

**Characterization of CAP-D2 and CAP-D3
condensin subunits in *Arabidopsis thaliana***

**Dissertation
zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)**

der

Naturwissenschaftlichen Fakultät III
Agrar- und Ernährungswissenschaften,
Geowissenschaften und Informatik

der Martin-Luther-Universität Halle-Wittenberg

vorgelegt von

Celia Maria Municio Diaz
Geb. am 30.06.1987 in Madrid, Spain

Gutachter

Prof. Dr. Mieke van Lijsebettens
Prof. Dr. Andreas Houben

Verteidigt am

21.01.2019

Halle/Saale

Acknowledgements

It's going to be a long list of acknowledgments! During the thesis, as in the "out-of-the-lab" life, I couldn't have survived (not an exaggeration) or advanced without the people that were helping and supporting me, so many thanks to all of you!

To Veit Schubert and Andreas Houben, my supervisors, for all your support during the process. To Veit for being ALWAYS so calm, even when I was freaking out because some results weren't as expected, and to Andreas for giving the opportunity to do the thesis in your group and pushing when it was necessary.

To Oda, Katrin and Karla because they are our Lab mothers. You helped me with a smile, even if I asked you to help me purifying a kg of plasmid, making an immuno I thought was impossible or seeding hundreds of pots. You pay attention to every detail and know where EVERYTHING is! Thank you for taking such good care of me!

To Dmitri, I was soooooo scared of you at the beginning, but then you help me sooooo much with the experiments and gave me many good advices. By the way, I used more than once the Western blot Buffers labeled with "ONLY DEMIDOV". To Jörg, even though I had to listen to your music for 3 years and half. Your comments were always welcome, you are so helpful and cheerful; the youngest person I know!

To my lab mates Mateusz (trash music) and Steven (heavy/dark/black? metal music) and office mate Susan (actually cool music) I learned a lot with you and it was always fun. Also thank you for listening to my many many "it doesn't work" and helping me solving it when possible and comforting me when not. To Ishii, my scientific big brother, for all your patience in the lab, good food and moments in the club, travels to Japan, Italy, Munich... for everything.

To Dandan, Alevtina and Inna. And to all the other members of the CSF group, there has been so many of you coming and going!! Which meant many goodbyes but also many goodbye breakfast!!

To the CHIP-ET group, it was a pleasure to discover Europe with you! Especially to Klaus and Wojciech for hosting me in Regensburg and helping me getting my first nice results. And Mieke, thank you for all the effort you made in the coordination of the CHIP-ET group and hosting me in Gent.

To Andrea Bräutigam, the group of Frederik Coppens in VIB (Gent) and Etienne Kornobis, for all your help and patience with the bioinformatic analysis. Thanks to you I'm not scared anymore of the "black screen" and I can even enjoy doing things on R.

To Britt Leps, because you are a life saver!!! Thank you so much for all your help in everything, I think I cannot name a procedure in which you weren't involved!

To the club people!! Because living in Gatersleben without you would have been difficult. Gatersleben might not have a disco to dance musica latina, but has Paride, Isa, Francesca, Adonis...Or a theater to watch movies, but has Jonathan. Or restaurants with food from all around the world, but has Jay, Nagu, Marco, Andre, Ishii... and again, Jonathan. Or a biergarten (actually, there might be one...) for summer after work, but has Ishii, Nadine, Marco... and again, Jonathan. Or party venue for bachelor parties, but has many couples getting married and keen to celebrate it. You make it all possible, I enjoyed a lot, thank you!

Maria! Porque aunque llegaste la última, llegaste cuando más te necesitaba!!

A Etienne porque después de diez años, four countries, trois langues y un cacahuete, siempre has estado aquí para ayudarme en TODO, apoyarme y cuidarme. Muchas gracias amor!

Table of contents

| | |
|--|-----------|
| List of abbreviations | i |
| 1. Introduction | 1 |
| 1.1 The chromatin structure and organization in the nucleus | 1 |
| 1.1.1 The chromatin landscape in plants | 1 |
| 1.1.2 Spatial chromatin organization | 4 |
| 1.1.3 <i>Arabidopsis thaliana</i> genome organization | 6 |
| 1.2 The SMC complexes | 9 |
| 1.2.1 The cohesin complex | 10 |
| 1.2.2 The condensin complexes | 12 |
| 2. Aims | 17 |
| 3. Materials and methods | 18 |
| 3.1 Plant material, transformation and growing conditions | 18 |
| 3.1.1 <i>Arabidopsis</i> plant material and stable transformation | 18 |
| 3.1.2 <i>Arabidopsis</i> protoplast isolation and transformation | 19 |
| 3.1.3 <i>Arabidopsis</i> cell suspension culture transformation | 19 |
| 3.1.4 <i>Nicotiana benthamiana</i> transient transformation | 20 |
| 3.2 General methods used to characterized the condensin subunits | 20 |
| 3.2.1 Genomic DNA isolation and PCR | 20 |
| 3.2.2 Total RNA isolation, cDNA synthesis and quantification of RNA | 20 |
| 3.2.3 Cloning and construct generation | 21 |
| 3.2.3.1 Condensin subunit EYFP-fusion constructs | 22 |
| 3.2.3.2 Condensin I Bimolecular Fluorescence Complementation (BiFC) constructs | 23 |
| 3.2.3.3 CAP-D2 recombinant protein expression construct | 23 |
| 3.2.3.4 CAP-D2 and CAP-D3 affinity purification constructs | 23 |
| 3.2.3.5 CAP-D2 and CAP-D3 promoter-GUS reporter lines | 24 |
| 3.2.3.6 <i>CAP-D2</i> CRISPR-Cas 9 constructs | 24 |
| 3.2.4 DNA sequence analysis | 24 |
| 3.2.5 Total protein extraction and Western blot | 24 |
| 3.2.6 CAP-D2 and CAP-D3 affinity purification and mass spectrometry | 25 |
| 3.2.7 <i>CAP-D2</i> and <i>CAP-D3</i> promoter analysis (β -glucuronidase activity assay) | 26 |

| | |
|--|----|
| 3.2.8 Nuclei isolation and flow cytometry | 26 |
| 3.2.9 Slide preparation with flow-sorted nuclei..... | 27 |
| 3.2.10 Preparation of squashed Arabidopsis roots..... | 27 |
| 3.2.11 Probe preparation and Fluorescence <i>in situ</i> Hybridization (FISH) | 27 |
| 3.2.12 Indirect immunofluorescence labeling | 28 |
| 3.2.13 Microscopy and image analysis..... | 28 |
| 3.2.14 Gene and protein identification numbers..... | 29 |
| 3.3 CAP-D3 characterization..... | 30 |
| 3.3.1 Chromosome territory quantification..... | 30 |
| 3.3.2 Preparation of <i>cap-d3</i> mutants nuclei with preserved 3D structure | 30 |
| 3.3.3 Epigenetic landscape of the <i>cap-d3</i> mutants..... | 31 |
| 3.3.3.1 Distribution of DNA 5-methyl-cytosine, histone H3 methylation and acetylation ..31 | |
| 3.3.3.2 Centromeric DNA methylation..... | 31 |
| 3.3.4 <i>cap-d3</i> RNA-seq and transcriptome analysis | 31 |
| 3.4 CAP-D2 characterization..... | 32 |
| 3.4.1 Anti-CAP-D2 antibody production..... | 32 |
| 3.4.2 Condensin I Bimolecular Fluorescence Complementation (BiFC)..... | 32 |
| 3.4.3 Arabidopsis condensin I localization in protoplasts..... | 33 |
| 3.4.4 Meiotic analysis of the <i>cap-d2</i> T-DNA insertion mutant..... | 33 |
| 3.4.5 CRISPR-Cas 9 <i>in vitro</i> assay and the generation of <i>cap-d2</i> mutants | 33 |
| 4. Results | 36 |
| 4.1 General characterization of CAP-D2 and CAP-D3 condensin subunits..... | 36 |
| 4.1.1 Putative regulators of <i>CAP-D2</i> and <i>CAP-D3</i> expression | 36 |
| 4.1.2 <i>CAP-D2</i> and <i>CAP-D3</i> expression in Arabidopsis | 39 |
| 4.1.3 Localization of condensin I specific subunits and CAP-D3 | 42 |
| 4.1.4 CAP-D2 and CAP-D3 interacting proteins..... | 46 |
| 4.2 CAP-D3 characterization..... | 47 |
| 4.2.1 CAP-D3 organizes chromatin during interphase | 47 |
| 4.2.2 Epigenetic landscape in the <i>cap-d3</i> mutants | 51 |
| 4.2.3 The effect of CAP-D3 on transcription | 53 |
| 4.3 CAP-D2 characterization..... | 55 |
| 4.3.1 <i>cap-d2</i> T-DNA insertion mutant analysis..... | 55 |
| 4.3.2 Generation of CRISPR/Cas-based <i>cap-d2</i> mutants..... | 56 |

| | |
|---|----|
| 5. Discussion | 59 |
| 5.1 Regulation of <i>CAP-D2</i> and <i>CAP-D3</i> expression in Arabidopsis. | 59 |
| 5.2 Condensin I localizes in the cytoplasm during interphase | 61 |
| 5.3 The Arabidopsis condensin I and II subunit composition is similar to other eukaryotes | 62 |
| 5.4 <i>CAP-D3</i> organizes the chromatin during interphase..... | 63 |
| 5.4.1 <i>CAP-D3</i> and its influence in euchromatin | 63 |
| 5.4.2 <i>CAP-D3</i> and its role in heterochromatin organization | 65 |
| 5.5 Arabidopsis <i>CAP-D2</i> C-terminal domain structure differs from human <i>CAP-D2</i> | 69 |
| 6. Outlook | 70 |
| 7. Summary | 72 |
| 8. Zusammenfassung | 73 |
| 9. References | 74 |
| 10. Curriculum vitae | 85 |
| 11. Eidesstattliche Erklärung / Declaration under Oath | 88 |
| 12. Appendix | 89 |

List of abbreviations

| | |
|-------------------|---|
| 3C | Chromosome Conformation Capture |
| 3D | Three-Dimensional |
| 4C | Chromosome Conformation Capture-on-Chip |
| 5C | Chromosome Conformation Capture Carbon Copy |
| AGI | Arabidopsis Genome Initiative |
| Agrobacterium* | <i>Agrobacterium tumefaciens</i> |
| Arabidopsis* | <i>Arabidopsis thaliana</i> |
| BAC | Bacteria Artificial Chromosome |
| BiFC | Bimolecular Fluorescence Complementation |
| bp | Base pair |
| <i>C. elegans</i> | <i>Caenorhabditis elegans</i> |
| CAP | Chromosome Associated Protein |
| Cas9 | CRISPR associated protein 9 |
| CTCF | CCCTC-Binding Factor |
| cDNA | Copy DNA |
| COP9 | CONstitutive Photomorphogenesis 9 |
| CRISPR | Clustered Regularly Interspaced Short Palindromic Repeats |
| CRWN | CROwded Nuclei |
| CSN (3 /4 /5b) | Constitutive photomorphogenesis 9 SigNalosome (subunits 3 /4 /5b) |
| CT1Bp | Chromosome Territory 1 Bottom part |
| Drosophila* | <i>Drosophila melanogaster</i> |
| DAPI | 4',6-DiAmino-2-PhenylIndole, dihydrochloride |
| DCC | Dosage Compensation Complex |
| DEG | Differentially Expressed Genes |
| DIC | Differential Interference Contrast |
| DNA | DeoxyriboNucleic Acid |
| dNTPs | DeoxyNucleotide TriphosPhates |
| DTT | DiThioThreitol |
| <i>E. coli</i> | <i>Escherichia coli</i> |
| EDTA | EthyleneDiamineTetraAcetic acid |
| EYFP | Enhanced Yellow Fluorescent Protein |
| FACS | Fluorescence Activated Cell Sorting |
| FISH | Fluorescence In Situ Hybridization |
| gDNA | Genomic DNA |
| GUS | β -GlucUronidaSe |
| HEPES | 4-(2-HydroxyEthyl)-1-PiperazineEthaneSulfonic acid |
| HPLC | High Performance Liquid Chromatography |
| HUB1 /2 | Histone monoUBquitination1 /2 |
| IgG | Immunoglobulin G |
| IP | ImmunoPrecipitation |
| IPTG | IsoPropyl- β -D-ThioGalatopyranoside |
| LAC | chromosome with Loops And a Chromocenter |
| LINC | LInker of Nucleoskeleton and Cytoeskeleton |
| Kb | Kilo base |

| | |
|-----------------------|--|
| Mb | Mega base |
| MES | 2-(N-morpholino)ethanesulfonic acid |
| MORC | MicrORChidia |
| MS | Murashige and Skoog |
| MSMO | Murashige and Skoog basal salts with Minimal Organics |
| <i>N. benthamiana</i> | <i>Nicotiana benthamiana</i> |
| NCBI | National Center for Biotechnology Information |
| NEBD | Nuclear Envelope Break Down |
| NHEJ | Non-Homologous End Joining |
| NOR | Nucleolus Organizing Region |
| PAM | Protospacer-Adjacent Motif |
| PBS | Phosphate Buffered Saline |
| PCA | Principal Component Analysis |
| PCR | Polymerase Chain Reaction |
| PEG | PolyEthylene Glycol |
| PPT | PhosPhinoTricine |
| RB | RetinoBlastoma |
| RBR | RetinoBlastoma-Related |
| RNA | RiboNucleic Acid |
| ROI | Region Of Interest |
| RT-PCR | Reverse Transcription Polymerase Chain Reaction |
| SCF | Skp-Cullin-Fbox (Ubiquitin ligase) |
| SDW | Sterile Distilled Water |
| SgRNA | Single guide RNA |
| SIM | Structure Illumination Microscopy |
| SMC | Structural Maintenance of Chromosomes |
| SSC | Saline-Sodium citrate |
| TAIR | The Arabidopsis Information Resource |
| TAD | Topologically Associating Domain |
| TBST | Tris Buffered Saline with Tween 20 |
| TE | Transposable Element |
| UHR-QTOF | UltraHigh Resolution Quadrupole Time-Of-Flight |
| UTR | UnTranslated Region |
| X-Glu | 5-Bromo-4-Chloro-3-indolyl- β -D-Glucopyranoside |
| YEB | Yeast Extract Broth |

*The species name is written how it appears more often written in the literature. However, the use of italics was avoided not to mistake it with the genus name.

1. Introduction

1.1 The chromatin structure and organization in the nucleus

In eukaryotes the DNA is packaged in the nucleus together with proteins and RNA as chromatin. The chromatin basic structural unit is the nucleosome, which consists of 146 bp of DNA wrapped around a histone octamere core (Luger et al., 1997), formed by two of each histone molecules H2A, H2B, H3 and H4, and spaced by 10-50 bp of linker DNA. The spatial organization and constitution of the chromatin impact on processes that have DNA as substrate, like DNA repair, replication, recombination and transcription.

Plants, due to their sessile nature, require likely more mechanism of regulation to respond effectively to environmental conditions and stresses than animals (Huey et al., 2002). In addition, the development in plants takes place mainly post-embryonically leading to a strict regulation of developmental genes. Thus, the accessibility of regulatory proteins to DNA and interactions between loci and regulatory sequences are of special importance for plants. This regulation occurs in two ways, modification of chromatin properties and of its higher order structure.

1.1.1 The chromatin landscape in plants

Traditionally, chromatin has been divided into two types attending to the level of transcriptional activity, compaction and gene content: euchromatin and heterochromatin. Euchromatin is gene-dense, transcriptionally active and lowly condensed. In contrast, heterochromatin has a high content of repetitive sequences, a low transcriptional activity and is highly condensed. Heterochromatin can be further divided into constitutive and facultative heterochromatin. The former is permanently condensed and cytologically visible in interphase as e.g. chromocenters. It is present around centromeres, telomeres and nucleolus organizing regions. Facultative heterochromatin involves regions that, in specific cells or through development, become compact and transcriptionally inactive. First cytological observations of heterochromatin and euchromatin were made by Heitz (1928) (for review see Passarge, 1979). Besides this “traditional” chromatin classification, recent studies have categorized the chromatin into further subclasses. In *Arabidopsis*, four (Roudier et al., 2011) or nine (Sequeira-Mendes et al., 2014) chromatin signatures can be differentiated. These categories are still divided into euchromatin and heterochromatin, but are defined by specific combinations of epigenomic features, such as DNA methylation, histone modifications and histone variants.

The properties of chromatin can be modified by changing its composition: replacement of canonical histones by other histone variants; post-translational modifications of the histone tails; methylation of the cytosine nucleotide of the DNA; and by chromatin remodelers (Fig. 1).

The incorporation of different histone variants into the nucleosomes changes the attributes of the chromatin (Fig. 1a). In *Arabidopsis*, the canonical histone H3.1 is enriched in transcriptionally silent areas of the genome, the histone variant H3.3 in active chromatin (Stroud et al., 2012) and the centromere-specific histone H3 variant CENH3 replaces the canonical histone in the centromeric chromatin (Talbert et al., 2002). Other examples are the histone H2A variant H2A.W, which localizes in heterochromatin and promotes its condensation (Yelagandula et al., 2014), and the phosphorylated variant H2A.Z, which is important for DNA repair (Lang et al., 2012).

Histone tails can be modified post-translationally by acetylation, phosphorylation, methylation and ubiquitylation, among other modifications (Fig. 1b) (Bannister & Kouzarides, 2011; Fuchs et al., 2006). Acetylation and phosphorylation, both, reduce the positive charge of the histones leading to more relaxed chromatin structure and facilitating the access to the DNA. Thus, acetylation of histone H3 and H4 is related to increased transcription, while deacetylation correlates with gene repression (Wang et al., 2014). However, the phosphorylation of the serine residues 10 and 28 of H3 (H3S10 and H3S28) is associated with cell cycle-dependent chromosome condensation. In plants, these marks are enriched in the pericentromeric chromatin during mitosis and meiosis II, where they are involved in centromeric cohesion. During meiosis I these marks are evident along the whole chromosome (Gernand et al., 2003). Methylation can occur in the lysine residues K4, 9, 27 and 36 of histone H3 and in K20 of H4 as mono-, di- or trimethylation, and their distribution and function can vary between eukaryotes (Feng & Jacobsen, 2011; Fuchs et al., 2006). In *Arabidopsis*, H3K4 in its three methylated forms is present in genes and promoters, therefore marking active chromatin although only H3K4me3 is correlated with active transcription (Zhang et al. 2009). H3K27me3 localizes also in euchromatin but it acts as a transcription silencing mark at individual loci (Zhang et al. 2007). Its localization also differs between plants (present only in single genes) and animals (present in large domains) (Zhang et al. 2007). H3K9me2 is enriched in *Arabidopsis* pericentromeric heterochromatin (Soppe et al., 2002); it acts as the major epigenetic mark for heterochromatin in plants, while the situation is different in other organism, like mammals, where H3K9me3 is the main heterochromatic mark.

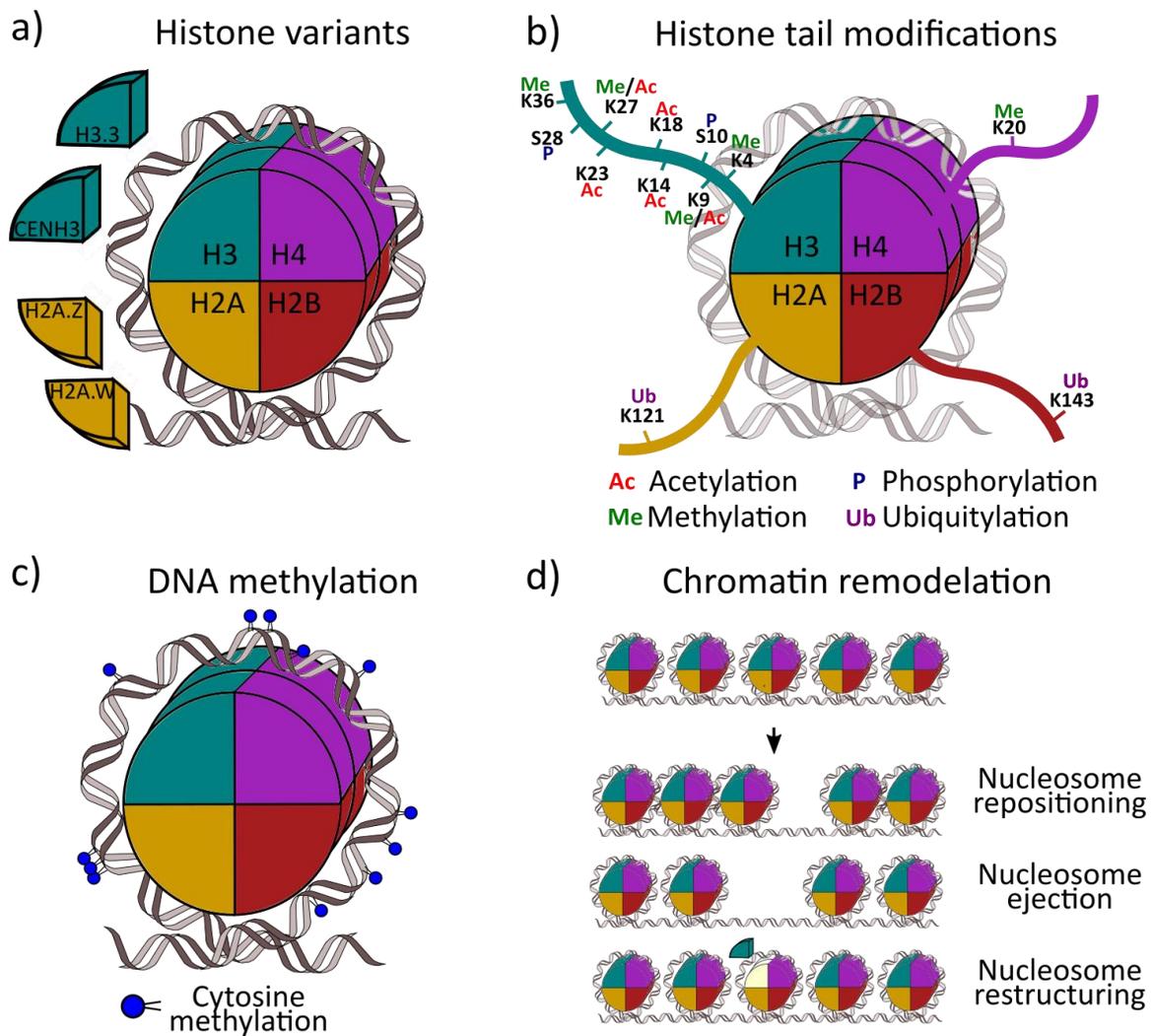


Figure 1. Chromatin modifications. The pictures represent a nucleosome (histone core wrapped by DNA) in the first three cartoons and a string of five nucleosomes in the last one. Four ways of changing the chromatin properties are depicted: a) Replacement of canonical histones (H3 and H2A) by other histone variants. b) Modification of certain residues of the histone tails by acetylation, phosphorylation, methylation and ubiquitylation. In black letters are written the amino acidic residues and in colors the modifications. c) Methylation of cytosine residues of the DNA. d) Reposition, ejection or restructuring of the nucleosomes by chromatin remodelers. Only histone variants and histone modifications mentioned in the text are depicted.

DNA methylation occurs in cytosine residues and can be present in three nucleotidic contexts, CG, CHG and CHH, where H can be an A, T or C nucleotide (Fig. 1c) (Feng & Jacobsen, 2011). In *Arabidopsis*, cytologically, DNA methylation mainly localizes to heterochromatic regions (Fransz et al., 2002). In heterochromatin, DNA methylation occurs in all three cytosine contexts and is present in transposable elements (TEs), tandem repeat sequences and long inactive gene bodies associated to H3K9me2 and H3K27me1 (Sequeira-Mendes et al., 2014) ensuring that these regions

remain inactive. CG methylation also occurs in transcribed genes, in which case it is associated to H3K4me1, H2Bub and H3K36me3 (Sequeira-Mendes et al., 2014).

Chromatin remodelers are proteins with ATPase activity. They affect the chromatin structure by ejecting, moving or restructuring the nucleosomes to expose genomic DNA to other proteins (Fig. 1d). The remodelers recognize histone modifications or are recruited by transcription factors. They alter the nucleosome position and assembly, which leads to more or less densely packed chromatin enforcing the repression or promoting the activation of genes (Clapier & Cairns, 2009). In Arabidopsis, chromatin remodelers are important in the regulation of developmental transitions and hormonal pathways (Gentry & Hennig, 2014).

1.1.2 Spatial chromatin organization

The nucleosome fiber, also known as the “beads-on-a-string” fiber, is the lowest level of chromatin configuration. This fiber can adopt higher-order structures by packaging the DNA more tightly, and regulating the accession of proteins to DNA, until reaching the level of metaphase chromosomes, the configuration with the highest compaction. The spatial folding of chromatin allows or impedes interactions between loci and regulatory sequences several kilo bases (Kb) apart, and thus influence their expression. Therefore, the spatial genome organization is a further level regulating the access to the DNA (Gibcus & Dekker, 2013).

Initial knowledge about the organization of the nucleus derived from cytological observations. Carl Rabl’s theory of the structure of the interphase nucleus stated already that during interphase: i) each chromosome occupies a distinct subnuclear domain, later known as chromosome territory (CT) (Boveri, 1909); and ii) the telomeres and centromeres cluster at opposite nuclear poles reflecting the anaphase chromosome configuration of the preceding mitosis (Rabl configuration) (Rabl, 1885 for review see Cremer & Cremer, 2010). Nevertheless, the Rabl configuration is not present in all organisms. In humans, the disposition of the chromosomes during interphase is not random, the gene dense chromosomes are located in the interior, while gene poor chromosomes are at the nuclear periphery (Boyle et al., 2001).

The development of the molecular Chromosome Conformation Capture (3C) technology allowed the analysis of contact frequencies between two genomic sequences at interphase, a “one-to-one” approach (Dekker et al., 2002). 3C can study long-range interactions, those between chromatin regions far apart in the same chromosome or between different ones, like the interaction between

gene enhancers and promoters. In recent years the improvement of 3C-based techniques facilitated the study of the three-dimensional (3D) genome organization. The 4C “one-to-all” approach allows the study of interactions between one genomic sequence and the rest of the genome. The 5C “many-to-many” method studies the interactions between multiple selected sequences and the “all-to-all” Hi-C method allows the analysis of genome wide interactions (Denker & de Laat, 2016; Lieberman-Aiden et al., 2009).

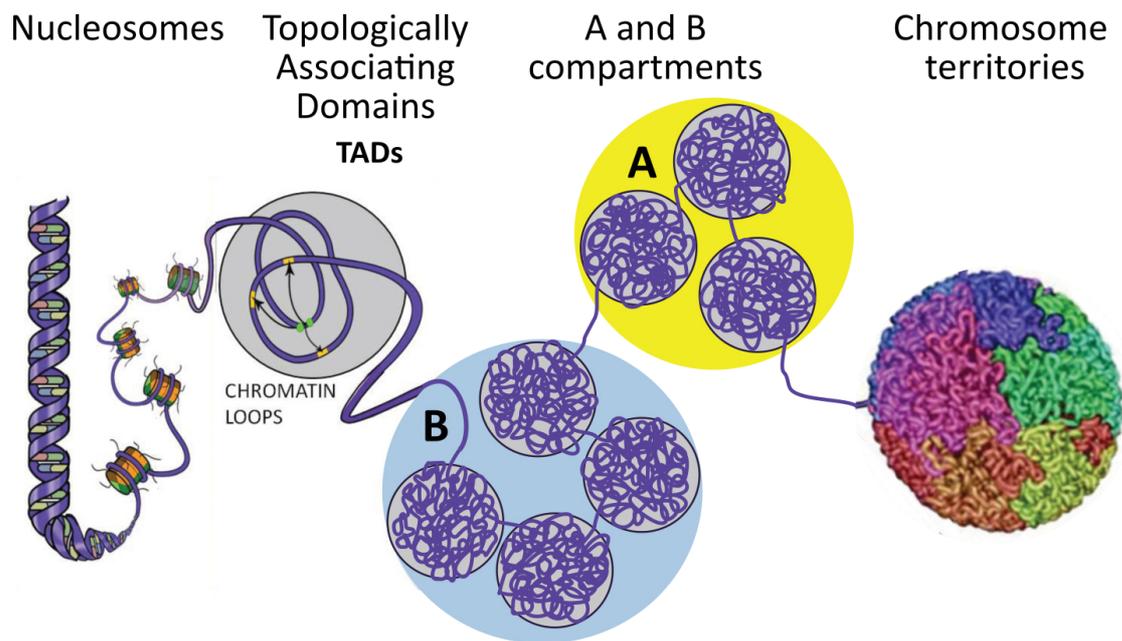


Figure 2. Interphase chromatin organization of the animal genome. The first level of chromatin organization is represented by nucleosomes. They are formed by DNA wrapped around a histone core. The next level is realized by topologically associated domains (TADs), which represent the basic unit of the higher order organization. They are regions of the genome in which interactions between regulatory elements (green) and loci (yellow) occur more often than interactions with adjacent regions. TADs with euchromatic and heterochromatic characteristics group together, defining the compartments A and B, respectively. Finally, groups of A and B compartment form the chromosome territories. Picture modified from Ea et al., (2015).

Interactions maps performed with 3C-based methods propose that topologically associating domains (TADs) are the basic unit of genome organization in animals (Fig. 2) (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012). The TADs are megabase-sized (200-kilobase to 1 Mb) local chromatin interaction domains. Interestingly, TADs are conserved between different cell types and even across species (Dixon et al., 2012) but internal contacts within each TAD are variable (Nora et al., 2012). In mammals, TAD boundaries are enriched in binding sites for the insulator protein CCCTC-Binding Factor (CTCF) (Dixon et al., 2012). In *Drosophila melanogaster*, in addition to CTCF,

other insulator proteins have been identified at TAD boundaries, including CP190 and Beaf-32 (Sexton et al., 2012). Disruption of the TAD boundaries leads to transcriptional misregulation (Nora et al., 2012). Indeed, in addition to being structural components, TADs are as well functional units. These regions of a chromosome are characterized by frequent interactions between genes and regulatory elements. Such interactions are less frequent with loci in neighboring domains. There is a common expression pattern of genes within the same TAD, suggesting that the physical confinement of genes and regulatory sequences within the TADs could coordinate their expression (Nora et al., 2012; Symmons et al., 2014). One of the proposed model for TAD formation is the loop-extrusion model (Dekker & Mirny, 2016), in which a loop-extruding factor attaches to chromatin and actively starts moving through the fiber creating a loop. This loop is enlarged until the extruding-factor arrives at two CTCF sites (one on each side of the loop) in the same orientation. That explains the enrichment of CTCF sites at the TAD boundaries. One of the complexes that have been proposed as a looping-extruding factor is the cohesin, which together with CTCF also helps the formation of interactions within TADs (see section 1.2.1) (Sofueva et al., 2013; van Ruiten & Rowland, 2018; Yuen & Gerton, 2018).

TADs with similar properties group together to form the next level of organization: the compartments A and B (Fig. 2). These compartments are defined as gene-rich, transcriptionally active and hyper-accessible to DNase I, or as gene-poor, transcriptionally silent and resistant against DNase I, respectively (Lieberman-Aiden et al., 2009). A and B compartments are more dynamic than the TADs and they are not conserved between different cell types (Dixon et al., 2015). The last level of organization is the chromosome territory, which is constituted by groups of A and B compartments (Fig. 2).

1.1.3 *Arabidopsis thaliana* genome organization

The karyotype of *Arabidopsis* ($2n=10$) presents five different chromosomes in which chromosomes 1 and 5 are the largest and metacentric. Chromosome 3 is medium-sized and submetacentric, and the chromosomes 2 and 4 are smaller and acrocentric. In the ecotype Columbia-0, used in this study, the short arms of the chromosomes 2 and 4 harbor the 45S rDNA-containing Nucleolus Organizing Region (NOR). The 5S rDNA loci map to chromosomes 3, 4 and 5 (Fransz et al., 1998)(Fig. 3a). The repetitive sequences are clustered mainly within the pericentromeric heterochromatin and the NOR (*Arabidopsis* Genome Initiative, 2000).

In plants as in animals, the arrangement of the chromosomes during interphase does not depend on genome size. Some species with large genomes, as wheat (*Triticum aestivum*), barley (*Hordeum vulgare*) and oat (*Avena sativa*) present the Rab1 configuration, while others like maize (*Zea mays*) do not (Santos & Shaw, 2004; Schubert & Shaw, 2011). Arabidopsis has a relatively small genome of 125 Mb (Arabidopsis Genome Initiative, 2000) and does not present the Rab1 configuration during interphase (Fransz et al., 2002).

During interphase, the heterochromatin is visible as bright DAPI-stained structures located near the nuclear periphery and the nucleolus. They are called chromocenters and comprise the centromeric and pericentromeric regions (Fig. 3b), the NOR and the 5S rDNA. Accordingly, DNA methylation, which is associated with transcriptionally silent DNA, is mainly present in the chromocenters (Fig. 3c). Acetylation of histones H3 and H4, an epigenetic mark associated with transcriptional activity, co-localizes with euchromatin (Fransz et al., 2002). The NOR and the 5S rDNA localize together with the centromeres of the corresponding chromosome in the same chromocenter. Therefore, in a diploid nucleus 10 chromocenters should be visible (10 centromeres and the 5S and 45S rDNA associated to them), but instead, nuclei often show 8 or 9 chromocenters, meaning that association of chromocenters occurs (Fransz et al., 2002; Schubert et al., 2012). The telomeres localize outside of the chromocenters, in the vicinity of the nucleolus and often are associated (Fig. 3d) (Fransz et al., 2002; Schubert et al., 2012). This telomere arrangement is also present in meiocytes. During the meiotic interphase the telomeres cluster around the nucleolus facilitating the homologous association of chromosomes (Armstrong et al., 2001).

The chromosomes occupy discrete spatial chromosome territories (Fig. 3e), but contrary to what happen in mammals, there is no preferential positioning of the chromosomes within the nucleus (Pecinka et al., 2004). All possible combinations of homologous and heterologous positioning of the chromosomes are present in the interphase nuclei. However, preferential homologous chromosome pairing was observed for chromosomes 2 and 4 which pair more often than at random. This is due to the attachment of their corresponding NORs to the nucleolus (Pecinka et al., 2004).

Based on cytological observations, the rosette model was proposed to explain the organization of the Arabidopsis nucleus during interphase. It states that the heterochromatin forms chromocenters that are located at the nuclear periphery and euchromatic loops emanate from

them (Fig. 3f) (Fransz et al., 2002). This cytological model is supported by computer simulations (de Nooijer et al., 2009) and Hi-C data (Feng et al., 2014; Liu et al., 2016).

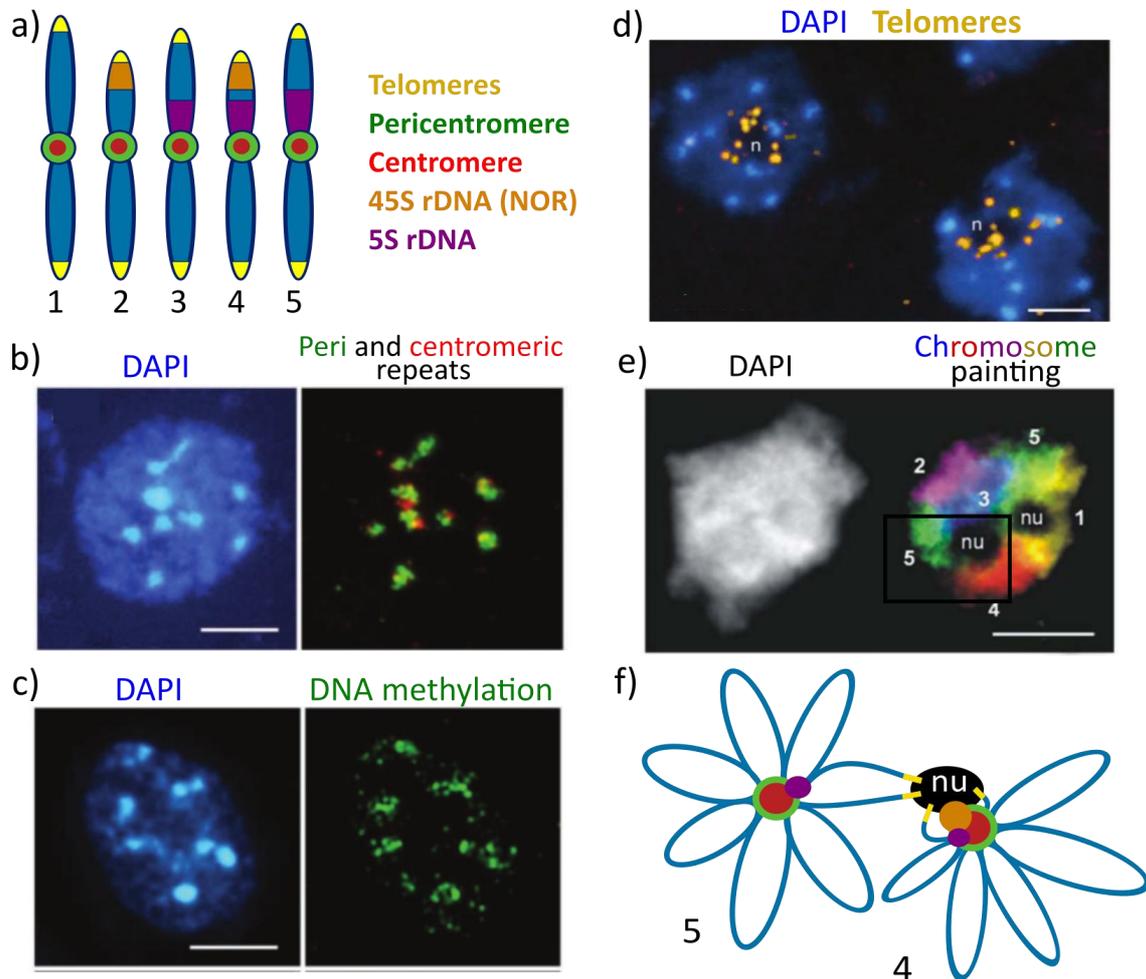


Figure 3. Arabidopsis genome organization. a) Ideogram representing the 5 different chromosomes of *Arabidopsis thaliana* ecotype Columbia-0. The telomeres are depicted in yellow and the peri- and centromeric region in green and red respectively. The 5S rDNA (purple) is present on chromosomes 3, 4 and 5 and the 45S rDNA, which forms the Nucleolar Organizing Region (NOR) maps to chromosomes 2 and 4 (orange). b) FISH in an *Arabidopsis* nucleus with the peri- and centromeric repeats in green and red, respectively. The signals localize in the chromocenters. c) Immunolocalization against 5-methyl-cytosine (DNA methylation) in green. In *Arabidopsis* the heterochromatic regions are highly methylated. d) FISH against the telomeric repeats (yellow). In *Arabidopsis*, the telomeres cluster around the nucleolus (n). e) FISH with different BACs resulting in the “painting” of each chromosome in a different color (chromosome painting). Each chromosome occupies a discrete region, the chromosome territory. The numbers indicate the chromosome number and nu is the nucleolus. f) Rosette model representation of chromosomes 5 and 4 from figure 3e (black box). The color key is the same as in 3a. According to this model, in *Arabidopsis* the chromosomes are organized as chromatin loops emanating from the heterochromatin. Pictures (b-d) are modified from Fransz et al., 2002 and picture (d) from Pecinka et al., 2004.

Recently, interaction maps based on 3C-based techniques have been described for *Arabidopsis* (Feng et al., 2014; Grob et al., 2013; Liu et al., 2016; Wang et al., 2015). Interestingly, compared to animals, in *Arabidopsis* TADs are absent (Feng et al., 2014; Wang et al., 2015). Since in vertebrates

TAD boundaries are enriched in CTCF-binding sites, the absence of TADs in Arabidopsis could be explained by the absence of CTCF in plants. However, TADs exist in rice and their boundaries are enriched in a motif recognized by transcription factors (Liu et al., 2017). Instead of TADs, gene bodies are proposed to be the basic packing unit in Arabidopsis (Liu et al., 2016). On the other hand, the separation between an A (euchromatic) and a B (heterochromatic) compartment is similar to animals (Grob et al., 2013; Liu et al., 2016). The strongest interactions, inter-chromosomal interactions, occur between telomeres and between peri- and centromeric regions. This supports the cytological observations of centromere association and telomere clustering (Fransz et al., 2002; Schubert et al., 2012). Nevertheless, most of the interactions occur intra-chromosomally and within the same arm. But the contact frequency between two loci decreases with the genomic distance. According to its interactions, the chromosome arms can be divided into a proximal region, that interacts with itself and with the pericentromere, and a distal region, that interacts with itself and the telomeric regions (Feng et al., 2014; Grob et al., 2013). No clustering (increased interactions) of highly expressed genes was observed (Feng et al., 2014; Liu et al., 2016). This further confirms the absence of distinct clustered transcription factories in Arabidopsis, as suggested by the finding of a relatively homogeneous distribution of RNA polymerase II within the euchromatin (Schubert & Weisschart, 2015). In addition, it has been shown by FISH that euchromatin segments bearing low or high expressing genes do not reveal different association frequencies (Schubert et al., 2014).

