-E/H-SFRRNVVS-C-NH_2)

Targeting N-terminus of NEAPs



Antibodies detect proteins with expected MW in Western blotting





→ Characterization of protein synthesis in *neap* single, double and triple mutants
→ Available for immunoprecipitation and Co-IP to detect new NEAP interactors

Conclusions and Perspectives

- Transcription factor AtbZIP18 interact with AtNEAPs
- New mutants obtained showing a reduced silique size
- Antibodies generated, suitable for protein detection

Future work:

- 3D image analysis for nuclear morphology in *neap1/2/3*, single *bZIP18* and *neap1/2/3;bzip18* mutants
- Immunoprecipitation of AtNEAPs to identify new partners by Mass Spectrometry

Gwénaëlle DETOURNE, third year PhD student
Co-tutelle between Oxford (UK) and Clermont-Ferrand (F)
Contact: 15059427@brookes.ac.uk

Nigel Groome Studentship

Journal of Experimental Botany Advance Access published September 14, 2016

Journal of Experimental Botany
doi:10.1093/jxb/erw332

RESEARCH PAPER

A novel family of plant nuclear envelope-associated proteins

Vidya Pawar[1,2], Axel Poulet[1,2], Gwénaëlle Détourné[1,2], Christophe Tatout[2], Emmanuel Vanrobays[2], David E Evans[1] and Katja Graumann[1,*]

[1] Department of Biological and Medical Sciences, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, OX3 0BP, UK
[2] UMR CNRS 6293 INSERM U 1103 Clermont Université, GReD, Avenue des Landais 63171 Aubière Cedex, France

* Correspondence: kgraumann@brookes.ac.uk

Received 19 April 2016; Accepted 17 August 2016

Editor: Steve Penfield, John Innes Centre

Abstract

This paper describes the characterisation of a new family of higher plant nuclear envelope-associated proteins (NEAPs) that interact with other proteins of the nuclear envelope. In the model plant *Arabidopsis thaliana*, the family consists of three genes expressed ubiquitously (*AtNEAP1–3*) and a pseudogene (*AtNEAP4*). NEAPs consist of extensive coiled-coil domains, followed by a nuclear localisation signal and a C-terminal predicted transmembrane domain. Domain deletion mutants confirm the presence of a functional nuclear localisation signal and transmembrane domain. AtNEAP proteins localise to the nuclear periphery as part of stable protein complexes, are able to form homo- and heteromers, and interact with the SUN domain proteins AtSUN1 and AtSUN2, involved in the linker of nucleoskeleton and cytoskeleton (LINC) complex. An *A. thaliana* cDNA library screen identified a putative transcription factor called AtbZIP18 as a novel interactor of AtNEAP1, which suggest a connection between NEAP and chromatin. An *Atneap1 Atneap3* double-knockout mutant showed reduced root growth, and altered nuclear morphology and chromatin structure. Thus AtNEAPs are suggested as inner nuclear membrane-anchored coiled-coil proteins with roles in maintaining nuclear morphology and chromatin structure.

Key words: *Arabidopsis thaliana*, chromatin, higher plant, inner nuclear membrane, LINC complex, nuclear envelope, nucleoskeleton, nuclear lamina, nucleus.

Introduction

The nuclear envelope (NE) in opisthokonts is closely associated with the proteins of the nuclear lamina and chromatin (Crisp *et al.*, 2006; Evans *et al.*, 2014). Proteins integral to the inner nuclear membrane (INM) link chromatin, the lamina, and nuclear membranes. A key protein family involved in this process is the SUN domain protein family. SUN proteins interact with Klarsicht-Anc1-Syne1 homology (KASH) domain proteins in the outer nuclear membrane (ONM), linking to the cytoskeleton, and to lamins in the nucleus, as part of the linker of nucleoskeleton and cytoskeleton (LINC) complex that spans the NE (Tzur *et al.*, 2006). Interactions of the NE, lamina, and chromatin play important roles in meiosis and mitosis, in chromatin positioning and silencing, in positioning nuclei, and in signalling (Okada *et al.* 2005; Gonzalez-Suarez *et al.* 2009; Dechat *et al.* 2010; Smith *et al.*, 2015). Previous studies have shown that SUN proteins are also present in plants and that they are also part of LINC complexes associating with plant-specific KASH proteins at the ONM and plant-specific nuclear filamentous proteins at the nuclear face of the NE (Moriguchi *et al.*, 2005; Graumann *et al.* 2010; Murphy *et al.* 2010; Oda and Fukuda 2011; Graumann and Evans 2011; Graumann *et al.*, 2014).

A nuclear lamina has been described underlying and closely associated with the INM in metazoan (Gruenbaum *et al.*, 2005) and has been suggested to be present in plant nuclei as well (Fiserova *et al.*, 2009). While the lamina of animal cells has been well characterised, that of plants is much less well described. The lamina of animal cells is comprised of lamins, type-5 intermediate filament proteins, and lamin-associated proteins (reviewed by Wilson and Berk, 2010). Sequence homologues of mammalian lamins are not found in plants (Brandizzi *et al.*, 2004; Meier, 2007; Graumann and Evans, 2010; Koreny and Field, 2016). However, there is a meshwork of proteins underlying and attached to the plant INM (Minguez and Moreno Diaz de la Espina, 1993; Masuda *et al.*, 1997; Fiserova *et al.*, 2009; Ciska and Moreno Díaz de la Espina, 2013; Sakamoto and Takagi, 2013). Proteins that may be components of this plant 'lamina' include the NMCPs (nuclear matrix constituent proteins), also known as LINC (little nuclei) and CRWN (crowded nuclei) in Arabidopsis (Masuda *et al.*, 1993; Ciska *et al.*, 2013; Ciska and Moreno Díaz de la Espina, 2013; Sakamoto and Takagi, 2013). They have multiple coiled-coil domains, form filamentous dimers and function in control of nuclear size, shape, and heterochromatin organisation (Dittmer *et al.*, 2007; van Zanten *et al.*, 2011, 2012; Sakamoto and Takagi, 2013; Wang *et al.*, 2013). Recently, interactions between AtCRWN1 and AtSUN1 and AtSUN2 have been suggested, arguing in favour of an association between the plant lamina and the nuclear envelope (Graumann, 2014). AtKAKU4, a putative lamina component, has also been shown to be localised at the inner nuclear membrane and interacts with AtCRWN1 and AtCRWN4 (Goto *et al.*, 2014). AtKAKU4 has been shown to affect nuclear shape and size.

In this study we describe members of a higher plant-specific family of nuclear-localised coiled-coil proteins that interact with SUN domain proteins at the nuclear periphery, and suggest a role as putative bridges between the NE and chromatin.

Materials and methods

Seed stocks, plant growth, and T-DNA mutants

All *A. thaliana* transfer (T)-DNA insertion lines were ordered from the Nottingham Arabidopsis Stock Centre (NASC, UK) or the Arabadopsis Biological Resource Centre (ABRC; Ohio), with the exception of the GABI-kat lines which were ordered from Bielefeld University (Germany). T-DNA lines were of the ecotype Col-0 and were established as homozygous lines. Seeds were germinated as described in Graumann *et al.* (2014) and grown in long-day conditions (16 h light, 8 h dark at 18 °C). Genotyping PCR was used for identification of homozygous T-DNA insertion lines and semi-quantitative RT-PCR to confirm absence of the corresponding mRNA (Supplementary Table S1 at *JXB* online).

The SAIL_846_B07 homozygous line (*Atneap1*) was crossed with WiscDsLoxHS086_02C (*Atneap3*), and their *Atneap1 Atneap3* double-heterozygous offspring were allowed to self-pollinate. Their seeds were collected and 24 seedlings were screened, several *Atneap1 Atneap3* homozygous mutant plants were identified and their progeny were phenotyped (Supplementary Table S1).

Membrane yeast two-hybrid system

The split-ubiquitin-based membrane yeast two-hybrid (MYTH) system (Snider *et al.*, 2010a, b) was used essentially as described by Graumann *et al.* (2014), using the same bait and prey purchased from DualsystemsBiotech (http://www.dualsystems.com). Prey constructs were cloned in the pPR3N (2μ, TRP1, AmpR) vector and bait constructs were cloned in the pBT3N (CEN, LEU2, KanR) vector. *AtNEAP* cDNA were fused to chimeric primers having 35 base pairs complementary to the linearised bait or prey plasmid on the 5′ ends, and 18 base pairs complementary to the N-terminus of *AtNEAP* cDNA on the 3′ end. *AtNEAP* cDNA were cloned in plasmids by 'gap-repair' homologous recombination in yeast (Oldenburg *et al.*, 1997). After digestion by SfI1, prey or bait plasmids and cDNA were co-transformed into yeast in the 1:3 vector:insert ratio, and successfully transformed clones were selected on test medium. Clones were then subjected to colony PCR, followed by extraction of the plasmid DNA and sequencing. *AtNEAP* containing bait vectors were verified for self-activation and only *AtNEAP1* and *AtNEAP2* baits that did not self-activate were used. Bait and prey vectors were allowed to co-transform in yeast. Presence of interaction was analysed by yeast growth on test medium (TM: YNB without Leu, Trp, Ade, and His) at 30 °C for more than 48 h. The controls were grown on permissive medium (PM: YNB without Leu and Trp) in identical conditions as the test medium. Clones were verified by colony PCR. The *A. thaliana* cDNA library containing 3.6 million fragments (DualsystemsBiotech) cloned into the prey vector pDSL-Nx (2μ, TRP1, AmpR) was screened for novel interactors using the *AtNEAP1* bait. The library consisted of cDNA from 6-d-old etiolated seedlings as well as seedlings exposed to blue and far-red light. A positive control prey included the yeast ER resident protein Ost1 fused to the Nub portion of yeast ubiquitin in the pOst1–NubI (2μ, TRP1, AmpR) vector. Transformants from the screen were allowed to grow on highly restrictive medium (YNB without Leu, Trp, His, Ade), and as a back-up on low-stringency restrictive medium (YNB without Leu, Trp, His). Plasmid DNA was extracted from yeast colonies that grew on highly restrictive medium and sent for sequencing.

Phylogenetic reconstruction and evolution rate

AtNEAP coding sequences were used for phylogenic reconstruction and substitution rate calculation. Selected protein sequences were aligned with MUSCLE multiple sequence alignment (http://www.ebi.ac.uk/Tools/msa/muscle) and maximum-likelihood analysis was performed with FastTree (http://www.microbesonline.org/fasttree) using default parameters. ω (the ratio of non-synonymous/synonymous substitution rates) was determined using Codeml from the PaML package (Yang, 2007).

RNA-seq data mining

Already published RNA-seq datasets from the wild-type Col-0 ecotype were used in order to monitor the expression of *AtNEAP*s. The Illumina RNA-seq data are available at the NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi) under accession numbers SRR1463325 and SRR1463326 for epidermal cells from 10-d-old cotyledons, SRR1042766 and SRR1042767 for primary roots from 7-d-old seedlings, and SRR826283 from 10-d-old seedlings for guard cells. Reads from RNA-seq libraries were mapped onto the candidate gene sequences allowing no mismatches using TOPHAt v. 2.0.14 (Kim *et al.*, 2013), using standard settings and maximum of multi-hits set at 1, minimum intron length set at 15 bp, and maximum intron length set as 6000 bp. Reads were summed up for each gene using HTseq-count with the overlap resolution mode set as intersection non-empty and with no strand-specific protocol (Anders *et al.*, 2015). Transcription levels were normalised to *SAND* as for RT-qPCR and expressed in reads per kilobase of exon model (RPKM) per million mapped reads.

Domain prediction

Coiled-coil domains were predicted using SMART COILS, PairCoil2 and Marcoil (Lupas *et al*, 1991; Dolerenzi and Speed, 2002; McDonnell *et al.*, 2006, Letunic *et al.*, 2012). Nuclear

localization signals (NLSs) were predicted using cNLS mapper and NLStradamus (Kosugi *et al.*, 2009; Nguyen Ba *et al.*, 2009). TM domains were predicted using ARAMEMNON and DAS (Cserzo *et al.*, 1997; Schwacke *et al.*, 2003).

Cloning and fluorescent protein fusions

The coding sequences of *AtNEAP1*, *AtNEAP2*, *AtNEAP3*, *AtbZIP18*, and domain deletion mutants were amplified using the gene-specific primers listed in Supplementary Table S2. Gateway *attB* flanking sequences were added to each of the constructs and gateway was technology used for cloning in pDONR207, and afterwards into expression vectors pCAMBIA 1300, pK7CWG2, and pK7WGC2, as described by Graumann *et al.* (2014). The primers used to generate the domain deletions *AtNEAP3ΔCC1* (aa13-93 deleted), *AtNEAP3ΔCC2* (aa124-185 deleted), *AtNEAP3ΔNLS* (aa239-264 deleted), and *AtNEAP3ΔTM* (aa314-333 deleted) are listed in Supplementary Table S2. Supplementary Table S3 lists all expression vectors created in this study.

Transient expression and confocal microscopy

Leaves of 5–6-week-old *Nicotiana benthamiana* were infiltrated with agrobacteria carrying expression vectors for transient expression, as described by Sparkes *et al.* (2006) and Graumann *et al.* (2014). All AtNEAP-fluorescent protein (FP) fusions were infiltrated at an OD of 0.1 and SUN-FP fusions were infiltrated at an OD of 0.03 together with p19 at an OD of 0.05. Tissues were imaged using a Zeiss (Welwyn Garden City, UK) LSM 510 META or an inverted LSM 510 confocal laser scanning microscope fitted with 40×, 63×, and 100× oil immersion objectives.

Fluorescence recovery after photobleaching (FRAP) was used to investigate the mobility of AtNEAP 1–3 fluorescent protein fusions at the NE. FRAP was carried out as described by Graumann *et al.* (2007). Briefly, transiently expressing *N. benthamiana* lower epidermal leaf cells were treated with Latrunculin B to immobilise the nucleus and then imaged with a 514-nm laser to excite the yellow fluorescent protein (YFP). Scanning transmission was kept low and bleaching was performed at 100% transmission. The fluorescence was monitored in a constant-sized region of interest (ROI) pre- and post-bleach. The raw data were converted to percentages and mobile fractions, and half-times were calculated as described by Graumann *et al.* (2007, 2010). Students *t*-test was used for statistical analysis; 30 nuclei per sample were photobleached.

In planta *protein interaction studies*

Acceptor photobleaching fluorescence resonance energy transfer (apFRET) was used to detect *in planta* protein interactions. ApFRET was performed as described in Graumann *et al.* (2010) and Graumann (2014). Briefly, transiently expressing *N. benthamiana* leaf tissue was imaged as described in the previous section. YFP was excited with 514-nm light and cyan fluorescent protein (CFP) with 458-nm light. The YFP laser transmission was kept low during scanning to avoid photobleaching but was set at 100% during the bleach. Five pre-bleach and five post-bleach scans were carried out in a constant-sized ROI. Fluorescence intensity was monitored in the ROI and analysed using Microsoft Excel. For each sample, approximately 35 nuclei were used. Student's *t*-test was carried out for statistical analysis. FRET efficiency is given as percentage CFP fluorescence increase, expressed as mean ± standard error of mean (SEM) compared to a non-bleached control region.

Western blotting

Total protein was extracted from infiltrated and non-infiltrated *N. benthamiana* leaves. Leaf material was immediately frozen in liquid nitrogen and ground with a mortar and pestle. Ground material was collected in liquid nitrogen-cooled 15-ml tubes, to which 1 ml of protein extraction buffer (100 mM Tris, pH 6.8, 4.5 M urea, 1 M thiourea, 2% CHAPS, 0.5% Triton X-100, 10mM DTT, 1% Sigma protease inhibitor cocktail, 0.52 µl ml^{-1} benzonase, 0.2 mM phenylmethylsulfonyl fluoride) was added. Protein was precipitated using ice-cold acetone, protein extract, and trichloroacetic acid (8:1:1) and centrifuged for 15 min at 10 000 *g*. The pellet was washed 2× with ice-cold acetone and dried before suspension in 100 µl of 1× SDS buffer containing DTT and 8 M urea. The sample was separated on an 8% SDS-PAGE gel and transferred to an Immobilon-P PVDF membrane (Millipore, Livingstone, UK) before blocking with 5% milk PBST and detection with Abcam (Cambridge, UK) rabbit GFP antibody diluted 1:3000 in 5% milk PBST at 4 °C. Detection was with a goat anti-rabbit Cy5 conjugated antibody (Jackson Immunoresearch, Newmarket, UK) and imaged using a Bio-Rad ChemiDoc^{tm} imaging system.

Phenotype of AtNEAP1*,* AtNEAP3*, and* AtNEAP1/3 *mutants*

T-DNA lines were obtained for *AtNEAP1* (SAIL_846_B07, NASC number CS837770) and *AtNEAP3* (WiscDsLoxHs086_02C). For general observation of phenotype, seeds were germinated and grown in 16 h light at 21 °C and 8 h dark at 18 °C. Seedlings at 10–12 d old were transplanted to Levington F2S compost mixed with perlite in 5 × 5-cm pots. Wild-type and mutant plants were grown simultaneously in controlled conditions, and germination efficiency, plant vigour, and fertility were carefully observed. For root growth analysis, seedlings were grown on half-strength MS agar on vertical plates and scanned on days 3, 7, 10, and 14 after germination. Nuclear morphology and chromatin organisation was determined with NucleusJ, as described by Poulet *et al.* (2015).

Results

NEAPs in the plant kingdom

Four members of a family of proteins, designated AtNEAP for *Arabidopsis thaliana* nuclear envelope-associated protein (At3g05830, AtNEAP1; At5g26770, AtNEAP2; At1g09470, AtNEAP3; and At1g09483, AtNEAP4) were initially identified in a bioinformatics screen searching for the presence of coiled-coil domains, a nuclear localisation signal, and a C-terminal hydrophobic domain (based on a previous description of AtNEAP1 by Lu, 2011; Supplementary Fig. S1). Among this small protein family, AtNEAP1 has been annotated as a nuclear intermediate filament (IF) -like protein in the UniProtKB/TrEMBL database.

The NEAPs identified in *A. thaliana* are members of the gene family HOM03D003386 (PLAZA 3.0, http://bioinformatics.psb.ugent.be/plaza/versions/plaza_v3_dicots/), with 54 members for which no function is assigned. Additional analysis of the gene family reveals representatives in the magnoliophyta and gymnosperms (Supplementary Fig. S2). The cladogram is organised with AtNEAP1, 2, 3, and 4 forming adjacent sub-branches clustering with other crucifer homologues (*Brassica rapa*, *Capsella rubella*, *Arabidopsis lyrata*, and *Thellungiella parvula*); monocot NEAPs group together in three sub-branches. The gymnosperm *Picea abies* has two representatives while the basal angiosperm *Amborella trichopoda* has a single representative (AtR_00045G00720).

Three members of the family, AtNEAP1–3 display similar size (349, 335, and 336 amino acids, respectively) while AtNEAP4 is smaller (112 amino acids; Fig. 1). AtNEAP4 shares the highest sequence homology with the C-terminus

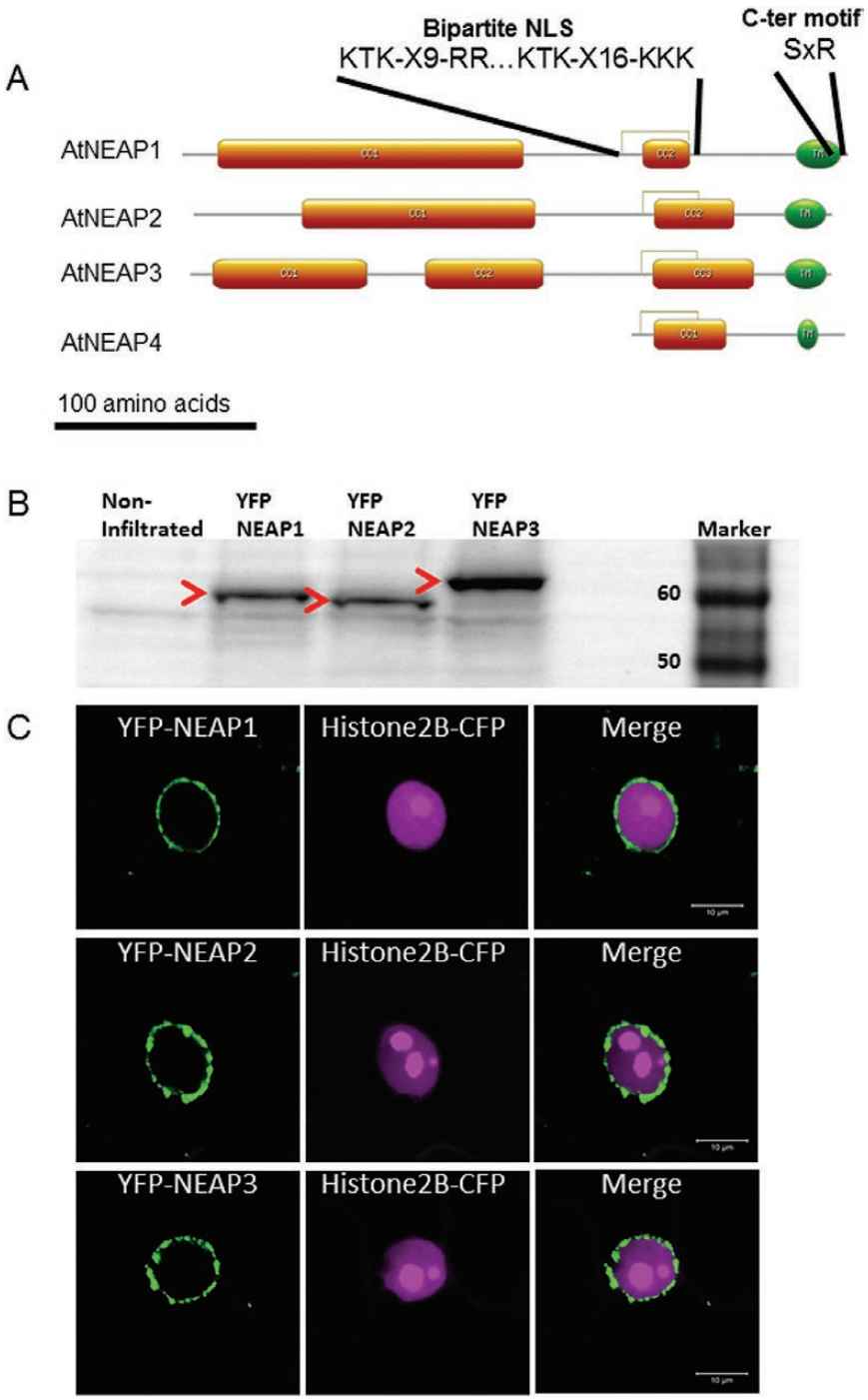


Fig. 1. AtNEAP structure and expression. (A) Schematic representation of AtNEAP1, 2, 3, and 4 protein organisation, showing coiled-coil domains (orange rectangles), NLS (grey boxes), and transmembrane domain (green ovals). The sequence and position of the bipartite NLS and conserved C-terminal motif are indicated. (B) Western blot of protein extracts from *N. benthamiana* leaves transiently expressing YFP-AtNEAP1-3 in the presence of p19, resolved on an 8% SDS-PAGE gel and detected with an anti-YFP antibody. Red arrows highlight the NEAP bands. YFP-AtNEAP1 and YFP-AtNEAP2 have a relative molecular mass of approximately 60 kDa while YFP-AtNEAP3 is approximately 65 kDa. The net relative molecular masses of AtNEAP1, 2, and 3 were approximately 34, 33 and 38k Da, respectively. (C) Confocal micrographs showing N-terminal YFP fusions of AtNEAP proteins (green) and histone H2B-CFP (magenta) transiently expressed in *N. benthamiana* leaf epidermal cells in the presence of p19. All three NEAPs localize at the nuclear periphery surrounding chromatin, which is labelled by histone H2B-CFP. Scale bars = 10 μm.

of AtNEAP3 and may be a truncated gene duplication of the common ancestor of *AtNEAP3* and *4*. This hypothesis is strongly supported by the fact that AtNEAP4 is associated with the AtNEAP3 sub-branch in the phylogenetic data (Supplementary Fig. S2). Analysis of the expression levels (Supplementary Fig. S3) and evolution rates of the AtNEAPs using the PAML software (Yang, 2007) suggests that AtNEAP4 is a pseudogene, as analysis of the *AtNEAP* orthologous show an increase of the accumulation of non-synonymous mutations in *AtNEAP4* (Supplementary Fig. S3). Therefore, this study was focused on AtNEAP1, AtNEAP2, and AtNEAP3. AtNEAPs 1–3 show a

characteristic domain structure (Fig. 1A, Supplementary Fig. S1), with a variable, long N-terminal domain containing two or three coiled-coils (aa54–184 and aa221–266 in AtNEAP1; aa54–185 and aa220–298 in AtNEAP2; aa13–93, aa124–185 and aa220–306 in AtNEAP3) predicted using SMART, COILS, PairCoil2, and Marcoil (Lupas *et al.*, 1991; Delorenzi and Speed, 2002; McDonnell *et al.*, 2006; Letunic *et al.*, 2012) and a conserved C-terminus comprising an NLS, and hydrophobic domain close to the C-terminus (Fig. 1A). The bipartite NLS predicted by cNLS mapper and NLStradamus (Kosugi *et al.*, 2009; Nguyen Ba *et al.*, 2009) is located at the N-terminus of the coiled-coil domain nearest to the C-terminus (KTK-X9-RR and KTK-X16-KKK; aa239–264 in AtNEAP1, aa238–263 in AtNEAP2, and aa239–264 in AtNEAP3; Fig. 1A and Supplementary Fig. S1). The C-terminals of AtNEAPs 1–3 end in a characteristic motif ending in the hydrophobic domain (aa324–345 in AtNEAP1; aa311–331 in AtNEAP2; and aa314–333 in AtNEAP3) followed by SxR, where x is K (AtNEAP1, AtNEAP2) or R (AtNEAP3). In monocots this is typically xKR, where x is either A or T. The TM domains of AtNEAP1–3 show a high level of sequence homology (Supplementary Fig. S1).

Western blot analysis of YFP-AtNEAP1, YFP-AtNEAP2, and YFP-AtNEAP3 proteins expressed transiently in *N. benthamiana,* indicates that the relative molecular masses of YFP-AtNEAP1 and YFP-AtNEAP2 were approximately 60 kDa, while YFP-AtNEAP3 was larger at 65 kDa, giving AtNEAP1 and AtNEAP2 a relative molecular mass of 34 kDa and AtNEAP3 a mass of 38 kDa, smaller than their predicted masses of 41, 38, and 39 kDa (Fig. 1B). Extraction of AtNEAPs required the presence of a high concentration of urea and of detergent (CHAPS and Triton x-100), indicating that the NEAPs are highly insoluble and this may explain the aberrant molecular mass obtained.

Expression data gained from Genevestigator (Toufighi *et al.*, 2005) and from RNA-seq data mining reveals that *AtNEAP1* and *AtNEAP2* are expressed at medium levels in most tissues, including the primary root, leaf epidermis, and guard cells. *AtNEAP3* is expressed at low levels in the leaf epidermis and guard cells but at higher levels in the primary root (Supplementary Fig. S4).

NEAPs localise to the nuclear periphery

Localisation of the NEAP family to the nucleoplasm or inner nuclear envelope was suggested by the presence of a bipartite NLS (Fig. 1) and confirmed using fluorescent protein fusions in transient expression in *N. benthamiana* leaves. YFP-NEAPs 1–3 localise to the nuclear periphery, surrounding chromatin labelled with histone H2B-CFP (Fig. 1C). Transiently expressed YFP-AtNEAP1–3 were also used to study the mobility of the proteins at the NE by FRAP in *N. benthamiana* leaves (Fig. 2). YFP-AtNEAP1 and YFP-AtNEAP2 have significantly lower (P<0.001) mobile fractions (20.6 ± 1.8% and 17.7 ± 1.5%, respectively) compared to YFP-AtNEAP3 (46.9 ± 5.3%; Fig. 2). Similarly, the half-time is significantly higher (P<0.05) for YFP-NEAP3 (9.5 ± 3.5 s) then YFP-AtNEAP1 and YFP-AtNEAP2 (3.6 ± 0.17 and 2.3 ± 2.4 s, respectively). The significant differences in mobility

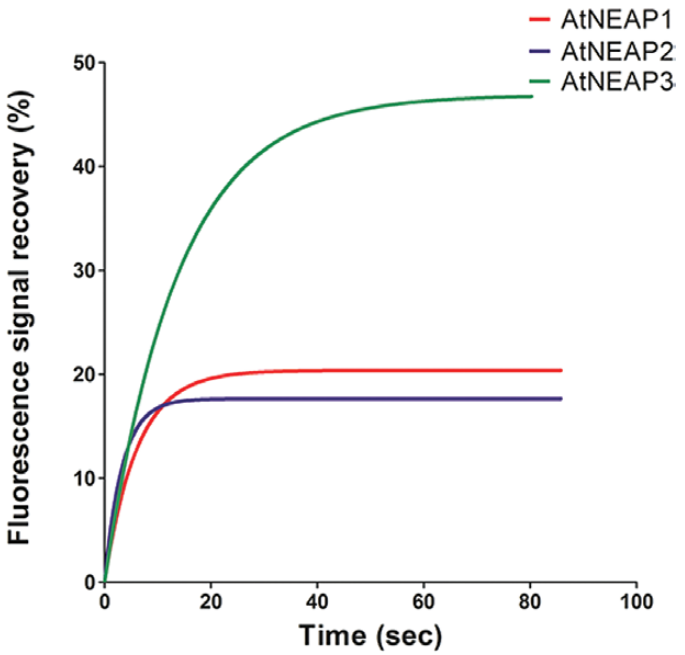


Fig. 2. Fluorescence recovery curves of YFP-AtNEAP1 (red), YFP-AtNEAP2 (blue), and YFP-AtNEAP3 (green) obtained after photobleaching *in planta*. Time zero denotes the time of bleaching. Results of an unpaired *t*-test showed that the maximum fluorescence recovery of YFP-AtNEAP3 was significantly (P<0.0001) higher than YFP-AtNEAP1 and YFP-AtNEAP2.

of AtNEAP3 suggest that binding interactions differ between the AtNEAP homologues, with AtNEAP1 and AtNEAP2 being most tightly bound. The mobile fractions of AtNEAP1 and AtNEAP2 are comparable to other NE proteins, for example the AtSUNs and NE-associated proteins such as AtCRWN1 (Graumann 2014; Graumann *et al.*, 2014).

Domain function was studied using deletion and truncation mutants of AtNEAP3 (Fig. 3). Deletion of the first coiled-coil domain (aa13–93; YFP-AtNEAP3ΔCC1) resulted in nucleoplasmic fluorescence, while deletion of the second coiled-coil domain (aa124–185; YFP-AtNEAP3ΔCC2) had no effect on localisation (Fig. 3A, B). Deletion of the NLS (aa239–264; YFP-AtNEAP3ΔNLS) resulted in cytoplasmic fluorescence (Fig. 3A, B). Finally, deletion of the predicted TM domain (aa314–333; YFP-AtNEAP3ΔTM) resulted in nucleoplasmic fluorescence (Fig. 3A, B). The presence of CC1 and the TM domain are therefore important in localising the protein from the nucleoplasm to the nuclear periphery, while the presence of the NLS is required to target the protein to the nucleus from the cytoplasm. Interestingly, co-expression of the domain deletions with full-length CFP-AtNEAP3 resulted in co-localisation at the NE (Fig. 3C). This suggests that NE-localised CFP-AtNEAP3 can interact with all four domain deletion mutants and 'rescue' them to the NE.

AtNEAP proteins interact to form homomers and heteromers

The effect on localisation of the AtNEAP3 domain deletion mutants upon co-expression with full-length AtNEAP3 suggests the possibility that AtNEAP3 is able to interact with itself. To test interactions between the AtNEAPs, apFRET

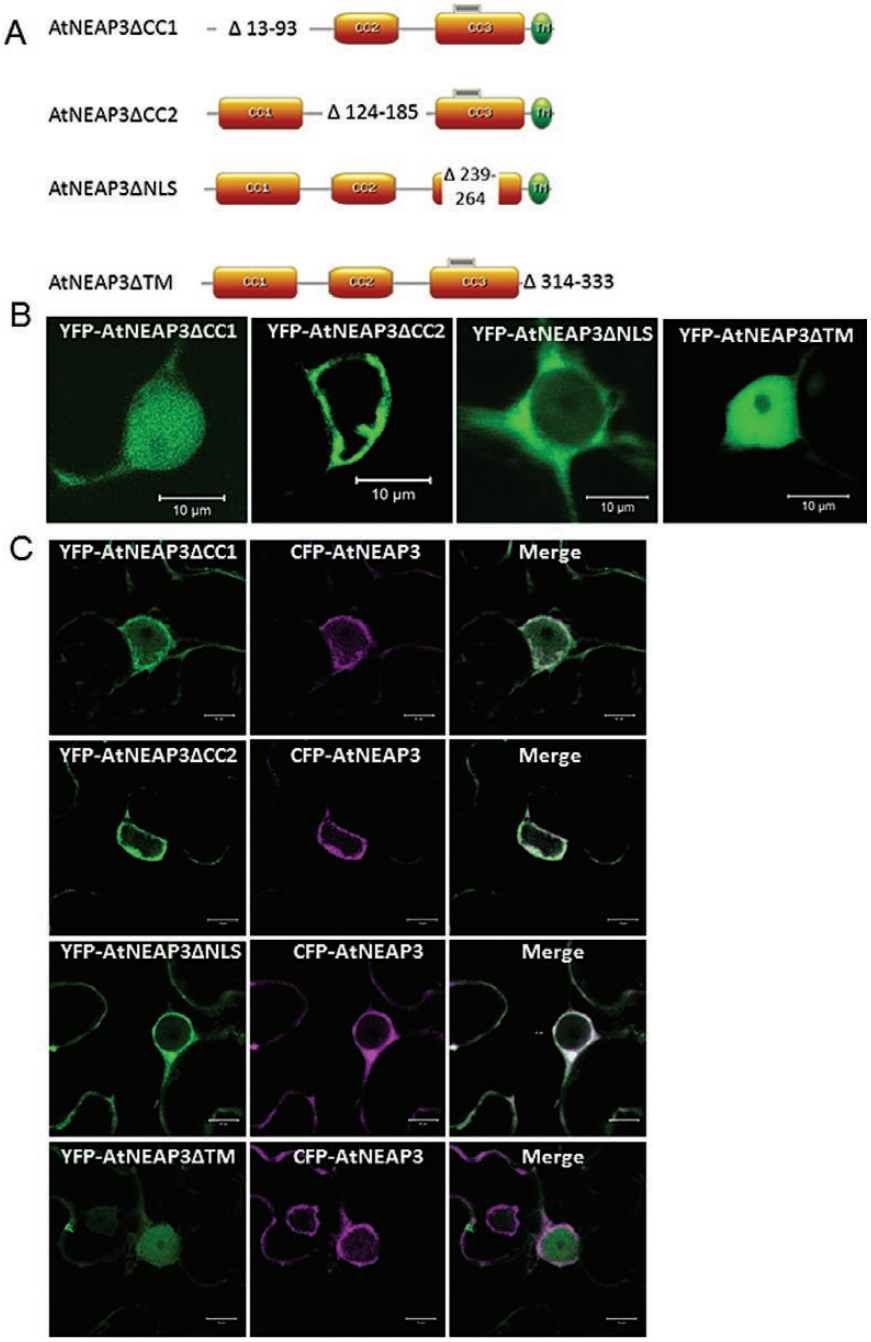


Fig. 3. Domain deletion mutants of AtNEAP3. (A) Schematic representation of domain deletion constructs AtNEAP3ΔCC1, AtNEAP3ΔCC2, AtNEAP3ΔNLS, and AtNEAP3ΔTM highlighting which amino acids are not present in the respective constructs. (B, C) Domain deletion constructs were fused to YFP at the N-terminus and transiently expressed in *N. benthamiana* leaf epidermal cells in the presence of p19. (B) Confocal micrographs of single expression showing nuclear localisation of YFP-AtNEAP3ΔCC1 and YFP-AtNEAP3ΔTM, cytoplasmic localisation of YFP-AtNEAP3ΔNLS, and nuclear rim localisation of YFP-AtNEAP3ΔCC2. (C) Confocal micrographs of full-length CFP-AtNEAP3 co-expressed with the domain deletion mutants showing that mutant localisation is rescued to the NE. Scale bars = 10μm.

and MYTH were used. Firstly, YFP- and CFP-fusions of the NEAPs were co-expressed transiently to show that all AtNEAPs co-localised at the nuclear periphery (Fig. 4A).

