



Analysis of the response of the nucleus of *Arabidopsis thaliana* to heat stress

Eduardo Muñoz Díaz

► To cite this version:

Eduardo Muñoz Díaz. Analysis of the response of the nucleus of *Arabidopsis thaliana* to heat stress. Vegetal Biology. Université de Perpignan, 2024. English. NNT : 2024PERP0001 . tel-04554274

HAL Id: tel-04554274

<https://theses.hal.science/tel-04554274>

Submitted on 22 Apr 2024

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

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

THÈSE DE DOCTORAT

Analysis of the response of the nucleus of *Arabidopsis thaliana* to heat stress

Présentée en vue de l'obtention grade de docteur
en **Biologie** de l'**Université de Perpignan Via Domitia**

Ecole Doctorale n°305 – Énergie Environnement (ED305)

Présentée par **Eduardo MUÑOZ DÍAZ**

Dirigée par **Julio SÁEZ-VÁSQUEZ**

Soutenue publiquement **22 JANVIER 2024** devant le jury composé de

Mme. Cecile RAYNAUD, Directeur de recherche CNRS INRAE IPS2, Paris	Rapporteuse
Mme. Catharina MERCHANTE BERG, Chercheuse IHSM-UMA-CSIC, Málaga	Rapporteuse
M. Robert BLANVILLAIN, Professeur associé CNRS CEA INRA IRIG-LPCV, Grenoble	Examineur
M. Anthony HENRAS, Directeur de recherche UMR5077 CNRS-UT3, Toulouse	Examineur
M. Jean-Philippe REICHHELD, Directeur de recherche UMR 5096 CNRS-UPVD, Perpignan	Examineur
M. Julio SÁEZ-VÁSQUEZ, Directeur de recherche UMR 5096 CNRS-UPVD, Perpignan	Directeur de thèse

ACKNOWLEDGEMENTS

First of all, I would like to thank the jury of my thesis for having taken the time to read and evaluate this piece of work, as well as the members of my CSI. I would also like to mention the director of the LGDP Olivier Panaud and the members of the ED305.

I have to give special thanks to Julio Sáez Vásquez for giving me the opportunity to participate in this project. He has been able to help me and guide me throughout these three years. Moreover, I would like to remark the excellent contribution of Anne de Bures, as she kindly helped me whenever I required it. This could also be applied to the rest of the members of the team: Pascale Comella, for her knowledge and support, Sophie Brando, for always telling me how nice my perfume is, Tommy Darriere, not only professionally but also personally, Sara Newmann, as my international partner, and the most recent incorporation, Erwan Lemesle. Besides that, thank you to all the LGDP members for having helped me, even when I did not speak French. Special thanks to Dom, Claire, Julie, Christel, Jacinthe, Christelle, Michèle, Elisabeth, Jean-Jacques and Laëtitia. I have to include the coolest cleaning lady, Bea, since she always had the best topics to chat about.

Moreover, I have to mention the best office of the LGDP, which has been evolving throughout these three years. Starting with Émilie, Clément, Arnaud and myself, it grew including Abirami, Juline and Lucas. I need to emphasize Émilie for her emotional support and her role as restaurant partner. Not only did she teach me verlan, but she also helped me to escape from the lab life. To conclude with the professional section of these acknowledgements, I have to include the amazing team in Santiago de Chile. They assisted me in the lab, and they also became my main support during my stay in Santiago. Immense thanks to Diego, Sofi, Ariel, and Isabel, among others. You made my trip to Santiago one of the most remarkable experiences in the last years. I need to include Pablo too. Even though he is not part of the team in the lab, he was (and is) one of the coolest and nicest guys I met there.

Moving on to the more personal section of the acknowledgements, I have to thank the extremely support of my mother, my father, my sister and my grandma during these three years. Thank you mum for being always there and making sure I was alright. I could have not finished this project without her. She has also reminded me to feel confident and proud of myself. Thank you to my dad and my sister, who have been next to me in these three years. Thank you to Marina, who is also “enjoying” her PhD journey. We may give each other advice, we may complain for hours or we may just give up and try to deal with it. Two other friends who also knew how to comfort me are Paloma and Lorena. They have heard my 10-minute-long WhatsApp audios complaining about everything, even though they have no clue what I do in the lab. Thank you, Lorena, for always bringing the TikToks

and memes. They have been key to deal with these intense years. Besides that, thank you Paloma for the emotional support and the little stays in Barcelona.

I have to mention some people I met in Perpignan, who definitely made my 3-year experience nicer. First of all, I need to thank Ana and Jorge. Even though you moved out of Perpignan in 2021, you are probably one of the most remarkable friendships I made so far. Special thanks go to Géraud, for the beers and bars; Guillaume and Kiki, for the festivals, interviews and memes; and Eben, for sharing our struggles. I need to include Khaled, who I recently met. You have been (and are) one the nicest people I met in Perpignan. It feels like I've known you all my life.

Finally, I have to thank myself in the last place, since I managed to finish this stage of my life abroad and alone.

I still don't know what
I'm doing with my life

INDEX

ABSTRACT.....	11
RÉSUMÉ	12
PREFACE.....	14
1. INTRODUCTION.....	17
1.1. Response of plants to heat stress.....	18
1.1.1. Heat Shock Factors and Heat Shock Proteins	18
1.1.2. Regulatory RNAs and epigenetic marks under heat stress	19
1.1.3. Heat stress and post-translational modifications in plants.....	21
1.1.4. ROS signaling and phytohormone response.....	22
1.1.5. Heat stress and light stress.....	23
1.2. DNA-dependent RNA polymerases in eukaryotes.....	24
1.3. Nuclear dynamics: Formation of bodies and trafficking in plant nuclei	27
1.4. The nucleolus: structure and function.....	47
1.4.1. Structure of the nucleolus	47
1.4.2. Nucleolar-Associated Chromatin Domains.....	48
1.4.3. Functions of the nucleolus	49
1.4.3.1. Ribosome biogenesis.....	49
1.4.3.1.1. <i>Transcription of rRNA genes</i>	50
1.4.3.1.2. <i>Processing of pre-rRNA molecules</i>	54
1.4.3.1.3. <i>Formation of pre-ribosomal particles</i>	55
1.4.3.2. Other nucleolar functions in plants	56
1.4.3.3. Nucleolar sequestration of proteins under abiotic stresses in human cells.....	58
1.4.3.3.1. <i>The Nucleolar Detention Signal (NoDS)</i>	58
1.4.3.3.2. <i>IGS-derived lncRNAs</i>	59
1.4.3.3.3. <i>The detention centre</i>	59
2. MATERIALS AND METHODS.....	63
2.1. Acronyms, abbreviations and chemical formulas.....	64
2.2. Genetic nomenclature.....	64
2.3. Biological material.....	64
2.4. Cultures, media and growth conditions.....	65
2.4.1. Cultures, media and growth conditions for Arabidopsis	65
2.4.2. Cultures, media and growth conditions for bacteria.....	66
2.5. Primers used in this project	66
2.6. DNA analysis techniques	66
2.6.1. Plasmid DNA isolation	66

2.6.2. Genomic DNA isolation from Arabidopsis	68
2.6.3. DNA synthesis by Polymerase Chain Reaction (PCR).....	69
2.6.4. Restriction of DNA molecules	69
2.6.5. Quantification of nucleic acids	70
2.6.6. Nucleic acids electrophoresis.....	70
2.6.7. DNA purification from agarose gels.....	70
2.6.8. DNA sequencing	70
2.7. RNA analyses techniques.....	71
2.7.1. Isolation of total RNA.....	71
2.7.2. cDNA synthesis.....	71
2.7.3. Real-Time PCR (qPCR)	72
2.8. Cloning methods	74
2.8.1. Gateway technology	74
2.8.2. Generation of pCambia1300/35S-35S _{pro} :NoDS _{LAS1} -GFP-NLS by restriction enzyme digestion	75
2.8.3. Subcloning into pGEM-T-Easy.....	75
2.8.4. <i>In-silico</i> cloning strategies	76
2.9. Protein methods	76
2.9.1. Total protein extraction from Arabidopsis	76
2.9.2. Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE).....	77
2.9.3. Protein detection by immunoblot.....	77
2.9.4. FLAG immunoprecipitation.....	79
2.10. Transformation methods	79
2.10.1. Bacterial transformation	79
2.10.1.1. Transformation of E. coli by heat shock	79
2.10.1.2. Transformation of Agrobacterium tumefaciens by heat shock.....	80
2.10.2. Stable Arabidopsis transformation by floral dipping	80
2.11. Confocal microscopy.....	81
2.11.1. Observation of fluorophore-tagged molecules.....	81
2.11.2. Observation of proteins by immunofluorescence.....	81
2.11.3. Observation of RNA molecules by Fluorescence <i>In-Situ</i> Hybridization (FISH).....	82
2.12. Image processing.....	83
3. RESULTS.....	85
3.1. Quantitative analysis of the nuclear proteome and transcriptome under heat stress in Arabidopsis.....	86
3.1.1. Quantitative proteomic profiling of Arabidopsis nuclei reveals distinct protein accumulation kinetics upon heat stress.....	87

3.1.2. Comparative analysis of the transcriptome and nuclear proteome of Arabidopsis under heat stress.....	159
3.1.3. Overall results	162
3.2. Behaviour of nucleolar components to heat stress in Arabidopsis.....	163
3.2.1. The nucleolus and heat stress in Arabidopsis.....	164
3.2.2. Goals and experimental approach	166
3.2.3. Results.....	167
3.2.3.1. Subcellular distribution of NRPA3 ^m -FLAG-HA under heat stress.....	167
3.2.3.2. Response of NUC1, FIB2 and C/D box snoRNAs to high temperatures.....	167
3.2.3.3. Behaviour of the fusion protein GFP-OLI2 under heat stress.....	170
3.2.3.4. LAS1-GFP stays in the nucleolus under heat stress	172
3.2.3.5. Arabidopsis NUC2 is induced by heat stress.	175
3.2.3.7. NUC1 surrounds FIB2-YFP and GFP-OLI2 in the nucleolus.....	176
3.2.4. Overall results	179
3.3. Addressing the nucleolar sequestration of proteins under heat stress in Arabidopsis.....	180
3.3.1. Nucleolar sequestration of proteins in Arabidopsis	181
3.3.1.1. Human-like NoDS in Arabidopsis.....	181
3.3.1.2. Transcription of Arabidopsis IGS under heat stress.....	181
3.3.2. Goals and experimental approach	188
3.3.3. Results.....	189
3.3.3.1. Analysis of LAS1 _{pro} :LAS1-GFP (Col-0) plants.....	189
3.3.3.2. Subcellular localisation of GFP-NoDS _{LAS1} and NoDS _{LAS1} -GFP at 22 and 37°C	189
3.3.3.3. Generation and analysis of 35S _{pro} :NoDS _{LAS1} -GFP-NLS plants	191
3.3.3.4. Arabidopsis IGS transcripts under heat stress.....	192
3.3.3.5. IGS-derived RNAs: sense or antisense transcripts	197
3.3.3.6. RNA Pol I may interact with the IGS transcripts under heat stress in Arabidopsis	197
3.3.4. Overall results	199
4. DISCUSSION AND FUTURE PERSPECTIVES.....	202
4.1. Significant accumulation of major nucleolar proteins after heat stress in Arabidopsis	203
4.2. Atypical distribution patterns of nucleolar proteins during and after heat stress in Arabidopsis	213
4.3. Nucleolar targeting of NoDS _{LAS1} and accumulation of IGS transcripts under heat stress	219
4.4. Concluding remarks.....	227
5. ANNEXES	229
5.1. Annexe 1. Impact of repeated heat/light stress exposure on nucleolus-based regulatory processes over generations (Ribostress)	230
5.2. Annexe 2. Further analysis regarding the IGS transcripts	249

5.2.1. Arabidopsis IGS transcripts in different tissues and developmental stages	249
5.2.2. Analysis of the accumulation of the IGS transcripts under biotic stress	249
6. REFERENCES.....	252

FIGURE INDEX

Figure 1.1. Exposure of plants to heat stress.....	20
Figure 1.2. Simplified structure of Arabidopsis HSFs.....	21
Figure 1.3. Representation of the plant nucleus and nucleolus.....	48
Figure 1.4. Structure and function of Arabidopsis ribosomes.....	50
Figure 1.5. 45S rDNA organisation and expression in Arabidopsis.....	52
Figure 1.6. Structure of Arabidopsis IGS.....	53
Figure 1.7. Representation of the different steps of ribosome biogenesis in Arabidopsis	56
Figure 1.8. Processing steps of pre-rRNA molecules.....	58
Figure 1.9. Nucleolar sequestration of proteins in human cells.....	61
Figure 2.1. Primer efficiency for IGS accumulation.....	73
Figure 2.2. Plasmids employed in this project.....	76
Figure 3.1. Comparison of the nuclear proteome and the transcriptome under heat stress in Arabidopsis.....	160
Figure 3.2. Transcriptomic profile under heat stress of the ten established groups	161
Figure 3.3. Effect of high temperatures in Arabidopsis nucleoli.....	165
Figure 3.4. NRPA3m-FLAG-HA forms punctuate structure under heat stress.....	168
Figure 3.5. Subcellular distribution of NUC1 before, during and after heat stress	169
Figure 3.6. Behaviour of FIB2-YFP before, during and after heat stress in Arabidopsis.....	171
Figure 3.7. Distribution of C/D box snoRNAs under normal conditions, heat stress and recovery period in Arabidopsis.....	173
Figure 3.8. Behaviour of GFP-OLI2 under heat stress.....	174
Figure 3.9. Nucleolar distribution of LAS1-GFP under heat stress in Arabidopsis.....	175
Figure 3.10. Heat stress promotes the expression and accumulation of NUC2-FLAG-HA.....	176
Figure 3.11. Co-localisation of NUC1, FIB2-YFP and GFP-OLI2 in <i>35Spro:FIB2-YFP</i> (Col-0) and <i>OLI2pro:GFP-OLI2 (oli2-1)</i> seedlings.....	178
Figure 3.12. Possible NoDS-containing proteins in Arabidopsis.....	186

Figure 3.13. RT-PCR analyses of the accumulation of IGS transcripts under heat stress in Arabidopsis from Kawach, (2019) Master's report.....	187
Figure 3.14. Subcellular distribution of NoDSLAS1-GFP and GFP_NoDSLAS1.....	191
Figure 3.15. Subcellular distribution of NoDSLAS1-GFP-NLS in Arabidopsis.....	192
Figure 3.16. Accumulation of the IGS under heat stress in Arabidopsis.....	194
Figure 3.17. Sequence analysis of the IGS transcripts.....	195
Figure 3.18. RNA FISH to detect IGS transcripts under heat stress.....	196
Figure 3.19. Strand-specific RT-PCR of IGS transcripts in Arabidopsis under heat stress.....	198
Figure 3.20. Detection of IGS transcripts after FLAG immunoprecipitation in <i>NRPA3pro:NRPA3m-FLAG-HA (nrpa3)</i> seedlings.....	199
Figure 4.1. Echevarría-Zomeño et al., (2016) proteome and our nuclear proteome under high temperatures in Arabidopsis.....	204
Figure 4.2. Arabidopsis nuclear and nucleolar proteomes.....	205
Figure 4.3. GO annotation analysis of the 156 proteins with decreasing protein abundance in the nucleus.....	208
Figure 4.4. Arabidopsis RNA Pol I subunits detected in the nuclear proteome	211
Figure 4.5. Heat map with the differences in nuclear abundance of <i>NOP2</i> paralogues before, during and after heat stress.....	211
Figure 4.6. Relationship between changes in the transcriptome and nuclear proteome under heat stress.....	212
Figure 4.7. Schematic representation of the behaviour of nucleolar components under heat stress.....	215
Figure 4.8. Domain organisation, IDRs analysis and 3D structure of major nucleolar proteins.....	218
Figure 4.9. Analysis of the sequence and structure of Arabidopsis LAS1.....	223
Figure 4.10. Model of rRNA IGS transcription at normal conditions and heat stress.....	225
Figure 5.1. Experimental design for the evaluation of high temperatures and intense light over generations.....	230
Figure 5.2. Phenotypic analysis of heat- or light-stressed plants.....	236
Figure 5.3. Comparative analysis of the dry seed weight.....	242
Figure 5.4. Comparative analysis of the bolting date	248
Figure 5.5. Further analysis with the IGS transcripts.....	250

TABLE INDEX

Table 1.1. Subunit composition according to Ream et al., (2015)	25
Table 2.1. Primers used in this project.....	67
Table 2.2. Separating and stacking gels preparation	78
Table 3.1. List of the different C/D box snoRNAs analysed by RNA FISH.....	172
Table 3.2. 118 NoDS-containing proteins (R-R-L/I-X ₍₀₋₁₀₎ -L- Φ-L/V-X ₍₀₋₁₀₎ -L- Φ-L/V) in Arabidopsis proteome.....	182
Table 3.3. 15 Arabidopsis proteins involved in DNA and/or RNA metabolism containing a human-like NoDS.....	185
Table 4.1. Statistical criteria applied to create the EG, LG, TG, EPG, LPG and RG	207

ABSTRACT

The nucleus is the compartment in eukaryotic cells containing most of the genetic information. It is also the site of diverse nuclear bodies, such as the nucleolus or Cajal Bodies. The movement of non-nuclear proteins into the nucleus has been characterized in plants in response to heat stress, which is one of the most prominent abiotic stresses nowadays. However, the nuclear proteome under conditions of heat stress is not yet well-characterized in *Arabidopsis thaliana* (Arabidopsis). Thus, one of the main goals of my PhD project is the analysis of the nuclear proteome before, during and after heat stress. That way, different kinetics have been described according to the relative protein abundance in the nucleus throughout heat stress and the recovery period. This analysis identified crucial nucleolar proteins significantly accumulating in the nucleus in the recovery phase after exposure to high temperatures. Since Arabidopsis nucleoli are affected structural- and functionally upon heat stress, this accumulation of nucleolar proteins in the recovery phase aims the recovery of the normal nucleolar structure and function. The second part of my PhD project includes the response of the nucleolus to heat stress in Arabidopsis. The subcellular distribution of different nucleolar components has been examined during and after heat stress. These components include fibrillarin 2 (FIB2), nucleolin 1 (NUC1) or the RNA polymerase I subunit NRPA3, among others. These nucleolar components exhibit atypical distribution patterns in the nucleolus upon and after heat stress. Finally, the last part of my PhD project involves the existence of the post-translational event of nucleolar sequestration of proteins in Arabidopsis. Two components of this phenomenon have been analysed. On the one hand, the role of a protein motif termed nucleolar detention signal (NoDS), which has been characterised in human cells, has been analysed in Arabidopsis under heat stress. This motif is responsible of the reversible immobilisation of proteins in the nucleolus upon acidosis or heat shock in human cells. NoDSs have been found in certain proteins in the proteome of Arabidopsis. The NoDS of one of those proteins (LAS1) has been experimentally tested for the first time in Arabidopsis, promoting nucleolar localisation under heat stress. On the other hand, the accumulation of the IGS transcripts has also been analysed in Arabidopsis under heat stress, observing a peak of accumulation under heat stress.