In short, to regulate gene expression a flexible 3D arrangement of the genome as well as a dynamic chromatin composition (e.g. modification of histones, histone variants, methylation of DNA) are required. In Arabidopsis, gene repositioning from the nuclear interior to the periphery has been observed upon transcriptional activation by light stimulus (Feng et al., 2014). During seedling development, light also causes a massive reorganization of the heterochromatin into chromocenters (Bourbousse et al., 2015).

1.2 The SMC complexes

Structural Maintenance of Chromosomes (SMC) proteins are present from prokaryotes to eukaryotes (Cobbe & Heck, 2004). They are essential for chromosome structure and dynamics, gene regulation and DNA repair. In eukaryotes six SMC proteins are conserved and they form the core of three different complexes: the cohesin complex, involved in sister chromatids cohesion and chromosome segregation; the condensin complexes, involved in mitotic and meiotic

chromosome formation; and the SMC5/SMC6 complex which role is on DNA repair and replication (Jeppsson et al., 2014). The SMC proteins are long coiled-coils with a globular ATPase “head” domain at one end and a hinge domain at the other end. Each complex consist of a V-shaped heterodimer formed by two SMC proteins linked by its hinge domain and a kleisin protein connecting the ATPase heads and thus, forming a closed tripartite structure (ring). The complex is completed by one or two accessory proteins containing HEAT-repeats, these repeats are involved in protein-protein interactions (Fig. 4) (Neuwald & Hirano, 2000).

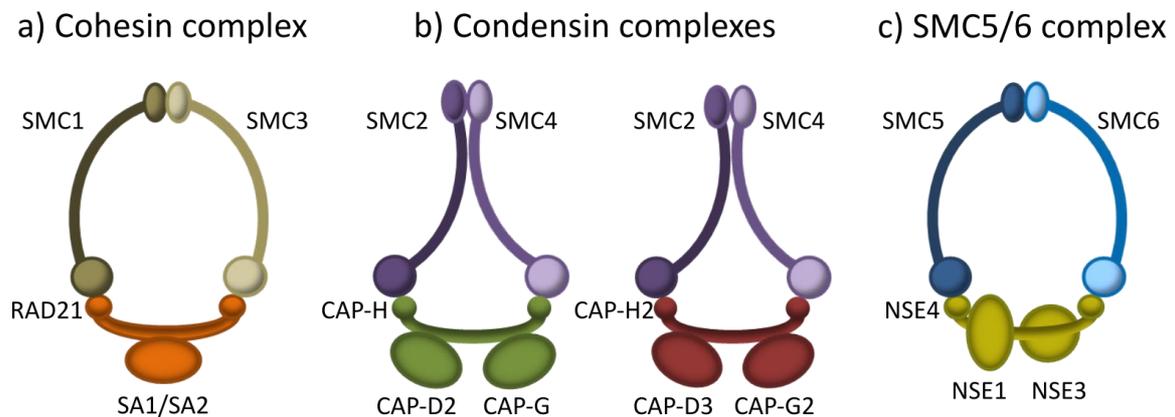


Figure 4. The SMC complexes of vertebrates. The three complexes share a basic structure. The core is formed by a heterodimer of SMC proteins, (a) SMC1-SMC3 in the case of cohesin, (b) SMC2-SMC4 for the condensin and (c) SMC5-SMC6 for the SMC5/6 complex. The cohesin ring is closed by the kleisin RAD21, and has SA1/SA2 as the accessory protein linked to the complex. The condensin complex is present in two variants, condensin I and II, which differ in the kleisin and the accessory proteins: CAP-H, CAP-D2 and CAP-G in condensin I, and CAP-H2, CAP-D3 and CAP-G2 in condensin II. In the SMC5/6 complex, the accessory proteins are NSE4, NSE1 and NSE3.

1.2.1 The cohesin complex

The cohesin complex contains a SMC1-SMC3 heterodimer connected by the α -kleisin RAD21 in vertebrates (Scc1 in budding yeast) and the adjacent HEAT-repeat subunit SA1 or SA2/STAG1 or STAG2 (Scc3) (Fig. 4a). The canonical role of the cohesin complex is the cohesion of sister chromatids during mitosis and meiosis, which ensures an accurate chromosome segregation.

In vertebrates, cohesin is loaded onto chromosomes during telophase by the NIPBL-MAU2 complex (Scc2-Scc4 in yeast). During G1, the loading is counteracted by the proteins WAPL-PDS5, which remove the cohesin complex from the chromosomes, creating a dynamic loading-removal of cohesin. The cohesin binding is not stable until SMC3 is acetylated by ESCO1 and ESCO2 (Eco1 and Eco2) and protected by the protein Sororin during the S phase. Cohesin maintains sister chromatids together as they are formed in S-phase and assists the repair of DNA double strand breaks that occur during DNA replication. At the beginning of mitosis cohesin is removed from the

chromosome arms but persists at the centromeres. At anaphase, the protease separase cleaves the RAD21 subunit, allowing the segregation of the sister chromatids to each respective pole (Fig. 5) (Jeppsson et al., 2014; Seitan & Merkschlager, 2012). In yeast, the cohesin is not released from the chromosome arms at the beginning of mitosis; instead, it maintains the cohesion until anaphase, when it is then released following the Scc1 cleavage (Marston, 2014). During meiosis, RAD21 or Scc1 are replaced by the meiosis-specific kleisin Rec8 (Parisi et al., 1999). In this case, the cohesion is resolved in two steps; at anaphase I cohesin is released from the chromosome arms but persists at the centromeres, allowing the segregation of the homologous chromosomes (reductional segregation). The cohesin is then released from the centromeres at anaphase II permitting sister chromatids segregation (equational segregation)(Watanabe & Nurse, 1999).

Cohesin works as an intermolecular linker for sister chromatids cohesion by trapping two different DNA molecules in *trans*, but it also functions as an intramolecular bridge, forming loops in the chromatin during interphase. Cohesin, thus, is also important for genome organization and gene regulation (Sofueva et al., 2013; van Ruiten & Rowland, 2018; Yuen & Gerton, 2018). As an intramolecular bridge (connecting two loci in *cis*), cohesin can inhibit or promote transcription by affecting long-range interactions. On the one hand, together with CTCF they have an insulator function, blocking the effect of enhancers on promoters (Wendt et al., 2008). On the other hand, cohesin interacts with Mediator, a transcriptional activator, creating loops between enhancers and promoters (Kagey et al., 2010).

In Arabidopsis, the cohesin and cohesin-related proteins are conserved (Schubert, 2009). There are four members of the kleisin subunit, SYN1-SYN4. SYN1 is needed for cohesion during meiosis, while SYN2-4 are essential in somatic tissues (Schubert, 2009). The relevance of cohesin for the normal development in Arabidopsis is proven by the embryo lethality of *smc1*, *smc3*, *scc3*, *scc2* and *syn3* homozygous mutants and endosperm defects of *ctf7* (ESCO1 homolog). Moreover, cohesin and proteins related to the establishment of cohesion are necessary for normal fertility and chromosome segregation in Arabidopsis (Bolaños-Villegas et al., 2013; Chelysheva et al., 2005; Liu et al., 2002; Sebastian et al., 2009; Singh et al., 2013).

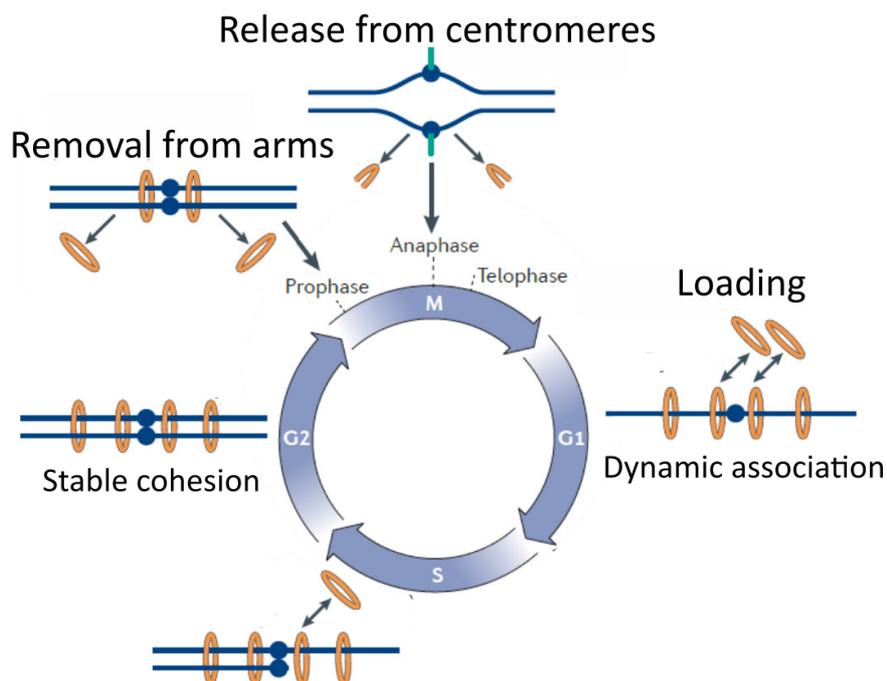


Figure 5. Vertebrate cohesin cycle. Cohesins become loaded during the G1 phase onto the chromosomes. This association is dynamic, and there is a loading-removal of the cohesins from the chromosomes. During S phase, cohesins holding the newly synthesized sister chromatids become more stably bound, and during G2 the association is stable. At the beginning of mitosis, the cohesins become removed from the chromosome arms, but not from the centromeres. Finally, during anaphase the cohesin rings become cleaved and released from the centromeres allowing the segregation of the sister chromatids. Modified from Jeppsson et al. (2014).

1.2.2 The condensin complexes

Higher eukaryotes have two condensin complexes, the condensin complex I and II. In yeast, there is one condensin complex analogous to condensin I (Freeman et al., 2000), and even bacteria and archaea have a condensin-like complex (Hirano, 2012). This conservation across all domains of life stresses the importance of this complex, whose principal role is the segregation of the genetic material.

Condensin I and condensin II share a core formed by SMC2 and SMC4 and differ in the accessory proteins. In condensin I, CAP-H is the kleisin linking the SMC subunits, and CAP-D2 and CAP-G are the HEAT-repeat proteins. In condensin II, CAP-H2 is the kleisin and the accessory proteins are CAP-D3 and CAP-G2 (Ono et al., 2003)(Fig. 4b). This composition of the complexes is conserved in higher eukaryotes except in *Drosophila*, where condensin II has only four subunits since no CAP-G2 has been described (Herzog et al., 2013; Ono et al., 2003).

Both condensin complexes show a distinct subcellular localization during the cell cycle. During interphase, condensin I is cytoplasmic and condensin II nuclear (Hirota et al., 2004; Ono et al.,

2004). In mitosis both complexes localize along the chromosome arms in an alternate fashion and both are enriched in the centromeres (Ono et al., 2003, 2004; Savvidou et al., 2005). Condensins associate to the chromosomes in an ATP-binding manner that does not require ATP hydrolysis (Hudson et al., 2008), i.e., condensins need to bind ATP, but not to hydrolyze it to associate to the chromosomes.

In budding yeast, all condensin subunits are essential for the cell viability. Depletion of condensins causes a cell division block due to incomplete anaphase (Freeman et al., 2000). SMC2 and SMC4 were first described in *Xenopus laevis* egg extracts as essential for chromosome condensation (Hirano & Mitchison, 1994). Recent studies show also that only six factors (the core histones, three histone chaperones, topoisomerase II and condensin I) are enough to assemble DNA in a chromatid-like structure *in vitro* (Shintomi et al., 2015). Surprisingly, in vertebrates and *Drosophila*, the depletion of condensin subunits delays mitosis and causes segregation problems, but do not prevent the formation of chromosomes. Nonetheless, those chromosomes without condensins show an aberrant morphology, chromosome bridges in anaphase and other segregation defects (Gerlich et al., 2006; Hartl et al., 2008b; Hirota et al., 2004; Hudson et al., 2003; Ono et al., 2003, 2004; Savvidou et al., 2005). The bridges and the impaired segregation are due to entanglements between the chromosomes that have not been resolved before metaphase (Hartl et al., 2008b; Ono et al., 2013). Depletion of one or the other complex produces different chromosome morphologies. Depletion of condensin I produces short fuzzy chromosomes while depletion of condensin II produces long curly chromosomes (Green et al., 2012; Ono et al., 2003). Therefore, the condensins are more important for the individualization, shape and rigidity of the chromosomes, than for their compaction.

According to the model for vertebrate condensin-mediated chromosome formation, during interphase condensin I locates in the cytoplasm and condensin II in the nucleus. In S phase, condensin II is involved in the resolution of the sister chromatids. In prometaphase it becomes stably bound to the chromosomes and compacts them axially by creating long-range chromatin loops. After the nuclear envelope break-down, condensin I binds to the chromosomes in a dynamic way and mediates frequent short-range interactions between the chromatin loops, compacting the chromosomes laterally and fully resolving them (Fig. 6) (Green et al., 2012; Hirano, 2012). Recent studies support this model. A single condensin molecule is capable of creating a DNA loop in an ATP-hydrolysis dependent manner (Ganji et al., 2018). Moreover, in

prometaphase, condensin II creates a helical central scaffold from which 200-400 Kb outer loops emanate. Those loops are split into smaller 80 kb condensin I-mediated loops after the nuclear envelope break-down. Condensin II is centrally located and more stable, while condensin I is more peripheral (Gibcus et al., 2018; Walther et al., 2018). The loops are not attached at specific loci, i.e. they are variable (Gibcus et al., 2018). This observation explains why in mitotic chromosomes specific sequences for condensin attachment is not evident, and why condensin enrichment is mainly found at repetitive sequences such as at centromeres, tRNA and rRNA (Kim et al., 2013; Kim et al., 2016; Piazza et al., 2013).

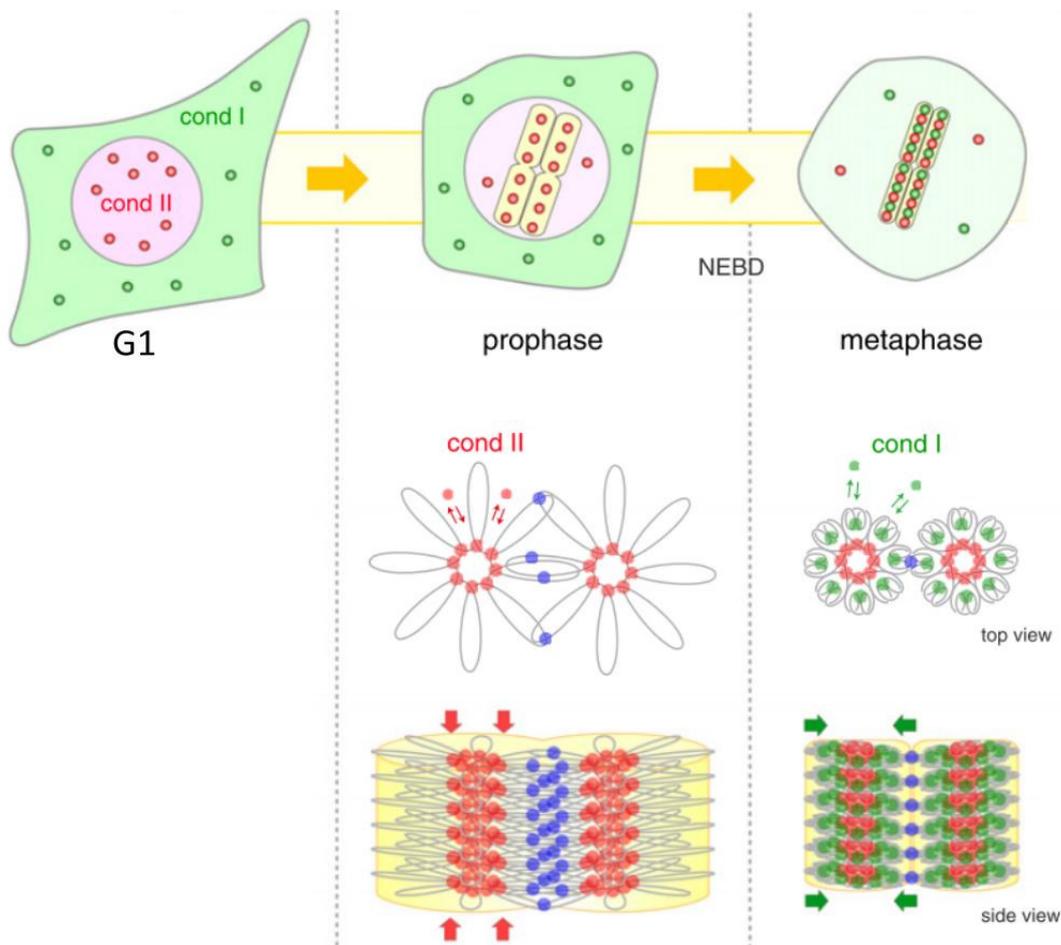


Figure 6. Condensin localization and function in vertebrates. During interphase, condensin I (green) and condensin II (red) localize in the cytoplasm and the nucleus, respectively. In S phase, condensin II starts to localize within the chromosomes helping to resolve the sister chromatids. At prophase, it accumulates in the chromosomes and mediates the axial compaction of the chromosomes. Cohesins (blue) mediate the cohesion of the sister chromatids. After nuclear envelope break-down (NEBD), condensin I gains access to the chromosomes and compacts them laterally. During metaphase condensin I and II localize to the mitotic chromosomes in an alternate fashion. Modified from Hirano, 2012.

The location of condensin II in the nucleus suggests an interphase-specific function which differs from the function during the mitotic chromosome formation. Accordingly, during interphase, condensin II is also involved in gene expression and chromatin organization (Wallace & Bosco, 2013).

The best examples for gene regulation are found in *Caenorhabditis elegans*, which encodes a third condensin complex, the Dosage Compensation Complex (DCC). The DCC ensures the equal expression of X chromosome-linked genes in hermaphrodites (two X chromosomes) and males (one X chromosome). In *Drosophila*, CAP-D3 together with the RetinoBlastoma protein (RB), regulate gene clusters involved in tissue-specific programs (Longworth et al., 2012). And in mouse and human, condensin II localizes to the promoters of active genes and is required for normal gene expression (Downen et al., 2013; Yuen et al., 2017).

Condensin II also organizes chromatin during interphase (Wallace & Bosco, 2013). In *Drosophila*, condensin II promotes the formation chromosome territories, ensures the individualization of the chromosomes and the dispersion of repetitive sequences (Bauer et al., 2012; Hartl et al., 2008b).

Although condensin I has been repeatedly reported to be only cytoplasmatic during interphase, some studies also address its presence in the nucleus during interphase. Budding yeast condensin, which is analogous to condensin I, is present in the nucleus during interphase. It localizes to centromeres and RNA polymerase III transcribed genes, such as the tRNA and the 5S rDNA, and it is essential for rDNA condensation and transmission (D'Ambrosio et al., 2008b; Freeman et al., 2000). In human, a subpopulation of CAP-D2 is nuclear during interphase (Schmiesing et al., 2000). *Drosophila* CAP-D2 is predominantly nuclear during interphase and is required for the resolution of sister chromatids (Savidou et al., 2005). In chicken, condensin I is needed for the correct condensation of the rDNA and a heterochromatic region of the chromosome Z, and its depletion affects gene expression, suggesting a role in transcription (Zhang et al., 2016).

In *Arabidopsis*, condensins have not been studied as widely as in other organism, but the components for both condensin complexes are present (Schubert, 2009; Smith et al., 2014). In contrast to other organisms, *Arabidopsis* has two SMC2 family members, SMC2A and SMC2B. Both proteins must have redundant functions since single mutants are viable but the double mutant is embryo lethal (Siddiqui et al., 2003). SMC4 mutants are also inviable, showing the importance of condensin for normal plant development (Siddiqui et al., 2006). Condensin I and II subunits have a

different subcellular location in Arabidopsis. CAP-H is present in the cytoplasm during interphase (Fujimoto et al., 2005) while the condensin II subunits CAP-H2 and CAP-D3 are mainly nuclear. However, CAP-H2 was mainly detected in the nucleolus while CAP-D3 is absent from it (Fujimoto et al., 2005; Schubert et al., 2013). CAP-H2, CAP-H and SMC4 localize in the chromosomes during mitosis (Fujimoto et al., 2005; Smith et al., 2014). As in other organism, the chromosomes condense in Arabidopsis condensin mutants, but they present abnormal shapes and segregation defects (anaphase bridges and chromatin threads between the chromosomes) (Smith et al., 2014). CAP-D2 is needed for the normal organization of the centromeres and the rDNAs in meiotic chromosomes (Smith et al., 2014). Like in *Drosophila*, condensin II is involved in the organization of chromatin during interphase. Arabidopsis CAP-D3 prevents centromeric heterochromatin associations and induces chromatin compaction. However, the condensin I protein CAP-D2 participates also in both processes (Schubert et al., 2013).

Overall, the role of the condensins as organizers of the nucleus and chromosomes seem to be conserved in Arabidopsis. Besides its structural functions, condensin is also involved in the response to DNA damage caused by boron (Sakamoto et al., 2011), in the silencing of pericentromeric transposons and in correct gene expression (Wang et al., 2017).

2. Aims

Plants need a strict regulation of transcription to respond effectively to environmental changes. Transcription and its regulation occur during interphase, when chromatin is more relaxed and proteins can access the DNA. Also during interphase, the spatial folding of the chromatin allows interactions between loci and regulatory sequences. The analysis of the nuclear organization during interphase is thus required to understand the regulation of the transcription.

The condensin complexes are conserved in all eukaryotes. Their roles in shaping chromosomes and organizing the chromatin during interphase have been widely studied in yeast and animals. However plant condensins remain largely unknown. This study is a continuation of the work of Schubert et al. (2013) about the condensin subunits CAP-D2 and CAP-D3 in *Arabidopsis thaliana* with special focus on:

- i) A general characterization of CAP-D2 and CAP-D3 considering their expression pattern and cellular localization to understand how similar *Arabidopsis* condensins are to other eukaryotes condensins.
- ii) The specific role of CAP-D3 organizing the nucleus. Schubert et al. (2013) showed that CAP-D3 affects the organization of the chromosome territories and centromeric regions. Here we intend to decipher the effects of CAP-D3 on the organization of other repetitive regions and determine the relationships between CAP-D3 euchromatin organization, the epigenetic landscape and transcription.
- iii) The role of CAP-D2 as a nuclear organizer studying in depth the phenotype of plants defective for CAP-D2.

3. Materials and methods

3.1 Plant material, transformation and growing conditions

3.1.1 Arabidopsis plant material and stable transformation

Arabidopsis thaliana (L.) Heynh was used as the model plant of this study. All lines and control plants are in Columbia-0 (Col-0) background. The Arabidopsis T-DNA insertion lines were obtained from the European Arabidopsis Stock Center, except for *cap-d3* SALK (SAIL_826_B06), *cap-d3* SALK (SALK_094776) and *cap-d2-1* (SALK_044796), which were previously described and selected in our laboratory (Schubert et al., 2013), and for the double mutant *hub1-3/hub2-1* (Fleury et al., 2007) which was kindly donated by Dr. Mieke Van Lijsebettens (VIB, Ghent, Belgium). Seeds were sown in soil and germinated under short day conditions (16h dark/8h light, 18-20 °C) and then transferred to long day conditions (16h light/ 8h dark, 18-20°C) before bolting. The lines were genotyped by PCR using the primers listed in the Appendix Table 1. The presence of the T-DNA was further confirmed by sequencing (Sequencing platform, IPK, Gatersleben, Germany). All T-DNA accession numbers and loci are listed in Table 1.

Table 1. T-DNA insertion lines used in this study.

| Name | Locus | Seed Stock Number | Description |
|-------------------------|----------------------|-------------------|-------------------------|
| <i>hub 1-3</i> | At2g44950 | GABI-276D08 | (Fleury et al., 2007) |
| <i>hub 1-4</i> | At2g44950 | SALK_122512 | (Fleury et al., 2007) |
| <i>hub 1-5</i> | At2g44950 | SALK_044415 | (Y. Liu et al., 2007) |
| <i>hub1_760</i> | At2g44950 | SALK_037760 | This study |
| <i>hub1_867</i> | At2g44950 (promoter) | SALK_119867 | This study |
| <i>hub 2-1</i> | At1g55250 | GABI-634H04 | (Fleury et al., 2007) |
| <i>hub 2-2</i> | At1g55250 | SALK_071289 | (Y. Liu et al., 2007) |
| <i>hub 1-3 /hub 2-1</i> | At2g44950 /At1g55250 | | (Fleury et al., 2007) |
| <i>rbr 1-3</i> | At3g12280 | GABI-170G02 | (Ebel et al., 2004) |
| <i>rbr_029</i> | At3g12280 | SALK_096029 | This study |
| <i>rbr_478</i> | At3g12280 | SALK_071478 | This study |
| <i>csn3-1</i> | At5g14250 | SALK_000593 | (Dohmann et al., 2008b) |
| <i>csn3-2</i> | At5g14250 | SALK_106465 | (Dohmann et al., 2008b) |
| <i>csn4-1</i> | At5g42970 | SALK_043720 | (Dohmann et al., 2008b) |
| <i>csn4-2</i> | At5g42970 | SALK_053839 | (Dohmann et al., 2008b) |
| <i>csn5b-1</i> | At1g71230 | SALK_007134 | (Dohmann et al., 2005) |
| <i>csn5b-2</i> | At1g71230 | SALK_030493 | This study |
| <i>csn5b-3</i> | At1g71230 | SALK_036658 | This study |
| <i>cap-d3 SAIL</i> | At4g15890 | SAIL_826_B06 | (Schubert et al., 2013) |
| <i>cap-d3 SALK</i> | At4g15890 | SALK_094776 | (Schubert et al., 2013) |
| <i>cap-d2-1</i> | At3g57060 | SALK_077796 | (Schubert et al., 2013) |
| <i>cap-d2-2</i> | At3g57060 | SALK_044796 | This study |
| <i>cap-d2-3</i> | At3g57060 | SALK_065716 | This study |

Arabidopsis stable transformants were generated by the floral dip method (Clough & Bent, 1998). For selection of primary transformants, the seeds were sterilized and plated on ½ Murashige and Skoog (MS) basal medium (Sigma) supplemented with the adequate antibiotics when required and grown in a growth chamber under long day conditions.

3.1.2 Arabidopsis protoplast isolation and transformation

Isolation and transformation of Arabidopsis leaf protoplasts were performed as described in Yoo et al. (2007). In brief, well expanded Arabidopsis rosette leaves were collected and cut into thin strips with a razor blade. The leaf strips were incubated in an enzyme solution (1.5 % CelluloseR10 (Duchefa Biochemie) 0.4 % Macerozyme R10 (Duchefa Biochemie) until the protoplasts were released. Then, the protoplast suspension was filtered through a 100 µm nylon mesh and the protoplasts were precipitated by centrifugation 1 min at 200 g at 4 °C. After being washed with W5 buffer (154 mM NaCl, 125 mM CaCl₂, 5 mM KCl, 2 mM MES), the protoplasts were resuspended in MMG buffer (0.4 M Mannitol, 15 mM MgCl₂, 4 mM MES) to a concentration of 2x10⁵ protoplast/ml followed by PEG-transformation with 10 µg plasmid DNA per 100 µl protoplasts. The protoplasts were maintained in W1 buffer (0.5 M Mannitol, 20 mM KCl, 4 mM MES) at room temperature in darkness and analyzed the following days.

3.1.3 Arabidopsis cell suspension culture transformation

The Arabidopsis ecotype 'Landsberg erecta' cell suspension (PSB-D) was grown by shaking in an orbital shaker at 130 rpm at 25 °C in the dark in MSMO medium with adequate antibiotics for selection. PSB-D cells were transformed and upscaled as previously described (Van Leene et al., 2011) in collaboration with the group of Prof. Klaus Grasser (University of Regensburg, Germany). For transformation, the PSB-D cells were co-cultivated with *Agrobacterium tumefaciens* containing the constructs of interest in MSMO with 0.2 mM acetosyringone (Sigma). The transformed mixture was transferred to MSMO medium with Vancomycin (500 µg/ml, Duchefa) and Carbenicillin (500 µg/ml, Duchefa) to eliminate the *Agrobacterium* cells from the culture, and Kanamycin (50 mg/ml, Duchefa) for the construct's selection. Every week for the next 3 weeks the cell culture was transferred to increasing volumes of fresh MSMO with the 3 antibiotics (Vancomycin, Carbenicillin and Kanamycin) and then grown for another 3 weeks, in increasing volumes of MSMO with only the selection antibiotic to increase the cell mass. The cell suspension was collected by centrifugation and stored at -80 °C.

3.1.4 *Nicotiana benthamiana* transient transformation

The transient transformation of *N. benthamiana* leaf cells was carried out as described in Sparkes et al. (2006). *Agrobacterium* strain GV3101 carrying the constructs of interest was grown in YEB medium with suitable antibiotics to an OD₆₀₀ of 1 and resuspended in infiltration medium (10 mM MES, 10 mM MgCl₂, pH 5.6, 3.3 mM acetosyringone). *N. benthamiana* leaves of 2 to 3 weeks old plants were infiltrated with the *Agrobacterium* suspension using a syringe without needle and analyzed 2 to 4 days later. When co-infiltration of more than one construct was required, the *Agrobacterium* cultures were mixed to a 1:1 ratio each before resuspension in infiltration medium.

3.2 General methods used to characterized the condensin subunits

3.2.1 Genomic DNA isolation and PCR

Genomic DNA was isolated from leaf material. The leaves were frozen, grinded in liquid nitrogen and resuspended in DNA Extraction buffer (100 mM Tris, 0.7 M NaCl, 0.05 M EDTA pH 8.0). The DNA was extracted from this suspension with chloroform:isoamyl alcohol (24:1 ratio) followed by precipitation and washing with isopropanol and ethanol. Then, the DNA was resuspended in sterile distilled water (SDW).

Routine PCRs, as genotyping PCRs or RT-PCR, were performed with 1 µl gDNA or cDNA in a PCR mixture containing 1.25 U Taq DNA Polymerase in 1X Buffer (Qiagen), 0.8 mM dNTPs mix (Bioline) and 0.4 µmol of each primer.

3.2.2 Total RNA isolation, cDNA synthesis and quantification of RNA

Total RNA for regular procedures such as expression checking by RT-PCR or to generate cDNA templates for cloning, was extracted from leaves and flower buds with the TRIzol method (Life Technologies). Total RNA for *CAP-D2* and *CAP-D3* transcript quantification was extracted from leaves, roots, 7 days-old seedlings and flower buds with RNeasy Plant Mini kit (Qiagen) following manufacturer's instructions. All RNA samples were treated with TURBO DNase (Thermo Fisher Scientific) and tested for DNA contamination by PCR. Reverse transcription was performed using 250 ng of total RNA and the RevertAid H Minus First Strand cDNA Synthesis kit (Thermo Fischer Scientific), with oligo(dT)₁₈ primers, according to the manufacturer's instructions. The quality of the cDNA was checked with a PCR targeting EF1B mRNA (Elongation factor 1β) with the primer pair EF1BF 5'-AAACCTACATCTCCGGGATCAATT-3' / EF1BR 5'-ACAGAAGACTTTCCACTCTCTTTAG-3'.

Quantitative RT-PCRs for *CAP-D2* and *CAP-D3* transcripts were done in triplicates and from three independent biological samples using SYBR™ Green PCR Master Mix (Thermo Fischer Scientific) in a 7900HT Fast Real-Time PCR System (Applied Biosystems). For each reaction, 0.5 µl of cDNA template and 0.6 mM primers (Table 2) were used in 10 µl. *PPA2* and *At4g26410* (Kudo et al., 2016) were used as reference genes for data normalization and the data were analyzed with the Double Delta Ct method (Livak & Schmittgen, 2001).

Table 2. List of primers used for *CAP-D2*, *CAP-D3* and reference gene transcript quantification

| Primer name | Sequence 5'-3' |
|--------------------|------------------------|
| D2QRT2_F | CCACCCAAGAGAACAATGGC |
| D2QRT2_R | TGCACACTCCCAATCAGAT |
| D3QRT1_F | AGAATGACGTACAAGGGCTAGA |
| D3QRT1_R | ATCGCCAGCCCATGTAGAAG |
| PP2A_F | TAACGTGGCCAAAATGATGC |
| PP2A_R | GTTCTCCACAACCGCTTGGT |
| At4G26410_F | GAGCTGAAGTGGCTTCCATGAC |
| At4G26410_R | GGTCCGACATACCCATGATCC |

3.2.3 Cloning and construct generation

DNA fragments for cloning were produced using a high fidelity DNA polymerase (KOD Hot Start DNA Polymerase, Merck Millipore). Two types of cloning were used, traditional cloning with restriction enzymes (Thermo Fisher Scientific) and Quick ligation Kit (NEB), and Gateway cloning using the Gateway LR Clonase II enzyme mix (Invitrogen) following the manufacturer instructions. Below are the specifications for the generation of each construct and the primers are listed in Table 3. All the constructs were transformed in DH5α *E. coli* cells (NEB), isolated using QIAprep Spin Miniprep Kit (Qiagen) and confirmed by sequencing (Sequencing platform, IPK, Gatersleben, Germany). All the constructs generated are listed in Appendix Table 2.

Table 3. Primer sequences and usage. Linker sequences are in lower case letters, restriction sites are underlined and genomic sequences are written in upper case.

| Name | Sequence 5'-3' | Use |
|---------------|--|-------------------------------------|
| CAPG_pEnt_f | acgt <u>GTCGAC</u> ATGGGCGAAGAATCAGAAATC | CAP-G cloning into pEntry |
| CAPG_pEnt_r | atta <u>GCGGCCGC</u> gaTTCATCTGAATCATCTGCTGT | CAP-G cloning into pEntry |
| CAPH_pentry_f | actg <u>GTCGAC</u> ATGGATGAATCCTTAACTCCA | CAP-H cloning into pEntry |
| CAPH_pentry_r | atta <u>GCGGCCGC</u> gagGGCAAGGTGTATTGTTAGATCA | CAP-H cloning into pEntry |
| D2CtSalI_F | act <u>GTCGAC</u> taGGTTCTGTTGAGAAGAATCTG | CAP-D2 Ct cloning into pET23a |
| D2CtNotI_R | tatt <u>GCGGCCGC</u> ACTTCTACTTCCTGACCT | CAP-D2 Ct cloning into pET23a |
| D2SgRNA1F | attgATCACTATCTGCTGGAAGAC | Sg1 protospacer for pEn-Chimera |
| D2SgRNA1R | aaacGTCTTCCAGCAGATAGTGAT | Sg1 protospacer for pEn-Chimera |
| D2SgRNA2F | attgTGTTATTCAGCGGTTCTCCG | Sg2 protospacer for pEn-Chimera |
| D2SgRNA2R | aaacCGGAGAACCGCTGAATAACA | Sg2 protospacer for pEn-Chimera |
| D2SgRNA3F | attgGATCCATCAATGGAAGAATC | Sg3 protospacer for pEn-Chimera |
| D2SgRNA3R | aaacGATTCTTCCATTGATGGATC | Sg3 protospacer for pEn-Chimera |
| D2SgRNA4F | attgCAGAGTCATCGAGCAGCATC | Sg4 protospacer for pEn-Chimera |
| D2SgRNA4R | aaacGATGCTGCTCGATGACTCTG | Sg4 protospacer for pEn-Chimera |
| D2-392F | gtgc <u>GTCGAC</u> CTCAAAGCTTTTCTGCTTC | CAP-D2 promotor cloning into pEntry |
| D2-1156F | gtgc <u>GTCGACT</u> GGTACTGAAGCTAAGAAGG | CAP-D2 promotor cloning into pEntry |
| D2ProR | gaag <u>GCGGCCGC</u> TTTTTCTAGAGAGAGAGAGA | CAP-D2 promotor cloning into pEntry |
| D2Int1R | caat <u>GCGGCCGC</u> TCAGAAAGGTCAAAGGATAC | CAP-D2 promotor cloning into pEntry |
| D2Int2R | aaat <u>GCGGCCGC</u> TTTTTCTCCCTCGTGCTG | CAP-D2 promotor cloning into pEntry |
| D3-474F | gtgc <u>GTCGAC</u> ATTTTGTTGTCTAGAATTTG | CAP-D3 promotor cloning into pEntry |
| D3-1318F | gtgc <u>GTCGAC</u> TTTTCTCTGTTCAATAG | CAP-D3 promotor cloning into pEntry |
| D3ProR | taat <u>GCGGCCGC</u> GGCGATTCTCTACTGATAGA | CAP-D3 promotor cloning into pEntry |

3.2.3.1 Condensin subunit EYFP-fusion constructs

The 3942 bp and 4245 bp long cDNA sequences of *CAP-D3* and *CAP-D2* respectively, were synthesized and cloned into pEntr 1A (Invitrogen) by DNA-Cloning-Service (Hamburg, Germany). An intron of *Nicotiana tabacum* RubisCo was introduced after the first 1000 bp of both *CAP-D2* and *CAP-D3* synthesized sequences to avoid the potential toxic effect of long DNA sequences on bacteria. The inclusion of this intron impedes the complete transcription of long proteins that could be detrimental for bacteria growth. The 3153 bp and the 2013 bp long cDNA sequences of *CAP-H* and *CAP-G* respectively, were amplified from flower buds cDNA with the primer pairs CAPH_pentry_f/CAPH_pentry_r and CAPG_pEnt_f/CAPG_pEntr_r (Table 3) respectively, and cloned between the *SalI* and *NotI* sites of the pEntr 1A plasmid (Invitrogen).

Once in the pEntr 1A plasmid, the coding sequences of the genes of interest were subcloned into pGWB641 and pGWB642 plasmids (Nakamura et al., 2010) using Gateway cloning (Invitrogen). The generated expression cassettes contained the proteins of interest fused to EYFP C-terminally for the pGWB641 constructs (CAP-D2_EYFPc, CAP-D3_EYFPc, CAP-G_EYFPc and CAP-H_EYFPc) or N-terminally for the pGWB642 constructs (CAP-D2_EYFPn and CAP-D3_EYFPn) and both were under the control of the cauliflower mosaic virus 35S promoter. As a control (Control_EYFPc), a plasmid containing only the EYFP under the 35S promoter was generated.

3.2.3.2 Condensin I Bimolecular Fluorescence Complementation (BiFC) constructs

The coding sequences of *CAP-D2*, *CAP-G* and *CAP-H* previously cloned in the pEntr 1A plasmid (described above) were subcloned using Gateway cloning (Invitrogen) in the SPYNE and SPYCE plasmids (Walter et al., 2004). The final constructs CAP-D2_SPYNE, CAP-G_SPYNE and CAP-H_SPYNE have the proteins of interest fused upstream of the N-terminal part of EYFP (amino acids 1-155) and CAP-D2_SPYCE, CAP-G_SPYCE and CAP-H_SPYCE to the C-terminal part of the EYFP (amino acids 156-239); all sequences are under control of the 35S promoter.

3.2.3.3 CAP-D2 recombinant protein expression construct

The sequence between 2743 and 4248 bp of *CAP-D2* (C-terminal 500 amino acids) was amplified from Arabidopsis flower buds cDNA with the D2CtSall_F and D2CtNotI_R primers (Table 3). The fragment was cloned between the *Sall* and *NotI* restriction sites of the pET23a(+) plasmid (Novagen) resulting in a pEt23_CAP-D2_Ct construct which contains the cassette T7 promoter::T7 tag-CAP-D2Ct-His tag.