The co-localisation of the AtNEAPs was used to measure apFRET efficiency (E_F; Fig. 4B). There was no significant increase (P>0.1) in AtNEAP1-CFP fluorescence post-YFP-AtNEAP1 bleach, indicating that AtNEAP1 does not interact with itself in this system (Fig. 4B). However, both AtNEAP2 and AtNEAP3 showed a significant (P<0.001) interaction with themselves with AtNEAP2-CFP (E_F=21.3±1.7%) and AtNEAP3-CFP (E_F=18.4±1.9%), respectively (Fig. 4B). Furthermore, bleaching YFP-AtNEAP3 also led to a significant (P<0.001) increase in fluorescence of co-expressed AtNEAP1-CFP and AtNEAP2-CFP with calculated E_F values of 16.6±1.5% and 18.6±1.4%, respectively (Fig. 4B). Bleaching YFP-AtNEAP1 also led to a significant (P<0.001) increase in co-expressed fluorescence of AtNEAP2-CFP (E_F=10.2±1.1%; Fig. 4B). Thus AtNEAP1, AtNEAP2, and AtNEAP3 interact with each other *in planta*, although AtNEAP1 does not strongly self-interact.

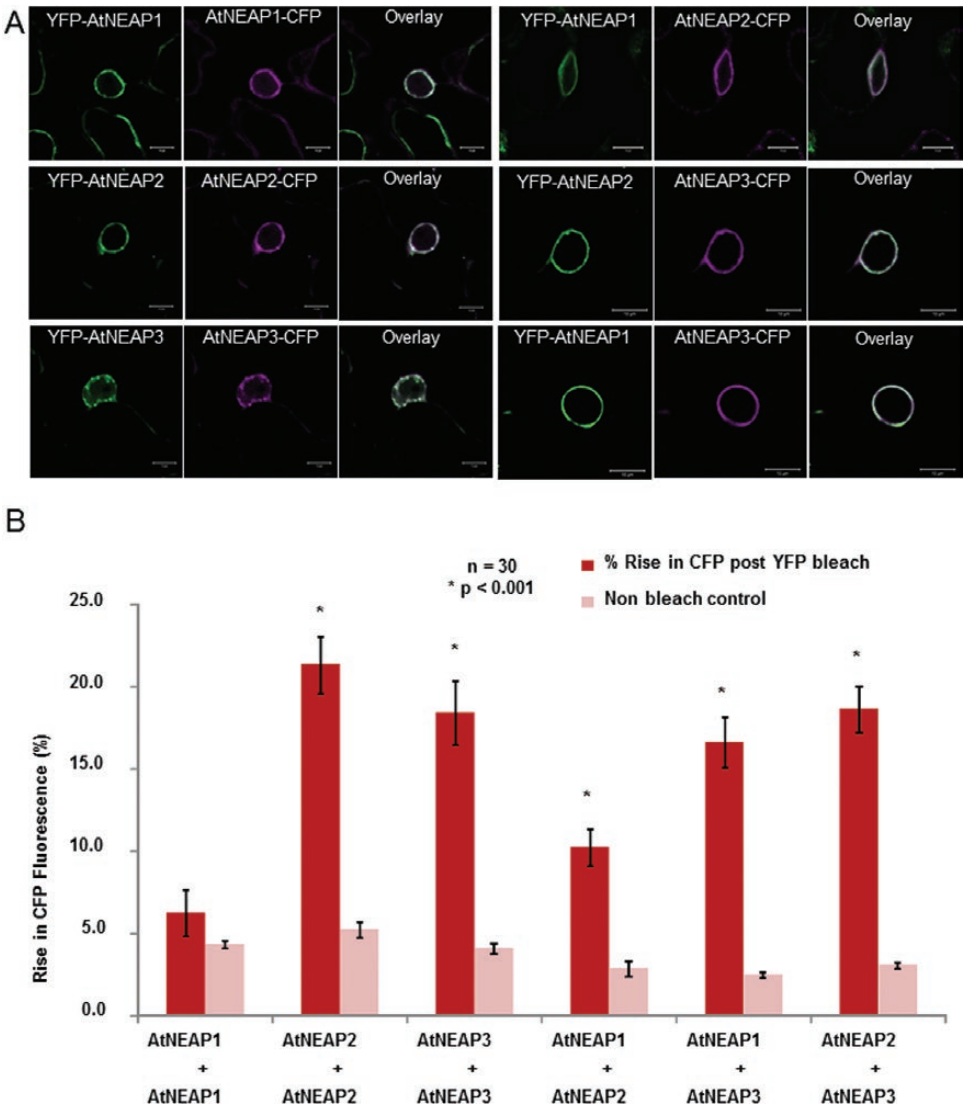


Fig. 4. Interactions between AtNEAPs as measured by apFRET. (A) Confocal micrographs of transiently co-expressed YFP- and CFP-AtNEAPs demonstrating co-localisation at the nuclear periphery. Scale bars = 10μm. (B) apFRET of co-localised AtNEAPs; efficiency (E_F) measured as changes in CFP fluorescence in a bleached (red) vs non-bleached (pink) region of YFP fluorescence. A significant increase in CFP fluorescence indicates interaction *in planta*. In each case the upper partner is YFP-NEAP while the lower partner is CFP-NEAP. Values are percentage mean ± standard error of the mean and compared to an unbleached control region (n=30). A paired t-test was performed between the bleached and non-bleached populations (*P<0.001).

As apFRET relies on transient expression of proteins, we also employed a membrane yeast two hybrid (MYTH) system to confirm the NEAP–NEAP interactions identified by apFRET. Two bait vectors containing *AtNEAP1* and *AtNEAP2* were used. *AtNEAP3* was discarded as it activates detection in the absence of prey. When yeast containing the AtNEAP1 bait was transformed with AtNEAP1, AtNEAP2, and AtNEAP3 prey vectors, all transformations successfully yielded colonies on restrictive medium (Supplementary Fig. S5A), confirming the AtNEAP1–AtNEAP2 and AtNEAP1–AtNEAP3 interactions identified *in vivo*.

AtNEAP proteins interact with other nuclear envelope proteins

The mobility studies indicated that YFP-AtNEAP1 and YFP-AtNEAP2 have similar mobile fractions to the SUN domain proteins. As the SUN proteins are a well-characterised group of NE proteins and part of the nucleo-cytoskeletal bridging complexes, we wanted to explore the possibility as to whether the AtNEAPs can associate with AtSUNs. For this, combinations of N-terminal YFP fusions of AtSUN1 and AtSUN2 were co-expressed with N-terminal CFP fusions of AtNEAP1, AtNEAP2, and AtNEAP3 in *N. benthamiana* leaves, which revealed that AtNEAPs co-localise with AtSUN1 and AtSUN2 at the NE (Fig. 5A).

In planta interactions between co-localised CFP-NEAPs and YFP-SUNs were tested using apFRET. Bleaching YFP-AtSUN1 led to a significant (P<0.005) increase in fluorescence of co-expressed CFP-AtNEAP1, CFP-AtNEAP2, and CFP-AtNEAP3 with average values of E_F of 6.9 ± 0.7%, 7.8 ± 0.7%, and 3.9 ± 0.4%, respectively (Fig. 5B). Similarly, bleaching YFP-AtSUN2 led to a significant (P<0.0001) increase in fluorescence of co-expressed CFP-AtNEAP1, CFP-AtNEAP2, and CFP-AtNEAP3 with average E_F values of 18.4 ± 1.4%, 14.4 ± 0.9%, and 26.9 ± 1.9%, respectively.

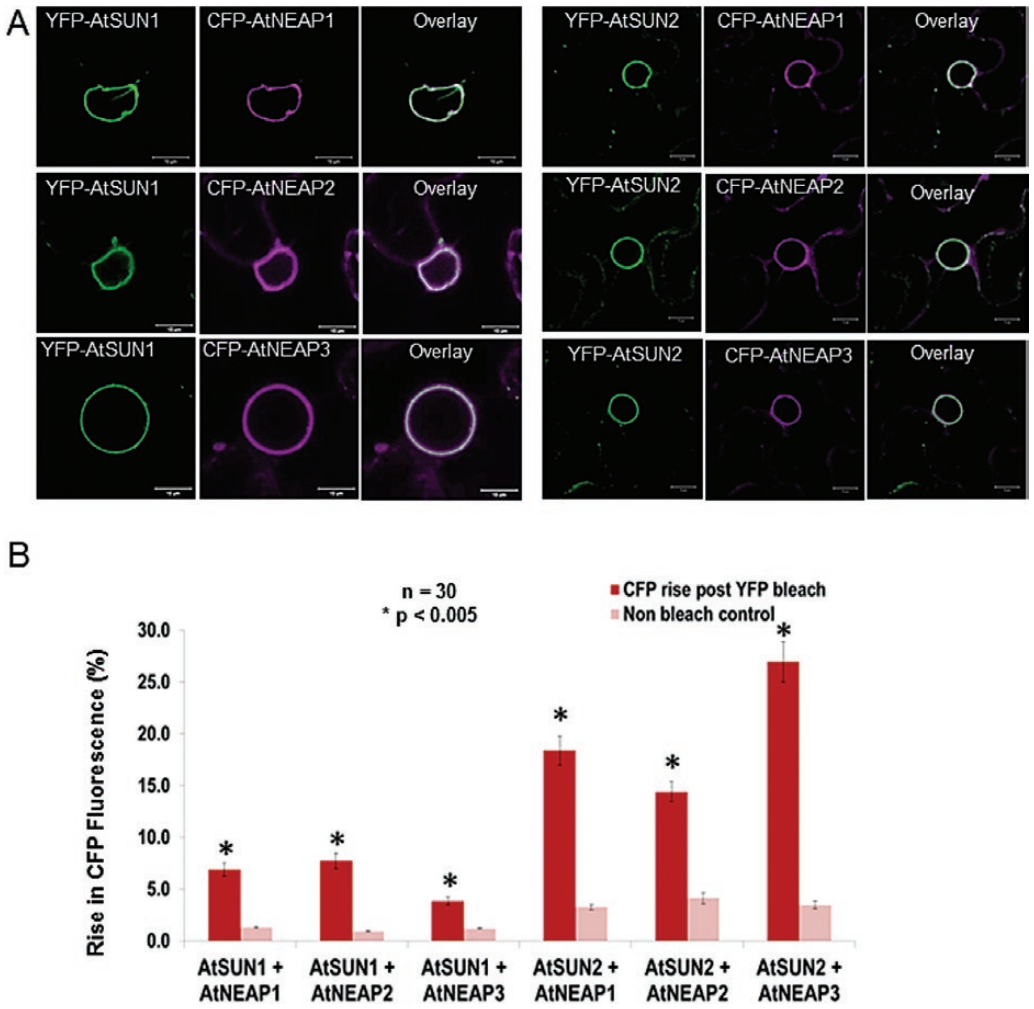


Fig. 5. *In planta* interactions between AtSUNs and AtNEAPs. (A) Confocal micrographs showing the co-localisation of N-terminal YFP (green) fusions of AtSUN1 and AtSUN2 with N-terminal CFP (magenta) fusions of AtNEAP1, AtNEAP2, and AtNEAP3 expressed transiently in the presence of p19. Scale bars = 10 μm. AtSUNs and AtNEAPs co-localised at the NE. (B) apFRET of co-localised AtSUNs and AtNEAPs; changes in efficiency (E_F), measured as CFP fluorescence, in a region of bleached (red) vs a control non-bleached (pink) region of YFP fluorescence. A significant increase in CFP fluorescence indicates interaction *in planta*. In each case, the upper partner is a N-terminal YFP-AtSUN construct, while the lower partner is a N-terminal CFP-NEAP construct. Values are expressed as percentage mean ± SEM (n=30). A paired t-test was performed between the bleached and non-bleached populations (*P<0.005).

This shows that all three NEAPs can interact with AtSUN1 and AtSUN2 *in planta*.

Interactions between SUNs and NEAPs were also confirmed using MYTH. Yeast containing AtNEAP1 and AtNEAP2 bait were transformed with AtSUN1 and AtSUN2. The growth of colonies on restrictive medium confirmed the interactions of AtNEAP2 with AtSUN1 and AtSUN2, but an interaction was not detected with AtNEAP1 (Supplementary Fig. S5B). The ability of NEAPs and SUNs to interact with each other indicates that AtNEAPs may also be associated with nucleo-cytoskeletal bridging complexes in plants.

AtNEAP1 interacts with a transcription factor

The MYTH assay was also employed to screen the *A. thaliana* cDNA library for novel AtNEAP1 interaction partners. Briefly, 3.6 million cDNA fragments were screened for interactors of AtNEAP1 bait, and 25 colonies were selected and sent for sequencing. Nine of the 25 colonies sequenced returned a single gene, *At2g40620*, a basic-leucine zipper (*AtbZIP28*) transcription factor. In order to confirm its nuclear localisation, fluorescent protein fusion of the coding sequence of *AtbZIP18* under the CaMV 35S promoter was expressed transiently in *N. benthamiana*. YFP-AtbZIP18 was localised to the nucleoplasm and cytoplasm (Fig. 6A). When co-expressed with YFP-AtbZIP18, CFP-AtNEAP1 failed to accumulate at the nuclear periphery and was found to co-localise with the YFP-bZIP18 in the nucleoplasm (Fig. 6B). The nucleoplasmic co-localisation with YFP-bZIP18 was also seen with AtNEAP2 and AtNEAP3 (data not shown). AtbZIP18 is therefore a potential *in vivo* interaction partner for the AtNEAPs.

Functional analysis of NEAPs

To investigate putative functions of the AtNEAPs, we used T-DNA knockout lines and focused on the tissues where

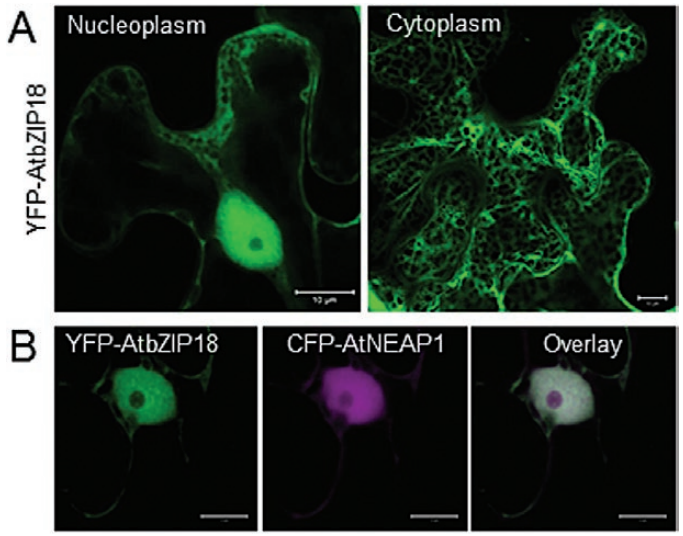


Fig. 6. Subcellular localisation of AtbZIP18. Confocal micrographs showing YFP-AtbZIP18 expressed transiently in *N. benthamiana* leaves in the presence of p19, (A) during single expression, showing the protein is localised in the nucleoplasm and cytoplasm; and (B) co-expression with CFP-AtNEAP1, showing the two proteins co-localised in the nucleoplasm. Scale bars = 10μm.

AtNEAP1-3 appeared to be expressed at higher levels, namely the primary root, leaf epidermis, and guard cells (Supplementary Fig. S4). All single and the double *neap* mutant lines germinated normally (95–100% germination). No significant difference was observed in root growth in the single insertion lines but, in contrast, the *atneap1 atneap3* double-knockout showed significantly reduced primary root growth from day 2 to day 8 post-germination (Fig. 7A–C). Nuclear morphology and chromatin organisation of pavement cells and guard cells for the cotyledon epidermis were examined for the *atneap1*, *atneap3*, and *atneap1 atneap3* mutants. Nuclear volume appeared increased in the pavement cells of all three mutants (Fig. 7D), And chromocentre volume was decreased in all mutants in both pavement and guard cells (Fig. 7E). In addition, the *atneap3* single-mutant also had a reduced relative heterochromatin fraction in both cell types (Fig. 7F), while the number of chromocentres appeared to be increased in pavement cells (Fig. 7G). The latter indicates that chromocentre organisation is disrupted in the *atneap3* single-mutant with smaller but more numerous chromocentres, indicating some impact on nuclear organisation. RT-PCR showed that both single-mutants were complete knockout mutants (Fig. 7H).

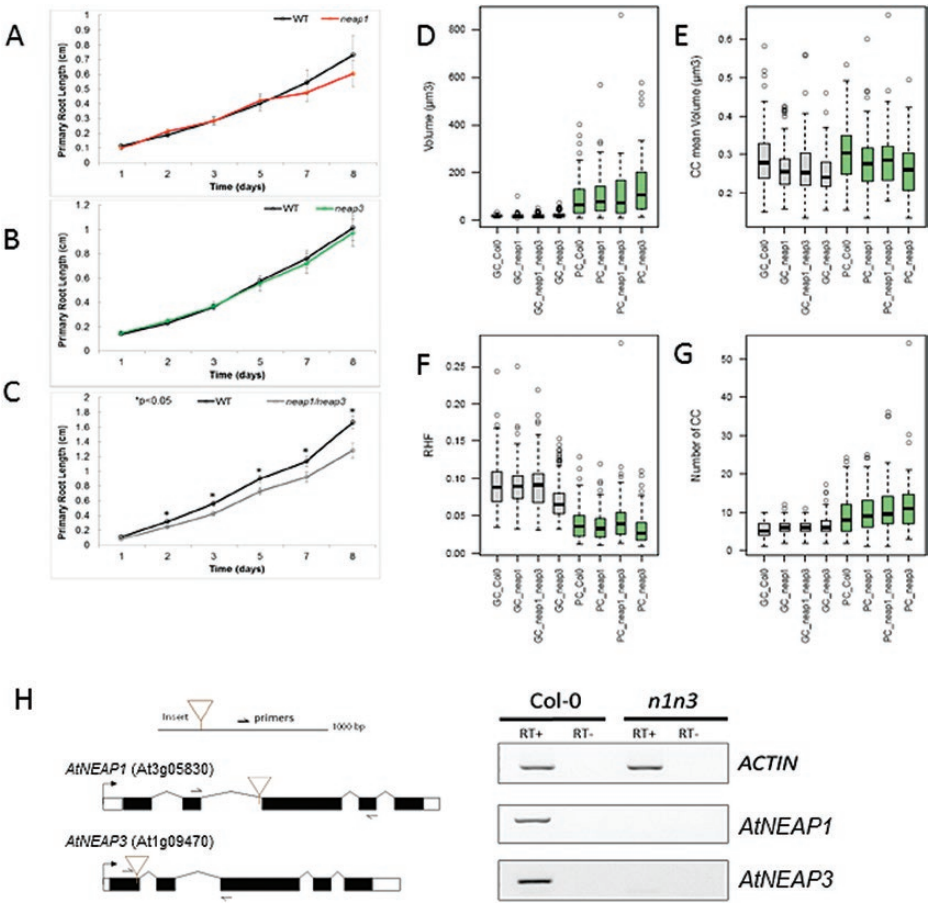


Fig. 7. Analysis of AtNEAP1 and AtNEAP3 T-DNA knockout mutants. (A–C) Root growth assays comparing primary root length of (A) *neap1*, (B) *neap3*, and (C) *neap1/3* double-mutant lines versus the wild-type Col-0 in 1–8-d-old seedlings. Values are shown as mean ± standard error of the mean, and an unpaired *t*-test was performed where $*P<0.05$ was statistically significant ($n=30$). (D–G) Box plots showing the nuclear volume (D), chromocentre (CC) volume (E), relative heterochromatic fraction (RHF; F), and number of chromocentres (G) for mutant and wild-type nuclei of guard cells (GC) and pavement cells (PC) of 10-d-old cotyledons of wild-type (Col 0) and mutants (*atneap1*, *atneap3*, and *atneap1 atneap3*). (H) Schematic of *AtNEAP1* and *AtNEAP3* genes with insertion sites and locations of the primers used for RT-PCR. (I) RT-PCR performed on Col 0 and *atneap1 atneap3*. Negative controls (RT–) are presented where no MMLV-reverse transcriptase was added. An actin gene was used as the positive control.

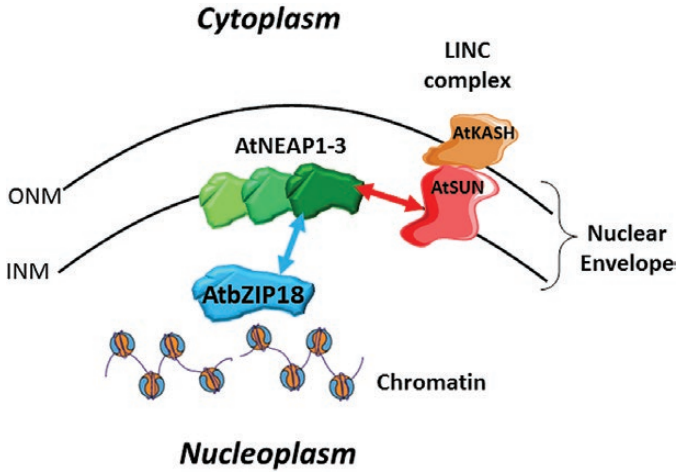


Fig. 8. Schematic representation of the NEAPs localised in the inner nuclear membrane (INM) and interacting with AtSUN, a component of the LINC complex, and AtbZIP18, a putative transcription factor linked to chromatin. ONM, outer nuclear membrane.

Discussion

The members of the family designated AtNEAP1–4 and characterised in this paper are plant-specific proteins associated with the inner nuclear envelope. Structurally, they are predominantly coiled-coil proteins, with an active NLS and a predicted C-terminal transmembrane domain; together these localise the proteins at the INM, predicted to be orientated with the coiled-coil domains in the nucleoplasm. As a full proteome of the plant INM is yet to be identified, the AtNEAPs are part of a small group of characterised plant INM proteins. Another well-characterised group of INM proteins are the SUN domain proteins, which are part of the nucleocytoskeletal bridging complexes. The ability of AtNEAPs to interact with AtSUN1 and AtSUN2 strongly indicates that AtNEAPs are also part of these LINC complexes and may be involved in some LINC functions (Fig. 8). Interestingly, plant LINC complex components such as AtSUNs and AtCRWN have been shown to regulate nuclear morphology (Dittmer *et al.*, 2007; Graumann *et al.*, 2014; Poulet *et al.*, 2015). The nuclear morphology changes observed in the AtNEAP1 and AtNEAP3 knockout mutants support the hypothesis that the AtNEAPs may also be involved in this process. Similarly, a reduction in primary root length, as observed here for the AtNEAP1-AtNEAP3 double-knockout, have previously been reported for the plant KASH protein AtTIK, also an interactor with AtSUNs (Graumann *et al.*, 2014). This raises the question whether AtTIK, AtNEAP1, and AtNEAP3 may be part of LINC complexes involved in cellular events that affect root growth. Certainly, *AtNEAP1* and *AtNEAP3* are expressed in this tissue and *AtNEAP3* is expressed at higher levels than in other tissues. Interestingly, AtNEAP3 in particular appears to have a function in chromatin organization, judging by the mutant's effects on chromocentre organisation and the relative hetrochromatic fraction. While it remains unclear how AtNEAP3 is linked to chromatin, a more direct association with chromatin has been identified for AtNEAP1. The interaction of AtNEAP1 with chromatin is suggested by the identification by MYTH of a DNA-binding leucine zipper transcription factor, AtbZIP18, as an interaction partner, and by altered localisation of CFP-AtNEAP1 resulting from co-expression with YFP-AtbZIP18. This is the first evidence in plants that LINC complexes are also associated with chromatin. The functional significance of the AtNEAP1–AtbZIP18 interaction will be explored in future studies.

Our *in vivo* and *in planta* interaction data show that all three AtNEAPs are able to homomerise and heteromerise. All NEAPs have extensive coiled-coils and it could be hypothesised that they play a role in mediating NEAP–NEAP interactions – although, at least for AtNEAP3, the first coiled-coil domain is not required as the YFP-AtNEAP3ΔCC1 mutant relocalises together with full-length CFP-AtNEAP3 at the NE. Interestingly, AtNEAP1 and AtNEAP2 appear to be more tightly anchored at the INM than AtNEAP3, indicating that they might be involved in different binding or protein complexes. This is also supported by the different expression patterns of AtNEAP1–3, where AtNEAP1 and 2 appear more highly expressed than AtNEAP3. The observation that all three AtNEAPs have reduced mobility at the NE, comparable to other NE and NE-associated proteins, indicates that they are functional components of the NE.

Taken together, the AtNEAPs are a novel family of nuclear-envelope proteins and our identification and initial characterisation of the NEAP family adds one more component to the rapidly developing story of the plant nuclear envelope, nucleoskeleton, and chromatin interactome, and will provide a basis for further understanding of the way in which the plant nucleus is structured and functions.

Supplementary data

Supplementary data are available at *JXB* online.

Figure S1. Multiple sequence alignment of AtNEAPs1–4 and prediction of functional domains.

Figure S2. Maximum-likelihood phylogenetic tree of the NEAPs gene family.

Figure S3. NEAP coding sequences used for phylogenic reconstruction, and substitution rate calculations.

Figure S4. Expression profile of AtNEAP1–4 mRNA obtained from RNA-seq and Genevestigator.

Figure S5. Split ubiquitin membrane yeast two hybrid (MYTH) assay.

Table S1. Primers used to genotype *Atneap* and *Atbzip18* insertion mutants

Table S2. Primers used to clone *AtNEAP*- and *AtbZIP18*-coding DNAs

Table S3. Fluorescent protein fusions constructed in this study.

Acknowledgements

KG is a Leverhulme Trust Early Career Fellow. The work was supported by the CNRS, INSERM, Blaise Pascal and Oxford Brookes Universities. We acknowledge the contribution of Ting Lu to the early stages of this project during his MPhil studies.

References

Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics **31** 166–169.

Brandizzi F, Irons SL, Evans DE. 2004. The plant nuclear envelope: new prospects for a poorly understood structure. New Phytologist **163,** 227–246.

Ciska M, Masuda K, Moreno Diaz de la Espina S. 2013. Lamin-like analogues in plants: the characterization of NMCP1 in *Allium cepa*. Journal of Experimental Botany **64,** 1553–1564.

Ciska M, Moreno Diaz de la Espina S. 2013. NMCP/LINC proteins: putative lamin analogs in plants? Plant Signalling and Behaviour **8** e26669.

Crisp M, Liu Q, Roux K, Rattner JB, Shanahan C, Burke B, Stahl PD, Hodzic D. 2006. Coupling of the nucleus and cytoplasm: role of the LINC complex. The Journal of Cell Biology **172,** 41–53.

Cserzo M, Wallin E, Simon I, von Heijne G, Elofsson A. 1997. Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Engineering **10,** 673–676.

Dechat T, Adam SA, Taimen P, Shimi T, Goldman RD. 2010. Nuclear lamins. Cold Spring Harbor Perspectives in Biology **2,** a000547.

Delorenzi M, Speed T. 2002. An HMM model for coiled-coil domains and a comparison with PSSM-based predictions. Bioinformatics **18,** 617–625.

Dittmer TA, Stacey NJ, Sugimoto-Shirasu K, Richards EJ. 2007. *LITTLE NUCLEI* genes affecting nuclear morphology in *Arabidopsis thaliana*. Plant Cell **19,** 2793–2803.

Evans DE, Pawar V, Smith SJ, Graumann K. 2014. Protein interactions at the higher plant nuclear envelope: evidence for a linker of nucleoskeleton and cytoskeleton complex. Frontiers in Plant Sciences **5,** 183.

Fiserova J, Kiseleva E, Goldberg MW. 2009. Nuclear envelope and nuclear pore complex structure and organization in tobacco BY-2 cells. The Plant Journal **59,** 243–255.

Gonzalez-Suarez I, Redwood AB, Perkins SM, *et al*. 2009. Novel roles for A-type lamins in telomere biology and the DNA damage response pathway. The EMBO Journal **28,** 2414–2427.

Goto C, Tamura K, Fukao Y, Shimada T, Hara-Nishimura I. 2014. The novel nuclear envelope protein KAKU4 modulates nuclear morphology in Arabidopsis. The Plant Cell **26,** 2143–2155.

Graumann K. 2014. Evidence for LINC1-SUN associations at the plant nuclear periphery. PLoS One **9,** e93406.

Graumann K, Evans DE. 2010. Plant SUN domain proteins: components of putative plant LINC complexes? Plant Signalling and Behaviour **5,** 154–156.

Graumann K, Evans DE. 2011. Nuclear envelope dynamics during plant cell division suggest common mechanisms between kingdoms. Biochemical Journal **435,** 661–667.

Graumann K, Irons SL, Runions J, Evans DE. 2007. Retention and mobility of the mammalian lamin B receptor in the plant nuclear envelope. Biology of the Cell **99,** 553–562.

Graumann K, Runions J, Evans DE. 2010. Characterization of SUN-domain proteins at the higher plant nuclear envelope. The Plant Journal **61,** 134–144.

Graumann K, Vanrobays E, Tutois S, Probst AV, Evans DE, Tatout C. 2014. 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,** 6499–6512.

Gruenbaum Y, Margalit A, Goldman RD, Shumaker DK, Wilson KL. 2005. The nuclear lamina comes of age. Nature Reviews Molecular Cell Biology **6,** 21–31.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biology **14,** R36.

Koreny L, Field MC. 2016. Ancient eukaryotic origin and evolutionary plasticity of nuclear lamina. Genome Biology and Evolution in press, doi:10.1093/gbe/evw087.

Kosugi S, Hasebe M, Tomita M, Yanagawa H. 2009. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. Proceedings of the National Academy of Sciences, USA **106,** 10171–10176.

Letunic I, Doerks T, Bork P. 2012. SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Research **40,** D302–D305.

Lu T. 2011. *Identification of a plant nuclear envelope associated protein and its characterisation*. MPhil Thesis, Oxford Brookes University, UK.

Lupas A, Van Dyke M, Stock J. 1991. Predicting coiled coils from protein sequences. Science **252,** 1162–1164.

Masuda K, Takahashi S, Nomura K, Arimoto M, Inoue M. 1993. Residual structure and constituent proteins of the peripheral framework of the cell nucleus in somatic embryos from *Daucus carota* L. Planta **191,** 532–540.

Masuda K, Xu ZJ, Takahashi S, Ito A, Ono M, Nomura K, Inoue M. 1997. Peripheral framework of carrot cell nucleus contains a novel protein predicted to exhibit a long alpha-helical domain. Experimental Cell Research **232,** 173–181.

McDonnell AV, Jiang T, Keating AE, Berger B. 2006. Paircoil2: improved prediction of coiled coils from sequence. Bioinformatics **22,** 356–358.

Meier I. 2007. Composition of the plant nuclear envelope: theme and variations. Journal of Experimental Botany **58,** 27–34.

Minguez A, Moreno Diaz de la Espina S. 1993. Immunological characterization of lamins in the nuclear matrix of onion cells. Journal of Cell Science **106,** 431–439.

Moriguchi K, Suzuki T, Ito Y, Yamazaki Y, Niwa Y, Kurata N. 2005. Functional isolation of novel nuclear proteins showing a variety of subnuclear localizations. The Plant Cell **17,** 389–403.

Murphy SP, Simmons CR, Bass HW. 2010. 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 Biology **10,** 269.

Nguyen Ba AN, Pogoutse A, Provart N, Moses AM. 2009. NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. BMC Bioinformatics **10,** 202.

Oda Y, Fukuda H. 2011. Dynamics of Arabidopsis SUN proteins during mitosis and their involvement in nuclear shaping. The Plant Journal **66,** 629–641.

Okada Y, Suzuki T, Sunden Y, *et al*. 2005. Dissociation of heterochromatin protein 1 from lamin B receptor induced by human polyoma virus agnoprotein: role in nuclear egress of viral particles. EMBO Reports **6,** 452–457.

Oldenburg KR, Vo KT, Michaelis S, Paddon C. 1997. Recombination-mediated PCR-directed plasmid construction in vivo in yeast. Nucleic Acids Research **25,** 451–452.

Poulet A, Arganda-Carreras I, Legland D, Probst A, Andrey P, Tatout C. 2015. NucleusJ: anImageJ plugin for quantifying 3D images of interphase nuclei. Bioimage Informatics **31,** 1144–1146.

Sakamoto Y, Takagi S. 2013. LITTLE NUCLEI 1 and 4 regulate nuclear morphology in *Arabidopsis thaliana*. Plant and Cell Physiology **54,** 622–633.

Schwacke R, Schneider A, van der Graaff E, Fischer K, Catoni E, Desimone M, Frommer WB, Flugge UI, Kunze R. 2003. ARAMEMNON, a novel database for *Arabidopsis* integral membrane proteins. Plant Physiology **131,** 16–26.