Keywords: nucleoli, nucleus, temperature, stress, Arabidopsis

RÉSUMÉ

Le noyau est un compartiment cellulaire chez les eucaryotes qui contient la plupart de l'information génétique. Il contient également divers corps nucléaires, comme le nucléole ou les Corps de Cajal. Le mouvement de protéines non-nucléaires vers le noyau a été caractérisé dans les plantes, en réponse au stress thermique, qui est l'un des stress abiotiques le plus remarquables de nos jours. Cependant, le protéome nucléaire, dans les conditions de stress thermique, n'est pas encore complètement caractérisé chez *Arabidopsis thaliana* (*Arabidopsis*). Donc, un des objectifs de mon projet de thèse est l'analyse du protéome nucléaire avant, pendant et après le stress thermique. Ainsi, j'ai décrit et corrélié différentes cinétiques ont été décrites, selon l'abondance des protéines dans le noyau, pendant le stress thermique et la période de récupération. Cette analyse a permis d'identifier des protéines nucléolaires essentielles qui s'accumulent dans le noyau dans la période de récupération qui suit l'exposition à des températures élevées. Puisque le nucléole, chez *Arabidopsis*, est structurellement et fonctionnellement affecté par le stress thermique, cette accumulation de protéines nucléolaires dans la période de récupération est destinée à restaurer la structure et de la fonction normales du nucléole. La seconde partie de mon projet consiste d'étudier la réponse du nucléole à de hautes températures chez *Arabidopsis*. La distribution subcellulaire de différents facteurs nucléolaires a été examinée pendant et après le stress thermique. Parmi ces derniers, la fibrillarine 2 (FIB2), la nucléoline 1 (NUC1) ou la sous-unité de l'ARN polymérase I NRPA3, parmi d'autres. Ces composants nucléolaires montrent une distribution atypique dans le nucléole pendant et après le stress thermique. Finalement, la dernière partie de mon projet s'intéresse au phénomène de séquestration nucléolaire de protéines chez *Arabidopsis*, régulé au niveau post-traductionnel. Deux éléments de ce phénomène ont été analysés. D'un côté, la fonction d'un motif protéique nommé signal de détention nucléolaire [Nucleolar Detention Signal (NoDS) en anglais], caractérisée dans les cellules humaines, a été analysée chez *Arabidopsis* pendant le stress thermique. Ce motif est le responsable de l'immobilisation réversible de protéines dans le nucléole de l'acidose ou de choc thermique. Ce motif est présent dans plusieurs protéines du protéome chez *Arabidopsis*. J'ai mis en évidence et validé expérimentalement le NoDS d'une de ces protéines (LAS1) pour la première fois chez *Arabidopsis*. Ce motif permet l'adressage de la protéine LAS1 vers le nucléole lors de stress thermique. D'autre part, j'ai également montré que la transcription des IGS se caractérise par un pic d'accumulation de transcrits pendant le stress thermique chez *Arabidopsis*.

Mots clés : nucléole, noyau, température, stress, *Arabidopsis*

PREFACE

The existence of the nucleus conforms one of the main differences between eukaryotic and prokaryotic cells. It is the site of most of the genetic information in eukaryotic cells. Structurally, the nucleus is separated from the cytoplasm by a double membrane with pores. This membrane also allows a functional separation, as DNA replication and transcription happens in the nucleus, and protein translation occurs in the cytoplasm. The nucleus also possesses different types of nuclear bodies, such as the nucleolus or Cajal Bodies. The nucleolus is the most prominent nuclear body, whose structure and function have been studied. It is composed of the fibrillar centre, surrounded by the dense fibrillar component, which are embedded in the granular component. This tripartite distribution is tightly linked with its main function, which is the ribosome biogenesis. Apart from being a “ribosome factory”, the nucleolus also exerts other functions, one of them being a stress sensor.

The overall goal of my thesis project is to dive into the effect of heat stress on the nucleus and nucleolus. High temperatures are one of the main abiotic stressors, having an impact on plant growth and development. Thus, my thesis project has been divided in three different sections: (i) the analysis of the nuclear proteome of *Arabidopsis thaliana* under heat stress, (ii) the response in the subcellular distribution of nucleolar components to heat stress in *Arabidopsis thaliana*, and (iii) the analysis of the protein sequestration in the nucleolus of *Arabidopsis thaliana* upon heat stress. Thus, my PhD project is structured according to these three sections, as well as the canonical chapters.

The chapter “Introduction” is composed of four sections. The first one is the response of plants to heat stress. The second section briefly touches eukaryotic RNA polymerases. The third section consists of a review article published in *Frontiers of Plant Science* in August 2022 titled “Nuclear dynamics: Formation of bodies and trafficking in plant nuclei”. Finally, the last section of the introductory chapter deepens into the structure and function of the nucleolus. This chapter is followed by “Materials and methods”. After that, the chapter “Results” is composed of three sections. The first section contains the analysis of the nuclear proteome of *Arabidopsis thaliana* under heat stress. The results of this section have been compiled in an article which has been sent to *The Plant Journal* for publication. The results of this paper are completed with the analysis of the transcriptome of *Arabidopsis thaliana* before, during and after heat stress. The second section revolves around the behaviour of major nucleolar components under heat stress in *Arabidopsis thaliana*. The third section of this chapter dives into the nucleolar sequestration of proteins under heat stress in *Arabidopsis thaliana*. The second and third sections possess a short specific introduction with previous results, the goals and experimental approaches, the results, and a short conclusion. Afterwards, in the chapter “Discussion and future perspectives”, the results shown in the chapter “Results” are discussed in a more integrative way, taking into consideration published results. It also contains certain approaches and ideas for further characterisation of the presented results. Finally, two different annexes are

presented: (i) the impact of repeated heat/light stress exposure on nucleolus-based regulatory processes over generations (in the frame of the ANR project Ribostress) and (ii) further analysis regarding the IGS transcripts. The second annexe includes the analyses of the accumulation of IGS transcripts in different plant tissues and developmental stages, as well as under biotic stress (as part of an Ecos-Sud collaboration).

The figures and tables have been labelled according to the chapter they are located, employing the format “Figure/Table x.yy”. Whereas “x” represents the chapter the corresponding figure is included in, “yy” establishes the order of the figures in that particular chapter. For instance, Figure 3.12 represents a figure in the chapter 3 (“Results”), being the 12th figure in that chapter. Similarly, Table 2.1 indicates the 1st table of chapter 2 (“Materials and Methods”). This has been applied to the figures and tables shown exclusively in this manuscript. Figures and tables corresponding to the review article “Nuclear dynamics: Formation of bodies and trafficking in plant nuclei” and the submitted article “Quantitative proteomic profiling of Arabidopsis nuclei reveals distinct protein accumulation kinetics upon heat stress” do not follow this format.

1. INTRODUCTION

1.1. Response of plants to heat stress

Plants are sessile organisms and, therefore, they are exposed to a variety of stresses. A stress condition can be described as an adverse scenario (internal or external) which is going to impact the ability of the plant to produce biomass. The severity of the stress varies according to its intensity and its duration. Thus, plants are able to recover and/or adapt the stressor, when exposed to a light and/or short stress, or they may die when the stress conditions are more pronounced. Plants are exposed to two types of stresses: biotic and abiotic stresses. Biotic stressors are biological entities, including fungi, viruses, and other pests, which threaten plant growth and development. These biological agents cause severe effects to plants, leading to plant death. On the other hand, abiotic stresses are environmental agents that negatively affect plant growth and development. Abiotic stressors include high and low temperatures, drought, high-intensity light, exposure to toxic compounds, or high salinity, among others (reviewed by Verma et al., 2013; reviewed by Gull et al., 2019; reviewed by Umar et al., 2022).

The exposure to high temperatures or heat stress is one of the main stressors nowadays, as the average temperature is increasing worldwide (reviewed by Lippmann et al., 2019). Heat stress consists of temperatures 10-15°C above normal conditions (reviewed by Firmansyah and Argosubekti, 2020). This increase in the ambient temperature has a clear impact on plant physiology. This includes a decrease in the photosynthetic efficiency, disrupting the thylakoid membranes of the photosystems. The overall structure of the chloroplasts is also impacted by high temperatures (reviewed by Zhao et al., 2021). Heat stress also increases the fluidity of the lipid membranes, leading to the loss of electrolytes. In addition, the negative impact upon exposure to heat stress has been addressed in seed germination (reviewed by Song et al., 2014). Moreover, heat stress triggers the expression and accumulation of certain proteins and RNA species, as well as the presence of certain epigenetic marks and post-translational modifications. The role of reactive oxygen species (ROS) and phytohormones in the response to heat stress has also been documented (Figure 1.1).

1.1.1. Heat Shock Factors and Heat Shock Proteins

Heat Shock Factors (HSFs) are transcription factors which induce the expression of heat stress responsive genes by associating with a specific sequence of their promoter termed heat shock element. Three classes of HSFs can be distinguished: class A, B and C. The structure of the members of each class of HSFs is highly conserved in *Arabidopsis thaliana* (Arabidopsis). Class A HSFs are characterised by the presence of a DNA binding domain in the N-terminus, the oligomerisation domain HR-A/B, characterised by the presence of hydrophobic residues, and the C-terminal nuclear localisation signal (NLS), nuclear export signal (NES) and the AHA motifs, which are involved in the

transcription activation. In contrast, the AHA motif is lacking in class B and C HSFs in Arabidopsis (Figure 1.2; Nover et al., 2001). Functionally, class A HSFs are key players in the response to heat stress. For instance, HsfA1a and HsfA1b play a key role in the earlier stages of the response to heat stress in Arabidopsis, whereas HsfA2 mainly participates upon long-term heat stress and recovery conditions (reviewed by Kotak et al., 2007). Moreover, some class A HSFs are responsive to ROS in Arabidopsis (Pérez-Salamó et al., 2014; Giesguth et al., 2015). Class B HSFs have been classified as transcriptional repressors of heat stress responsive genes (Czarnecka-Verner et al., 2004; Ikeda et al., 2011). In contrast, the function of class C HSFs is still not clear (reviewed by Nover et al., 2001; reviewed by Baniwal et al., 2004; reviewed by Wang et al., 2023). Besides that, other transcription factors involved in the response to heat stress are the multiprotein-bringing factor 1c (MBF1c), and the NAC, the WRKY, the MYB and the basic leucine zipper (bZIP) families (reviewed by Zhao et al., 2021; reviewed by Wang et al., 2023).

HSFs promote the expression of Heat Shock Proteins (HSPs), the best-known heat stress responsive genes. HSPs can be classified according to their molecular weight: HSP100, HSP90, HSP70, HSP60 and small HSPs. HSPs act as molecular chaperones, preventing the denaturation of proteins upon heat stress. Under normal conditions, HSPs associate to HSFs, preventing them from transcribing heat stress responsive genes. Under heat stress, HSPs no longer bind to HSFs, as they exert their role as molecular chaperones (reviewed by Kotak et al., 2007; reviewed by Qu et al., 2013; reviewed by Wang et al., 2023). In addition, other proteins have been observed to be induced in the response to heat stress, including ubiquitins, dehydrins or late embryogenesis abundant (LEA) proteins (reviewed by Bokszczanin and Fragkostefanakis, 2013).

1.1.2. Regulatory RNAs and epigenetic marks under heat stress

In addition to proteins, several non-coding RNAs (ncRNAs) participate in the response to heat stress and the acquisition of thermotolerance. Some of them include microRNAs (miRNAs), which are small RNAs whose targets are messenger RNAs (mRNAs). These miRNAs are able to degrade the mRNAs or repress its translation. For instance, miR156 is induced after exposure to heat stress, leading to the expression of heat stress responsive genes. This miRNA has been characterised as a key player in the memory to heat stress (Stief et al., 2014). Moreover, miR398 is also transcribed upon heat stress, as HsfA1b and HsfA7b bind to its promoter in Arabidopsis. miR398 reduces the amount of transcripts of copper/zinc superoxide dismutase 1 and 2 (CSD1 and 2, respectively), which act as negative regulators in the expression of HSFs (Guan et al., 2013). Similarly, the transcription of miR165/166 is also induced by heat stress. These miRNAs regulate HsfA1 in a transcriptional and translational manner (Li et al.,

2023). Moreover, the *trans*-acting small interfering RNAs (ta-siRNAs) HEAT-INDUCED TAS1 TARGET1 (HTT1) and HTT2 have been characterised as player in the acquisition of thermotolerance (Li et al., 2014).

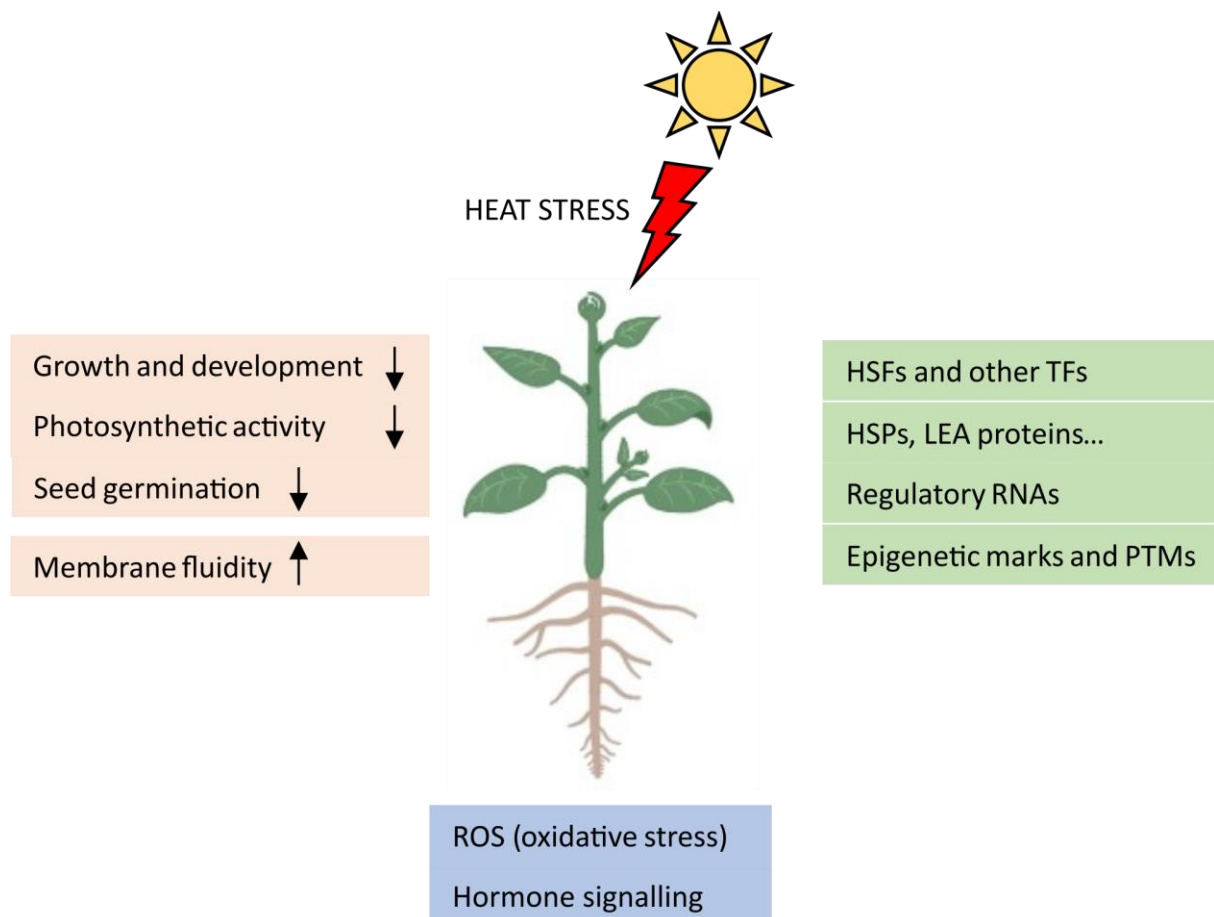


Figure 1.1. Exposure of plants to heat stress. The illustration shows some of the physiological effects of heat stress (orange), the diverse changes in the transcriptomic and proteomic profiles (green) and the ROS and hormonal signalling pathways (blue) which are involved in the heat stress response.