3.2.3.4 CAP-D2 and CAP-D3 affinity purification constructs

The cDNA sequences of *CAP-D3* and *CAP-D2* were synthesized and cloned into pCambia 2300 35S GS-Ct, kindly donated by Prof. Klaus Grasser (University of Regensburg, Germany), by DNA-Cloning-Service (Hamburg, Germany) resulting in the constructs pCambia2300_CAP-D2_GS and pCambia2300_CAP-D3_GS. The constructs contain the coding sequence of CAP-D2 and CAP-D3, respectively, under the 35S promoter and a GS-tag fused to the C-terminal part of the protein. As explained above, the synthesized sequences contain one intron of the *Nicotiana tabacum* RubisCo sequence.

3.2.3.5 CAP-D2 and CAP-D3 promoter-GUS reporter lines

Different lengths of the promoter regions of both *CAP-D2* and *CAP-D3* were cloned between the *SalI* and *NotI* restriction sites of the pEntr 1A plasmid (Invitrogen). The sequences were amplified from leaf gDNA with the primer pairs D2-1156F/D2ProR for the Pro4_D2 fragment, D2-1156F/D2Int1R for Pro5_D2, D2-1156F/D2Int2R for Pro6_D2, D2-392F/D2ProR for Pro7_D2, D2-392F/D2Int1R for Pro8_D2, D2-392F/D2Int2R, for Pro9_D2, D3-1318F/D3ProR for Pro10_D3 and D3-474F/D3ProR for Pro11_D3 (Table 3). In total six versions of long/short promoters including/excluding the first and second intron were cloned for *CAP-D2* (Pro4_D2 to Pro9_D2) and two, long/short promoters, for *CAP-D3* (Pro10_D3 and Pro11_D3). The fragments were subcloned upstream of the GUS reporter gene in the pGWB633 plasmid (Nakamura et al., 2010) using Gateway cloning (Invitrogen).

3.2.3.6 CAP-D2 CRISPR-Cas 9 constructs

Four protospacer sequences were designed to target *CAP-D2*. Due to the small size of the protospacers (20 nucleotides), they were generated by oligo annealing (Table 3) and cloned first into pEn-Chimera and then subcloned into pDE-CAS9 (Fauser et al., 2014) following the protocol described in Schiml et al., 2016. The final constructs (pDeCas Sg1, pDeCas Sg2, pDeCas Sg3 and pDeCas Sg4) include an expression cassette with the SpCas9 protein under the ubiquitin4-2 promoter from *Petroselinum crispum* (PcUbi4-2 promoter) and the specific Single guide RNA (SgRNA) under the Arabidopsis ubiquitin AtU6-26 promoter.

3.2.4 DNA sequence analysis

Sequence alignment, editing and chromatogram checking were done with BioEdit v7.2.6.1 and Genome Compiler v0.6.0 (Genome Compiler Corporation). The later was also used for sequence annotation and *in silico* cloning to generate maps of the constructs and plasmids. Primers were designed with Primer3Plus (Untergasser et al., 2007) and NetPrimer (Premier Biosoft).

3.2.5 Total protein extraction and Western blot

Isolated protoplast or grinded plant leaf material were resuspended in 100-300 µl of protein extraction buffer (56 mM Na₃CO₃, 56 mM DTT, 2% SDS, 12% Sucrose, 2 mM EDTA, bromophenol blue), incubated 20 min at 65 °C and centrifuged at high speed. Then, the supernatant containing the soluble total protein was used for Western blot.

For Western blot, the extracted proteins were separated in 10% polyacrylamide gels (Schägger & von Jagow, 1987) and blotted into Immobilon-FI membranes (Millipore). The membranes were incubated in blocking solution (5% skim milk in TBST) for 30 min to reduce non-specific binding of the antibodies. Two types of Western blot detection were used, fluorescent and chemiluminescent detection. In the first case, the membranes were incubated overnight at 4 °C with the adequate primary antibody in TBST buffer: mouse anti-His-tag (1:2000, Millipore, 05-949), rabbit anti-CAP-D2 serum (1:1000) or rabbit anti-GFP conjugated with Alexa 488 (1:1000, Chromotek, PABG1); and then with a secondary antibody: anti-mouse IgG IRDye 680RD (1:10000, LI-COR, 926-32222) or anti-rabbit IgG IRDye800CW (1:5000, LI-COR, 925-32213). The membranes were imaged using a LI-COR Odyssey Imager (LI-COR). For T7-tag detection, chemiluminescent detection was used: the membranes were first incubated with anti-T7-tag conjugated with alkaline phosphatase (1:10000, Merck, #6999) for 1 h, and then with phosphate-activity buffer (100mM Tris, 100 mM NaCl, 1 mM MgCl₂, 0.33 mg/ml nitro blue tetrazolium and 0.17 mg/ml 5-bromo-4-chloro-3-indolyl phosphate) in the dark until the signals were visible.

3.2.6 CAP-D2 and CAP-D3 affinity purification and mass spectrometry

The constructs pCambia2300_CAP-D2_GS and pCambia2300_CAP-D3_GS were transformed into a PSB-D Arabidopsis cell suspension as described above. Transformation of PSB-D cell suspension and co-immunoprecipitation were done in collaboration with the group of Prof. Dr. Klaus Grasser (Plant Chromatin Group, Institute of Plant Science, University of Regensburg). CAP-D2-GS and CAP-D3-GS were affinity purified following the protocol described in Dürr et al. (2014). In brief, 15 g of cell material were resuspended and sonicated in extraction buffer (25 mM HEPES, 100 mM NaCl, 0.05% NP-40, 1 mM DTT, 2 mM MgCl₂, 5 mM EGTA, 10 % glycerol and protease inhibitors). The protein solution was incubated with magnetic beads conjugated with IgG (IgG presents high affinity binding to the GS-tag). After applying a magnet to the mixture, the beads pulled-down CAP-D2-GS or CAP-D3-GS and interacting proteins separating them from the rest of the protein mixture, which was washed away. The purified proteins were eluted in 0.1 M glycine and analyzed by mass spectrometry.

For mass spectrometry, the eluted proteins were separated in a 10 % polyacrylamide gel and digested with trypsin. Mass spectrometry and data analysis were performed according to Antosz et al. (2017). Briefly, peptides were separated on an Ultimate 3000 RSLC nano System (ThermoScientific) by reversed-phase chromatography using a Repronil-Pur Basic C18 nano column

(Dr. Maisch GmbH) in a linear gradient of 4 to 40% acetonitrile in 0.1% formic acid. The LC System was coupled to a MaXis plus UHR-QTOF system (Bruker Daltonics) via a nanoelectrospray source (Bruker Daltonics). Data-dependent acquisition of tandem mass spectrometry (MS/MS) spectra by collision-induced dissociation fragmentation was performed using a dynamic method with a fixed cycle time of 3 s (Compass 1.7; Bruker Daltonics). Protein Scape 3.1.3 (Bruker Daltonics) in connection with Mascot 2.5.1 (Matrix Science) facilitated database searching of the NCBI nr database

Three independent affinity purifications were performed; a MASCOT score of minimum 100 and the presence in at least two of the purifications were considered as criteria for reliable protein identification. The experimental background (contaminating proteins that co-purified with the unfused GS-tag) and non-specific interactions (proteins that co-purified independently of the bait used) were subtracted. The list of non-specific Arabidopsis proteins is based on 543 affinity purification experiments using 115 different baits (Van Leene et al., 2014).

3.2.7 CAP-D2 and CAP-D3 promoter analysis (β -glucuronidase activity assay)

Arabidopsis stable transformants carrying the β -glucuronidase (GUS) gene fused to different versions of *CAP-D2* and *CAP-D3* promoters (promoter-GUS reporter constructs) were selected in ½ MS (Sigma) with 16 mg/L PPT (Duchefa). One month after sowing, the plantlets were stained for GUS activity according to Jefferson et al. (1987) with small modifications. Arabidopsis plantlets were collected in 15 ml tubes containing 1 % X-Glu (5-Bromo-4-Chloro-3-indolyl- β -D-Glucopyranoside; Duchefa) in 0.1 M phosphate buffer (pH 7.0). To facilitate the penetration of the solution in the material, the tubes with the plant material and the staining solution were vacuumed in a speed vacuum for 5 min and incubated overnight at 37 °C. Next day, the staining solution was replaced by 70 % ethanol and incubated 20 min at 60 °C. This step was repeated until the chlorophyll was removed. The stained material was preserved in 70 % ethanol at 4 °C and analyzed under a stereo microscope.

3.2.8 Nuclei isolation and flow cytometry

Arabidopsis nuclei of differentiated cells were isolated and flow-sorted according to their ploidy level as previously described (Weisshart et al., 2016). Leaf material was fixed in 4 % formaldehyde in Tris buffer (10 mM Tris, 10 mM NaEDTA, 100 mM NaCl, 0.1 % Triton-X, pH 7.5) for 20 min on ice under vacuum. The leaves were washed in Tris buffer and chopped with a razor blade in

chromosome isolation buffer (15 mM Tris, 2 mM Na₂EDTA, 0.5 mM spermin, 80 mM KCl, 15 mM 2-mercaptoethanol, 1% Triton X-100). The suspension was filtered through a 35 µm mesh and used, after 4'6-Diamino-2-Phenylindole (DAPI, Thermo Fischer Scientific) staining, for Fluorescence Activated Cell Sorting (FACS) in a BD INFLUX Cell Sorter (BD Bioscience). The nuclei were sorted based on their DNA content in 2C, 4C, 8C and 16C ploidy fractions. 1C is the DNA content of a haploid not replicated nucleus.

3.2.9 Slide preparation with flow-sorted nuclei

Twelve µl of sorted nuclei and the same amount of sucrose buffer (10 mM Tris, 50 mM KCl, 2 mM MgCl-6H₂O, 5% sucrose, pH 8.0) were placed over a clean glass slide for Structured Illumination Microscopy (SIM) or over a clean high precision coverslip for PhotoActivated Localization Microscopy (PALM) and air dried at room temperature overnight. The slides and coverslips were directly used or stored at -20 °C until further use (Weisshart et al. 2016).

3.2.10 Preparation of squashed Arabidopsis roots

Arabidopsis seeds were grown on a wet filter paper and fixed after 3-4 days in 4 % paraformaldehyde in phosphate-buffered saline buffer (1xPBS buffer). The seedlings were washed in 1xPBS buffer and digested for 30 min at 37°C in an enzyme mix (0.5 % pectolyase (Sigma), 0.5 % cytohelicase (Sigma), 0.35 % cellulose (Calbiochem), 0.35 % cellulose (Duchefa) in 1xPBS buffer. After removal of the enzyme and 1xPBS washing, the root tips were transferred to a clean slide and squashed between coverslip and slide. The liquid nitrogen frozen coverslip was lifted and the slide directly used for immunostaining.

3.2.11 Probe preparation and Fluorescence *in situ* Hybridization (FISH)

The probes were generated using four different methods: (i) by PCR for the 180 bp centromeric repeat (pAL probe; Martinez-Zapater et al. 1986), (ii) from a plasmid for the 5S rDNA probe (pCT4.2; Campell et al. 1992), (iii) from BACs containing the 45S rDNA repeats (BAC T15P10), and (iv) and for painting part of Chromosome Territory 1 Bottom (CT1Bp) from BACS arranged in contigs (BACs F11P17 to F12B7). The BACs were obtained from the Arabidopsis Biological Resource Center (Ohio, USA). The probes were labeled with modified dUTPs conjugated with Texas-red (Invitrogen) or Alexa-488 (Invitrogen) by nick-translation. The FISH was performed as previously described (Pecinka et al. 2004). The slides were rinsed in 2xSSC, treated with pepsin (50 µg / ml in 0.01 M HCl) for 2 min at 38 °C, washed in 2xSSC, post-fixed in 4% formaldehyde in 2xSSC for 10

min at room temperature and washed again in 2xSSC. The slides were then dehydrated by washing in an ethanol series (70 %, 80 % and 96 % ethanol) for 2 min each except for the slides with the nuclei embedded in acrylamide. Per slide 1 μ l of labeled probe was used, except for CT1bottom painting in which 18 μ l of pooled BAC probes per slide were used. All the probes were precipitated with ethanol, resuspended in 20 μ l of HB50 (20 % dextranosulfat, 50 % formamid, 50 mM phosphate buffer in 2xSSC) and denaturated for 5 min at 90 °C. The probes were hybridized on the slide for 3 min at 80 °C, and then incubated at 37 °C overnight in a wet chamber. After FISH the slides were first washed three times in 50 % formamid in 2xSSC and then in 2xSSC at 42 °C, dehydrated in an ethanol series, embedded and stained with DAPI in antifade (Vectaschild; Vectorslab).

3.2.12 Indirect immunofluorescence labeling

Nuclei and chromosome preparations were washed in 1xPBS and incubated for 30 min at 37 °C in a moist chamber with 30 μ l blocking buffer (4 % BSA, 0.1 % Tween-20 in 1xPBS) to reduce non-specific antibody binding. After three washes in 1xPBS, the slides were incubated with the primary antibody diluted in antibody buffer (1 % BSA, 0.1 % Tween-20 in 1xPBS) overnight at 4 °C. Next day, the slides were washed in 1xPBS again and incubated with the secondary antibody in antibody buffer for 1 h at 37 °C. After incubation, the preparations were washed in 1xPBS, dehydrated in an ethanol series (70 %, 90 % and 96 % ethanol for 2 min each) and counterstained with DAPI in antifade (Vectashield; Vectorslab). All primary and secondary antibodies, and the dilutions used are listed in Table 4.

Immunolocalization of 5-methyl-cytosine requires an initial DNA denaturation of the specimen. Slides with sorted nuclei were denaturated in 70 % formamid in 2xSSC for 2 min at 70 °C. The preparations were dehydrated in ice cold 70 % and 96 % ethanol for 5 min each, and then air-dried. Subsequent blocking and antibody incubations were carried as described above.

3.2.13 Microscopy and image analysis

Wide field fluorescence microscopy was used to evaluate and image the nuclei and chromosome preparations with an Olympus BX61 microscope (Olympus) and an ORCA-ER CCD camera (Hamamatsu). When higher resolution was needed to analyze substructures and spatial arrangements of immunosignals and chromatin, a super-resolution fluorescence microscope Elyra

PS.1 (Zeiss) was used with the ZEN software. This microscope allows applying spatial Structured Illumination Microscopy (SIM), which enhances the lateral resolution up to ~120 nm.

Processing and analysis of microscopic image stacks were done with ZEN (Zeiss), Adobe Photoshop CS5 (Adobe) and Imaris 8.0 (Bitplane) software. Gray-scale pictures taken by light microscopy were colored and combined by Adobe Photoshop.

Plant leaf tissues cause high autofluorescence derived from chlorophyll and other secondary metabolites. In order to accurately detect the signal of interest in BiFC experiments, *N. benthamiana* leaf preparations were analyzed under a confocal laser scanning microscope (Zeiss LSM 510 META), which is able to decompose the complex emission spectra of the sample and separate the unique spectral signature of the desired fluorophore, in this case EYFP.

Table 4. Antibodies and their dilutions used for immunolocalization.

| Antibody name | Primary/Secondary | Dilution used | Reference |
|--------------------------|-------------------------------------|---------------|-----------------------------|
| Anti-CAP-D3 | Primary; Guinea Pig | 1:200 | Schubert et al. 2013 |
| Anti-CAP-D2 serum | Primary; Rabbit | 1:200 | This study |
| Anti-H3K27me3 | Primary; Rabbit | 1:100 | Merck 07-449 |
| Anti-H3K9me1 | Primary; Rabbit | 1:200 | Merck 07-395 |
| Anti-H3K9me2 | Primary; Rabbit | 1:100 | Merck 07-441 |
| Anti-H3K4me3 | Primary; Rabbit | 1:200 | Merck 07-473 |
| Anti-H3K9ac | Primary; Rabbit | 1:500 | Nobus Biological NBP2-44095 |
| Anti-H3K14ac | Primary; Rabbit | 1:1000 | Merck 07-353 |
| Anti-H3K18ac | Primary; Rabbit | 1:1500 | Abcam ab1191 |
| Anti-H3K9+14+18+23+27 ac | Primary; Rabbit | 1:500 | Abcam ab47915 |
| Anti-5methylcytosine | Primary; Mouse | 1:100 | Abcam ab10805 |
| Anti-GFP 488 | Primary conjugated with Dylight 488 | 1:1000 | Rockland 200-341-215 |
| Anti-Rabbit 488 | Secondary | 1:100 | Dianova 711-545-152 |
| Anti-Rabbit Rhodamine | Secondary | 1:300 | Jackson 111-025-003 |
| Anti Mouse 488 | Secondary | 1:50 | Molecular probes A11001 |
| Anti Guinea Pig 488 | Secondary | 1:100 | Molecular probes A11073 |

3.2.14 Gene and protein identification numbers

Sequence data from this study can be found in The Arabidopsis Information Resource (TAIR) or National Center for Biotechnology Information (NCBI) databases under the following gene identification numbers: *CAP-D2*, AT3G57060; *CAP-D3*, AT4G15890; *CAP-G*, AT5G37630; *CAP-H*, AT2G32590.

The protein sequences for CAP-D2 can be found in the InterPro database under the following identification numbers: *Arabidopsis thaliana* F4J246; *Homo sapiens* Q15021; *Mus musculus*

Q8K2Z4; *Xenopus laevis* Q9YHY6; *Drosophila melanogaster* Q9VAJ1 and *Gallus gallus* A0A1D5PW66.

3.3 CAP-D3 characterization

3.3.1 Chromosome territory quantification

One fourth of the Arabidopsis Chromosome Territory 1 bottom (CT1Bp) was determined by FISH as described above on sorted nuclei from *cap-d3 SAIL*, *cap-d3 SALK* and wild-type control plants. The centromeric 180 bp repeat was included in the FISH as a control. CT1Bp signal was quantified on 16-bit gray scale microscopic images using ImageJ v1.50i (Schneider et al., 2012). The images were taken from preparations of flow-sorted nuclei. Since this technique flattens the nuclei, they were considered as two-dimensional. Within each datasets all the images were treated the same way after using the same acquisition parameters. For the CT1Bp signal image dataset, the background was subtracted with the option 'Rolling ball' set at 25 pixels and the delimitation of the region of interest (ROI) with the RenyiEntropy threshold. For the nuclear area image dataset (measured based on DAPI staining), the background was not subtracted and the nuclear area was delimited as a ROI also with the RenyiEntropy threshold. The area of each ROI was automatically measured by the program.

3.3.2 Preparation of *cap-d3* mutants nuclei with preserved 3D structure

Arabidopsis nuclei were embedded in acrylamide to preserve their 3D structure following the procedure described by Kikuchi et al., 2005 with modifications. The nuclei were isolated as described above from wild-type plants and the *cap-d3 SAIL* and *cap-d3 SALK* mutants. Twelve μ l of nuclei suspension were mixed on a slide with 6 μ l of active 15 % acrylamide embedding medium (15 % acrylamide/bisacrylamide (29:1), 15 mM PIPES, 80 mM KCl, 20 mM NaCl, 2 mM EDTA, 0.5 mM EGTA, 0.5 mM spermidine, 0.2 mM spermine, 1 mM DTT, 0.32 M sorbitol, 2 % APS and 2 % Na₂SO₃). A coverslip was carefully placed on top of the acrylamide-nuclei mixture and let to polymerize 30 min at room temperature. The coverslip was then removed letting a thin pad of nuclei embedded in acrylamide on the slide that was directly used for FISH with the centromeric 180 bp repeat. The reconstruction of spatial Arabidopsis nuclei was generated from SIM image stacks using the Imaris software.

3.3.3 Epigenetic landscape of the *cap-d3* mutants

The epigenetic landscape of *cap-d3* mutants was evaluated with two different methods.

3.3.3.1 Distribution of DNA 5-methyl-cytosine, histone H3 methylation and acetylation

The distribution of 5-methyl-cytosine (5MC), H3K27me₃, H3K9me₁, H3K9me₂, H3K4me₃, H3K9ac, H3K14ac, H3K18ac, H3K9+14+18+23+27 ac was evaluated on flow-sorted 4C nuclei from control plants and the *cap-d3 SAIL*, *cap-d3 SALK* mutants and by immunofluorescence performed as described above.

3.3.3.2 Centromeric DNA methylation

Southern blot analyses were performed to compare CpG centromeric DNA methylation between *cap-d3* mutants and control plants. Five µg of gDNA from Arabidopsis leaves was digested with either *HpaII* or *MspI* (Thermo Fischer Scientific). Both enzymes recognize the same restriction sites, but *MspI* is insensitive to methylation while *HpaII* is sensitive. The digestion was run in 1% agarose gel over night for a complete separation of the fragments. The gel was denatured (1.5 M NaCl, 0.5 M NaOH) and neutralized (1.5 M NaCl, 1 M Tris, 1 mM EDTA) before being blotted onto a positively charge nylon membrane (Hybond- XL, Amersham) with high-salt buffer (20X SSC). After the transfer, the membrane was incubated in herring DNA and hybridized with ³²P-labelled centromeric 180 bp repeat pAL probe in CHURCH buffer overnight at 65 °C. Next day, the membrane was washed in decreasing concentrations of SSC (2x, 0.5x and 0.1x) and the signals were detected by autoradiography using an Imager-Plate and scanned with a phosphorimager. The Arabidopsis centromeric pAL probe was generated by PCR and ³²P-labeled according to manufacturer's instructions (Deca-Label DNA labelling Kit, Thermo Scientific).

3.3.4 *cap-d3* RNA-seq and transcriptome analysis

cap-d3 SAIL, *cap-d3 SALK* and control (Col-0) seeds were sown in soil under short day conditions. RNA was extracted with the RNeasy Plant Mini Kit (Qiagen) from 50 mg of pooled 4 weeks-old plantlets cut above the root. For each of the three Arabidopsis genotypes five independent RNA extractions were performed and the RNA integrity of the samples was measured in a 2100 Bioanalyzer (Agilent). The four RNA samples of each genotype with the highest RIN (RNA Integrity Number) were used for library preparation and RNA sequencing (NGS platform, IPK, Gatersleben, Germany). The libraries were prepared with TruSeq RNA Library Kit (Illumina) unstranded and were sequenced in a HiSeq2000 system.

The quality of the RNA-seq reads were assessed with FastQC v0.11.4 (Babraham Bioinformatics) and adaptors trimmed with Trimmomatic v0.32 (Bolger et al., 2014). After a second quality check in FastQC, the reads were aligned with GSNAP v2016-05-25 (Wu & Nacu, 2010) against the Arabidopsis TAIR10 genome and the gene counts calculated with HTseq v0.6.1 (Anders et al., 2015). Differential expression analyses were performed using the DESeq2 1.14.0 Bioconductor package (Love et al., 2014). Genes were considered differentially expressed (DEG) when they had a Benjamini-Hochberg-adjusted-*P* value ≤ 0.05 and a \log_2 -fold change ≤ -1 or ≥ 1 . Genes detected as differentially expressed for both *cap-d3* mutants were considered as the genes associated to CAP-D3 defective proteins independently of the specific mutation. Gene enrichment was analyzed with PLAZA 3.0 (Proost et al., 2015) and agriGO v1.2 (Du et al., 2010). The analysis of the transcription factors present in *cap-d3* DEG was performed with the Arabidopsis Transcription Factor Database (AtTFDB) from the Arabidopsis Gene Regulatory Information Server (AGRIS; Yilmaz et al., 2011).

3.4 CAP-D2 characterization

3.4.1 Anti-CAP-D2 antibody production

The construct pEt23_CAP-D2_Ct, which contains the last 500 amino acids of CAP-D2 fused to a His-tag, was transformed into *E. coli* BL21 cells and the expression of the transgene induced with 1 mM IPTG (isopropyl- β -D-thiogalactopyranoside, Sigma-Aldrich). The recombinant protein was purified with agarose beads that bind specifically to the His-tag (Ni-NTA Agarose, Qiagen) following the purification hybrid method from the ProBond purification system (Thermo Fisher Scientific). The purified recombinant protein was used to produce an anti-CAP-D2 polyclonal antibody in rabbit (Phytoantibodies group, IPK, Gatersleben, Germany). Two rabbits were immunized with the recombinant protein and the anti-serum collected after four immunizations. The anti-CAP-D2 serum was not further purified. It was tested for sensitivity and specificity, and used directly for Western blot and immunolocalization.

3.4.2 Condensin I Bimolecular Fluorescence Complementation (BiFC)

For confirmation of interactions found between condensin I subunits they were further examined by BiFC. This technique is based in the restoration of a split fluorescent protein upon interaction of the two proteins being tested (Hu et al., 2002): the proteins of interest are fused each one to the N-terminal or the C-terminal part of EYFP, the constructs are co-transformed in *N. benthamiana* and if the proteins interact with each other, the fluorescence of the restored EYFP will be visible.

N. benthamiana 3 to 4 weeks-old plants were transiently transformed with all the possible pair combinations between CAP-D2, CAP-G and CAP-H SPYNE (N-terminal EYFP) and SPYCE (C-terminal EYFP) constructs that could restore the split EYFP protein. The transcription factor bZIP63, known to form homodimers, was used as a positive control (Walter et al., 2004) and SPYNE and SPYCE empty vectors as a negative control. Plants were analyzed 2-4 days after infiltration.

3.4.3 Arabidopsis condensin I localization in protoplasts

The protoplasts were isolated and transformed as described above with the constructs CAP-D2_EYFPc, CAP-G_EYFPc, CAP-H_EYFPc and control_EYFPc. Following a fixation in 4% formaldehyde in 1xPBS and washes in 1xPBS, the protoplasts were centrifuged at 400 rpm for 5 min (Shandon CytoSpin3, GMI inc) onto a microscopic slide. The slides were directly used for immunostaining against EYFP.

3.4.4 Meiotic analysis of the *cap-d2* T-DNA insertion mutant

The fixation and slide preparation of Arabidopsis pollen mother cells were performed according to the method described by Sanchez Moran et al., (2001). *cap-d2-1* T-DNA insertion line mutant flower buds were collected and fixed in Carnoy's fixative (6:3:1, ethanol:chloroform:acetic acid). To obtain chromosomes spreads, fixed flower buds were washed in 3:1 (ethanol:acetic acid) and 10 mM citrate buffer (10 mM citric acid, 10 mM sodium citrate, pH 4.5) before digestion in an enzyme mix (0.3 % pectolyase (Sigma), 0.3 % cytohellicase (Sigma), 0.3 % cellulose (Duchefa)) in citrate buffer for 2 h at 37 °C. The anthers were dissected, placed on a slide with a drop of citrate buffer and macerated soften with a needle until a homogeneous mix was obtained. Ten µl of 60 % acetic acid were dropped on the mix and the slide was placed over a warming plate at 42 °C for 1 min. Another 10 µl of 60 % acetic acid were added off the plate and the mix surrounded with 3:1 fixative. The slides were air-dried and used for FISH against the 5SrDNA and the 45SrDNA as described above.

3.4.5 CRISPR-Cas 9 *in vitro* assay and the generation of *cap-d2* mutants

The CRISPR/Cas system is based on two components, a nuclease (Cas9) that cleaves a double-stranded DNA target, and a single-guide RNA (sgRNA) that determines the target specificity. The Cas9 protein produces a double strand break in the target DNA which will be repaired by non-homologous end joining (NHEJ). NHEJ is an error prone repair system that can introduce small indels when repairing the DNA, and therefore can be used for mutagenesis (Fauser et al., 2014).

The *Streptococcus pyogenes* Cas9 nuclease (SpCas9), codon-optimized for Arabidopsis was used (Fauser et al., 2014). The sgRNA comprises three parts: a 20 nucleotide long protospacer complementary to the DNA target; a protospacer-adjacent motif (PAM), 3 nt needed for the binding and cleavage by Cas9; and a trans-activating CRISPR RNA (tracrRNA) that recruits the Cas9 protein. For *S. pyogenes* the PAM sequence is NGG, hence a protospacer can be designed to target any *CAP-D2* region ending NGG.

The efficiency of the 4 SgRNAs (Sg1 to Sg4) guiding the Cas9 protein to its target was tested *in vitro* based on the Guide-it sgRNA *In Vitro* Screening System (Takara Bio). In this system, a template containing the SgRNA target is created by PCR and mixed with the SgRNA of interest and Cas9 nuclease. The SgRNA guides the Cas9 nuclease to the template, which cleaves it. The SgRNAs were produced *in vitro*. A DNA fragment containing the T7 promoter and the protospacer sequence was generated by PCR (primers in Table 5), and then transcribed with T7 pol mix (HiScribe T7 high yield RNA, NEB) following manufacturer's protocol. The template was generated by amplifying a fragment of approximately 2 kb containing the SgRNA target by PCR. Thirty ng of the purified SgRNA, 30 nM *S. pyogenes* Cas9 nuclease (NEB) and 100 ng of the template were mixed and incubated 1 h at 37 °C. The mixture was incubated 10 min at 70 °C to stop the reaction and run in a 1% agarose gel to separate and visualize the cleaved products.

Table 5. Primers used for in vitro Cas9 assay and *cap-d2* mutants screening. Underlined sequences correspond to T7 promoter and bold letters to the protospacer sequence.

| Primer name | Sequence 5'-3' |
|--------------|---|
| T7_SgRNA1_F | <u>GCGGCCTCTAATACGACTCACTATAGGG</u> ATCACTATCTGCTGGAAGAC |
| T7_SgRNA2_F | <u>GCGGCCTCTAATACGACTCACTATAGGG</u> TGTTATTCAGCGTTCTCCG |
| T7_SgRNA3_F | <u>GCGGCCTCTAATACGACTCACTATAGGG</u> GATCCATCAATGGAAGAATC |
| T7_SgRNA4_F | <u>GCGGCCTCTAATACGACTCACTATAGGG</u> CAGAGTCATCGAGCAGCATC |
| SgRNA_R | GTTTTAGAGCTAGAAATAGCA |
| Sg2_screen_F | CAAAGTCGTGGGGCTCTATC |
| Sg2_screen_R | GGTGGCCTCAAGGTTTTCTT |
| Sg3_screen_F | GAAGTTTGCGAAGCCAGAAC |
| Sg3_screen_r | AGTTTTCACATACCGCCACT |
| Sg4_screen_f | ACGGCCATTCCTCGTTATTT |
| Sg4_screen_r | ACGTGGAGGAAAGTAGGTGT |

The constructs pDeCas Sg2, pDeCas Sg3 and pDeCas Sg4 were transformed into Arabidopsis and T1 plants were selected on ½ MS plates supplemented with 16 mg/L PPT (Duchefa). Next plant generations (from T2 on) were grown in ½ MS plates and screened for mutations by PCR and

sequencing. gDNA was isolated from single plants leaves and a fragment containing the SgRNA target amplified by PCR (Table 5) was used for sequencing (Sequencing platform, IPK, Gatersleben, Germany). The analysis of mutations generated by CRISPR-Cas9 was done with TIDE v2.0.1 (Brinkman et al., 2014), a software which reconstructs the indels produced by genome editing tools from the sequence chromatogram.

4. Results

4.1 General characterization of CAP-D2 and CAP-D3 condensin subunits

4.1.1 Putative regulators of CAP-D2 and CAP-D3 expression

Previous *in silico* analysis with the perturbation tool of GENEVESTIGATOR suggest an influence of the histone monoubiquitination 1 (HUB1), the retinoblastoma-related protein (RBR) and the COP9 signalosome (CSN) subunits CSN3, CSN4 and CSN5, on CAP-D2 and CAP-D3 transcription (Table 6) (Schubert et al., 2013). As the *cap-d2* and *cap-d3* mutants show centromere association during interphase (Schubert et al., 2013), the putative regulators were examined for a similar phenotype (Fig. 7b). Two to five T-DNA insertion mutant lines for CSN3, CSN4, CSN5b, RBR, HUB1 and HUB2 were selected, the T-DNA insertions confirmed by sequencing and analyzed for centromere association during interphase (Fig. 7a). For all mutants the analysis was done on the same material and ploidy level, based on flow-sorted 4C nuclei isolated from expanded rosette leaves. The centromere association was evaluated by counting the number of FISH centromere (pAL) signals per nucleus.

Table 6. In silico gene expression analysis of CAP-D2 and CAP-D3 co-regulated genes using GENEVESTIGATOR. Fold-change expressions of CAP-D2 and CAP-D3 in the perturbations are indicated. Modified from Schubert et al., (2013).

| Protein | Perturbation | CAP-D2 | | CAP-D3 | |
|---|----------------|-------------|---------|-------------|---------|
| | | Fold-change | P value | Fold-change | P Value |
| COP9 signalosome subunit 3 (CSN3) | T-DNA mutation | 2.73 | <0.001 | 5.11 | <0.001 |
| COP9 signalosome subunit 4 (CSN4) | T-DNA mutation | 2.76 | <0.001 | 5.27 | <0.001 |
| COP9 signalosome subunit 5 (CSN5) | T-DNA mutation | 2.72 | <0.001 | 4.43 | <0.001 |
| Retinoblastoma-related protein (RBR) | RNAi depletion | 1.38 | 0,007 | 1.68 | 0.008 |
| Histone Monoubiquitination 1 (HUB1) | EMS mutation | -2.12 | <0.001 | -1.45 | 0.016 |

The CSN is an eight-subunit complex that regulates cullin-RING E3 ubiquitin ligases, which in turn, regulate multiple processes by targeting proteins for proteasome degradation (Hotton & Callis, 2008). Therefore, the CSN complex is involved in plant development, photomorphogenesis, DNA repair and biotic stress (Schwechheimer & Isono, 2010; Stratmann & Gusmaroli, 2012). Loss of function mutants for Arabidopsis CSN subunits show seedling growth arrest and G2 phase delay (Dohmann et al., 2008a). Among the selected T-DNA insertion, *csn3-1* and *csn5b-3* presented the seedling arrest phenotype (Dohmann et al., 2008a,2008b), and therefore, were not included in the study. For the other *csn* lines (*csn3-2*, *csn4-1*, *csn4-2*, *csn5b-2* and *csn5b-1*), the growth was similar to wild-type plants. Two homozygous plants for each line were analyzed, except for *csn5b-1*,

where only one homozygote could be confirmed. Significant differences (P value < 0.05) were observed in the number of centromeric pAL signals after FISH for *csn3-2*, *csn4-1*, *csn4-2* and *csn5b-2* mutants compared to wild-type (Fig. 7c - e).

The retinoblastoma protein (Rb) is a key regulator of the cell cycle which is conserved in metazoans (Miskolczi et al., 2007). Rb binds to the E2F transcription factor and acts as repressor impeding the G1/S progression. In Arabidopsis, the retinoblastoma-related protein (RBR) is involved in cell proliferation, differentiation and responses to the environment (Harashima & Sugimoto, 2016), and homozygous *rbr* mutants are gametophytic lethal (Ebel et al., 2004). Heterozygous plants could be recovered from two T-DNA lines which were not described previously: *rbr-478*, with the T-DNA insertion at the end of the gene; and *rbr-029*, in the 3'UTR. Only the *rbr-029* mutant presented differences compared to wild-type in the distribution of centromeric pAL signals in interphase nuclei (Fig. 7f).

Histone monoubiquitination 1 (HUB1) and its homolog HUB2 are E3 ubiquitin ligases that monoubiquitinate histone H2B (H2Bub1), which functions as a epigenetic mark of active transcription in Arabidopsis (Feng & Shen, 2014). Arabidopsis *hub1* and *hub2* mutants have reduced levels of H2Bub1 (Liu et al., 2007) and the G2/M transition is blocked (Fleury et al., 2007). This block possibly causes vegetative growth defects of the mutant. Mutant plants are smaller than wild-type, they have modified leaf shape and reduced rosette biomass, root growth and seed dormancy (Fleury et al., 2007). We obtained homozygous plants for all the selected T-DNA lines: *hub1-3*, *hub1-4*, *hub1-5*, *hub1-760*, *hub1-867*, *hub2-1*, *hub2-2* and the double mutant *hub1-3/hub2-1*. At least two plants were analyzed per line, but no differences in the number of centromeric signals between the mutants and wild-type were detected (Fig. 7g, h).

In total, 5 out of 15 mutants selected as putative transcriptional regulators of CAP-D2 and CAP-D3 showed a deviation in the number of centromeric pAL signals compared to wild-type nuclei. Nevertheless, none of the mutants reached the level of centromeric association observed in the *cap-d3* and *cap-d2* mutants, where more than 80% of the nuclei showed 1 to 6 signals (Schubert et al., 2013)(Fig. 7b). The *csn* mutants presented less centromeric association than the *cap-d* mutants, and the percentage of nuclei with less than 6 centromeric pAL signals was less than 25% (Fig. 7c - e).

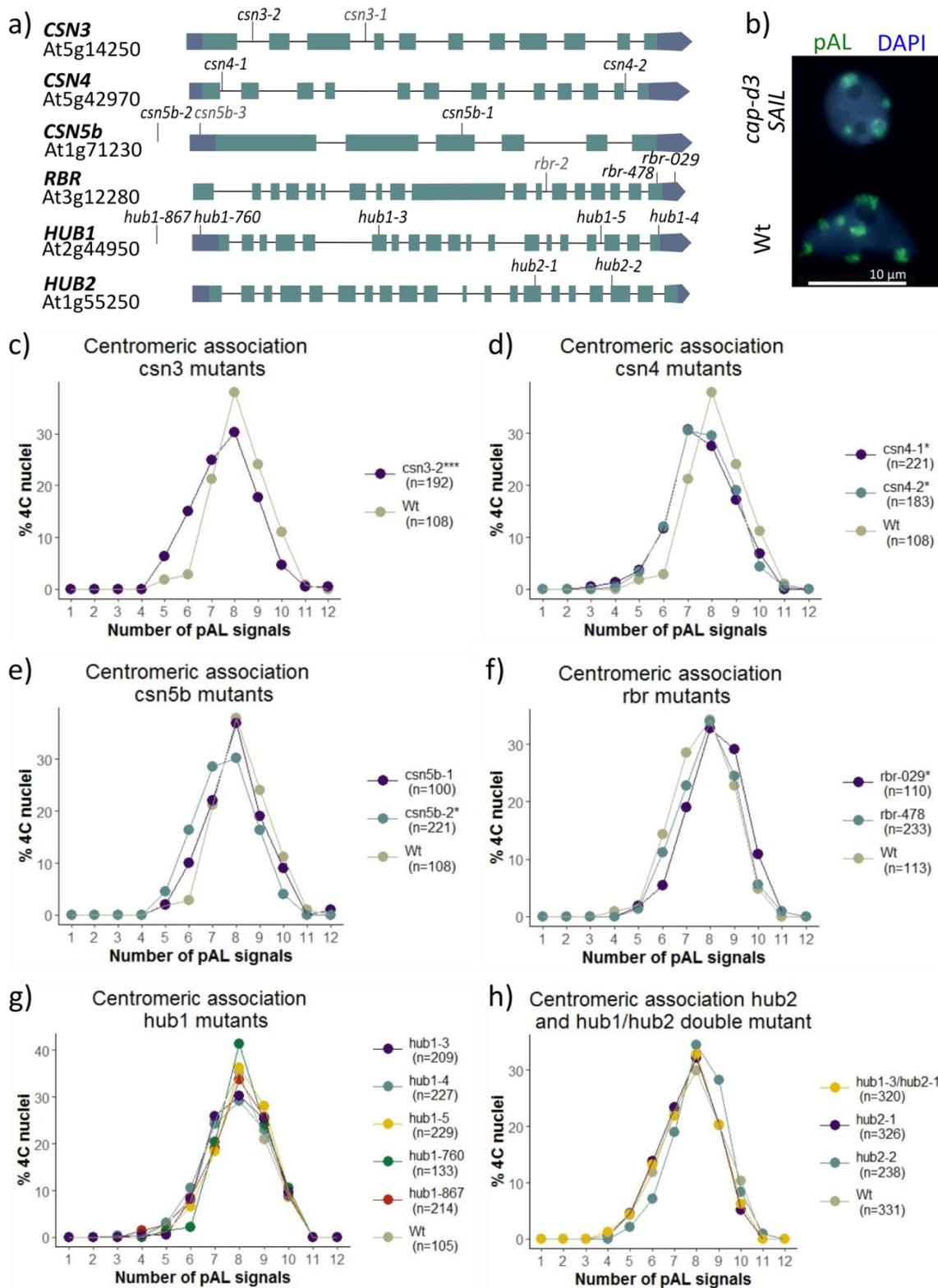


Figure 7. Degree of centromere association during interphase in putative transcriptional regulators of *CAP-D2* and *CAP-D3*. a) Gene structure models of *CSN3*, *CSN4*, *CSN5b*, *RBR*, *HUB1* and *HUB2*. Blue boxes represent the exons, lines the introns, and dark blue boxes the 5' and 3' UTR. The T-DNA insertion lines analyzed for each gene and its insertion site

is represented above the model. Names in light-grey are mutant lines where the insertion could not be confirmed or which were lethal. b) Example of centromeric clustering in the *cap-d3 SAIL* mutant compared to wild-type (Wt). The number of centromeres was determined using the pAL repeat as FISH probe (in green). The diagrams (c-h) represent the centromeric pAL signal frequencies in 4C nuclei of the *csn3*, *csn4*, *csn5b*, *rbr*, *hub1* single and the *hub2* and *hub1-hub2* double mutants. n is the number of nuclei analyzed. Statistical analyses were performed with the Mann-Whitney Rank Sum test or the Kruskal-Wallis one way analysis of variance on Rank and Dunn's method. * and *** represent P values < 0.05 and 0.001, respectively.