Smith S, Galinha C, Desset S, Tolmie F, Evans D, Tatout C, Graumann K. 2015. Marker gene tethering by nucleoporins affects gene expression in plants, Nucleus **6,** 471–478.

Snider J, Kittanakom S, Curak J, Stagljar I. 2010*a*. Split-ubiquitin based membrane yeast two-hybrid (MYTH) system: a powerful tool for identifying protein-protein interactions. Journal of Visualized Experiments **36,** e1698.

Snider J, Kittanakom S, Damjanovic D, Curak J, Wong V, Stagljar I. 2010*b*. Detecting interactions with membrane proteins using a membrane two-hybrid assay in yeast. Nature Protocols **5,** 1281–1293.

Sparkes IA, Runions J, Kearns A, Hawes C. 2006. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. Nature Protocols **1,** 2019–2025.

Toufighi K, Brady SM, Austin R, Provart NJ. 2005. The Botany Array Resource: e-Northerns, Expression Angling, and promoter analyses. The Plant Journal **43,** 153–163.

Tzur YB, Margalit A, Melamed-Book N, Gruenbaum Y. 2006. Matefin/SUN-1 is a nuclear envelope receptor for CED-4 during *Caenorhabditis elegans* apoptosis. Proceedings of the National Academy of Sciences, USA **103,** 13397–13402.

Van Zanten M, Carles A, Li Y, Soppe WJ. 2012. Control and consequences of chromatin compaction during seed maturation in *Arabidopsis thaliana*. Plant Signalling and Behaviour **7,** 338–341.

Van Zanten M, Koini MA, Geyer R, *et al*. 2011. Seed maturation in *Arabidopsis thaliana* is characterized by nuclear size reduction and increased chromatin condensation. Proceedings of the National Academy of Sciences, USA **108,** 20219–20224.

Wang H, Dittmer TA, Richards E. 2013. Arabidopsis CROWDED NUCLEI (CRWN) proteins are required for nuclear size control and heterochromatin organization. BMC Plant Biology **13**

Wilson KL, Berk JM. 2010. The nuclear envelope at a glance. Journal of Cell Sciences **123,** 1973–1978.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution **24,** 1586–1591.

Figure **S1**

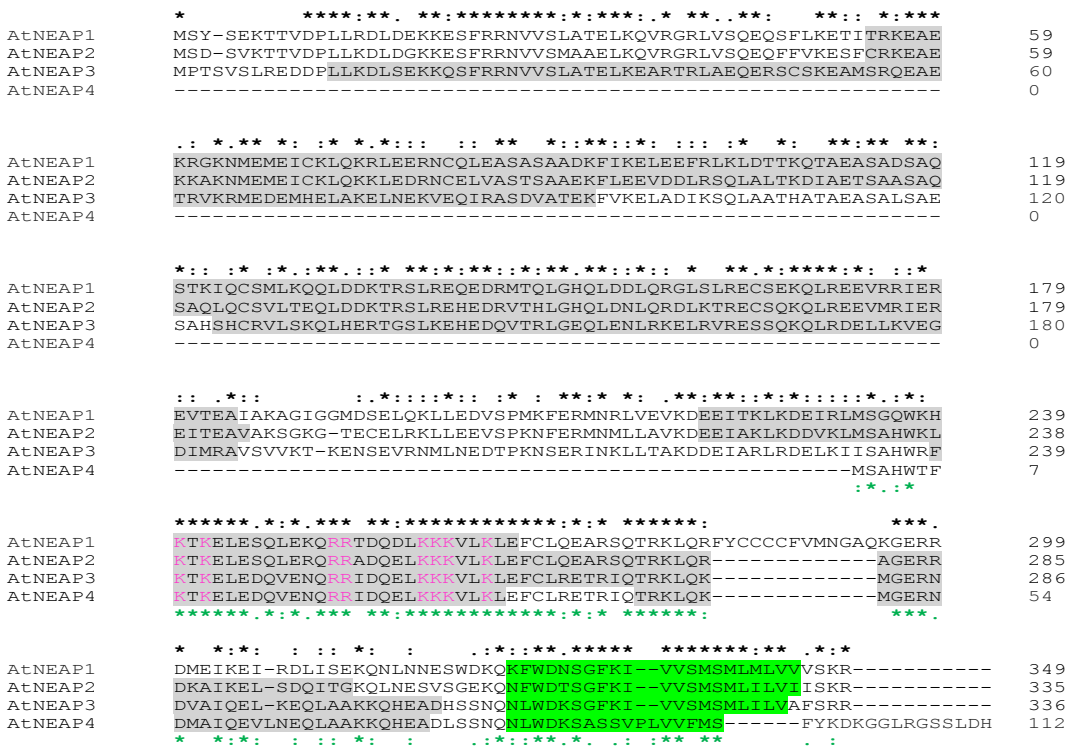


Figure S1 Clustal Omega (1,2.1) multiple sequence alignment of AtNEAPs 1-4 and prediction of functional domains. * donates consensus sequence between AtNEAPs 1-3; green * represents consensus between AtNEAPs 1-4. Coiled coil domains highlighted in grey, NLS pink and TM domains in green. Coiled coil domains were predicted using SMART, COILS, PairCoil2 and Marcoil (Lupas et al, 1991; Dolerenzi and Speed, 2002; McDonnell et al., 2006, Letunic et al., 2012). NLS predicted using cNLS mapper and NLSstradamus (Kosugi et al., 2009; Nguyen Ba at al., 2009_. TM domains were predicted using ARAMEMNON and DAS (Cserzo et al., 1997; Schwacke et al., 2003).

Figure S2

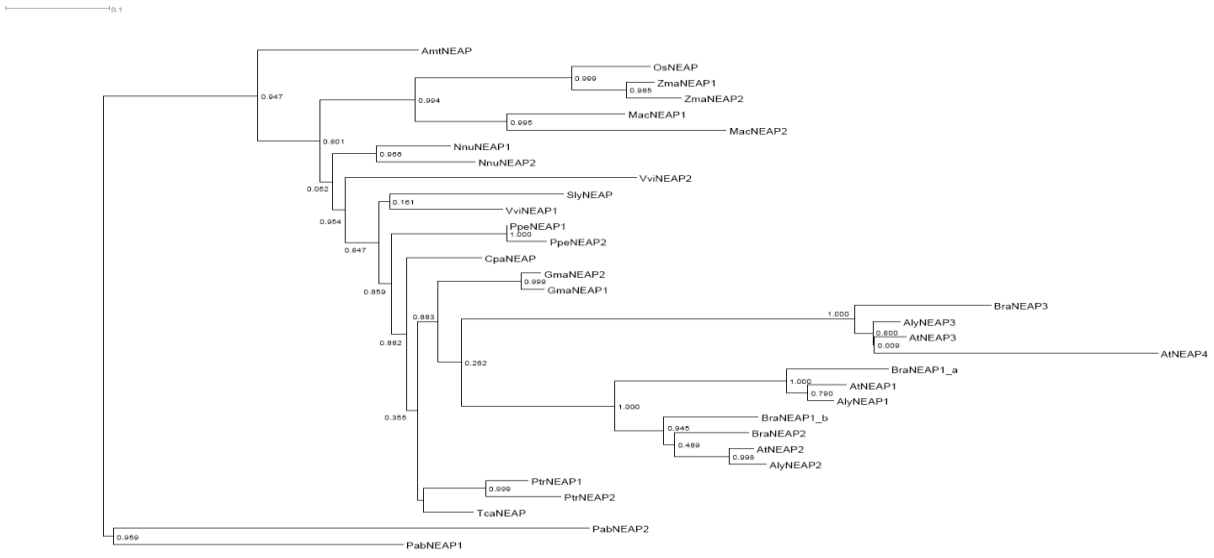
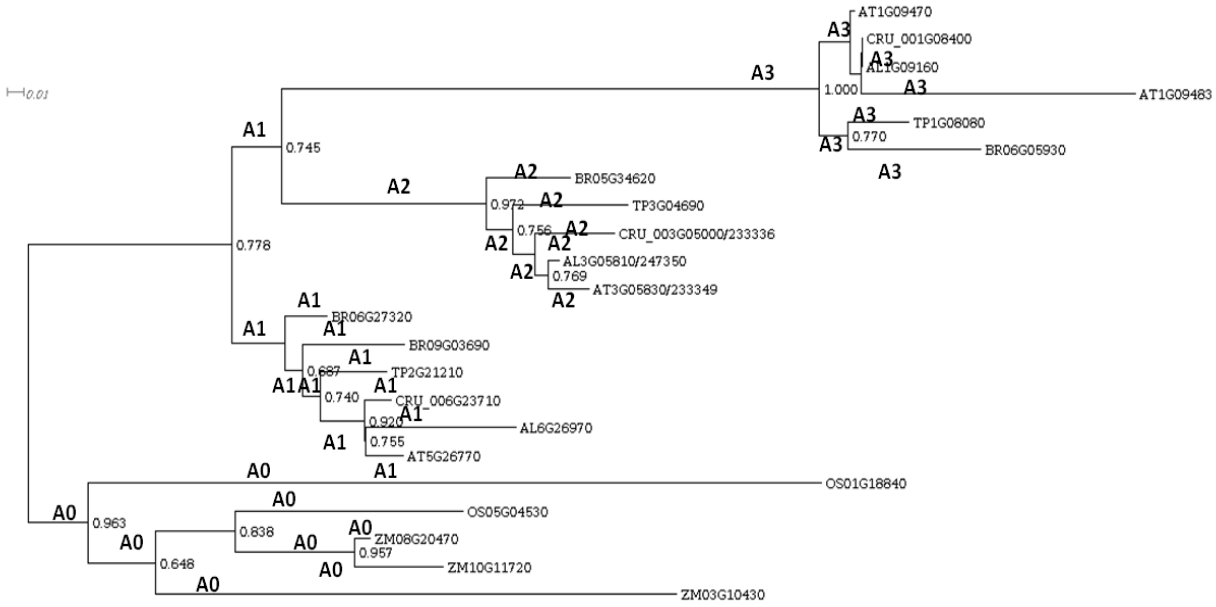


Figure S2 Maximum likelihood (ML) phylogenetic tree. The ML tree was inferred using proteins sequences from dicot species including *Arabidopsis thaliana* (AtNEAP1, AtNEAP2, AtNEAP3 and AtNEAP4 respectively [GenBank: NP_001189818, NP_568487, NP_172418 and NP_683289]), *Arabidopsis lyrata* (AlyNEAP1, AlyNEAP2 and AlyNEAP3 respectively [GenBank: XP_002882437, XP_002874325 and XP_002892507]) *Brassica rapa* (BraNEAP1_a, BraNEAP1_b, BraNEAP2 and BraNEAP3 respectively [GenBank: CDY27738, CDY19756, XP_009151160 and XP_00914824]), *Carica papaya* (CpaNEAP [IdPlaza: CP00048G02160]), *Glycine max* (GmaNEAP1 and GmaNEAP2 respectively [GenBank: XP_003555780 and XP_003536028]), *Nelumbo nuciferagi* (NnuNEAP1 and NnuNEAP2 respectively [GenBank: XP_010276551 and XP_010270974]), *Prunus persica* (PpeNEAP1 and PpeNEAP2 respectively [Genbank: XP_007205454 and XP_007205455]), *Populus trichocarpa* (PtrNEAP1 and PtrNEAP2 respectively [Genbank: XP_006382272 and XP_002318927], *Solanum lycopersicum* (SlyNEAP: [GenBank: XP_004230355]), *Theobroma cacao* (TcaNEAP: [Genbank: XP_007031253]), *Vitis vinifera* (VviNEAP1 and VviNEAP2 respectively [GenBank: CBI39661 and XP_002280405]), monocot species including *Zea mays* (ZmaNEAP1 and ZmaNEAP2 respectively [GenBank: NP_001149106 and XP_008661538]), *Oryza sativa* (OsNEAP: [Genbank NP_001054577]), *Musa acuminata* (MacNEAP1 and MacNEAP2 respectively [Genbank: XP_009382460 and XP_009411597), basal angiosperm species including *Amborella trichopoda* (AmtNEAP :[Genbank : XP_006840319]), and gymnosperm species including *Picea abies* (PabNEAP1 and PabNEAP2 respectively [congenie: MA_136804g0010 MA_902507g0010]) . Numbers above branches indicate bootstrap values.

Figure S3



	hypothesis	ωA0	ωa1	ωa2	ωa3	ωa4	l		2Δl	p value
h0	H0: ωA0 = ωA1 = ωA2 = ωA3 = ωA4	0.17489	= ωA0	= ωA0	= ωA0	= ωA0	-6495.969814	Na	NA	
h1	H1: ωA0 = ωA1 = ωA2 ≠ ωA3 = ωA4	0.17049	= ωA0	= ωA0	0.20053	= ωA3	-6495.767548	H0 vs H1	0.202266	0.6528
h2	H2: ωA0 = ωA1 = ωA2 ≠ ωA3 ≠ ωA4	0.17025	= ωA0	= ωA0	0.09985	0.84655	-6487.539788	H0 vs H1	8.430026	0.0036
h3	H3: ωA0 ≠ ωA1 ≠ ωA2 ≠ ωA3 = ωA4	0.16185	0.16890	0.20727	0.20171	= ωA3	-6495.379931	H0 vs H1	0.589883	0.4424
h4	H4: ωA0 ≠ ωA1 ≠ ωA2 ≠ ωA3 ≠ ωA4	0.1617	0.1686	0.2078	0.1002	0.8515	-6487.137043	H0 vs H1	8.832771	0.0029
h5	H5: ωA0 = ωA1 = ωA2 =ωA3 ≠ ωA4	0.16116	= ωA0	= ωA0	= ωA0	0.84888	-6488.941673	H0 vs H5	7.028141	0.0080

Figure S3 NEAP coding sequences were used for phylogenic reconstruction and substitution rate calculation. Selected protein sequences were aligned with MUSCLE multiple sequence alignment and maximum likelihood analysis was performed with FastTree using default parameters. ω (the ratio of nonsynonymous/synonymous substitution rates) was determined using Codeml from the PaML package NEAP4 has an rate evolution rate (0.8) higher than the other NEAPs, which have rates equal or equivalent to expressed genes.
This shows an accumulation of non synonymous mutations in the *AtNEAP4* sequence, which implies a possible decrease in selection pressure, and a pseudogeneisation of *AtNEAP4*

Figure S4

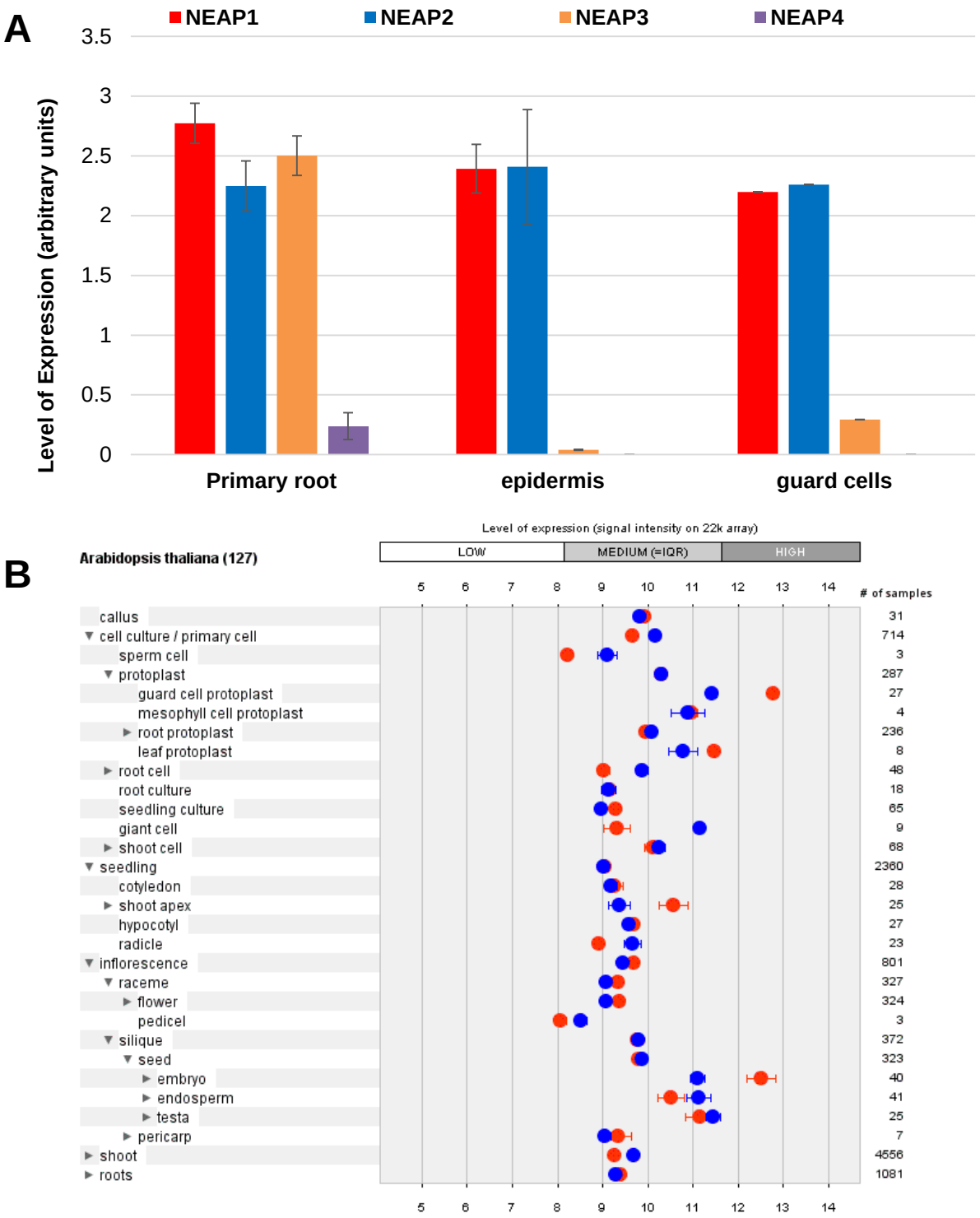


Figure S4 Expression profile of *AtNEAP1, 2, 3* and *4* mRNA obtained from RNA-Seq and Genevestigator (http://www.ncbi.nlm.nih.gov/geo/browse/; Toufighi et al., 2005). A) RNA-Seq data shows that *AtNEAP1* and *AtNEAP2* are expressed at higher levels in all three tissues while *AtNEAP3* is also highly expressed in root but not in episermis and guard cells. No or low expression of *AtNEAP4* suggests it may be a pseudogene. B) Genevestigator microarray data shows *AtNEAP1* (red) and *AtNEAP2* (blue) expressed at medium levels in all tissues with higher levels of AtNEAP1 expression in guard cells and embryos.

Figure S5

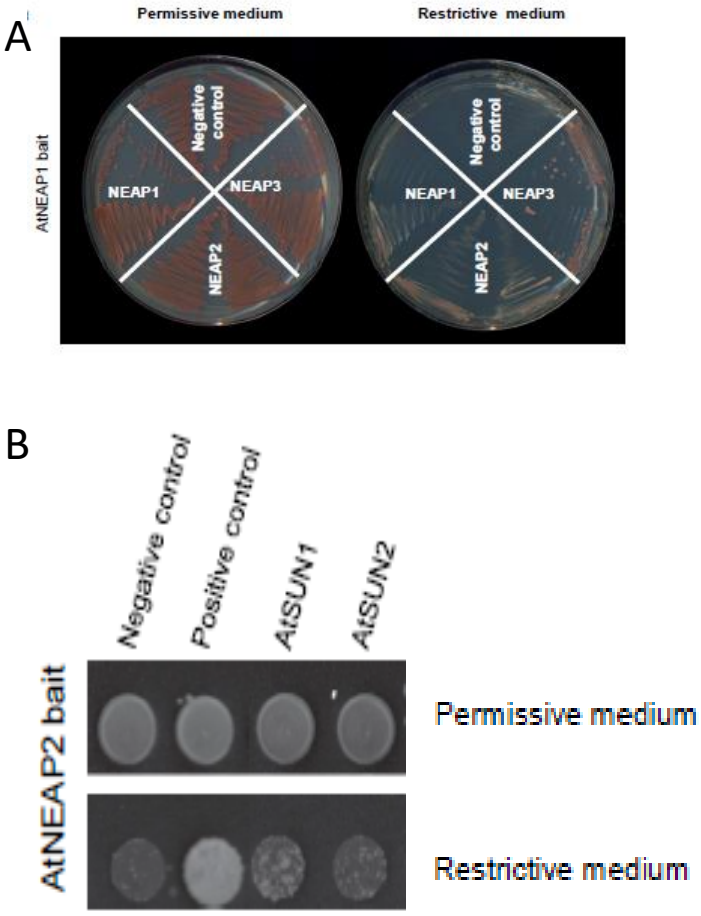


Figure S5 Split ubiquitin membrane yeast two hybrid (MYTH) assay. A) Colonies of yeast showing transformed yeast cells grown on permissive and restrictive medium indicating successful bait prey interaction for AtNEAPs 1-3. B) Yeast colonies containing AtNEAP2 bait transformed with AtSUN1 and AtSUN2 prey vectors grown on permissive and restrictive medium. Transformation with empty prey vector was used as a negative control; the positive control was an ER protein fused to the N-terminus of the split ubiquitin molecule.

Table S1 lists all primers used to genotype *Atneap* and *Atbzip18* insertion mutants

Primer name	Forward/ Reverse	Primer description	Tm (˚C)	Sequence
LPNEAP1	F	NEAP1_SAIL846_B07	49	CTCTGCAGCTTTCTTGTCTGG
RPNEAP1	R	NEAP1_SAIL846_B07	47	AGCTTGAAGCTTCTGCATCTG
LB3_SAIL	F	SAIL left border	55	TAGCATCTGAATTTCATAACCAATCTCGATACAC
LPNEAP2	F	NEAP2_SALK_012087	43	TTTGATTCGATGCTTATGCAG
RPNEAP2	R	NEAP2_SALK_012087	47	AGAAGCAGCACTTGTTTCTGC
LBb1.3_SALK	F	SALK left border	42	ATTTTGCCGATTTCGGAAC
Wisc_LPNEAP2	F	WiscDsLoxHs194_12D	47	TACCATATCAGAGCGGGATTG
Wisc_RPNEAP2	R	WiscDsLoxHs194_12D	45	TTGTTGCTCGAACTGTTGTTG
WiscHS_LB	F	WiscDsLoxHs left border	55	TGATCCATGTAGATTTCCCGGACATGAAG
178C02_LPN2a	F	NEAP2_GABI178C02: insertion:chr5 9409102	47	TGCACCTGAGATTCAAGTTCC
178C02_RPN2a	R	NEAP2_GABI178C02: insertion:chr5 9409102	48	TGCTTTGGTAGGGTCAGAAATC
178C02_LPN2b	F	NEAP2_GABI178C02: insertion:chr5 9409081	43	CGCTTTTGAAAGATTTGGATG
178C02_RPN2b	R	NEAP2_GABI178C02: insertion:chr5 9409081	49	GCTTCAGTTATCTCACGCTCG
589B02_LPN2a	F	NEAP2_GABI_589B02: insertion:chr5 9409811	45	AAAGGGCCATTGATTACCAAG
589B02_RPN2a	R	NEAP2_GABI_589B02: insertion:chr5 9409811	45	AGAAATTCGGAAGGGAAAGAC
589B02_LPN2b	F	NEAP2_GABI_589B02: insertion:chr5 9409738	47	AGCGAGGTTTTAGACTTTCCG
589B02_RPN2b	R	NEAP2_GABI_589B02: insertion:chr5 9409738	47	CCTTTTCAGCAGCAGAAGTTG
GABI_8474	F	GABI right border	50	ATAATAACGCTGCGGACATCTACATTTT
LPNEAP3	F	NEAP3_WiscDsLoxHs086_02C	50	TTCCTACCAAACCCAGAAACC
RPNEAP3	R	NEAP3_WiscDsLoxHs086_02C	50	TCAGCCAATTCCTTCACAAAC
LPNEAP4	F	NEAP4_SAIL_1239_G02	50	TTCACTCCAATGAAATCGAGC
RPNEAP4	R	NEAP4_SAIL_1239_G02	50	TTGTTCTTCTGGATCAGGTGG

Table S2 lists all primers used to clone *AtNEAP* and *AtbZIP18* coding DNAs

Primer name	Forward/ Reverse	Primer description	Tm (˚C)	Sequence
FNEAP1 (FTL5)	F	binds first 24 bp of AtNEAP1	63	ATGTCTTATTCTGAAAAAACGACG
RNEAP1(RTL5)	R	binds last 24 bp of AtNEAP1 (minus stop)	56	TCTCTTGGAGACTACCACTAACAT
RT_FNEAP1	F	binds bp 150-177 spanning the first intron of AtNEAP1	54	GAGACCATTACTAGAAAAGAAGCAGAG
RT_RNEAP1	R	binds bp 834-857 spanning the last intron of AtNEAP1	54	CAACAACAATAAAACCTCTGCAGC
FNEAP2	F	binds first 25bp of AtNEAP2	63	ATGTCGGATTCCGTCAAAACGACGG
RNEAP2	R	binds last 25bp of AtNEAP2 (minus STOP)	56	TCTTTTGGAGATAATAACTAATATC
FNEAP3	F	binds first 20bp of AtNEAP3	48	ATGCCAACTTCTGTTAGTCT
RNEAP3	R	binds last 19bp of AtNEAP3 (minus STOP)	47	ACGCCTAGAAAACGCAACT
FN3dCC1a	F	binds bp 280 to 299 of NEAP3, CC1 ends at 279	47	TTTGTGAAGGAATTGGCTG
FN3dCC1b	F	same as FN3dCC1a but adds beginning of NEAP3 as overhang	47	ATGCCAACTTCTGTTAGTCTAAGAGAGGATGATCCT-TTTGTGAAGGAATTGGCTG
RN3dCC2a	R	binds bp 351 to 369 of NEAP3, CC2 starts at 370	49	ATGTGCTGATTCAGCTGAC
RN3dCC2b	R	same as RN3dCC2a, but adds region after CC2 as overhang	49	GGTCTTGACCACTGATAC-ATGTGCTGATTCAGCTGAC
FN3dCC2a	F	binds bp 556 to 574 of NEAP3, CC2 ends at 555	48	GTATCAGTGGTCAAGACC
FN3dCC2b	F	same as FN3dCC2a, but adds region before CC2 as overhang	48	GTCAGCTGAATCAGCACAT-GTATCAGTGGTCAAGACC
RN3dNLSa	R	binds bp 701 to 718, NLS starts at 719	49	AAACCTCCAGTGAGCCG
RN3dNLSb	R	same as RN3dNLSa, but adds region after NLS as overhang	49	CTTGTATTGCCACATCGTT-AAACCTCCAGTGAGCCG
FN3dNLSa	F	binds bp 856 to 875 of NEAP3, NLS ends at 855	47	AACGATGTGGCAATACAAG
FN3dNLSb	F	same as FN3dNLSa, but adds region before NLS as overhang	47	CGGCTCACTGGAGGTTT-AACGATGTGGCAATACAAG
RN3dTMa	R	binds bp 913 to 933 of NEAP3, TM domain starts at 934	50	TTGGTTGCTAGAATGATCAGC
RN3dTMb	R	same as RN3dTMa, but adds last 12 bp after TM as overhang	50	ACGCCTAGAAAA-TTGGTTGCTAGAATGATCAGC
GWRN3dTM	R	GW primer binds end of delTM sequence cloned using RN3dTMa &b, addsattB2 sequence	50	GGGGACCACTTTGTACAAGAAAGCTGGGTC-ACGCCTAGAAAATTGGTTGC
FNEAP4	F	binds bp 1 - 18 of AtNEAP4	57	ATGTCGGCTCATTGGACG
RNEAP4	R	binds bp 316 - 336 of AtNEAP4	56	ATGATCAAGACTTGAACCACG
RNEAP4a	R	binds bp 308 - 329 of AtNEAP4	58	AGACTTGAACCACGTAATCCAC
attL1_FNEAP4	F	gateway primer, binds bp1 to 20 of AtNEAP4 with attB1 sequence	60	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCCGCCA-ATGTCGGCTCATTGGACGTT
RNEAP4_attL2	R	gateway primer, binds last 27 of AtNEAP4 with attB2 sequence	60	GGGGACCACTTTGTACAAGAAAGCTGGGTC-ATGATCAAGACTTGAACCACGTAATCC
FbZIP18	F	binds bp 1 to 21 of bZIP18	60	ATGGAGGATCCTTCTAACCCACA
RbZIP18	R	binds last 23 bp of bZIP18 (minus STOP)	57	AGTGCTGCTGCTTTCACTGAC
attL1_FbZIP	F	gateway primer, binds bp1 to 23 of AtbZIP18 with attB1 sequence	61	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCCGCCA-ATGGAGGATCCTTCTAACCCACA
RbZIP_attL2	R	gateway primer, binds last 20 bp of AtbZIP18 with attB2 sequence	61	GGGGACCACTTTGTACAAGAAAGCTGGGTC-CATAGTGCTGCTGCTTTCAC

Table S3 lists all fluorescent protein fusions constructed in this study

Construct	Gateway destination Vector	Bacterial resistance*
35S-YFP-NEAP2	pB7WGY2	spectinomycin
35S-NEAP2-CFP	pK7CWG2	spectinomycin
35S-CFP-NEAP2	pK7WGC2	spectinomycin
35S-YFP-NEAP3	pB7WGY2	spectinomycin
35S-NEAP3-CFP	pK7CWG2	spectinomycin
35S-CFP-NEAP3	pK7WGC2	spectinomycin
35S-NEAP3ΔCC1-CFP	pK7CWG2	spectinomycin
35S-NEAP3ΔCC2-CFP	pK7CWG2	spectinomycin
35S-NEAP3ΔNLS-CFP	pK7CWG2	spectinomycin
35S-NEAP3ΔTM-CFP	pK7CWG2	spectinomycin
35S-YFP-NEAP3ΔCC1	pCambia1300-casetteA	kanamycin
35S-YFP-NEAP3ΔCC2	pCambia1300-casetteA	kanamycin
35S-YFP-NEAP3ΔNLS	pCambia1300-casetteA	kanamycin
35S-YFP-NEAP3ΔTM	pCambia1300-casetteA	kanamycin
35S-SUN2ΔSUN-YFP	pCambia1300-casetteB	kanamycin
35S-SUN2ΔCC-YFP	pCambia1300-casetteB	kanamycin
35S-SUN2ΔN-YFP	pCambia1300-casetteB	kanamycin
35S-YFP-bZIP18	pCambia1300-casetteA	Kanamycin

The Plant Cell, Vol. 29: 1773–1793, July 2017, www.plantcell.org © 2017 ASPB.

Arabidopsis ATRX Modulates H3.3 Occupancy and Fine-Tunes Gene Expression

Céline Duc,[a] Matthias Benoit,[a,b] Gwénaëlle Détourné,[a,c] Lauriane Simon,[a] Axel Poulet,[a,c] Matthieu Jung,[d] Alaguraj Veluchamy,[e,f] David Latrasse,[e] Samuel Le Goff,[a] Sylviane Cotterell,[a] Christophe Tatout,[a] Moussa Benhamed,[e,f] and Aline V. Probst[a,1]

[a] GReD, Université Clermont Auvergne, CNRS, INSERM, 63001 Clermont-Ferrand, France
[b] The Sainsbury Laboratory, University of Cambridge, Cambridge CB2 1LR, United Kingdom
[c] Department of Biological and Medical Sciences, Oxford Brookes University, Oxford OX3 0BP, United Kingdom
[d] Department of Functional Genomics and Cancer, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, 67404 Illkirch, France
[e] Institute of Plant Sciences Paris Saclay IPS2, CNRS, INRA, Université Paris-Sud, Université Evry, Université Paris-Saclay, 91405 Orsay, France
[f] Division of Biological and Environmental Sciences and Engineering, King Abdullah University of Science and Technology, Thuwal, 23955-6900, Kingdom of Saudi Arabia

ORCID IDs: 0000-0001-7426-4541 (C.D.); 0000-0002-3958-3173 (M.B.); 0000-0001-8328-7059 (G.D.); 0000-0001-7285-9437 (L.S.); 0000-0002-3415-8960 (A.P.); 0000-0002-3272-7322 (M.J.); 0000-0002-5349-5794 (A.V.); 0000-0003-2149-0128 (D.L.); 0000-0002-9991-6435 (S.L.G.); 0000-0001-9665-5009 (S.C.); 0000-0001-5215-2338 (C.T.); 0000-0002-4181-1702 (M.B.); 0000-0001-9534-8058 (A.V.P.)

Histones are essential components of the nucleosome, the major chromatin subunit that structures linear DNA molecules and regulates access of other proteins to DNA. Specific histone chaperone complexes control the correct deposition of canonical histones and their variants to modulate nucleosome structure and stability. In this study, we characterize the *Arabidopsis thaliana* Alpha Thalassemia-mental Retardation X-linked (ATRX) ortholog and show that ATRX is involved in histone H3 deposition. Arabidopsis ATRX mutant alleles are viable, but show developmental defects and reduced fertility. Their combination with mutants of the histone H3.3 chaperone HIRA (Histone Regulator A) results in impaired plant survival, suggesting that HIRA and ATRX function in complementary histone deposition pathways. Indeed, ATRX loss of function alters cellular histone H3.3 pools and in consequence modulates the H3.1/H3.3 balance in the cell. H3.3 levels are affected especially at genes characterized by elevated H3.3 occupancy, including the 45S ribosomal DNA (45S rDNA) loci, where loss of ATRX results in altered expression of specific 45S rDNA sequence variants. At the genome-wide scale, our data indicate that ATRX modifies gene expression concomitantly to H3.3 deposition at a set of genes characterized both by elevated H3.3 occupancy and high expression. Together, our results show that ATRX is involved in H3.3 deposition and emphasize the role of histone chaperones in adjusting genome expression.

INTRODUCTION

Gene regulation in the eukaryotic genome requires a controlled balance between packaging the large linear DNA molecules and permitting regulated access to protein complexes involved in DNA transcription, replication, and repair. Within nucleosomes, the basic building blocks of chromatin, the double-stranded DNA helix wraps around octamers of histone proteins. Canonical histones are deposited on newly synthesized DNA to maintain nucleosomal density following passage of the replication fork. These canonical histones can then be replaced with specific histone variants to modify nucleosome composition, stability, higher-order chromatin organization, and DNA accessibility in a site-specific manner (Talbert and Henikoff, 2010). Most eukaryotes express variants of the canonical histone H3.1, such as the replacement variant H3.3, as well as tissue-specific H3 variants (Talbert et al., 2012). In mammals, H3.3 deposition is associated with dynamic chromatin regions such as transcriptionally active genes and regulatory regions, with high nucleosome turnover, and DNA accessibility. The plant model *Arabidopsis thaliana* also encodes H3.1 and H3.3 proteins that differ by only four amino acids, as well as H3.3-like variants expressed in specific reproductive tissues (Ingouff et al., 2010; Okada et al., 2005). Genome-wide studies revealed preferential enrichment of H3.1 at heterochromatic regions and of H3.3 at active genes, promoters, and telomeric repeats (Stroud et al., 2012; Wollmann et al., 2012; Vaquero-Sedas and Vega-Palas, 2013; Shu et al., 2014).