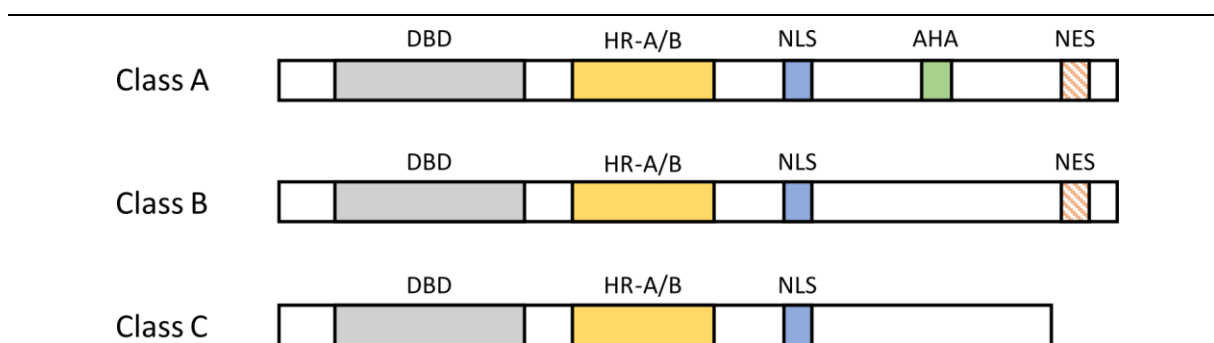


Figure 1.2. Simplified structure of Arabidopsis HSFs. Class A HSFs (top) possess the activation motif AHA in their C-terminus, which has not been identified in class B (middle) and C (bottom) HSFs. The presence of a nuclear export signal (NES) has been reported in some class A and B HSFs (striped). DBD: DNA Binding Domain (grey); HR-A/B: oligomerisation domain (yellow); NLS: Nuclear Localisation Signal (blue); AHA: activation domain (green); NES: Nuclear Export Signal (orange). Figure based on Nover et al., (2001).

Not only are the transcriptomic and proteomic profiles of plants modified during heat stress, but there are also certain epigenetic marks that are affected by heat stress. The effect of the exposure to high temperatures has been addressed in plants regarding DNA methylation or histone modifications (reviewed by Verma et al., 2013; reviewed by Zhao et al., 2021). DNA methylation is the addition of a methyl group ($-\text{CH}_3$) to the cytosine in the DNA. In Arabidopsis, the methylation state changes during and after heat stress, as heat-stressed plants possess more methylation marks (Korotko et al., 2021). Besides that, certain mutants in genes indirectly involved in DNA methylation, such as *DCL3* or *NRPD2*, are heat sensitive, establishing a link between DNA methylation and the response to heat stress (Popova et al., 2013). On the other hand, histone modifications affect the activation of genes involved in the response to heat stress. DNA and histone octamers (two copies of H2A, H2B, H3, and H4) compose the nucleosome. For instance, the histone chaperones ANTI-SILENCING FUNCTION 1A/B (ASF1A/B) promote the expression of HSFs and HSPs via histone acetylation in Arabidopsis (Weng et al., 2014). Certain histone modifications are associated with the heat stress memory. This is the case of H3K4me2/3 in HSF2, which increases upon exposure to heat stress (Lämke et al., 2016).

1.1.3. Heat stress and post-translational modifications in plants

Other than histone acetylation, the link between other post-translational modifications (PTMs) and the heat stress response has been examined. These PTMs include phosphorylation, ubiquitination and SUMOylation (reviewed by Han et al., 2022). Phosphorylation is one of the main PTMs, consisting in the addition of a phosphate group by a kinase. As a result, phosphorylation negatively or positively influences the the function and/or localisation of a protein (reviewed by Nardoizzi et al., 2010). Several

examples in the literature portray the relationship between protein phosphorylation and heat stress. For instance, the cyclin-dependent kinase CDC2 phosphorylates Arabidopsis HSFs, inhibiting the binding to Heat Shock Elements (Reindl et al., 1997). Similarly, the phosphorylation of HSFA1a by Arabidopsis CaM-binding protein kinase 3 (CBK3), modulating the association with Heat Shock Elements (Liu et al., 2008). On the other hand, Arabidopsis mitogen-activated protein kinase 6 (MPK6) is induced by heat stress. This kinase phosphorylates HSFA2, promoting the nuclear localisation of HSFA2 under heat stress (Evrard et al., 2013). Secondly, ubiquitination consists in the reversible association of the C-terminal of a ubiquitin to an amino acid, commonly lysine. As a result, proteins can become monoubiquitinated (a single ubiquitin) multiubiquitinated (several monoubiquitins) or polyubiquitinated (oligoubiquitination; Sadowski and Sarcevic, 2010). Under heat stress, the main ubiquitination event is polyubiquitination, which targets misfolded proteins to the 26S proteasome for degradation (reviewed by Han et al., 2022). Lastly, SUMOylation is the attachment of a small ubiquitin-like modifier (SUMO) to a protein. Upon high temperatures, the levels of SUMOylation increase, phenomenon termed as SUMO stress response in eukaryotes (Lewicki et al., 2015). Besides that, SUMOylation has been linked to thermotolerance in Arabidopsis via the HSFA1 family (Hammoudi et al., 2021).

1.1.4. ROS signaling and phytohormone response

Moreover, heat stress leads to the production of reactive oxygen species (ROS), causing a secondary stress: the oxidative stress. These ROS include singlet oxygen ($^1\text{O}_2$), superoxide anion radical ($\text{O}_2^{\bullet-}$), hydroxyl radical (OH) and hydrogen peroxide (H_2O_2). This oxidative stress is characterised by the inactivation of enzymes, protein degradation, and membrane and pigment damage. However, the correct proportion of ROS production and scavenging activity allows the role of ROS as signalling molecules during the response to heat stress (reviewed by Awasthi et al., 2015; reviewed by Fortunato et al., 2023; reviewed by Hendrix et al., 2023). HSFs, such as HSFA4a or HSFA5, act as a positive and negative regulators of the ROS signalling, promoting the activation of transcription factors via the mitogen-activated protein kinase (MAPK) pathway (reviewed by Qu et al., 2013).

The role of various phytohormones during the response to heat stress has been analysed. For instance, it has been documented how abscisic acid (ABA) induces the expression of HSFA6b in the response to heat stress (Huang et al., 2016). Similarly, the ETHYLENE RESPONSE FACTORS 95 AND 97 (ERF95 and 97, respectively) bind to the promoter of HSFA2, promoting the expression of heat responsive genes (Huang et al., 2021a). Moreover, the levels of salicylic acid increase in the heat stress response, promoting the expression of HSPs (Clarke et al., 2004). Besides that, mutants in genes

involved in the synthesis and/or signalling pathway of phytohormones display low tolerances to heat stress (reviewed by Devireddy et al., 2021).

1.1.5. Heat stress and light stress

Even though plants can be exposed to controlled stress conditions for experimental purposes, multiple stressors co-exist in nature. In other words, heat stress in nature is often accompanied by other stresses, such as drought and/or high light stresses. Besides that, the existence of additional stresses, i.e., biotic stressors, is also common (reviewed by Suzuki et al., 2014). Thus, the combined effect of heat and light stresses has been documented in the literature (Hewezi et al., 2008; reviewed by Janda et al., 2021). Exposure to high-intensity light or light stress leads to photodamage. There is a subsequent imbalance in the energy distribution between the photosystems, having a detrimental effect on the photosynthetic efficiency. For that, plants have evolved a series of responses to deal with light stress. These responses include the production of chloroplastic ROS, the synthesis of anthocyanins or the modulation in the opening of stomata (reviewed by Shi et al., 2022).

1.2. DNA-dependent RNA polymerases in eukaryotes

DNA transcription is carried out by DNA-dependent RNA polymerases (RNA polymerases). Most eukaryotes possess three main RNA polymerases: RNA polymerase I, II and III (RNA Pol I, II and III, respectively). RNA Pol I, located in the nucleolus, is responsible of the transcription of the 5.8S, 18S and 25S (28S in mammals) rRNA genes. RNA Pol II is present in the nucleoplasm and performs the transcription of protein-coding genes into pre-mRNA, as well as certain ncRNAs (miRNAs, small nucleolar RNAs, small nuclear RNAs and long non-coding RNAs). RNA Polymerase III, also present in the nucleoplasm, is involved in the transcription of 5S rDNA genes, transfer RNA (tRNAs), certain regulatory RNAs and retrotransposons (Guilfoyle and Dietrich, 1987; Ream et al., 2015; reviewed by Yang et al., 2023). Moreover, the organelle RNA polymerases have been also described: the chloroplastic (in plants) and the mitochondrial RNA polymerases. Two chloroplastic RNA polymerases have been characterised: the plastic-encoded and the nuclear-encoded RNA polymerases (PEP and NEP, respectively). Interestingly, *Arabidopsis* possesses two additional RNA polymerases: RNA polymerases IV and V (RNA Pol IV and V, respectively). These two polymerases have evolved from RNA Pol II (Ream et al., 2015). Whereas RNA Pol IV is in charge of the synthesis of the 24-nt long siRNAs (Panda et al., 2020), the RNA Pol V generates long non-coding RNAs (lncRNAs) involved in the de novo DNA methylation (Wierzbicki et al., 2008; Borges and Martienssen, 2015). Structurally, yeast RNA Pol I, II and III possess 14, 12 and 17 subunits, respectively. Homology analyses in *Arabidopsis* revealed the existence of 15 subunits in RNA Pol I, 19 subunits in RNA Pol II, 23 subunits in RNA Pol III, 16 subunits in RNA Pol IV, and 18 subunits in RNA Pol V (Table 1.1; Ream et al., 2009, 2015).

Three different stages can be differentiated in DNA transcription: initiation, elongation and termination. Even though eukaryotic RNA polymerases have common components in these three stages, some particularities are found. In the initiation phase, RNA polymerases require General Transcription Factors (GTFs) to recognise the specific sequence in the promoter (TATA box), so that the double-stranded DNA helix is able to unwind (reviewed by Tyagi, 2001; reviewed by Yang et al., 2023). The initiation factors for transcription via RNA Pol II in eukaryotes are TFIIA, TFIIB, TFIIC, TFIID, TFIIIE, TFIIF and TFIIH, which act sequentially in the initiation process. Similarly, TFIIA, TFIIB and TFIIC are required for DNA transcription by RNA Pol III. In contrast, the initiation factors needed for transcription by RNA Pol I vary among species. For instance, the upstream activating factor (UAF), the core factor (CF) and the RNA polymerase I transcription factor RRN3 are necessary in budding yeast. Nevertheless, the general recognition factors implicated in the transcription via RNA Pol IV and V remain mainly uncharacterised (reviewed by Yang et al., 2023).

Table 1.1. Subunit composition according to Ream et al., (2015)

The colours indicate conserved subunits among the different RNA polymerases. For instance, At3g22320 encodes a subunit of RNA Pol I, II, III and IV, coloured in brown,

RNA Pol I		RNA Pol II		RNA Pol III		RNA Pol IV		RNA Pol V	
Name	Accession	Name	Accession	Name	Accession	Name	Accession	Name	Accession
NRPA1	At3g57660	NRPB1	At4g35800	NRPC1	At5g60040	NRPD1	At1g63020	NRPE1	At2g40030
NRPA2	At1g29940	NRPB2	At4g21710	NRPC2	At5g45140	NRPD2	At3g23780	NRPE2	At3g23780
NRPA3	At1g60850	NRPB3	At2g15430	NRPC3	At1g60620	RPD2b	At3g18090	NRPE3a	At2g15430
NRPA4	-	NRPB4	At5g09920	NRPC4	At5g62950	NRPD3	At2g15430	NRPE3b	At2g15400
NRPA5	At3g22320	NRPB5	At3g22320	NRPC4-like	At3g28956	NRPD4	At4g15950	NRPE4	At4g15950
NRPA6a	At5g51940	NRPB5-like	At5g57980	NRPC5	At3g22320	NRPD5	At3g22320	NRPE5	At3g57080
NRPA6b	At2g04630	NRPB6a	At5g51940	NRPC6a	At5g51940	NRPD6a	At5g51940	NRPE5-like	At2g41340
NRPA7	At1g75670	NRPB6b	At2g04630	NRPC6b	At2g04630	NRPD6b	At2g04630	NRPE5-like	At3g54490
NRPA8a	At1g54250	NRPB7	At5g59180	NRPC7	At1g06790	NRPD7	At3g22900	NRPE6a	At5g51940
NRPA8b	At3g59600	NRPB7-like	At4g14520	NRPC8a	At1g54250	NRPD8a	At1g54250	NRPE6b	At2g04630
NRPA9	At3g25940	NRPB8a	At1g54250	NRPC8b	At3g59600	NRPD8b	At3g59600	NRPE7	At4g14660
NRPA10	At1g11475	NRPB8b	At3g59600	NRPC9a	At4g07950	NRPD9a	At3g16980	NRPE8a	At1g54250
NRPA11	At2g29540	NRPB9a	At3g16980	NRPC9b	At1g01210	NRPD9b	At4g16265	NRPE8b	At3g59600
NRPA12	At5g41010	NRPB9b	At4g16265	NRPC10	At1g11475	NRPD10	At1g11475	NRPE9a	At3g16980
NRPA13	At3g13940	NRPB10	At1g11475	NRPC11	At2g29540	NRPD11	At3g52090	NRPE9b	At4g16265
NRPA14	At5g64680	NRPB10-like	At1g61700	NRPC12	At5g41010	NRPD12	At5g41010	NRPE10	At1g11475
		NRPB11	At3g52090	NRPC13	At3g49000			NRPE11	At3g52090
		NRPB12	At5g41010	NRPC14a	At4g25180			NRPE12	At5g41010
		NRPB12-like	At1g53690	NRPC14b	At5g09380				
				NRPC15	At5g49530				
				NRPC16	At5g23710				
				NRPC17C	At4g01590				
				NRPC17b	At4g35680				

Additionally, DNA transcription is modulated by regulatory sequences: enhancers and silencers. The former are short sequences that act as activators of DNA transcription. They act in *cis* or in *trans*, in both directions and at great distances from the target gene. On the other hand, silencers are repressors of DNA transcription in the vicinity of the target genes (reviewed by Tyagi, 2001; reviewed by Kolovos et al., 2012).

For transcription elongation, accessory elongation factors are required for an efficient elongation from RNA Pol I, II and III (Anderson et al., 2011; Girbig et al., 2021; reviewed by Chen et al., 2018). For instance, the polymerase-associated factor 1 complex (PAF1C) acts as an elongation factor with Pol I and II in yeast (Zhang et al., 2009; Xu et al., 2017). Homology analyses reveal the existence of plant orthologues of these genes in plants. Whereas RNA Pol IV interacts with the single-subunit RNA-dependent RNA polymerase RDR2 during elongation, RNA Pol V requires the recruitment of ARGONAUTE 4 and 6 (AGO4 and 6, respectively), and the domains rearranged methyltransferase 2 (DRM2; reviewed by Yang et al., 2023). However, the elongation carried out by RNA Pol IV and V is not as efficient as RNA Pol II. Regarding RNA Pol IV, it has been proposed that the residues present in its surface impedes the correct recruitment of elongation factors (Huang et al., 2021b). In contrast, elongation by RNA Pol V has a higher fidelity despite its lower efficiency (Marasco et al., 2017).

In the termination stage, the nascent RNA molecule separates from the elongation machinery in different ways depending on the RNA polymerase (reviewed by Yang et al., 2023). Termination in RNA Pol I transcription slightly differs among eukaryotes. In yeast, the termination step is mediated by the action of three sequential proteins: RNase three 1 (Rnt1), the termination factors RNA polymerase I enhancer binding protein (Reb1) and NTS1 silencing protein 1 (Ntsi1), and the exonuclease ribonucleic acid trafficking 1 (Rat1), which degrades the nascent transcript from the 5' to the 3'. This last step is known as the torpedo mechanism, which concludes with the release of the transcript (Kawauchi et al., 2008). In humans, the release of the RNA molecule in the last step is not carried out by the torpedo mechanism (Németh et al., 2013). On the other hand, the type of transcript is going to determine the players of the termination step by RNA Pol II. For mRNAs, the multiple-subunit cleavage and polyadenylation factor-cleavage factor complex (CPF-CF) cleaves at the polyA site (reviewed by Casañal et al., 2017). Then, two models have been proposed for the release of the transcripts: the torpedo mechanisms and the allosteric model (reviewed by Proudfoot, 2016). For ncRNAs, the termination step by RNA Pol II varies among species (reviewed by Yang et al., 2023). In contrast, the termination of the transcription by RNA Pol III happens spontaneously at polythymidine sites in eukaryotes (Cozzarelli et al., 1983; Cloix et al., 2003). Transcription termination in RNA Pol IV transcription follows a three-step atypical model (Fukudome et al., 2021). On the other hand, the termination by RNA Pol V remains uncharacterised.

1.3. Nuclear dynamics: Formation of bodies and trafficking in plant nuclei

Article published in *Frontiers of Plant Science*



OPEN ACCESS

EDITED BY

Ming Luo,

South China Botanical Garden (CAS),
China

REVIEWED BY

Yuda Fang,

Shanghai Jiao Tong University, China
Kentaro Tamura,

University of Shizuoka, Japan

*CORRESPONDENCE

Eduardo Muñoz-Díaz
eduardo.munoz-diaz@univ-perp.fr
Julio Sáez-Vásquez

saez@univ-perp.fr

SPECIALTY SECTION

This article was submitted to
Plant Cell Biology,a section of the journal
Frontiers in Plant Science

RECEIVED 01 July 2022

ACCEPTED 04 August 2022

PUBLISHED 23 August 2022

CITATION

Muñoz-Díaz E and Sáez-Vásquez J
(2022) Nuclear dynamics: Formation of
bodies and trafficking in plant nuclei.
Front. Plant Sci. 13:984163.

doi: 10.3389/fpls.2022.984163

COPYRIGHT

© 2022 Muñoz-Díaz and

Sáez-Vásquez. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Nuclear dynamics: Formation of bodies and trafficking in plant nuclei

Eduardo Muñoz-Díaz^{1,2*} and Julio Sáez-Vásquez^{1,2*}

¹Centre National de la Recherche Scientifique (CNRS), Laboratoire Génome et Développement des Plantes, UMR 5096, Perpignan, France, ²Univ. Perpignan Via Domitia, Laboratoire Génome et Développement des Plantes, UMR 5096, Perpignan, France

The existence of the nucleus distinguishes prokaryotes and eukaryotes. Apart from containing most of the genetic material, the nucleus possesses several nuclear bodies composed of protein and RNA molecules. The nucleus is separated from the cytoplasm by a double membrane, regulating the trafficking of molecules in- and outwards. Here, we investigate the composition and function of the different plant nuclear bodies and molecular clues involved in nuclear trafficking. The behavior of the nucleolus, Cajal bodies, dicing bodies, nuclear speckles, cyclophilin-containing bodies, photobodies and DNA damage foci is analyzed in response to different abiotic stresses. Furthermore, we research the literature to collect the different protein localization signals that rule nucleocytoplasmic trafficking. These signals include the different types of nuclear localization signals (NLSs) for nuclear import, and the nuclear export signals (NESs) for nuclear export. In contrast to these unidirectional-movement signals, the existence of nucleocytoplasmic shuttling signals (NSSs) allows bidirectional movement through the nuclear envelope. Likewise, nucleolar signals are also described, which mainly include the nucleolar localization signals (NoLSs) controlling nucleolar import. In contrast, few examples of nucleolar export signals, called nucleoplasmic localization signals (NpLSs) or nucleolar export signals (NoESs), have been reported. The existence of consensus sequences for these localization signals led to the generation of prediction tools, allowing the detection of these signals from an amino acid sequence. Additionally, the effect of high temperatures as well as different post-translational modifications in nuclear and nucleolar import and export is discussed.