4.1.2 CAP-D2 and CAP-D3 expression in Arabidopsis

Based on *in silico* analysis done with the Arabidopsis eFP Browser 2.0, CAP-D2 and CAP-D3 have a similar expression pattern. Both are highly expressed in shoot apex, roots, flower buds and vegetative rosette leaves. Their expression is lower in cotyledons, rosette leaves after bolting, mature flowers, siliques and embryos (Appendix Fig. 1). To corroborate the *in silico* analysis we assessed the transcription of both genes in seedlings, rosette leaves before bolting, roots and flower buds by quantitative real-time RT-PCR. Transcription values were normalized to the geometric mean of the house keeping genes PP2A and At4g26410 and relative to the transcription in seedlings. Both genes, CAP-D2 and CAP-D3, were transcribed in the 4 tested samples. The highest transcription for both genes was observed in flower buds, the lowest in seedlings (Fig. 8a). The transcription of *CAP-D2* is 25.6, 14.8 and 3.5 times higher in flower buds, roots and leaves, respectively, than in seedlings (Fig. 8b). Similarly, the *CAP-D3* transcription is 18.3, 9.4 and 4.4 times higher in flower buds, roots and leaves respectively, than in seedlings (Fig. 8b).

The activity of the *CAP-D2* and *CAP-D3* promoters was evaluated in Arabidopsis transgenic lines expressing different versions of the promoters fused to the β -glucuronidase (GUS) reporter gene (Fig. 8c). Six presumed promoters of different length were analyzed for *CAP-D2* and 2 for *CAP-D3*. The promoter region of *CAP-D2* contains two putative E2F binding sites at -345 bp and -114 bp upstream of the start of the gene (Schubert et al., 2013). In addition, there is a gene at position -391 bp (At3G57062) of 146 bp length and unknown function. Considering the positions of At3G57062 and the E2F sites, two promoter lengths were analyzed, a long promoter that comprises 1156 bp upstream the start of *CAP-D2* (Pro4), and a short promoter of 391 bp (Pro7), which includes only the E2F sites. In addition, we designed the different promoters with or without intron 1 and 2. In Arabidopsis, promoter proximal introns can enhance the expression of a gene by a mechanism known as Intron Mediated Enhancement (IME) (Rose et al., 2008). The putative enhancing ability of *CAP-D2* introns was analyzed *in silico* with the web tool IMETER (Parra et al., 2011). The IMETER score is positively correlated to the enhancing ability of an intron. For

CAP-D2 the two first introns have positive IMETER scores (Fig. 8d), meaning they are likely to enhance expression. These introns were included in the analysis in combination with the long and short promoters of *CAP-D2*: Pro5 (long promoter) and Pro8 (short promoter) include Intron1; and Pro6 (long promoter) and Pro9 (short promoter) include Intron1 and Intron2 (Fig. 8c).

The promoter region of *CAP-D3* contains also two putative E2F binding sites at -397 bp and -84 bp (Schubert et al., 2013) and a short gene (At4G15885) of unknown function at position -474 bp. The IMETER scores of all *CAP-D3* introns were negative and therefore unlikely to enhance expression (Fig. 8d). Therefore, for *CAP-D3* the introns were not considered and only a long promoter at -1318 bp (Pro10) and a short promoter at -474 bp (Pro11) from the start of the gene were analyzed (Fig. 8c).

T1 transgenic plants with the different versions of *CAP-D2* and *CAP-D3* promoters were stained for GUS analysis, except for Pro4 where no positive plants could be isolated. The *CAP-D2* promoter version Pro5 (n=7) was active (blue staining) on stipules (small organs at the base of the leaves), leaf vascular tissue and root tips; Pro6 (n=6) had weak activity on root tips; all Pro7 plants (n=22) showed GUS-staining in leaf vascular tissue and root tips and six plants presented also staining in the apical meristem and 16 in stipules; all Pro8 plants (n=23) presented GUS activity in apical meristem and root tips, and 16 of them also in leaf vascular tissue. Pro9 (n=6) showed activity in roots, and 3 plants also weakly in the apical meristem (Fig. 8e). Therefore, all *CAP-D2* promoter versions were active in roots, but the staining was stronger in the short promoter versions (Pro7, Pro8 and Pro9) than in the long ones (Pro5 and Pro6). Short promoters showed activity on the apical meristem; versions that include the second intron (Pro6 and Pro9) lost the staining on leaf vascular tissue. *CAP-D3* Pro10 (n=5) showed no activity (no staining), and for Pro11 (n=18), 8 plants showed activity in apical meristem and root tips. For both, *CAP-D2* and *CAP-D3*, the expression can be driven more effectively by the short promoter, which contains the E2F sites.

Taken together quantitative real-time RT-PCR and GUS activity staining demonstrated that, *CAP-D2* and *CAP-D3* are highly expressed in meristematic tissues (root tips, flower buds, apical meristem) and leaves. The low transcription observed in seedlings could be due to a low amount of meristematic tissue in the sample since complete one-week old seedlings were used.

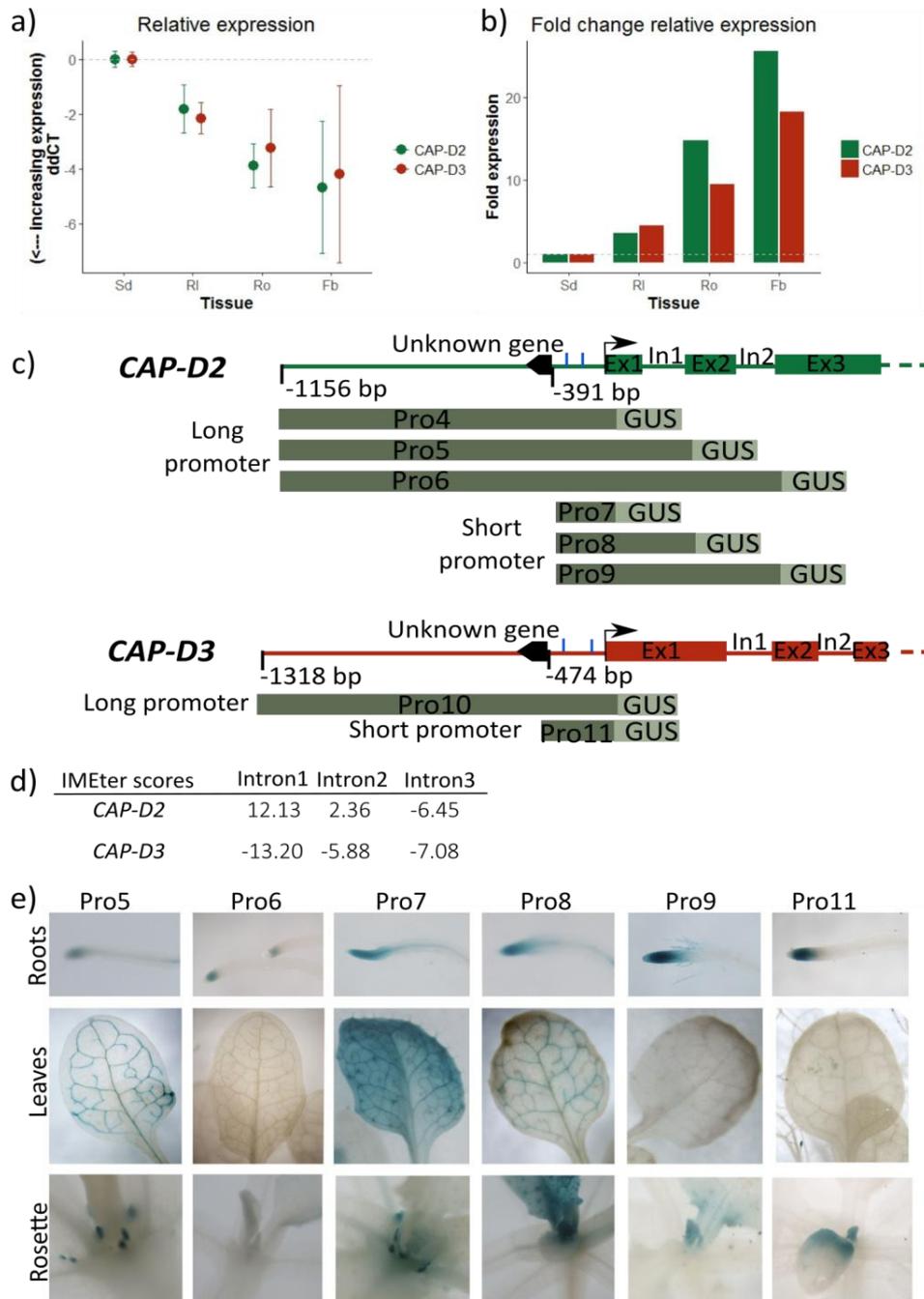


Figure 8. *CAP-D2* and *CAP-D3* expression and analysis of the promoter regions. a) Relative gene transcription. ddCt values, lower values indicate higher transcription. Error bars represent the standard deviation between three biological replicates; b) Fold change transcription. a) and b) values were normalized to the geometric mean of the house keeping genes *PP2A* and *At4g26410* and relative to the transcription in seedlings. Sd: Seedlings; RI: Rosette leaves; Ro: Roots; Fb: Flower buds. c) Representation of the promoters of *CAP-D2* (green) and *CAP-D3* (red). Exons are represented as boxes, introns and upstream sequences as a line, the start of the coding region is marked by an arrow and the black fill-arrows represent the unknown genes. Blue lines represent the position of the E2F binding sites. Below the different versions tested of the promoters fused to the GUS gene are shown. d) IMETER scores for the three first *CAP-D2* and *CAP-D3* introns. High values represent high enhancing ability of an intron on the expression of a gene. e) Histochemical GUS staining in roots, leaves and the rosette centers of plants transformed with the indicated promoter versions (Pro5-9, 11).

4.1.3 Localization of condensin I specific subunits and CAP-D3

In order to localize CAP-D2 and CAP-D3 proteins, Arabidopsis wild-type plants were transformed with constructs containing the coding region of either gene fused at its C-terminus to enhanced yellow fluorescence protein (EYFP) and under the control of the 35S promoter (CAP-D2_EYFPc and CAP-D3_EYFPc). In both cases, the detection of the proteins *in vivo* or by immunolocalization with anti-GFP antibodies (also detecting EYFP) was not possible. Constructs with EYFP at the N-terminus (CAP-D2_EYFPn and CAP-D3_EYFPn) were also tested with the same result. Arabidopsis protoplasts have been used previously to examine the localization of proteins, including the condensin subunits CAP-H and CAP-H2 (Fujimoto et al., 2005). Therefore, the localization of the condensin I specific subunits CAP-D2, CAP-G and CAP-H was also analyzed in Arabidopsis protoplasts. Constructs including the coding region of *CAP-D2*, *CAP-G* and *CAP-H* fused to EYFP (CAP-D2_EYFPc, CAP-G_EYFPc and CAP-H_EYFPc) were transiently expressed in Arabidopsis mesophyll protoplasts. To improve the visualization of the reporter construct, the protoplasts were fixed and the EYFP-fusion proteins immunolocalized with anti-GFP antibodies. In addition, they were counterstained with DAPI to see clearly the nucleus. All these three condensin I complex subunits are present in the cytoplasm and nucleus (Fig. 9a). The cytoplasmic localization of CAP-H was previously described (Fujimoto et al., 2005), but not yet its nuclear localization. In the cytoplasm GFP-negative, but DAPI-positive round organelles were also visible. We identified them as chloroplasts, since the protoplasts were derived from leaf tissue. As controls were used untransformed protoplasts (negative EYFP control) and protoplasts transformed with Control_EYFPc (positive EYFP control, Appendix Table 2). The free EYFP (positive control) localized in the cytoplasm and nucleus (Fig. 9a). Western blot analysis on CAP-D2_EYFPc transformed protoplasts confirmed that the CAP-D2_EYFP protein was intact and that the visible localization corresponds to the fusion proteins (187 kDa, red arrow), and not to free EYFP (27 kDa, green arrow) (Fig. 9b). Hence, the CAP-D2_EYFPc construct is functional in transiently transformed protoplasts, although the protein is not visible in stable Arabidopsis transformants.

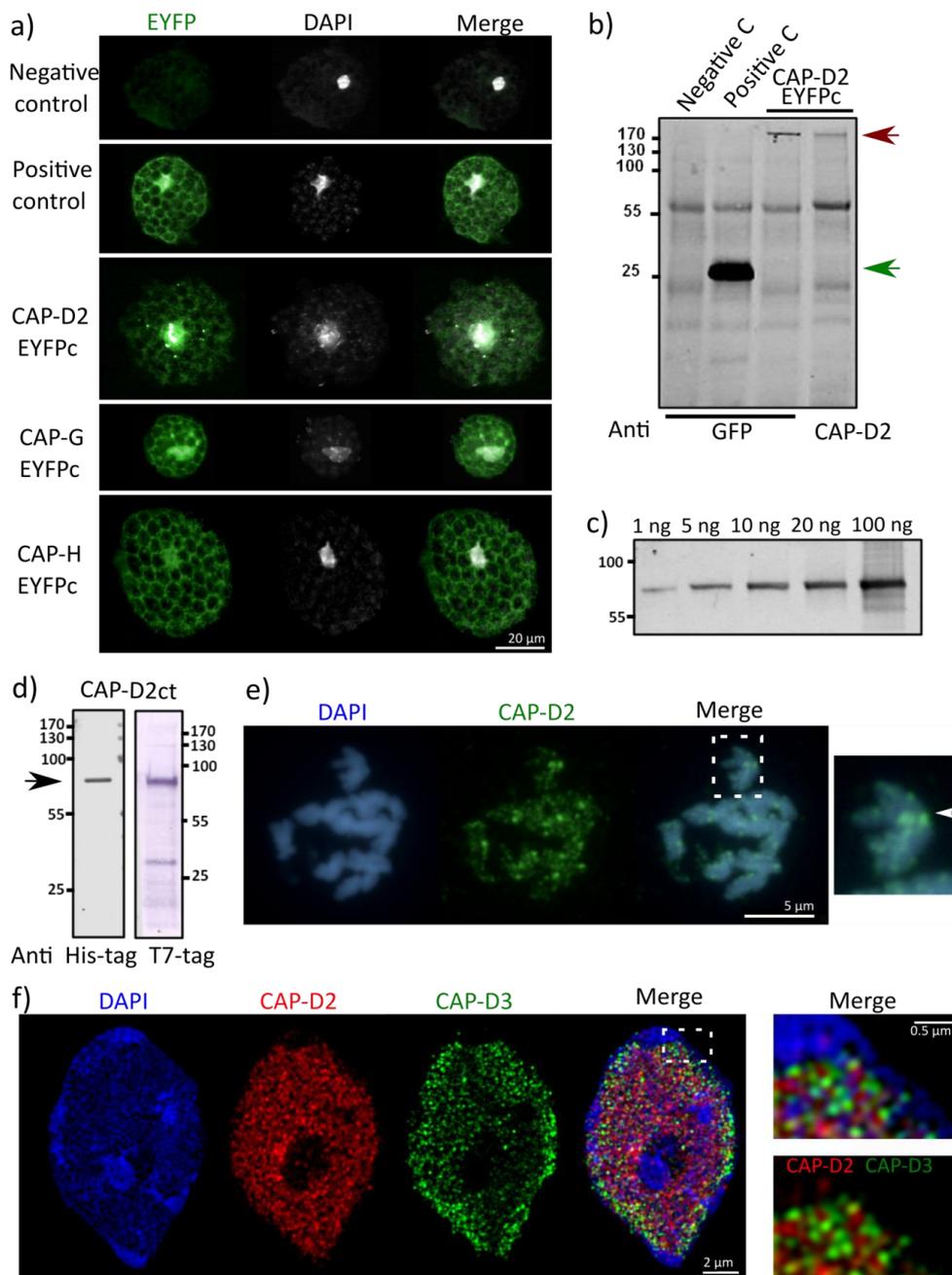


Figure 9. Localization of the condensin subunits CAP-D2, CAP-D3, CAP-G and CAP-H in Arabidopsis. a) Untransformed protoplasts (Negative control) and transformed with the condensin I specific subunits CAP-D2, CAP-G and CAP-H fused C-terminally to EYFP, and free EYFP (Positive control). b) Western blot analysis on total protein extracts from protoplasts untransformed (Negative C), transformed with free EYFP (Positive C) or with CAP-D2-EYFPc (CAP-D2 EYFPc). Detection was made with anti-GFP antibodies or anti-CAP-D2 serum. The intense band of 27 kDa (green arrow) corresponds to free EYFP. The bands of 187 kDa (red arrow) correspond to CAP-D2_EYFP fusion protein. c) Sensitivity test for anti-CAP-D2 serum. Western blot on different amounts (1 – 100 ng) of CAP-D2_Ct recombinant protein (56.3 kDa) against anti-CAP-D2 serum. d) Western blot on CAP-D2_Ct recombinant proteins against anti-His-tag and anti-T7-tag. Black arrow marks the band containing the CAP-D2_Ct recombinant protein e) CAP-D2 immunolocalization on mitotic metaphase chromosomes of Arabidopsis. Dashed white square marks the enlarged region on the right. The white arrow indicates the centromeric region. f) SIM image of CAP-D2 and CAP-D3 immunocolocalization on an 8C interphase nucleus. White dashed square marks the enlarged region on the right. The enlarged region is depicted with (merge) and without the DAPI staining (CAP-D2 CAP-D3).

To address the subcellular localization of CAP-D2, rabbit polyclonal antibodies were generated against recombinant proteins containing the last 500 amino acids of CAP-D2 (CAP-D2 Ct). The recombinant protein produced in *E. coli* had a weight of ~80 kDa although the expected weight was 59.92 kDa including the T7- and His-tags on the N-t and C-termini, respectively. Western blot analysis against both tags confirmed the integrity of the recombinant protein (Fig. 9d). The difference in molecular weight between the expected and actual size could be due to the high proportion of acidic amino acids in the recombinant protein, 18.5% amino acids are negatively charged while only 11.1% are positively charged (Guan et al., 2015). The sensitivity of the serum was tested with a range of different quantities of CAP-D2ct recombinant proteins. The outcome was that the CAP-D2 antiserum can detect amounts as low as 1 ng of recombinant protein (Fig. 9c). The specificity of the serum was tested on a total protein extract from CAP-D2_EYFPc transformed protoplasts (Fig. 9b) and from Arabidopsis wild-type leaves (data not shown). The CAP-D2 antiserum is specific, it detects the CAP-D2 fusion protein from protoplasts but not the CAP-D2 protein from wild-type leaves. This is possibly due to a low amount of protein (below Western blot detection limit) on leaves compared with the protoplasts. In protoplast there is an overexpression of the construct since it is under the control of the 35S promoter.

Immunolocalization of CAP-D2 on mitotic chromosomes shows a distribution along the sister chromatids and an enrichment in the centromeric region (Fig. 9e). This localization is similar to what has been previously described in *Drosophila* (Savvidou et al., 2005), human (Ono et al., 2003, 2004) and budding yeast (D'Ambrosio et al., 2008a). Double immunolabelling of CAP-D2 and CAP-D3 in interphase nucleus shows a similar localization of both proteins in euchromatin and absence from the nucleolus and chromocenters (heterochromatin) (Fig. 9f). As visible in the further enlarged area (Fig. 9f), CAP-D2 and CAP-D3 co-localize only partially. The localization of condensin II in the nucleus and condensin I in the cytoplasm during interphase has been widely described in vertebrates (Gerlich et al., 2006; Hirota et al., 2004; Ono et al., 2004). On the other hand, the localization of CAP-D2 (condensin I) in the nucleus during interphase has been previously observed in chicken and *Drosophila* (Savvidou et al., 2005; Zhang et al., 2016). These observations imply a role of CAP-D2 in the organization of chromatin during interphase.

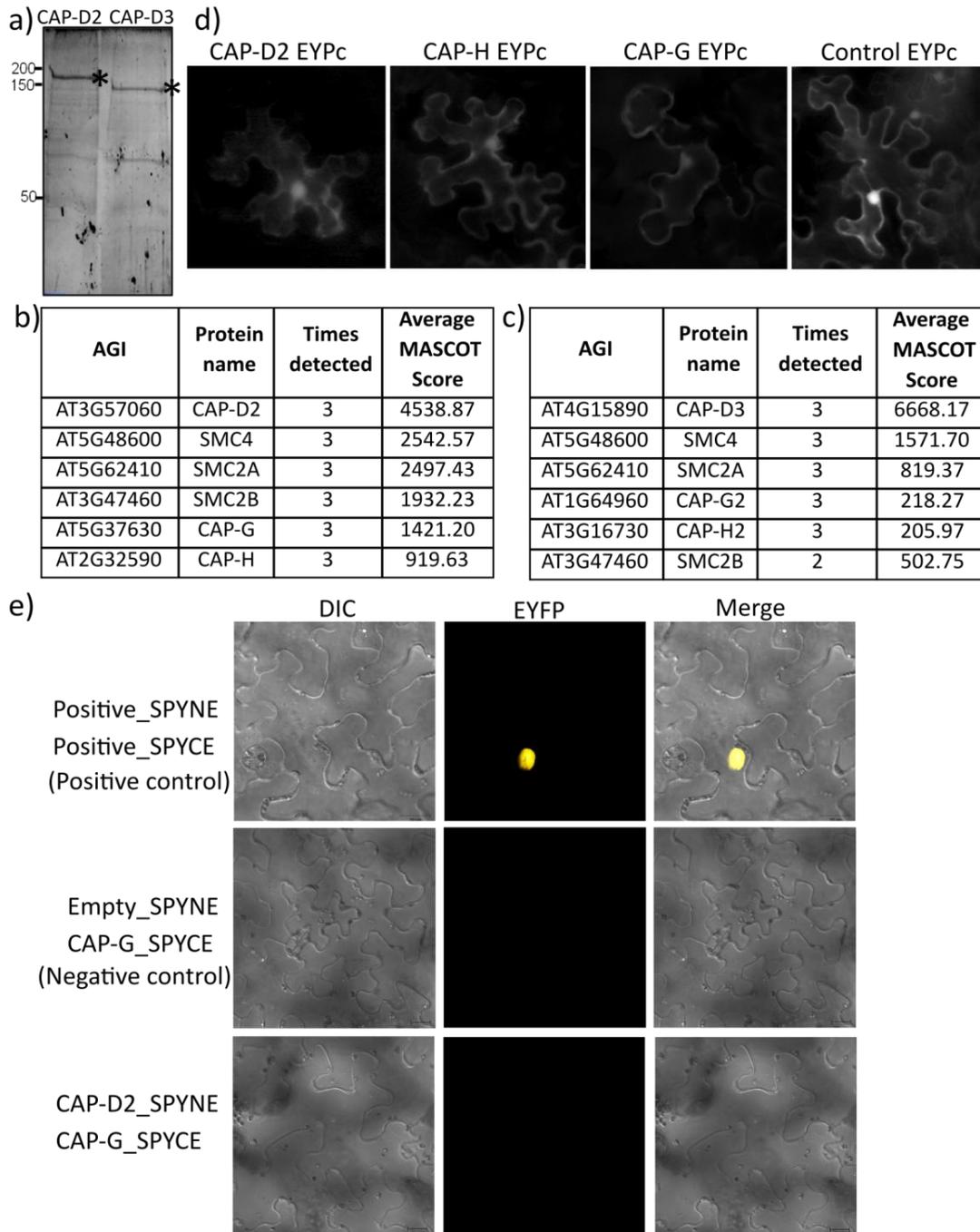


Figure 10. CAP-D2 and CAP-D3 interact with the other putative condensin subunits in Arabidopsis. a) Affinity purification of CAP-D2 and CAP-D3 GS fusion proteins. Coomassie staining of co-purified proteins separated in a polyacrylamide gel. The asterisks indicate CAP-D2-GS (176 kDa) and CAP-D3-GS (163 kDa) proteins respectively. b - c) Identified condensin subunits co-purifying with CAP-D2 (b) and CAP-D3 (c) GS fusion proteins analyzed by mass spectrometry. The number of times detected in three independent affinity purification and the average MASCOT scores are specified. d) Localization of CAP-D2, CAP-H and CAP-G EYFP fusion proteins and EYFP alone (control EYFPc) in *N. benthamiana* leaf epidermal cells. e) Differential Interference Contrast (DIC) and confocal images of *N. benthamiana* leaf epidermal cells infiltrated with the indicated construct pairs to perform BiFC. SPYNE constructs contain the EYFP N-terminal part, while SPYCE constructs contain the C-terminal part of the EYFP. Only the positive control shows BiFC.

4.1.4 CAP-D2 and CAP-D3 interacting proteins

A previous study confirmed the presence of both condensin complex components in *Arabidopsis* (Smith et al., 2014), but whether the complexes are formed by the same subunits as in non-plant species is unknown. To determine the specific composition of each complex, CAP-D2 and CAP-D3 were fused to a GS-tag, and affinity-purified from *Arabidopsis* PSB-D suspension cultured cells in collaboration with the research group of Prof. Klaus Grasser (University Regensburg) (Fig. 10a). The proteins interacting with CAP-D2-GS and CAP-D3-GS were identified by mass spectrometry (Appendix Table 3-4). All putative subunits of the condensin I complex, SMC2A and B, SMC4, CAP-H and CAP-G, co-purified with CAP-D2-GS in the three affinity purifications performed. They all presented high MASCOT scores (Fig. 10b). Similarly, all the putative subunits of the condensin II complex, SMC2A, SMC4, CAP-H2 and CAP-G2, were detected in the three affinity purifications performed for CAP-D3-GS (Fig. 10c). Both complexes are specific, meaning that only condensin I specific subunits were found among the CAP-D2 co-purified proteins and only condensin II specific subunits for CAP-D3. These results confirm that the composition of the condensin complexes in *Arabidopsis* is homologous to other organisms like mammals, chicken and *C. elegans* (Hirano, 2012; Onn et al., 2007). Additionally, among the proteins which co-purified with CAP-D2, there were other interesting proteins such as SMC3, which is part of the cohesin complex; the chromatin remodeling factors CHR17 and CHR19; CUL1, a subunit of the SCF ubiquitin ligase complex; HDC1, a histone deacetylase and ELO3, a histone acetyltransferase from the elongator complex. Among the proteins which co-purified with CAP-D3 were two nucleosome assembly proteins (NAP); CSN1, a subunit of the COP9 signalosome; the helicase BRAHMA; ELO3, from the elongator complex; and NERD, involved in DNA methylation. These proteins co-purifying with CAP-D2 and CAP-D3 need confirmation of their interactions by other methods, but are interesting because of their roles in chromatin remodeling, transcription and protein degradation. All of them have roles related to condensin as a chromatin organizer.

Direct interaction between CAP-D2, CAP-G and CAP-H was tested by Bimolecular Fluorescence Complementation (BiFC) in *N. benthamiana*. First, the localization of CAP-D2, CAP-G and CAP-H in *N. benthamiana* leaf epidermal cells was analyzed using the same constructs previously used in *Arabidopsis* protoplasts (Fig. 9a). The localization of the three proteins is similar to that observed in protoplasts, being present in the nucleus and the cytoplasm (Fig. 10d). Each protein was fused to the N (SPYNE constructs) or C (SPYCE constructs) terminal part of EYFP and transformed in *N. benthamiana*. All possible pairs that could restore the EYFP fluorescence were

tested: CAP-D2-SPYNE/CAP-G-SPYCE, CAP-G-SPYNE/CAP-H-SPYCE, CAP-D2-SPYNE/CAP-H-SPYCE, CAP-G-SPYNE/CAP-D2-SPYCE, CAP-H-SPYNE/CAP-D2-SPYCE, CAP-H-SPYNE/CAP-G-SPYCE and the positive control Positive-SPYNE/Positive-SPYCE. Unfortunately, none of the combinations could restore the EYFP fluorescence as seen for the positive control (Fig. 10e, only one possible combination shown as example).

4.2 CAP-D3 characterization

4.2.1 CAP-D3 organizes chromatin during interphase

CAP-D3 involvement in the organization of the chromatin during interphase was previously described in *Arabidopsis*. *cap-d3* mutants present centromeric clustering and chromosome territory dispersion (Schubert et al., 2013). In *Drosophila*, CAP-D3 is also involved in the formation of compact chromosome territories (Hartl et al., 2008b). To further study the involvement of CAP-D3 in chromatin organization during interphase, we used two *cap-d3* mutants described previously (Schubert et al., 2013). Both mutants are T-DNA insertion lines with the insertion in the first exon (*cap-d3 SAIL*) and fourth exon (*cap-d3 SALK*) (Fig. 11a). In both cases there is a partial transcription of the gene that could lead to a truncated but partially functional protein (Schubert et al., 2013). Homozygous *cap-d3 SAIL* and *cap-d3 SALK* mutant plants are viable. The former shows a smaller habit than wild-type, but the latter is similar to wild-type plants (Fig. 11b). Both show a reduced seed setting (Schubert et al., 2013).

To confirm the centromeric clustering and chromosome territory dispersion phenotypes in both mutants, a FISH experiment was performed with probes specific for the centromeric repeat pAL and part of the Chromosome Territory 1 Bottom arm (CT1Bp) on flow sorted 4C nuclei (Fig. 11c, d). In addition to the number of centromeric pAL signals per nucleus, the area of CT1Bp signal and the nucleus were measured. The median area of the CT1Bp signal was 3.9, 4.7 and 4.7 μm^2 for *cap-d3 SAIL*, *cap-d3 SALK* and wild-type, respectively (Fig. 11e). No significant differences were found, thus we could not confirm the chromosome territory dispersion phenotype of *cap-d3* mutants described in Schubert et al. (2013). In addition, no significant difference was found in the nuclear area between the *cap-d3* mutants and wild-type plants (Fig. 11e).

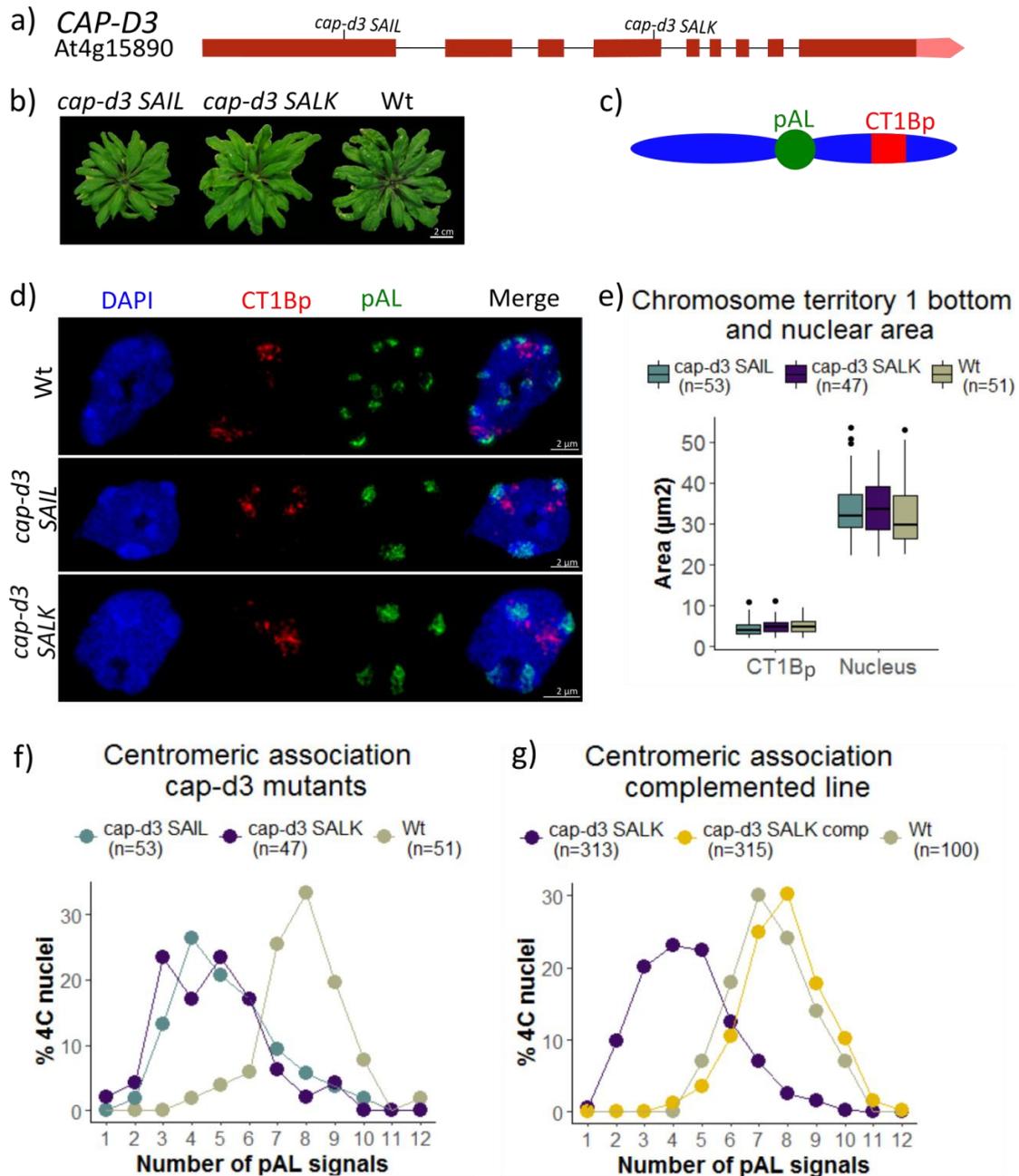


Figure 11. Confirmation of the Arabidopsis *cap-d3* mutant phenotypes. a) Gene structure model of *CAP-D3*: Red boxes represent exons, lines the introns and the lighter red box the 3'UTR. The T-DNA insertion sites of *cap-d3 SAIL* and *cap-d3 SALK* lines are also indicated. b) Rosette pictures of homozygous *cap-d3 SAIL*, *cap-d3 SALK* and wild-type (Wt) plants. c) Schematic representation of chromosome 1 and the localization of the Chromosome Territory 1 Bottom part (CT1Bp) and centromeric (pAL) probes (pAL marks all 10 centromeres present in the nuclei). d) SIM image of a FISH with CT1Bp and pAL probes on 4C nuclei. e) Box plot diagram of CT1Bp signal and the nucleus area sizes of *cap-d3 SAIL*, *cap-d3 SALK* and Wt nuclei. The boxes indicate upper and lower quartiles and the black bar the median. f) and g) pAL signals frequencies in 4C nuclei of *cap-d3 SAIL*, *cap-d3 SALK*, *cap-d3 SALK complemented* and Wt. n is the total number of nuclei analyzed from two different plants in e) and f), and from three different plants in g).

On the other hand, the centromeric association phenotype could be confirmed. In both *cap-d3* mutants the nuclei showed a lower number of centromeric pAL signals than wild-type (Fig. 11d). Around 80% of *cap-d3* mutants nuclei showed less than six centromeric pAL signals, while in wild-type the nuclei with less than six pAL signals were 12% (Fig. 11f). To verify that the mutation in the *CAP-D3* gene is indeed responsible for the centromeric clustering, complementation experiments were carried out. Complementation consists in re-introducing a working copy of the gene in a mutant plant, and checking if the phenotype reverts to the wild-type level. Thus, *cap-d3* *SALK* mutant plants were transformed with *CAP-D3_EYFPc* constructs, containing the coding region of *CAP-D3* fused to EYFP under the control of the 35S promoter.

The centromeric association phenotype was evaluated in *cap-d3* *SALK* complemented plants by FISH with the centromeric repeat pAL and compared to *cap-d3* *SALK* mutants and wild-type. Only 15% of the complemented nuclei had less than 6 centromeric signals, which is similar to wild-type (Fig. 11g). Thus, the association phenotype in *cap-d3* *SALK* complemented plants, reverted to the wild-type level confirming that *CAP-D3* is responsible for the centromere association present in the mutants.

Beside centromeres, in Arabidopsis, the 45S and 5S rDNA are heterochromatin-associated sequences. In differentiated nuclei, 45S rDNA tends to associate but the 5S rDNA loci are often separated (Berr & Schubert, 2007). To examine whether *CAP-D3* affects the organization of the heterochromatin, the distribution of the 45S and 5S rDNA loci was analyzed by FISH in the *cap-d3* mutants (Fig. 12a). The 45S rDNA forms the nucleolar organizing region (NOR) which locates close to the nucleolus during interphase (Berr et al., 2006; Fransz et al., 2002). Arabidopsis chromosomes 2 and 4 harbor the 45S rDNA. Therefore, in a 4C nucleus a maximum of eight 45S rDNA signals could be evident. In wild-type, 37% of the nuclei showed three 45S rDNA signals, while in both mutants over 40% of the nuclei showed two signals (Fig. 12b). The 5S rDNA loci are on chromosomes 3, 4 and 5. Thus, a 4C nucleus should have a maximum of 12 signals. Over 70% of the nuclei showed between six and ten 5S rDNA signals in *cap-d3* mutants and wild-type plants (Fig. 12c). Therefore, *cap-d3* mutants present a higher association of the 45S rDNA than wild-type, but the number of 5S rDNA signals remains unaffected.

Arabidopsis centromeres are positioned at the nuclear periphery (Fang & Spector, 2005; Fransz et al., 2002). The centromere position in the *cap-d3* mutants was assessed by embedding nuclei in acrylamide to preserve their 3D structure followed by FISH with the centromeric repeat pAL

(3D-FISH). For each genotype, *cap-d3 SAIL*, *cap-d3 SALK* and wild-type, 10 nuclei were analyzed. Optical sections (Z-stacks) were taken for each nucleus followed by image restoration (Fig. 12d). The position of the centromeres was analyzed in each section. In all the cases, the centromeric signals were at the periphery of the nucleus and the centromeric clustering was still visible in *cap-d3* mutants. In consequence, we did not observe a different centromere positioning between wild-type and the *cap-d3* mutants.