To ensure the incorporation of the adequate histone type at the right time and genomic location, specialized proteins called histone chaperones associate with histones during their shuttling from cytoplasm to nucleoplasm and deposit histones on DNA.

[1] Address correspondence to aline.probst@uca.fr.
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: Aline V. Probst (aline.probst@uca.fr).
www.plantcell.org/cgi/doi/10.1105/tpc.16.00877

These histone chaperones operate in a coordinated network. For the assembly of H3-H4 dimers into nucleosomes, several chaperone complexes have been characterized in eukaryotes. The CAF-1 (Chromatin Assembly Factor 1) complex deposits H3 in a DNA-synthesis-coupled manner during DNA replication and repair (Tagami et al., 2004; Smith and Stillman, 1989) and is composed of the subunits FASCIATA1 (FAS1), FAS2, and MULTICOPY SUPPRESSOR OF IRA1 in Arabidopsis (Kaya et al., 2001). The HIR complex incorporates H3 histones in a DNA synthesis-independent manner throughout the cell cycle or in resting cells. In Arabidopsis, this complex is conserved and composed of HISTONE REGULATOR A (HIRA), UBINUCLEIN 1/2 (UBN1/UBN2), and CALCINEURIN BINDING PROTEIN1 (CABIN1) (Nie et al., 2014; Duc et al., 2015). In mammals, these chaperone complexes show clear variant specificity, with CAF-1 depositing the canonical histone H3.1, while the HIR complex assembles H3.3 (Ricketts et al., 2015; Tagami et al., 2004). Little is known so far concerning such a specificity of the different complexes for H3 histone variants in plants, except that the Arabidopsis HIR complex binds H3.3 (Nie et al., 2014). In mammals, additional proteins that function as histone chaperones have been described, such as the DEK (Sawatsubashi et al., 2010) and the ATRX/DAXX (Alpha Thalassemia-mental Retardation X-linked syndrome/Death-domain Associated protein) heterocomplex, which shows specificity for H3.3 (Drané et al., 2010; Goldberg et al., 2010; Lewis et al., 2010; Wong et al., 2010).

ATRX/DAXX and HIR show differential chromatin binding patterns in mammals (He et al., 2015; Pchelintsev et al., 2013) and are known to deposit histone H3.3 at distinct genomic regions. Indeed, while HIRA deposits H3.3 at genic regions (Ray-Gallet et al., 2011; Pchelintsev et al., 2013; Goldberg et al., 2010), ATRX/DAXX incorporates this histone variant at pericentromeric repeats, telomeres, endogenous retroviral elements, and silenced imprinted alleles (Elsässer et al., 2015; Voon et al., 2015; Filipescu et al., 2013; Gokhman et al., 2013; Xue et al., 2003; McDowell et al., 1999; Goldberg et al., 2010; Lewis et al., 2010). ATRX has the capacity to bind lysine 9 (K9) methylated histone tails via its ATRX-DNMT3-DNMT3L (ADD) N-terminal domain (Iwase et al., 2011; Eustermann et al., 2011). This function is thought to contribute to the specific deposition pattern of H3.3 histones. Besides the ADD domain, ATRX contains a C-terminal SWI/SNF2-like ATPase motif found in chromatin remodeling proteins. These proteins use the energy of ATP to modulate histone-DNA interactions within nucleosomes and contribute to a wide range of cellular processes, including recombination (Alexeev et al., 2003), DNA replication (Collins et al., 2002), and histone exchange (Mizuguchi et al., 2004; Konev et al., 2007). The human ATRX was described about two decades ago due to various disorders associated with mutations in the *ATRX* gene, such as the X-linked-thalassemia mental retardation syndrome, characterized by several developmental abnormalities (Gibbons et al., 1995a, 1995b). Besides its role as a component of a histone chaperone complex, several studies analyzed the effects of ATRX loss of function in mammals and defined additional roles for this protein, including telomere maintenance (notably by protecting them from replication fork stalling), DNA replication (defects in ATRX leading to a prolongation of the S-phase), and heterochromatin silencing (reviewed in Clynes et al., 2013). Indeed, several recent studies connect ATRX/DAXX with silencing of retrotransposons and satellite sequences via the incorporation of H3.3 and the recruitment of histone methyltransferase activity (Sadic et al., 2015; Elsässer et al., 2012; Voon et al., 2015; He et al., 2015).

The *ATRX* gene has been conserved through evolution, but intriguingly, in invertebrates such as *Drosophila melanogaster*, ATRX is split into two proteins, dADD1 (CG8290), which harbors a homolog of the human ATRX ADD domain (López-Falcón et al., 2014; Alekseyenko et al., 2014), and dATRX/XNP (X-linked Nuclear Protein), which contains the SWI/SNF-like ATPase domain (Bassett et al., 2008; Emelyanov et al., 2010). The dATRX/XNP protein shares similar functions with its mammalian counterpart; however, XNP isoforms localize to chromosome arms and heterochromatin regions (Bassett et al., 2008). ATRX null alleles are embryo-lethal in human and mouse (Garrick et al., 2006), and semilethal in fly (Lee et al., 2007). In plants, despite the identification of a putative ATRX-encoding gene in Arabidopsis (Shaked et al., 2006; Otero et al., 2014), the involvement of this protein in histone H3 deposition and chromatin function has not been analyzed yet.

In this study, we investigated the role of ATRX in histone H3 deposition and chromatin function in Arabidopsis. We showed that Arabidopsis plants harboring *atrx* mutant alleles are viable but show reduced vigor and fertility. Genetic analyses using a combination of mutants in the different histone H3 incorporation systems suggest that ATRX functions as a histone H3.3 chaperone. ATRX controls histone H3.3 cellular pools and chromatin incorporation and thereby fine-tunes gene expression. Genomic sites with medium to high H3.3 occupancy, including the 45S rDNA loci, show decreased H3.3 upon loss of ATRX. Moreover, at 45S rDNA, ATRX controls histone H3 occupancy and H3.3 levels, rRNA gene transcription, and variant dosage. This study characterized ATRX as a player in the plant chaperone network.

RESULTS

Identification and Characterization of the Arabidopsis ATRX Ortholog

Using the conserved domains of the human ATRX, we confirmed the identification of At1g08600, also called CHROMATIN REMODELLING20 (Shaked et al., 2006; Otero et al., 2014), as the unique Arabidopsis ATRX ortholog. Plant ATRX proteins shared similarities with mammalian and invertebrate ATRX proteins, although they formed a distinct monophyletic group (Figure 1A). As in mammals, and contrary to Drosophila that expresses the two distinct proteins dADD1 (López-Falcón et al., 2014; Alekseyenko et al., 2014) and dATRX/XNP, the Arabidopsis ATRX harbors simultaneously the N-terminal ADD domain, which contains a pocket for recognition of histone H3 tails (Iwase et al., 2011), and the C-terminal helicase domain (Figure 1B). The plant protein was shorter than its mammalian counterpart, devoided of the large central region involved in DAXX interaction (Tang et al., 2004) (Figure 1B), and localized to the nucleus when transiently expressed in tobacco (*Nicotiana benthamiana*) leaves (Figure 1C). Both the plant and the human ADD domains contained the GATA and PHD zinc finger helices and shared 36.8% similarity for this

domain as can be seen in the 3D model (Supplemental Figures 1A and 1D). The ADD domain is known to interact with histone H3 tails and preferentially with methylated H3K9 (Iwase et al., 2011). The C-terminal region of the Arabidopsis ATRX contained the DExDc and the HELICc subdomains constituting the ATPase domain required for chromatin remodeling and characteristic of the SNF2 family proteins (46.2% and 54.8% similarity, respectively, for DExDc and HELICc domains with HsATRX; Supplemental Figures 1B and 1C). In a yeast two-hybrid assay, Arabidopsis ATRX interacts with the canonical histone H3.1 and the variant H3.3 (Figure 1D) and ATRX can immunoprecipitate Arabidopsis H3.3 when both proteins are transiently coexpressed in tobacco leaves (Supplemental Figures 1E to 1G). Taken together, the Arabidopsis ATRX protein shares major features with its human counterpart and interacts with H3, leading us to explore its potential role in histone dynamics and assembly.

We first quantified transcript levels of *ATRX* by qRT-PCR, which showed expression in all tested plant tissues (Supplemental Figure 1H). Two different mutant alleles in which T-DNAs are inserted in an exon (Figure 2A) were obtained and homozygous mutant plants established. RT-PCR and qRT-PCR analyses revealed that the *atrx-1* allele is a knockout mutant (Figures 2B and 2C; Supplemental Figures 2A and 2B) as the remaining transcript would encode a short truncated protein devoid of the ADD and helicase domains (Supplemental Figure 2C). The *atrx-2* allele produced transcripts comprising part of the *ATRX* 5′ and 3′ regions (Figure 2B; Supplemental Figures 2B and 2C), suggesting that truncated versions of the ATRX protein containing the majority of the ADD and/or part of the HELICc domain could be produced.

Plants carrying either mutant allele were viable and showed no obvious vegetative abnormalities compared with wild-type plants (Figure 2D), except slightly reduced rosette surface and root growth (Figures 2E and 2F). In addition, *atrx* mutant siliques showed fewer viable seeds and an increased number of aborted seeds, and more unfertilized ovules were scored in *atrx-2* mutant plants (Figures 2G and 2H; Supplemental Figure 2H). The *atrx* anthers develop normally but Alexander staining revealed reduced pollen content in both mutants (Figure 2I). Given that in mammals, ATRX mutations are associated with replication defects (Leung et al., 2013; Huh et al., 2012), we performed flow cytometry analysis of nuclei from dissected whole cotyledons. For the *atrx-1* knockout allele, we noticed broader peaks in the flow cytometry profile (Supplemental Figure 2D). The latter might be caused by delayed S-phase progression or replication defects. While we found no hypersensitivity to the DNA replication inhibitor

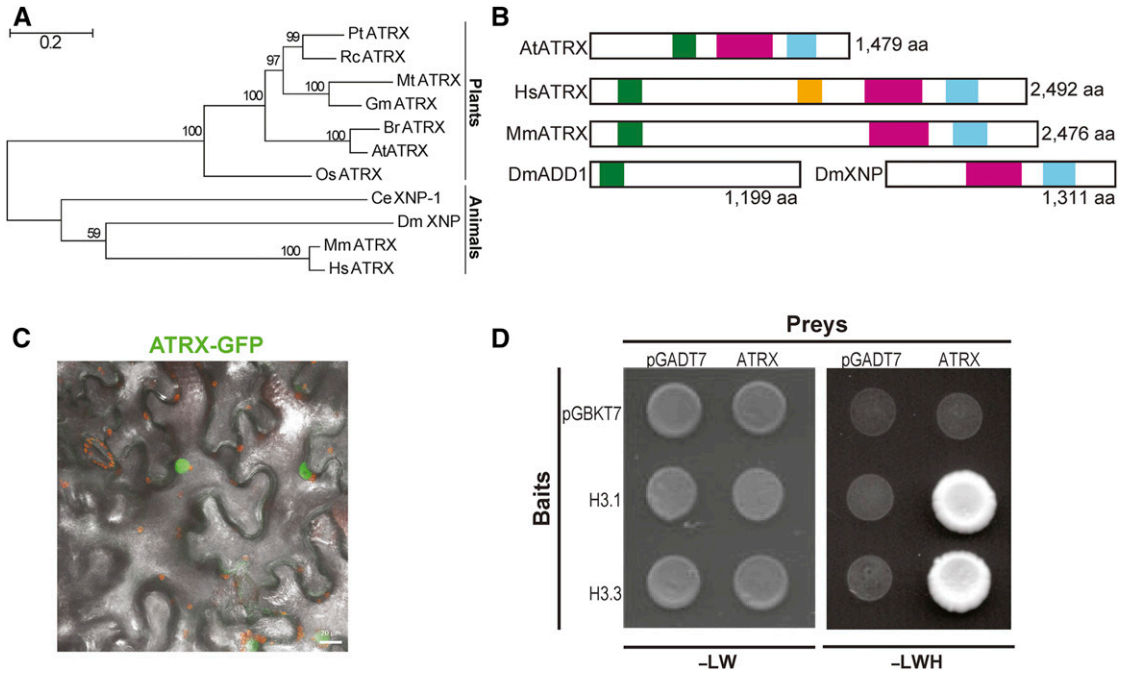


Figure 1. Characterization of the ATRX Ortholog in Arabidopsis.

(A) Phylogenetic tree of ATRX proteins. Pt, *Populus trichocarpa*; Rc, *Ricinus communis*; Mt, *Medicago truncatula*; Gm, *Glycine max*; At, *Arabidopsis thaliana*; Br, *Brassica rapa*; Os, *Oryza sativa*; Mm, *Mus musculus*; Hs, *Homo sapiens*; Dm, *Drosophila melanogaster*; Ce, *Caenorhabditis elegans*. Bar = 0.2 substitutions/per site. The percentage of trees in which the associated taxa clustered together is shown next to the branches.

(B) Functional domains of ATRX proteins. The ADD (ATRX-DNMT3-DNMT3L) domain is displayed as a green box, the DAXX-I (DAXX-interacting) domain as an orange box, the DEXDc (DEAD-like helicase superfamily) domain as a pink box, and the HELICc (HELICase superfamily C-terminal) domain as a blue box.

(C) Merged maximum intensity projection of confocal fluorescence and bright-field images of *N. benthamiana* leaves transiently expressing ATRX-GFP fusion (green) proteins. Chlorophyll fluorescence appears in red. Bar = 20 μm.

(D) Interaction of ATRX with histones H3.1 and H3.3 in a yeast two-hybrid assay. Photographs were taken after 3 d of yeast cell growth on leucine-tryptophan-/yeast nitrogen base medium (-LW) or on the selective leucine-tryptophan-histidine-yeast nitrogen base (-LWH) medium. The pGADT7 (prey) and pGBKT7 (bait) empty vectors were used as negative controls.

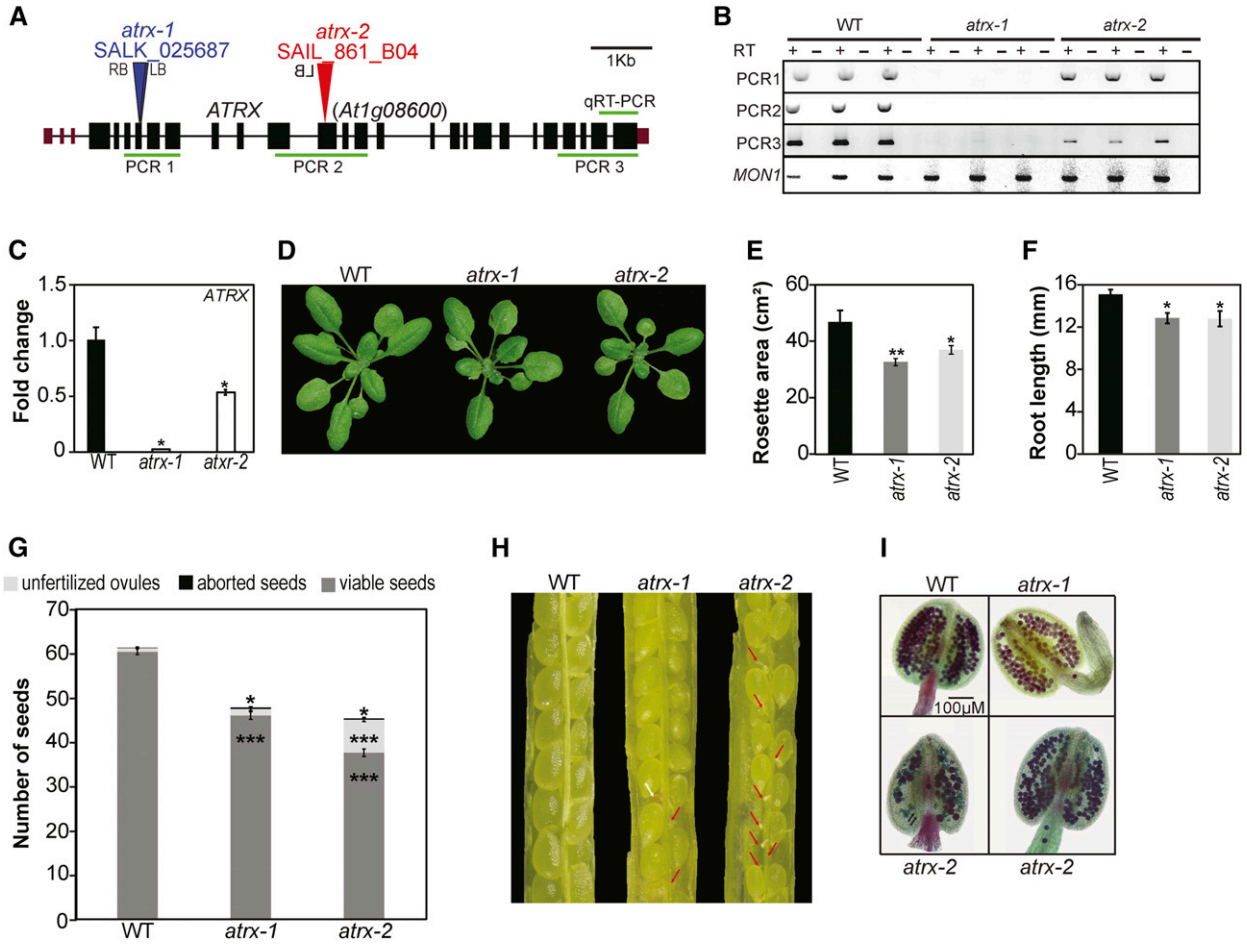


Figure 2. Characterization of the Arabidopsis *atrx* Mutant Alleles.

(A) Gene structure of Arabidopsis *ATRX*. Exons, black rectangles; untranslated regions, purple rectangles; introns, lines; T-DNA insertion, triangle; LB, left border; RB, right border.

(B) Analysis of *ATRX* transcripts produced in mutants with the *atrx-1* or *atrx-2* alleles by RT-PCR on three biological replicates consisting of independent pools of about twenty 2.5-week-old in vitro-grown whole plantlets sampled at the same time. The amplified regions are displayed by green lines in **(A)**. *MON1* (*At2g28390*) was used as a control.

(C) Mean expression of *ATRX* in *atrx* mutants analyzed by qRT-PCR in the same samples than in **(B)**. Transcript levels in the wild type were set to 1. The analyzed region is displayed in **(A)**.

(D) Representative 3-week-old wild-type and *atrx* mutant plantlets grown on soil.

(E) Quantification of rosette surface area (in cm^2) of *atrx* mutant plants. Mean rosette area is shown for at least six 2-week-old plants for each genotype. Student's *t* test; **P < 0.01.

(F) Root length quantification (in mm) of *atrx* mutant plants. Mean root length is calculated from at least four 5-d-old in vitro-grown plants for each genotype. A representative experiment out of three independent ones is displayed. Student's *t* test; *P < 0.05.

(G) Quantification of seed content in *atrx* mutant siliques compared with the wild type. Mean seed number was calculated from at least 30 siliques pooled from four plants per genotype grown at the same time. Student's *t* test compared with the wild type; *P < 0.05 and ***P < 0.001.

(H) Representative dissected siliques from *atrx* mutants. Red arrows indicate unfertilized ovules and the white arrow an aborted seed.

(I) Representative anthers out of four independent 4-week-old plants grown at the same time, for which pollen viability was assessed by Alexander staining. Both *atrx-1* and *atrx-2* anthers show reduced pollen content, and *atrx-2* anthers contain nonviable pollen (green color) indicated by black arrows.

Error bars of all panels represent SE of the mean.

hydroxyurea (HU) (Supplemental Figure 2E), we noted an increased number of nuclei in early compared with late S-phase after 5-ethynyl-2′-deoxyuridine staining of root apexes in the mutant (Supplemental Figure 2F). This might be indicative of a slower progression through early S-phase. Furthermore, a previous study has implicated ATRX in the DNA damage response using an RNAi line (Shaked et al., 2006). We confirmed moderate sensitivity to γ-irradiation in our *atrx* mutant alleles using a DNA-damage sensitivity assay, which monitors emergence of true leaves after seed irradiation (Supplemental Figure 2G).

Taken together, *atrx* mutant alleles are viable but display moderate growth, reproductive, and DNA replication defects.

Epistatic Relationships between Players in the H3 Incorporation Pathways

Since ATRX is an essential component of a histone chaperone complex in several organisms and is highly conserved through evolution, we hypothesized that ATRX may play a similar role in Arabidopsis. Hence, we investigated the genetic interaction between known Arabidopsis histone H3 chaperone complexes and ATRX, by crossing mutants of the HIR (Nie et al., 2014; Duc et al., 2015) and the CAF-1 complex (Kaya et al., 2001) to each *atrx* mutant allele.

For the *hira-1* mutants that lack the central subunit of the HIR complex (Duc et al., 2015), we obtained in the F2 generation only one double mutant plant from the cross with *atrx-1* (n = 111) and none with *atrx-2* (n = 115; Supplemental Tables 1A and 1B), although the respective loci are genetically unlinked. We confirmed the distorted segregation by analyzing the F3 progeny (Supplemental Tables 2A and 2B). The few *hira-1 atrx-1* double mutants recovered were sterile and severely affected in growth in contrast to the *hira-1* and *atrx-1* single mutant sister plants (Figure 3A). Furthermore, different combinations of F2 plants heterozygous for one mutation and homozygous for the other showed normal flowers (Supplemental Figures 3A and 3B), but reduced number of pollen grains (Supplemental Figures 3C and 3D) and, for certain combinations, reduced seed set compared with single mutants (Figure 3B; Supplemental Figure 3E). Finally, to confirm that the observed lethality of *hira-1 atrx* double mutations is indeed caused by loss of ATRX and not by an unlinked mutation, we analyzed the F1 progeny of a cross between an *atrx-1*/*atrx-1 hira-1*/*HIRA* plant and an *atrx-2*/*atrx-2 hira-1*/*HIRA* plant. Out of the 43 genotyped plants transheterozygous for *atrx* (i.e., *atrx-1*/*atrx-2*), none was homozygous for *hira-1* (Supplemental Table 1C). Jointly, this shows that the simultaneous loss of HIRA and ATRX impairs plant viability or causes severe developmental defects in the surviving plants. For the crosses with mutants for two other subunits of the HIR complex, UBN2 and CABIN1, we obtained viable double mutants with expected or reduced frequency, respectively (Supplemental Tables 1D to 1G, 2C, and 2D), and that displayed a reduced leaf surface compared with the corresponding single mutant for a member of the HIR complex (Figures 3C to 3E; Supplemental Figures 3F to 3H).

To explore the relationship with the CAF-1 complex, which is thought to be involved in replication-coupled histone deposition, we crossed both *atrx* alleles with the *fas2-5* mutant, a knockout mutant for the second largest subunit of the CAF-1 complex (Duc et al., 2015). Double mutants were obtained with the expected frequencies (Supplemental Table 3). When we looked closer at the development of the double mutant plants, we noticed flowers similar to *fas2-5* (Supplemental Figures 3I and 3J), but with aggravated anther shapes (Supplemental Figures 3K and 3L), and further reduction of the already low seed set (Figures 3F and 3G; Supplemental Figures 3M and 3N).

Taken together, this analysis reveals aggravated developmental defects and severe growth deficiencies or lethality when combining *atrx* mutations with plants deficient in CAF-1 and HIR complexes, respectively. In particular, the lethality for mutant combinations within the replication-independent system of H3 incorporation suggests that Arabidopsis ATRX could play a complementary role in histone H3.3 variant deposition in plants.

Loss of ATRX Affects Histone Pools and H3 Occupancy

Based on the above results, we reasoned that loss of Arabidopsis ATRX function might alter histone H3 pools. We used different protein extraction protocols (Durut et al., 2014; Honda et al., 1966) to recover histones from distinct cellular fractions (see Methods for details) of the wild type, *hira-1*, and the knockout *atrx-1* mutant. In comparison to wild-type plants, non-nucleosomal and total H3 amounts were significantly reduced in *atrx-1*, and non-nucleosomal and nuclear fractions were affected in *hira-1* (Figures 4A and 4B). The reduction observed in *atrx-1* occurred without changes in expression levels of the H3.1- and H3.3-encoding genes (Supplemental Figure 4A). Consistent with the association of H3 and H4 proteins as heterodimers and heterotetramers (Banks and Gloss, 2004), non-nucleosomal and total H4 levels were concomitantly reduced in *atrx-1* (Supplemental Figures 4B and 4C). These results show that loss of both HIRA and ATRX affects H3 and H4 histone pools.

We further investigated the possible contribution of ATRX to histone deposition by analyzing histone H3 occupancy at specific genomic sites. For this purpose, we combined H3-ChIP (chromatin immunoprecipitation) with quantitative PCR on wild-type, *atrx-1*, and *hira-1* plants. Since the combination of *atrx* and *hira-1* mutations was lethal, we hypothesized that ATRX affects incorporation of the replacement variant H3.3. Therefore, we first measured H3 occupancy in the 3′ regions of three transcriptionally active genes (*UBC28*, *UEV1C*, and *HXK1*) that have different expression levels (Duc et al., 2015) and were previously shown to be enriched in H3.3 (Stroud et al., 2012). While H3 occupancy was lower in *hira-1* for two of the three genes (Duc et al., 2015), it was reduced only at *HXK1* in *atrx-1* (Figure 4C). Given that ATRX binds to telomeres and subtelomeric regions of human chromosomes (Law et al., 2010), we analyzed three genes in subtelomeric regions. Only a minor reduction of H3 occupancy was found at *At5g67640* (Figure 4D). In mammals, ATRX is involved in H3.3 deposition at pericentromeric regions (Voon et al., 2015). We therefore determined H3 occupancy at two centromeric regions: the 180-bp repeats and the *106B* long terminal (LTR)-like repeats (Thompson et al., 1996; Fransz et al., 1998); and at two pericentromeric regions: *TSI* (Transcriptionally Silent Information) (Steimer et al., 2000) and at the 5S ribosomal DNA loci; as well as at *Mule* (Mutator-like, *At2g15810*), a DNA transposon located on a chromosome arm. While histone H3 occupancy was reduced in *hira-1* at *TSI*, *106B*, and *Mule* (Duc et al., 2015), it was not affected in *atrx-1* (Figure 4E). We further checked whether the H3K9me2 mark, the major plant heterochromatin signature present in pericentromeric regions and in patches of heterochromatin on chromosome arms (Bernatavichute et al., 2008), was maintained in *atrx-1*. Indeed, once normalized to H3, H3K9me2 levels were unchanged at the heterochromatic repeats *TSI*, *106B*, and *180bp* as well as at the *Mule* transposon or at *At5g67640*, a gene situated in the sub-telomeric region (Supplemental Figure 4D). This finding suggests that ATRX does not affect the setting or maintenance of this H3K9me2 repressive histone mark at several heterochromatic sequences. In agreement with this observation, silencing of the centromeric and pericentromeric *180bp*, *106B*, and *TSI* repeat sequences (Supplemental Figure 4E) and of several transposable elements (list of tested loci available in Methods) was unchanged in *atrx-1*,

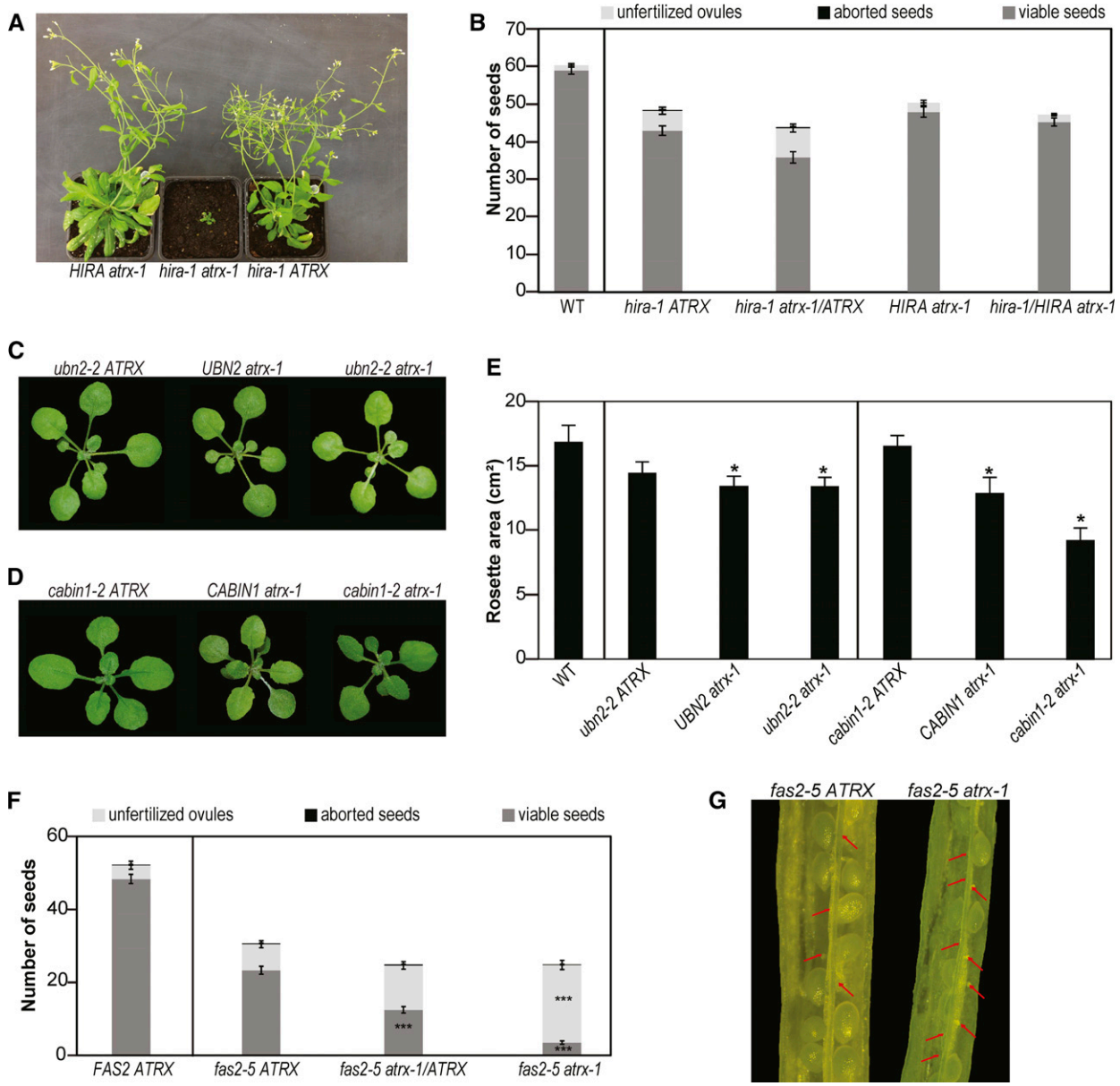


Figure 3. Epistatic Relationship between ATRX and CAF-1 or HIR Histone Chaperone Complexes.

(A) Representative 7-week-old F3 sister plants grown on soil.

(B) Quantification of seed content in *hira-1*, *hira-1 atrx-1/ATRX*, *atrx-1*, and *atrx-1 hira-1/HIRA* mutant siliques. Mean seed number was calculated from at least 30 siliques pooled from four plants grown at the same time. A Student's *t* test revealed no statistically significant difference between the genotypes.

(C) and **(D)** Representative 3-week-old F3 plants derived from crosses of *atrx-1* alleles with *ubn2-2* **(C)** or *cabin1-2* **(D)** grown on soil.

(E) Quantification of total rosette surface area of the F3 sister plants described in **(C)** and **(D)**. Mean rosette area was calculated from at least six 2-week-old plants for each genotype. Student's *t* test in comparison to *ubn2-2 ATRX* for *UBN2 atrx-1* and *ubn2-2 atrx-1* mutants and in comparison to *cabin1-2 ATRX* for *CABIN1 atrx-1* and *cabin1-2 atrx-1* mutants; *$P < 0.05$.

(F) Quantification of seed content in siliques of *FAS2 ATRX*, *fas2-5 ATRX*, *fas2-5 atrx-1/ATRX*, and *fas2-5 atrx-1* F2 sister plants. Mean for seed number was calculated from at least 30 siliques pooled from four plants grown at the same time. Student's *t* test in comparison to *fas2-5 ATRX*; ***$P < 0.001$.

(G) Representative dissected siliques from *fas2-5 ATRX* and *fas2-5 atrx-1* F2 sister plants. Red arrows indicate unfertilized ovules.

Error bars of all panels represent SE of the mean.