KEYWORDS

nucleoplasm, nucleolus, stress, localization signals, non-coding RNAs, nuclear bodies

Introduction: Cell compartmentalization and the nucleus

Cell compartmentalization allows the physical separation of molecules and metabolic reactions within the cell. In particular, plants possess and exert a large number of biochemical routes and metabolites because of their sessile character (Solymosi and Schoefs, 2019). Thus, compartmentalization is essential in plant cells for their correct functioning. During evolution, compartmentalization appeared as the distinction between eukaryotes and prokaryotes, since prokaryotic cells lack membrane-bound organelles. Interestingly, plant cells possess an exclusive organelle, the chloroplast, whose best-known function consists in obtaining energy through photosynthesis (Alberts et al., 2002; Lunn, 2007; Solymosi and Schoefs, 2019).

Nevertheless, the nucleus can be conceived as the organelle distinguishing eukaryotic and prokaryotic cells. The nucleus contains most of the genetic material, excluding the mitochondrial and (in plants) the chloroplastic genomes (Figure 1). Functionally, it separates the DNA replication and DNA transcription taking place in the nucleoplasm from the protein translation in the cytosol. The nucleus also has a protective effect on the genetic material. Structurally, it comprises the nucleoplasm and the nuclear envelope (NE). The nucleoplasm contains the chromatin and the nuclear bodies, and it is also the site of several enzymes involved in the metabolism of DNA and RNA. On the other hand, the NE delimitates the nucleoplasm from the cytoplasm. It is composed of the outer nuclear membrane and the inner nuclear membrane (ONM and INM, respectively), forming the perinuclear space in-between. Whereas the ONM is in contact with the endoplasmic reticulum in the cytoplasm, the INM associates with the nuclear lamina, which is involved in several nuclear functions in animals cells (Taddei et al., 2004; Guo and Fang, 2014). The nucleoplasm and the cytoplasm are in contact through thousands of nuclear pore complexes (NPCs) located along the NE. In addition, actively transcribed chromatin is often found interacting with the NPC, whereas inactive chromatin is associated with the nuclear lamina in animals and yeast (Taddei et al., 2004; Németh and Längst, 2011) or with the periphery of the nucleolus in eukaryotic cells (Hicks, 2013; Padeken and Heun, 2014; Pontvianne et al., 2016; Picart-Piccolo et al., 2020; Castel and Chae, 2021).

Nuclear bodies are dynamic structures composed of proteins and RNA molecules involved in related functions (Mao et al., 2011; Petrovská et al., 2015). They are present in the nucleoplasm and/or the nucleolus. Indeed, the best-known nuclear body in eukaryotic cells is the nucleolus (Mao et al., 2011; Sleeman and Trinkle-Mulcahy, 2014; reviewed by Sáez-Vásquez and Medina, 2008). Even though animals, yeast and plant cells share certain nuclear bodies, there are others that

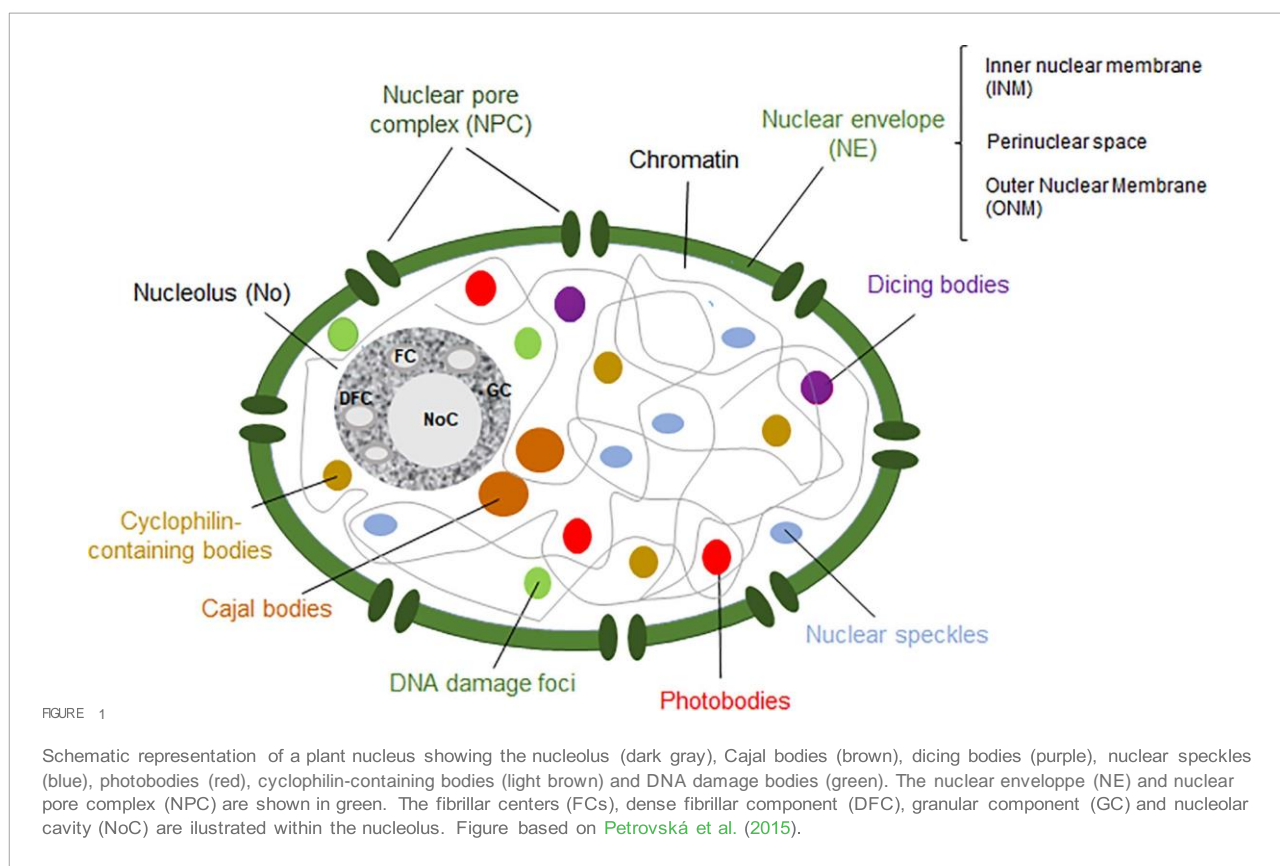
are exclusive to each cell type. On the one hand, nuclear bodies in human cells include the nucleolus, promyelocytic leukemia nuclear bodies, nuclear speckles, paraspeckles, Cajal bodies (CBs), Sam68 bodies, as well as other non-characterized nuclear bodies (Fong et al., 2013). On the other hand, plant nuclear bodies include the nucleolus, CBs, dicing bodies (D-bodies), nuclear speckles, cyclophilin-containing speckles, photobodies and DNA damage foci (Figure 1; Guo and Fang, 2014; Petrovská et al., 2015; Emenecker et al., 2020). Nuclear bodies can be distinguished according to their composition (Table 1) and/or in terms of their assembly mechanism. Three different models have been proposed to describe the formation of nuclear bodies. Firstly, a stochastic series of events may lead to the assembly of nuclear bodies, in which the structural process is mainly random. Moreover, nuclear bodies can also be formed by a coordinated mechanism, in which each element is incorporated into the nuclear body after the other following a tightly sequential order. In this model, only one or two assembly pathways exist to form the nuclear body. Lastly, nuclear bodies can also follow the seedling assembly mechanism, in which one of their components acts as a seed to initiate and nucleate the formation of the nuclear body (Mao et al., 2011; Matera et al., 2011; Guo and Fang, 2014). Whereas the formation of the nucleolus is governed by a seedling mechanism, where the nascent rRNAs act as the “seed” (Karpen et al., 1988), CBs have been observed to follow a stochastic events among the different components (Kaiser et al., 2008).

In this review, we will consider the different plant nuclear bodies, from their protein and RNA composition to the functions and processes they are involved in. Moreover, the dynamics of nuclear bodies upon different stressors in animals, yeast and plants will be also addressed. In addition, the features of the different protein signals that govern nuclear and nucleolar import and export will be detailed. This review also focuses on how heat stress and main post-translational modifications (PTMs) modulate nuclear import and export.

Nuclear bodies

Nucleolus

The nucleolus is the most prominent subnuclear structure in eukaryotic cells. It is considered to be a nuclear body because of the presence of protein and RNA molecules. The nucleolus possesses a tripartite composition distributed in a vectorial fashion: fibrillar center (FC), dense fibrillar component (DFC) and granular component (GC), although in yeast the nucleolus exhibits a bipartite organization (Sáez-Vásquez and Gadal, 2010). The FCs are low-density areas surrounded by the DFC, which is embedded in the GC. Moreover, it is also common to find nucleolar cavities in plant nucleoli (Jordan, 1984; Sirri et al., 2008). Moreover, the plant nucleoli differentiate



between two types of FCs: homogeneous (similar to animals) and heterogeneous. Homogeneous FCs, composed of a fibrous loose material, are small and numerous, generally abundant in cells actively producing ribosomes. In contrast, heterogeneous FCs are large and scarce, associated with low translation rates ([Sáez-Vásquez and Medina, 2008](#)).

The nucleolus is largely known for its role in the biogenesis of ribosomes ([Mélèse and Xue, 1995](#)). Ribosome biogenesis begins with the transcription of rRNA genes (rDNA) into pre-rRNA molecules. The 5.8S, 18S and 25S (28S in mammals) rRNA genes are located in tandem in polycistronic rDNA units (called 35S, 45S, and 47S rDNA in yeast, plants and mammals, respectively) and transcribed by the RNA polymerase I (RNA Pol I). A fourth rRNA gene, the 5S, is transcribed by the RNA Polymerase III (RNA Pol III; [Campbell et al., 1992](#)). RNA Pol I and III are multimeric enzymatic complexes composed of up to ~15 subunits ([Guilfoyle and Dietrich, 1987](#); [Haag and Pikaard, 2007](#); [Fernández-Tornero et al., 2013](#); [Ream et al., 2015](#)).

Transcribed pre-rRNAs (47S/45S/35S and 5S) undergo several processing steps by endo- and exonucleases to form the mature 5S, 5.8S, 18S and 25S (28S in mammals) rRNAs. In plants, this include endonucleases RTL2 (ribonuclease 3-like protein 2) and 5¹-3¹ and 3¹-5¹ exoribonuclease activities from XRN2 and the exosome, respectively (reviewed by [Sáez-Vásquez and Delseny, 2019](#)). In the processing of rRNA, both C/D

and H/ACA small nucleolar ribonucleoproteins (snoRNPs) also play a central role in the modifications of the rRNAs. On the one hand, C/D snoRNPs mediate 2¹-O-methyl ribose methylation of rRNAs, in which fibrillarin has been described in many species as the methyltransferase. In contrast to animals and yeast, two different genes encode fibrillarin in *Arabidopsis thaliana*, referred as Arabidopsis (AtFIB1 and AtFIB2; [Barneche et al., 2000](#); [Pih et al., 2000](#)). Approximately 120 sites experiencing 2¹-O-methyl ribose methylation have been described in Arabidopsis ([Azevedo-Favory et al., 2021](#)). On the other hand, H/ACA snoRNPs mediate 5-ribosyluracil pseudouridylation of rRNAs. In Arabidopsis, dyskerin conforms the catalytic subunit of the H/ACA complex ([Maceluch et al., 2001](#)). Dyskerin is encoded by a single gene in Arabidopsis, AtNAP57. In addition, rRNA molecules can be subjected to base methylations, such as m⁷G, m⁶A, m³U, m⁵C or Ac⁴C ([Sloan et al., 2017](#); [Taoka et al., 2018](#)).

The mature 5S, 5.8S and 25S rRNAs (28S rRNA in mammals), along with large ribosomal proteins (RPLs), form the large ribosomal particle (60S), while the small ribosomal particle (40S) contains the 18S rRNA plus small ribosomal proteins ([Fromont-Racine et al., 2003](#); [Korostelev and Noller, 2007](#); [Weis et al., 2015](#); [Sáez-Vásquez and Delseny, 2019](#)).

The number of nucleolar proteins in Arabidopsis is significantly lower than in humans ([Palm et al., 2016](#);

TABLE 1 Plant nuclear bodies.

Nuclear body	Protein components	RNA components	Function
Nucleolus (No)	Nucleolin Fibrillarin ...	rRNAs C/D snoRNAs H/ACA snoRNAs	Ribosome biogenesis Regulation of the cell cycle Stress response
Cajal bodies (CBs)	Coilin U2B Fibrillarin Dyskerin AGO4, DCL3	snRNAs C/D snoRNAs H/ACA snoRNAs scaRNAs mRNAs	Modification of sRNAs Formation of spliceosomal particles Gene silencing (Arabidopsis)
Dicing bodies (D-bodies)	DCL1 HYL1 AGO1 DRB1 HEN1	-	Gene silencing
Nuclear speckles	SR-rich proteins snRNPs non-snRNPs Transcription factors 3' processing factors Cyclophilins	snoRNAs pre-mRNAs	Formation of spliceosomal particles
Cyclophilin-containing bodies	Cyclophilins (BypRS64)	-	Protein folding Plant development and signaling
Photobodies	Phytochromes Cryptochromes	-	Storage of active phytochromes Degradation of phytochromes
DNA damage foci	γH2AX RBR1 RAD51, RAD54 E2F	-	DNA damage response (DDR)

The proteins and RNAs listed are considered to be major (abundance) components and normally used as compartmental markers.

Montacé et al., 2017). In Arabidopsis, the vast majority are proteins involved in the transcription and processing of the rRNAs (Leung et al., 2003; Pendle et al., 2005; Montacé et al., 2017). Nucleolin is the most abundant non-ribosomal protein in the nucleolus in eukaryotic cells (Ginisty et al., 1999; Tajrishi et al., 2011; Durut and Sáez-Vásquez, 2015). Even though it is required for ribosome biogenesis, nucleolin participates in other functions such as DNA replication, mRNA stability and translation or maintenance of the chromatin (Roger et al., 2003; Kim et al., 2005; Takagi et al., 2005; Angelov et al., 2006). In Arabidopsis, two different nucleolins are found (AtNUC-L1/NUC1 and AtNUC-L2/NUC2), showing structural homology with animal and yeast nucleolins (Pontvianne et al., 2007; reviewed by Durut and Sáez-Vásquez, 2015).

The assembly and organization of the nucleolus are governed by liquid–liquid phase separation (LLPS; Lafontaine et al., 2021). LLPS consists of the spontaneous demix of a solution into several phases that coexist. Thus, the three nucleolar subdomains (FC, DFC, and GC) behave as three different coexisting liquid phases (reviewed by Emenecker et al., 2020). Many nucleolar proteins, such as fibrillarin or nucleolin, are capable of condensing through LLPS, feature designated as multivalency (reviewed by Banani et al., 2017). These proteins possess Gly-Arg-rich (GAR) domains as well

as intrinsically disordered regions (IDRs), which have been observed to promote LLPS. Interestingly, the protein content of the DFC and GC generate immiscibility owing to disfavorable interaction. Thus, the nucleolus can be seen as a multilayered condensate whose formation is governed by LLPS (reviewed by Lafontaine et al., 2021).

During mitosis, the nucleus disappears in the majority of the eukaryotes, including animal and plants. The nucleolus disassembles in the early mitosis, becoming completely lost in the prometaphase. Nevertheless, some nucleolar components appear to be associated with the periphery of the chromosomes during the metaphase and anaphase as sheath-like structures. In the telophase, the sheath-like material forms the perinucleolar bodies (PNBs), which are recruited by the nucleolar organizer regions (NORs). This fact, along with the transcription of the rRNA genes, promotes the synthesis of new nucleoli in the daughter cells (Ochs et al., 1985). Moreover, transcription from Alu elements generates the so-called aluRNAs, which are necessary to maintain the nucleolar integrity (Caudron-Herger et al., 2015). These elements have not yet been described in plants. Exceptionally, many fungi, including yeast, undergo closed mitosis in which the nuclear structures are present throughout mitosis (Asakawa et al., 2016).