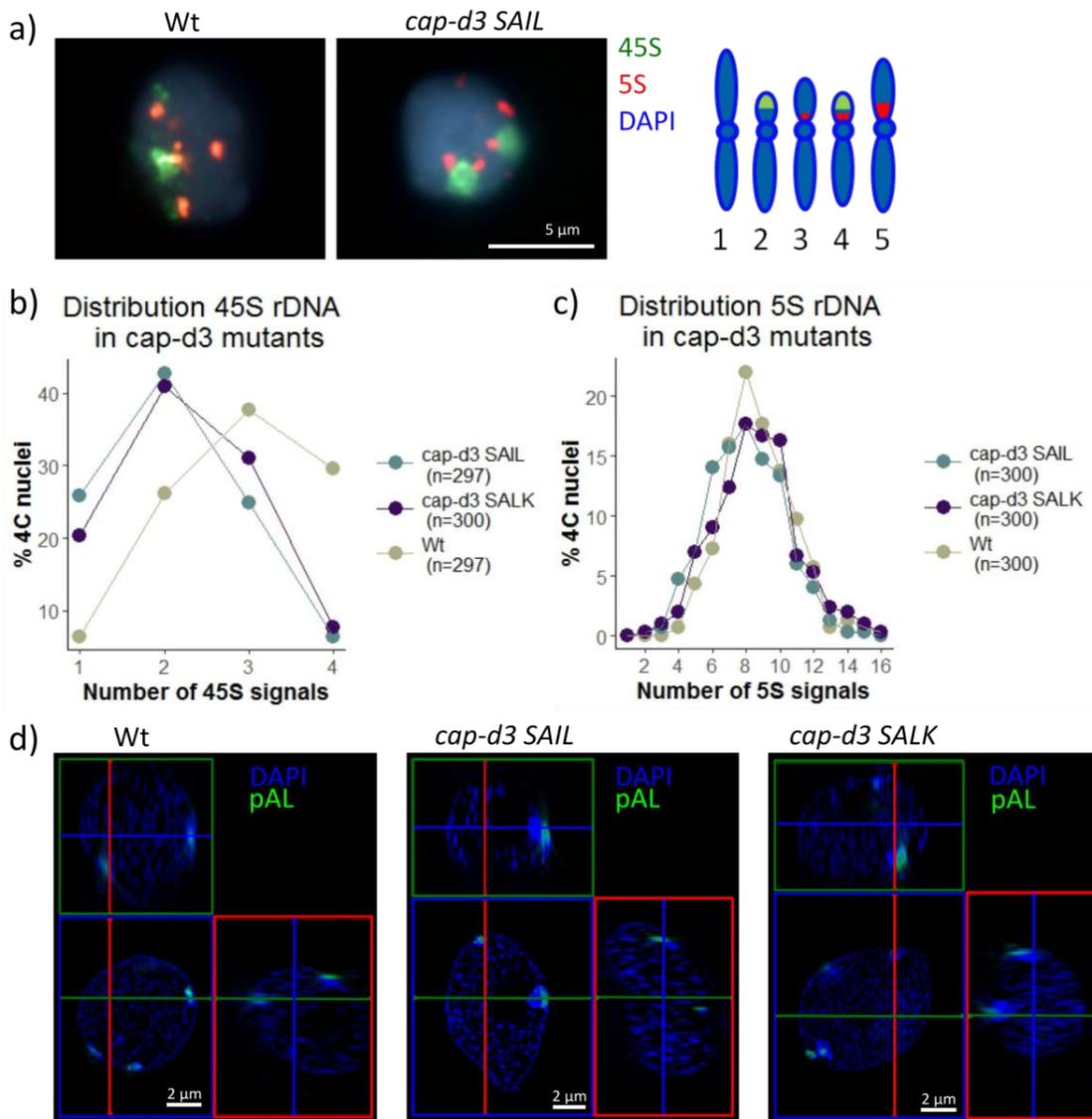


Figure 12. Influence of CAP-D3 on the association of 45S and 5S rDNA and the spatial centromere arrangement in Arabidopsis interphase nuclei. a) FISH with 45S rDNA and 5S rDNA on 4C nuclei of wild-type (Wt) and *cap-d3 SAIL* mutants. The ideogram (right) represents the Arabidopsis chromosomes showing the localization of the 45S and 5S rDNA. b) and c) Frequency of 45S and 5S rDNA signals in 4C nuclei of Wt and *cap-d3* mutants. n is the total number of nuclei analyzed from three different plants. d) SIM orthogonal view of FISH with the centromeric repeat (pAL) on structurally preserved acrylamide-embedded nuclei of Wt and *cap-d3* mutants. Blue, green and red rectangles show x-y, x-z and y-z optical cross-sections, respectively.

4.2.2 Epigenetic landscape in the *cap-d3* mutants

DNA can be methylated at cytosine as 5-methyl-cytosine (5mC). The methylation of DNA is associated with heterochromatin formation and therefore, it is found at the chromocenters (Fransz et al., 2002). Mouse embryonic stem cells depleted of condensin show a reduction of 5mC (Fazio & Panning, 2010). In order to decipher whether such an effect can also be observed in plants, immunolocalization against 5mC was performed to evaluate the distribution of methylated DNA in *cap-d3* mutants compared to wild-type. In both *cap-d3* mutants and wild-type the distribution was similar regarding the 5mC signal localizing in the chromocenters (Fig. 13a). The Arabidopsis centromeric region is highly DNA methylated in the CpG context (Martinez-Zapater et al., 1986). The use of methylation sensitive enzymes and Southern blot hybridization allows a more precise determination of the relative DNA methylation level of the centromeric repeats. *HpaII* and its isoschizomer *MspI* cleave the same CCGG sequence, but *HpaII* is methylation sensitive while *MspI* is not. *MspI* will cut when the middle cytosine is methylated (CCmGG). In wild-type, the centromeric repeats are highly methylated and are thus digested by *MspI* (Fig. 13b). The ladder-like pattern corresponds to the monomer, dimer, trimer and higher orders of centromeric repeats. As expected, *HpaII* does not cut in wild-type DNA. In both *cap-d3* mutants, the hybridization pattern is similar to wild-type. Thus, the relative level of CCGG methylation did not change in *cap-d3* mutants (Fig. 13b).

CAP-D3 prevents the clustering of heterochromatin, but the CAP-D3 protein itself localizes in euchromatic region during interphase. Both types of chromatin are characterized by specific histone modifications marks (Fuchs et al., 2006). To evaluate a possible functional association between histone modifications and CAP-D3, the distribution pattern of different histone marks was compared between *cap-d3* mutants and wild-type. Specific marks for heterochromatin (histone H3K9me1, H3K9me2) and euchromatin (histone H3K4me3, H3K27me3) were tested by indirect immunostaining. In addition, the H3 acetylation marks H3K9ac, H3K14ac, and H3K18ac as well as H3K9+14+18+23+27ac were evaluated. Histone acetylation relaxes chromatin allowing different protein complexes to access the DNA (Wang et al., 2014). Thus, acetylation is associated with transcription, and hypoacetylation with transcriptional repression. In flow-sorted 4C wild-type nuclei, H3K4me3 localizes in euchromatin and it is absent from chromocenters and the nucleolus; in *cap-d3* mutants the localization is identical (Fig. 13c). H3K9me1 is a heterochromatin-specific histone modification that localizes in the chromocenters in both *cap-d3* mutants and wild-type (Fig. 13c). Finally, the acetylation mark H3K14ac localizes mainly in the euchromatin

(transcriptionally active chromatin) of wild-type nuclei, but also in the mutants (Fig. 13c). The other histone modifications tested (H3K27me3, H3K9me2, H3K9ac, H3K18ac and H3K9+14+18+23+27ac) followed also the same pattern in wild-type and the *cap-d3* mutants (Appendix Fig. 2). In short, we could not see obvious differences in the distribution patterns of the different histone marks between wild-type and the *cap-d3* mutants.

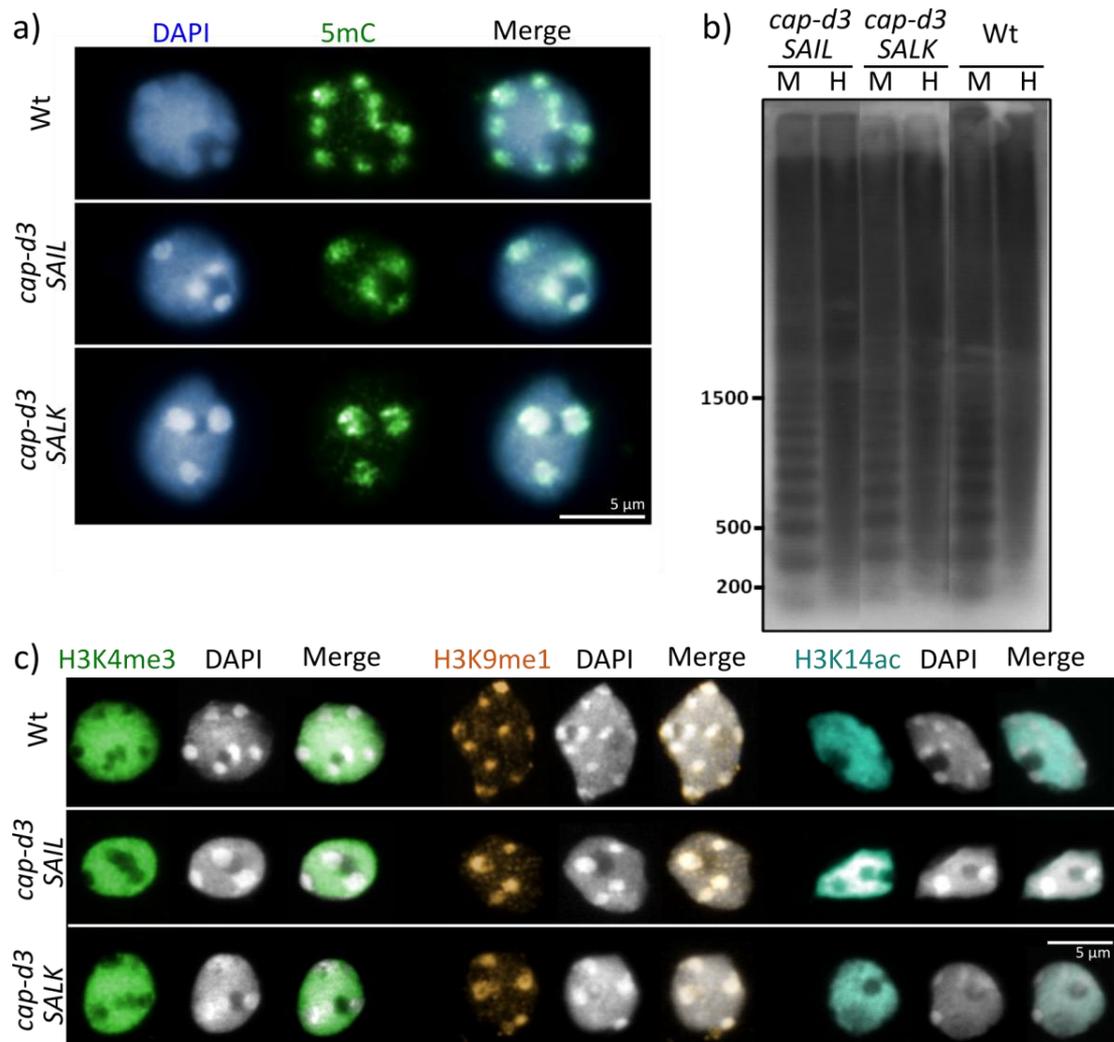


Figure 13. Epigenetic landscape of *cap-d3* mutants in interphase nuclei of *Arabidopsis*. a) 5-methyl-cytosine immunolocalization in 4C nuclei of wild-type (Wt) and *cap-d3* mutants. b) Southern blot of *cap-d3* mutants and Wt genomic DNA digested with *Hpa*II (H) or *Msp*I (M) and hybridized with the p^{32} -labelled centromeric repeat pAL. c) Immunolocalization of histone H3K4me3, H3K9me1 and H3K14ac on 4C nuclei of Wt and the *cap-d3* mutants.

4.2.3 The effect of CAP-D3 on transcription

To assess if the clustering of the pericentromeric interphase chromatin in *cap-d3* mutants affects gene transcription, the transcriptome of both mutants was compared to wild-type. RNA-sequencing was performed in 4 samples (pooled 4 weeks-old plantlets) for each genotype. Principal component analysis (PCA) of the samples allowed a clear separation between wild-type and *cap-d3* mutants (Fig. 14a). *cap-d3 SALK* mutant plants present an intermediate phenotype between wild-type and *cap-d3 SAIL* plants. They have the same size as wild-type plants but as *cap-d3 SAIL*, they present reduced fertility and centromere clustering (Schubert et al., 2013). In agreement, *cap-d3 SALK* samples cluster together in the PCA between wild-type and *cap-d3 SAIL* (Fig. 14a). To test whether a major transcriptional change between the genotypes occurred, the distribution (frequency) of the expressed genes for each genotype was over-laid (Fig. 14b). The transcriptional profiles were almost identical between genotypes, i.e., a high number of genes have very low transcription (few counts) while few genes are highly transcribed. We could therefore exclude an important change in the transcriptional program between *cap-d3* mutants and wild-type plantlets.

Nonetheless, after differential expression analysis, we could observe alterations between the *cap-d3* mutants and wild-type transcriptomes. The genes with at least 2-fold change transcription and a $p\text{Adj} \leq 0.05$ were considered as differentially expressed genes (DEG) between two genotypes. The smallest difference was observed between *cap-d3 SAIL* vs. *cap-d3 SALK* (74 DEG), and the highest between *cap-d3 SAIL* vs. Wt (398 DEG). *cap-d3 SALK* vs. Wt was intermediate (97 DEG)(Fig. 14c). Both *cap-d3* mutants show centromere and 45S rDNA clustering, but *cap-d3 SAIL* plants show growth defects that are absent in *cap-d3 SALK* plants. To separate the individual effect of each allele, in further analysis only the DEG shared by both mutants when compared to wild-type were considered. These 83 genes, common to the *cap-d3* mutation independently of the specific alleles, are subsequently referred to as “*cap-d3* DEG” (Fig. 14d and Appendix Table 5). Of the *cap-d3* DEG, 57 genes are up-regulated and 26 down-regulated compared to wild-type (Table 7). These genes are distributed throughout all chromosomes arms (Fig. 14e). According to their Gene Ontology (GO) enrichment, the *cap-d3* DEG are mainly involved in transcription, particularly in biological processes affecting the response to water, stimuli and stress (Table 8). In agreement with their role in transcription, 13 out of the 83 *cap-d3* DEG are transcription factors (Appendix Table 5).

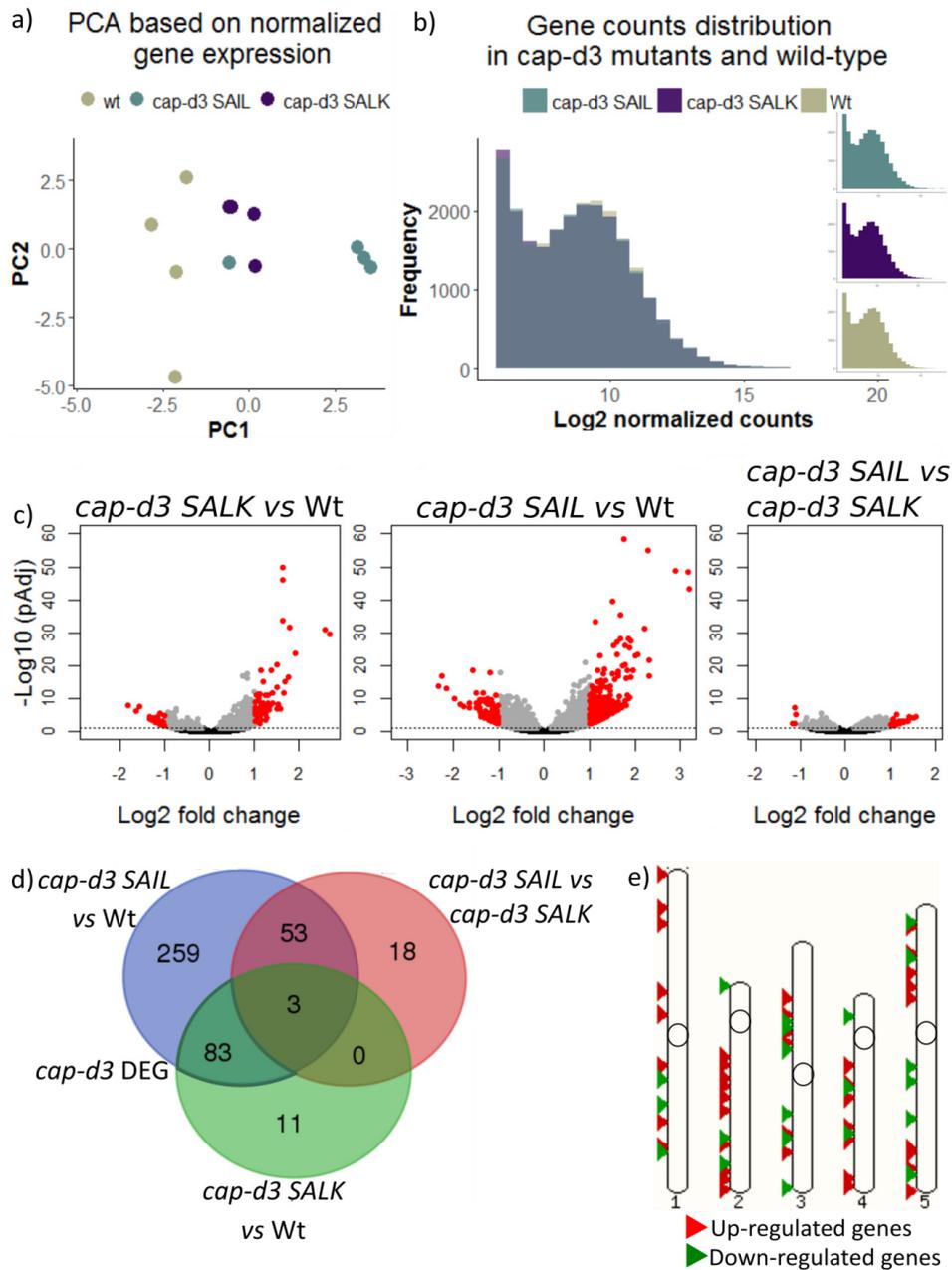


Figure 14. Transcriptome analysis of *cap-d3* mutants and wild-type plantlets. a) Principal component analysis of both *cap-d3* mutants and wild-type (Wt) based on normalized transcription of all genes. b) Gene count distribution in *cap-d3* mutants and Wt. The individual transcription profiles for each genotype are over-laid. For sake of representation, genes with less than 4 counts were excluded, their high frequency would make the other peaks not visible. On the right each histogram is depicted individually. c) Volcano plots showing transcriptome comparisons between *cap-d3 SALK*, *cap-d3 SAIL* and Wt. The horizontal dotted line corresponds to $p_{Adj} = 0.05$. Genes below are depicted in black and above in grey. In red are genes differentially expressed (DEG) at a threshold of 2 fold change (i.e., up-regulated: ≥ 1 Log2 FC, or down-regulated: ≤ -1 Log2 FC) and with a $p_{Adj} \leq 0.05$. p_{Adj} is the p-value corrected for multiple testing with the Benjamini-Hochberg adjustment. d) Venn diagram showing the DEG across the three comparisons. Each circle comprises all the DEG genes of one comparison and the intersections between circles are the common DEG. For example: the blue circle represents the *cap-d3 SAIL* vs Wt DEG, which are 398, of those: 83 are the same as in *cap-d3 SALK* vs Wt, 53 are the same as in *cap-d3 SAIL* vs *cap-d3 SALK*, 3 are differentially expressed in all comparison and 259 are only present in *cap-d3 SAIL* vs Wt. e) Ideogram of Arabidopsis chromosomes showing the position of the 83 *cap-d3* DEG along the chromosomes.

The influence of CAP-D3 directly in transcription is moderate, there are no major transcriptome changes and the number of DEG between both mutants and wild-type is relatively low. However, the DEG involvement in plant response to stress and the proportion of transcription factors indicate that CAP-D3 could have an indirect role in transcription.

Table 7. Number of differentially expressed genes (DEG) in *cap-d3 SAIL* vs Wt, *cap-d3 SALK* vs Wt and the intersection (common genes) between both sets (*cap-d3*). All DEG are at least 2 fold change transcribed and $p_{Adj} \leq 0.05$.

| | <i>cap-3 SAIL</i> vs Wt | <i>cap-d3 SALK</i> vs Wt | <i>cap-d3</i> |
|----------------------|-------------------------|--------------------------|---------------|
| DEG | 398 | 97 | 83 |
| Up-regulated genes | 271 | 67 | 57 |
| Down-regulated genes | 127 | 30 | 26 |

Table 8. Gene ontology (GO) categories enriched in the 83 *cap-d3* DEG. FDR (False Discovery Rate): p-value adjusted for multiple testing.

| Ontology | GO term | Description | FDR |
|--------------------|------------|----------------------------------|--------|
| Biological process | GO:0009414 | Response to water deprivation | 0.0001 |
| Biological process | GO:0009415 | Response to water | 0.0001 |
| Biological process | GO:0042221 | Response to chemical stimulus | 0.0021 |
| Biological process | GO:0009628 | Response to abiotic stimulus | 0.0033 |
| Biological process | GO:0050896 | Response to stimulus | 0.01 |
| Biological process | GO:0006950 | Response to stress | 0.041 |
| Molecular function | GO:0030528 | Transcription regulator activity | 0.024 |
| Molecular function | GO:0003700 | Transcription factor activity | 0.047 |

4.3 CAP-D2 characterization

4.3.1 *cap-d2* T-DNA insertion mutant analysis

CAP-D2 has been previously described to be involved in Arabidopsis chromosome territory organization (Schubert et al., 2013) and meiotic chromosome structure (Smith et al., 2014). In the first study a T-DNA insertion mutant was used, the line *cap-d2-1* (Fig. 15a), while in the study of Smith et al. (2004) a CAP-D2 RNAi line was employed. Only heterozygous individuals of the T-DNA insertion line *cap-d2-1* are viable. The growth of *cap-d2-1* plants is similar to wild-type, but they present reduced fertility (Schubert et al., 2013). To evaluate the effect of this allele during meiosis, the structure of meiotic chromosomes was analyzed in pollen mother cell spreads. In addition, to identify the chromosomes FISH with labeled 45S and 5S rDNA was performed. As described by (Fransz et al., 1998) wild-type metaphase I bivalents show the characteristic pattern of 45S and 5S

rDNA signals. The 45S rDNA probe labeled the bivalents of chromosome 2 and 4. 5S rDNA signals were found on the bivalents of chromosomes 3, 4 and 5 (Fig. 15b, c). Notably, in the *cap-d2-1* mutants the pattern of the 45S rDNA signal is different compared to wild-type. During diplotene, in addition to the 5S rDNA signal, one chromosome of bivalent 3 shows an extra 45S rDNA signal (arrow Fig. 15b), while in the bivalent 2 one 45S rDNA signal is missing. In metaphase I chromosomes 3 and 2 form a tetravalent (Fig. 15b), meaning they have some homologous region that allows them to pair. In conclusion, line *cap-d2-1* has a heterozygous translocation involving the 45S rDNA between chromosome 2 and 3 (Fig. 15c). Chromosome translocations associated to T-DNA insertions are a common phenomenon in Arabidopsis (Clark & Krysan, 2010). The translocation may induce impaired gametes that can be responsible for the reduced fertility of the *cap-d2-1* plants. Since it cannot be distinguished unambiguously whether the reduced fertility and nuclear phenotype of line *cap-d2-1* is caused by the T-DNA insertion or by the translocation, we excluded it from future experiments.

Other T-DNA insertion lines for *CAP-D2* were analyzed, *cap-d2-2* and *cap-d2-3* (Fig. 15a), but the T-DNA insertion could not be confirmed in neither of them.

4.3.2 Generation of CRISPR/Cas-based *cap-d2* mutants

During interphase, condensin I is required to maintain the compaction of the rDNA and the Zrep heterochromatic region of chromosome Z in chicken (Zhang et al., 2016). Moreover, we found the localization of CAP-D2 in Arabidopsis interphase nuclei. To study the function of CAP-D2 in chromatin organization, a stable Arabidopsis *cap-d2* mutant line was generated via CRISPR/Cas genome engineering. Of the possible protospacers (short sequence complementary to the target region) for *CAP-D2*, four were selected and cloned (Sg1-4, Fig. 15a). Sg1 targets the beginning of the 4th exon, a region highly conserved across CAP-D2 orthologues (Appendix Fig. 3). Sg2 targets a low conserved region at the beginning of the 6th exon, but downstream there is a highly conserved area (Appendix Fig. 3). Sg3 and Sg4 are at the end of the gene, in the 15th and 16th exon, respectively, and in a non-conserved area. The functionality of sgRNAs was tested *in vitro* by incubating them with their target DNA and the recombinant Cas9 protein (Fig. 15d). If cleavage products can be detected after electrophoresis, it means the sgRNA guides the Cas9 protein effectively to its target. Sg1 had a low activity, there was only a partial digestion of the target, but for the other Sg, the digestion was almost complete (Fig. 15d).

Consequently, Sg2, Sg3 and Sg4 were transformed into Arabidopsis. T1 transgenic plants carrying the Sg+Cas9 insert were grown on selective media. The screening for possible mutations was performed from T2 generation on (Table 9). For the screening, a DNA fragment flanking the Sg target was amplified by PCR and directly sequenced (Fig. 15e). Sg4 yielded only one T1 positive plant and in the T2, the three plants analyzed were of wild-type genotype. The number of plants analyzed for Sg4 is very low, so we focused on Sg2 and Sg3 since here the T1 selection was more promising (Table 9). For Sg2 one plant out of 15 was heterozygous and for Sg3, nine out of 10 plants analyzed were heterozygous in T2. Since the analysis of the genotype is based on sequencing, heterozygous, in this context refers to a genotype that is not like wild-type and does not have a clear homozygous mutation. In the sequence chromatogram it can be seen as regular unique peaks until the mutation, and from there on as a mixture of not well defined double or triple peaks (Fig. 15e *cap-d2 Sg3 Hz*). It cannot be distinguished between heterozygous (one wild-type allele and a mutant allele), trans-heterozygous (each allele with a different mutation) and chimeric (each cell has different mutations). The Sg3 Hz plants were analyzed in T3 to allow the segregation of the mutations. In T3 we could obtain two homozygous plants for the same mutation: one base-pair deletion before the PAM sequence (Fig. 15e). This point mutation creates a premature stop codon four amino acids downstream, which in turn produces a truncated CAP-D2 protein of 1266 aa (149 aa smaller than the wild-type version). This small truncation does not seem to affect the growth of the *cap-d2 sg3* mutant plants, which are similar to wild-type (data not shown). A more exhaustive analysis of this mutant will be required to determine if the truncated protein is completely or partially functional. Preliminary results of the centromeric distribution in *cap-d2 sg3* mutant do not show differences compared to wild-type. The low number of only heterozygous plants obtained with the Sg2 RNA, which is in the middle of the gene, could suggest that a bigger truncation of the CAP-D2 protein is lethal.

Table 9. Arabidopsis Sg transformants and genotypes of the screened plants. Wt: genotype like wild-type; Hz: a mutation is present and can be trans-heterozygous, chimeric or heterozygous; Ho: Single homozygous mutation.

| Arabidopsis transformants | T1 Plants selected on media | T2 Plants screened for mutations (Wt:Hz:Ho) | T3 Plants screened for mutations (Wt:Hz:Ho) |
|---------------------------|--------------------------------|--|--|
| Sg2 | 5 | 15 (14:1:0) | - |
| Sg3 | 3 | 10 (1:9:0) | 23 (9:19:2) |
| Sg4 | 1 | 3 (3:0:0) | - |

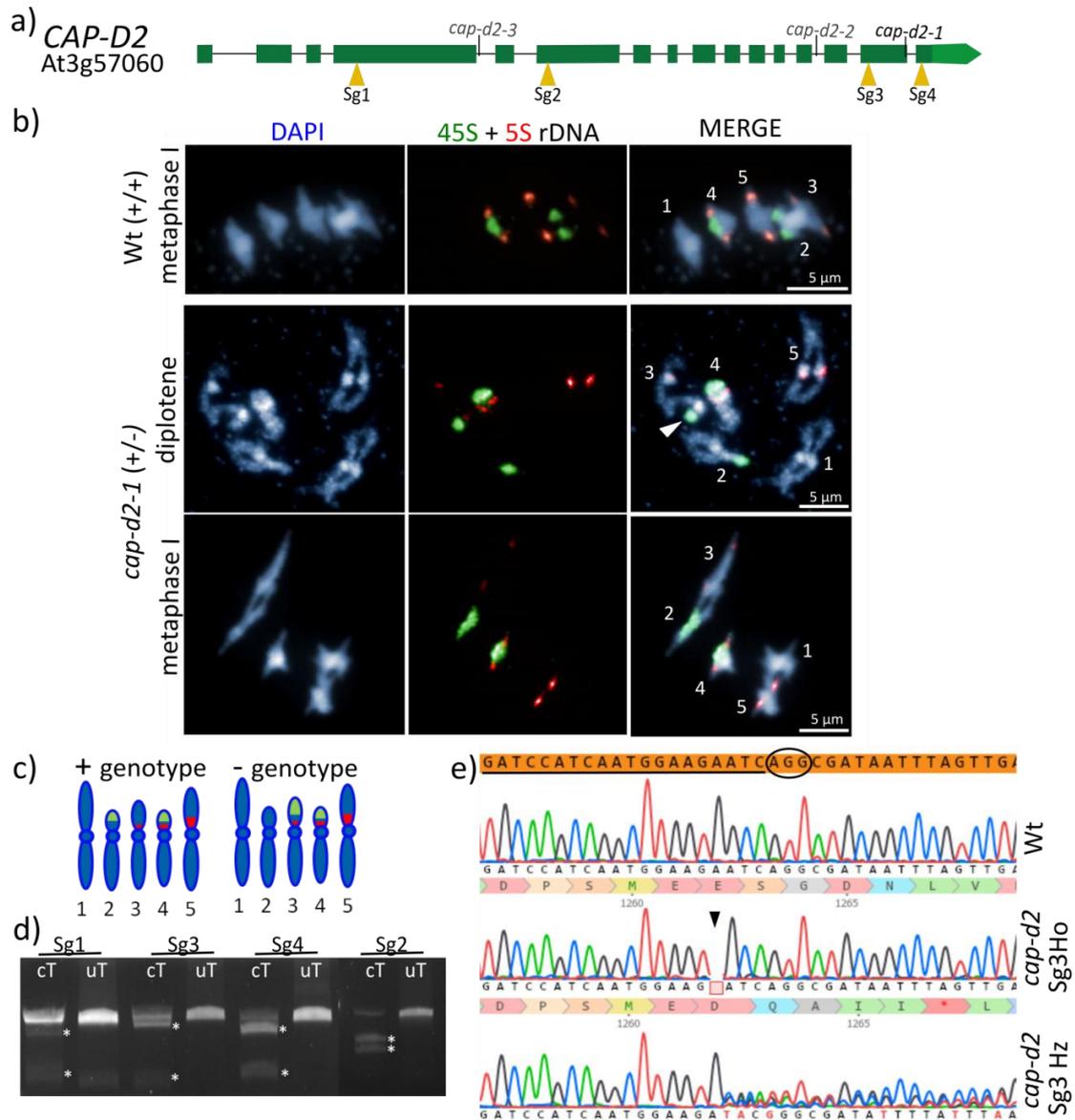


Figure 15. *cap-d2* T-DNA insertion line analysis and generation of a CRISPR/Cas9-based *cap-d2* mutant. a) Schematic representation of *CAP-D2* gene model. Green boxes represent the exons, lines the introns and the light green box is the 3' UTR. T-DNA insertion sites are depicted over the model; names in grey are lines which could not be confirmed. Yellow triangles below the model are the sgRNA genomic regions targeted. b) FISH with the 45S and 5S rDNA probes on pollen mother cells of wild-type (Wt) and *cap-d2-1* plants. White numbers indicate the bivalent chromosome number and the white arrow shows a 45S rDNA signal on one of the chromosome 3 homologues. c) Ideogram representing the Arabidopsis chromosomes, and showing the localization of the 5S and 45S rDNA in Wt (+ genotype) and translocation (- genotype). Wt plants are homozygote for + genotype (+/+) while *cap-d2-1* plants are heterozygous for the translocation (+/-). d) Agarose gel showing the products of the *in vitro* cas9 assay. cT: the template incubated with the SgRNA and the cas9 nuclease, the white asterisks indicate the cut products. uT: uncut template (template incubated only with the cas9 nuclease). e) *cap-d2* Sg3 mutant genotyping. In orange is the *CAP-D2* genomic sequence target for Sg3 (underlined) and the PAM sequence (circled); below are the sequencing chromatogram and amino acid translation of a Wt plant, a *cap-d2* Sg3 plant homozygous for a 1 bp deletion (black triangle) and a heterozygous *cap-d2* Sg3 plant.

5. Discussion

5.1 Regulation of *CAP-D2* and *CAP-D3* expression in Arabidopsis.

Based on *in silico* analysis, it was proposed that HUB1, RBR and the CSN subunits CSN3, CSN4 and CSN5 could influence the transcription of *CAP-D3* and *CAP-D2* (Schubert et al., 2013). In plants defective for these putative regulators, we would expect a deregulation of *CAP-D2* and *CAP-D3* expression associated with a deviation of the centromere distribution in interphase nuclei, like observed for *cap-d2* and *cap-d3* mutants. Our FISH analysis of mutants demonstrated that no clustering of interphase centromeres occurs in the putative regulator mutants. Thus, HUB1, RBR and the CSN subunits seem not to affect the *CAP-D2* and *CAP-D3* expression. The deviation from the prediction could be due to the following reasons:

i) The putative regulators do not affect the expression of *CAP-D2* and *CAP-D3*. However, as the transcript level of *CAP-D2* and *CAP-D3* was not measured in the mutants, we cannot discard that the putative regulators do not affect the transcription of *CAP-D2* and *CAP-D3*. Or if they do, there could be another level of translational regulation.

ii) The analyzed T-DNA mutants are not loss-of-function alleles (knockout) and the remaining protein is enough to maintain a normal expression level of *CAP-D2* and *CAP-D3*. However, the mutants analyzed for the CSN subunits were previously described as knockouts, and the loss of one subunit impairs the CSN function, though the seedling arrest phenotype was not observed in all the mutants (Dohmann et al., 2008a). The *hub-1* and *hub-2* mutants analyzed were earlier described as knockdown alleles, meaning that a region of the gene is still transcribed. However, histone H2B mono-ubiquitination is absent in those mutants (Liu et al., 2007). *rbr* mutants are described as gametophytic lethal (Ebel et al., 2004). Our *rbr* mutants analyzed had the T-DNA insertion in the last exon or at the 3' UTR, and only heterozygous plants could be recovered. Hence, the analyzed *rbr* mutants had most likely a completely functional RBR protein.

iii) In the *csn* and *rbr* mutants, *CAP-D2* and *CAP-D3* transcription is predicted to be up-regulated, which does not mean it translates into higher protein levels. Therefore, the phenotype could be indistinguishable from wild-type. On the other hand, in *Drosophila*, overexpression of the condensin II subunit *CAP-H2* induces separation (dispersion) of the polytene chromosome components and centromere dispersal (Bauer et al., 2012; Buster et al., 2013; Hartl et al., 2008a).

iv) The putative regulators influence the expression of *CAP-D2* and *CAP-D3* at other developmental stages or cell-cycle phase but not in the somatic interphase nuclei analyzed. *HUB1* is ubiquitously expressed (Liu et al., 2007), but *CSN* and *RBR* are expressed in a cell-cycle-specific manner. *CSN* is expressed in meristematic tissue from late S phase until mid-mitosis (Dohmann et al., 2008b), and *RBR* is a core cell cycle regulator (Harashima & Sugimoto, 2016). Therefore, we cannot exclude that *CSN* or *RBR* affect the *CAP-D2* and *CAP-D3* gene expression in other phases of the cell cycle or tissues other than the analyzed.

Indeed, *CAP-D2* and *CAP-D3* are mainly expressed in dividing tissues. The *CAP-D2* transcription level is consistently higher than that of *CAP-D3*. Both are highly expressed in meristems and mitotically active tissues (flower buds, roots) but lowly expressed in non-cycling tissues (leaves). Similarly, the condensin subunits *CAP-H* and *SMC2* are highly expressed in active tissues (Fujimoto et al., 2005; Liu et al., 2002; Siddiqui et al., 2003). Our GUS staining results agree with the previously described *SMC2* expression pattern (Siddiqui et al., 2003).

Introns, when affecting the expression of a gene, often enhance its expression by increasing the transcript amount or by inducing the expression in specific tissues (Heckmann et al., 2011; Parra et al., 2011; Rose et al., 2008). Nonetheless, the second intron of *CAP-D2* could have intragenic regulatory sequences repressing the expression in non-dividing tissues. This is supported herein by the loss of GUS reporter expression in leaves of the Pro6 and Pro9 transgenic plants compared with Pro5, Pro7 and Pro8 plants, which do not carry the second intron. The second intron of the *AGAMOUS* gene is also responsible to inhibit the GUS reporter expression in vegetative tissues, and drives its correct expression in flowers (Sieburth & Meyerowitz, 1997). Moreover, our quantitative RT-PCR results show the transcription of *CAP-D2* in leaves, although at a low level.

Sequences of 391 bp or 474 bp upstream of the start of *CAP-D2* or *CAP-D3*, respectively, are sufficient to drive the expression of GUS. Longer fragments (>1000 bp) do not improve the expression of the reporter gene. Interestingly, *CAP-D2* and *CAP-D3* minimal promoters (short promoters) contain two previously predicted E2F binding sites (Schubert et al., 2013). E2F is a transcriptional activator of genes important for the cell cycle progression. Together with RBR and a dimerization partner, they control the transition from G1 to S phase. E2F sites are also present in the Arabidopsis *SMC2* promoter (Siddiqui et al., 2003). In mouse, CNAP1 (*CAP-D2*) is also a target of E2F (Verlinden et al., 2005). Taken together, the expression patterns, the features of their

promoters and comparisons with other organism, it is plausible that in Arabidopsis, *CAP-D2* and *CAP-D3* are transcriptionally regulated during the cell cycle.

5.2 Condensin I localizes in the cytoplasm during interphase

The localization of CAP-D2 in the mitotic chromosomes of Arabidopsis presents a similar pattern as described for condensin I in other organisms, i.e. in the chromosome axis and enriched in the centromere (Ono et al., 2003, 2004; Savvidou et al., 2005; Schmiesing et al., 2000). During interphase, the most commonly described localization of condensin I is exclusively in the cytoplasm (Gerlich et al., 2006; Hirano, 2012; Hirota et al., 2004; Ono et al., 2004). However, some reports described the localization of condensin I additionally within the nucleus during interphase in *Drosophila*, human cell cultures and chicken (Savvidou et al., 2005; Schmiesing et al., 2000; Zhang et al., 2016). In Arabidopsis protoplasts and *N. benthamiana* epidermal leaves we observed EYFP-fusion proteins of the condensin I specific subunits CAP-D2, CAP-G and CAP-H in the cytoplasm as well as the nucleus. In addition, super-resolution microscopy identified CAP-D2 after immunolabelling with specific antibodies exclusively within the euchromatin of flow-sorted Arabidopsis nuclei intermingling with CAP-D3-positive nuclear subregions. Hence, by using two different methods, EYFP-fusion proteins and immunolocalization, we observed CAP-D2 within the nucleus during interphase. Interestingly, in organisms in which CAP-D2 is described to be mostly cytoplasmatic during interphase (human, mouse), it contains a bipartite nuclear localization signal that is absent in organisms where CAP-D2 is present in the cytoplasm as well as in the nucleus during interphase (chicken, *Drosophila*, Arabidopsis and *S. cerevisiae*)(Ball et al., 2002)(Appendix Fig. 3).

In order to localize CAP-D2 and CAP-D3 *in vivo*, we also generated Arabidopsis stable transformants carrying CAP-D2 or CAP-D3 tagged at its N- or C-terminus to EYFP, and stable *cap-d3* mutants carrying CAP-D3-EYFP. In all constructs, CAP-D2 and CAP-D3 fusion EYFP-proteins were under the regulation of the 35S promoter. The 35S promoter is a constitutive promoter that produces overexpression of the genes it regulates. However, in plants positive for the reporter constructs it was not possible to visualize the fusion proteins directly, or to detect them by anti-GFP-immunolocalization. The constructs are functionally active, since they work in Arabidopsis protoplasts and *N. benthamiana* transient transformations. In addition, the CAP-D3-EYFP construct was able to complement the centromeric phenotype of *cap-d3* *SALK* mutants. Similar problems have been described for GFP-PATRONUS1 Arabidopsis transformants (Zamariola et al., 2014).

These authors suggested that the reason behind could be the low expression or stability of the PATRONUS protein due to the presence of an APC/C degradation box. However, in CAP-D2 no APC/C degradation box is present. The detection of CAP-D2 in leaves from wild-type Arabidopsis plants by Western blot was also not possible, even though we detected CAP-D2 signals by immunolocalization in sorted nuclei. This is possibly due to a low protein level in wild-type leaves since the transcript level in leaves is very low. By Western blot the CAP-D2 protein was detectable in protoplasts only when overexpressed (CAP-D2-EYFP is under the control of the 35S promoter). Similarly, in *Drosophila* the detection of condensin from extracts of non-dividing tissues was also not possible (Cobbe et al., 2006).

5.3 The Arabidopsis condensin I and II subunit composition is similar to other eukaryotes

Protein immunoprecipitation (IP) from flower bud extracts already confirmed the presence of the subunits for condensin I and condensin II in Arabidopsis (Smith et al., 2014). Nonetheless, these IPs were performed with anti-SMC4, which would target both condensin complexes, and therefore could not determine the exact composition of condensin I and II. Here we present additional data to support that in Arabidopsis the two condensin complexes are present, and their protein composition is the same as described for other organisms (Hirano, 2012). Notably, Arabidopsis is the only species in which two SMC2 homologs have been predicted and described (Cobbe & Heck, 2004; Siddiqui et al., 2003). Both, SMC2A and B can act redundantly, but *SMC2A* accounts for most of the *SMC2* transcript pool (Siddiqui et al., 2003). Both SMC2A and B interact with the rest of the condensin subunits in vegetative and somatic tissues (Smith et al., 2014; this study).