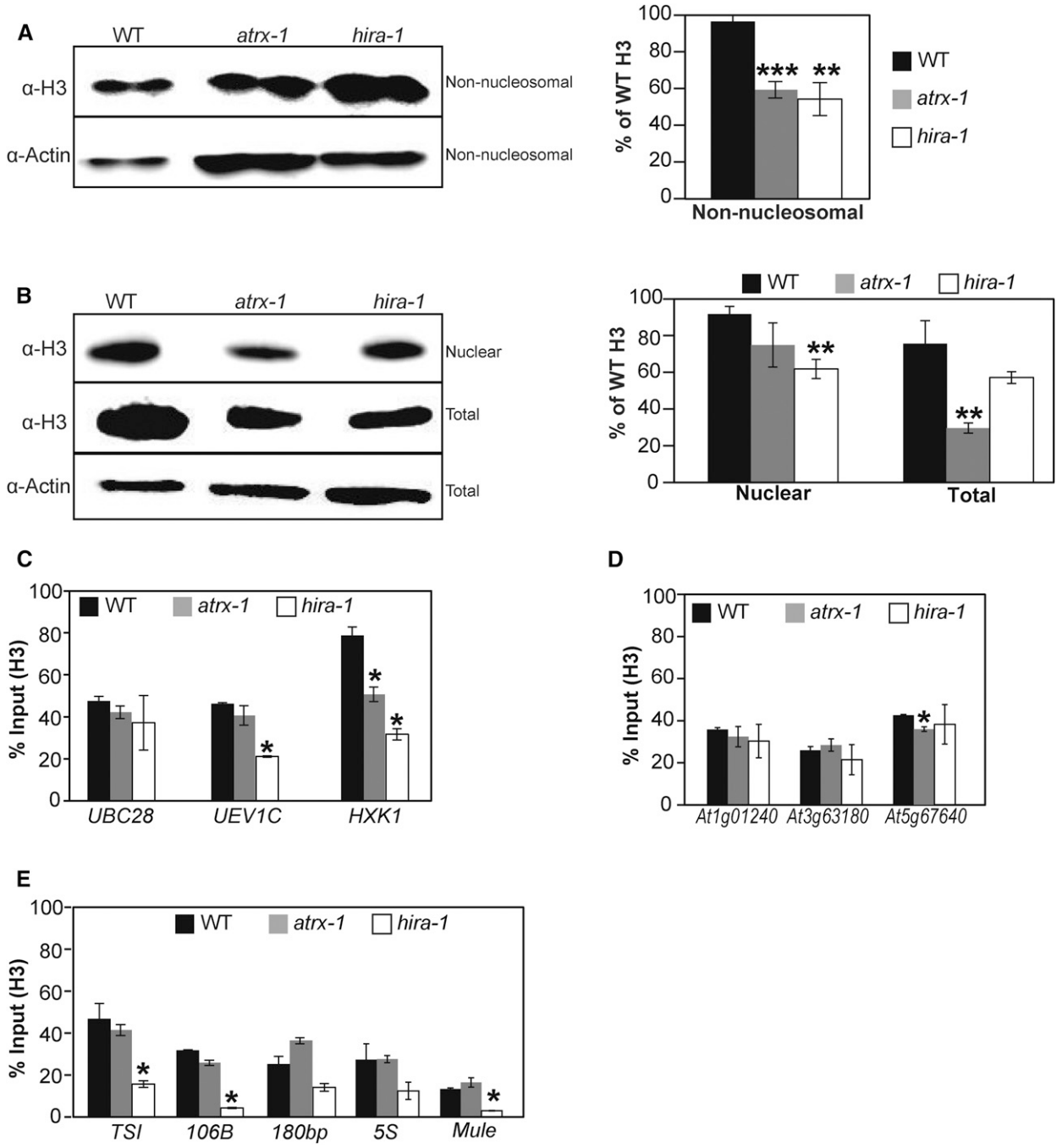


Figure 4. Effects of ATRX Loss on Histone Pools and Nucleosome Occupancy.

(A) and **(B)** Left: Histone H3 protein levels quantified by immunoblot in non-nucleosomal fractions **(A)** and nuclear and total extracts **(B)**. Twenty **(A)** or six **(B)** micrograms of proteins were loaded per lane. Right: Quantification of H3 band intensities normalized to Actin from two independent experiments, each comprising two biological replicates for each genotype, consisting of independent pools of 1 g of 2.5-week-old in vitro-grown plantlets collected at the same time, and several blots. Student's *t* test compared with the wild type; **P < 0.01 and ***P < 0.001.

(C) to **(E)** Histone H3 occupancy at three active genes (*UBC28*, *UEV1C*, and *HXK1*) **(C)**, at three genes situated in subtelomeric regions (*At1g01240*, *At3g63180*, and *At5g67640*) **(D)**, and at centromeric and pericentromeric repeats (*TSI*, *106B*, *180bp*, and ribosomal 5S rDNA loci) and at a transposon on a chromosome arm (*At2g15810*, *Mule*) **(E)** was assessed by H3-ChIP qPCR in three biological replicates consisting of pools of 1 g of 2.5-week-old in vitro-grown wild-type, *atrx-1*, and *hira-1* mutant plants. Student's *t* test compared with the wild type; *P < 0.05.

Error bars of all panels represent SE of the mean.

except for the *Mule* transposon (Supplemental Figure 4F) and three Gypsy LTR transposons GP2NLTR, Atlantys2_I, and Athila2_I that were reactivated in *atrx-1* in our transcriptome analysis presented below (Supplemental Data Set 1).

Therefore, loss of ATRX affects histone pools and has a moderate influence on histone H3 occupancy at certain active genes. In contrast to mammals, ATRX in Arabidopsis is required neither for H3 incorporation at the tested pericentromeric regions nor to maintain transcriptional silencing at repeats and transposable elements.

Arabidopsis ATRX Influences Histone H3.3 Deposition

Arabidopsis canonical histones H3.1 and variants H3.3 differ by only four amino acids, and no variant-specific antibodies are available so far in plants. Hence, to follow specifically each H3 histone type, we used plants expressing FLAG-HA-tagged versions of H3.1 (epitope-tagged H3.1, eH3.1) or H3.3 (eH3.3) under the control of their endogenous promoters. We first validated chromatin incorporation of eH3.1 and eH3.3 proteins by ChIP using a FLAG antibody. As expected, eH3.1 was preferentially enriched at heterochromatic loci (*180bp*, *TSI*, and the *Ta3* retrotransposon localized in the pericentromeric region of chromosome 1), while transcriptionally active genes were enriched in eH3.3 (*UEV1C*, *UBC28*, and *HXK1*) (Figure 5A). We then crossed the eH3.1 and eH3.3 transgenic lines with plants carrying the *atrx-1* allele. From the progeny, we selected sister plants with and without the *atrx-1* mutation that expressed sufficient eH3.1 or eH3.3 for further analysis. We chose two lines for each genotype with similar transgene expression levels (Supplemental Figures 5A and 5B) to ensure that changes in eH3.1 and eH3.3 abundance between wild-type and *atrx-1* mutant plants were caused by ATRX loss and not by differential expression of the transgene.

We first determined the amount of eH3.1 and eH3.3 in total and nuclear histone pools by immunoblot. Since the α-H3 antibody recognized endogenous and epitope-tagged H3.1 and H3.3, blots were hybridized with a α-H4 antibody for normalization. The quantification of H4 levels confirmed the moderate reduction of the H3-H4 pool in the total protein extract in both *atrx-1* eH3.1 and *atrx-1* eH3.3 lines (Figures 5B and 5C). The eH3.1/H4 ratio was increased in the total extract of *atrx-1* eH3.1 mutants compared with the wild type (Figure 5B), while the eH3.3/H4 ratio was reduced in both total and nuclear extracts of *atrx-1* eH3.3 compared with the wild type (Figure 5C). Based on these results, we conclude that storage and deposition of H3.3 are affected by the absence of ATRX.

To analyze the chromatin distribution of H3.1 or H3.3, a ChIP coupled to quantitative PCR was performed on plants expressing eH3.1 or eH3.3. A ChIP performed with the H3 antibody revealed that *atrx-1* eH3.1 and *atrx-1* eH3.3 display similar H3 enrichment patterns (Supplemental Figures 5C and 5D) as previously observed for *atrx-1* (Figures 4C to 4E). Then, to determine potential changes in the incorporation of eH3.1 and eH3.3, we performed a ChIP with a FLAG antibody using wild-type plants as a control for background noise (Figures 5D and 5E). We evaluated eH3.1 and eH3.3 levels at heterochromatic regions (*TSI*, *106B*, and *180bp*) and at *5S* rDNA. In *atrx-1* mutants, the eH3.1 abundance was maintained at *TSI*, *180bp*, and *5S* rDNA regions; only *106B* repeats displayed a higher abundance of eH3.1 (Figure 5D). In contrast to eH3.1, eH3.3 was significantly less enriched at *TSI* and *5S* rDNA regions in *atrx-1* mutants (Figure 5E). Furthermore, analysis of three transcriptionally active genes (*UBC28*, *UEV1C*, and *HXK1*) and three genes in subtelomeric regions revealed a significant reduction of eH3.3 occupancy at these targets except for *At5g67640* (Figure 5F).

Taken together, our results from immunoblot and ChIP-qPCR analyses were consistent with the hypothesis that Arabidopsis ATRX is involved in H3.3 deposition.

ATRX Controls Gene Expression and H3.3 Distribution

To determine how the distribution of H3.3 is affected by ATRX loss of function at a genome-wide scale, we performed ChIP-seq on chromatin prepared from whole seedlings expressing the eH3.3 transgene. The genome-wide distribution of eH3.3 along the chromosomes (Supplemental Figure 6A, left) showed the expected profile in the wild type (Stroud et al., 2012; Wollmann et al., 2012; Vaquero-Sedas and Vega-Palas, 2013; Shu et al., 2014), the profile being similar in *atrx-1* mutants (Supplemental Figure 6A, right). We therefore assessed changes in eH3.3 peak summit occupancy between the wild type and *atrx-1* using DANPOS2 (Chen et al., 2013). We identified 54,836 genomic peaks with significantly different eH3.3 occupancy in *atrx-1* mutants compared with the wild type (Supplemental Data Set 2, smt_diff_FDR < 0.05), out of which 66% correspond to a reduced occupancy in *atrx-1* (Figure 6A), as expected for a H3.3 chaperone. Upon ATRX loss, reduced eH3.3 occupancy was observed at regions of medium to high eH3.3 occupancy, while increased eH3.3 occupancy was found at regions of low eH3.3 occupancy (Figure 6B). This suggests that ATRX affects regions with different levels of H3.3 occupancy in a distinct manner. Given that in mammals, ATRX/DAXX was shown to control silencing of retrotransposons notably through the incorporation of H3.3 (Elsässer et al., 2015), we analyzed changes in eH3.3 incorporation at transposable elements (TEs). In Arabidopsis, TEs are primarily localized in the pericentromeric regions, which are H3.3 poor (Supplemental Figure 6A, left). Only 29% of TEs showed differential H3.3 occupancy, and 43% of these TEs presented reduced H3.3 occupancy in *atrx-1* (Supplemental Figure 6B and Supplemental Data Set 3). Since 70% of the genomic peaks with eH3.3 occupancy changes localized to gene bodies, we further focused on the effect of ATRX loss on H3.3 incorporation at genes. Among the 38,645 peaks located in 19,396 gene bodies, 75% corresponded to a reduced H3.3 occupancy in *atrx-1*, confirming the variations in eH3.3 occupancy observed by ChIP-qPCR (Figure 6C; Supplemental Figures 6C and 6D). The distribution was similar for the few (0.07%) intergenic regions (Supplemental Figure 6E and Supplemental Data Set 4) that showed differential eH3.3 occupancy. Therefore, ATRX appears to preferentially control H3.3 occupancy at genic sites.

H3.3 occupancy is known to be correlated with gene expression (Stroud et al., 2012; Wollmann et al., 2012; Shu et al., 2014). Hence, we analyzed how the defective H3.3 incorporation upon ATRX loss influences gene expression by performing RNA-seq on rRNA-depleted total RNAs prepared in triplicates from whole 2.5-week-old wild-type and *atrx-1* seedlings. We identified 456 down- and 346 upregulated genes in *atrx-1* compared with wild-type plants

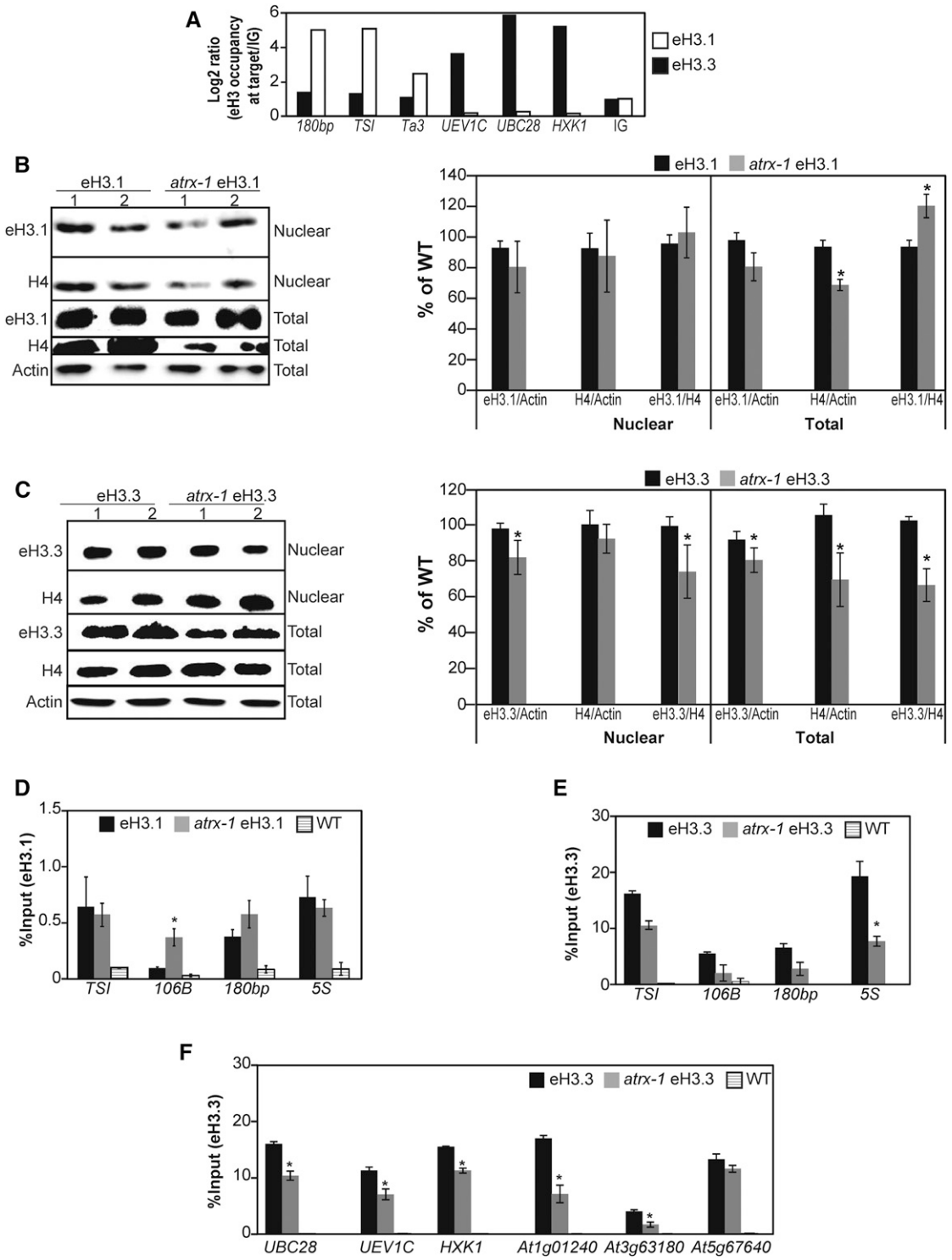


Figure 5. Effect of ATRX Loss on Incorporation and Balance of the Canonical Histone H3.1 and the H3.3 Variant.

(A) Canonical eH3.1 and variant eH3.3 occupancy at heterochromatic repeats (*180bp*, *TSI*, and *Ta3*) and at three active genes (*UBC28*, *UEV1C*, and *HXK1*) assessed by ChIP-qPCR in one biological replicate sampled at the same time for each genotype and consisting of a pool of 0.5 g of 2.5-week-old in vitro-grown plants. The eH3.1 and eH3.3 occupancy is normalized to the occupancy at an intergenic region (IG; set to 1).

(B) and **(C)** Left: Histone eH3.1 **(B)** and eH3.3 **(C)** protein levels quantified by immunoblot in nuclear fractions and total extracts prepared with Honda buffer (Honda et al., 1966). Six micrograms of proteins were loaded per lane. Right: Quantification of H4 and eH3.1 **(B)** or eH3.3 **(C)** band intensities relative to Actin

($|\log_2$fold change$|>0$ and FDR < 0.1; Figure 6D; Supplemental Data Set 5). Enriched Gene Ontology (GO) terms were determined to uncover modulated cellular components (Figure 6E). While terms such as response to stress as well as abiotic and chemical stimuli were enriched among upregulated genes, the term "ribosome" emerged as strongly enriched in downregulated genes (71 genes out of 456 corresponding to the GO term "ribosome"). Compared with all the genes expressed in the RNA-seq data set (Figure 6F), downregulated genes in *atrx-1* belonged to medium and highly expressed genes (84% with transcript levels >100 reads per kilobase per million mapped reads [RPKM]), while 88% of upregulated genes showed transcript levels between 1 and 100 RPKM (Figure 6G). This suggests a role for ATRX in the positive regulation of highly expressed genes.

We then asked whether eH3.3 occupancy changes induced by ATRX loss are associated with gene expression changes. For this purpose, we first looked at the *ATRX* locus itself, as no *ATRX* transcripts were observed downstream of the T-DNA insertion site in *atrx-1* mutants (Supplemental Figure 2B). Indeed, eH3.3 occupancy is significantly reduced in *atrx-1* compared with the wild type in this region (Figure 6H). This prompted us to analyze in detail genes that showed significantly increased or decreased eH3.3 occupancy in *atrx-1* and which were differentially expressed in the RNA-seq data set. It appears that 278 out of 456 downregulated genes (Figure 6I) and 267 out of 346 upregulated genes (Figure 6J) presented significantly altered eH3.3 occupancy in our ChIP-seq data, and both sets of genes showed similar H3.3 occupancy distribution in wild-type plants. The vast majority of the downregulated genes (88%) showed reduced eH3.3 occupancy in *atrx-1* (Figure 6I) and were characterized by a medium to high gene expression level (Figure 6K), suggesting that modulation of the expression of these genes may be associated with an ATRX-dependent deposition of H3.3. Unexpectedly, most upregulated genes also showed reduced H3.3 occupancy (Figure 6J). These genes differed from the downregulated ones by their low expression level (Figure 6K), implying that, at this set of genes, H3.3 occupancy does not positively regulate gene expression and that other mechanisms dominate.

To conclude, ATRX controls H3.3 occupancy and fine-tunes gene expression of moderately to highly expressed genes.

Arabidopsis ATRX Maintains H3 Occupancy and Expression at 45S rDNA Genes

Many riboprotein-coding genes were downregulated in our transcriptome analysis in *atrx-1* mutants, suggesting altered ribosome biogenesis in the mutant. Besides riboproteins, major components of the ribosome are the structural *18S*, *5.8S*, and *25S* rRNAs encoded by repetitive arrays of 45S ribosomal genes, which are organized in nucleolus organizer regions (NORs) that are highly enriched in H3.3 (Shi et al., 2011; Shu et al., 2014). We therefore analyzed histone occupancy by H3-ChIP in *atrx-1* mutants and found that histone occupancy was reduced at the 45S rDNA loci (Figure 7A). Given that the formation of a repressive chromatin environment is required to silence 45S rRNA genes since only 10 to 27% of 45S rDNA genes are transcribed (Pontes et al., 2003; Lawrence et al., 2004), we asked whether reduced histone occupancy leads to altered 45S rRNA expression and nuclear organization. We found that despite reduced histone H3 occupancy, 45S pre-rRNA levels (Figure 7B) were not significantly altered in *atrx* potentially due to compensation by increased H3K9me2 enrichment relative to H3 at the 45S rDNA loci (Figure 7C). Furthermore, in *atrx-1*, the repetitive 45S rDNA arrays remained clustered into highly condensed chromatin regions termed chromocenters (Supplemental Figure 7A). The *atrx-1* nuclei showed an increased number of NORs (Supplemental Figure 7B), which were associated with the nucleolus, in agreement with increased endoreduplication and chromocenter numbers observed in *atrx* mutant cotyledons (Supplemental Figures 7C and 7D).

While the quantity of 45S pre-rRNA transcripts was not affected, altered chromatin structure could influence the regulation of the different 45S sequence variants. These variants (VAR1 to 4) are defined by small sequence variations in the 3′ external transcribed spacer region and can be separated by size after amplification with a primer pair common to all variant types (Pontvianne et al., 2010) (Figure 7D, top panel). In our conditions, mainly VAR2 and VAR3 were transcribed in wild-type plants (Abou-Ellail et al., 2011; Pontvianne et al., 2010), while VAR2 was predominantly expressed in both *atrx-1* and *atrx-2* mutant alleles (Figure 7D, bottom panel; Supplemental Figure 7E). To explain this observation, we determined the relative abundance of the rDNA variants, as well as their enrichment in H3 or H3K9me2, by submitting ChIP samples to PCR (Figures 7E and 7F). In the wild type, VAR1 is enriched in H3 and H3K9me2 when compared with the ChIP input DNA, while VAR3 is depleted (Figure 7F). This finding is in agreement with previous observations showing that VAR1 is transcriptionally silent in wild-type plants (Pontvianne et al., 2012). When comparing wild-type and *atrx-1* mutants, we noticed a similar pattern for VAR1, but found that VAR2 was more abundant in the *atrx-1* genome than in wild-type plants, while VAR3 was underrepresented (Figure 7F). This illustrates a change in the relative variant abundance in *atrx-1* mutants, while 45S rDNA copy numbers remained unchanged (Supplemental Figure 7H). Furthermore, the

Figure 5. (continued).

in total extracts from two independent experiments comprising in total four biological replicates of pools of 1 g of 2.5-week-old in vitro-grown plants and several blots. Student's *t* test compared with the wild type; *P < 0.05.

(D) and **(E)** Histone eH3.1 **(D)** and eH3.3 **(E)** occupancy at heterochromatic repeats (*TSI*, *106B*, and *180bp*) and at the *5S* ribosomal DNA loci assessed by ChIP-qPCR in three biological replicates sampled at the same time and consisting of pools of 1 g of 2.5-week-old in vitro-grown eH3.1, *atrx-1* eH3.1, and wild-type plants. Student's *t* test compared with the wild type; *P < 0.05.

(F) Histone eH3.3 occupancy at three active genes (*UBC28*, *UEV1C*, and *HXK1*) and at three genes situated in subtelomeric regions (*At1g01240*, *At3g63180*, and *At5g67640*) assessed by ChIP-qPCR in the same plant material as in **(E)**. Student's *t* test compared with the wild type; *P < 0.05.

Error bars of all panels represent SE of the mean.

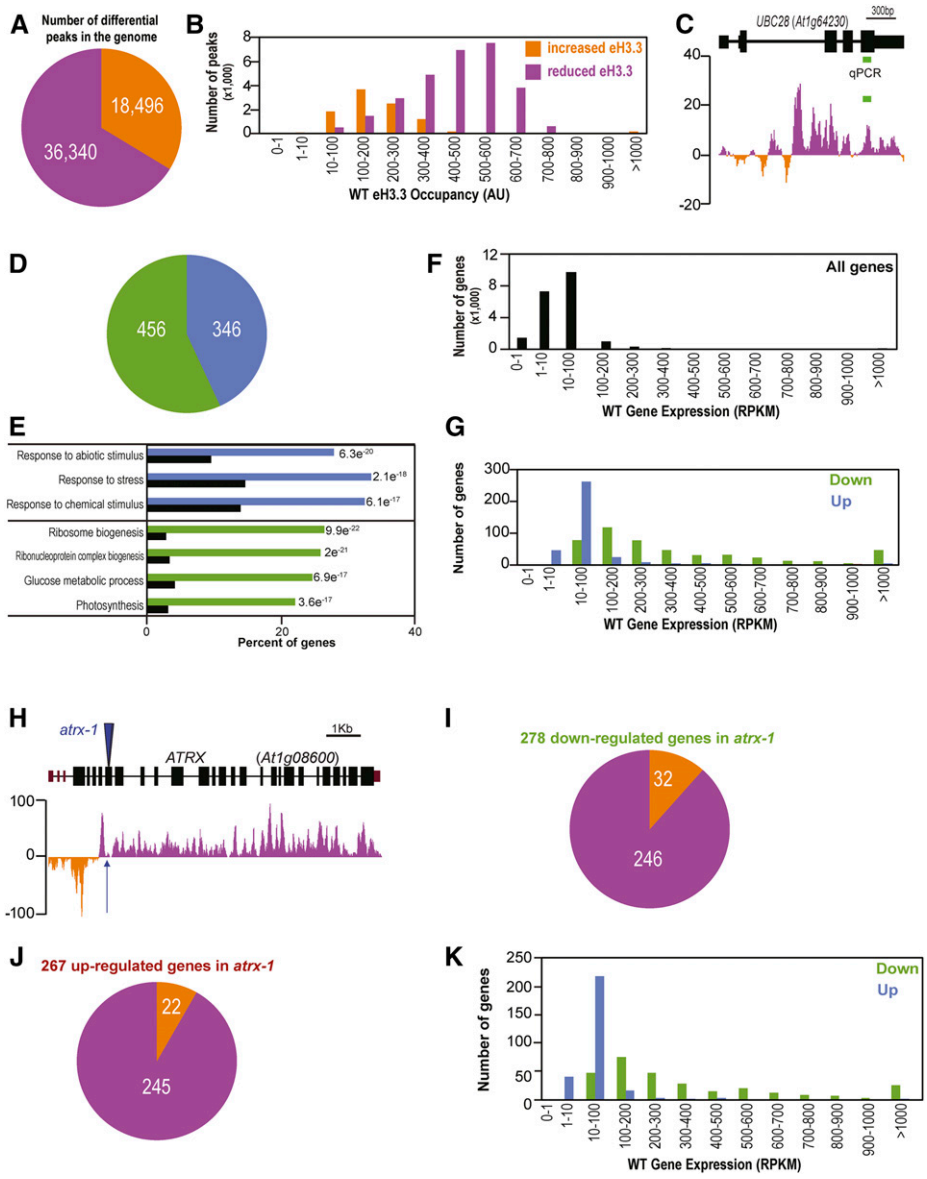


Figure 6. Effects of ATRX Loss on Gene Expression and Histone Variant H3.3 Occupancy.

(A) Pie chart showing the number of peaks situated in the Arabidopsis genome with increased (orange) or reduced (violet) eH3.3 occupancy in *atrx-1* relative to the wild type. Numbers of peaks are indicated in the pie chart. ChIP-seq was performed in two biological replicates for each genotype sampled at the same time consisting of pools of 2 g of 2.5-week-old in vitro-grown eH3.3 and *atrx-1* eH3.3 plants. Peak occupancy was determined with DANPOS2.

(B) Distribution of gene-localized peaks according to their increased (orange) or reduced (violet) eH3.3 occupancy in the wild-type ChIP-seq data set. Peak occupancy was determined with DANPOS2 (AU, arbitrary units).

(C) Differential eH3.3 summit occupancy calculated at single nucleotide resolution in the ChIP-seq data sets in comparison to ChIP-qPCR data from Figure 5F. The active gene *UBC28* is presented. Peaks with higher eH3.3 occupancy in *atrx-1* compared with the wild type are shown in orange (negative values), while those with lower eH3.3 occupancy are in violet (positive values). The region amplified in ChIP-qPCR is displayed by a green line. Reduced eH3.3 occupancy was monitored by ChIP-qPCR at this locus (Figure 5F).

(D) Pie chart showing the number of differentially expressed genes. Upregulated and downregulated genes in *atrx-1* compared with the wild type are displayed in blue and green, respectively, and numbers of differentially expressed genes are indicated in the pie chart. RNA-seq was conducted on total RNAs extracted from three biological replicates sampled at the same time consisting of pools of around twenty 2.5-week-old in vitro-grown wild-type and *atrx-1* plants.

(E) Top GO terms for upregulated and downregulated genes in *atrx-1* compared with the wild type. The P values are indicated. Terms for upregulated genes are in blue, while those for downregulated genes are in green. Background levels are shown in black. e, exponent (scientific format, in which "e" multiplies the preceding number by 10 to the -nth power).

H3/input and the H3K9me2/input ratios revealed reduced H3 and H3K9me2 enrichment at VAR2 in *atrx-1* (Supplemental Figures 7F and 7G). Together, this is consistent with the increased VAR2 transcript levels observed in the *atrx-1* mutant.

We then used the *atrx-1* eH3.1 and *atrx-1* eH3.3 lines to analyze whether the reduced H3 occupancy is caused by loss of H3.1 or H3.3. We first confirmed the reduced H3 occupancy at 45S rDNA in *atrx-1* compared with the respective wild-type lines (Supplemental Figures 7I and 7L). While no significant change in eH3.1 occupancy was observed at 45S rDNA loci (Figure 7G), eH3.3 occupancy was decreased (Figure 7J), suggesting that deficient H3.3 deposition caused the reduced nucleosomal occupancy at 45S rDNA loci in *atrx-1* mutants. When we looked at eH3.1 and eH3.3 enrichment at specific 45S variants, we noticed, compared with input, less eH3.3 enrichment at VAR1 as well as more eH3.3 and less eH3.1, respectively, at VAR2 in wild-type and *atrx-1* plants (Figures 7H, 7I, 7K, and 7L), in agreement with their relative expression patterns in these lines (Supplemental Figures 7J, 7K, 7M, and 7N). In the *atrx-1* eH3.1 and *atrx-1* eH3.3 lines, VAR2 abundance was similar in *atrx-1* DNA compared with the wild type (Figures 7H, 7I, 7K, and 7L), likely because these two lines are issued from a cross between plants expressing the epitope-tagged histones and an *atrx-1* mutant plant that led to homogenization. Nevertheless, VAR2 remained predominantly expressed in *atrx-1* eH3.1 and *atrx-1* eH3.3, respectively, compared with eH3.1 and eH3.3 plants (Supplemental Figures 7J, 7K, 7M, and 7N). This suggests that increased VAR2 abundance at the RNA level in *atrx-1* mutants results not only from increased VAR2 abundance at the DNA level, but may also be the consequence of ATRX loss of function.

To summarize, ATRX loss affects H3.3 occupancy and H3K9me2 enrichment at 45S rDNA repeats and leads to a modified variant dosage and altered 45S rRNA variant expression.

DISCUSSION

We investigated the role of ATRX in the model plant Arabidopsis using two independent mutant alleles in the single *ATRX* gene generated by T-DNA insertions. For both mutant alleles, and in contrast to loss of the mammalian ATRX that is embryo lethal, we found that Arabidopsis *atrx* mutants are viable. The *atrx* plants show reduced vigor, reduced vegetative growth, and reduced fertility, features that correlate well with the reduced riboprotein-coding gene expression.

Both *atrx* alleles, as well as transheterozygotes, result in lethality in combination with HIR complex mutations and affect rDNA expression, but dissimilarities concerning the seed set have been observed between the two *atrx* alleles. This might be explained by differences in the number of subsequent generations homozygous for the *atrx* mutation, by the presence of potential partial proteins still produced in the *atrx-2* allele, or by a second mutation. Normal seed set in F1 *ATRX/atrx-2* (Supplemental Figure 2H) ruled out the hypothesis of a partial ATRX protein acting as a dominant negative. The analysis of F1 *atrx-1/atrx-2* plants revealed a seed set similar to *atrx-1* but different from the *atrx-2* mutant (Supplemental Figure 2H). This rather supports the hypothesis of a second mutation in the *atrx-2* mutant allele, which seems to enhance the fertility defects, such as those observed in combination with *hira-1* (Supplemental Figure 3E). Hence, the knockout allele *atrx-1* was chosen for in-depth molecular analysis.

ATRX is a multifunctional protein carrying both an ADD domain involved in histone tail binding and a SWI/SNF helicase domain implicated in chromatin remodeling. Therefore, ATRX is likely to play a broader role in chromatin structure not necessarily linked to its potential histone chaperone function. Indeed, the moderate sensitivity to γ-irradiation observed in the mutant plants might be attributed to reduced nucleosomal occupancy as well as to the loss of chromatin remodeling activity (Shaked et al., 2006). One of the phenotypes we noted in *atrx-1* mutants is an aberrant cell cycle profile with enlarged peaks in the 2C and 4C populations, which could be explained by an S-phase defect. While we found no hypersensitivity to HU, which inhibits replication by diminishing the available nucleotide pool, we noticed more early S-phase nuclei relative to late ones in root tips, which could suggest a delayed progression through early S-phase. The binding of mammalian ATRX to genic and intergenic sites rich in variable number tandem repeats that can form G-quadruplexes (Law et al., 2010) may imply that ATRX is also needed for the resolution of such structures interfering with proper DNA replication in Arabidopsis. Further studies are needed to explore fully the role of ATRX in the replication process in plants.

ATRX as a Player in the Arabidopsis H3.3 Chaperone Network

Lethality and developmental defects observed in combination with mutations for HIR or CAF-1 histone chaperone complexes,

Figure 6. (continued).

(F) and **(G)** Distribution of all genes **(F)** and of differentially expressed genes in *atrx-1* (**[G]**, in blue for upregulated genes and in green for downregulated genes) according to their expression levels in the wild type RNA-seq data set (RPKM).

(H) Differential eH3.3 summit occupancy calculated at single nucleotide resolution at the *ATRX* locus from the ChIP-seq data set. Peaks with higher eH3.3 occupancy in *atrx-1* compared with the wild type are shown in orange, while those with lower eH3.3 occupancy are in violet. Peak occupancy was determined with DANPOS2. The blue arrow indicates the T-DNA insertion location for the *atrx-1* mutant.

(I) Pie chart showing distribution of genes with differential summit eH3.3 occupancy determined with DANPOS2 in *atrx-1* eH3.3 in the ChIP-seq data, for the 278 downregulated genes identified by DeSeq2 in *atrx-1* in the RNA-seq data. The 246 genes with reduced eH3.3 occupancy in *atrx-1* relative to the wild type are in violet. The 32 genes with increased eH3.3 occupancy in *atrx-1* relative to the wild type are in orange.

(J) Pie chart showing distribution of genes with differential summit eH3.3 occupancy determined with DANPOS2 in *atrx-1* eH3.3 in the ChIP-seq data, for the 267 upregulated genes identified by DeSeq2 in *atrx-1* in the RNA-seq data. The 245 genes with reduced eH3.3 occupancy in *atrx-1* relative to the wild type are in violet. The 22 genes with increased eH3.3 occupancy in *atrx-1* relative to the wild type are in orange.

(K) Distribution of genes downregulated (green) and upregulated (blue) in *atrx-1* according to their expression levels in the wild-type RNA-seq data set. Differential expression was determined with DeSeq2.