Cajal bodies and histone locus bodies

Cajal bodies are among the best-characterized nuclear bodies in animal, yeast and plant cells. They were discovered by Ramon y Cajal along with other nuclear bodies (reviewed by Nizami et al., 2010). These dynamic structures are able to fuse and divide. Moreover, they are also associated with the nucleolus, moving in or out of it (Andrade et al., 1991; Beven et al., 1995; Boudonck et al., 1999). Whereas the absence of CBs causes developmental abnormalities and lethality in animals (Liu et al., 2009; Walker et al., 2009; Strzelecka et al., 2010; Kanno et al., 2016), CBs are not essential for plant viability (Collier et al., 2006; Nizami et al., 2010). The protein and RNA content of these bodies is very diverse and extensive. The main component of CBs is coilin (Collier et al., 2006), followed by small nuclear RNPs (snRNPs) involved in the processing of pre-mRNAs, such as U2B (Beven et al., 1995); and small nucleolar RNPs (C/D and H/ACA snoRNP) involved in the processing of rRNAs, tRNAs and snRNAs (Ogg and Lamond, 2002; Kannan et al., 2008), as well as signaling pathways (Love et al., 2017). In plants, proteins involved in gene silencing are also part of CBs, such as AGO4 or DCL3 (Li et al., 2006; Pontes and Pikaard, 2008). Coilin is required for the formation of CBs, as shown in knockout and knockdown mutants in some species, i.e., *Arabidopsis thaliana* or *Mus musculus* (Tucker et al., 2001; Collier et al., 2006). The structure of the coilin shows homology across several species: it possesses two nuclear localization signals (NLS), one predicted nucleolar localization signal (NoLS), an N-terminal globular domain and a C-terminal Tudor-like structure (Makarov et al., 2013). Furthermore, three RNA species are localized in CBs: snRNAs, snoRNAs and small CB-specific RNAs (scaRNAs). Interestingly, plant CBs also contain poly(A) RNAs, such as mRNAs (Kim et al., 2010; Niedojadlo et al., 2014). Contrary to other types of nuclear bodies, CBs are dynamic because of a continuous exchange of their components (reviewed by Nizami et al., 2010).

Because of the diverse composition of CBs, these nuclear bodies take part in numerous functions. One of the most important processes involving CBs is the formation of spliceosomal particles (snRNPs). After being synthesized in the nucleoplasm, they are translocated to the cytosol to interact with Sm proteins. After methylation of the 5' of the snRNAs, the snRNP complex moves back into the nucleus (Suzuki et al., 2010). Another function revolving CBs is the modification of small RNAs (sRNAs). The presence of C/D box snoRNAs and scaRNAs mediates the 2'-O-methyl ribose methylation of snRNAs, whereas H/ACA box snoRNAs promote 5-ribosyluracil pseudouridylation of RNA molecules in CBs (Bassett, 2012). The possible role of CBs in telomerase activity has been hypothesized. In *Arabidopsis*, the telomerase interacts with dyskerin, which is a component of CBs. Moreover, the telomerase in invertebrates possesses a domain

that leads to accumulation in CBs (reviewed by Love et al., 2017).

Some functions attributed to CBs are specific to plants, i.e., the nonsense-mediated mRNA decay (NMD), a quality control mechanism for premature terminated mRNA molecules. Whereas this process takes place in the cytosol in human cells, the nonsense-mediated mRNA decay might occur in the nucleolus in plants. The nucleolar localization of the exon junction complex, mRNA molecules and Up-frameshift factors in plants sparked the idea of CBs involved in the nonsense-mediated mRNA decay. Nevertheless, this hypothesis must be fully demonstrated (Pendle et al., 2005; Trinkle-Mulcahy, 2009; Bassett, 2012). Another plant-specific function of these bodies is gene silencing. Several components of the gene-silencing machinery have been observed to co-localize with components of CBs (Li et al., 2006, 2008; Fujioka et al., 2007).

Histone locus bodies are another type of nuclear body involved in the processing of histone pre-mRNA, as they are associated with histone-coding genes. In fact, these bodies resemble CBs in terms of structure and composition. HLBs were first discovered in *Drosophila melanogaster* and human cells, even though they were considered to be CBs (Frey and Matera, 1995; Liu et al., 2006; Bongiorno-Borbone et al., 2008; Ghule et al., 2008; Nizami et al., 2010). As CBs, HLBs also contain coilin. The difference resides in the fact that coilin is not essential for the assembly of HLBs, in contrast to CBs (Love et al., 2017). However, these bodies have not been described in plants.

Dicing bodies

MicroRNAs (miRNAs) are a type of RNase III-dependent sRNAs involved in gene silencing. They are transcribed by the RNA Polymerase II (RNA Pol II) as pri-miRNAs, which are processed by the DCL1-HYL-SE complex into a duplex miRNA. Then, these duplexes associate with ARGONAUTE proteins to form the RISC complex in order to exert their function (reviewed by Liu et al., 2012). In *Arabidopsis*, DCL1 was found to form round structures in the nucleus. These bodies were able to diffuse around the nucleoplasm, but they were not associated with the nucleolus. In addition, HYL1 also forms aggregates in the nucleus, co-localizing with DCL1 bodies. Similarly, SE forms aggregates in the nucleus. However, they do not always co-localize with DCL1 and HYL, as SE is also found in nuclear speckles (Fang et al., 2004; Fang and Spector, 2007; Song et al., 2007). What is more, the DCL1-HYL1 bodies are different from CBs owing to the absence of coilin. Thus, these DCL1-HYL1 bodies were named D-bodies (Fang and Spector, 2007). Other proteins co-localizing with D-bodies include AGO1, HEN1, DRB1, and PIF4 (reviewed by Emenecker et al., 2020). The formation of D-bodies is also governed by LLPS. It was observed that SE forms droplets, followed by the presence of HYL, DCL1 and pri-pre-miRNAs. The absence of SE inhibits the formation

of D-bodies, which indicates that these bodies are formed via SE-phase separation (Xie et al., 2021).

Nuclear speckles and paraspeckles

Nuclear speckles constitute another type of common nuclear body present in animal and plant cells (Reddy et al., 2012). These bodies are located in the interchromatin space, and they store splicing factors, as well as snRNPs, non-snRNPs, transcription factors and 3¹ processing factors (Lamond and Spector, 2003). These speckles are normally found near active transcription sites, where pre-mRNA molecules have been also found forming fibers (Spector and Lamond, 2011).

Serine/arginine (SR)-rich proteins are splicing proteins involved in recognition of pre-mRNA introns and in the assembly of the spliceosome (Lorković et al., 2004). The arginine/serine (RS)-rich motif present in these proteins, apart from having an NLS, has been experimentally demonstrated to be responsible for the accumulation in the nuclear speckles (Tillemans et al., 2005). The number, size and shape of nuclear speckles in plant nuclei vary according to the metabolic stage, transcriptional activity or cell type. For instance, actively transcribing cells have numerous small nuclear speckles, whereas inhibition of transcription leads to the formation of larger and less numerous nuclear speckles (Reddy et al., 2012). Interestingly, the co-localization of SR-rich proteins within nuclear speckles also depends on the cell type and/or environmental conditions of the plant cell. What is more, it has been observed that the co-localization of proteins in nuclear speckles does not imply physical interaction among them (Lorković et al., 2008; Reddy et al., 2012). There is a continuous interchange of components between the nuclear speckles and the nucleoplasm (Rausin et al., 2010). The presence of Arabidopsis SR31, SR1 and atSRp30 in nuclear speckles was demonstrated (Fang et al., 2004). Even though the main components of these bodies are SR proteins, the precise composition of the nuclear speckles continuously changes (Reddy et al., 2012).

Paraspeckles have been described in animals and are composed of non-coding RNA molecules and proteins. They have not been described in plants (Spector and Lamond, 2011; Reddy et al., 2012).

Cyclophilin-containing bodies

Cyclophilins are a family of proteins that are present in many organelles in plant cells (Singh et al., 2020). They are believed to be involved in protein folding, possibly mediating the assembly of the spliceosome. Recently, it has been observed that cyclophilins constitute versatile proteins that exert a wide array of functions in plant development and signaling (Schmid, 1995;

Lorković et al., 2004; Singh et al., 2020). It was observed that CypRS64, a member of the cyclophilins in Arabidopsis, formed certain bodies in the nucleus named cyclophilin-containing bodies. This protein contains three different domains: (i) the PPIase motif, (ii) the KRS motif, and (iii) the RS/SP domain. The localization of CypRS64 in the cyclophilin-containing bodies requires both the KRS and the RS/SP domains. It is also known that cyclophilins interact with SR proteins, which form part of the nuclear speckles. When CypRS64 was co-expressed with one of its interactors, CypRS64 translocated into the nuclear speckles. It has been hypothesized that this re-localization allows the gathering of different proteins involved in the same process. In addition, the phosphorylation of the CypRS64-interacting proteins is necessary in order to associate with CypRS64 (Lorković et al., 2004).

Photobodies

Phytochromes (phys) are photoreceptors responsible for the red (R) and far-red (FR) sensing (Schäfer et al., 1972). They possess inactive and active conformations, referred as R light-absorbing Pr and FR light-absorbing Pfr forms, respectively. In Arabidopsis, there are five types of phys (phyA–phyE). Among them, the most prominent in Arabidopsis are phyA, which senses R, FR and blue light, and phyB, which responds to R light (van Buskirk et al., 2012). Interestingly, the conversion from Pr to Pfr leads to the translocation of phys from the cytosol into the nucleus (Kircher et al., 1999; Kim et al., 2000; Rockwell et al., 2006). Not only are these receptors located in the nucleus upon light excitation, but they also form nuclear bodies, named photobodies (Yamaguchi et al., 1999). The formation of photobodies occurs during the dark-to-light transition. Photobodies containing both phyA and phyB can be observed a few minutes after exposure to R light. These photobodies, named “early photobodies,” disappear after 1 h of exposure. After 2 h of exposure to R light, novel photobodies called “late photobodies” are formed. PhyA is no longer present in the “late photobodies,” as they have been degraded because of continuous exposure to light (Kircher et al., 1999; Yamaguchi et al., 1999; Kim et al., 2000; Bauer et al., 2004). The size and number of phyB photobodies are determined by the amount of phyB in the Pfr form under continuous exposure to R light. Upon high-intensity R light, the Pfr form is predominant, leading to the formation of large photobodies. On the other hand, dim R light generates smaller and more dispersed photobodies, because of the conversion of Pfr into Pr (Chen et al., 2003). In addition, some phyB photobodies may contain cryptochromes, which are blue light receptors (reviewed by van Buskirk et al., 2012). Apart from photoreceptors, the composition of the photobodies includes transcription factors, such as Constans (CO) and the B-box transcription factor 28 (BBX28),

or the E3 ligase constitutively photomorphogenic 1 (COP1; Liu et al., 2014, 2020).

Structurally, phyB is able to form either homodimers or heterodimers. Each monomer possesses a N-terminal domain to sense light, and a C-terminal domain to allow dimerization (Clack et al., 2009; Nagatani, 2010). The C-terminus of phyB is required for the formation of photobodies independently of light. The proline-rich domain (PRD) domain, in the C-terminus, might contain an NLS or be able to bind a nuclear protein. Notably, the whole C-terminal domain (PRD and histidine kinase-related domain (HKRD) subdomains) is required for the formation of photobodies under normal conditions (Matsushita et al., 2003; Chen et al., 2005). It has been proposed that in the inactive Pr form, the C-terminal domain of phyB is hidden and masked by the N-terminal domain. The transition to the active Pfr form allows the exposure of the NLS in the C-terminal domain to form the photobodies (Fankhauser and Chen, 2008). Regarding the function of photobodies, several hypotheses have been proposed due to the heterogeneous composition: (i) they may act as a storage site for active photoreceptors, (ii) they could be sites of protein degradation, since many proteins are localized in photobodies prior to degradation, or (iii) they could act in transcriptional regulation, as many transcriptional regulators are present (reviewed by van Buskirk et al., 2012). Using a nucleolus-tethering system (NoTS) to dive into the assembly of the photobodies, it was observed that any of the components of the photobodies is able to trigger the formation and assembly of these bodies. Thus, the formation of the photobodies follows a stochastic pathway, also referred as self-organized assembly (Liu et al., 2014).

DNA damage foci

Because of the sessile character of plants, they are highly exposed to several adverse conditions that lead to DNA damage. However, the mutation rate is very low due to the existence of reparation mechanisms. One of these is called the DNA damage response (DDR), which is highly conserved among animals and plants. The DDR starts with the activation of the protein kinases ataxia telangiectasia mutated (ATM) by double strand breaks (DSBs), and ATM- and Rad3-related (ATR) proteins by single strand (SS) DNA. Then, the suppressor of gamma-response 1 (SOG1) is phosphorylated, promoting the transcription of DNA repair genes and the regulation of the cell cycle. However, there is a SOG1-independent DDR, which involves E2F-retinoblastoma-related protein 1 (RBR1) complexes (reviewed by Nisa et al., 2019).

DNA damage foci appear at sites of double-stranded damage in animals, yeast and plants. It has been hypothesized that LLPS, similar to the nucleolus, governs the formation of these bodies (reviewed by Emenecker et al., 2020). The protein components of DNA damage foci are involved in DDR.

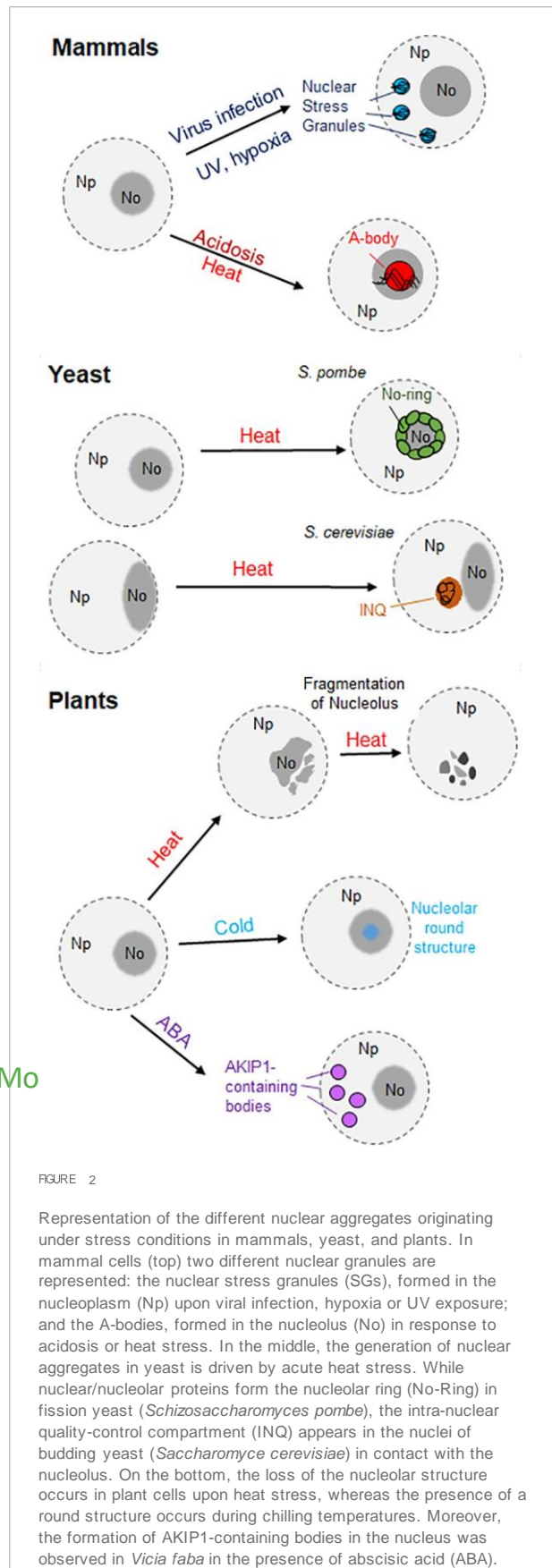
One of those proteins is the phosphorylated histone H2AX (γ H2AX). This histone variant accumulates at DNA damage sites, becoming an excellent marker for DNA damage foci (Löbrich et al., 2010). In Arabidopsis, the E2F transcription factors also form DNA damage foci that co-localize with γ H2AX. In addition, RBR1 co-localizes with E2Fa in DNA damage foci, recruiting proteins involved in DNA repair (Lang et al., 2012; Biedermann et al., 2017). Other components of the DNA damage foci include radiation-sensitive (RAD) proteins, such as RAD54 (Hirakawa and Matsunaga, 2019) or RAD51 (Kurzbaue et al., 2012). However, according to Singh et al., RAD51 did not co-localized with γ H2AX. Moreover, the gamma-tubulin complex component 3-interacting protein 1 (GIP1), involved in the maintenance of the nuclear structure and organization, forms nuclear foci that co-localize with γ H2AX foci (Singh et al., 2022).

Nuclear granules and bodies under stress

Eukaryotic cells are often exposed to unfavorable conditions, such as extreme temperatures or hypoxia. These stresses activate different types of cellular responses to mitigate and/or fight the adverse conditions (Audas et al., 2016). These stressors have been demonstrated to have an impact on the composition, shape, size and number of nuclear bodies. For instance, the nucleolus undergoes reversible changes in response to low and high temperatures. Plant nucleoli show speckled structures upon incubation at 37°C (Hayashi and Matsunaga, 2019). Moreover, they start to disaggregate and disassemble after longer exposure to 37°C (Darriere et al., 2022). On the other hand, chilling temperatures lead to the formation of a round structure in the nucleolus. It was also observed that both low and high temperatures inhibit the accumulation of newly synthesized rRNA in the nucleolus (Figure 2; Hayashi and Matsunaga, 2019).

High temperatures induce the enlargement of CDK2-containing nuclear speckles, whereas cold treatment inhibits their formation (Kitsios et al., 2008; reviewed by Reddy et al., 2012). CBs also respond to heat shock, since they disappear upon exposure to high temperatures. Nevertheless, they reappear once the heat stress stops (Boudonck et al., 1999). As coilin plays a role in some signaling pathways in plant cells, the involvement of CBs and/or coilin has been suggested in the perception and response to stresses (Love et al., 2017).

Different cellular bodies are formed in response to abiotic stress. In the cytosol of plant cells, stress granules and heat stress granules appear upon short- and long-term exposure to heat stress, respectively. They can also be differentiated according to their protein and RNA composition (reviewed by Maruri-López et al., 2021). Interestingly, the generation of nuclear aggregates upon diverse stimuli was observed in human cells



and yeast (Figure 2). The nuclear stress granules appear in the nucleoplasm of mammalian cells after exposure to different stimuli, such as hypoxia or UV exposure (Figure 2). However, their existence in plant cells remains uncharacterized (Gaete-Argel et al., 2021; reviewed by Biamonti and Vourc'h, 2010).