We could not confirm the interaction between CAP-H, CAP-G and CAP-D2 by BiFC. This could be due to a number of reasons like a low co-expression of the constructs in *N. benthamiana*, low protein stability, an unbalanced ratio of the translated proteins that are not enough to visualize an interaction, suboptimal conditions for the interaction of the proteins, and/or no interaction in the specific cell type used. Another reason could be that contrary to other organisms, in Arabidopsis, CAP-D2, CAP-G and CAP-H do not interact directly but indirectly via an unknown protein. However, the last explanation is unlikely, due to the conserved homology between all the condensin I complexes among different organisms (Hirano, 2012; Onn et al., 2007).

In human cells and *Drosophila*, CAP-D3 interacts with RB and promotes the correct localization of condensin II in the chromosomes (Longworth et al., 2008). In *Arabidopsis*, this interaction seems not to be conserved, since we could not detect RBR among the proteins that co-purified with CAP-D3. In the human cell culture, Cdc20, a component of the anaphase-promoting complex E3 ubiquitin ligase, interacts and regulates CAP-H2 (Kagami et al., 2017). In *Drosophila*, CAP-H2 also interacts and is regulated by SCF^{Slimb} (Buster et al., 2013). The Skp-cullin-F-box (SCF) is an E3 ubiquitin ligase regulated by CSN (Hotton & Callis, 2008). Among the proteins that co-purified with CAP-D2 and CAP-D3 there were components of the ubiquitin-26S proteasome pathway. CULLIN 1 co-purified with CAP-D2 and CSN1 with CAP-D3 in all the replicates. CSN3 and CSN4 also co-purified with CAP-D3 in the three triplicates but also in 3 out of 115 of the non-specific proteins affinity purifications (data not shown). CULLIN1 and CULLIN3 were present in two of the CAP-D3 triplicates. This data suggest that in *Arabidopsis*, ubiquitination could be involved in the regulation of the condensins.

A screen for functional partners of condensin in yeast identified, among others, two chromatin remodeling proteins and a histone deacetylase, as collaborators of condensin for chromosome condensation (Robellet et al., 2014). In this screen they looked for mutations that produced cell lethality when combined with *cut3* (SMC4) mutation, i.e. both proteins are needed for cell viability but do not necessarily interact. However, we also found chromatin remodeling enzymes (CHR17, CRH19 and BRAHMA), histone chaperones (NAP1;1 and NAP1;2), a histone deacetylase (HDC1) and a histone acetyltransferase (ELO3) in the affinity purification experiments for CAP-D2 and CAP-D3. All of them are chromatin modifiers important for plant development (Gentry & Hennig, 2014; Perrella et al., 2013; Skylar et al., 2013).

5.4 CAP-D3 organizes the chromatin during interphase

5.4.1 CAP-D3 and its influence in euchromatin

In *Drosophila*, CAP-D3 and CAP-H2 are needed to form compact chromosomes (Bauer et al., 2012; Hartl et al., 2008b). In embryonic stem cells of mice, condensin is also required to maintain a compact chromatin structure during interphase (Fazio & Panning, 2010). On the other hand, in *C. elegans*, the depletion of SMC4, CAP-G2 or HCP-6 (CAP-D3) does not change the chromosome volume (Lau et al., 2014). In *Arabidopsis*, previous studies suggested also an influence of CAP-D3 in the formation of the chromosome territories at interphase based on FISH experiments (Schubert

et al., 2013). In *cap-d3 SAIL* and *cap-d3 SALK* mutants, the top arm of chromosome 1 appears dispersed during interphase, and the bottom arm of chromosome 4 is dispersed in *cap-d3 SAIL* but not *cap-d3 SALK* mutants (Schubert et al., 2013). Using FISH probes against a part of chromosome 1 bottom arm, we could not detect an increase of the hybridization signal area in *cap-d3* mutants compared to wild-type plants. The differences could be explained by labeling only one fourth of the chromosome arm by FISH, while in the previous study the whole chromosome arm was visualized, and by the different ways to quantify the dispersion of the interphase chromatin.

It has been proposed that the condensins via maintaining the chromatin condensation, consequently also maintain the nuclear shape and size (George et al., 2014). In agreement, the depletion of CAP-H2 in *Drosophila* and SMC2, CAP-H2 and CAP-D3 in human cell cultures, increases the nuclear size (George et al., 2014). In embryonic stem cells of mice, the depletion of SMC2 causes chromatin decondensation as well as an increase of the nuclear volume (Fazio & Panning, 2010). We could not observe a change in the nuclear area of the *cap-d3* mutants compared to wild-type plants, supporting that there is no chromosome territory dispersion in *cap-d3* mutants. However, these results do not exclude that CAP-D3 is involved in the organization of the chromosome territories. The mutants used in the analysis (*cap-d3 SAIL* and *cap-d3 SALK*) have knockdown alleles, meaning that there is still a truncated transcript that can produce a partially functional CAP-D3 protein (Schubert et al., 2013).

Besides its structural function, condensin II has also been described to influence transcription. In embryonic stem cells of mice, CAP-H2 and cohesin are enriched at active enhancers and promoters of pluripotency genes. Both are needed for their normal expression and the cohesin effect on transcription is more prominent than the condensin effect (Downen et al., 2013). On the other hand, another study with embryonic stem cells of mice and human cell cultures, reported an enrichment of CAP-H2 and CAP-D3 at TFIIC (transcription factor associated to polymerase III) binding sites, and to a lesser extent, at promoters of housekeeping genes (Yuen et al., 2017). CAP-H2 knock-down reduces the expression of housekeeping genes involved in RNA-processing, and translation (Yuen et al., 2017). In *Drosophila*, CAP-D3 together with RB regulates clusters of genes which are involved in programs of specific cell-types, like genes involved in innate immunity (Longworth et al., 2012). The authors suggest that CAP-D3 and RB could protect these gene clusters to make them accessible in case of infection. Accordingly, CAP-D3 deficient flies die more quickly in response to a bacterial infection than wild-type organisms (Longworth et al., 2012). Even

though *Arabidopsis cap-d3* mutants do not show major transcriptional changes based on our comparative RNA-seq analysis, CAP-D3 might still affect the expression of genes involved in transcription and response to stress. This conclusion arise from our observation of *Arabidopsis cap-d3* mutants dying more quickly from stress, like pathogen infection, than wild-type plants (preliminary results).

A recent study in *Arabidopsis* condensin mutants showed that SMC4, but not CAP-D3, is important to maintain the repression of pericentromeric retrotransposons independently of DNA methylation (Wang et al., 2017). Accordingly, we did not observe a higher transcription of retrotransposons in any of the *cap-d3* mutants. Moreover, the protein coding genes up-regulated in the *scm4* mutants are distributed all over the genome (Wang et al., 2017) as we also observed in both *cap-d3* mutants. The genes up-regulated in the *scm4* mutants are mainly involved in flower development, reproductive processes and DNA repair. In contrast, we observed in *cap-d3* mutants a differential expression of genes involved in transcription and stress response. This difference could be due to the observation of combined effects of both condensin complexes I and II in *scm4* mutants, while in *cap-d3* mutants only the condensin complex II is compromised.

5.4.2 CAP-D3 and its role in heterochromatin organization

CAP-H2 promotes the dispersion of heterochromatic sequences in *Drosophila* during interphase (Bauer et al., 2012; Buster et al., 2013). Similarly, in *Arabidopsis*, depletion of CAP-D3 produces the clustering of the centromeres at interphase (Schubert et al., 2013). We confirmed this phenotype in the *cap-d3* mutants. CAP-D3 depletion also resulted in the clustering of the 45S rDNA loci but not of the 5S rDNA sites. Given the proximity of the 5S rDNA to the pericentromeric heterochromatin and that both the centromeres and the 45S rDNA cluster in *cap-d3* mutants, we expected also a clustering of the 5S rDNA loci. In other systems like in protoplasts of *Arabidopsis*, the 45S rDNA and the 5S rDNA behave differently during the reassembly of heterochromatin (Tessadori et al., 2007). 5S and 45S rDNA are transcribed by different RNA polymerases, the RNA polymerase III and RNA polymerase I, respectively (Layat et al., 2012). Therefore, the different clustering behaviors of both rDNA repeats in *cap-d3* mutants could be due to their different structural and functional properties. Moreover, condensin of fission yeast, which is similar to condensin I, binds to polymerase III genes (tRNA and 5S rDNA) and mediates their localization near the centromeres (Iwasaki et al., 2010).

The structure of chromocenters has been widely used to study extensive chromatin rearrangements and heterochromatin formation in *Arabidopsis*. The organization of the heterochromatin in chromocenters may be affected by stress (Probst & Mittelsten Scheid, 2015), cell de-differentiation (Tessadori et al., 2007), seed maturation (van Zanten et al., 2011), seedling development (Bourbousse et al., 2015), endosperm formation (Baroux et al., 2007) and the ploidy level (Schubert et al., 2006, 2012). The chromocenters are highly DNA methylated (Fransz et al., 2002) and hypomethylation results in a reduction of their size and dispersion of the pericentromeric repeats (Soppe et al., 2002). DNA hypomethylation mutants have also a reduction of histone H3K9 methylation (Soppe et al., 2002). Other proteins important for chromocenter maintenance are the MicroRNA ATPases (MORC) (Moissiard et al., 2012) and the nuclear periphery components Linker of Nucleoskeleton and Cytoskeleton complex (LINC)(Poulet et al., 2017) and CRWded Nuclei proteins (CRWN also known as Little Nuclei)(Wang et al., 2013). *morc* mutants show decondensation of the pericentromeric heterochromatin without changes in DNA and histone methylation (Moissiard et al., 2012). Nuclei deficient in components of the LINC complex present a reduced nuclear volume and decondensed chromocenters. Such chromocenters do not localize at the nuclear periphery, but in a more internal position during interphase (Poulet et al., 2017). Nuclei deficient in the CRWN proteins present a reduced nuclear volume and a reduced number of chromocenters due to clustering (Wang et al., 2013). The nuclear and chromocenter phenotype which we observed in the *cap-d3* mutants is different to that previously described. There is a chromocenter clustering instead of decondensation, and they localize at the nuclear periphery. The nuclear area does not change compared to wild-type, and the general degree of DNA and histone methylation is unaffected. Moreover, hypomethylation, *linc* and *morc* mutants do not show transcriptional silencing of centromeric and pericentromeric repeats, and of silenced genes (Moissiard et al., 2012; Poulet et al., 2017). In contrast, CAP-D3 has little effect on silencing. *cap-d3* mutants do not show an increased transcription of transposable elements (Wang et al., 2017). MORC, CRWN and LINC proteins localized close to the chromocenters. MORC proteins, localize in foci adjacent to the chromocenters (Moissiard et al., 2012). The CRWN1 and CRWN4 proteins localize in the nuclear periphery (Sakamoto & Takagi, 2013), and the LINC complex in the nuclear envelope (Tatout et al., 2014). Conversely, CAP-D3 influences the organization of the chromocenters but localizes exclusively in euchromatin during interphase. Therefore, CAP-D3 has mainly a structural role during interphase and affects the clustering of chromocenters without localizing them.

Computer simulations modeling of Arabidopsis chromosomes as polymers predict that the position of the chromocenters in the nucleus is due to non-specific interactions (de Nooijer et al., 2009). Three types of chromosomes were modeled for the simulations: a linear chromosome with a chromocenter, a chromosome with loops and a chromocenter (LAC), and a Rosette chromosome from which loops emanate from the chromocenter (Fig. 16a). With all chromosome models, the chromocenters localized in the periphery, as described *in vivo* (Fransz et al., 2002), but only the chromosomes with loops (the LAC and the Rosette chromosomes) form chromosome territories (de Nooijer et al., 2009). On the one hand, only the Rosette model explains the central position of the nucleolus, but does not explain the association of chromocenters observed *in vivo*, since in the model no chromocenter clustering occurs (Fig. 16a). On the other hand, the LAC model does not explain the position of the nucleoli, and the chromocenters associate in 1 or 2 clusters (Fig. 16a). Then, the authors propose that only the loops that emanate from the chromocenters suppress chromocenter clustering (de Nooijer et al., 2009). Indeed, depletion-attraction forces predict that big particles in an environment crowded with small particles will tend to cluster together (Marenduzzo et al., 2006) (Fig. 16b). This situation can be applied to the nucleus where the chromocenters act as big particles and the euchromatin as small particles. If nothing prevents their association, the chromocenters will cluster. In *cap-d3* mutants we observed chromocenter clustering but not chromosome territory dispersion and it localizes in euchromatin. CAP-D3 could help making the loops more stable (stiff) but does not intervene in the base of the loops, presumably close to the chromocenters (Fig. 16c). Then, in absence of CAP-D3, the loops will not be “strong” enough to prevent the clustering of the chromocenters but the chromosome territories will stay intact (the loops are still formed)(Fig. 16c). It could also explain why CAP-D3 localizes in euchromatin and is not enriched in chromocenters or close to them. During mitosis, CAP-D3 is needed to confer the rigidity of the chromosome arms (Green et al., 2012).

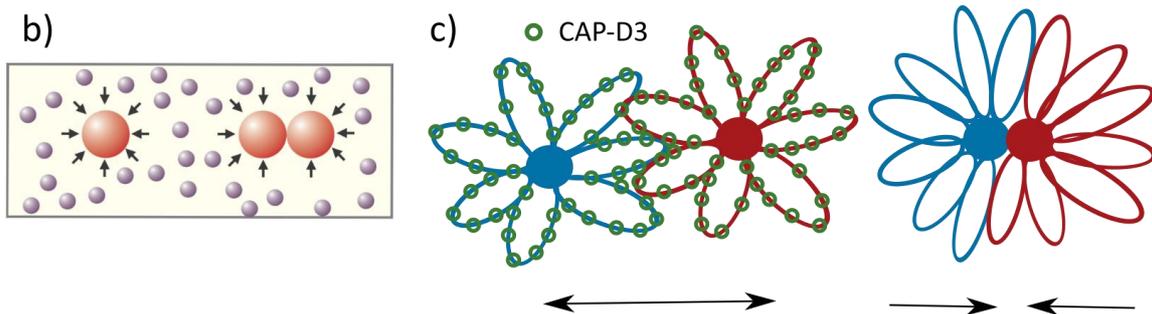
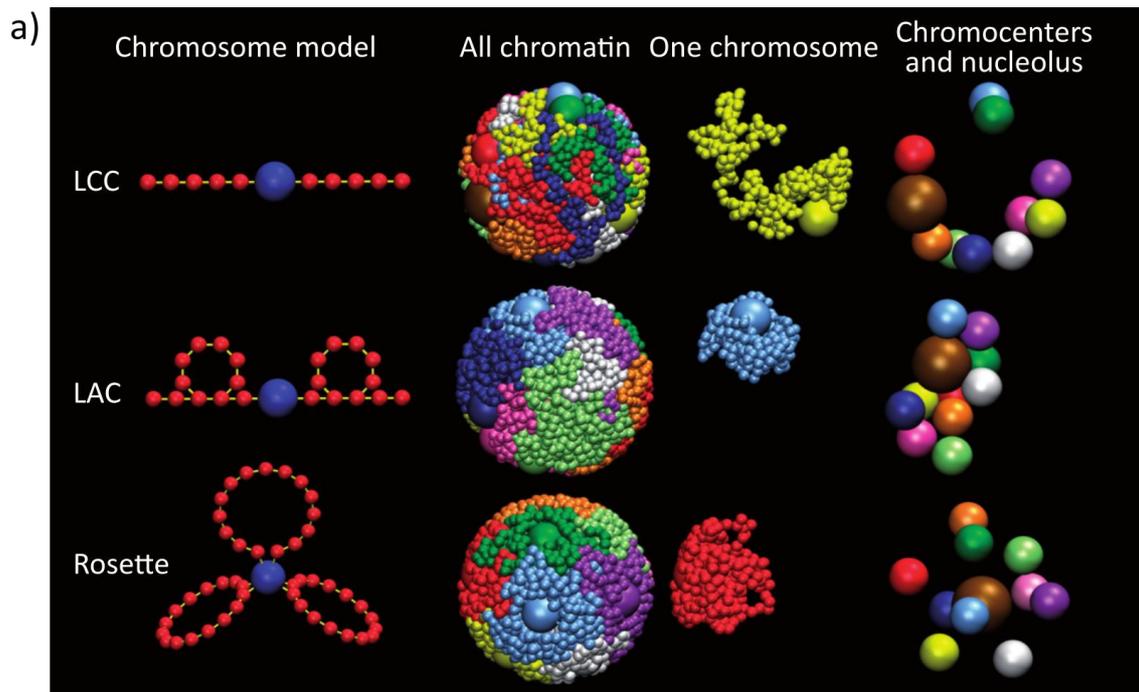


Figure 16. Computer simulated nuclear organization in Arabidopsis and the possible function of CAP-D3. a) Computer simulations of Arabidopsis nuclear organization using three different chromosome models. The Linear Chain with Chromocenter (LCC) chromosome simulation shows disorganized chromosomes. The chromocenters cluster in the nuclear periphery and the nucleolus (brown sphere) localizes also in the periphery. The Looped Arms with Chromocenters (LAC) and the Rosette chromosome models form chromosome territories. In the LAC model the chromocenters cluster in the nuclear periphery and the nucleolus is in a peripheral position as well. In the Rosette model the chromocenters are in the periphery but do not cluster like *in vivo* where 8 or 9 chromocenters are visible. The nucleolus adopts a central position like *in vivo*. Picture modified from de Nooijer et al., (2009). b) Schemata representing depletion-attraction forces. The small spheres representing small particles, macromolecules and euchromatin bombard the bigger spheres (large complexes, organelles, chromocenters...) from all sides (left drawing). When two big spheres meet, the forces that the small particles exert upon them keep them together (right drawing). Picture modified from Marenduzzo et al., (2006). c) Possible CAP-D3 function model. The Rosette chromosome model is assumed, two chromosomes are represented in blue and in red. CAP-D3 (green circles) localizes in euchromatin making the chromatin loops rigid. This rigid loops keep the chromocenters separated (left drawing). In absence of CAP-D3 the chromatin loops are not strong enough to counter balance the depletion-attraction forces and therefore, the chromocenters cluster (right).

5.5 Arabidopsis CAP-D2 C-terminal domain structure differs from human CAP-D2

In humans, the last 113 amino acids of CAP-D2 contain a chromosome-targeting domain which is essential for nuclear localization and chromosome association, but not for the interaction with the other components of the complex (Ball et al., 2002). However, the *cap-d2 sg3* mutant has a truncated analogous region, but the plants show no deviating phenotype. Therefore, in Arabidopsis, this chromosome-targeting domain must be in another region or may not be essential. In Arabidopsis CAP-D2 is present in the nucleus during interphase. It might not need a nuclear localization signal and the chromosome association could be different than in humans. We could not obtain any homozygous plants with a bigger truncation, suggesting that CAP-D2 is essential.

In Arabidopsis the T-DNA insertion line *cap-d2-1* is only viable in heterozygosity (Schubert et al., 2013). This line also presents a translocation involving the 45S rDNA (this study). The lethality in homozygosity of the *cap-d2-1* line is possibly due to the translocation and not to the mutation of the *CAP-D2* gene. The CRISPR-Cas9 induced *cap-d2 sg3* mutation is upstream of the *cap-d2-1* mutation and is viable in homozygosity. Therefore, the lethality of the *cap-d2-1* line is not due to the mutation but caused by the translocation in the same chromosome as the T-DNA insertion that involves the 45s rDNA.

6. Outlook

Due to time limitations and new data derived from this study, subsequent questions need to be addressed in the future:

- i) The effect of CSN and RB in *CAP-D2* and *CAP-D3* expression in new Arabidopsis mutants for these putative regulators and mitotically active tissues.
- ii) The functionality of the E2F sites on *CAP-D2* and *CAP-D3* promoters. The activity could be checked by mutating the sites in a GUS reporter expression constructs.
- iii) The localization of *CAP-D2* and *CAP-D3* in Arabidopsis transgenic plants using EYFP-fusion proteins under the regulation of the native promoters. Also, the complementation of *cap-d3* *SALK* mutant plants with a *CAP-D3*-EYFP fusion protein under the control of the native promoter. The use of the native promoters will ensure a regulation of the protein closer to wild-type conditions than the 35S promoter.
- iv) Confirmation of the direct interaction between *CAP-H*, *CAP-G* and *CAP-D2* that could not be assessed by the BiFC experiments. A yeast two-hybrid assay could determine the direct interaction between the proteins. However, the yeast two-hybrid assay involves the expression of the proteins in a heterologous system like the BiFC in *N. benthamiana* and could cause similar problems.
- v) Measurement of chromosome territory dispersion in *cap-d3* mutants using FISH against the complete chromosome territory 1 top. This will allow a direct comparison with previous results.
- vi) In *cap-d3* mutants, genes involved in stress response are differentially expressed compared to wild-type plants. To evaluate the function of *CAP-D3* in stress, the response of *cap-d3* mutants to stress should be assessed.
- vii) If *CAP-D3* maintains the chromatin loops rigidity, the interactions between the loops should change in *cap-d3* mutants. The use of Hi-C on *cap-d3* mutants compared to wild-type plants could account for different contacts between the chromatin loops.
- viii) Analysis of the centromere organization in the *cap-d2 sg3* mutant by FISH to evaluate the function of *CAP-D2* in the organization of interphase chromatin. To analyze other *cap-d2* alleles, the number of plants analyzed for the *Sg2* transformants should be increased to obtain homozygous or heterozygous mutant plants. If no homozygous or

heterozygous plants could be obtained, then an inducible *cap-d2* mutant plant could be generated.

7. Summary

In this work we present new findings about the condensin subunits CAP-D2 and CAP-D3 in *Arabidopsis thaliana*.

- i) *Arabidopsis* condensin CAP-D2 and CAP-D3 present characteristics common to other eukaryotic condensins. 1) They are expressed in mitotically active tissues. 2) CAP-D2 localizes along the sister chromatids and centromeres of mitotic chromosomes. 3) CAP-D2 and CAP-D3 interact with the other putative subunits of condensin I and II respectively.

In *Arabidopsis*, a fragment of less than 500 bp containing E2F sites is enough to drive the expression of both CAP-D2 and CAP-D3. Also, CAP-D2 EYFP-fusion proteins and the generation of a specific antibody against CAP-D2 showed its localization in the cytoplasm and nucleus during interphase. During interphase CAP-D2 and CAP-D3 localize intermingled exclusively in euchromatin.

- ii) Plants defective for CAP-D3 do not present changes in the nuclear volume, chromosome territory dispersion and epigenetic landscape compared to wild-type plants. However, CAP-D3 is needed for the organization of the centromeric regions and 45S rDNA during interphase and the correct expression of genes involved in stress.
- iii) Commercial *cap-d2-1* T-DNA insertion line presents a translocation between chromosome 3 and 2 that involves the 45SrDNA loci. Hence, a new *cap-d2* mutant, *cap-d2 sg3*, was generated using CRISPR-Cas9 technology. In *cap-d2 sg3* mutants, CAP-D2 presents a truncation of the last 149 amino acids but the plants show no obvious growing defects.

8. Zusammenfassung

In dieser Arbeit werden folgende neue Erkenntnisse über die Kondensin-Untereinheiten CAP-D2 und CAP-D3 in *Arabidopsis thaliana* vorgestellt.

- i) *Arabidopsis* CAP-D2 und CAP-D3 weisen Eigenschaften auf, die anderen Eukaryoten-Kondensinen gemeinsam sind. 1) Sie werden in mitotisch aktiven Geweben exprimiert. 2) CAP-D2 Proteine lokalisieren entlang der Schwesterchromatiden und Zentromere mitotischer Chromosomen. 3) CAP-D2 und CAP-D3 interagieren mit den anderen mutmaßlichen Untereinheiten von Kondensin I bzw. Kondensin II.

In *Arabidopsis* ist ein Fragment von weniger als 500 bp, das E2F-Stellen enthält, ausreichend, um die Expression von sowohl CAP-D2 als auch CAP-D3 zu steuern. Auch CAP-D2 EYPF-Fusionsproteine und die Erzeugung eines spezifischen Antikörpers gegen CAP-D2 zeigten eine Lokalisierung im Cytoplasma und im Interphase-Zellkern. Während der Interphase befinden sich CAP-D2 und CAP-D3 im Euchromatin.

- ii) Pflanzen, die für CAP-D3 defekt sind, zeigen im Vergleich zu Wildtyppflanzen keine Veränderungen des Kernvolumens, der Verteilung der Chromosomenterritorien und der epigenetischen Landschaft. CAP-D3 wird jedoch für die Organisation der Zentromerregionen und 45S-rDNA während der Interphase sowie für die korrekte Expression von Genen benötigt, welche stressreguliert sind.
- iii) Die *cap-d2-1*-T-DNA-Insertionslinie besitzt eine Translokation zwischen den Chromosomen 3 und 2, dabei ist der 45S rDNA Locus involviert. Eine neue *cap-d2*-Mutante, *cap-d2 sg3*, wurde unter Verwendung der CRISPR-Cas9-Technologie erzeugt. *cap-d2 sg3* zeigt eine Verkürzung der letzten 149 Aminosäuren. Diese Mutantpflanzen zeigen keine offensichtlichen Wachstumsdefekte.

9. References

- Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, *31*(2), 166–169.
- Antosz, W., Pfab, A., Ehrnsberger, H. F., Holzinger, P., Köllen, K., Mortensen, S. A., Grasser, K. D. (2017). The Composition of the Arabidopsis RNA Polymerase II Transcript Elongation Complex Reveals the Interplay between Elongation and mRNA Processing Factors. *The Plant Cell*, *29*(4).
- Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, *408*(6814), 796–815.
- Armstrong, S. J., Franklin, F. C., & Jones, G. H. (2001). Nucleolus-associated telomere clustering and pairing precede meiotic chromosome synapsis in *Arabidopsis thaliana*. *Journal of Cell Science*, *114*(23), 4207–4217.
- Ball, A. R., Schmiesing, J. A., Zhou, C., Gregson, H. C., Okada, Y., Doi, T., & Yokomori, K. (2002). Identification of a Chromosome-Targeting Domain in the Human Condensin Subunit CNAP1/hCAP-D2/Eg7. *Molecular and Cellular Biology*, *22*(16), 5769–5781.
- Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Research*, *21*(3), 381–395.
- Baroux, C., Pecinka, A., Fuchs, J., Schubert, I., & Grossniklaus, U. (2007). The Triploid Endosperm Genome of *Arabidopsis* Adopts a Peculiar, Parental-Dosage-Dependent Chromatin Organization. *The Plant Cell*, *19*(6), 1782–1794.
- Bauer, C. R., Hartl, T. A., & Bosco, G. (2012). Condensin II Promotes the Formation of Chromosome Territories by Inducing Axial Compaction of Polyploid Interphase Chromosomes. *PLoS Genetics*, *8*(8), e1002873.
- Berr, A., Pecinka, A., Meister, A., Kreth, G., Fuchs, J., Blattner, F. R., Schubert, I. (2006). Chromosome arrangement and nuclear architecture but not centromeric sequences are conserved between *Arabidopsis thaliana* and *Arabidopsis lyrata*. *The Plant Journal*, *48*(5), 771–783.
- Berr, A., & Schubert, I. (2007). Interphase chromosome arrangement in *Arabidopsis thaliana* is similar in differentiated and meristematic tissues and shows a transient mirror symmetry after nuclear division. *Genetics*, *176*(2), 853–863.
- Bolaños-Villegas, P., Yang, X., Wang, H. J., Juan, C. T., Chuang, M. H., Makaroff, C. A., & Jauh, G. Y. (2013). Arabidopsis CHROMOSOME TRANSMISSION FIDELITY 7 (AtCTF7/ECO1) is required for DNA repair, mitosis and meiosis. *The Plant Journal*, *75*(6), 927–940.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120.
- Bourbousse, C., Mestiri, I., Zabulon, G., Bourge, M., Formiggini, F., Koini, M. A., Barneche, F. (2015). Light signaling controls nuclear architecture reorganization during seedling establishment. *Proceedings of the National Academy of Sciences*, *112*(21), E2836–E2844.
- Boveri, T. (1909). Die Blastomerenkerne von *Ascaris megalo ephala* und die Theorie der Chromosomenindividualität. *Archiv Für Zellforschung*, *3*(181–268).
- Boyle, S., Gilchrist, S., Bridger, J. M., Mahy, N. L., Ellis, J. A., & Bickmore, W. A. (2001). The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Human Molecular Genetics*, *10*(3), 211–219.
- Brinkman, E. K., Chen, T., Amendola, M., & Van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Research*.

- Buster, D. W., Daniel, S. G., Nguyen, H. Q., Windler, S. L., Skwarek, L. C., Peterson, M., Rogers, G. C. (2013). SCF5limbubiquitin ligase suppresses condensin II-mediated nuclear reorganization by degrading Cap-H2. *Journal of Cell Biology*, 201(1), 49–63.
- Campbell, B. R., Song, Y., Posch, T. E., Cullis, C. A., & Town, C. D. (1992). Sequence and organization of 5S ribosomal RNA-encoding genes of *Arabidopsis thaliana*. *Gene*, 112, 225–228.
- Chelysheva, L., Diallo, S., Vezon, D., Gendrot, G., Vrielynck, N., Belcram, K., Grelon, M. (2005). AtREC8 and AtSCC3 are essential to the monopolar orientation of the kinetochores during meiosis. *Journal of Cell Science*, 118, 4621–32.
- Clapier, C. R., & Cairns, B. R. (2009). The Biology of Chromatin Remodeling Complexes. *Annual Review of Biochemistry*, 78(1), 273–304.
- Clark, K. A., & Krysan, P. J. (2010). Chromosomal translocations are a common phenomenon in *Arabidopsis thaliana* T-DNA insertion lines. *The Plant Journal : For Cell and Molecular Biology*, 64(6), 990–1001.
- Clough, S. J., & Bent, A. F. (1998). Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal*, 16(6), 735–743.
- Cobbe, N., & Heck, M. M. S. (2004). The Evolution of SMC Proteins: Phylogenetic Analysis and Structural Implications. *Molecular Biology and Evolution*, 21(2), 332–347.
- Cobbe, N., Savvidou, E., & Heck, M. M. S. (2006). Diverse mitotic and interphase functions of condensins in *Drosophila*. *Genetics*, 172(2), 991–1008.
- Cremer, T., & Cremer, M. (2010). Chromosome territories. *Cold Spring Harbor Perspectives in Biology*, 2(3).
- D'Ambrosio, C., Schmidt, C. K., Katou, Y., Kelly, G., Itoh, T., Shirahige, K., & Uhlmann, F. (2008a). Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. *Genes and Development*, 22(16), 2215–2227.
- D'Ambrosio, C., Schmidt, C. K., Katou, Y., Kelly, G., Itoh, T., Shirahige, K., & Uhlmann, F. (2008b). Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. *Genes & Development*, 22, 2215–2227.
- de Nooijer, S., Wellink, J., Mulder, B., & Bisseling, T. (2009). Non-specific interactions are sufficient to explain the position of heterochromatic chromocenters and nucleoli in interphase nuclei. *Nucleic Acids Research*, 37(11), 3558–3568.
- Dekker, J., & Mirny, L. (2016). The 3D Genome as Moderator of Chromosomal Communication. *Cell*, 164(6), 1110–1121.
- Dekker, J., Rippe, K., Dekker, M., & Kleckner, N. (2002). Capturing chromosome conformation. *Science*, 295(5558), 1306–1311.
- Denker, A., & de Laat, W. (2016). The second decade of 3C technologies: detailed insights into nuclear organization. *Genes & Development*, 30(12), 1357–82.
- Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J. E., Lee, A. Y., Ren, B. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature*, 518(7539), 331–336.
- Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398), 376–380.
- Dohmann, E. M. N., Kuhnle, C., & Schwechheimer, C. (2005). Loss of the CONSTITUTIVE PHOTOMORPHOGENIC9 signalosome subunit 5 is sufficient to cause the cop/det/fus mutant phenotype in *Arabidopsis*. *The Plant Cell*, 17(7), 1967–1978.
- Dohmann, E. M. N., Levesque, M. P., De Veylder, L., Reichardt, I., Jürgens, G., Schmid, M., & Schwechheimer, C. (2008a). The *Arabidopsis* COP9 signalosome is essential for G2 phase progression and genomic stability. *Development*, 135(11).

- Dohmann, E. M. N., Levesque, M. P., Isono, E., Schmid, M., & Schwechheimer, C. (2008b). Auxin responses in mutants of the Arabidopsis CONSTITUTIVE PHOTOMORPHOGENIC9 signalosome. *Plant Physiology*, *147*(3), 1369–79.
- Downen, J. M., Bilodeau, S., Orlando, D. A., Hübner, M. R., Abraham, B. J., Spector, D. L., & Young, R. A. (2013). Multiple structural maintenance of chromosome complexes at transcriptional regulatory elements. *Stem Cell Reports*, *1*(5), 371–378.
- Du, Z., Zhou, X., Ling, Y., Zhang, Z., & Su, Z. (2010). agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Research*, *38* (Web Server issue), W64–70.
- Dürr, J., Lolas, I. B., Sørensen, B. B., Schubert, V., Houben, A., Melzer, M., Grasser, K. D. (2014). The transcript elongation factor SPT4/SPT5 is involved in auxin-related gene expression in Arabidopsis. *Nucleic Acids Research*, *42*(7), 4332–47.
- Ea, V., Baudement, M. O., Lesne, A., & Forné, T. (2015). Contribution of topological domains and loop formation to 3D chromatin organization. *Genes*, *6*(3), 734–750.
- Ebel, C., Mariconti, L., & Gruissem, W. (2004). Plant retinoblastoma homologues control nuclear proliferation in the female gametophyte. *Nature*, *429*(6993), 776–780.
- Fang, Y., & Spector, D. L. (2005). Centromere positioning and dynamics in living Arabidopsis plants. *Molecular Biology of the Cell*, *16*(12), 5710–8.
- Fausser, F., Schiml, S., & Puchta, H. (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. *The Plant Journal*, *(79)*, 348–359.
- Fazio, T. G., & Panning, B. (2010). Condensin complexes regulate mitotic progression and interphase chromatin structure in embryonic stem cells. *Journal of Cell Biology*, *188*(4), 491–503.
- Feng, J., & Shen, W. H. (2014). Dynamic regulation and function of histone monoubiquitination in plants. *Frontiers in Plant Science*, *5*, 1–9.
- Feng, S., Cokus, S. J., Schubert, V., Zhai, J., Pellegrini, M., & Jacobsen, S. E. (2014). Genome-wide Hi-C Analyses in Wild-Type and Mutants Reveal High-Resolution Chromatin Interactions in Arabidopsis. *Molecular Cell*, *55*, 694–707.
- Feng, S., & Jacobsen, S. E. (2011). Epigenetic modifications in plants: An evolutionary perspective. *Current Opinion in Plant Biology*, *14*(2), 179–186.
- Fleury, D., Himanen, K., Cnops, G., Nelissen, H., Boccardi, T. M., Maere, S., Van Lijsebettens, M. (2007). The Arabidopsis thaliana Homolog of Yeast BRE1 Has a Function in Cell Cycle Regulation during Early Leaf and Root Growth. *The Plant Cell*, *19*(2), 417–432.
- Fransz, P., Armstrong, S., Alonso-Blanco, C., Fischer, T. C., Torres-Ruiz, R. A., & Jones, G. (1998). Cytogenetics for the model system Arabidopsis thaliana. *The Plant Journal*, *13*(6), 867–876.
- Fransz, P., de Jong, J. H., Lysak, M., Castiglione, M. R., & Schubert, I. (2002). Interphase chromosomes in Arabidopsis are organized as well defined chromocenters from which euchromatin loops emanate. *Proceedings of the National Academy of Sciences*, *99*(22), 14584–14589.
- Freeman, L., Aragon-Alcaide, L., & Strunnikov, A. (2000). The condensin complex governs chromosome condensation and mitotic transmission of rDNA. *Journal of Cell Biology*, *149*(4), 811–824.
- Fuchs, J., Demidov, D., Houben, A., & Schubert, I. (2006). Chromosomal histone modification patterns – from conservation to diversity. *Trends in Plant Science*, *11*(4), 199–208.
- Fujimoto, S., Yonemura, M., Matsunaga, S., Nakagawa, T., Uchiyama, S., & Fukui, K. (2005). Characterization and dynamic analysis of Arabidopsis condensin subunits, AtCAP-H and AtCAP-H2. *Planta*, *222*(2), 293–300.
- Ganji, M., Shaltiel, I. A., Bisht, S., Kim, E., Kalichava, A., Haering, C. H., & Dekker, C. (2018). Real-time imaging

of DNA loop extrusion by condensin. *Science*, 360(6384), 102–105.