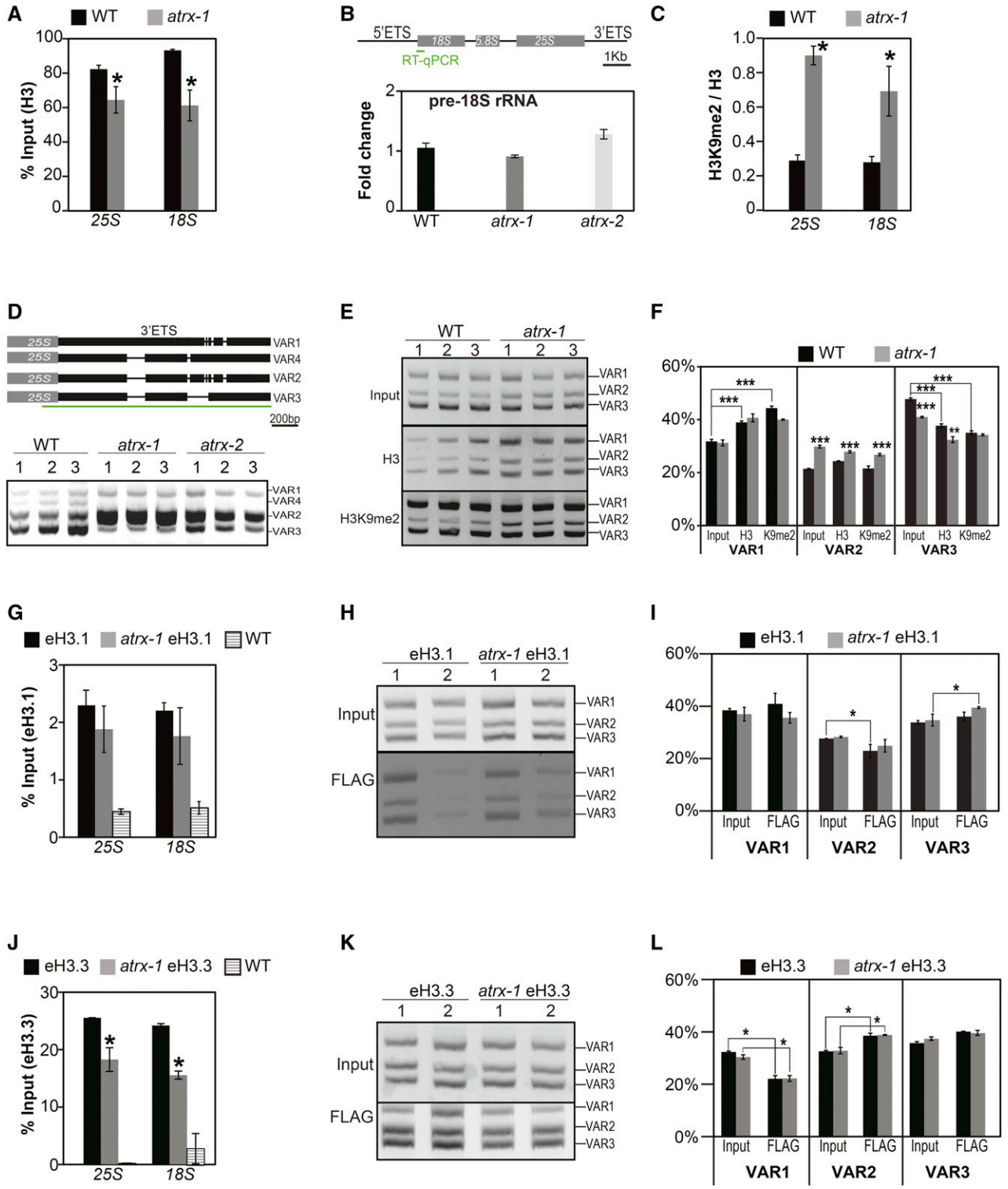


Figure 7. Effects of ATRX Loss on Nucleosome Occupancy and Balance between Histone Variants at 45S rDNA Loci.

(A) Mean histone H3 occupancy at *25S* and *18S* rDNA loci assessed by H3-ChIP qPCR from three biological replicates consisting of independent pools of 1 g of 2.5-week-old in vitro-grown wild-type and *atrx-1* mutant plants collected at the same time. Student's *t* test; *P < 0.05.

(B) Quantification of pre-rRNA transcripts by qRT-PCR. Top: 45 rDNA unit with the *18S*, *5.8S*, and *25S* rRNA genes (gray boxes) with positions of the external transcribed spacers (5′ETS and 3′ETS). The region amplified in qRT-PCR is displayed by a green line. Bottom: Quantification of pre-rRNA transcripts by

along with the reduced amounts of H3.3 histones in different cellular pools and in chromatin, strongly suggest that ATRX functions as a histone H3.3 chaperone in Arabidopsis. The most severe phenotypes were observed in crosses with mutants for the HIR complex that is involved in the replication-independent histone deposition in mammals (Tagami et al., 2004) and that interacts with H3.3 in Arabidopsis (Nie et al., 2014). Therefore, in the absence of a functional HIR complex, ATRX may be required as an alternative pathway of H3.3 deposition, recalling observations in Drosophila: *hira* and *xnp* single mutants are viable and reveal only moderate defects in H3.3 deposition, while double mutants die during larval development and are unable to assemble H3.3 into chromatin (Schneiderman et al., 2012). In *hira* mutants, H3 occupancy is affected at many genomic locations including active genes as well as transcriptionally repressed centromeric and pericentromeric repeats or transposons (this study; Duc et al., 2015). By contrast, in *atrx* mutants, H3 occupancy changes are moderate compared with *hira* and the majority of loci with differential H3.3 occupancy identified in the ChIP-seq data set (70%) situate in genes. Therefore, ATRX is rather implicated in deposition of the histone variant H3.3 in genes and may function as a fail-safe mechanism for impaired HIR-mediated H3.3 deposition, which is consistent with the recent finding that a complete loss of H3.3 through the disruption of the three H3.3-encoding genes is lethal in Arabidopsis (Wollmann et al., 2017). Within the mammalian ATRX/DAXX complex, DAXX binds histone H3.3 and determines the variant specificity of the heterocomplex (Liu et al., 2012; Elsässer et al., 2012), while ATRX guides the heterocomplex to specific genomic sites through interaction with methylated histone tails. Arabidopsis ATRX contains a conserved ADD domain and interacts with H3 histones in Y2H assays and in planta. To date, no DAXX homolog could be identified in the Arabidopsis genome based on protein sequence homology (this study; Otero et al., 2014), consistent with the absence of the mammalian DAXX binding domain in Arabidopsis ATRX. The Arabidopsis ATRX therefore might affect H3.3 deposition by interacting with a functional homolog of DAXX or with yet unknown partners. The identification of these binding partners will be an interesting avenue to better understand ATRX function in Arabidopsis. An alternative hypothesis that cannot be completely excluded is that ATRX affects H3.3 occupancy in an indirect manner, e.g., by modulating chromatin organization due to its remodeling activity or via changes in the enrichment in other core histones. Indeed, in mammals, ATRX coimmunoprecipitates with macroH2A, and ATRX loss leads to increased macroH2A enrichment at specific genomic sites (Ratnakumar et al., 2012). Furthermore, genome-wide analysis of the distribution of ATRX and DAXX proteins in embryonic stem cells revealed only a relatively small percentage of common binding sites (He et al., 2015), suggesting that ATRX and DAXX may have independent functions at certain targets.

ATRX and Heterochromatin

In mammals, ATRX binds to repetitive satellite sequences and is involved in heterochromatin formation and silencing of certain transposable elements (He et al., 2015; Sadic et al., 2015; Voon et al., 2015). This is in agreement with the observation of H3.3 enrichment at pericentromeric domains in mouse embryonic stem cells (Drané et al., 2010; Goldberg et al., 2010; Wong et al., 2010) and at specific classes of transposable elements (Elsässer et al., 2015; He et al., 2015; Voon et al., 2015). In *atrx* mutant plants, however, transcriptional silencing of repetitive elements or transposons is not impaired, as monitored by qRT-PCR and RNA-seq. This discrepancy might be resolved by the fact that H3.3 marks active chromatin and telomeres in Arabidopsis, while repetitive elements and transposons are globally enriched in H3.1 and not in H3.3 (Shu et al., 2014; Vaquero-Sedas and Vega-Palas, 2013; Stroud et al., 2012; Wollmann et al., 2012; this study). Furthermore, in mammals, the localization of ATRX to heterochromatin requires binding to methylated H3K9 via the ADD domain (Dhayalan et al., 2011; Iwase et al., 2011; Eustermann et al., 2011). While we found that the overall structure of the ADD domain is conserved in Arabidopsis ATRX and that ATRX can bind histones, the amino acids Y203, Y204, and Q219, critical for binding to H3K9me3 (Iwase et al., 2011), are not conserved (Supplemental Figure 1A). Therefore, it can be envisaged that, in plants, ATRX may not be targeted specifically to heterochromatin but may rather localize to regions with ongoing nucleosome

Figure 7. (continued).

qRT-PCR on cDNA prepared from three biological replicates consisting of independent pools of about twenty 2.5-week-old in vitro-grown wild-type, *atrx-1*, and *atrx-2* plantlets collected at the same time. Wild type was set to 1.

(C) Mean ratio of histone H3K9me2/H3 occupancy at *25S* and *18S* rDNA loci assessed by ChIP-qPCR from three biological replicates consisting of independent pools of 1 g of 2.5-week-old in vitro-grown wild-type and *atrx-1* mutant plants collected at the same time. Student's *t* test; *P < 0.05.

(D) Analysis of 45S rRNA variants. Top: Schema of 3′ETS rRNA gene variants. Black lines joining rectangles indicate deletions in sequences of each rRNA gene variant. The green line represents the amplified region presented in the bottom panel. Bottom: RT-PCR on cDNA, prepared from three biological replicates consisting of independent pools of about twenty 2.5-week-old in vitro-grown wild-type, *atrx-1*, and *atrx-2* plants grown and collected at the same time. Detection of pre-rRNA transcripts was performed with primers enclosing the external transcribed 45S spacer and detecting all variants. The amplified region is displayed by a green line in the top panel.

(E) Analysis of 45S variant abundance in the input, H3-ChIP, and H3K9me2-ChIP (K9me2) fractions presented in **(A)** and **(C)**.

(F) Quantification of relative band intensities presented in **(E)**. Student's *t* test; **P < 0.01 and ***P < 0.001.

(G) and **(J)** Mean histone eH3.1 **(G)** and eH3.3 **(J)** occupancy at *25S* and *18S* rDNA loci from three biological replicates consisting of independent pools of about 1 g of 2.5-week-old in vitro-grown eH3, *atrx-1* eH3, and wild-type plants assessed by ChIP-qPCR.

(H) and **(K)** Analysis of the relative 45S variant abundance in the input and Flag-ChIP samples in **(G)** and **(J)**.

(I) and **(L)** Quantification of relative band intensities presented in **(H)** and **(K)**. Student's *t* test; *P < 0.05.

Error bars of all panels represent SE of the mean.

displacement (Schneiderman et al., 2009) or bind similarly to the HIR complex directly to nucleosome-free gaps in DNA to direct histone deposition (Schneiderman et al., 2012; Ray-Gallet et al., 2011).

ATRX and Chromatin Organization at 45S rDNA Loci

Chromatin modifications are known to modulate rRNA expression in various species (Sandmeier et al., 2002; McStay and Grummt, 2008; Grummt and Längst, 2013). Indeed, in most eukaryotes, 45S rRNA genes are present in excess, resulting in selective silencing of a subset of genes by epigenetic mechanisms involving histone modifications (Preuss and Pikaard, 2007; Pontvianne et al., 2012, 2013). In Arabidopsis, the 45S rDNA loci are highly enriched in H3.3, and its presence correlates with transcriptional activity of the loci (Shi et al., 2011; Shu et al., 2014). Loss of ATRX affects H3.3 occupancy at 45S rDNA loci, thereby leading to reduced nucleosomal occupancy. This is in agreement with genome-wide studies in mammals, which identified ATRX at ribosomal genes (Law et al., 2010; He et al., 2015). Notably, H3K9me2 levels are globally increased at 45S rDNA in *atrx-1*, and we could speculate that the increased amount of this mark compensates for reduced nucleosomal occupancy. The existence of 45S rDNA sequence variants in the Arabidopsis genome that differ by small indels (Pontvianne et al., 2010) allows to evidence changes in rRNA gene choice. Indeed, it has been previously reported that VAR1, which is the most abundant at the DNA level, is silenced in Col-0 wild-type plants. This is in agreement with higher histone occupancy and H3K9 methylation, along with reduced H3.3 occupancy when normalized to input, the opposite being observed for VAR3, corroborating their expression profiles. While the relative abundance of VAR1 rDNA as well as its enrichment in H3K9me2/H3 is unchanged, expression shifts to increased VAR2 and reduced VAR3 rRNA in the *atrx* mutant. This might be partly explained by increased VAR2 and decreased VAR3 abundance in the *atrx-1* genome, as well as reduced histone occupancy and a lower H3K9me2 enrichment at VAR2. Changes in variant abundance have previously been observed in different mutants, such as *nucleolin2* (Durut et al., 2014), *atxr5 atxr6* (Mohannath et al., 2016; Pontvianne et al., 2012), or CAF-1 mutants (Mozgová et al., 2010; Pavlištová et al., 2016), and have been suggested to occur through gene conversion or selective rereplication coupled to intra-NOR recombination. Since 45S rRNA genes may form complex DNA structures prone to replication defects (Muchová et al., 2015), we could speculate that replication defects in *atrx* contribute to changes in relative 45S rDNA variant abundance. Given the downregulation of many ribosomal proteins in *atrx-1* concomitant with different 45S rDNA variant expression, it would be of interest to explore further whether this observation hints toward communication between the different transcriptional machineries involved in ribosome synthesis (Laferté et al., 2006).

Role of ATRX in Genomic H3.3 Distribution and Gene Expression

The amount of the variant histone H3.3 relative to H4 is reduced in all analyzed cellular pools, likely due to the combination of a reduced chaperone-bound histone fraction, which is shuttled for degradation, and deficient H3.3 deposition. At the nucleosomal level, H3.3 occupancy is reduced in *atrx* at a genome-wide level in particular within genes, pointing out a role for ATRX in histone deposition. Most of these genes with reduced H3.3 occupancy belong to classes with medium to elevated H3.3 levels, suggesting that ATRX contributes to H3.3 deposition at regions globally rich in H3.3 and corresponding to key metabolic players such as genes coding for ribosomal proteins. Differential H3.3 enrichments were globally modest in comparison to the H3.3 differences observed at the *ATRX* locus, which is equally affected in all cell types. This was expected since differences in H3.3 incorporation are likely to occur in a stochastic manner in agreement with our genetic analysis showing that ATRX is not the only pathway involved in H3.3 deposition. Alternatively, H3.3 incorporation may be mediated by ATRX only in specific cell types.

H3.3 occupancy has been globally correlated with gene expression even though it might not be indispensable for transcription (Wollmann et al., 2012, 2017; Stroud et al., 2012). For genes belonging to classes with medium to high expression, we observed that downregulation of gene expression is associated with reduced H3.3 occupancy in *atrx-1*, suggesting a role for ATRX in favoring gene expression at these loci. By contrast, the genes that are upregulated in *atrx-1* also predominantly display reduced H3.3 occupancy. Several explanations for this can be put forward, such as that H3.3 loss rather reflects reduced nucleosomal occupancy resulting from increased transcription and polymerase II abundance at these genes. However, a remarkable feature of these upregulated genes is their low expression while being enriched in H3.3, a characteristic recently described for certain subsets of genes (Liu et al., 2016). Hence, these genes are likely controlled by mechanisms different from those regulating constitutively expressed genes. Interestingly, upregulated genes notably correspond to genes associated with stress.

In conclusion, ATRX plays a major role in the maintenance of H3.3 variant occupancy at a genome-wide level at genes and at the 45S rRNA gene loci, which are H3.3-rich regions. Loss of ATRX affects H3.3 histone pools and H3.3 deposition and therefore modulates the H3.1/H3.3 balance in the cell. At 45S rRNA gene loci, ATRX loss results in altered 45S rRNA gene choice and at a genome-wide level in gene expression changes. Hence, the Arabidopsis ATRX protein shares some properties with its mammalian and invertebrate counterparts and functions in H3.3 deposition as part of a complementary pathway to HIR-mediated nucleosome assembly.

METHODS

Plant Material

Mutant *Arabidopsis thaliana* lines were obtained from the Nottingham Arabidopsis Stock Center and/or were gifts from other laboratories. We used the following mutant Arabidopsis lines *atrx-1* (SALK_025687), *atrx-2* (SAIL-861-B04), *fas2-5* (SALK_147693) (Duc et al., 2015), *hira-1* (WiscDsLox362H05) (Ingouff et al., 2010), *ubn2-2* (GABI_018D02), *cabin1-2* (SALK_099927), (Duc et al., 2015), the double mutant *atxr5* (SALK_130607) *atxr6* (SAIL_181_D09) (Jacob et al., 2009), *atr-2* (SALK_032841) (Rounds and Larsen, 2008), and the triple mutant *ku80* (FLAG_DMT5) *xrcc1* (SALK_125373) *xpf* (N3819) (Charbonnel et al., 2011) kindly provided by S. Amiard (GReD). Except *xrcc1 ku80 xpf*, all mutants are in the Columbia

background. Plants were grown on soil in a growth chamber under 16-h-light/8-h-dark cycles at 22°C (white fluorescent tubes, 100 to 130 μmol m^{-2} s^{-1}). The eH3.1 and eH3.3 lines were generated by transcriptional fusion of the genomic fragment containing the promoter and the genomic coding region of either *HTR9* or *HTR5* (stop codon excluded) with the FLAG-HA tag and the *OCS* (*octopine synthase*) terminator using classic cloning with restriction enzymes and the Gateway technology. Monolocus homozygous lines were selected based on segregation of hygromycin resistance also encoded in the transgene, and further crosses to *atrx-1* were performed. For in vitro culture, seeds were surface-sterilized and sown on germination medium containing 0.8% (w/v) agar, 1% (w/v) sucrose, and Murashige and Skoog salts (M0255; Duchefa Biochemie). After 2 d of stratification at 4°C in the dark, plants were grown under 16-h-light/8-h-dark cycles at 23°C (white fluorescent tubes, 120 μmol m^{-2} s^{-1}).

Phylogenetic Analysis

To identify the Arabidopsis ATRX ortholog, we performed interspecies BLAST searches with the mammalian protein sequence. Conserved domains were aligned using the program Muscle with the ESPript (Easy Sequencing in PostScript) program (Robert and Gouet, 2014) with default settings. Trees were constructed from protein sequences (Supplemental Data Set 6). The evolutionary history was inferred by using the maximum likelihood method based on the JTT matrix-based model (Jones et al., 1992). The tree with the highest log likelihood (−15825.7621) was used. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Distance bootstrap analyses consisted of 1000 replicates. Initial trees for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model and then selecting the topology with a superior log likelihood value. The tree is drawn to scale, with branch lengths measured in number of substitutions per site. The analysis involved 11 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 1002 positions in the final data set. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

Yeast Two-Hybrid Assay

The full-length cDNAs encoding histones H3.1 and H3.3 were cloned into the pGBKT7 vector as baits and the sequence encoding full-length Arabidopsis ATRX was cloned into the pGADT7 vector as prey. Vectors pGADT7 and pGBKT7 were respectively transformed into *Saccharomyces cerevisiae* Y187 and AH109 strains based on the manufacturer's instructions of the MatchMaker III GAL4 two-hybrid system (Clontech). Screening and interaction studies between preys and baits were performed by mating compatible yeast strains following Clontech's instructions. Interactions between ATRX and histones H3.1 and H3.3 were determined by growing transformants on media depleted of leucine, tryptophan, and histidine following the manufacturer's recommendations.

Confocal Images of ATRX-GFP in *Nicotiana benthamiana*

The full-length cDNA of *ATRX* was cloned behind the cauliflower mosaic virus 35S promoter in translational fusion with GFP at the C terminus into a plasmid using the Gateway technology (see Supplemental Table 4 for primer sequences). This plasmid was transiently expressed by agro-infiltration in *N. benthamiana* leaves with the viral silencing suppressor p19. Infected leaves were sectioned 3 to 5 d after infiltration. Infiltrated *N. benthamiana* leaves were imaged with Zeiss LSM 880 microscope (Carl Zeiss Vision) with a 40× objective.

Documentation of Phenotypes

Images of dissected siliques and flowers were taken using a Leica binocular and the LAS 3.6 software (Leica Microsystems), with 0.63× and 2× magnification, respectively. Viability of mature pollen grains was assayed as described (Alexander, 1969). Anthers from stained flowers were isolated and photographed using a Zeiss Axioplan microscope and Axiovision 4.2 software (Carl Zeiss Vision). For seed counting, we distinguished viable green seeds from unfertilized ovules as described (Duc et al., 2015). For rosette area measurement, digital photographs of double mutants and their respective single mutant sister plants were processed with the ImageJ (Fiji) software to measure rosette area (total area of rosette leaves).

RNA Extraction and RT-PCR

RNAs were extracted with Tri-Reagent (Euromedex) according to the manufacturer's instructions, then treated with RQ1 DNase I (Promega) and purified with phenol-chloroform extraction. Reverse transcription was done either with oligo$(dT)_{15}$ or with random hexamers using M-MLV reverse transcriptase (Promega). For analysis of *ATRX* expression in *atrx* alleles, reverse transcription was done with the ATRX_RT_Rev primer (5′-GGGACCCGTTGAACTTCCTCCC-3′) combined with random hexamers. Obtained cDNAs were diluted 1:3 and used in PCR (Promega Flexi) or in quantitative PCR with the LightCycler 480 SYBR Green I Master kit on the Roche LightCycler 480. Transcript levels of interest were normalized to *MON1* (*At2g28390*) (Czechowski et al., 2005) or to *UBC28* (*At1g64230*) for Supplemental Figure 2, using the comparative threshold cycle method. RT-qPCR histograms show means of transcript levels ± SE obtained for two independent PCR amplifications of three biological replicates. The *y* axis shows the fold change relative to the wild type (wild type set to 1) after normalization to *MON1* (*At2g28390*) expression. For analysis of silencing release, the following loci were tested: *Ta3* and *Athila* LTR pericentromeric retrotransposons, *Mule* transposons *At1g40097* and *At1g43280*, the *Mu1* (*At4g08680*) transposons and *SINE* retrotransposon *AtSN1* (*At3TE63860*) (Vaillant et al., 2006), the *CACTA1* transposon and the *T5L23.26 CACTA-like* transposon (located at the heterochromatin knob of chromosome 4) (Ono et al., 2006), and the *COPIA78* LTR transposon (Pecinka et al., 2010).

Analysis of 45S rDNA and rRNA Variants

For analysis of 45S rDNA expression, cDNA synthesis was performed on 2.5 μg of DNaseI-treated RNA with the 3allrRNAvar primer using M-MLV reverse transcriptase (Promega). Obtained cDNAs were diluted 1:2 and used in PCR (Promega Flexi) with 25 to 40 cycles. For the relative abundance of each class of rDNA and rRNA variants, PCR was performed with 5allrRNAvar/3allrRNAvar primers. Quantification of variants was performed on non-saturated signals using Multi Gauge software (Fujfilm).

ChIP Analysis

The 2.5-week-old plantlets were formaldehyde cross-linked, and ChIP was performed as previously described (Bowler et al., 2004) with minor modifications: Chromatin was sheared with the Diagenode Bioruptor (10 cycles of 30 s on and 1.5 min off). Protein A-coupled Dynabeads (Invitrogen) or Anti-FLAG M2 Magnetic beads (Sigma-Aldrich; M8823) were used, and the sonicated chromatin was precleared in presence of Invitrogen magnetic beads for 3 h, before immunoprecipitation with the anti-H3 antibody (Abcam; ab1791, lots GR242835-1, GR265016-1, and GR172700-1), with the anti-H3K9me2 antibody (Abcam; ab1220, lot GR166768-3), or with the Anti-FLAG M2 Magnetic beads (Sigma-Aldrich; M8823, lot SBL1128V). DNA was quantified using qPCR and normalized relatively to input.

Protein Extraction and Immunoblot Analysis

To recover the non-nucleosomal fraction of histones present in the cytoplasm and in the nucleoplasm, proteins were extracted from 100 mg of plantlets according to Durut et al. (2014). For the total extract corresponding to the histones present in the cytoplasm, nucleoplasm, chromatin, and the nucleus, proteins were recovered from 1 g of plantlets with Honda buffer (2.5% Ficoll 400, 5% Dextran T40, 0.4 M sucrose, 25 mM Tris-HCl, pH 7.4, 10 mM $MgCl_2$, 10 mM β-mercaptoethanol, 0.5 mM PMSF, and complete Protease Inhibitor Cocktail Tablets [Roche]) (Honda et al., 1966). Briefly, ground tissues were homogenized in 15 mL of Honda buffer and filtered through a double layer of Miracloth (Millipore). After removal of an aliquot that constitutes the "total" extract, Triton X-100 was added at a 0.5% concentration and samples were submitted to rotation on a wheel at cold for 15 min. After 5 min centrifugation at 1500*g*, pellets were washed with Honda buffer containing 0.1% Triton X-100. These nucleus-rich preparations were centrifuged at 1500*g* for 5 min, and recovered pellets were washed with Honda buffer without Triton X-100. Nuclei were recovered after 5 min centrifugation at 1500*g* to constitute the "nuclear" fraction. Immunoblots were probed with the anti-H3 antibody (Abcam, ab1791, lots GR242835-1, GR265016-1, and GR172700-1; 1/3000), anti-HA antibody (Abcam; ab9110, lot GR235874-6; 1/1000), or with the anti-H4 antibody (Abcam; ab10158, lots GR264160-1 and GR273051-1; 1/1000). Equal loading was confirmed with an anti-Actin antibody (Sigma-Aldrich; clone 10-B2, A0480, lot 054M4805V; 1/3000). Primary antibodies were revealed by incubation with anti-rabbit (Abliance; BI2407, lot 14052; 1/3000) or anti-mouse (Abliance; BI2413C, lot 13101; 1/3000) secondary antibodies. Immunoblot chemiluminescence was revealed using ECL protein gel blotting detection reagents (GE Healthcare Biosciences). Densitometric analysis of immunoreactive protein bands was performed on nonsaturated signals using Multi Gauge software (Fujfilm) and H3, H4, and HA levels normalized to Actin with the wild type set to 100% and Actin being used as a loading control.

RNA-Seq Library Construction

RNAs were isolated from 2.5-week-old in vitro-grown plantlets using Tri-Reagent (Euromedex) and then treated with RQ1 DNase I (Promega) and purified with phenol-chloroform extraction. RNA-seq libraries were prepared using the TruSeq Stranded Total RNA sample preparation kit with Ribo-Zero Plant (Illumina) following the manufacturer's protocol. All six samples, along with their inputs, were pooled together and run on one lane, each on a separate flow cell of an Illumina HiSeq 4000 for paired-end sequencing (GEO accession number GSE87918).

ChIP-Seq Assay

ChIP-seq was performed on chromatin from 2.5-week-old in vitro-grown plantlets sheared to mononucleosomes using Anti-FLAG M2 Magnetic beads (Sigma-Aldrich). Preparation of ChIP-seq samples and library construction were performed as described (Veluchamy et al., 2016). Single-end sequencing using Illumina GAIIx with a read length of 75 bp was performed on two eH3.3 and two *atrx-1* eH3.3 ChIP samples, along with their inputs, at the IPS2 Transcriptomic platform (Paris-Saclay; GEO accession number GSE87918).

Analysis of ChIP-Seq Data

The quality of reads was checked using FASTQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/) with a cutoff Phred score of 20. Reads with minimum length of 36 bp were kept after trimming using Trimmomatic. Parameters for Trimmomatic were set as follows: minimum length of 36 bp, mean Phred quality score greater than 30, leading and trailing base removal with base quality below 3, and sliding window of 4:15. Reads were then mapped onto TAIR10 using Bowtie with the unique read mapping parameter (Langmead and Salzberg, 2012). Peak detection was performed using MACS2 (Zhang et al., 2008). To analyze changes in location and occupancy of eH3.3 binding position, we used DANPOS2 (Chen et al., 2013). DANPOS2 settings were used as follows: minimum read depth, 5; window size, 50 bp; minimal distance between peaks, 100 bp. To analyze occupancy, we used unique reads mapping to a position and masked transposable element and low complexity regions. Analysis and visualization of the data were performed using the IGB browser (Integrated Genome Browser; Nicol et al., 2009). Mean and SD of the coverage depth were calculated and plotted using Qualimap.

Gene Ontology

Analysis for GO term enrichment was performed using the Agrigo tool (http://bioinfo.cau.edu.cn/agriGO/) (Du et al., 2010) with TAIR10 annotations. Down- and upregulated genes were analyzed separately.

Primers

All primer sequences are listed in Supplemental Table 4.

Accession Numbers

The sequencing data corresponding to RNA-seq and ChIP-seq libraries have been deposited in the Gene Expression Omnibus under accession number GSE87918. Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession numbers: *At1g65470* (*FAS1*), *At5g64630* (*FAS2*), *At3g44530* (*HIRA*), *At1g21610* (*UBN1*), *At1g77310* (*UBN2*), *At1g64230* (*UBC28*), *At4g29130* (*HXK1*), *At2g36060* (*UEV1C*), *At2g15810* (*Mule*), *At1g08600* (*ATRX*), and *At2g28390* (*MON1*); and for the ATRX homologs: *Populus trichocarpa* (XP_002319663), *Ricinus communis* (XP_015576552), *Medicago truncatula* (XP_003590986), *Glycine max* (XP_003555577), *Arabidopsis thaliana* (AtATRX, NP_001184937), *Brassica rapa* (XP_018515512), *Oryza sativa* (XP_015614509), *Mus musculus* (MmATRX, AAC08741), *Homo sapiens* (HsATRX, NP_000480), *Drosophila melanogaster* (DmXNP, NP_651398 et DmADD1, NP_725094), and *Caenorhabditis elegans* (Q9U7E0).

Supplemental Data

Supplemental Figure 1. Evolutionary conservation of the ATRX chaperones and expression of Arabidopsis *ATRX*.

Supplemental Figure 2. Analysis of *atrx* mutant alleles.

Supplemental Figure 3. Genetic interaction between CAF-1 and HIR histone chaperones and ATRX.

Supplemental Figure 4. Effects of ATRX loss on H3K9me2 enrichment and expression of several silent loci.

Supplemental Figure 5. Effects of ATRX loss on H3.1 and H3.3 levels.

Supplemental Figure 6. Analysis of effects of ATRX loss on transcription by RNA-seq and H3.3 occupancy by ChIP-seq.

Supplemental Figure 7. Effects of ATRX loss at 45S rDNA loci.

Supplemental Table 1. Segregation analysis in F2 populations from crosses between *atrx* alleles and mutants in the HIR complex.

Supplemental Table 2. Segregation analysis in F3 populations from crosses between *atrx* alleles and mutants in the HIR complex.

Supplemental Table 3. Segregation analysis in F2 populations from crosses between *fas2-5* and *atrx* alleles.

Supplemental Table 4. Primer list.

Supplemental Data Set 1. List of transposons differentially expressed in *atrx-1*.

Supplemental Data Set 2. List of the differential peaks identified in *atrx-1* eH3.3.

Supplemental Data Set 3. List of the differential peaks identified in *atrx-1* eH3.3 located in transposons.

Supplemental Data Set 4. List of the differential peaks identified in *atrx-1* eH3.3 located in intergenic regions.

Supplemental Data Set 5. List of genes differentially expressed in *atrx-1*.

Supplemental Data Set 6. Alignments used to generate the phylogeny presented in Figure 1A.

ACKNOWLEDGMENTS

We thank S. Amiard for seeds and help with the 5-ethynyl-2′-deoxyuridine detection, HU, and γ-irradiation experiments; P. Pouchin and N. Goué for help with bioinformatics analyses; A. Lhomme for technical assistance; E. Vanrobays for the NLS-GFP construct; O. Mittelsten Scheid for critical reading of the manuscript; and J. Sáez-Vásquez and F. Pontvianne for valuable advice with 45S rDNA variant analysis and for critical reading of the manuscript. The RNA-seq was performed by the IGBMC Microarray and Sequencing platform, a member of the "France Génomique" consortium (ANR grant ANR-10-INBS-0009). This work was supported by doctoral stipends from the Region Auvergne, ANR grants 'Dynam'Het' ANR-11 JSV2 009 01 and 'SINUDYN' ANR-12 ISV6 0001. This work was further supported by the CNRS, INSERM, and Université Clermont Auvergne.

AUTHOR CONTRIBUTIONS

C.D. and A.V.P. designed the research. C.D., M.B., G.D., L.S., D.L., S.L.G., A.P., S.C., and A.V.P. performed experiments. C.D., A.P., and M.J. gathered and analyzed RNA-seq data, and C.D., D.L., A.V., and M.B. gathered and analyzed ChIP-seq data. C.D. and A.V.P. analyzed the data. A.P. contributed analysis tools. C.D., C.T., and A.V.P. wrote the article.

Received November 28, 2016; revised May 24, 2017; accepted June 28, 2017; published July 6, 2017.

REFERENCES

Abou-Ellail, M., Cooke, R., and Sáez-Vásquez, J. (2011). Variations in a team: major and minor variants of *Arabidopsis thaliana* rDNA genes. Nucleus **2:** 294–299.

Alekseyenko, A.A., Gorchakov, A.A., Zee, B.M., Fuchs, S.M., Kharchenko, P.V., and Kuroda, M.I. (2014). Heterochromatin-associated interactions of Drosophila HP1a with dADD1, HIPP1, and repetitive RNAs. Genes Dev. **28:** 1445–1460.

Alexander, M.P. (1969). Differential staining of aborted and non-aborted pollen. Stain Technol. **44:** 117–122.

Alexeev, A., Mazin, A., and Kowalczykowski, S.C. (2003). Rad54 protein possesses chromatin-remodeling activity stimulated by the Rad51-ssDNA nucleoprotein filament. Nat. Struct. Biol. **10:** 182–186.

Banks, D.D., and Gloss, L.M. (2004). Folding mechanism of the (H3-H4)2 histone tetramer of the core nucleosome. Protein Sci. **13:** 1304–1316.

Bassett, A.R., Cooper, S.E., Ragab, A., and Travers, A.A. (2008). The chromatin remodelling factor dATRX is involved in heterochromatin formation. PLoS One **3:** e2099.

Bernatavichute, Y.V., Zhang, X., Cokus, S., Pellegrini, M., and Jacobsen, S.E. (2008). Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in *Arabidopsis thaliana*. PLoS One **3:** e3156.

Bowler, C., Benvenuto, G., Laflamme, P., Molino, D., Probst, A.V., Tariq, M., and Paszkowski, J. (2004). Chromatin techniques for plant cells. Plant J. **39:** 776–789.

Charbonnel, C., Allain, E., Gallego, M.E., and White, C.I. (2011). Kinetic analysis of DNA double-strand break repair pathways in Arabidopsis. DNA Repair (Amst.) **10:** 611–619.

Chen, K., Xi, Y., Pan, X., Li, Z., Kaestner, K., Tyler, J., Dent, S., He, X., and Li, W. (2013). DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. Genome Res. **23:** 341–351.

Clynes, D., Higgs, D.R., and Gibbons, R.J. (2013). The chromatin remodeller ATRX: a repeat offender in human disease. Trends Biochem. Sci. **38:** 461–466.

Collins, N., Poot, R.A., Kukimoto, I., García-Jiménez, C., Dellaire, G., and Varga-Weisz, P.D. (2002). An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. Nat. Genet. **32:** 627–632.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol. **139:** 5–17.

Dhayalan, A., Tamas, R., Bock, I., Tattermusch, A., Dimitrova, E., Kudithipudi, S., Ragozin, S., and Jeltsch, A. (2011). The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. Hum. Mol. Genet. **20:** 2195–2203.