Moreover, the amyloid bodies (A-bodies) are formed in the nucleoplasm of human cells in response to various stimuli such as hypoxia, heat stress or acidosis (Figure 2). The protein content of these A-bodies [also referred as Detention Center in Audas et al. (2012a)], is heterogeneous, but all of the proteins share a protein motif known as an amyloid-converting motif (ACM). In addition, the presence of lncRNAs derived from the ribosomal intergenic spacer (IGS) was observed in the A-bodies (Audas et al., 2016). Regarding fungi, the formation of nucleolar rings is attributed to *S. pombe*. Upon acute heat stress, nuclear and nucleolar proteins accumulate in the periphery of the nucleolus (Gallardo et al., 2020). Another example constitutes the “intra-nuclear quality-control compartment” located in the nucleus of *S. cerevisiae* upon heat stress. This nuclear structure, located close to the nucleolus, contains misfolded cytosolic and nuclear proteins (Figure 2; Kaganovich et al., 2008; reviewed by Gallardo et al., 2021).

In Arabidopsis, abiotic stress modulates the composition of certain nuclear bodies. For instance, early flowering 3 (ELF3), a component of the evening complex, forms nuclear speckles in response to high temperatures. The C-terminal prion domain of ELF3 is responsible for this behavior (Jung et al., 2020). In contrast, low and high temperatures promote the disaggregation of phyB photobodies in Arabidopsis Col-0 and *Ler* ecotypes. This disaggregation occurs because of the transition from the active Pfr form to the inactive Pf form of phyB photobodies (Legris et al., 2016; Hahm et al., 2020). On the other hand, the recruitment of the RNA-binding proteins UBA2a and UBA2b to nuclear speckles in Arabidopsis is enhanced upon exposure to abscisic acid, a hormone that mediates the response to some abiotic stresses, such as salinity or drought (Bove et al., 2008). Similarly, the RNA-binding protein AKIP1 forms a plant-specific nuclear body called “AKIP1-containing bodies” in fava bean (*Vicia faba*) upon exposure to abscisic acid (Figure 2; Li et al., 2002).

In order to exert their function, nuclear proteins, synthesized in the cytosol, need to cross the NE through the NPC. As mentioned before, the NPCs are embedded in the nuclear membrane, creating a channel between the cytoplasm and the nucleoplasm. These cylindrical structures constitute the largest macromolecular complexes present in eukaryotic cells. Morphologically, each NPC is composed by a cytoplasmic and a nucleoplasmic ring, connected by eight spokes. A basket-like structure has been observed in the nucleoplasmic side

of the NPC, whereas some fibrillar structures are present in the cytoplasmic face. The main component of the NPC are proteins known as nucleoporins, which are partially conserved in eukaryotes (Nigg, 1997; Stewart, 2007; Hicks, 2013; Petrovská et al., 2015). The NPC regulates protein movement from the cytosol into the nucleus and vice versa. Small molecules can cross the NPC by diffusion. In contrast, larger molecules such as proteins need to be actively translocated in order to cross the NPC (Hicks, 2013). Generally, the exclusion size of the nucleus ranges from 30 to 60 kDa, in which passive transport through the NPC is possible. However, smaller proteins have been observed to cross the NPC by active transport (Timney et al., 2016). Particular amino acid sequences and/or arrangements have been implicated in the active transport of proteins between the cell nucleus and the cytoplasm (Figure 3).

Nuclear import

The nuclear import mechanism includes the movement of proteins from the cytosol into the nucleus. The molecular players in this mechanism have been characterized in eukaryotes. In animals, there are different nuclear import pathways depending on the protein–protein interactions (Stewart, 2007).

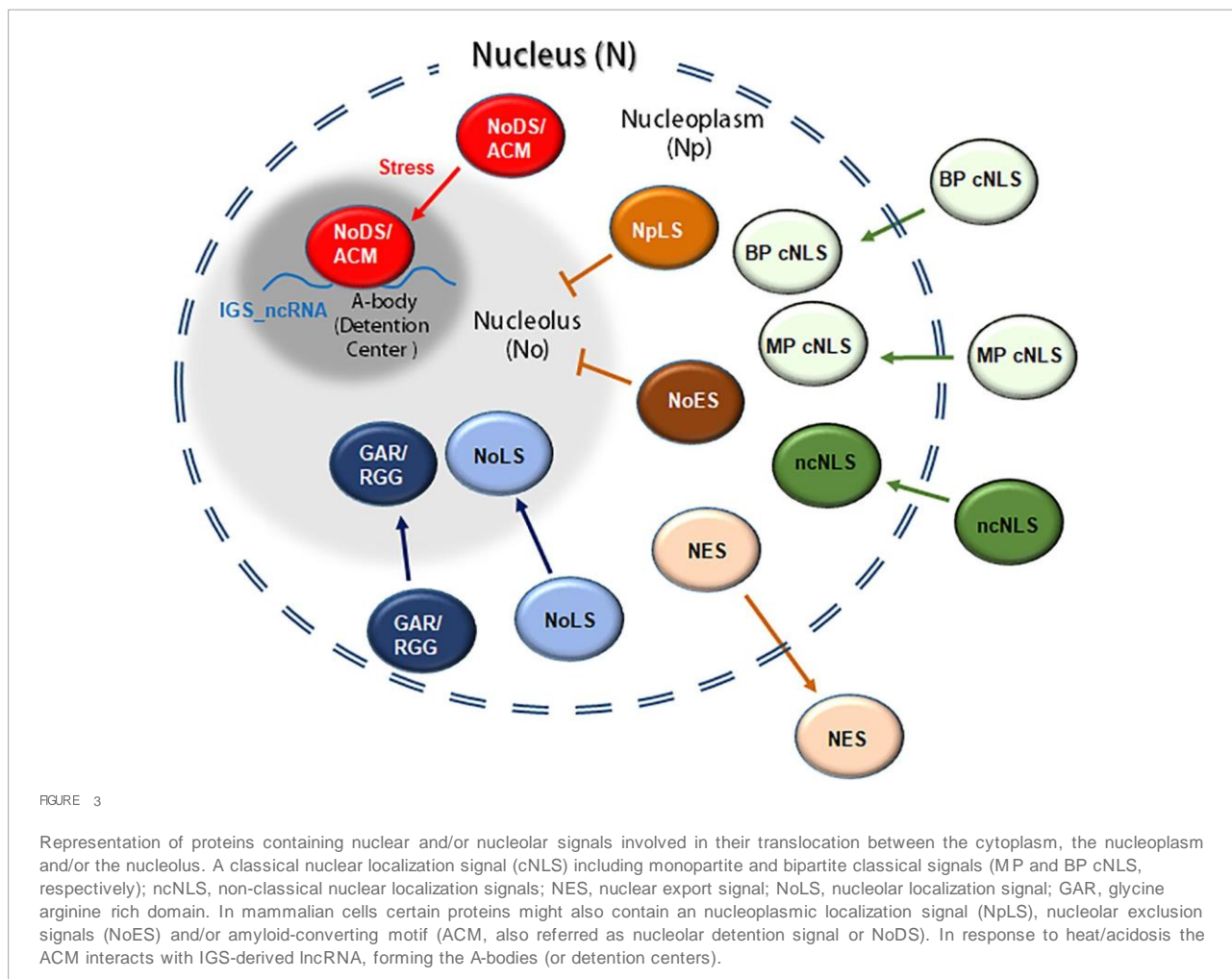
The classical nuclear import pathway includes the cytosolic importin β . This protein interacts with the target nuclear protein through an adaptor protein, importin α , which recognizes a specific motif of the target nuclear protein. The complex importin α –nuclear protein–importin β migrates into the nucleoplasm via interaction with proteins from the NPC. In the nucleus, Ran-GTP promotes the dissociation of importin α , importin β and the target nuclear protein. Importin α moves back into the cytoplasm by interaction with the β -karyopherin CAS and Ran-GTP. At the same time, importin β , along with Ran-GTP, migrates back to the cytoplasm. The hydrolysis of Ran-GTP into Ran-GDP by Ran GTPase-activating protein allows its dissociation from importin β (Smith et al., 1998; Sazer and Dasso, 2000; Hicks, 2013). Finally, the nuclear transport factor 2 (NTF2) mediates the re-importing of Ran-GDP into the nucleus (Kutay and Bischoff, 1997; Ribbeck et al., 1998). In plants, homologs of the components of the machinery have been identified. For instance, Arabidopsis possesses orthologs of importin α , such as At-IMP α and AtKAP α , and importin β (Ballas and Citovsky, 1997; Hübner et al., 1999; Tamura and Hara-Nishimura, 2014). Moreover, orthologs of importin α and importin β have also been found in rice (Matsuki et al., 1998). This suggests that the nuclear import pathway is mostly conserved between animals and plants.

The specific sequence of the target nuclear protein recognized by importin α is known as the nuclear localization signal (NLS). Unlike other localization signals, such as mitochondrial and plastid signals, NLSs are not proteolytically

removed after nuclear import, allowing nuclear proteins to participate in more than one round of nuclear transport. These localization signals can be found in the N- and/or C-terminus, as well as within the protein (Martoglio and Dobberstein, 1998; Lu et al., 2021). Adam et al. described the first NLS in the simian virus 40 large T-antigen (SV40), comprising seven hydrophobic residues ($^{126}\text{PKKKRKV}^{132}$). Sequence analysis of other nuclear proteins revealed the presence of NLSs. These signals can be classified in different groups according to their structure and composition. The first class includes the classical NLSs (cNLSs), which are the most characterized. This class can also be subdivided into two categories: monopartite and bipartite cNLSs. Monopartite cNLSs are composed of 4–8 basic amino acids, at least four of them being positively charged (lysine or arginine). The consensus sequence for this subgroup of cNLSs is K-K/R-X-K/R, where X represents any residue (Dingwall and Laskey, 1991; Lange et al., 2010; Lu et al., 2021). One of the most notorious and best-characterized members of the monopartite cNLSs subgroup is SV40. Functional analysis of this NLS revealed that the third lysine ($^{126}\text{PKKKRKV}^{132}$) is necessary for the correct nuclear localization of the protein (Kalderon et al., 1984). The second subgroup within the cNLSs is the bipartite cNLSs, having two clusters of positively charged amino acid residues separated by a spacer of 9–12 residues. The consensus sequence of this subgroup is R/K-X_(9–12)-K-R-X-K, where X represents any amino acid. The protein nucleoplasmin exemplifies the possession of a bipartite cNLS ($^{155}\text{KRPAATKKAGQA}^{170}$), where the two positively charged clusters are underlined; Smith et al., 1995; Lange et al., 2010; Lu et al., 2021). Lange et al. demonstrated the importance of the length of the spacer in the bipartite cNLSs, it being crucial in the interaction with importin α . By testing different spacer lengths of a bipartite cNLS, they concluded that the longer the spacer is, the less nuclear accumulation is observed.

The second class of NLSs is known as non-classical NLSs (ncNLSs). This type comprises NLSs whose composition varies from positively charged residues. The best-known ncNLSs are the proline-tyrosine (PY) ncNLSs, which are composed of 20–30 residues with a basic or hydrophobic N-terminus and a common C-terminal motif ([basic/hydrophobic]-X_n-R/H/K-X_(2–5)-P-Y, where X represents any amino acid; Wang et al., 2012; Mallet and Bachand, 2013). Nevertheless, other ncNLSs cannot be represented as a consensus sequence, such as the ribosomal protein L23a (Jäkel and Görlich, 1998). Additionally, Lu et al. designated a third class of miscellaneous NLSs, including (i) proteins with a potential NLS, predicted *in silico*, that do not lead to nuclear localization; (ii) NLSs recognized upon protein dimerization; (iii) cryptic NLSs, where a stimulus is necessary for the translocation into the nucleus; and (iv) proteins with multiple NLSs, all of them required for nuclear import.

In plants, several NLSs have been characterized both *in silico* and experimentally. Examples include the E3



ubiquitin-protein ligase COP1 in Arabidopsis, a repressor of the photomorphogenesis. This protein exhibits nuclear localization owing to the presence of a bipartite cNLS (²⁹⁴RKKRIHAQFNDLQECYLQKRRQLA³¹⁷; Stacey and Von Arnim, 1999). Another example is the Arabidopsis transcriptional elongation regulator MINIYO, which possesses two NLSs in its sequence. One of them is a monopartite cNLS located in the N-terminus (²⁵⁴LKKRKH²⁵⁹), whereas the other is a bipartite cNLS present in the C-terminus (¹⁴⁰¹RKRHREGMMLDLLRYKK¹⁴¹⁷; Contreras et al., 2019). Arabidopsis RTL2 contains a bipartite cNLS in the C-terminal portion (³⁷¹KKAESSAYHMRALRK³⁸⁷; Comella et al., 2008). In maize, three different NLSs are present in the protein R. Two of them are monopartite cNLSs (¹⁰⁰CDRRAAPARP¹⁰⁹ located in the N-terminus, and ⁴¹⁹MSERKRREKL⁴²⁸ found within the sequence). The third NLS is a Mat α 2-type NLS, named after the unusual NLS of the Mat α 2 protein in yeast (Hall et al., 1984), located in the C-terminus (⁵⁹⁸MISESLRKAICKR⁶¹⁰; Shieh et al., 1993; Hicks et al., 1995). Finally, the first 43 amino acids of the *Brassica napus* 60S ribosomal protein L13-1 are sufficient to target this protein to the nucleus. The NLS is likely to be present

between the residues 29 and 43 (²⁹RKTRRRVARQKKAVK⁴³, Table 2; Sáez-Vásquez et al., 2000).

Nuclear export

In contrast to nuclear import, the mechanisms that govern the movement of proteins from the nucleoplasm to the cytoplasm have been much less characterized. There are few examples of transporters of proteins from the nucleoplasm into the cytoplasm. One of them is Exportin1 (CRM1), which is known to specifically interact with nuclear proteins in order to translocate them into the cytoplasm (Fornerod and Ohno, 1997; Ossareh-Nazari et al., 1997). For this, CRM1 interacts with a specific sequence of the nuclear protein, as well as with Ran-GTP. This complex interacts with the NPC, crossing the NE and reaching the cytosol. The complex dissociates when Ran-GTP is hydrolyzed into Ran-GDP (Bischoff and Görlich, 1997; Neville et al., 1997). This nuclear transporter has also been found in Arabidopsis, named AtXPO1. Similarly, AtXPO1 interacts with a specific motif of the nuclear proteins and

TABLE 2 Nuclear localization signals.

Protein	Organism	Sequence ¹	Type	References
Nuclear localization signals (NLS)				
COP1	<i>Arabidopsis thaliana</i>	294 <u>KKRIHAQFNDLQECYLQKRRQLA</u> ³¹⁷	BP cNLS	Stacey and Von Arnim (1999)
MINIYO	<i>Arabidopsis thaliana</i>	254 <u>LKKRKH</u> ²⁵⁹	MP cNLS	Contreras et al. (2019)
		1401 <u>RKRHREGMMLDLLRYKK</u> ¹⁴¹⁷	BP cNLS	
RTL2	<i>Arabidopsis thaliana</i>	371 <u>KKAESSAYHMRALRK</u> ³⁸⁷	BP cNLS	Comella et al. (2008)
Hsf1	<i>Arabidopsis thaliana</i>	230 <u>KEKKSFLGLDVGRKRR</u> ²⁴⁵	BP cNLS	Evrard et al. (2013)
Coilin	<i>Arabidopsis thaliana</i>	175 <u>KRKK</u> ¹⁷⁸	MP cNLS	Makarov et al. (2013)
		264 <u>KKAKR</u> ²⁶⁸	MP cNLS	
Protein R	<i>Zea mays</i>	100 CDRRAAPARP ¹⁰⁹	MP cNLS	Hicks et al. (1995)
		419 MSERKRREKL ⁴²⁸	MP cNLS	
		598 MISESLRKAICKR ⁶¹⁰	ncNLS	
60S ribosomal protein L13-1	<i>Brassica napus</i>	29 <u>RKTRRRVARQKKA</u> VK ⁴³	N/S	Sáez-Vásquez et al. (2000)
Nucleoplasmin	<i>Xenopus laevis</i>	155 <u>KRPAATKKAGQAKKKK</u> ¹⁷⁰	BP cNLS	Dingwall et al. (1988)
ERK5	<i>Homo sapiens</i>	505 <u>RKPVTAQERQREERKRRRRQERA</u> KEREKRRQERE ⁵³⁹	BP cNLS	Kondoh et al. (2006)
CCTα	<i>Homo sapiens</i>	12 <u>RKRRK</u> ¹⁶	MP cNLS	Taneva et al. (2012)
SV40	Simian virus	126 <u>PKKKRKV</u> ¹³²	MP cNLS	Adam et al. (1989)
N protein	Porcine reproductive and respiratory syndrome virus	10 <u>KRRK</u> ^{13 7}	MP cNLS	Rowland and Yoo (2003)
		41 <u>PGKKNKK</u> ⁴	MP cNLS	

cNLS, classical nuclear localization signal; ncNLS, non-classical nuclear localization signal; MP, monopartite; BP, bipartite; N/S, non-specified. ¹The two positively charged amino acid clusters of the BP cNLSs are underlined.

with the protein Ran in order to export the nuclear proteins (Haasen et al., 1999). It has been suggested that the nuclear proteins react with phenylalanine and glycine (FG)-repeats of the nucleoporins from the NPC in order to reach the cytoplasm (Stutz et al., 1996).