- Gentry, M., & Hennig, L. (2014). Remodelling chromatin to shape development of plants. *Experimental Cell Research*, 321(1), 40–46.
- George, C., Bozler, J., Nguyen, H., & Bosco, G. (2014). Condensins are Required for Maintenance of Nuclear Architecture. *Cells*, 3(3), 865–882.
- Gerlich, D., Hirota, T., Koch, B., Peters, J.-M., & Ellenberg, J. (2006). Condensin I Stabilizes Chromosomes Mechanically through a Dynamic Interaction in Live Cells. *Current Biology*, 16(4), 333–344.
- Gernand, D., Demidov, D., & Houben, A. (2003). The temporal and spatial pattern of histone H3 phosphorylation at serine 28 and serine 10 is similar in plants but differs between mono- and polycentric chromosomes. *Cytogenetic and Genome Research*, 101(2), 172–6.
- Gibcus, J. H., & Dekker, J. (2013). The Hierarchy of the 3D Genome. *Molecular Cell*, 49(5), 773–782.
- Gibcus, J. H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Dekker, J. (2018). A pathway for mitotic chromosome formation. *Science*, 6135.
- Green, L. C., Kalitsis, P., Chang, T. M., Cipetic, M., Kim, J. H., Marshall, O., Hudson, D. F. (2012). Contrasting roles of condensin I and condensin II in mitotic chromosome formation. *Journal of Cell Science*, 125(6),
- Grob, S., Schmid, M. W., Luedtke, N. W., Wicker, T., & Grossniklaus, U. (2013). Characterization of chromosomal architecture in Arabidopsis by chromosome conformation capture. *Genome Biology*, 14(11), R129.
- Guan, Y., Zhu, Q., Huang, D., Zhao, S., Jan Lo, L., & Peng, J. (2015). An equation to estimate the difference between theoretically predicted and SDS PAGE-displayed molecular weights for an acidic peptide. *Scientific Reports*, 5), 1–11.
- Harashima, H., & Sugimoto, K. (2016). Integration of developmental and environmental signals into cell proliferation and differentiation through RETINOBLASTOMA-RELATED 1. *Current Opinion in Plant Biology*, 29, 95–103.
- Hartl, T. A., Smith, H. F., & Bosco, G. (2008a). Chromosome Alignment and Transvection Are Antagonized by Condensin II. *Science*, 322, 1384–1387.
- Hartl, T. A., Sweeney, S. J., Knepler, P. J., & Bosco, G. (2008b). Condensin II resolves chromosomal associations to enable anaphase I segregation in Drosophila male meiosis. *PLoS Genetics*, 4(10), 18–22.
- Heckmann, S., Lermontova, I., Berckmans, B., De Veylder, L., Bäumllein, H., & Schubert, I. (2011). The E2F transcription factor family regulates CENH3 expression in Arabidopsis thaliana. *The Plant Journal*, 68(4), 646–56.
- Herzog, S., Jaiswal, S. N., Urban, E., Riemer, A., Fischer, S., Heidmann, S. K., Heidmann, S. K. (2013). Functional Dissection of the Drosophila melanogaster Condensin Subunit Cap-G Reveals Its Exclusive Association with Condensin I. *PLoS Genetics*, 9(4), e1003463.
- Hirano, T. (2012). Condensins: Universal organizers of chromosomes with diverse functions. *Genes and Development*, 26(15), 1659–1678.
- Hirano, T., & Mitchison, T. J. (1994). A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro. *Cell*, 79(3), 449–458.
- Hirota, T., Gerlich, D., Koch, B., Ellenberg, J., & Peters, J.-M. (2004). Distinct functions of condensin I and II in mitotic chromosome assembly. *Journal of Cell Science*, 117(26), 6435–6445.
- Hotton, S. K., & Callis, J. (2008). Regulation of cullin RING ligases. *Annual Review of Plant Biology*, 59, 467–89.
- Hu, C. D., Chinenov, Y., & Kerppola, T. K. (2002). Visualization of interactions among bZIP and Rel family

- proteins in living cells using bimolecular fluorescence complementation. *Molecular Cell*, 9(4), 789–798.
- Hudson, D. F., Ohta, S., Freisinger, T., MacIsaac, F., Sennels, L., Alves, F., Earnshaw, W. C. (2008). Molecular and Genetic Analysis of Condensin Function in Vertebrate Cells. *Molecular Biology of the Cell*, 19(2), 327–331.
- Hudson, D. F., Vagnarelli, P., Gassmann, R., & Earnshaw, W. C. (2003). Condensin is required for nonhistone protein assembly and structural integrity of vertebrate mitotic chromosomes. *Developmental Cell*, 5(2), 323–336.
- Huey, R. B., Carlson, M., Crozier, L., Frazier, M., Hamilton, H., Harley, C., Kingsolver, J. G. (2002). Plants Versus Animals: Do They Deal with Stress in Different Ways? *Integrative and Comparative Biology*, 42(3), 415–423.
- Iwasaki, O., Tanaka, A., Tanizawa, H., Grewal, S. I. S., & Noma, K.-I. (2010). Centromeric localization of dispersed Pol III genes in fission yeast. *Molecular Biology of the Cell*, 21(2), 254–65.
- Jefferson, R. A., Kavanagh, T. A., & Bevan, M. W. (1987). GUS fusions: B-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO Journal*, 6(13), 3901–3907.
- Jeppsson, K., Kanno, T., Shirahige, K., & Sjögren, C. (2014). The maintenance of chromosome structure: positioning and functioning of SMC complexes. *Nature Reviews Molecular Cell Biology*, 15(9), 601–614.
- Kagami, Y., Ono, M., & Yoshida, K. (2017). Plk1 phosphorylation of CAP-H2 triggers chromosome condensation by condensin II at the early phase of mitosis. *Scientific Reports*, 7(1), 5583.
- Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D., van Berkum, N. L., Jobdeckerumassmededu, J. D. (2010). Mediator and Cohesin Connect Gene Expression and Chromatin Architecture. *Nature*, 467(7314), 430–435.
- Kikuchi, S., Kishii, M., Shimizu, M., & Tsujimoto, H. (2005). Centromere-specific repetitive sequences from *Torenia*, a model plant for interspecific fertilization, and whole-mount FISH of its interspecific hybrid embryos. *Cytogenetic and Genome Research*, 109(1–3), 228–235.
- Kim, J. H., Zhang, T., Wong, N. C., Davidson, N., Maksimovic, J., Oshlack, A., Hudson, D. F. (2013). Condensin I associates with structural and gene regulatory regions in vertebrate chromosomes. *Nature Communications*, 4, 1–14.
- Kim, K.-D., Tanizawa, H., Iwasaki, O., & Noma, K. (2016). Transcription factors mediate condensin recruitment and global chromosomal organization in fission yeast. *Nature Genetics*, 48(10), 1242–1252.
- Kudo, T., Sasaki, Y., Terashima, S., Matsuda-Imai, N., Takano, T., Saito, M., Yano, K. (2016). Identification of reference genes for quantitative expression analysis using large-scale RNA-seq data of *Arabidopsis thaliana* and model crop plants. *Genes & Genetic Systems*, 91(2), 111–125.
- Lang, J., Smetana, O., Sanchez-Calderon, L., Lincker, F., Genestier, J., Schmit, A.-C., Chabouté, M.-E. (2012). Plant γ H2AX foci are required for proper DNA DSB repair responses and colocalize with E2F factors. *New Phytologist*, 194(2), 353–363.
- Lau, A. C., Nabeshima, K., & Csankovszki, G. (2014). The *C. elegans* dosage compensation complex mediates interphase X chromosome compaction. *Epigenetics & Chromatin*, 7(1), 31.
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragozy, T., Telling, A., Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (New York, N.Y.)*, 326(5950), 289–93.
- Liu, C., Cheng, Y.-J., Wang, J.-W., & Weigel, D. (2017). Prominent topologically associated domains differentiate global chromatin packing in rice from *Arabidopsis*. *Nature Plants*, 3(9), 742–748.
- Liu, C. M., McElver, J., Tzafrir, I., Joosen, R., Wittich, P., Patton, D., Meinke, D. (2002). Condensin and cohesin

- knockouts in *Arabidopsis* exhibit a titan seed phenotype. *The Plant Journal*, 29(4), 405–415.
- Liu, C., Wang, C., Wang, G., Becker, C., Zaidem, M., & Weigel, D. (2016). Genome-wide analysis of chromatin packing in *Arabidopsis thaliana* at single-gene resolution. *Genome Research*, 26,
- Liu, Y., Koornneef, M., & Soppe, W. J. J. (2007). The absence of histone H2B monoubiquitination in the *Arabidopsis* hub1 (rdo4) mutant reveals a role for chromatin remodeling in seed dormancy. *The Plant Cell*, 19(2), 433–44.
- Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta C_T}$ Method. *METHODS*, 25, 402–408.
- Longworth, M. S., Herr, A., Ji, J. Y., & Dyson, N. J. (2008). RBF1 promotes chromatin condensation through a conserved interaction with the Condensin II protein dCAP-D3. *Genes and Development*, 22(8), 1011–1024.
- Longworth, M. S., Walker, J. A., Anderssen, E., Moon, N. S., Gladden, A., Heck, M. M. S., Dyson, N. J. (2012). A shared role for RBF1 and dCAP-D3 in the regulation of transcription with consequences for innate immunity. *PLoS Genetics*, 8(4), e1002618.
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15.
- Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, 389(6648), 251–260.
- Marenduzzo, D., Finan, K., & Cook, P. R. (2006). The depletion attraction: An underappreciated force driving cellular organization. *Journal of Cell Biology*, 175(5), 681–686.
- Marston, A. L. (2014). Chromosome segregation in budding yeast: Sister chromatid cohesion and related mechanisms. *Genetics*, 196(1), 31–63.
- Martinez-Zapater, J. M., Estelle, M. A., & Somerville, C. R. (1986). A highly repeated DNA sequence in *Arabidopsis thaliana*. *Mol Gen Genet*, 204, 417–423.
- Miskolczi, P., Lendvai, Á., Horváth, G. V., Pettkó-Szandtner, A., & Dudits, D. (2007). Conserved functions of retinoblastoma proteins: From purple retina to green plant cells. *Plant Science*, 172(4), 671–683.
- Moissiard, G., Cokus, S. J., Cary, J., Feng, S., Billi, A. C., Stroud, H., Jacobsen, S. (2012). MORC Family ATPases Required for Heterochromatin Condensation and Gene Silencing. *Science*, 336(6087), 1448–1451.
- Nakamura, S., Mano, S., Tanaka, Y., Ohnishi, M., Nakamori, C., Araki, M., Ishiguro, S. (2010). Gateway Binary Vectors with the Bialaphos Resistance Gene, bar, as a Selection Marker for Plant Transformation. *Bioscience, Biotechnology, and Biochemistry*, 74(6), 1315–1319.
- Neuwald, A. F., & Hirano, T. (2000). HEAT Repeats Associated with Condensins, Cohesins, and Other Complexes Involved in Chromosome-Related Functions. *Genome Research*, 1445–1452.
- Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., Heard, E. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, 485(7398), 381–385.
- Onn, I., Aono, N., Hirano, M., & Hirano, T. (2007). Reconstitution and subunit geometry of human condensin complexes. *The EMBO Journal*, 26(4), 1024–1034.
- Ono, T., Fang, Y., Spector, D. L., & Hirano, T. (2004). Spatial and Temporal Regulation of Condensins I and II in Mitotic Chromosome Assembly in Human Cells. *Molecular Biology of the Cell*, 15, 3296–3308.
- Ono, T., Losada, A., Hirano, M., Myers, M. P., Neuwald, A. F., & Hirano, T. (2003). Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. *Cell*, 115(1), 109–21.
- Ono, T., Yamashita, D., & Hirano, T. (2013). Condensin II initiates sister chromatid resolution during S phase.

Journal of Cell Biology, 200(4), 429–441.

- Parisi, S., McKay, M. J., Molnar, M., Thompson, M. A., van der Spek, P. J., van Drunen-Schoenmaker, E., Kohli, J. (1999). Rec8p, a Meiotic Recombination and Sister Chromatid Cohesion Phosphoprotein of the Rad21p Family Conserved from Fission Yeast to Humans. *Molecular and Cellular Biology*, 19(5), 3515–3528.
- Parra, G., Bradnam, K., Rose, A. B., & Korf, I. (2011). Comparative and functional analysis of intron-mediated enhancement signals reveals conserved features among plants. *Nucleic Acids Research*, 39(13), 5328–5337.
- Passarge, E. (1979). Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *American Journal of Human Genetics*, 31(2), 106–115.
- Pecinka, A., Schubert, V., Meister, A., Kreth, G., Klatte, M., Lysak, M. A., Schubert, I. (2004). Chromosome territory arrangement and homologous pairing in nuclei of *Arabidopsis thaliana* are predominantly random except for NOR-bearing chromosomes. *Chromosoma*, 113(5), 258–269.
- Perrella, G., Lopez-Vernaza, M. A., Carr, C., Sani, E., Gossele, V., Verduyn, C., Amtmann, A. (2013). Histone Deacetylase Complex1 Expression Level Titrates Plant Growth and Abscisic Acid Sensitivity in *Arabidopsis*. *The Plant Cell*, 25(9), 3491–3505.
- Piazza, I., Haering, C. H., & Rutkowska, A. (2013). Condensin: Crafting the chromosome landscape. *Chromosoma*, 122(3), 175–190.
- Poulet, A., Duc, C., Voisin, M., Desset, S., Tutois, S., Vanrobays, E., Tatout, C. (2017). The LINC complex contributes to heterochromatin organisation and transcriptional gene silencing in plants. *Journal of Cell Science*, 130(3), 590–601.
- Probst, A. V., & Mittelsten Scheid, O. (2015). Stress-induced structural changes in plant chromatin. *Current Opinion in Plant Biology*, 27, 8–16.
- Proost, S., Van Bel, M., Vanechoutte, D., Van de Peer, Y., Inzé, D., Mueller-Roeber, B., & Vandepoele, K. (2015). PLAZA 3.0: an access point for plant comparative genomics. *Nucleic Acids Research*, 43(Database issue), D974–81.
- Rabl, C. (1885). Über zelltheilung. *Morph Jb*, 10, 214–330.
- Robellet, X., Fauque, L., Legros, P., Mollereau, E., Janczarski, S., Parrinello, H., Bernard, P. (2014). A Genetic Screen for Functional Partners of Condensin in Fission Yeast. *G3: Genes/Genomes/Genetics*, 4(2), 373–381.
- Rose, A. B., Elfersi, T., Parra, G., & Korf, I. (2008). Promoter-Proximal Introns in *Arabidopsis thaliana* Are Enriched in Dispersed Signals that Elevate Gene Expression. *THE PLANT CELL ONLINE*, 20(3), 543–551.
- Roudier, F., Ahmed, I., Bérard, C., Sarazin, A., Mary-Huard, T., Cortijo, S., Colot, V. (2011). Integrative epigenomic mapping defines four main chromatin states in *Arabidopsis*. *The EMBO Journal*, 30(10), 1928–1938.
- Sakamoto, T., Inui, Y. T., Uraguchi, S., Yoshizumi, T., Matsunaga, S., Mastui, M., Fujiwara, T. (2011). Condensin II Alleviates DNA Damage and Is Essential for Tolerance of Boron Overload Stress in *Arabidopsis*. *The Plant Cell*, 23(9), 3533–3546.
- Sakamoto, Y., & Takagi, S. (2013). LITTLE NUCLEI 1 and 4 Regulate Nuclear Morphology in *Arabidopsis thaliana*. *Plant and Cell Physiology*, 54(4), 622–633.
- Sanchez Moran, E., Armstrong, S. J., Santos, J. L., Franklin, F. C., & Jones, G. H. (2001). Chiasma formation in *Arabidopsis thaliana* accession Wassileskija and in two meiotic mutants. *Chromosome Research*, 9(2), 121–8.
- Santos, A. P., & Shaw, P. (2004). Interphase chromosomes and the Rab1 configuration: Does genome size

- matter? *Journal of Microscopy*, 214(2), 201–206.
- Savidou, E., Cobbe, N., Steffensen, S., Cotterill, S., & Heck, M. M. S. (2005). Drosophila CAP-D2 is required for condensin complex stability and resolution of sister chromatids. *Journal of Cell Science*, 118, 2529–2543.
- Schägger, H., & von Jagow, G. (1987). Tricine-Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis for the Separation of Proteins in the Range from 1 to 100 kDa. *Analytical Biochemistry*, 166(2), 368–379.
- Schimpl, S., Fauser, F., & Puchta, H. (2016). CRISPR/Cas-mediated site-specific mutagenesis in Arabidopsis thaliana using Cas9 nucleases and paired nickases. In *Methods in Molecular Biology* (Vol. 1469, pp. 111–122).
- Schmiesing, J., Gregson, H., Zhou, S., & Yokomori, K. (2000). A human condensin complex containing hCAP-C-hCAP-E and CNAP1, a homolog of Xenopus XCAP-D2, colocalizes with phosphorylated histone H3 during the early stage of mitotic chromosome condensation. *Molecular and Cellular Biology*, 20(18), 6996–7006.
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Meth*, 9(7), 671–675.
- Schubert, I., & Shaw, P. (2011). Organization and dynamics of plant interphase chromosomes. *Trends in Plant Science*, 16(5), 273–281.
- Schubert, V. (2009). SMC proteins and their multiple functions in higher plants. *Cytogenetic and Genome Research*, 124(3–4), 202–214.
- Schubert, V., Berr, A., & Meister, A. (2012). Interphase chromatin organisation in Arabidopsis nuclei: Constraints versus randomness. *Chromosoma*, 121(4), 369–387.
- Schubert, V., Klatt, M., Pecinka, A., Meister, A., Jasencakova, Z., & Schubert, I. (2006). Sister chromatids are often incompletely aligned in meristematic and endopolyploid interphase nuclei of Arabidopsis thaliana. *Genetics*, 172(1), 467–475.
- Schubert, V., Lermontova, I., & Schubert, I. (2013). The Arabidopsis CAP-D proteins are required for correct chromatin organisation, growth and fertility. *Chromosoma*, 122(6), 517–533.
- Schubert, V., & Weisshart, K. (2015). Abundance and distribution of RNA polymerase II in Arabidopsis interphase nuclei. *Journal of Experimental Botany*, 66(6), 1687–1698.
- Schwechheimer, C., & Isono, E. (2010). The COP9 signalosome and its role in plant development. *European Journal of Cell Biology*, 89(2–3), 157–162.
- Sebastian, J., Ravi, M., Andreuzza, S., Panoli, A. P., Marimuthu, M. P. A., & Siddiqi, I. (2009). The plant adherin AtSCC2 is required for embryogenesis and sister-chromatid cohesion during meiosis in Arabidopsis. *The Plant Journal*, 59(1), 1–13.
- Seitan, V. C., & Merckenschlager, M. (2012). Cohesin and chromatin organisation. *Current Opinion in Genetics and Development*, 22(2), 93–100.
- Sequeira-Mendes, J., Araguez, I., Peiro, R., Mendez-Giraldez, R., Zhang, X., Jacobsen, S. E., Gutierrez, C. (2014). The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. *The Plant Cell*, 26(6), 2351–2366.
- Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Cavalli, G. (2012). Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. *Cell*, 148(3), 458–472.
- Shintomi, K., Takahashi, T. S., & Hirano, T. (2015). Reconstitution of mitotic chromatids with a minimum set of purified factors. *Nature Cell Biology*, 17(8), 1014–1023.
- Siddiqi, N. U., Rusyniak, S., Hasenkampf, C. A., & Riggs, C. D. (2006). Disruption of the Arabidopsis SMC4

- gene, AtCAP-C, compromises gametogenesis and embryogenesis. *Planta*, 223(5), 990–997.
- Siddiqui, N. U., Stronghill, P. E., Dengler, R. E., Hasenkampf, C. a., & Riggs, C. D. (2003). Mutations in Arabidopsis condensin genes disrupt embryogenesis, meristem organization and segregation of homologous chromosomes during meiosis. *Development*, 130(14), 3283–3295.
- Sieburth, L. E., & Meyerowitz, E. M. (1997). Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically. *The Plant Cell*, 9(3), 355–365.
- Singh, D. K., Andreuzza, S., Panoli, A. P., & Siddiqi, I. (2013). AtCTF7 is required for establishment of sister chromatid cohesion and association of cohesin with chromatin during meiosis in Arabidopsis. *BMC Plant Biol*, 13(1), 117.
- Skylar, A., Matsuwaka, S., & Wu, X. (2013). ELONGATA3 is required for shoot meristem cell cycle progression in Arabidopsis thaliana seedlings. *Developmental Biology*, 382(2), 436–445.
- Smith, S. J., Osman, K., & Franklin, F. C. H. (2014). The condensin complexes play distinct roles to ensure normal chromosome morphogenesis during meiotic division in Arabidopsis. *The Plant Journal : For Cell and Molecular Biology*, 80(2), 255–68.
- Sofueva, S., Yaffe, E., Chan, W.-C., Georgopoulou, D., Vietri Rudan, M., Mira-Bontenbal, H., Hadjur, S. (2013). Cohesin-mediated interactions organize chromosomal domain architecture. *The EMBO Journal*, 32(24), 3119–3129.
- Soppe, W. J. J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M. S., Fransz, P. F. (2002). DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. *The EMBO Journal*, 21(23), 6549–59.
- Sparkes, I. A., 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(4), 2019–2025.
- Stratmann, J. W., & Gusmaroli, G. (2012). Many jobs for one good cop - The COP9 signalosome guards development and defense. *Plant Science*, 185–186, 50–64.
- Stroud, H., Otero, S., Desvoyes, B., Ramirez-Parra, E., Jacobsen, S. E., & Gutierrez, C. (2012). Genome-wide analysis of histone H3.1 and H3.3 variants in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America*, 109(14), 5370–5.
- Symmons, O., Uslu, V. V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. *Genome Research*, 24(3), 390–400.
- Talbert, P. B., Masuelli, R., Tyagi, A. P., Comai, L., & Henikoff, S. (2002). *Centromeric localization and adaptive evolution of an Arabidopsis histone H3 variant*. *The Plant cell* (Vol. 14). American Society of Plant Biologists.
- Tatout, C., Evans, D. E., Vanrobays, E., Probst, A. V., & Graumann, K. (2014). The plant LINC complex at the nuclear envelope. *Chromosome Research*, 22(2), 241–252.
- Tessadori, F., Chupeau, M.-C., Chupeau, Y., Knip, M., Germann, S., van Driel, R., Gaudin, V. (2007). Large-scale dissociation and sequential reassembly of pericentric heterochromatin in dedifferentiated Arabidopsis cells. *Journal of Cell Science*, 120(7), 1200–1208.
- Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., & Leunissen, J. A. M. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research*, 35.
- Van Leene, J., Eeckhout, D., Cannoot, B., De Winne, N., Persiau, G., Van De Slijke, E., De Jaeger, G. (2014). An improved toolbox to unravel the plant cellular machinery by tandem affinity purification of Arabidopsis protein complexes. *Nature Protocols*, 10(1), 169–187.
- Van Leene, J., Eeckhout, D., Persiau, G., Slijke, E. Van De, Geerinck, J., Isterdael, G. Van, De Jaeger, G. (2011).

- Isolation of Transcription Factor Complexes from Arabidopsis Cell Suspension Cultures by Tandem Affinity Purification. In L. Yuan & S. E. Perry (Eds.), *Plant transcription factors. Methods in Molecular Biology* (Vol. 754, pp. 195–218).
- van Ruiten, M. S., & Rowland, B. D. (2018). SMC Complexes: Universal DNA Looping Machines with Distinct Regulators. *Trends in Genetics*.
- van Zanten, M., Koini, M. A., Geyer, R., Liu, Y., Brambilla, V., Bartels, D., Soppe, W. J. J. (2011). Seed maturation in Arabidopsis thaliana is characterized by nuclear size reduction and increased chromatin condensation. *Proceedings of the National Academy of Sciences*, 108(50), 20219–20224.
- Verlinden, L., Eelen, G., Beullens, I., Van Camp, M., Van Hummelen, P., Engelen, K., Verstuyf, A. (2005). Characterization of the condensin component Cnap1 and protein kinase melk as novel E2F target genes down-regulated by 1,25-dihydroxyvitamin D3. *Journal of Biological Chemistry*, 280(45), 37319–37330.
- Wallace, H. A., & Bosco, G. (2013). Condensins and 3D Organization of the Interphase Nucleus. *Current Genetic Medicine Reports*, 1(4), 219–229.
- Walter, M., Chaban, C., Schütze, K., Batistic, O., Weckermann, K., Näke, C., Kudla, J. (2004). Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. *The Plant Journal*, 40(3), 428–438.
- Walther, N., Hossain, M. J., Politi, A. Z., Koch, B., Kueblbeck, M., Ødegård-Fougner, Ø., Ellenberg, J. (2018). A quantitative map of human Condensins provides new insights into mitotic chromosome architecture. *The Journal of Cell Biology*, jcb.201801048.
- Wang, C., Liu, C., Roqueiro, D., Grimm, D., Schwab, R., Becker, C., Weigel, D. (2015). Genome-wide analysis of local chromatin packing in Arabidopsis thaliana. *Genome Research*, 25(2), 246–256.
- Wang, H., Dittmer, T. a, & Richards, E. J. (2013). Arabidopsis CROWDED NUCLEI (CRWN) proteins are required for nuclear size control and heterochromatin organization. *BMC Plant Biology*, 13(1), 200.
- Wang, J., Blevins, T., Podicheti, R., Haag, J. R., Tan, E. H., Wang, F., & Pikaard, C. S. (2017). Mutation of Arabidopsis SMC4 identifies condensin as a corepressor of pericentromeric transposons and conditionally expressed genes. *Genes & Development*, 31(15), 1601–1614.
- Wang, Z., Cao, H., Chen, F., & Liu, Y. (2014). The roles of histone acetylation in seed performance and plant development. *Plant Physiology and Biochemistry*, 84, 125–133.
- Watanabe, Y., & Nurse, P. (1999). Cohesin Rec8 is required for reductional chromosome segregation at meiosis. *Nature*, 400(6743), 461–464.
- Weisshart, K., Fuchs, J., & Schubert, V. (2016). Structured Illumination Microscopy (SIM) and Photoactivated Localization Microscopy (PALM) to Analyze the Abundance and Distribution of RNA Polymerase II Molecules on Flow-Sorted Arabidopsis Nuclei. *Bio-Protocol*, 6(3).
- Wendt, K. S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Peters, J. M. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*, 451(7180), 796–801.
- Wu, T. D., & Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics (Oxford, England)*, 26(7), 873–81.
- Yelagandula, R., Stroud, H., Holec, S., Zhou, K., Feng, S., Zhong, X., Berger, F. (2014). The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in arabidopsis. *Cell*, 158(1), 98–109.
- Yilmaz, A., Mejia-Guerra, M. K., Kurz, K., Liang, X., Welch, L., & Grotewold, E. (2011). AGRIS: the Arabidopsis Gene Regulatory Information Server, an update. *Nucleic Acids Research*, 39(Database), D1118–D1122.
- Yoo, S.-D., Cho, Y.-H., & Sheen, J. (2007). Arabidopsis mesophyll protoplasts: a versatile cell system for

- transient gene expression analysis. *Nature Protocols*, 2(7), 1565–1572.
- Yuen, K. C., & Gerton, J. L. (2018). Taking cohesin and condensin in context. *PLOS Genetics*, 14(1), e1007118.
- Yuen, K. C., Slaughter, B. D., & Gerton, J. L. (2017). Condensin II is anchored by TFIIC and H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters. *Science Advances*, 3(6), e1700191.
- Zamariola, L., De Storme, N., Vannerum, K., Vandepoele, K., Armstrong, S. J., Franklin, F. C. H., & Geelen, D. (2014). SHUGOSHINS and PATRONUS protect meiotic centromere cohesion in *Arabidopsis thaliana*. *Plant Journal*, 77(5), 782–794.
- Zhang, T., Paulson, J. R., Bakhrebah, M., Kim, J. H., Nowell, C., Kalitsis, P., & Hudson, D. F. (2016). Condensin I and II behaviour in interphase nuclei and cells undergoing premature chromosome condensation. *Chromosome Research*, 24(2), 243–269.
- Zhang, X., Bernatavichute, Y. V., Cokus, S., Pellegrini, M., & Jacobsen, S. E. (2009). Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biology*, 10(6), 1–14.
- Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y. V., Pellegrini, M., Goodrich, J., & Jacobsen, S. E. (2007). Whole-Genome Analysis of Histone H3 Lysine 27 Trimethylation in *Arabidopsis*. *PLoS Biology*, 5(5), e129.

10. Curriculum vitae

PERSONAL DETAILS

Name Celia Muncio Diaz
Nationality Spanish
Date of birth June 30 1987
Present address 188 Avenue Henri Barbusse, 92140 Clamart, France

CONTACT INFORMATION

Telephone: +33 783542442
E-mail munciadc@gmail.com

EDUCATION

Nov 2013-Aug 2017 **Leibniz Institut of Plant Genetics and Crop Plant Research (IPK)** Gatersleben, Germany.
PhD Student in the group Chromosome Structure and Function.
Supervisors: Prof. Andreas Houben and Prof. Veit Schubert

Jan 2012-Sep 2012 **Complutense University of Madrid**, Madrid, Spain.
Master thesis "The role of CAND1 in crossover formation in the meiosis of *Arabidopsis thaliana*". Supervisor: Monica Pradillo Orellana and Juan Luis Santos Coloma.

Oct 2011- Sep 2012 **Autonomous University of Madrid**, Madrid, Spain.
Master in Cell Biology and Genetics.

Sep 2005- Jun 2010 **University of Salamanca**, Salamanca, Spain
Bachelor's degree in Biology.

STAYS IN ANOTHER RESEARCH CENTERS AND OTHER RESEARCH EXPERIENCE

October 6-14 2016 **VIB Ghent**, Belgium.
Research groups of Mieke Van Lijsebettens and Frederik Coppens.

April 15-30 2015 - January 19-30 2015 **University of Regensburg**. Regensburg, Germany
Plant chromatin Group. Supervisor: Prof. Dr. Klaus Grasser.

Jan 2011-Aug 2012 **Matís ltd. Icelandic Food and Biotech R&D**. Reykjavík, Iceland.
Research student. Project "Cloning and expression of a *Rodothermus marinus* chitinase and *Arthobacter* condroitinase in *Pichia pastoris*"
Supervisor: Ólafur H. Friðjónsson and Guðmundur Óli Hreggviðsson.

Jun 2010-Jul 2010 **Institute of Microbiology and Biochemistry (IMB)**. Salamanca, Spain.
Undergraduate student in Cell Wall and Morphogenesis in Yeast Group. Supervisor: César Roncero Maillo.

- Jan 2009-May 2009** **University of Iceland. Faculty of life and environmental sciences.**
Reykjavík, Iceland.
Undergraduate student. Project “Diversity of *Nostoc* cyanobionts from the lichen *Peltigera membranacea*”. Supervisor: Prof. Ólafur S. Andr sson.
- Sep 2008-Dec 2008** **University of Iceland. Faculty of life and environmental sciences.**
Reykjavík, Iceland.
Undergraduate student. Research project “*Vibrio cholerae* at geothermal locations along the Icelandic coast”. Supervisor: Eva Benediktsd ttir.

ORAL PRESENTATIONS

Celia Muncio Characterizing the *Arabidopsis* CAP-D2 and CAP-D3 condensin subunits. Workshop in the Centre of Plant Structural and Functional Genomics (30-31 January 2017) Olomouc, Czech Republic.

Celia Muncio Condensin dynamics in *Arabidopsis* interphase nuclei. European Workshop Series on Plant Chromatin (25-26 June 2015), Uppsala (Sweden)

POSTER PRESENTATIONS

Muncio C., Antosz W., Houben A., Grasser K., Schubert V. Characterization of *Arabidopsis* CAP-D2 and CAP-D3 condensin subunits. 21st International Chromosome Conference (10-13 July 2016) Foz do Iguazu, Brazil.

Muncio C., Antosz W., Houben A., Grasser K., Schubert V. Characterization of *Arabidopsis* CAP-D2 and CAP-D3 condensin subunits. Plant Science Student Conference (4-7 July 2016) IPK, Gatersleben, Germany

Muncio C., Antosz W., Houben A., Grasser K., Schubert V. Characterizing the *Arabidopsis* CAP-D2 and CAP-D3 condensin subunits. Epigenetic and chromatin regulation of plant traits Conference (14-15 January 2016) Strasbourg, France.

Muncio C., Antosz W., Houben A., Grasser K., Schubert V. Functions of the condensin subunits CAP-D2 and CAP-D3 in *Arabidopsis thaliana*. Plant Science Student Conference (2-5 June 2015) Leibniz-Institute of Plant Biochemistry, Halle, Germany.

Muncio C., Houben A., Schubert V. Influence of condensin regulatory elements on *Arabidopsis* interphase nuclei. Plant Molecular Cytogenetics in Genomic and Postgenomic Era (23-24 September 2014) University of Silesia, Katowice, Poland.

Muncio C., Houben A., Schubert V. Condensin dynamics in *Arabidopsis* interphase nuclei. Plant Science Student Conference (2-5 June 2014) IPK, Gatersleben, Germany.

CONFERENCES ATTENDED

- 10-13 Jul 2016** “21st International Chromosome Conference” Foz do Iguazu, Brazil.
- 4-7 Jul 2016** “Plant Science Student Conference” IPK, Gatersleben, Germany.
- 14-15 Jan 2016** “Epigenetic and Chromatin Regulation of Plant Traits” Strasbourg, France.
- 25-26 Jun 2015** “4th European Workshop Series on Plant Chromatin” Uppsala, Sweden.

- 2-5 Jun 2015** **“Plant Science Student Conference”** Leibniz-Institute of Plant Biochemistry (IPB), Halle, (Germany).
- 23-24 Sept 2014** **“Plant Molecular Cytogenetics in Genomic and Postgenomic Era”** University of Silesia, Katowice, Poland.
- 2-5 Jun 2014** **“Plant Science Student Conference”** IPK, Gatersleben, Germany.
- 29-30 Aug 2013** **“3rd European Workshop Series on Plant Chromatin”** Madrid, Spain.

PUBLICACION LIST

Badaeva E, Ruban A S, Aliyeva-Schnorr L, Municio C, Hesse S, Houben A. In situ hybridisation to plant chromosomes. In: Liehr T (Ed.): Fluorescence In Situ Hybridization (FISH): Application Guide, 2nd edition. : Springer (2016) DOI 10.1007/978-3-662-52959-1

Haley BJ, Chen A, Grim CJ, Clark P, Díaz CM, Taviani E, Hasan NA, Sancomb E, Elnemr WM, Islam MA, Huq A, Colwell RR & Benediktsdóttir E. (2012). *Vibrio cholerae* in an Historically Cholera-Free Country. *Environ Microbiol Rep.* 1;4(4):381-389. DOI: 10.1111/j.1758-2229.2012.00332.x

Celia Municio Diaz

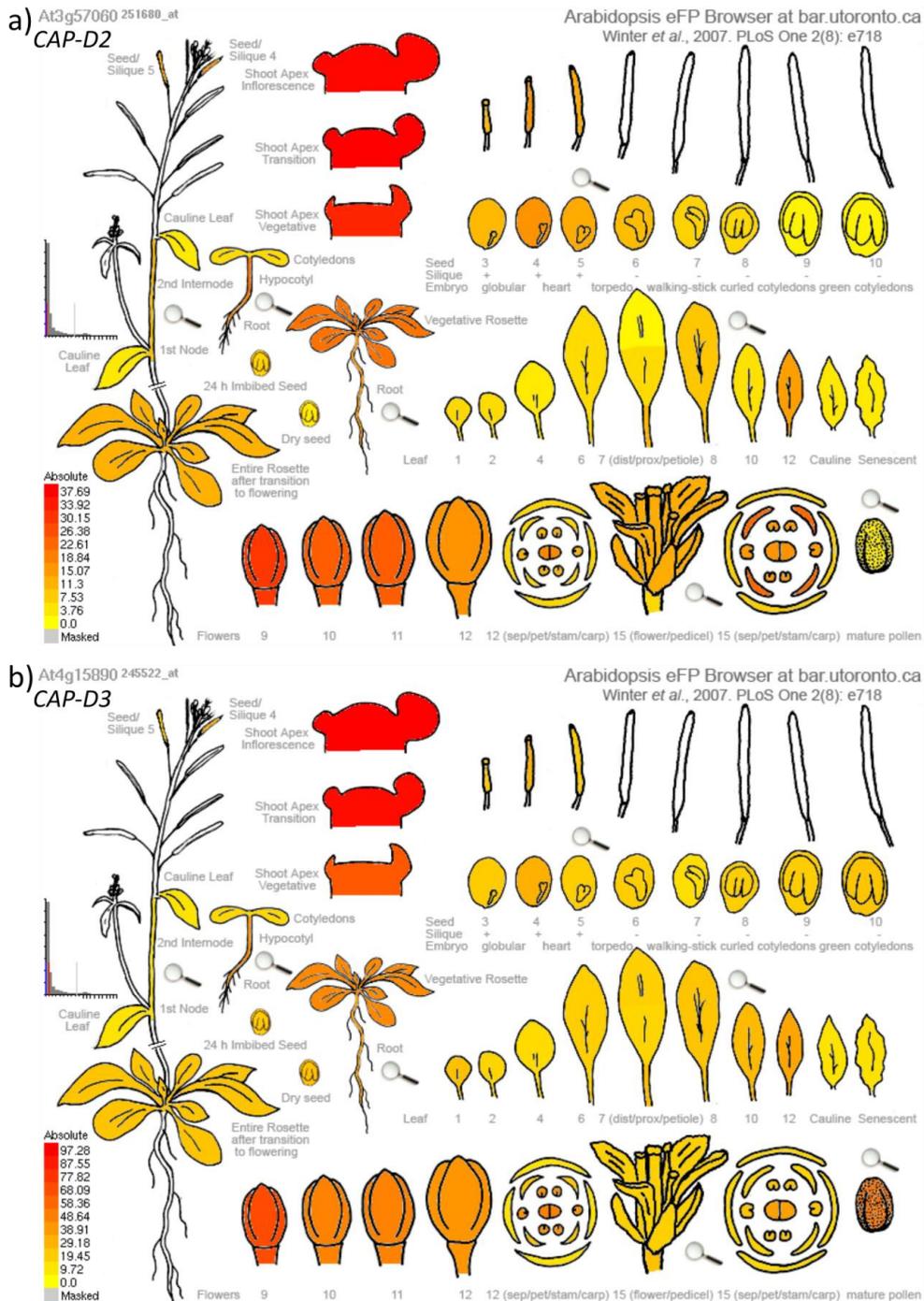
11. Eidesstattliche Erklärung / *Declaration under Oath*

Ich erkläre an Eides statt, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

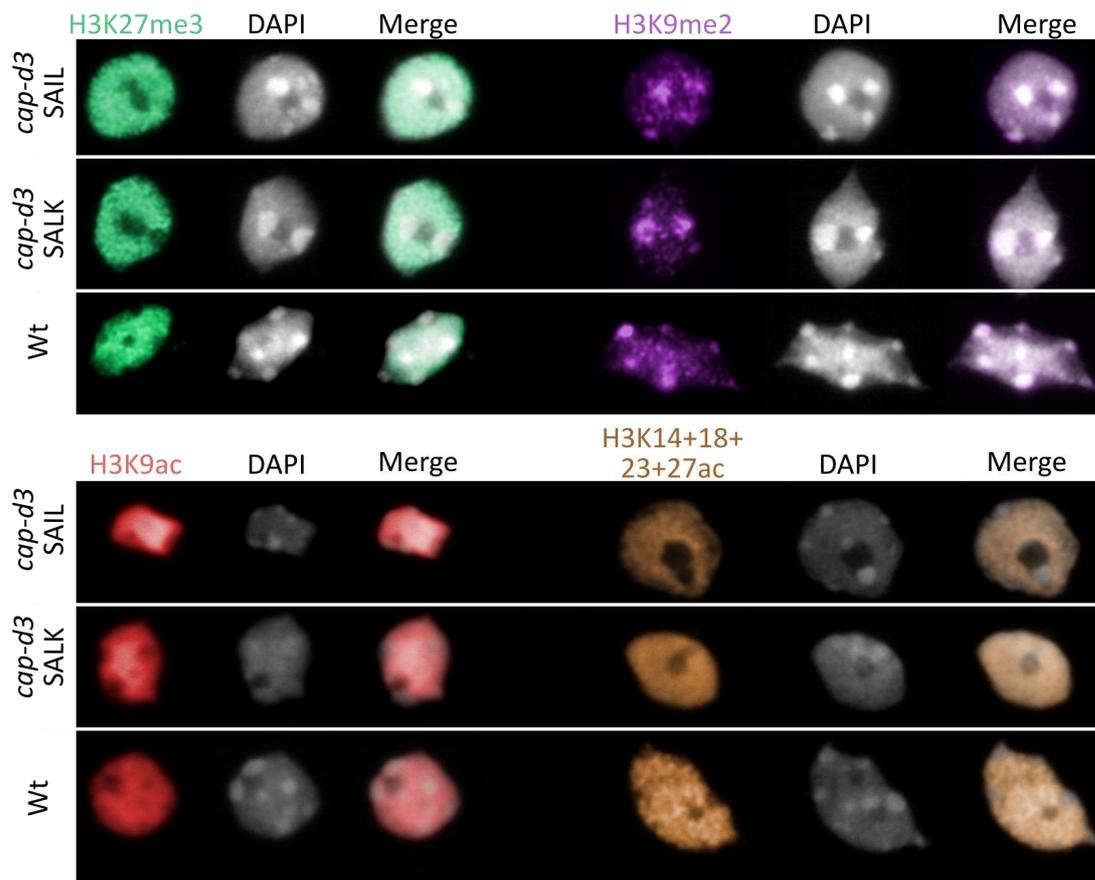
I declare under penalty of perjury that this thesis is my own work entirely and has been written without any help from other people. I used only the sources mentioned and included all the citations correctly both in word or content.

Celia Municio Diaz

12. Appendix



Appendix Fig. 1. Expression of CAP-D2 and CAP-D3 in Arabidopsis. *In silico* analysis performed with Arabidopsis eFP Browser . a) Expression of CAP-D2 in different organs and developmental stages. b) Expression of CAP-D3 in different organs and developmental stages



Appendix Fig. 2. Immunolocalization of histone modifications in *cap-d3* mutants and wild-type plants. Immunolocalization in 4C nuclei of *cap-d3* SAIL, *cap-d3* SALK and wild-type (Wt) with antibodies against histone H3K27me3 (euchromatic); H3K9me2 (heterochromatic); H3K9ac and with an antibody that recognizes H3K14+18+23+27ac.