Drané, P., Ouararhni, K., Depaux, A., Shuaib, M., and Hamiche, A. (2010). The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. Genes Dev. **24:** 1253–1265.

Du, Z., Zhou, X., Ling, Y., Zhang, Z., and Su, Z. (2010). agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res. **38:** W64–W70.

Duc, C., Benoit, M., Le Goff, S., Simon, L., Poulet, A., Cotterell, S., Tatout, C., and Probst, A.V. (2015). The histone chaperone complex HIR maintains nucleosome occupancy and counterbalances impaired histone deposition in CAF-1 complex mutants. Plant J. **81:** 707–722.

Durut, N., et al. (2014). A duplicated NUCLEOLIN gene with antagonistic activity is required for chromatin organization of silent 45S rDNA in Arabidopsis. Plant Cell **26:** 1330–1344.

Elsässer, S.J., Huang, H., Lewis, P.W., Chin, J.W., Allis, C.D., and Patel, D.J. (2012). DAXX envelops a histone H3.3-H4 dimer for H3.3-specific recognition. Nature **491:** 560–565.

Elsässer, S.J., Noh, K.-M., Diaz, N., Allis, C.D., and Banaszynski, L.A. (2015). Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. Nature **522:** 240–244.

Emelyanov, A.V., Konev, A.Y., Vershilova, E., and Fyodorov, D.V. (2010). Protein complex of Drosophila ATRX/XNP and HP1a is required for the formation of pericentric beta-heterochromatin in vivo. J. Biol. Chem. **285:** 15027–15037.

Eustermann, S., Yang, J.-C., Law, M.J., Amos, R., Chapman, L.M., Jelinska, C., Garrick, D., Clynes, D., Gibbons, R.J., Rhodes, D., Higgs, D.R., and Neuhaus, D. (2011). Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. Nat. Struct. Mol. Biol. **18:** 777–782.

Filipescu, D., Szenker, E., and Almouzni, G. (2013). Developmental roles of histone H3 variants and their chaperones. Trends Genet. **29:** 630–640.

Fransz, P., Armstrong, S., Alonso-Blanco, C., Fischer, T.C., Torres-Ruiz, R.A., and Jones, G. (1998). Cytogenetics for the model system *Arabidopsis thaliana*. Plant J. **13:** 867–876.

Garrick, D., Sharpe, J.A., Arkell, R., Dobbie, L., Smith, A.J.H., Wood, W.G., Higgs, D.R., and Gibbons, R.J. (2006). Loss of Atrx affects trophoblast development and the pattern of X-inactivation in extraembryonic tissues. PLoS Genet. **2:** e58.

Gibbons, R.J., Picketts, D.J., and Higgs, D.R. (1995a). Syndromal mental retardation due to mutations in a regulator of gene expression. Hum. Mol. Genet. **4:** 1705–1709.

Gibbons, R.J., Picketts, D.J., Villard, L., and Higgs, D.R. (1995b). Mutations in a putative global transcriptional regulator cause X-linked mental retardation with α-thalassemia (ATR-X syndrome). Cell **80:** 837–845.

Gokhman, D., Livyatan, I., Sailaja, B.S., Melcer, S., and Meshorer, E. (2013). Multilayered chromatin analysis reveals E2f, Smad and Zfx as transcriptional regulators of histones. Nat. Struct. Mol. Biol. **20:** 119–126.

Goldberg, A.D., et al. (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell **140:** 678–691.

Grummt, I., and Längst, G. (2013). Epigenetic control of RNA polymerase I transcription in mammalian cells. Biochim. Biophys. Acta **1829:** 393–404.

He, Q., Kim, H., Huang, R., Lu, W., Tang, M., Shi, F., Yang, D., Zhang, X., Huang, J., Liu, D., and Songyang, Z. (2015). The Daxx/Atrx complex protects tandem repetitive elements during DNA hypomethylation by promoting H3K9 trimethylation. Cell Stem Cell **17:** 273–286.

Honda, S.I., Hongladarom, T., and Laties, G.G. (1966). A new isolation medium for plant organelles. J. Exp. Bot. **17:** 460–472.

Huh, M.S., Price O'Dea, T., Ouazia, D., McKay, B.C., Parise, G., Parks, R.J., Rudnicki, M.A., and Picketts, D.J. (2012). Compromised genomic integrity impedes muscle growth after Atrx inactivation. J. Clin. Invest. **122:** 4412–4423.

Ingouff, M., Rademacher, S., Holec, S., Soljić, L., Xin, N., Readshaw, A., Foo, S.H., Lahouze, B., Sprunck, S., and Berger, F. (2010). Zygotic resetting of the HISTONE 3 variant repertoire participates in epigenetic reprogramming in Arabidopsis. Curr. Biol. **20:** 2137–2143.

Iwase, S., Xiang, B., Ghosh, S., Ren, T., Lewis, P.W., Cochrane, J.C., Allis, C.D., Picketts, D.J., Patel, D.J., Li, H., and Shi, Y. (2011). ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. Nat. Struct. Mol. Biol. **18:** 769–776.

Jacob, Y., Feng, S., LeBlanc, C.A., Bernatavichute, Y.V., Stroud, H., Cokus, S., Johnson, L.M., Pellegrini, M., Jacobsen, S.E., and Michaels, S.D. (2009). ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. Nat. Struct. Mol. Biol. **16:** 763–768.

Jones, D.T., Taylor, W.R., and Thornton, J.M. (1992). The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. **8:** 275–282.

Kaya, H., Shibahara, K.I., Taoka, K.I., Iwabuchi, M., Stillman, B., and Araki, T. (2001). FASCIATA genes for chromatin assembly factor-1 in Arabidopsis maintain the cellular organization of apical meristems. Cell **104:** 131–142.

Konev, A.Y., Tribus, M., Park, S.Y., Podhraski, V., Lim, C.Y., Emelyanov, A.V., Vershilova, E., Pirrotta, V., Kadonaga, J.T., Lusser, A., and Fyodorov, D.V. (2007). CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin in vivo. Science **317:** 1087–1090.

Laferté, A., Favry, E., Sentenac, A., Riva, M., Carles, C., and Chédin, S. (2006). The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components. Genes Dev. **20:** 2030–2040.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods **9:** 357–359.

Law, M.J., et al. (2010). ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. Cell **143:** 367–378.

Lawrence, R.J., Earley, K., Pontes, O., Silva, M., Chen, Z.J., Neves, N., Viegas, W., and Pikaard, C.S. (2004). A concerted DNA methylation/histone methylation switch regulates rRNA gene dosage control and nucleolar dominance. Mol. Cell **13:** 599–609.

Lee, N.G., Hong, Y.K., Yu, S.Y., Han, S.Y., Geum, D., and Cho, K.S. (2007). dXNP, a Drosophila homolog of XNP/ATRX, induces apoptosis via Jun-N-terminal kinase activation. FEBS Lett. **581:** 2625–2632.

Leung, J.W.-C., Ghosal, G., Wang, W., Shen, X., Wang, J., Li, L., and Chen, J. (2013). Alpha thalassemia/mental retardation syndrome X-linked gene product ATRX is required for proper replication restart and cellular resistance to replication stress. J. Biol. Chem. **288:** 6342–6350.

Lewis, P.W., Elsaesser, S.J., Noh, K.M., Stadler, S.C., and Allis, C.D. (2010). Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. Proc. Natl. Acad. Sci. USA **107:** 14075–14080.

Liu, C., Wang, C., Wang, G., Becker, C., Zaidem, M., and Weigel, D. (2016). Genome-wide analysis of chromatin packing in *Arabidopsis thaliana* at single-gene resolution. Genome Res. **26:** 1057–1068.

Liu, C.-P., Xiong, C., Wang, M., Yu, Z., Yang, N., Chen, P., Zhang, Z., Li, G., and Xu, R.-M. (2012). Structure of the variant histone H3.3-H4 heterodimer in complex with its chaperone DAXX. Nat. Struct. Mol. Biol. **19:** 1287–1292.

López-Falcón, B., Meyer-Nava, S., Hernández-Rodríguez, B., Campos, A., Montero, D., Rudiño, E., Vázquez, M., Zurita, M., and Valadez-Graham, V. (2014). Characterization of the Drosophila group ortholog to the amino-terminus of the alpha-thalassemia and mental retardation X-Linked (ATRX) vertebrate protein. PLoS One **9:** e113182.

McDowell, T.L., et al. (1999). Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. Proc. Natl. Acad. Sci. USA **96:** 13983–13988.

McStay, B., and Grummt, I. (2008). The epigenetics of rRNA genes: from molecular to chromosome biology. Annu. Rev. Cell Dev. Biol. **24:** 131–157.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.-H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science **303:** 343–348.

Mohannath, G., Pontvianne, F., and Pikaard, C.S. (2016). Selective nucleolus organizer inactivation in *Arabidopsis* is a chromosome position-effect phenomenon. Proc. Natl. Acad. Sci. USA **113:** 13426–13431.

Mozgová, I., Mokros, P., and Fajkus, J. (2010). Dysfunction of chromatin assembly factor 1 induces shortening of telomeres and loss of 45S rDNA in *Arabidopsis thaliana*. Plant Cell **22:** 2768–2780.

Muchová, V., Amiard, S., Mozgová, I., Dvořáčková, M., Gallego, M.E., White, C., and Fajkus, J. (2015). Homology-dependent repair is involved in 45S rDNA loss in plant CAF-1 mutants. Plant J. **81:** 198–209.

Nicol, J.W., Helt, G.A., Blanchard, S.G., Jr., Raja, A., and Loraine, A.E. (2009). The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics **25:** 2730–2731.

Nie, X., Wang, H., Li, J., Holec, S., and Berger, F. (2014). The HIRA complex that deposits the histone H3.3 is conserved in Arabidopsis and facilitates transcriptional dynamics. Biol. Open **3:** 794–802.

Okada, T., Endo, M., Singh, M.B., and Bhalla, P.L. (2005). Analysis of the histone H3 gene family in Arabidopsis and identification of the male-gamete-specific variant AtMGH3. Plant J. **44:** 557–568.

Ono, T., Kaya, H., Takeda, S., Abe, M., Ogawa, Y., Kato, M., Kakutani, T., Mittelsten Scheid, O., Araki, T., and Shibahara, K. (2006). Chromatin assembly factor 1 ensures the stable maintenance of silent chromatin states in Arabidopsis. Genes Cells **11:** 153–162.

Otero, S., Desvoyes, B., and Gutierrez, C. (2014). Histone H3 dynamics in plant cell cycle and development. Cytogenet. Genome Res. **143:** 114–124.

Pavlištová, V., Dvořáčková, M., Jež, M., Mozgová, I., Mokroš, P., and Fajkus, J. (2016). Phenotypic reversion in *fas* mutants of *Arabidopsis thaliana* by reintroduction of *FAS* genes: variable recovery of telomeres with major spatial rearrangements and transcriptional reprogramming of 45S rDNA genes. Plant J. **88:** 411–424.

Pchelintsev, N.A., McBryan, T., Rai, T.S., van Tuyn, J., Ray-Gallet, D., Almouzni, G., and Adams, P.D. (2013). Placing the HIRA histone chaperone complex in the chromatin landscape. Cell Reports **3:** 1012–1019.

Pecinka, A., Dinh, H.Q., Baubec, T., Rosa, M., Lettner, N., and Mittelsten Scheid, O. (2010). Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in Arabidopsis. Plant Cell **22:** 3118–3129.

Pontes, O., Lawrence, R.J., Neves, N., Silva, M., Lee, J.-H., Chen, Z.J., Viegas, W., and Pikaard, C.S. (2003). Natural variation in nucleolar dominance reveals the relationship between nucleolus organizer chromatin topology and rRNA gene transcription in Arabidopsis. Proc. Natl. Acad. Sci. USA **100:** 11418–11423.

Pontvianne, F., et al. (2010). Nucleolin is required for DNA methylation state and the expression of rRNA gene variants in *Arabidopsis thaliana*. PLoS Genet. **6:** e1001225.

Pontvianne, F., Blevins, T., Chandrasekhara, C., Feng, W., Stroud, H., Jacobsen, S.E., Michaels, S.D., and Pikaard, C.S. (2012). Histone methyltransferases regulating rRNA gene dose and dosage control in Arabidopsis. Genes Dev. **26:** 945–957.

Pontvianne, F., Blevins, T., Chandrasekhara, C., Mozgová, I., Hassel, C., Pontes, O.M., Tucker, S., Mokros, P., Muchová, V., Fajkus, J., and Pikaard, C.S. (2013). Subnuclear partitioning of rRNA genes between the nucleolus and nucleoplasm reflects alternative epiallelic states. Genes Dev. **27:** 1545–1550.

Preuss, S., and Pikaard, C.S. (2007). rRNA gene silencing and nucleolar dominance: insights into a chromosome-scale epigenetic on/off switch. Biochim. Biophys. Acta **1769:** 383–392.

Ratnakumar, K., et al. (2012). ATRX-mediated chromatin association of histone variant macroH2A1 regulates α-globin expression. Genes Dev. **26:** 433–438.

Ray-Gallet, D., Woolfe, A., Vassias, I., Pellentz, C., Lacoste, N., Puri, A., Schultz, D.C., Pchelintsev, N.A., Adams, P.D., Jansen, L.E., and Almouzni, G. (2011). Dynamics of histone H3 deposition in vivo reveal a nucleosome gap-filling mechanism for H3.3 to maintain chromatin integrity. Mol. Cell **44:** 928–941.

Ricketts, M.D., Frederick, B., Hoff, H., Tang, Y., Schultz, D.C., Singh Rai, T., Grazia Vizioli, M., Adams, P.D., and Marmorstein, R. (2015). Ubinuclein-1 confers histone H3.3-specific-binding by the HIRA histone chaperone complex. Nat. Commun. **6:** 7711.

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. **42:** W320–W324.

Rounds, M.A., and Larsen, P.B. (2008). Aluminum-dependent root-growth inhibition in Arabidopsis results from AtATR-regulated cell-cycle arrest. Curr. Biol. **18:** 1495–1500.

Sadic, D., Schmidt, K., Groh, S., Kondofersky, I., Ellwart, J., Fuchs, C., Theis, F.J., and Schotta, G. (2015). Atrx promotes heterochromatin formation at retrotransposons. EMBO Rep. **16:** 836–850.

Sandmeier, J.J., French, S., Osheim, Y., Cheung, W.L., Gallo, C.M., Beyer, A.L., and Smith, J.S. (2002). RPD3 is required for the inactivation of yeast ribosomal DNA genes in stationary phase. EMBO J. **21:** 4959–4968.

Sawatsubashi, S., et al. (2010). A histone chaperone, DEK, transcriptionally coactivates a nuclear receptor. Genes Dev. **24:** 159–170.

Schneiderman, J.I., Orsi, G.A., Hughes, K.T., Loppin, B., and Ahmad, K. (2012). Nucleosome-depleted chromatin gaps recruit assembly factors for the H3.3 histone variant. Proc. Natl. Acad. Sci. USA **109:** 19721–19726.

Schneiderman, J.I., Sakai, A., Goldstein, S., and Ahmad, K. (2009). The XNP remodeler targets dynamic chromatin in Drosophila. Proc. Natl. Acad. Sci. USA **106:** 14472–14477.

Shaked, H., Avivi-Ragolsky, N., and Levy, A.A. (2006). Involvement of the Arabidopsis SWI2/SNF2 chromatin remodeling gene family in DNA damage response and recombination. Genetics **173:** 985–994.

Shi, L., Wang, J., Hong, F., Spector, D.L., and Fang, Y. (2011). Four amino acids guide the assembly or disassembly of Arabidopsis histone H3.3-containing nucleosomes. Proc. Natl. Acad. Sci. USA **108:** 10574–10578.

Shu, H., Nakamura, M., Siretskiy, A., Borghi, L., Moraes, I., Wildhaber, T., Gruissem, W., and Hennig, L. (2014). Arabidopsis replacement histone variant H3.3 occupies promoters of regulated genes. Genome Biol. **15:** R62.

Smith, S., and Stillman, B. (1989). Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. Cell **58:** 15–25.

Steimer, A., Amedeo, P., Afsar, K., Fransz, P., Mittelsten Scheid, O., and Paszkowski, J. (2000). Endogenous targets of transcriptional gene silencing in Arabidopsis. Plant Cell **12:** 1165–1178.

Stroud, H., Greenberg, M.V., Feng, S., Bernatavichute, Y.V., and Jacobsen, S.E. (2013). Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. Cell **152:** 352–364.

Stroud, H., Otero, S., Desvoyes, B., Ramírez-Parra, E., Jacobsen, S.E., and Gutierrez, C. (2012). Genome-wide analysis of histone H3.1 and H3.3 variants in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **109:** 5370–5375.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell **116:** 51–61.

Talbert, P.B., et al. (2012). A unified phylogeny-based nomenclature for histone variants. Epigenetics Chromatin **5:** 7.

Talbert, P.B., and Henikoff, S. (2010). Histone variants–ancient wrap artists of the epigenome. Nat. Rev. Mol. Cell Biol. **11:** 264–275.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. **30:** 2725–2729.

Tang, J., Wu, S., Liu, H., Stratt, R., Barak, O.G., Shiekhattar, R., Picketts, D.J., and Yang, X. (2004). A novel transcription regulatory complex containing death domain-associated protein and the ATR-X syndrome protein. J. Biol. Chem. **279:** 20369–20377.

Thompson, H.L., Schmidt, R., and Dean, C. (1996). Identification and distribution of seven classes of middle-repetitive DNA in the *Arabidopsis thaliana* genome. Nucleic Acids Res. **24:** 3017–3022.

Vaillant, I., Schubert, I., Tourmente, S., and Mathieu, O. (2006). MOM1 mediates DNA-methylation-independent silencing of repetitive sequences in Arabidopsis. EMBO Rep. **7:** 1273–1278.

Vaquero-Sedas, M.I., and Vega-Palas, M.A. (2013). Differential association of Arabidopsis telomeres and centromeres with histone H3 variants. Sci. Rep. **3:** 1202.

Veluchamy, A., Jégu, T., Ariel, F., Latrasse, D., Mariappan, K.G., Kim, S.K., Crespi, M., Hirt, H., Bergounioux, C., Raynaud, C., and Benhamed, M. (2016). LHP1 regulates H3K27me3 spreading and shapes the three-dimensional conformation of the Arabidopsis genome. PLoS One **11:** e0158936.

Voon, H.P.J., Hughes, J.R., Rode, C., De La Rosa-Velázquez, I.A., Jenuwein, T., Feil, R., Higgs, D.R., and Gibbons, R.J. (2015). ATRX plays a key role in maintaining silencing at interstitial heterochromatic loci and imprinted genes. Cell Reports **11:** 405–418.

Wollmann, H., et al. (2017). The histone H3 variant H3.3 regulates gene body DNA methylation in Arabidopsis thaliana. Genome Biol. **18:** 94.

Wollmann, H., Holec, S., Alden, K., Clarke, N.D., Jacques, P.É., and Berger, F. (2012). Dynamic deposition of histone variant H3.3 accompanies developmental remodeling of the Arabidopsis transcriptome. PLoS Genet. **8:** e1002658.

Wong, L.H., McGhie, J.D., Sim, M., Anderson, M.A., Ahn, S., Hannan, R.D., George, A.J., Morgan, K.A., Mann, J.R., and Choo, K.H. (2010). ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. Genome Res. **20:** 351–360.

Xue, Y., Gibbons, R., Yan, Z., Yang, D., McDowell, T.L., Sechi, S., Qin, J., Zhou, S., Higgs, D., and Wang, W. (2003). The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies. Proc. Natl. Acad. Sci. USA **100:** 10635–10640.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. **9:** R137.

VERSION COURTE en LANGUE FRANÇAISE

Introduction

L'une des principales différences entre les procaryotes et les eucaryotes est la présence d'un noyau, qui permet l'encapsidation de l'ADN dans un compartiment spécialisé et l'isole du cytoplasme. La séparation est réalisée par une double membrane phospholipidique avec une membrane nucléaire interne (INM) et une membrane nucléaire externe (ONM), qui sont la continuité du réticulum endoplasmique (ER), un compartiment essentiel pour la maturation des protéines.

Le noyau est mobile et sa migration s'effectue principalement par le biais d'interactions entre nucléo- et cytosquelette, grâce aux propriétés de l'enveloppe nucléaire (NE), (Tamura and Hara-Nishimura, 2013; Zhou and Meier, 2014). Lors de la division cellulaire, le NE est nécessairement rompue par un mécanisme appelé NE Break Down (NEBD), processus dans lequel les composants du NE et du nucléosquelette sont impliqués (Murphy et al., 2010; Smoyer et Jaspersen, 2014). La structure du noyau joue un rôle fondamental pour la cellule, comme dans les réponses au stress, le développement cellulaire ou même la reproduction, et doit être dynamique pour adopter diverses formes et pour réguler l'expression génique (Ungricht and Kutay, 2017; Yang *et al*., 2017 ; Zhou and Meier, 2014). Cela implique l'interaction de composants structurels du noyau, y compris l'enveloppe et le nucléosquelette, avec des régions génomiques spécialisées (Pombo and Dillon, 2015). Ainsi, les caractéristiques principales des noyaux, en fonction du type de cellule, sont dues aux composants du NE et de la périphérie interagissant avec des régions génomiques spécifiques.

Les travaux de cette thèse ont été menés pour caractériser davantage les protéines localisées à la périphérie nucléaire chez *Arabidopsis thaliana* (*A. thaliana*), plus précisément les protéines associées à l'enveloppe nucléaire (AtNEAP). La famille des protéines AtNEAP

sont supposées jouer un rôle dans la forme du noyau et l'organisation de la chromatine, et font partie du réseau de protéines de la périphérie nucléaire (Pawar *et al.*, 2016). Après avoir présenté cette famille en détail à partir des travaux antérieurs sur les AtNEAP, les différents objectifs de cette thèse seront présentés.

Les "NUCLEAR ENVELOPE-ASSOCIATED PROTEINS"

I – Une nouvelle famille de protéines associées à l'INM

Sur la base d'un crible bio-informatique pour de nouvelles protéines du NE ayant les caractéristiques des KASH et contenant des CCD et un NLS, les protéines AtNEAP se sont révélées posséder également un domaine TM, (Pawar *et al*., 2016), ***Figure 1-6***. Une analyse phylogénétique avec des outils d'alignement de séquence (Poulet *et al*., 2016) ont révélé que la famille de gènes NEAP est apparue pour la première fois dans les gymnospermes et est absente des espèces archaïques comme les algues unicellulaires, ***Figure 1-7***. Les monocots et les eudicots forment des groupes monophylétiques avec une duplication génique spécifique à une espèce. Ainsi, lors de la spéciation des Brassicaceae dont *A. thaliana* est membre, un événement de duplication a abouti à trois gènes, NEAP1, NEAP2, NEAP3, (Poulet *et al*., 2016). AtNEAP4 étant spécifique à Arabidopsis, tronqué et transcrit à un niveau très bas, il a été considéré comme un pseudogène, (Poulet *et al*., 2016).

Les premiers travaux pour caractériser les AtNEAP ont été d'explorer leur localisation. À cette fin, une infiltration transitoire chez les plantes de *Nicotiana benthamiana* a été utilisée. Ces études ont été réalisées avec des constructions de protéines AtNEAP surexprimées de manière transitoire, fusionnées à un marqueur fluorescent. La localisation a été évaluée par microscopie confocale localisant les trois protéines à la périphérie nucléaire. Des expériences de FRAP ont montré qu'AtNEAP1 et AtNEAP2 sont plus étroitement liées à la périphérie nucléaire qu'AtNEAP3. De plus, les fractions mobiles de AtNEAP1 et AtNEAP2

sont comparables à d'autres protéines associées aux NE ou aux NE, telles que les protéines AtSUN. Des études sur les différents domaines protéiques de AtNEAP3 ont indiqué que les premiers domaines CCD et TM sont nécessaires pour la localisation à la périphérie nucléaire plutôt que dans le nucléoplasme. Il a aussi été démontré que le NLS était nécessaire pour cibler la protéine dans le noyau au lieu du cytoplasme, (Pawar *et al*., 2016).

Une deuxième étude a examiné si les AtNEAP étaient capables de former des homo- ou des hétérodimères en réalisant une expérience (apFRET) après co-infiltration de deux protéines AtNEAP fusionnées avec la protéine fluorescente jaune ou cyan (YFP / CFP) dans des feuilles de *Nicotiana Benthamiana*. Les résultats obtenus indiquent que AtNEAP1, AtNEAP2 et AtNEAP3 peuvent interagir les unes avec les autres et avec elles-mêmes. Comme ces études ont été réalisées avec des protéines surexprimées de manière transitoire, des études complémentaires ont également été menées à l'aide du système MYTH (Membrane Yeast Two Hybrid). Les interactions AtNEAP1-AtNEAP1, AtNEAP1-AtNEAP2 et AtNEAP1-AtNEAP3 ont été confirmées, bien que ces interactions aient été faibles en MYTH, (Pawar *et al.*, 2016). Pawar *et al*, 2016 ont identifié d'autres partenaires protéiques pour les AtNEAP. En effet, les expériences apFRET et MYTH ont montré que les trois AtNEAP interagissaient avec des composants du complexe LINC, AtSUN1 et AtSUN2, (Pawar *et al*., 2016).

II – Un partenaire pour AtNEAP : AtbZIP18, un lien avec la chromatine ?

Un crible MYTH utilisant AtNEAP1 comme appât a également révélé, entre autres, la protéine AtbZIP18 en tant que partenaire d'interaction, qui est un facteur de transcription (TF), (Gibalová *et al*., 2017; Pawar *et al*., 2016). Les études de localisation et de co-localisation montrent qu'AtbZIP18 est localisé dans le noyau et le cytoplasme. Lorsque AtNEAP1 est co-exprimé avec AtbZIP18, AtNEAP1 ne parvient pas à se localiser à la

périphérie nucléaire et co-localise avec AtbZIP18 dans le nucléoplasme (Pawar *et al*., 2016), indiquant une interaction potentielle *in-vivo* pour AtbZIP18 et AtNEAP1.

Ainsi, le facteur de transcription AtbZIP18 (Pawar *et al*., 2016) pourrait être un interactant de AtNEAP1. AtbZIP18 appartient à une grande famille de facteurs de transcription, appelée bZIP, impliquée dans un large éventail de mécanismes (Dröge-Laser *et al*., 2018). Chez *A. thaliana*, la famille de protéines est composée de 78 membres répartis en 13 groupes (A-M) (Dröge-Laser *et al*., 2018). La caractéristique principale est la présence d'un domaine BRLZ avec une région basique de liaison à l'ADN suivie d'un leucine zipper permettant la dimérisation des bZIP. AtbZIP18 appartient au groupe I, qui est lié à la réponse au stress, à la régulation du cycle cellulaire et à divers aspects du développement (Dröge-Laser *et al*., 2018). Il est impliqué dans le développement du pollen et a été davantage caractérisé par (Gibalová *et al*., 2017). AtbZIP18 se localise dans le RE et dans le noyau, mais est exclu du nucléole (Gibalová *et al*., 2017). Son patron d'expression est plutôt omniprésent, avec des niveaux d'expression plus élevés dans les grains de pollen matures, les noyaux d'embryon et les racines. Il se peut qu'AtbZIP18 soit redondant avec AtbZIP34, l'un de ses partenaires de liaison, et les deux jouent un rôle dans le gamétophyte mâle. AtbZIP18 et AtbZIP34 pourraient être des répresseurs car dans chaque mutant, il existe plus de gènes régulés à la hausse que régulés à la baisse. De plus, AtbZIP18 a un motif de répression EAR qui est impliqué dans l'inhibition de la transcription par la modification de la chromatine (Gibalová *et al*., 2017; Kagale and Rozwadowski, 2010). En effet, le motif EAR est un motif commun de répression transcriptionnelle actif recrutant des co-répresseurs tels qu'AtSIN3, AtSAP18 ou TOPLESS (TPL), qui interagissent avec AtHDA19 procédant à la déacétylation des histones et ainsi à la répression des gènes, (Kagale and Rozwadowski, 2011).

Des expériences de double hybrides chez la levure (Y2H) ont montré qu'AtbZIP18 pouvait également interagir avec AtbZIP61 et AtbZIP52, qui possèdent également un motif

EAR; augmentant le nombre possible d'hétérodimères (Gibalová *et al*., 2017) et par conséquent le nombre de gènes différents ciblés.

Il est donc particulièrement intéressant d'étudier l'interaction d'AtbZIP18 avec la famille de protéines AtNEAP afin d'établir un lien entre la chromatine et la périphérie nucléaire. De plus, comme AtbZIP18 serait un régulateur négatif de l'expression des gènes, cela pourrait aider à mieux comprendre pourquoi et comment l'hétérochromatine est recrutée à la périphérie nucléaire.

III – Objectifs du projet de recherche

Afin de pouvoir caractériser complètement les protéines AtNEAP, la génétique inverse est requise et différents mutants simples pour les trois gènes ont été sélectionnés, puis croisés pour obtenir des doubles et triples mutants (Pawar-Menon, thèse, 2015). Le premier ensemble de mutants disponibles comprenait un simple mutant *Atneap1* KO, un simple *Atneap3* KO, un double *Atneap1Atneap3* KO (Pawar *et al*., 2016) et un simple *Atneap2* «leaky» probablement knock-down (KD), voir ***Résultats Chapitre 3-I***.

L'objectif du travail présenté dans cette thèse était de caractériser le rôle, la fonction et les interactions de la famille de protéines AtNEAP, en s'appuyant sur des travaux antérieurs (Pawar-Menon, thèse, 2015; Pawar *et al*., 2016), qui indiquaient un emplacement à la périphérie nucléaire et suggéraient un rôle dans l'interaction de la chromatine avec l'enveloppe nucléaire et le nucléosquelette via le complexe LINC.

Au cours de cette thèse, des tentatives ont été menées pour générer un triple mutant *Atneap*, grâce à la technologie CRISPR/Cas9, afin de générer un nouveau simple mutant *Atneap2* KO croisé avec le double mutant *Atneap1Atneap3* KO déjà disponible. Enfin, le triple KO *Atneap1Atneap2Atneap3* et toutes les combinaisons de mutants simples et doubles de ce croisement ont été identifiés et sélectionnés. L'étude et la caractérisation des mutants sont décrites dans la section ***Résultats Chapitre 3***.

De nouveaux partenaires d'interaction d'AtNEAP ont été explorés à l'aide du Y2H classique et une confirmation de la localisation et une interaction avec AtbZIP18 par microscopie confocale à haute résolution (apFRET) ont été obtenues. De plus, la localisation *in vivo* chez *A. thaliana* a été confirmée à l'aide d'une technique d'expression transitoire appelée FAST. Enfin, un certain nombre d'outils a été développé pour les travaux futurs, notamment des anticorps spécifiques aux AtNEAP, un vecteur de complémentation pour les lignées mutantes et des protocoles d'extraction de protéines pour l'immunoprécipitation.

Les résultats des travaux seront examinés dans le contexte des connaissances actuelles sur la structure nucléaire des plantes et les constituants du NE, du nucléosquelette et de la chromatine. Enfin, de possibles futurs travaux seront discutés afin de déterminer plus complètement, d'une part, le rôle de cette famille de protéines dans le réseau protéique de la périphérie nucléaire, et, d'autre part, son implication dans la régulation de l'expression des gènes par interaction avec le facteur de transcription AtbZIP18.

Résultats

Ch3 - Caractérisation de la famille de protéines AtNEAP

La fonction de la famille de protéines AtNEAP à la périphérie nucléaire doit encore être éclaircie. Afin d'étudier le rôle potentiel de ces protéines, des mutants pour chaque gène de cette famille et des combinaisons de ces mutants, soient des mutants simples, doubles et triples, ont été précédemment obtenus. Les mutants étaient des lignées d'insertion d'ADN-T et des analyses de RT-PCR ont montré que le mutant *neap2*-ADN-T n'était probablement pas une perte de fonction, ***Figure 3-1***. Ainsi, le «triple mutant» était en réalité un mutant avec trois insertions d'ADN-T mais pas une triple perte de fonction. Des travaux ont donc été

entrepris pour générer un nouveau mutant knock-out (KO) *neap2* en utilisant la technique CRISPR/Cas9 sur des plantes sauvages (WT).

Créer un mutant KO pour AtNEAP2 via CRISPR / Cas9

L'enzyme Cas9 est capable de casser un ADN double brin à des sites particuliers reconnus par un petit ARN guide (sgRNA) qui recrute la Cas9. Ce système peut être adapté afin de choisir un site spécifique dans lequel une cassure est souhaitée. Ensuite, une réparation par la cellule sera nécessaire, principalement par jonction d'extrémités non homologues (NHEJ) chez les plantes (Schiml et al., 2017), laquelle provoque généralement des mutations dans le gène par addition ou suppression de une ou plusieurs base(s).

Pour le gène AtNEAP2, deux sites cibles différents ont été choisis et deux vecteurs spéciaux fabriqués, contenant la séquence du sgARN spécifique de l'un ou l'autre site choisi et contenant également la séquence du gène Cas9. Ensuite, les plantes ont été transformées via la bactérie *Agrobacterium tumefaciens* préalablement transformée avec les vecteurs conçus.

Les sites cibles choisis étaient au niveau du premier et troisième exon. A chaque site où la coupure Cas9 est attendue, un site de restriction est également présent. S'il y a une mutation dans le site ciblé, le site de restriction est détruit, ce qui entraîne une perte de capacité pour l'enzyme de restriction de reconnaître son site et de couper. C'est de cette manière qu'il a été possible de sélectionner des plantes portant des mutations dans le gène AtNEAP2, ***Figure 3-2***.

Avant d'obtenir un nouveau mutant homozygote pour AtNEAP2, plusieurs générations de plantes ont dû être sélectionnées. La première génération (T1) de plantes transformées a été traitée avec un antibiotique BASTA pour sélectionner les plantes ayant incorporé le vecteur. Les plantes positives sont alors montées en graines pour donner la génération suivante. Ensuite, les plantes T2 ont été criblées pour deux paramètres différents : la présence ou l'absence du transgène CRISPR/Cas9 par PCR avec des amorces spécifiques au

transgène; et la présence d'une mutation dans le gène AtNEAP2, par PCR, suivie d'une digestion enzymatique des produits de PCR, ***Figure 3-3***.

Les plantes ayant perdu le transgène CRISPR/Cas9 et présentant une bande complète pour le gène AtNEAP2 après digestion - c'est-à-dire un ADN résistant à la digestion par l'enzyme de restriction et donc présentant une mutation dans la séquence - ont été maintenues à croître jusqu'à la génération suivante. Les plantes T3 ont ensuite été soumises à la même forme de sélection, des mutants homozygotes potentiels étant attendus à ce stade. Enfin, la sélection a été effectuée sur la génération suivante (T4) et deux allèles mutants différents sur le quatrième exon ont été obtenus pour AtNEAP2, nommés respectivement *Atneap2-1* et *Atneap2-2*. Ces deux mutants ont une seule insertion de nucléotide dans l'exon central, un T en position 743 pour *Atneap2-1* et *Atneap2-2* étant une transhétérozygote probable, par exemple deux versions différentes de l'insertion à la même position, ***Figure 3-4***.