The specific region of the nuclear proteins that is recognized by CRM1/AtXPO1 is called the nuclear export signal (NES). The existence of NESs was initially described in the HIV-1 Rev protein (⁷⁵LPPLERLTLD⁸⁴), involved in the export of pre-mRNAs and mRNAs from the nucleus, and the heat-stable inhibitor (PKI) of cAPK (³⁸LALKLAGLDI⁴⁷; Fischer et al., 1995; Wen et al., 1995). These NESs are rather short and hydrophobic sequences where Leu is highly present (Nigg, 1997). The presence of NESs is mainly characterized in proteins involved in the export of RNA molecules from the nucleus. For instance, the translocation of the 5S rRNA to the cytosol is mediated by TFIIA in *Xenopus* oocytes, which contains a Rev-like NES (³⁵⁷SLVLDKLT³⁶⁵; Guddat et al., 1990). In *Arabidopsis*, a Rev-like NES is located in the C-terminus of RanBP1a (¹⁷¹DTAGLLEKLTVEETKTEKT¹⁹⁰; Haasen et al., 1999). Besides a bipartite cNLS, *Arabidopsis* RTL2 also possesses an NES in the N-terminus (⁷PEYNFPAITRCSLSNSLP²⁶). The presence of an NLS and NES allows RTL2 to move between the cytosol and the nucleus (Table 3; Comella et al., 2008). The characterization of later NESs allowed grouping the majority of the NESs into three different consensus sequences: ϕ -X_{1,2}-[Δ P]-

ϕ -[Δ P]_{2,3}- ϕ -[Δ P]- ϕ (class 1), ϕ -[Δ P]- ϕ -[Δ P]₂- ϕ -[Δ P]- ϕ (class 2), and ϕ -X-[Δ P]- ϕ -[Δ P]₃- ϕ -[Δ P]₂- ϕ (class 3), where ϕ represents large hydrophobic residues, X_{1,2} represents any one or two amino acids, [Δ P] represents any amino acid except proline, and [Δ P]_{2,3} represents any two or three amino acids except proline (Kosugi et al., 2014).

Bidirectional signals

Another type of **localization** signal that mediates nucleocytoplasmic trafficking has been described. In contrast to the NLSs or NESs, which exert a unilateral translocation of proteins, the nucleocytoplasmic shuttling signals (NSSs) allow both import and export of proteins to/from the nucleus in human cells. Many NSS-containing proteins interact with mRNA molecules. These motifs are longer and lack basic residues (Michael, 2000). The first NSS was found in the human heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1), which associates with pre-mRNA and mRNA molecules. The domain responsible for the bidirectional behavior of A1 is called M9, present in the C-terminus (see Table 3 for sequence; Siomi and Dreyfuss, 1995). Other examples include the C-terminal ZNS domain of the DAZ-associated protein 1 (DAZAP1; see Table 3 for sequence; Lin and Yen, 2006) or the C-terminal nuclear transport domain (NTD) of the human RNA helicase

TABLE 3 Nuclear export signals and nucleocytoplasmic shuttling signals.

Protein	Organism	Sequence	Reference
Nuclear export signals (NES)			
RanBP1a	<i>Arabidopsis thaliana</i>	¹⁷¹ DTAGLLEKLTVEETKTEKT ¹⁹⁰	Haasen et al. (1999)
RTL2	<i>Arabidopsis thaliana</i>	⁷ PEYNFPAITRCSLSNSLPHR ²⁶	Comella et al. (2008)
TFIIIA	<i>Xenopus laevis</i>	³⁵⁷ SLVLDKLT ³⁶⁵	Guddat et al. (1990)
PKI	<i>Homo sapiens</i>	³⁸ LALKLAGLD ⁴⁷	Wen et al. (1995)
N protein	Porcine reproductive and respiratory syndrome virus	¹⁰⁶ LPTHHTVRLIRV ¹¹⁷	Rowland and Yoo (2003)
Rev protein	HIV-1	⁷⁵ LPPLERLTLD ⁸⁴	Fischer et al. (1995)
Nucleocytoplasmic Shuttling Signals (NSS)			
hnRNP A1	<i>Homo sapiens</i>	³¹⁶ GNYNNQSSNFPGPMKGGNFGRRSSGPYGGGGQYFAKPRNQGGY ³⁵⁷	Siomi and Dreyfuss (1995)
DAZAP1	<i>Homo sapiens</i>	³⁸³ GPPAGSGSGFRGQNHNVQGFHPYRR ⁴⁰⁷	Lin and Yen (2006)
RNA helicase A	<i>Homo sapiens</i>	¹¹⁵¹ GSTRYGDGPRPPKMARYDNGSGYRRGGSSYSGGGYGGGYSSGG YGGSGYGGGSANSFRAGYGAGVGGGYRGVSRGGFRGNSGGDYRGPS GGYRGSGGFQRRGGGRGAYGTGY ¹²⁶⁰	Tang et al. (1999)

A (see **Table 3** for sequence; Tang et al., 1999). M9-dependent nuclear import is conferred by transportin. Transportin binds directly to the M9 domain, binds to nucleoporins and is sufficient for the movement of M9-containing cargo through the NPC and into the nucleus. This transport seems to be dependent on the Ran-GTP concentration (reviewed by Michael, 2000). To our knowledge, NSSs have not been described in plant proteins.

Moving into and out of the nucleolus

Nucleolar localization signal

Once in the nucleus, proteins can diffuse in the nucleoplasm or migrate into the nucleolus and/or other bodies. Unlike nuclear targeting, the mechanism and localization signals that regulate the translocation of proteins into the nucleolus remain highly unexplored (**Figure 3**). Over the years, several proteins have been predicted to possess a nucleolar localization signal (NoLS). These signals are rich in basic amino acids, especially Lys and Arg, preferentially located in the C- or N-terminus of proteins. Moreover, the NoLSs are predicted to be present in alpha helices or random coils located on the surface of the protein (Scott et al., 2010). For a peptide to achieve nucleolar localization it must be positively charged, formed exclusively of six or more arginines, and with an isoelectric point above 12.6 (Martin et al., 2015).

The existence of NoLSs have been predicted *in silico* and experimentally confirmed. For instance, both isoforms of the Arabidopsis ribosomal protein RPL23a (RPL23aA and RPL23aB) accumulate in the nucleolus owing to an NLS/NoLS region (Degenhardt and Bonham-Smith, 2008).

An NoLS was also found in the sequence of Arabidopsis coilin (²⁰²KKKKKKK²⁰⁸), as well as two NLSs (¹⁷⁵KRKK¹⁷⁸ and ²⁶⁴KKAKR²⁶⁸; Makarov et al., 2013). Likewise, the human coilin also presents one NoLS (¹⁶⁰KKNKRN¹⁶⁸), which was experimentally confirmed to be necessary for nucleolar localization (Hebert and Matera, 2000). Moreover, the sequence of the breast cancer autoantigen nucleolar GTP-binding protein 2 (NGP-1) contains two NoLSs. One of them is located in the N-terminus (see **Table 4** for sequence), whereas the second one is a C-terminal NoLS (see **Table 4** for sequence; Datta et al., 2015). Interestingly, the nucleocapsid (N) protein of the porcine reproductive and respiratory syndrome virus (PRRSV) exhibits both cytosolic and nucleolar localization. This dual accumulation is achieved by the presence of two NLSs (¹⁰KRRK¹³ and ⁴¹PGKKNK⁴⁷), one NoLS (⁴¹PGKKNKKNPEKPHFLATEDDVRHHFTPSER⁷²) and one NES (¹⁰⁶LPTHHTVRLIRV¹¹⁷; **Table 4**; Rowland and Yoo, 2003).

It is widely thought that the nucleolar localization of many proteins is the result of the association with nucleolar components, such as rRNA (Schmidt-Zachmann and Nigg, 1993; Scott et al., 2010). In mammal cells, many ribosomal components adopt nucleolar localization by interaction with B23, the major constituent of the granular component of the nucleolus (Borer et al., 1989). It was proposed that B23, because of its constant movement between the nucleolus and the cytosol, is able to shuttle NoLS-containing proteins into the nucleolus (Borer et al., 1989; Scott et al., 2010). Nevertheless, this hypothesis has not been confirmed. What has been observed is that the ADP-ribosylation factor GTPase-activating protein 1 (ARF GAP 1) accumulates in the nucleolus because of the interaction with B23 (Korgaonkar et al., 2005; Sirri et al., 2008). Interestingly, human nucleolin has not been demonstrated to possess an NoLS, but it does contain a bipartite

TABLE 4 Nucleolar localization signals and nucleolar exclusion signals.

Protein	Organism	Sequence	Reference
Nucleolar Localization Signals (NoLS)			
Coilin	<i>Arabidopsis thaliana</i>	202 KKKKKK ²⁰⁸	Makarov et al. (2013)
RPL23a ¹	<i>Arabidopsis thaliana</i>	33 KKDK ³⁶ 36 KKIR ³⁹	Degenhardt and Bonham-Smith (2008)
Coilin	<i>Homo sapiens</i>	160 KKNKRNL ¹⁶⁸	Hebert and Matera (2000)
NGP-1 ²	<i>Homo sapiens</i>	¹ MVKPKYKGRSTINPSKASTNPDRVQAGGQNMRRDRIIRLNM YRQKERRNSRGKIKPLQYQSTVASGTVARVEPNIKWFGNTRVIKQ SSLQKFQEEMD ¹⁰⁰ ⁶³¹ DEKIAYQKFLDKAKAKKFSAVRISKGLSEKIFAKPEEQRTLEED VDDRAPSCKGKKRKAQREEEQEHNSKAPRALTSKERRRAVRQQR KKVGVRYETHNVKNRNRNKKKTNDSEGQ KHKRKKFRQKQ ⁷⁰¹	Datta et al. (2015)
N protein	Porcine reproductive and respiratory syndrome virus	⁴¹ PGKKNKKKNPEKPHPLATEDDVRHHFTPSER ⁷²	Rowland and Yoo (2003)
Nucleolar Export Signals (NoES)			
GNL3L ³	<i>Homo sapiens</i>	²⁹² EVYLDKFIRLLDAPGIVPGPNSEVGTILRNCVHVQKLADPVTVPVET ILQRCNLEEISNYGVSGFQTTEHFLTAVAHRLGKKKGGGLYSQEQAAK AVLADWVSGKISFYPPATHLPTLSAEIVKEMTEVFDIEDTEQAN EDTMECLATGESDELLGDTDPLEMEIKLLHSPMTKIADAIENKTTVYKI GDLTGCTNPNRHQMGWAKRNVDRPKSNSMVDVCSVDRRSVLQRI METDP ⁵³¹	Meng et al. (2007)
NGP-1 ³	<i>Homo sapiens</i>	³⁴⁹ QYITLMRRIFLIDCPGVVYPSSEDTIVLKGVVQVEKIKSPEDHIG AVLERAKPEYISKYKIDSWENAEDFLEKLAFTGKLLKGGEPDLQTV GKMVLNDWQGRIPFFVKPPNAEPLVAPQLPSSSLEVVPAAQNNP GEEVTETAGEGESIEETEENSHCDANTEMQQLTRVRQNFQKINV VPQFSGDDLVPVEVSDLEEEESFSEEEEEQEQRDDAEESSEPEEE NVGNDTKAVIKAL DEKIAYQKFLDKAKAKKFS ⁶²⁰	Meng et al. (2007)
TdIF2/ERBP	<i>Homo sapiens</i>	⁴⁴¹ VLLVL ⁴⁴⁵	Fukada et al. (2019)

¹ These sequences are described as NLS/NoLS.

² The N- and C-terminal regions are responsible for the nucleolar accumulation; the sequence of the NoLS is not detailed.

³ These sequences are named nucleoplasmic localization signals (NpLSs).

cNLS (²⁵⁶KKKKEMANKSAPEAKKKK²⁷³). The structure of the nucleolin was analyzed to determine which domain is responsible for its nucleolar accumulation (Créancier et al., 1993; Schmidt-Zachmann and Nigg, 1993). It was proposed that the GAR domain located in its C-terminus is necessary for its localization in the nucleolus (Pellar and DiMario, 2003). The localization of a GAR-deleted nucleolin is mainly nuclear, decreasing its nucleolar accumulation. Nevertheless, there are other domains that are also necessary for the translocation of nucleolin into the nucleolus, such as the RNA recognition motif (RRM). Thus, it was proposed that it interacts with nucleolar components such as rRNA to accumulate in the nucleolus (Doron-Mandel et al., 2021; Okuwaki et al., 2021). Similarly, the N-terminal GAR domain present in the human fibrillarin was demonstrated to drive both nuclear and nucleolar accumulation (Snaar et al., 2000; Shubina et al., 2020). Whereas fibrillarin is located in the nucleolus, GAR-deleted fibrillarin is distributed in the nucleolus, nucleoplasm and cytoplasm. Moreover, the methylation of the arginine residues of the GAR domain positively regulates nuclear localization. However, this methylation decreases the nucleolar accumulation

of fibrillarin (Shubina et al., 2020). As previously stated, two nucleolin (AtNUC-L1/NUC1 and AtNUC-L2/NUC2) and fibrillarin (AtFIB1 and AtFIB2) protein genes were described in *Arabidopsis*. Both nucleolin and fibrillarin proteins contain GAR domains in the C-terminus or N-terminus sequences, respectively. In contrast, the N-terminal region of AtNUC-L1/NUC1 contains two potential bipartite NLSs, whereas there is only one NLS in AtNUC-L2/NUC2 (Barneche et al., 2000; Pontvianne et al., 2007).

Nucleoplasmic localization signal

Two different types of signals that prevent nucleolar accumulation have been characterized in humans (Meng et al., 2007; Fukada et al., 2019). On the one hand, Meng et al. coined the term nucleoplasmic localization signal (NpLS) to describe the regions of the guanine nucleotide-binding protein-like 3-like protein (GNL3L) and the nucleolar GTP-binding protein 2 (NGP-1) that prevented localization in the nucleolus and promoted nucleoplasmic accumulation.

Fukada et al. also identified a region in the terminal sequence of the deoxynucleotidyltransferase-interacting factor 2/estrogen receptor α -binding protein (TdiF2/ERBP) that led to the similar nucleoplasmic localization, called the nucleolar exclusion signal (NoES). Whereas the described NpLSs represent rather large regions (>200 residues), the NoES is only composed of five hydrophobic residues. Each of the NpLSs contains five hydrophobic residues, which indicates that the NpLS and NoES are structurally similar (Table 4; Fukada et al., 2019).

In terms of composition, NoLSs and NLSs have a similar composition, since both contain basic residues. In some cases, nucleolar signals were initially predicted to be nuclear signals because of their similarity. Thus, the experimental validation of predicted nucleolar signals, as well as nuclear signals, is fundamental to fully characterizing the nuclear and/or nucleolar localization of a protein. A classification of NLSs and NoLSs consists of (i) NLS-only signals, responsible for nuclear localization, (ii) NoLS-only signals, which determine localization exclusively in the nucleolus, and (iii) joint NLS–NoLS regions, which lead to accumulation of proteins in the nucleus and the nucleolus (Scott et al., 2010). Likewise, leucine is a common residue of NESs and NoESs. Nevertheless, the NoES characterized by Fukada et al. (⁴⁴¹VLLVL⁴⁴⁵) cannot be included in any of the consensus NESs described above.

Nuclear and nucleolar accumulation under heat stress: Amyloid-converting motif/nucleolar detention signal

Part of the heat stress response includes inhibition or induction of specific protein activities. For that, transcription factors (TFs) are responsible of transforming the perception of the stressor into the expression of key genes. More specifically, heat stress transcription factors (HSFs) play a central role in gene transcription under different abiotic stresses, including heat stress (Guo et al., 2016). Consequently, the nuclear proteome undergoes substantial changes upon high temperatures, promoting the accumulation of HSFs. For instance, *Arabidopsis* bZIP18 and bZIP52, which are present in the cytoplasm under normal conditions, accumulate in the nucleus under heat stress. This nuclear localization is provoked by the dephosphorylation of Ser residues (Wiese et al., 2021). In contrast, phosphorylation of *Arabidopsis* HsfA2 promotes nuclear accumulation under heat stress. The phosphorylated residue consists of threonine (Thr249) located close to a bipartite cNLS (²³⁰**KEKKSLFGLDVGRKRRLT**S²⁴⁹, where the NLS is in bold, and Thr249 is underlined; Evrard et al., 2013). Another example includes the *Arabidopsis* heat shock factor-binding protein (AtHSBP), which shuttles

from the cytosol to the nucleus in response to heat stress (Hsu et al., 2010). Nevertheless, HSFs are not the only proteins that accumulate into the nucleus upon heat stress. *Arabidopsis* heat-intolerance 4 (HIT4), involved in the release from transcriptional gene silencing, translocates from the chromocentres to the nucleolus under heat stress (Wang et al., 2015).

This nuclear translocation upon heat stress has also been described in human cells. As mentioned previously (see the section: Nuclear granules and bodies under stress), some proteins are retained in the nucleus, forming A-bodies in response to heat stress. This phenomenon was initially called nucleolar sequestration of proteins, including the heat shock protein 70 (Hsp70) or the E3 ubiquitin-protein ligase MDM2. The ACM, which is the peptide responsible for this behavior, was originally defined as a nucleolar detention signal (NoDS). These signals are characterized by the presence of an arginine motif (R-R-L/I) and two or more hydrophobic triplets (L- ϕ -L/V, where ϕ represents a hydrophobic residue). Likewise, there is a physical interaction between the NoDS and IGS- derived lncRNAs, transcribed in response to high temperatures (Mekhail et al., 2007; Audas et al., 2012a,b). These NoDS/ACM signals have not been functionally characterized in plants.

In silico prediction tools for nuclear and nucleolar signals

The existence of consensus sequences of localization signals, such as NLSs or NESs, allows their prediction using computational methods. There is a wide array of online platforms and *in silico* methods to predict the existence of NLSs from an amino acid sequence. For instance, “NLStradamus” is based on a Hidden Markov Model and used to find cNLSs from yeast sequences (Nguyen Ba et al., 2009). Cokol et al. (2000) created “predictNLS” by performing *in silico* mutagenesis of a library of 91 experimentally tested NLSs. Similarly, “NESmapper” allows the detection of NESs with high accuracy and a low false positive rate. To do that, every residue of the NES is considered, contributing independently and additively to the nuclear exportation (Kosugi et al., 2014). In the case of “NLSdb,” the server allows the prediction of NLSs and NESs from nine different species, including *Arabidopsis thaliana*, *Homo sapiens* and *Oryza sativa* (Bernhofer et al., 2018).