F4J246|ARABIDOPSIS ---MAPPFVFPQILRALEEDPEDNHRIFAQNPVDVTSLRPSDLEEFVKGVSFDSLSDREL 56
Q15021|HUMAN -MAPQMYEFHPLPSPEELLKSGGVN-QYVVQEVLSI-KHLPPQLRAFQAA----FRAQGP 53
Q8K2Z4|MOUSE -MSPHNFEHPLPSPEELLKSGGVN-QYVVREVLPV-KHLSSQLRAFQSA----FRAQGP 53
Q9YHY6|XENOPUS ---MTFHFIPLAFRDLKSGGIG-QYVVQEVLPV-RHVDAQFAAFQTS----FRTEAP 50
Q9VAJ1|DROSOPHILA MEESHDFQFVLPNADSLINSSG-D-QYVKEIFGA-QEIPAKLQECKRK----VHLGDP 53
A0A1D5PW66|CHICKEN -MAAAQCEFQLPLAADLLRGLDSSR-HYVVQEVLSV-RELPVAVAAFRAA----FRERGA 53
* : * * * . . : . :

F4J246|ARABIDOPSIS FCVEDQDVFDVRSVLRVFFSLPPSCKCNLVEESLRNSLVLLPNVDSISRVSVDQ-QEDDV 115
Q15021|HUMAN LAM--LQHFDTIYSLHHRSIDPGLKEDT-----LQFLIKVVSRRHSQELPA-ILDDT 103
Q8K2Z4|MOUSE LAI--LEHFDTVYSILHHRSIDPGLKEDT-----LEFLKVVSRHSQELSS-ILDDA 103
Q9YHY6|XENOPUS LCI--LQHFDTIYSLHHRSLDIAIKEDV-----LEVMMKVASRHANELPA-ILEDL 100
Q9VAJ1|DROSOPHILA FYI--FEHFDLYYSIEARGSDGASAQNLMR-----SF'DLLYLTVKELFQDLQPLLTASE 106
A0A1D5PW66|CHICKEN LAV--LQHFDTIYSLRHRFRALGTAAKEDA-----LELMMQVVTRHSNELST-ILNDS 103
: : : * * * : : : : : : : : : :

F4J246|ARABIDOPSIS P-IIDRITSHRNALKIYTFLLTVVMNEESHIS--SVETTKVAARGR-----KQIIQS 166
Q15021|HUMAN TLSGSDRNAHLNALKMNCYALIRLLESFETMASQTNLVDLIDLGKGGKARTKAAH----G 159
Q8K2Z4|MOUSE ALSGSDRAHLNALKMNCYALIRLLESFENMTSQTSLIDLDIGGKGRARAKATL----G 159
Q9YHY6|XENOPUS NLSVQRAAHLNALKMNCFIITQLIEAFAETYKASLGSVPSGKGGKAKSK-PE----G 155
Q9VAJ1|DROSOPHILA PMSNQQRNSYLNLTGMTLFLQVSTVKKINNSVQQ-AMRDQQLNV--QKRAKPSGLEQF 163
A0A1D5PW66|CHICKEN GLSHTDRDAHLNALKMNCYLLSGLLEAFEMAFKSGLVEVDPAGNKKSRTK-AS----G 158
: : * * : : : : : : :

F4J246|ARABIDOPSIS WNWEPRQGRMLNLIANSLEINLSLLFGSSDLDENYLSFIVKNSFTLFENATILK--DAET 224
Q15021|HUMAN FDWEERQPIQLQLLQQLDIRHLWNHSIIEEFVSLVTGCCYRLENPTINHQRNPT 219
Q8K2Z4|MOUSE FDWEERQPIQLQLLQQLDIRHLWNHSAIEEFVSLVTGCCYRLENPTISHQRNPT 219
Q9YHY6|XENOPUS FSWESERESILQALTHLQLDIRLWSMVVEEFVSMMTSCCYKMMENPNIVMAKNKST 215
Q9VAJ1|DROSOPHILA PNWEVKRKFVQLFNVLCPLEKLSPPVAEEDFINLLCDPCYRTIELPLRMD-NKHV 222
A0A1D5PW66|CHICKEN FSWEDEREPLRLQLLQQLDLRQLWSGLVVEEELVSLLTGCCYRILENSNIGLQRYRAT 218
.* * * . * : : * : : * : : * : : : : . : * : .

F4J246|ARABIDOPSIS KDALCRIIGASATKYHYIVQSCASIMHLIHKYDFAVVHIADAVARAESKYSOGLTAVTII 284
Q15021|HUMAN REAITHLLGVALTRYNHMLSATVKI IQMLQHFHEHLAPVLAASVSWATDYGMKSI VGEIV 279
Q8K2Z4|MOUSE KEAIAHLLGVALVRYNHMLSATVKI IQMLQHFHEHLPPVLTAVSLWATDYGMKSI VGEIV 279
Q9YHY6|XENOPUS REALGHLLGVTVKRYNHMLSASVKVIQLLQHFHEHLASVLTAVSLWATEYGMKPIVGEIM 275
Q9VAJ1|DROSOPHILA FDTI FQILGTSIKRFNQAMTFFVRILQILRGTEHAAHSVAAGILLHHEEYGISSVFSILI 282
A0A1D5PW66|CHICKEN REAAATHLLATALTHYDHMFSATLKITQMLQHFHEHVAVPFAQAVTLWAKEYGLKSI VGEILL 278
: : : : : : : : : : : : : : : . * : : : .

F4J246|ARABIDOPSIS RDIGRTDPKAYVKDTAGADNVGRFLVELADRLPKLMSTNVGVLV-PHFGGESYKIRNALV 343
Q15021|HUMAN REIGQKCPQELSRDPSGFKGFAAFTELAERVPAIILMSMCILL-DHLDGENYMMRNVAVL 338
Q8K2Z4|MOUSE REIGQKCPQELSRDTAGAKGFAAFTELAERI PAVLMANMCILL-DHLDGENYMMRNVAVL 338
Q9YHY6|XENOPUS REIGQKCSQDLSRESGSKAFATFTELAERI PAIMMPSISVLL-DYLDGENYMMRNSV 334
Q9VAJ1|DROSOPHILA KSIVDALRMD-SSDSVSKHFSNLAEFNSIAPSLIVPHLEKLAEDLLDCQSHTLRNCVL 341
A0A1D5PW66|CHICKEN REIGQKCPQELARDTSGIKGYAVFITELAEQIPALVLSNISVLL-RHLDGENYMMRNAIL 337
: * : : : : : * : : : * : : : * : : : * : : : * : : : * : : : *

F4J246|ARABIDOPSIS GVLGKLVAKAFNDVEGDMSSKSLRLRTKQAMLEILLERCRDVSAYTRSRVLQVWAELECE 403
Q15021|HUMAN AAMAEMVLQVLSGDQLEAAAR----DTRDQFLDILQAHGHVNSFVRSRVLQFLTRIVQQ 394
Q8K2Z4|MOUSE AAIAEMVLQVLSGDQLEESAR----ETRDQFLDILQAHGHVNSFVRSRVLQFLTRIVQQ 394
Q9YHY6|XENOPUS TVMGEMVVRVLSGDQLEAEAK----SSRDQFLDILQEHLDVNTYVRSQVCIQIYNRIVQE 390
Q9VAJ1|DROSOPHILA QIIGDVTVSELTSSEDLSEELK----EVRNEFLEHMAHILDISAHVRSKVLSTWHLKLTQ 397
A0A1D5PW66|CHICKEN SAMAEVLLQVLSGDQLEAAAR----GTRDNFLKTLQAHICDVNGFVRSRVLQFLTRIVQQ 393
: : : : : : : : : : : * : * : * : * : * : * : * : * : * : * : *

F4J246|ARABIDOPSIS HVSISGLWNEVA **SLSAGRL**EDKSAIVRKSALNLLIMLQHNPFQPLRIASFTEATLEQYK 463
Q15021|HUMAN KALPLTRFQAVVALAVGRADKSVLVCKNAIQLLASFLANNPFSCKLSADLAGPLQKET 454
Q8K2Z4|MOUSE KALPLTRFQAVVALAVGRADKSVLVCKNAIQLLASFLANNPFSCKLSADLAGPLQKET 454
Q9YHY6|XENOPUS KALPLSRFQSVVTLVGRFLDKSVNVCKNAIQLLASFLANNPFSCKLSADLAGPLQKET 450
Q9VAJ1|DROSOPHILA HAIPLNFLTRVLEEAIGRLEDKSSLVRRAMHLIKSALESNPYSSKLSIDELRAKHEHEV 457
A0A1D5PW66|CHICKEN KVLPLTQFLSVVSLAVGRADKSVVVKNAIQLLAAFLSNPNPFSCKLSADLAGPLQKET 453
: : : * * * * * : : : : * * * : * : : * : : : *

F4J246|ARABIDOPSIS RKLNELEPTEHASKESTS---DGE-----SCNGDGE----- 491
Q15021|HUMAN QKLQEMRAQRRATAASAVLDPEEWEAMLPELKSTLQQLLQLPQEGEEIPEQIANTETTE 514
Q8K2Z4|MOUSE QKLQEMRAQRRATAAALDPEEWDAMPELKSTLQQLLQLPQEGEG--DHQIADAETA 512
Q9YHY6|XENOPUS KKLKEMREKYQGPVVISPEEWEAMLPEVLEAFKILQESKEEEDI--ETEEIESSQ 508
Q9VAJ1|DROSOPHILA QAMEKLNVELEERK-QEEKLNDEFSSLAPELLPFIEENLTE----FPDMQFDKEESDE 511
A0A1D5PW66|CHICKEN QKLQEMKDRCREA--APTITPEEWEAMLPEVTAARQIILQPLQDEDE-DEEVLEVEETA 510
: : : . . : *

F4J246|ARABIDOPSIS ----IDDLHLET-----TTKIHQDLSLSDSCQENGEEISE----- 522
Q15021|HUMAN DVKGRYQLLAKASYKAIITREATGHFQSEPFSSHIDPEESEETRLNLI LGLIFKGA 574
Q8K2Z4|MOUSE EVKGRIRQLLAKASYKQAVLITREATSHFQSEPFSSHTEPEE--NSFLNLLGLIFKGA 569
Q9YHY6|XENOPUS HLREQILILLRRTSYKNSIRLTQKGIERFQEDPLFSEGDSEA---KSELGLEKIFTEKK 566
Q9VAJ1|DROSOPHILA TLMERITPLMRKKNYKDVIVLVRKVDVFLAGNQNMSLLKHEE-HCVVVLALLK-TYH--- 565
A0A1D5PW66|CHICKEN GTSEQITGMLRKNYKNAVRLTQKALCRFQGEPPFSSKKEE-NEEATILGILKRLTYGFC 569
* : : : *

F4J246|ARABIDOPSIS -----KDVSVPI-----GNVEQTKALIASLEAGLRFKSCM 553
Q15021|HUMAN ASTQEKNPRESTGMVVTGQTVCKNKNPMSDPEESRGNDELVKQEMLVQYLQDAYSFSRKI 634

Appendix Table 1. Primers used for genotyping the T-DNA insertion lines.

| T-DNA line | Primer name | Sequence 5'-3' |
|-------------------|--------------------|---------------------------|
| csn3-1 | SALK_000593_LP | TTGAGAGGAGCTCAAGACTCG |
| | SALK_000593_RP | CTTACCTGCCATTGTGAATG |
| csn3-2 | SALK_106465_LP | TGCTTTGACTCTGTGATCG |
| | SALK_106465_RP | AAGAACATCTGGATGCAATGC |
| csn4-1 | SALK_043720_LP | TTCGATCAATTTTCGATTGC |
| | SALK_043720_RP | AGATTTACATACCAAAGCGG |
| csn4-2 | SALK_053839_LP | CTAAGGCCACATCAGGTATGG |
| | SALK_053839_RP | CGGCTTGGATAAAAGATTGTG |
| csn5b-1 | SALK_007134_LP | AATCCCCGAAGTAACATTTTTG |
| | SALK_007134_RP | CATTACCCAGCAGTGGAGAAG |
| csn5b-3 | SALK_036658_LP | TCATGGCTATTTTCATTTGC |
| | SALK_036658_RP | CATCCTGAGCATTAACCCTTG |
| csn5b-2 | SALK_030493_LP | TCTCGAAGCACGAAAATTCAC |
| | SALK_030493_RP | TTTCATACGGAATCGATGGAG |
| hub 1-4 | SALK_122512_LP | TGGGAAAACATGGTATTGAGG |
| | SALK_122512_RP | AGCTCCGACAAGAACTCAGTG |
| hub 1-5 | SALK_044415_LP | TTTTCTGTTTCAGGGATGTCC |
| | SALK_044415_RP | TTGGCTATTTCCATTTCTCC |
| hub1_760 | SALK_037760_LP | AGAACACTCAACCTGCTGAGC |
| | SALK_037760_RP | ACTCGCAACAAAATCGATTG |
| hub1_867 | SALK_119867_LP | GAAGCTCGAATACAAATCCCC |
| | SALK_119867_RP | AAGGGTGACATTCAAGCATTG |
| hub 1-3 | GABI_276D08_LP | CGTCTTTCGAGAAACATCACC |
| | GABI_276D08_RP | TGGGTGTACAAAACCTTCTG |
| | MV06 | ATGGTAGAGCCAGCTGTTAAGAAG |
| hub 2-2 | SALK_071289_LP | CATGGTACCACATCCAAGGTC |
| | SALK_071289_RP | CCTCTTTAGGCCGATCAAAC |
| hub 2-1 | GABI_634H04_LP | CCGTTTTGTGCTTTTCTTGTC |
| | GABI_634H04_RP | TTGGTTCTGTGTCTGCATGTC |
| | MV07 | AGGATGATTGTACCTCCATTTTCAG |
| rbr_029 | SALK_096029_LP | AAACTAGGTTGTGCCATGGTG |
| | SALK_096029_RP | CCTCAAAGGACCTATCTGCC |
| rbr_478 | SALK_071478_LP | ATTTATTCCTGCCGTAAGCC |
| | SALK_071478_RP | CATAAAGTGTGTTTTGCGGC |
| rbr1-2 | SALK_002946_LP | CCATATAGGACAAAATGGGGG |
| | SALK_002946_RP | TACAGGACCTAGCTCCACCAG |
| cap-d2-3 | SALK_065716_LP | TTGAAGTTTTACATACCGCC |
| | SALK_065716_RP | TCTGCATCCTCATCAATCTCC |
| cap-d2-2 | SALK_044796_LP | GTAGAGGTTCCACTTTGCC |
| | SALK_044796_RP | AACATAACCCCTTTGGTCCAC |

| | | |
|-------------------------|-------------|----------------------------------|
| cap-d2-1 | SALK_077796 | AAAACCAAGACCATGGAATCC |
| | SALK_077796 | CAAATTGTCATAACTAAACCGGC |
| cap-d3 SALK | SALK_094776 | TGGTTTGAAAATGGTTGCTTC |
| | SALK_094776 | AGCGATAGAAGGAATCGAAGG |
| cap-d3 SAIL | SAIL_826B06 | TGAAGAAGGTGGATTTGATGC |
| | SAIL_826B06 | CGGAAATAGCTGAAACTGCAG |
| T-DNA primer SAIL lines | SAIL_LB3 | AGCATCTGAATTCATAACCAATCTCGATACAC |
| T-DNA primer SALK lines | LBb1.3 | ATTTTGCCGATTTTCGGAAC |
| T-DNA primer GABI lines | GABI 8474 | ATAATAACGCTGCGGACATCTACATTTT |

Appendix Table 2. List of constructs generated

| Name | Plasmid | Gene inserted |
|-----------------------|----------------|-------------------------------------|
| pCambia2300_CAP-D3_GS | pCambia 2300 | CAP-D3 cDNA full lenght |
| CAP-D3_pEntry | pEntry 1A | CAP-D3 cDNA full lenght |
| CAP-D3_EYFPc | pGWB641 | CAP-D3 cDNA full lenght (pEntry D3) |
| CAP-D3_EYFPn | pGWB642 | CAP-D3 cDNA full lenght (pEntry D3) |
| pEt23_CAP-D2_Ct | pET-23a (+) | CAP-D2 cDNA C-terminal 500 aa |
| pCambia2300_CAP-D3_GS | pCambia 2300 | CAP-D2 cDNA full lenght |
| CAP-D2_pEntry | pEntry 1A | CAP-D2 cDNA full lenght |
| CAP-D2_EYFPc | pGWB641 | CAP-D2 cDNA full lenght (pEntry D2) |
| CAP-D2_EYFPn | pGWB642 | CAP-D2 cDNA full lenght (pEntry D2) |
| CAP-D2_SPYNE | 35S SPYNE | CAP-D2 cDNA full lenght (pEntry D2) |
| CAP-D2_SPYCE | 35S SPYCE | CAP-D2 cDNA full lenght (pEntry D2) |
| pEnChi Sg1 | pEn-Chimera | Sg RNA 1 |
| pDeCas Sg1 | pDe-Cas9 | Sg RNA 1 |
| pEnChi Sg2 | pEn-Chimera | Sg RNA 2 |
| pDeCas Sg2 | pDe-Cas9 | Sg RNA 2 |
| pEnChi Sg3 | pEn-Chimera | Sg RNA 3 |
| pDeCas Sg3 | pDe-Cas9 | Sg RNA 3 |
| pEnChi Sg4 | pEn-Chimera | Sg RNA 4 |
| pDeCas Sg4 | pDe-Cas9 | Sg RNA 4 |
| pEntry_Pro4_D2 | pEntry 1A | Pro4_D2 |
| pEntry_Pro5_D2 | pEntry 1A | Pro5_D2 |
| pEntry_Pro6_D2 | pEntry 1A | Pro6_D2 |
| pEntry_Pro7_D2 | pEntry 1A | Pro7_D2 |
| pEntry_Pro8_D2 | pEntry 1A | Pro8_D2 |
| pEntry_Pro9_D2 | pEntry 1A | Pro9_D2 |
| pEntry_Pro10_D3 | pEntry 1A | Pro10_D3 |
| pEntry_Pro11_D3 | pEntry 1A | Pro11_D3 |
| 633_Pro4_D2 | pGWB633 | pEntry_Pro4_D2 |
| 633_Pro5_D2 | pGWB633 | pEntry_Pro5_D2 |
| 633_Pro6_D2 | pGWB633 | pEntry_Pro6_D2 |
| 633_Pro7_D2 | pGWB633 | pEntry_Pro7_D2 |
| 633_Pro8_D2 | pGWB633 | pEntry_Pro8_D2 |
| 633_Pro9_D2 | pGWB633 | pEntry_Pro9_D2 |
| 633_Pro10_D3 | pGWB633 | pEntry_Pro10_D3 |
| 633_Pro11_D3 | pGWB633 | pEntry_Pro11_D3 |
| CAP-G_pEntry | pEntry 1A | CAP-G cDNA full length |
| CAP-G_EYFPc | pGWB641 | CAP-G cDNA full length (pEntry G) |
| CAP-G_SYPNE | 35S SPYNE | CAP-G cDNA full length (pEntry G) |
| CAP-G_SPYCE | 35S SPYCE | CAP-G cDNA full length (pEntry G) |
| CAP-H_pEntry | pEntry 1A | CAP-H cDNA full length |

| | | |
|---------------|-----------|-----------------------------------|
| CAP-H_EYFPc | pGWB641 | CAP-H cDNA full length (pEntry H) |
| CAP-H_SPYNE | 35S SPYNE | CAP-H cDNA full length (pEntry H) |
| CAP-H_SPYCE | 35S SPYCE | CAP-H cDNA full length (pEntry H) |
| Control_EYFPc | pGWB641 | pEntry C (empty) |

Appendix Table 3. List of proteins co-purified with CAP-D2-GS. Only proteins present in the three affinity purifications and with no occurrence among the non-specific proteins are listed.

| AGI | Protein names | Average MASCOT Score |
|-----------|---|----------------------|
| AT3G57060 | CAP-D2 | 4538.87 |
| AT5G48600 | SMC4 | 2542.57 |
| AT5G62410 | SMC2A | 2497.43 |
| AT3G47460 | SMC2B | 1932.23 |
| AT5G37630 | CAP-G | 1421.20 |
| AT2G32590 | CAP-H | 919.63 |
| AT3G46740 | protein TOC75-3 | 641.57 |
| AT3G08943 | armadillo/beta-catenin-like repeat-containing protein | 551.70 |
| AT1G72560 | PAUSED | 476.33 |
| AT5G09840 | putative endonuclease or glycosyl hydrolase | 470.07 |
| AT2G20800 | NAD(P)H dehydrogenase B4 | 445.47 |
| AT2G02090 | CHR19 Chromatin remodelling 19 | 400.83 |
| AT5G64270 | putative splicing factor | 393.70 |
| AT4G24840 | Brefeldin A-sensitive Golgi protein-like | 389.23 |
| AT1G07810 | Ca ²⁺ -transporting ATPase | 386.87 |
| AT4G01100 | adenine nucleotide transporter 1 | 380.70 |
| AT3G60860 | guanine nucleotide exchange factor | 374.00 |
| AT4G19490 | protein VPS54 | 360.83 |
| AT3G54110 | uncoupling mitochondrial protein 1 | 352.93 |
| AT2G40730 | SCY1-like protein | 352.83 |
| AT4G02510 | translocase of chloroplast 159 | 341.17 |
| AT3G01280 | mitochondrial outer membrane protein porin 1 | 340.97 |
| AT4G05020 | NAD(P)H dehydrogenase B2 | 334.20 |
| AT5G41950 | tetratricopeptide repeat domain-containing protein | 332.53 |
| AT5G16930 | AAA-type ATPase family protein | 329.23 |
| AT4G01400 | oligomeric golgi complex subunit-like protein | 318.60 |
| AT4G02570 | CUL1 cullin 1 | 315.07 |
| AT5G16210 | HEAT repeat-containing protein | 301.20 |
| AT2G39260 | regulator of nonsense transcripts UPF2 | 269.63 |
| AT5G22770 | AP-2 complex subunit alpha-1 | 265.63 |
| AT5G18420 | CCR4-NOT transcription complex subunit | 265.13 |
| AT2G01690 | ARM repeat superfamily protein | 256.73 |
| AT3G45190 | SIT4 phosphatase-associated family protein | 253.40 |
| AT4G33650 | dynamamin like protein 2a | 250.40 |
| AT4G02350 | exocyst complex component sec15B | 247.90 |
| AT2G36200 | kinesin family member 11 | 246.07 |
| AT5G26760 | RPAP2 IYO MATE (RIMA) | 243.80 |
| AT1G04080 | pre-mRNA-processing factor 39 | 242.80 |
| AT3G62360 | carbohydrate-binding-like fold-containing protein | 241.90 |
| AT1G60200 | RNA-Binding protein 25 | 236.77 |

| | | |
|-----------|--|--------|
| AT4G32050 | neurochondrin family protein | 234.23 |
| AT1G22730 | putative topoisomerase | 226.80 |
| AT5G13850 | nascent polypeptide-associated complex subunit alpha-like protein 3 | 226.57 |
| AT3G55410 | 2-oxoglutarate dehydrogenase, E1 subunit-like protein | 218.97 |
| AT2G18330 | AAA-type ATPase-like protein | 218.27 |
| AT3G16830 | Topless-related 2 protein | 218.23 |
| AT2G22300 | calmodulin-binding transcription activator 3 | 213.63 |
| AT2G14120 | dynamamin-like protein | 204.30 |
| AT4G01990 | tetratricopeptide repeat-like superfamily protein | 203.83 |
| AT3G11710 | lysyl-tRNA synthetase | 202.50 |
| AT5G49830 | exocyst complex component 84B | 197.60 |
| AT2G27170 | SMC3 | 196.17 |
| AT1G48900 | signal recognition particle subunit SRP54 | 195.77 |
| AT1G60070 | adaptor protein complex AP-1, gamma subunit | 194.00 |
| AT2G27900 | coiled-coil protein | 193.10 |
| AT1G63810 | nucleolar protein | 192.43 |
| AT4G21150 | ribophorin II (RPN2) family protein | 187.57 |
| AT2G31810 | ACT domain-containing small subunit of acetolactate synthase protein | 186.10 |
| AT4G27500 | proton pump interactor 1 | 186.07 |
| AT1G73430 | putative conserved oligomeric golgi complex 3 | 182.73 |
| AT1G71270 | A. thaliana VPS52 homolog | 178.03 |
| AT5G19760 | mitochondrial substrate carrier family protein | 176.33 |
| AT3G54540 | ABC transporter F family member 4 | 175.57 |
| AT5G13110 | glucose-6-phosphate 1-dehydrogenase | 171.50 |
| AT5G08550 | GC-rich sequence DNA-binding factor | 167.73 |
| AT5G47480 | RGPR-related protein | 167.73 |
| AT5G65460 | kinesin like protein for actin based chloroplast movement 2 | 167.23 |
| AT5G14580 | polyribonucleotide nucleotidyltransferase | 163.70 |
| AT5G09420 | translocon at the outer membrane of chloroplasts 64-V | 162.17 |
| AT3G11400 | eukaryotic translation initiation factor 3g | 155.40 |
| AT5G08450 | HDC1 histone deacetylation complex 1 | 153.20 |
| AT5G11980 | conserved oligomeric Golgi complex component-related | 153.10 |
| AT1G79940 | translocation protein SEC63 | 152.73 |
| AT3G45970 | expansin-like A1 | 152.73 |
| AT2G42710 | ribosomal protein .1/L10 family protein | 147.03 |
| AT3G08030 | uncharacterized protein | 146.87 |
| AT5G18620 | CHR17 chromatin remodeling factor17 | 146.83 |
| AT5G50320 | ELO3 histone acetyltransferase | 146.53 |
| AT1G31780 | conserved oligomeric Golgi complex 7 | 144.90 |
| AT5G19400 | telomerase activating protein Est1 | 142.17 |
| AT5G10470 | geminivirus Rep-interacting motor protein | 138.43 |
| AT1G61040 | Plus-3 domain-containing protein | 137.57 |
| AT1G67930 | Golgi transport complex-related protein | 136.70 |

| | | |
|-----------|---|--------|
| AT4G39690 | homolog of yeasst mic60 protein | 136.10 |
| AT1G03860 | prohibitin 2 | 135.43 |
| AT1G19870 | protein IQ-domain 32 | 135.23 |
| AT3G16620 | translocase of chloroplast 120 | 133.90 |
| AT5G61970 | signal recognition particle subunit SRP68 | 133.60 |
| AT1G32380 | phosphoribosyl pyrophosphate synthetase II | 128.47 |
| AT1G78380 | glutathione S-transferase TAU 19 | 128.00 |
| AT3G59020 | armadillo/beta-catenin-like repeat-containing protein | 125.70 |
| AT4G10320 | similar to isoleucyl-tRNA synthetases | 121.20 |
| AT4G24550 | AP-4 complex subunit mu-1 | 121.07 |
| AT1G62740 | putative stress-inducible protein | 117.43 |
| AT1G11910 | aspartic proteinase | 117.33 |
| AT5G67500 | voltage dependent anion channel 2 | 116.70 |
| AT5G46750 | ARF-GAP domain 9 | 115.70 |
| AT2G19480 | putative nucleosome assembly protein 1;2 | 115.53 |
| AT5G66680 | putative dolichyl-di-phosphooligosaccharide-protein glycotransferase | 114.07 |
| AT5G03540 | exocyst subunit exo70 family protein A1 | 112.07 |
| AT5G42960 | Outer envelope pore 24B-like protein | 110.17 |
| AT3G46220 | E3 UFM1-protein ligase 1-like protein | 109.93 |
| AT2G27030 | calmodulin 5 | 109.70 |
| AT3G44330 | M28 Zn-peptidase nicastrin | 109.40 |
| AT3G23300 | S-adenosyl-Lmethionine-dependent methyltransferases superfamily protein | 107.67 |
| AT1G49040 | stomatal cytokinesis defective (SCD1) | 106.07 |
| AT2G05120 | nucleoporin, Nup133/Nup155-like protein | 103.40 |

Appendix Table 4. List of proteins co-purified with CAP-D3-GS. Only proteins present in the three affinity purifications and with no occurrence among the non-specific proteins are listed.

| AGI | Protein names | Average MASCOT Score |
|------------|--|-----------------------------|
| AT4G15890 | CAP-D3 | 6668.17 |
| AT5G48600 | SMC4 | 1571.70 |
| AT5G62410 | SMC2A | 819.37 |
| AT2G19480 | nucleosome assembly protein 1;2 | 496.57 |
| AT2G38770 | intron-binding protein aquarius | 440.83 |
| AT3G13290 | varicose-related protein | 416.07 |
| AT4G02510 | translocase of chloroplast 159 | 407.33 |
| AT3G54110 | uncoupling mitochondrial protein 1 | 399.83 |
| AT1G48900 | signal recognition particle subunit SRP54 | 399.30 |
| AT4G26110 | nucleosome assembly protein 1-like 1 | 397.30 |
| AT4G01990 | pentatricopeptide repeat-containing protein | 371.23 |
| AT4G32050 | Neurochondrin family protein | 331.57 |
| AT2G03510 | SPFH/Band 7/PHB domain-containing membrane-associated protein | 317.63 |
| AT4G01100 | adenine nucleotide transporter 1 | 309.07 |
| AT3G02200 | proteasome component (PCI) domain protein | 308.57 |
| AT2G42710 | ribosomal protein .1/L10 family protein | 304.23 |
| AT3G60860 | SEC7-like guanine nucleotide exchange family protein | 303.10 |
| AT3G01280 | mitochondrial outer membrane protein porin 1 | 299.57 |
| AT5G09840 | putative endonuclease or glycosyl hydrolase | 288.67 |
| AT4G21150 | ribophorin II (RPN2) family protein | 250.40 |
| AT5G18420 | CCR4-NOT transcription complex subunit | 245.07 |
| AT1G64960 | CAP-G2 | 218.27 |
| AT4G33510 | phospho-2-dehydro-3-deoxyheptonate aldolase 2 | 211.13 |
| AT5G58410 | HEAT/U-box domain-containing protein | 210.40 |
| AT5G19760 | mitochondrial substrate carrier family protein | 207.50 |
| AT3G16730 | CAP-H2 | 205.97 |
| AT1G14850 | nucleoporin 155 | 200.63 |
| AT1G06530 | tropomyosin-related protein | 198.07 |
| AT5G40770 | prohibitin 3 | 194.70 |
| AT3G02650 | pentatricopeptide repeat-containing protein | 186.60 |
| AT2G20800 | NAD(P)H dehydrogenase B4 | 184.30 |
| AT1G20960 | putative U5 small nuclear ribonucleoprotein helicase | 183.03 |
| AT2G45140 | VAP-like protein 12 | 182.27 |
| AT3G44330 | M28 Zn-peptidase nicastrin | 180.40 |
| AT5G13110 | glucose-6-phosphate 1-dehydrogenase | 177.80 |
| AT4G24550 | AP-4 complex subunit mu-1 | 174.97 |
| AT5G64270 | putative splicing factor | 170.03 |
| AT5G66680 | putative dolichyl-di-phosphooligosaccharide-protein glycotransferase | 168.17 |
| AT3G02090 | putative mitochondrial processing peptidase | 162.33 |
| AT2G39260 | regulator of nonsense transcripts UPF2 | 161.90 |
| AT5G07340 | calnexin homolog | 155.57 |

| | | |
|-----------|---|--------|
| AT2G33040 | ATP synthase subunit gamma | 155.37 |
| AT4G02150 | Importin subunit alpha-2 | 155.03 |
| AT5G50320 | ELO3 histone acetyl transferase | 154.50 |
| AT3G55620 | translation initiation factor IF6 | 153.50 |
| AT5G27970 | armadillo/beta-catenin-like repeat-containing protein | 153.37 |
| AT1G26460 | pentatricopeptide repeat-containing protein | 153.33 |
| AT1G55890 | pentatricopeptide repeat-containing protein | 152.80 |
| AT1G02370 | pentatricopeptide repeat-containing protein | 151.40 |
| AT5G11980 | putative CONSERVED OLIGOMERIC GOLGI COMPLEX 8 | 148.23 |
| AT5G15610 | proteasome component (PCI) domain protein | 146.13 |
| AT5G20490 | Myosin XI family protein with Dil domain | 146.00 |
| AT5G30510 | ribosomal protein S1 | 145.93 |
| AT4G17330 | hypothetical protein | 144.13 |
| AT5G15020 | SNL2 homolog of the transcriptional repressor SIN3 | 142.40 |
| AT2G20360 | NAD(P)-binding Rossmann-fold superfamily protein | 141.63 |
| AT2G31810 | ACETOLACTATE SYNTHASE SMALL SUBUNIT 1 | 140.57 |
| AT1G64880 | hypothetical Protein | 139.73 |
| AT5G16930 | AAA-type ATPase family protein | 134.10 |
| AT1G27090 | uncharacterized glycine-rich protein | 132.70 |
| AT1G71410 | SCYL2B | 132.50 |
| AT3G55005 | TONNEAU 1B. Involved in cortical microtubule organization | 129.97 |
| AT3G53130 | carotenoid epsilon-ring hydroxylase | 129.13 |
| AT3G49080 | RIBOSOMAL PROTEIN S9 M | 128.83 |
| AT2G43950 | chloroplast outer envelope protein 37 | 128.77 |
| AT2G26890 | GRAVITROPISM DEFECTIVE 2 | 126.27 |
| AT5G12470 | RER4 putative UvrABC system C protein | 125.97 |
| AT4G31810 | 3-hydroxyisobutyryl-CoA hydrolase-like protein 2 | 122.43 |
| AT3G46950 | mitochondrial transcription termination factor family protein | 119.33 |
| AT3G18790 | pre-mRNA-splicing factor ISY1 | 116.20 |
| AT1G23280 | MAK16 protein-like protein | 114.93 |
| AT1G67140 | HEAT repeat-containing protein | 113.27 |
| AT5G46750 | putative ADP-ribosylation factor GTPase-activating protein AGD9 | 110.80 |
| AT3G61140 | CSN1 COP9 signalosome 1 | 110.77 |
| AT2G46020 | ATP-dependent helicase BRAHMA | 109.03 |
| AT4G21800 | QQT2. Required for early embryo development. | 108.77 |
| AT2G15630 | pentatricopeptide repeat-containing protein | 108.30 |
| AT4G38600 | E3 ubiquitin-protein ligase UPL3 | 108.17 |
| AT2G40890 | putative cytochrome P450 | 103.80 |
| AT4G11260 | phosphatase SGT1b | 103.67 |
| AT1G50030 | FKBP12-rapamycin complex-associated protein | 102.90 |
| AT4G05020 | NAD(P)H dehydrogenase B2 | 101.63 |
| AT2G16485 | NERD (Needed for RDR2-independent DNA methylation) | 101.37 |
| AT2G16640 | putative chloroplast outer membrane protein | 100.20 |

Appendix Table 5. *cap-d3* DEG. List of differentially expressed genes common to *cap-d3* SALK vs. wild-type (Wt) and *cap-d3* SAIL vs Wt. Genes in bold are transcription factors.

| AGI | Name | Log2 fold-change SAIL vs Wt | Log2 fold-change SALK vs Wt |
|------------------|---|--|--|
| AT3G48360 | Bt2 | 3.20 | 2.68 |
| AT1G50040 | Unknown | 3.17 | 2.57 |
| AT5G25240 | Unknown | 2.32 | 1.52 |
| AT2G28120 | Otu1 | 2.30 | 1.78 |
| AT5G67480 | Bt4 | 2.22 | 1.51 |
| AT5G05440 | PYL5 | 2.07 | 1.21 |
| AT1G11260 | STP1 | 2.01 | 1.16 |
| AT4G36850 | PQ-loop repeat family protein | 1.97 | 1.28 |
| AT2G44910 | ATHB-4 | 1.90 | 1.64 |
| AT1G32170 | XTH30 | 1.90 | 1.04 |
| AT3G23550 | DTX18 (LAL5) | 1.87 | 1.28 |
| AT1G02380 | Unknown | 1.86 | 1.24 |
| AT2G20670 | Unknown | 1.84 | 1.68 |
| AT5G14120 | Major facilitator? | 1.80 | 1.40 |
| AT3G15630 | Unknown | 1.80 | 1.17 |
| AT2G42870 | PAR1(HLH1) | 1.78 | 1.12 |
| AT2G47440 | Tetratricopeptide repeat | 1.75 | 1.62 |
| AT1G02610 | RING/PHD zinc finger superfam. Prot? | 1.74 | 1.14 |
| AT5G22920 | CHYR1 (ATRZPF34) | 1.68 | 1.12 |
| AT5G19120 | Eukaryotic aspartyl protease family protein? | 1.67 | 1.64 |
| AT2G25200 | Unknown | 1.62 | 1.30 |
| AT5G57550 | XTH25, XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASE 25 | 1.62 | 1.13 |
| AT3G30122 | Pseudogene | 1.62 | 1.92 |
| AT2G36050 | OFP15, ARABIDOPSIS THALIANA OVATE FAMILY PROTEIN 15 | 1.62 | 1.37 |
| AT2G18700 | TPS11, TREHALOSE PHOSPHATASE/SYNTHASE 11 | 1.62 | 1.10 |
| AT4G17470 | Lipid methabolism | 1.56 | 1.09 |
| AT2G15880 | Leucine rich repeat | 1.56 | 1.16 |
| AT5G19190 | Unknown | 1.53 | 1.20 |
| AT4G17460 | HAT1/JAB/JAIBA | 1.52 | 1.63 |
| AT3G19680 | Unknown | 1.52 | 1.51 |
| AT4G38470 | STY46 | 1.51 | 1.03 |
| AT3G15450 | SEN5 | 1.45 | 1.10 |
| AT1G69570 | Hipoxia induced TF | 1.45 | 1.01 |
| AT5G35777 | Transposable element gene | 1.38 | 1.30 |
| AT2G17230 | EXL5 | 1.35 | 1.08 |
| AT2G20835 | Unknown | 1.31 | 1.12 |
| AT2G22770 | NAI1 | 1.30 | 1.26 |
| AT2G17740 | VGL | 1.28 | 1.59 |
| AT5G28145 | Transposable element gene | 1.27 | 1.03 |
| AT5G61590 | DEWAX | 1.25 | 1.04 |

| | | | |
|------------------|--|--------------|--------------|
| AT4G14680 | APS3/ATPS3 | 1.25 | 1.25 |
| AT1G70290 | ATTPS8, ATTPSC, TPS8 | 1.24 | 1.10 |
| AT5G06870 | ATPGIP2, PGIP2 | 1.23 | 1.12 |
| AT1G64200 | VHA-E3 | 1.22 | 1.02 |
| AT3G53232 | DEVIL 20, DVL20, ROTUNDIFOLIA LIKE 1, RTFL1 | 1.21 | 1.13 |
| AT2G23130 | AGP17 | 1.20 | 1.16 |
| AT3G25760 | AOC1/ERD12 | 1.18 | 1.01 |
| AT5G05600 | DOX | 1.17 | 1.11 |
| AT1G70700 | JAZ9, TIFY7 | 1.17 | 1.33 |
| AT5G56550 | OXS3 | 1.13 | 1.22 |
| AT3G15500 | ANAC55 | 1.13 | 1.12 |
| AT1G36370 | MSA1/SHM7 | 1.11 | 1.39 |
| AT4G27410 | ANAC72/RD26 | 1.09 | 1.35 |
| AT4G24015 | RHA4A | 1.09 | 1.05 |
| AT3G49580 | LSU1 | 1.05 | 1.11 |
| AT1G15125 | S-adenosyl-L-methionine-dependent methyltransferase | 1.04 | 1.03 |
| AT4G10910 | Unknown | 1.03 | 1.14 |
| AT5G63130 | Octicosapeptide/Phox/Bem1p family protein | -1.01 | -1.21 |
| AT1G60190 | PUB19 | -1.06 | -1.07 |
| AT3G62260 | Protein phosphatase 2C family protein | -1.09 | -1.07 |
| AT5G15190 | Unknown | -1.09 | -1.15 |
| AT4G25490 | ATCBF1/ DRE BINDING PROTEIN 1B | -1.09 | -1.07 |
| AT3G21150 | ATBBX32 | -1.15 | -1.05 |
| AT3G49710 | Pentatricopeptide repeat (PPR) superfamily protein | -1.17 | -1.28 |
| AT2G34655 | Unknown | -1.19 | -1.04 |
| AT5G05410 | DREB2 | -1.19 | -1.19 |
| AT3G44450 | BIC-1 | -1.21 | -1.08 |
| AT3G22540 | Unknown | -1.23 | -1.01 |
| AT2G34650 | PID/ABR | -1.24 | -1.07 |
| AT3G27210 | Unknown | -1.28 | -1.35 |
| AT2G40750 | ATWRKY54 | -1.29 | -1.17 |
| AT5G41400 | RING/U-box superfamily protein | -1.29 | -1.02 |
| AT5G38005 | Unknown | -1.32 | -1.08 |
| AT1G53080 | Legume lectin family protein | -1.35 | -1.20 |
| AT3G22840 | ELIP | -1.36 | -1.21 |
| AT4G04223 | Unknown | -1.37 | -1.83 |
| AT5G49480 | CP1 | -1.42 | -1.12 |
| AT2G01670 | NUDT17 | -1.48 | -1.58 |
| AT4G25480 | ATCBF3/DREB1A | -1.77 | -1.19 |
| AT1G70640 | octicosapeptide/Phox/Bem1p (PB1) domain-containing protein | -1.84 | -1.36 |
| AT4G15248 | BBX30 | -1.96 | -1.00 |
| AT3G44990 | XTH31,XTR8 | -2.25 | -1.25 |
| AT3G15310 | Transposable element gene | -2.32 | -1.64 |