Chaque mutant *Atneap2* a ensuite été croisé avec le double mutant *Atneap1Atneap3* et les générations T2 ont été génotypées afin de trouver le triple mutant *Atneap*. La probabilité étant seulement de 1/64 de le trouver dans cette génération, les plantes mutantes homozygotes pour deux gènes et hétérozygotes pour le troisième gène ont été sélectionnées afin d'obtenir facilement le triple mutant dans la génération suivante (1/4). Cela permettant également de vérifier que le triple mutant n'était pas létal.

Trois plantes mutantes triple *Atneap* ont été finalement obtenues avec l'allèle *Atneap2.1* et les études ultérieures ont été menées avec les mutants simple *Atneap2.1* et triple *Atneap1Atneap2.1Atneap3* (nommé *Atneap1/2/3*).

Entre temps, une caractérisation plus poussée du mutant *Atneap2.1* a été réalisée avec une analyse RT-PCR et a révélé qu'un transcrit était toujours produit mais contenait le site SacI muté, qui est facile à suivre en utilisant cette enzyme de restriction spécifique, ***Figure 3-5***. Des études *in silico* ont indiqué qu'un codon stop prématuré apparaît après le site

d'insertion, ce qui pourrait conduire à une protéine tronquée sans NLS, sans domaine CC2 et sans domaine TM (***Figure 3-6***, ***Annexe V***). Des études *in vivo* sur des plants de tabac, *N. benthamiana*, avec l'expression transitoire d'une protéine tronquée AtNEAP2 similaire (domaines NLS, TM et un seul CC manquants) fusionnée à un tag GFP, ont indiqué que la protéine tronquée n'était pas localisée dans le noyau et les cellules de tabac présentaient une faible et diffuse coloration GFP suggérant une protéine instable avec un renouvellement rapide, (données non présentées).

Ainsi, il semblerait que la protéine AtNEAP2 tronquée ne soit plus nucléaire et que la mutation induite par CRISPR/Cas9 ait créé un nouvel allèle KO *Atneap2* nommé *Atneap2.1*.

Conclusion Chapitre 3

Utiliser la technologie CRISPR/Cas9 et générer un triple mutant KO pour la famille AtNEAP a été un processus long mais néanmoins réussi. L'analyse du phénotype de croissance générale dans des conditions optimales n'a révélé aucun effet de perte de fonction d'AtNEAP2 uniquement, ni d'AtNEAP1, AtNEAP2 et AtNEAP3 simultanément, ***Figure 3-8 et 3-9***. Les défauts chez les mutants ont été observés uniquement lorsque l'on regarde les tissus reproducteurs, à savoir une taille réduite des siliques et un nombre réduit de graines viables. Ces résultats indiquent premièrement que des défauts majeurs pourraient survenir lors des étapes de synthèse et de fécondation des gamètes. C'est pourquoi la méiose et la formation d'embryons sont à l'étude grâce à une collaboration. Deuxièmement, bien qu'aucun défaut de croissance n'ait été observé dans des conditions de croissance optimales, il reste à étudier comment les plantes mutantes *Atneap* réagiraient à différentes conditions de stress.

Les résultats de ces études indiquent que l'élimination totale de cette famille de protéines n'a pas d'effet important sur le phénotype général des plantes poussant dans des conditions optimales. Une étude plus poussée de l'architecture nucléaire et de l'organisation de la chromatine a montré que la position du chromocentre était affectée, ***Figure 3-7 et 3-10***.

Afin de déterminer si ces changements de position de chromocentre ont un impact sur le silence transcriptionnel des séquences répétées, organisées en chromocentre, une analyse est en cours sur l'expression des répétitions satellites 180bp, de TSI et de trois gènes euchromatiques (UBC28, UEV1C et HXK1) dans le sauvage Col-0 et le triple mutant *Atneap1/2/3*. Cela nous aidera à décider de la meilleure stratégie à appliquer avant de procéder à une analyse complète du transcriptome au niveau du génome (RNA-seq). Tout cela dans le but d'examiner le rôle des protéines AtNEAP dans la transcription des gènes et l'organisation de la chromatine, et déterminer si cela se produit de manière site-spécifique ou non.

En parallèle, un vecteur de complémentation, pAtNEAP1::4xc-Myc-AtNEAP1, a été conçu pour les lignées *Atneap* mutantes, est en cours de synthèse et devrait être transformé dans les lignées Col-0 et mutante triple *Atneap1/2/3* via *A. tumefaciens* (voir ***méthodes III.6***). Les mesures des noyaux seront évaluées selon les mêmes procédures, voir le ***chapitre 3-III*** ci-dessus. Les lignées complémentées pourraient également être des outils intéressants pour l'avenir afin d'explorer les partenaires susceptibles d'interagir avec les AtNEAP, en utilisant l'immunoprécipitation (IP) avec un anticorps anti-Myc suivi d'un séquençage par spectrométrie de masse.

Ch4 - Interactome des protéines AtNEAP

Un deuxième objectif de recherche était d'identifier de nouveaux partenaires protéiques des protéines AtNEAP afin de mieux comprendre leur fonction possible à la périphérie nucléaire. Trois aspects ont été explorés : les criblages Y2H, la caractérisation d'un interactant connu de la protéine AtNEAP - AtbZIP18 - et la conception d'anticorps dirigés contre les protéines AtNEAP.

Comme détaillé dans (Meng et al., 2005), plusieurs méthodes existent pour étudier les interactions protéine-protéine en utilisant la levure dans les systèmes Y2H et MYTH (le système MYTH teste les interactions au niveau de la membrane cellulaire et convient particulièrement aux protéines membranaires) ou en utilisant des techniques biophysiques basées sur l'imagerie fluorescente telles que apFRET ou BiFC. (Pawar et al., 2016) ont montré que les protéines AtNEAP étaient capables de former des homo- et hétéro-dimères par le biais d'expériences apFRET. Dans cette étude, les interactions ont été explorées à l'aide du système classique Y2H, qui teste l'interaction des protéines dans le noyau (Fields et Song, 1989). Des expériences ont été menées entre des protéines AtNEAP et des protéines connues localisées à la périphérie nucléaire, ou entre des protéines AtNEAP et une banque d'ADNc *d'A. thaliana* afin de rechercher de nouveaux interactants au travers de nouveaux cribles.

Un précédent crible en MYTH, utilisant AtNEAP1 comme appât (Pawar et al., 2016), a révélé une protéine interactante, AtbZIP18, qui est un facteur de transcription (Gibalová et al., 2017). Cette étude a montré une co-localisation pour AtNEAP1 et AtbZIP18 dans le nucléoplasme (Pawar et al., 2016). C'est pourquoi des expériences apFRET ont été réalisées afin de confirmer cette interaction suggérée, ainsi que des constructions avec des domaines supprimés pour la protéine bZIP18 afin de caractériser le site d'interaction spécifique avec les protéines AtNEAPs par des expériences de Y2H.

Enfin, afin d'étudier l'interactome AtNEAP, des anticorps spécifiques anti-AtNEAP étaient nécessaires pour des analyses de pull-downs suivis de spectrométrie de masse. Ainsi, des anticorps ont été conçus, produits chez le lapin, vérifiés et testés sur différents extraits de protéines.

Conclusion Chapitre 4

Après plusieurs tentatives de Y2H à la recherche de partenaires AtNEAP, aucun nouveau partenaire pertinent n'a été identifié et les études MYTH précédentes identifiant

AtbZIP18 et AtMaMYB n'ont pas été confirmées, ***Figure 4-1 et 4-2***. Cependant, cela est probablement dû aux différences de propriétés des systèmes Y2H et MYTH. Les expériences apFRET réalisées ont confirmé l'interaction physique entre AtbZIP18 et AtNEAPs, ***Figure 4-5 et 4-6***, suggérée par MYTH pour AtbZIP18 + AtNEAP1 dans Pawar et al, 2016.

Les résultats obtenus ont également montré qu'AtbZIP18 est localisée à la périphérie nucléaire dans une expression transitoire chez *N. benthamiana* lors de la co-infiltration avec les protéines AtNEAPs, ***Figure 4-5***, alors que ces protéines ont été suggérées comme étant localisées dans le nucléoplasme par Pawar et al., 2016. En outre, Gibalová et al., 2017, ont montré qu'AtbZIP18 seul se trouvait dans le nucléoplasme et dans la région périnucléaire. Ainsi, cette observation pourrait indiquer que AtNEAPs et AtbZIP18 influencent leur localisation mutuellement. Une analyse récente des domaines supposés contenus dans la séquence de la protéine AtbZIP18 a révélé la présence de domaines CC proches du domaine BRLZ, qui pourraient être responsables de l'interaction avec les AtNEAP, ***Figure 4-3***. Un examen plus approfondi est maintenant nécessaire pour déterminer la fonction spécifique d'AtbZIP18 avec les AtNEAP à la périphérie nucléaire. En ce qui concerne AtMaMYB, seules des preuves préliminaires de MYTH suggèrent l'existence d'un petit réseau d'interaction entre les protéines AtSUN3, AtNEAP1, AtMaMYB et AtbZIP18, ***Annexe VII***. Bisa Andov, étudiante en thèse à Oxford Brookes University, étudiant principalement AtMaMYB, effectue d'autres expériences et évaluera *in vivo* la pertinence de ce réseau.

De plus, des anticorps spécifiques anti-AtNEAP, ***Figure 4-7 et 4-8***, ont été obtenus et la preuve de la spécificité a été démontrée en utilisant des protéines de fusion surexprimées chez la levure et le tabac, ***Figure 4-9***, (et éventuellement sur de l'extrait de protéines natives d'*A. Thaliana*, ***Figure 4-10***). Afin d'utiliser d'autres approches moléculaires permettant d'étudier l'interaction des AtNEAPs, des expériences préliminaires ont récemment été entreprises pour explorer l'utilisation de ces nouveaux anticorps pour l'IP et la co-IP à partir

d'extraits d'*A. Thaliana*. Certaines difficultés sont apparues aux premières étapes de l'extraction des protéines natives AtNEAP et de l'enrichissement des lysats en protéines de NE non solubles. Frances Tolmie (Oxford Brookes University) a récemment mis au point un protocole utilisant une méthode basée sur celle du groupe de Hank Bass, Florida State University, ayant récemment réussi l'immunoprécipitation de SUN2 et également co-immunoprécipité de nombreuses protéines périphériques (Gumber et al., 2019). Les protocoles d'extraction de protéines avec un enrichissement en protéines nucléaires doivent être appliqués aux protéines AtNEAP afin de produire du matériel végétal destiné à l'IP et à la spectrométrie de masse. D'autres approches incluent (i) les constructions 6xHis-AtNEAPs et GST-AtNEAPs, pour des expériences de pull-down du nickel et de la GST, respectivement, et (ii) l'établissement de nouvelles lignées transgéniques exprimant pAtNEAP1::4xc-Myc-AtNEAP1 (voir ***Chapitre 3 Conclusion***) dans un fond mutant triple-*Atneap* pour la complémentation et pour l'IP en utilisant un anticorps anti-Myc suivi d'un séquençage par spectrométrie de masse, ce qui ouvrira, à l'avenir, de nouvelles possibilités.

Discussion

A travers les différentes approches développées dans cette thèse, la fonction des protéines AtNEAP et leur réseau d'interaction à la périphérie nucléaire ont été davantage caractérisés. La génétique inverse utilisant la technologie CRISPR/Cas9 a permis d'obtenir un nouveau mutant triple KO *Atneap1, Atneap2, Atneap3* (*Atneap1/2/3*). Dans ces plantes dépourvues de versions fonctionnelles des trois paralogues AtNEAP, plusieurs caractéristiques phénotypiques ont été observées. Au niveau de la plante entière, des altérations phénotypiques ont été observées dans les tissus reproducteurs suggérant un rôle fonctionnel dans la méiose ou la formation d'embryons. Au niveau cellulaire, des changements dans l'organisation du noyau par rapport aux plantes WT Col-0 ont été

enregistrés et suggèrent un rôle des protéines AtNEAP dans la localisation et l'ancrage possible des chromocentres à la périphérie nucléaire. En parallèle, des approches moléculaires incluant du Y2H et des études de la localisation et co-localisation *in vivo* associées à apFRET ont confirmé la capacité des protéines AtNEAP à former des homo et hétérodimères, et à interagir avec les domaines du TF AtbZIP18. Une approche biochimique comprenant la génération d'anticorps spécifiques à AtNEAP a confirmé l'expression *in vivo*, et a révélé les propriétés fortement hydrophobes des protéines AtNEAP, ce qui a entraîné des difficultés pour l'extraction de ces protéines. Globalement, les résultats confirment le rôle des protéines AtNEAP dans l'ancrage du TF AtbZIP18 au niveau de l'INM afin de maintenir la morphologie nucléaire et l'organisation de la chromatine. Dans cette discussion générale, tout d'abord, des approches à court terme seront suggérées pour démontrer le rôle fonctionnel des protéines AtNEAP à la périphérie nucléaire. Deuxièmement, un mécanisme d'action potentiel des protéines AtNEAP dans l'ancrage des chromocentres à la périphérie nucléaire et un rôle à travers une interaction avec AtbZIP18 dans la répression des gènes seront discutés. Enfin, de futurs travaux (approches à long terme) et des hypothèses seront suggérés.

I - Rôle des protéines AtNEAPs dans l'ancrage des chromocentres à la périphérie nucléaire

Afin d'étudier la fonction des protéines AtNEAP chez *A. thaliana*, la génétique inverse a été utilisée. Il était important de générer un triple mutant KO comprenant un allèle perte de fonction pour AtNEAP2. Cet allèle n'était pas disponible au début de cette étude et il a fallu deux ans pour créer un nouveau mutant *Atneap2* via la technique CRISPR/Cas9 dans des plantes de type sauvage et ensuite d'introgresser le nouvel allèle mutant dans le double mutant *Atneap1, Atneap3* déjà disponible. Aucune plante mutante n'a pu être récupérée avec une mutation sur le premier site cible CRISPR/Cas9, idéalement situé au début du gène AtNEAP2 dans le premier exon. Cela pourrait s'expliquer par le fait que les mutations Cas9 se

produisent de manière aléatoire, tout comme le mécanisme de réparation chez Arabidopsis qui est préférentiellement NHEJ (Schiml et al., 2017). Néanmoins, une mutation au niveau du deuxième site cible CRISPR/Cas9 a été obtenue, située dans le troisième exon avant les domaines NLS et TM et il a été décidé de poursuivre le travail en utilisant cet allèle mutant. Cette insertion d'un simple nucléotide introduit un codon stop prématuré et conduirait à une protéine tronquée sans domaine NLS ni TM. Une expérience d'expression transitoire de la protéine AtNEAP2 tronquée a confirmé que cette protéine mutante n'est pas ciblée sur le noyau et qu'elle est faiblement exprimée, suggérant une certaine instabilité (turn-over rapide). Ensuite, des cribles phénotypiques préliminaires ont été effectués sur le triple mutant *Atneap1/2/3*.

Une première analyse a révélé un impact sur les siliques, dont la taille est considérablement réduite, et contenant un nombre réduit de graines viables, en corrélation avec un nombre plus élevé d'ovules non fécondés, le nombre total de graines n'étant pas affecté. Ce résultat semble correspondre au niveau plus élevé de transcription des protéines AtNEAP dans les graines, en particulier les embryons, (Pawar et al., 2016). Cette analyse a soulevé également la question d'un effet potentiel sur la méiose ou la formation d'embryons chez les mutants *Atneap*, qui fait actuellement l'objet d'une étude en collaboration avec le groupe de Monica Pradillo en Espagne.

Une analyse plus fine de la morphologie nucléaire et de l'organisation de la chromatine au sein de noyaux mutants a montré que les chromocentres sont plus internes étant donné que la distance entre ceux-ci et la périphérie nucléaire est augmentée. Même si cette expérience doit être répétée avec un nombre accru de noyaux dans les deux populations, cellules de garde et de pavement, ces résultats préliminaires pourraient indiquer un défaut de l'ancrage physique présumé des chromocentres à la périphérie nucléaire lorsque les protéines AtNEAP sont absentes. Un résultat similaire a été obtenu avec le triple mutant *Atsun1/4/5*, qui montre

également une décompaction des chromocentres et une libération du silencing des gènes au niveau de certaines séquences répétées (Poulet et al., 2016). Cependant, aucune décompaction des chromocentres n'a été observée dans le triple mutant *Atneap1/2/3* et les résultats préliminaires de RT-qPCR utilisant les mêmes séquences répétées que pour le triple *Atsun1/4/5* (180 bp, TSI) n'ont pas permis de détecter un quelconque défaut du silencing des gènes. Si les protéines AtNEAP participent à l'ancrage de régions chromatiniennes spécifiques à la périphérie nucléaire, cela ne semble pas affecter les séquences répétées ni la compaction des chromocentres. Jusqu'à présent, il reste à montrer comment une modification de la position du chromocentre affecte l'expression du génome ou est liée aux différences phénotypiques observées chez les plantes mutantes *Atneap*.

II - Les protéines AtNEAP interagissent avec le facteur de transcription AtbZIP18

AtbZIP18 a été identifié dans un crible MYTH utilisant AtNEAP1 comme appât (Pawar et al., 2016). AtbZIP18 est un facteur de transcription exprimé partout dans la plante, mais avec un niveau de transcription plus élevé dans le grain de pollen mature, (Gibalová et al., 2017), ***Figure 5-1***. Parmi les gènes AtNEAP, AtNEAP1 et AtNEAP2 sont plus transcrits que AtNEAP3 dans tous les tissus avec une transcription relativement plus forte dans les graines (***Annexe VII***, Pawar et al., 2016, Figure 4 supplémentaire). Dans ce travail, l'interaction des protéines AtNEAP avec AtbZIP18 a été confirmée *in vivo* par apFRET, ***Figure 4-6***. Il est à noter que pendant cette expérience, la localisation d'AtbZIP18, ***Figure 4-5***, était significativement différente des données précédentes (Pawar et al., 2016) où elle était nucléoplasmique et non seulement limitée à la périphérie nucléaire. Cela pourrait être dû à différents problèmes concernant l'expérience en elle-même. En effet, selon la date de l'observation, entre deux et cinq jours après l'infiltration, comme l'expression est transitoire, le niveau d'expression peut changer radicalement et provoquer une mauvaise localisation si

les protéines sont trop surexprimées. Dans ce cas, les protéines ont tendance à diffuser dans le nucléoplasme au lieu d'être limitées à la périphérie nucléaire. C'est probablement le cas dans la thèse de doctorat de Pawar-Menon, 2015, où AtNEAP1 était parfois périphérique, parfois nucléoplasmique. Enfin, dans Gibalova et al. 2017, il a été démontré qu'AtbZIP18 était enrichi à lui seul à la périphérie nucléaire. Il est donc logique que la co-localisation et l'interaction d'AtbZIP18/AtNEAP se produisent à la périphérie nucléaire, comme illustrées dans cette étude, ***Figure 4-5*** et ***Figure 4-6***, mais si les protéines de fusion sont trop surexprimées, cette interaction peut être vue dans le nucléoplasme.

Une analyse récente des domaines prédits de la famille de facteurs de transcription bZIP, y compris AtbZIP18, a révélé la présence de domaines CC chevauchant le domaine BRLZ impliqué dans l'interaction avec l'ADN et la dimérisation avec d'autres TF bZIP (Dröge-Laser et al., 2018). Les domaines CC sont connus pour être importants pour l'interaction protéine-protéine (PPI) et même si le domaine CC est commun à tous les TF bZIP (Dröge-Laser et al., 2018), dans le cas d'AtbZIP18, il semble être responsable de l'interaction avec les protéines AtNEAP d'après le résultat des expériences Y2H avec des délétions de domaines d'AtbZIP18. Cela pourrait suggérer la possibilité qu'AtbZIP18 puisse, soit interagir avec les AtNEAP et donc être connecté au NE, soit à des sites cibles spécifiques de l'ADN pour réguler la transcription, mais ne serait pas en mesure de faire les deux en même temps, ***Figure 5-2***. En effet, en liant AtbZIP18 via son domaine CC, AtNEAP pourrait masquer le motif EAR (LxLxL) et empêcher la fixation de co-répresseurs sur AtbZIP18, ***Figure 5-2***. Il est aussi envisageable que l'interaction AtNEAP/AtbZIP18 inhibe la dimérisation de AtbZIP18 et donc la liaison à l'ADN. Ensuite, en raison d'un stimulus extérieur ou d'un stress environnemental, AtbZIP18 pourrait être libéré de son ancrage au NE et pourrait déclencher une voie de répression des gènes, ***Figure 5-2***. Par conséquent, il serait

intéressant de rechercher si AtbZIP18, une fois lié à AtNEAP, est encore capable de se dimériser avec un autre AtbZIP et de se lier à l'ADN ou à des co-facteurs.

Alternativement, si l'interaction des protéines AtNEAP avec AtbZIP18 lié simultanément à l'ADN est possible, cela suggérerait que les AtNEAP pourraient être en partie responsables de l'ancrage de certains domaines chromatinien à la périphérie nucléaire en liant les TF, ***Figure 5-3***. AtbZIP18 est potentiellement un répresseur de transcription car sa perte de fonction conduit à la régulation à la hausse de 117 gènes sur 133 gènes exprimés de manière différentielle (Fold Change ≥2), (Gibalová et al., 2017). Une des explications possibles de cette activité répressive est la présence du motif EAR situé en 3' du domaine BRLZ dans AtbZIP18. Le motif EAR participe à l'inhibition de la transcription par des modifications de la chromatine et est un motif commun présent dans les TF impliqués dans la répression des gènes (Gibalová et al., 2017; Kagale et Rozwadowski, 2010), ***Figure 5-3***. En effet, le motif EAR est important pour l'interaction du TF avec les facteurs de remodelage de la chromatine. Ces co-répresseurs sont capables de recruter AtHDA19, qui est une histone dé-acétylase (HDAC), conduisant à la répression des gènes (Kagale et Rozwadowski, 2010, 2011).

AtbZIP18 fut le seul TF révélé par le système MYTH, mais de récentes expériences au laboratoire suggèrent qu'un autre facteur de transcription appelé AtMaMYB est également capable d'interagir avec AtNEAP1 et AtbZIP18 dans le système MYTH (Voisin et Vanrobays non publiés). Bien que ces nouvelles interactions doivent être confirmées *in vivo*, cela suggère que d'autres TF interagissent avec AtNEAP.

III - Travaux futurs et perspectives

Au cours des premières étapes de ce travail, un crible Y2H a été effectué mais n'a pas permis de détecter de nouveaux interactants. En outre, ni AtbZIP18 ni AtMaMYB n'ont été identifiés dans ces cribles. L'échec de la recherche de nouveaux partenaires pourrait

s'expliquer par le fait que ce système, comparé au système MYTH, n'est pas adapté aux besoins spécifiques des protéines TM telles que les AtNEAPs, ce qui conduit probablement à une mauvaise localisation des AtNEAP chez la levure. En effet, dans la thèse de doctorat de Pawar-Menon 2015, une expérience MYTH a révélé une faible interaction entre AtSUN1/AtNEAP1 et AtSUN2/AtNEAP1 qui a néanmoins été confirmée par apFRET (Pawar et al., 2016).

Ainsi, il serait intéressant de réaliser ces cribles Y2H, cette fois-ci avec les AtNEAPs ayant le domaine TM supprimé. Le fait que les résultats obtenus jusqu'à présent pour identifier les partenaires AtNEAP n'aient abouti qu'à l'identification de deux TF, AtbZIP18 et AtMaMYB, suggère que l'interaction entre les protéines AtNEAPs et la chromatine est indirecte par le biais de TF. S'il est possible que les protéines AtNEAPs interagissent avec d'autres composants de la périphérie nucléaire, cela devra être identifié par d'autres stratégies que le double-hybride.

Une telle stratégie pourrait être l'immunoprécipitation (IP) des protéines AtNEAPs, puis du séquençage par spectrométrie de masse (MS). Les nouveaux anticorps générés ont détecté avec succès les protéines AtNEAPs lors de l'analyse en WB, du moins lors de la sur-expression des protéines de fusion, ***Figure 4-9***, et il serait intéressant de valider l'absence de protéines AtNEAPs dans le triple *Atneap1/2/3*. À l'avenir, il faudra établir des protocoles IP adaptés aux protéines AtNEAPs très hydrophobes. Cela pourrait tout d'abord être testé sur les protéines AtNEAPs exprimées chez la levure, puis en utilisant une lignée exprimant la protéine AtNEAP1 marquée par c-Myc ou Flag-HA dans un fond triple *Atneap1/2/3*. Des IP préliminaires ont été réalisées mais ont échoué car l'extraction des protéines AtNEAPs s'est avérée très difficile, celles-ci pouvant rester associées à la membrane nucléaire et demeurer dans la fraction insoluble. En effet, dans l'étude présentée dans cette thèse, une vaste recherche de nouveaux interactants des protéines AtNEAPs a été lancée mais n'a pas abouti

en raison de l'échec des premières étapes du protocole lors de la préparation d'échantillons pour faire de l'IP en raison de l'insolubilité des AtNEAPs.

Il est frappant de noter que certaines des premières expériences visant à purifier la lamina présumée/nucléosquelette de la plante (Sakamoto et Takagi, 2013) ont également échoué à détecter AtNEAPs, AtbZIP18 ou AtMaMYB. Cela pourrait mettre en évidence les difficultés liées à l'insolubilité relative des protéines AtNEAPs dans les tampons standards et expliquer leur absence dans la liste des 660 protéines de la lamina brute (Sakamoto et Takagi, 2013). Plus récemment, (Goto et al., 2019) ont réalisé une vaste étude du protéome nucléaire d'*A. thaliana* avec analyses en MS. Parmi les 1541 protéines identifiées, certaines protéines du nucléosquelette ont été retrouvées, telles que AtCRWN1, AtCRWN4, AtKAKU4, AtSUN1, AtSUN2, mais pas AtNEAPs, AtbZIP18 ou AtMaMYB. Il semble qu'une optimisation plus poussée des protocoles d'extraction soit nécessaire pour pouvoir faire des analyses en MS et mieux définir ce qu'est exactement l'interactome des AtNEAPs *in vivo*.

Le fait que le triple mutant *Atneap1/2/3* ne montre aucun phénotype apparent, ni pendant la phase végétative, ni au niveau de la tige ou de la racine, suggère d'abord que les protéines AtNEAPs pourraient avoir une fonction redondante avec d'autres acteurs situés à la périphérie nucléaire. Alternativement, la fonction des AtNEAPs pourrait être liée à la réponse au stress et, par conséquent, les mutants pourraient présenter un phénotype induit par le stress. Ainsi, ce mutant, et tous les mutants simples et doubles, pourraient être mis au défi dans différentes conditions de stress telles que la chaleur, le froid, la sécheresse, la durée et l'intensité de sel ou de lumière. En effet, les plantes doivent s'adapter aux changements de luminosité et de température entre nuit et jour et en fonction des conditions météorologiques. Il a été démontré que lors de la photomorphogenèse chez les plantes, l'organisation de l'hétérochromatine est profondément réorganisée (Bourbousse et al., 2015) et par un choc thermique de 30 heures à 37 ° C, la décondensation de l'hétérochromatine est induite (Pecinka

et al., 2010; Tittel-Elmer et al., 2010). Par conséquent, la plasticité du génome est vraiment importante pour les plantes lors de changements rapides dépendant de l'environnement, qui sont particulièrement difficiles. Un défaut d'organisation de la chromatine et de positionnement des chromocentres pourrait être délétère à ces étapes critiques pour les plantes mutantes. C'est pourquoi une étude de mutants *Atneap* dans des conditions de stress pourrait donc révéler des effets sur les conditions de croissance et l'intégrité de la morphologie nucléaire. De plus, comme proposé dans les modèles (***Figures 5-2*** et ***5-3***), les protéines AtNEAPs pourraient faire partie d'une voie de signalisation, reliant AtbZIP18 dont la fonction est de réprimer des gènes cibles spécifiques à l'aide de AtHDA19. Les rôles de cette HDAC ont été étudiés et il a été rapporté que AtHDA19 contrôlait l'élongation des cellules des racines, modulait la germination des graines et était impliquée dans la réponse au stress causée par l'acide abscissique et le sel, (Chen et Wu, 2010; Chen et al., 2015). Ainsi, les rôles spécifiques de HDA19 pourraient être utiles afin de déterminer quel type de stress il serait intéressant d'appliquer au triple mutant *Atneap* ou au simple *Atbzip18*.

Jusqu'à présent, la caractérisation du triple mutant *Atneap* était principalement phénotypique. Une caractérisation moléculaire plus approfondie est nécessaire afin de mieux comprendre les rôles moléculaires de la famille de protéines AtNEAPs. L'analyse d'image de la périphérie nucléaire dans le triple mutant *Atneap* a révélé une différence de localisation des chromocentres étant donné qu'un éloignement de la périphérie nucléaire a été observé. Ceci fut également observé dans le triple mutant *Atsun1/4/5* (Poulet et al., 2016) et chez les mammifères présentant des laminopathies où la forme nucléaire est modifiée avec la lobulation de la NE, l'épaississement de la lame nucléaire et la perte d'hétérochromatine périphérique, (Mattout et al., 2006). Une modification du schéma de positionnement de l'hétérochromatine entraîne des altérations de la transcription et, dans ce but, l'analyse par RNAseq pourrait être pertinente pour révéler un relâchement du silencing et une transcription

sporadique au niveau des régions centromériques et péricentromériques ou une modification de l'expression génique.

Un réarrangement de l'hétérochromatine modifierait profondément *in fine* l'organisation de l'épigénome dans le triple mutant *Atneap*. Par conséquent, il pourrait être intéressant de réaliser une analyse de séquençage du génome complet après traitement au bisulfite, ainsi que du ChIP-seq des marques histones permissives et répressives telles que, respectivement, H3K4me3, H3K9Ac et H3K9me2, H3K27me1, H3K27me3, pour voir quels gènes sont affectés par les modifications présentes chez ce mutant. Enfin, pour mieux définir la carte d'interaction chromatine-chromatine au niveau du génome entier, des analyses Hi-C pourraient être effectuées pour examiner les modifications potentielles.

Toutes ces analyses à l'échelle du génome pourraient également être effectuées sur des plantes soumises au stress afin d'exacerber les effets mutants sur la plasticité du génome. En parallèle, il serait utile de croiser le mutant *Atneap1/2/3* avec le mutant *AtbZIP18* pour étudier leur schéma d'interaction et leur fonction dans le réseau de protéines à la périphérie nucléaire. En effet, les mutants simples *Atbzip18*, en plus de présenter un nombre accru de grains de pollen avortés et des défauts chez les vivants, montrent une globale augmentation de la régulation des gènes (voir ***paragraphe II*** ci-dessus). Des expériences de RNA-Seq comparant le triple *Atneap1/2/3* à un quadruple *Atneap1/2/3, Atbzip18* pourraient révéler quelle partie du transcriptome susceptible d'être modifiée dans les mutants *Atneap1/2/3* est médiée par AtbZIP18. De la même manière, des expériences de Chip-Seq pourraient être réfléchie de manière à étudier ces sites cibles dans une lignée mutante *Atneap1/2/3* transformée avec la construction AtbZIP18-GFP sous le contrôle de son propre promoteur. Une vaste étude utilisant la technologie DAP-seq (O'Malley et al., 2016) a permis de déterminer des motifs spécifiques de liaison à l'ADN de plusieurs TF, notamment AtbZIP18 et certains de ses partenaires du groupe I de TF, AtbZIP52 et 51. Ces motifs sont -TGACAGCTGT - avec une

confiance accrue pour le noyau -CAGCT- et cette information pourrait être utile pour découvrir les gènes cibles d'AtbZIP18 liés à la périphérie nucléaire. En effet, à partir des résultats de RNA-Seq, les gènes cibles putatifs d'AtbZIP18 peuvent être identifiés et les régions en amont peuvent être criblées pour ce motif commun. Ensuite, une analyse de la co-expression de ces gènes pourra être effectuée.

Globalement, ces approches multiples seraient très prometteuses pour mieux définir l'impact de la périphérie nucléaire sur l'expression des gènes et surtout pour mieux élucider le rôle des protéines AtNEAPs à la périphérie nucléaire, au sein du nucléosquelette et dans l'ancrage de la chromatine.

RESUME

Au cours de l'évolution, les cellules eucaryotes ont acquis une enveloppe nucléaire (NE) renfermant et protégeant le génome organisé en chromatine, une structure où l'ADN s'enroule autour de protéines histones. La NE est composé de deux membranes: du côté nucléoplasmique, la membrane nucléaire interne (INM) et du côté cytoplasmique, la membrane nucléaire externe. La NE permet la communication entre les deux compartiments par le biais des complexes de pores nucléaires et relie le cytosquelette au nucléosquelette via le complexe LINC (LInker of Nucleoskeleton to Cytoskeleton). Ainsi, le nucléosquelette associé à l'INM est nécessaire pour transmettre des signaux au noyau et induire des changements dans l'organisation de la chromatine et finalement dans l'expression des gènes.

Une nouvelle famille de protéines associées à l'enveloppe nucléaire (NEAP), proposées comme nouveaux composants du nucléosquelette de la plante, a récemment été mise en évidence dans la plante modèle *Arabidopsis thaliana*. Ces protéines sont codées par une famille de trois gènes et sont ciblées vers le noyau via un NLS où elles sont ancrées à l'INM via leur domaine transmembranaire C-terminal. Les protéines AtNEAPs possèdent également plusieurs longs domaines en spirale (coiled-coil) rappelant la structure des lamines chez les animaux. Cette thèse visait à réaliser une analyse fonctionnelle des AtNEAPs à l'aide de lignées mutantes T-DNA et CRISPR/Cas9. L'interactome AtNEAP a été étudié par des approches moléculaires (Yeast Two Hybrid), indiquant des interactions entre AtNEAPs pouvant former des homo- ou hétéro-dimères; ainsi que la localisation et la co-localisation *in vivo* couplées à de l'imagerie (apFRET), qui ont confirmé les interactions avec le facteur de transcription (TF) AtbZIP18. Les anticorps spécifiques à AtNEAP générés au cours de cette étude ont été utilisés pour confirmer l'expression *in vivo*. En outre, les résultats ont indiqué que les AtNEAPs font partie du nucléosquelette et jouent un rôle dans l'ancrage des TF à l'INM afin de maintenir la morphologie nucléaire et l'organisation de la chromatine.