On the other hand, the prediction of nucleolar signals is challenging. First of all, the only well-established nucleolar signal is the NoLS, since few nuclear export signals have been described (see the above NoESs and NpLSs). Moreover, NoLSs are considerably similar to NLSs, both containing arginine and lysine residues. The web server “NoD” appears to be the best tool to predict an NoLS. It predicts the presence of an NoLS from the protein sequence, using the human-trained artificial neural network. Even though “NoD” performs best using mammal and

mammalian-infecting viral proteins, it can also be used with plant and plant virus proteins (Scott et al., 2010, 2011).

Post-translational modifications and nuclear localization

Post-translational modifications consist of the addition of functional groups, or even cleavage, of certain domains. These modifications change the properties of proteins, promoting functional diversity. The best-known PTMs conform the addition of chemical groups, which can be reversible or irreversible, including phosphorylation, methylation, acetylation, or ubiquitination. These modifications have an impact on many aspects of the protein, i.e., activity, half-life, interaction with other molecules, or subcellular localization (Ramazi and Zahiri, 2021). More specifically, certain PTMs are able to modulate the accumulation of certain proteins in the nucleus.

Phosphorylation

Phosphorylation is the best-characterized PTM. The addition of a phosphate negatively or positively influences the function of a protein. To couple and uncouple phosphate groups, kinases and phosphatases are necessary, respectively. Phosphorylation can also enhance or inhibit nuclear transport through different mechanisms (Nardozi et al., 2010). The first case portrays the accumulation in the nucleus after protein phosphorylation. As exemplified by Nardozi et al., phosphorylation can occur in the NLS or in the surrounding sequence. Phosphorylation can also induce conformational changes in the NLS-containing protein, exposing the NLS and enhancing nuclear accumulation. For instance, phosphorylation upstream the cNLS of SV40 (¹¹⁰PSSDDEAAADSQHAAPPKKKRKVG¹³³, where the NLS is marked in bold, and the phosphorylation sites are underlined) enhances nuclear import (Xiao et al., 1997). On the other hand, phosphorylation of proteins lacking NLSs can also increase nuclear accumulation. Upon phosphorylation of the TEY domain of ERK1/2 (²³²LDQLNHILGILGSPSQEDL²⁵⁰, the nuclear transport signal sequence, where the phosphorylated residue is underlined) or the RS domain of ASF/SF2, these proteins shuttle from the cytosol into the nucleus (reviewed by Nardozi et al., 2010). What is more, the phosphorylation of the TEY domain of ERK5 promotes nuclear accumulation. This protein contains a bipartite cNLS and non-classical NES, in which phosphorylation benefits the nuclear import over cytoplasmic localization (Table 2; Kondoh et al., 2006; Nardozi et al., 2010).

Several examples in which phosphorylation enhances nuclear import in plants can be found in the literature. For

instance, the Arabidopsis ssDNA binding protein WHIRLY (WHY1) is present in the nucleus and in the chloroplast under normal conditions. When WHY1 is phosphorylated by the calcineurin B-like-interacting protein kinase 14 (CIPK14), it accumulates predominantly in the nucleus (Ren et al., 2017). Another example is the movement of 14-3-3 proteins from the cytosol to the nucleus upon phosphorylation by the cold-activated plasma membrane protein cold-responsive protein kinase 1 (CRPK1; Liu et al., 2017). The second scenario is the inhibition of the nuclear import upon phosphorylation. Nardozi et al., presented different cases and examples. On the one hand, phosphorylation of nucleoporins of the NPC can repress nuclear import (Porter and Palmenberg, 2009). On the other hand, the phosphorylation of the NLS provokes cytoplasmic accumulation. As mentioned above, Arabidopsis bZIP18 and bZIP52, which are present in the cytoplasm under normal conditions, accumulate in the nucleus under heat stress. This nuclear localization is provoked by the dephosphorylation of Ser residues (Wiese et al., 2021). In contrast, phosphorylation of Arabidopsis HsfA2 promotes nuclear accumulation under heat stress. The phosphorylated residue consists of threonine (Thr249) located close to a bipartite cNLS (²³⁰**KEKKSLFGLDVGRKRRLTST**²⁴⁹, where the NLS is in bold, and Thr249 is underlined; Evrard et al., 2013).

In lymphocyte T cells, the nuclear factor of activated T cells (NFAT) has an NLS and NES. Low Ca²⁺ levels cause phosphorylation of the SRR2 region, which overlaps with the NLS, promoting cytoplasmic localization of the NFAT. On the other hand, nuclear accumulation of the NFAT has been observed with high levels of Ca²⁺ owing to dephosphorylation (Kehlenbach et al., 1998; Belfield et al., 2006). In Arabidopsis, the phytochrome-interacting factor 7 (PIF7) accumulates in the cytosol under white-light conditions because of its phosphorylation. In contrast, shade exposure activates phosphatases, de-phosphorylating PIF7 and promoting its accumulation in the nuclear photobodies (Huang et al., 2018). Similarly, the activity of the transcription factor Brassinazole-resistant 1 (BZR1) is linked to its phosphorylation status. Brassinosteroid-insensitive-2- (BIN2-) mediated phosphorylation of BZR1 promotes nuclear export of BZR1 and cytosolic accumulation. Either de-phosphorylation or mutation of the putative phosphorylation sites in BZR1 results in nuclear accumulation (Ryu et al., 2007).

Acetylation and methylation

The addition of an acetyl group is catalyzed by acetyltransferases using acetyl CoA as a cofactor. In contrast, the removal of the acetyl group is catalyzed by deacetylases. Protein acetylation becomes necessary in certain situations, such as protein-protein interaction, chromatin stability or nuclear transport (Yang and Seto, 2008; Choudhary et al., 2009;

Xia et al., 2020). Some examples in the literature exemplify the role of acetylation in nuclear protein accumulation. For instance, the human tyrosyl-tRNA synthetase (TyrRS) becomes highly acetylated upon oxidative stress, promoting its nuclear accumulation (Cao et al., 2017). Similarly, the acetylation of the Lys10 of the human translational corepressor CtBP2 by the nuclear acetylase p300 is essential for its nuclear localization (Zhao et al., 2006). In contrast, the acetylation of the NLS of some proteins prevents nuclear accumulation. This is the case of the human tyrosine-protein kinase c-Abl, in which the acetylation of the Lys730 within its second NLS (⁷²⁸SSKRFLR⁷³⁴, in which the Lys730 is underlined) promotes cytosolic localization rather than nuclear import (di Bari et al., 2006).

From another perspective, the acetylation of a component of the nuclear import machinery has an impact on the actual nuclear import rate in Arabidopsis. For instance, acetylation of the Lys18 of Nup50, which promotes the dissociation of importin α from the nuclear protein, decelerates this dissociation, repressing the nuclear import (Matsuura and Stewart, 2005; Füßl et al., 2018).

Protein methylation is a reversible PTM that occurs in the nucleus. Even though many residues can be methylated, this PTM is most common in arginine and lysine (reviewed by Ramazi and Zahiri, 2021). The effects of methylation in nuclear accumulation are diverse and specific for each protein. On the one hand, the human Hsp70 localizes in the nucleus when the Lys561 is dimethylated, accumulating in the cytosol when unmethylated (Cho et al., 2012). On the other hand, the methylation of the Lys494 of Yes-associated protein (Yap) by a SET-domain-containing lysine methyltransferase prevents its nuclear import, remaining in the cytosol in mouse (Oudhoff et al., 2013).

SUMOylation and Ubiquitination

The attachment of a small ubiquitin-like modifier (SUMO) to proteins involves three different enzymes: (i) the activating enzyme or E1, (ii) the conjugating enzyme or E2, and (iii) the ligase or E3. This PTM is crucial in several processes, such as the regulation of the cell cycle, subcellular localization or transcription (reviewed by Hannoun et al., 2010). In the case of the human polo-like kinase 1 (PLK1), the SUMOylation of the Lys492, which is close to one of the NLSs, is essential for the nuclear accumulation of PLK1, as well as increasing its stability (Wen et al., 2017). Similarly, the SUMOylation of the Lys248 of the human X-linked zinc finger transcription factor ZIC3 is important for its nuclear retention (Chen et al., 2013). As with acetylation, the SUMOylation of the Lys909 of the yeast importin Kap114 is essential for its role in the nuclear import mechanism (Rothenbusch et al., 2012). In Arabidopsis, heat stress increases the amount of SUMOylated proteins in the

nucleus, suggesting that SUMOylation induces nuclear import (Saracco et al., 2007).

Ubiquitination is a reversible PTM in which the C-terminus of an active ubiquitin is attached to a protein. Even though ubiquitination can occur in all 20 amino acids, it is more frequent in lysine residues. Similar to SUMOylation, the ubiquitin junction requires three enzymes: (i) the activating enzyme or E1, (ii) the conjugating enzyme or E2, and (iii) the ligase or E3. This reversible modification is normally associated with protein degradation via ubiquitin-proteasome. However, some effects in nuclear translocation have been described (Lecker et al., 2006; Bhogaraju and Dikic, 2016; Ramazi and Zahiri, 2021). The ubiquitination of the Lys57 of the human cytidylyltransferase (CCT α) promotes cytosolic accumulation. This ubiquitination event occurs near the N-terminal NLS of CCT α , disrupting its interaction with importin α (Trotman et al., 2007). Similarly, the monoubiquitination of two lysine residues (Lys13 and Lys289) in PTEN is necessary for its correct nuclear import, where it exerts its role as tumor suppressor (Trotman et al., 2007). Moreover, ubiquitination is also attributed to nuclear export. Upon proteasome inhibition, ubiquitinated proteins accumulate in the cytosol. This accumulation results from the transport of the ubiquitinated proteins from the nucleus to the cytosol (Hirayama et al., 2018). On the other hand, how ubiquitination affects nuclear and/or nucleolar import and/or export in plants remains mainly uncharacterized.

Perspectives

The purpose of this review is to gather global information concerning nuclear bodies in plants and, in particular, to survey their composition (proteins and RNA) and the peptide or amino acid sequence/structure signals involved in their localization and assembly. We did not intend to present an exhaustive catalog of protein and molecular bases involved in the assembly of nuclear bodies in plants, but rather to establish the current state of the art of these bodies in response to environmental conditions in plants and, more specifically, in response to abiotic stresses. Comparison of the behaviors of conserved nucleolar bodies revealed certain functional and structural similarities in yeast, animal and plant cells. However, under specific environmental conditions, particular nuclear bodies are formed and/or reorganized distinctly. For instance, this pertains to the nucleolus under heat stress conditions. Thus, although under optimal growth conditions nuclear bodies might have similar functions, key differences might appear upon specific developmental and environmental conditions. This is particularly true for plants, which are sessile organisms subjected to major developmental programs (including seed germination and flowering) and constrained to adapt to or

resist stressful conditions (biotic and abiotic) to survive. The functional, structural and molecular clues of these bodies remain elusive and deserve further study to better understand the underlying molecular mechanism of nuclear bodies in plants.

Author contributions

EM-D and JS-V wrote the review. Both authors contributed to the article and approved the submitted version.

Funding

This work was supported by grants from the ANR, RiboStress 17-CE12-0026-01, MetRibo ANR-20-CE12-0024-01, and a Ph.D. fellowship from CNRS to EM-D (UMR5096-JULSAE-004). This study is set within the framework of the “Laboratoires d’Excellences (LABEX)” TULIP (ANR-10-LABX-41) and of the “Ecole Universitaire de Recherche (EUR)” TULP- GS (ANR-18-EURE-00019).

References

- Adam, S. A., Loblt, T. J., Mitchell, M. A., and Gerace, L. (1989). Identification of specific binding proteins for a nuclear location sequence. *Nature* 337, 276–279. doi: 10.1038/337276a0
- Alberts, B., Johnson, A., and Lewis, J. (2002). *The compartmentalization of cells, in molecular biology of the cell*, 4th Edn. New York, NY: Garland Science.
- Andrade, L. E. C., Chan, E. K. L., Raska, I., Peebles, C. L., Roos, G., and Tan, E. M. (1991). Human autoantibody to a novel protein of the nuclear coiled body: Immunological characterization and cDNA Cloning of p80-coilin. *J. Exp. Med.* 173, 1407–1419. doi: 10.1084/jem.173.6.1407
- Angelov, D., Bondarenko, V. A., Almagro, S., Menoni, H., Mongéard, F., Hans, F., et al. (2006). Nucleolin is a histone chaperone with FACT-like activity and assists remodeling of nucleosomes. *EMBO J.* 25, 1669–1679. doi: 10.1038/sj.emboj.7601046
- Asakawa, H., Yang, H. J., Hiraoka, Y., and Haraguchi, T. (2016). Virtual nuclear envelope breakdown and its regulators in fission yeast meiosis. *Front. Cell Dev. Biol.* 4:5. doi: 10.3389/fcell.2016.00005
- Audas, T. E., Audas, D. E., Jacob, M. D., Ho, J. J. D., Khacho, M., Wang, M., et al. (2016). Adaptation to stressors by systemic protein amyloidogenesis. *Dev. Cell* 39, 155–168. doi: 10.1016/j.devcel.2016.09.002
- Audas, T. E., Jacob, M. D., and Lee, S. (2012a). Immobilization of proteins in the nucleolus by ribosomal intergenic spacer noncoding RNA. *Mol. Cell* 45, 147–157. doi: 10.1016/j.molcel.2011.12.012
- Audas, T. E., Jacob, M. D., and Lee, S. (2012b). The nucleolar detention pathway: A cellular strategy for regulating molecular networks. *Cell Cycle* 11, 2059–2062. doi: 10.4161/cc.20140
- Azevedo-Favory, J., Gaspin, C., Ayadi, L., Montacé, C., Marchand, V., Jobet, E., et al. (2021). Mapping rRNA 2'-O-methylations and identification of C/D snoRNAs in *Arabidopsis thaliana* plants. *RNA Biol.* 18, 1760–1777. doi: 10.1080/15476286.2020.1869892
- Ballas, N., and Citovsky, V. (1997). Nuclear localization signal binding protein from *Arabidopsis* mediates nuclear import of agrobacterium VirD2 protein. *Proc. Natl. Acad. Sci. U.S.A.* 94, 10723–10728. doi: 10.1073/pnas.94.20.10723
- Banani, S. F., Lee, H. O., Hyman, A. A., and Rosen, M. K. (2017). Biomolecular condensates: Organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* 18, 285–298. doi: 10.1038/nrm.2017.7

Acknowledgments

We thank Carlos Fernández Tornero for the detailed critical reading of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Barneche, F., Steinmetz, F., and Echeverri la, M. (2000). Fibrillarin genes encode both a conserved nucleolar protein and a novel small nucleolar RNA involved in ribosomal RNA methylation in *Arabidopsis thaliana*. *J. Biol. Chem.* 275, 27212–27220. doi: 10.1074/jbc.M002996200

Bassett, C. L. (2012). Cajal bodies and Plant RNA metabolism. *CRC. Crit. Rev. Plant Sci.* 31, 258–270. doi: 10.1080/07352689.2011.645431

Bauer, D., Viczián, A., Kircher, S., Nobis, T., Nitschke, R., Kunkel, T., et al. (2004). Constitutive photomorphogenesis 1 and multiple photoreceptors control degradation of phytochrome interacting factor 3, a transcription factor required for light signaling in *Arabidopsis*. *Plant Cell* 16, 1433–1445. doi: 10.1105/tpc.021568

Belfield, J. L., Whittaker, C., Cader, M. Z., and Chawla, S. (2006). Differential effects of Ca²⁺ and cAMP on transcription mediated by MEF2D and cAMP-response element-binding protein in hippocampal neurons. *J. Biol. Chem.* 281, 27724–27732. doi: 10.1074/jbc.M601485200

Bernhofer, M., Goldberg, T., Wolf, S., Ahmed, M., Zaugg, J., Boden, M., et al. (2018). NLSdb-major update for database of nuclear localization signals and nuclear export signals. *Nucleic Acids Res.* 46, D503–D508. doi: 10.1093/nar/gkx1021

Beven, A. F., Simpson, G. G., Brown, J. W. S., and Shaw, P. J. (1995). The organization of spliceosomal components in the nuclei of higher plants. *J. Cell Sci.* 108, 509–518. doi: 10.1242/jcs.108.2.509

Bhogaraju, S., and Dikic, I. (2016). Ubiquitination without E1 and E2 enzymes. *Nature* 533, 43–44. doi: 10.1038/nature17888

Biamonti, G., and Vourc’h, C. (2010). Nuclear stress bodies. *Cold Spring Harb. Perspect. Biol.* 2:a000695. doi: 10.1101/cshperspect.a000695

Biedermann, S., Harashima, H., Chen, P., Heese, M., Bouyer, D., Sofroni, K., et al. (2017). The retinoblastoma homolog RBR1 mediates localization of the repair protein RAD51 to DNA lesions in *Arabidopsis*. *EMBO J.* 36, 1279–1297. doi: 10.15252/embj.201694571

Bischoff, F. R., and Görlich, D. (1997). RanBP1 is crucial for the release of RanGTP from importin β -related nuclear transport factors. *FEBS Lett.* 419, 249–254. doi: 10.1016/s0014-5793(97)01467-1

Bongiorno-Borbone, L., De Cola, A., Vernole, P., Finos, L., Barcaroli, D., Knight, R. A., et al. (2008). FLASH and NPAT positive but not coilin positive

- Cajal bodies correlate with cell ploidy. *Cell Cycle* 7, 2357–2367. doi: 10.4161/cc.6344
- Borer, R. A., Lehner, C. F., Eppenberger, H. M., and Nigg, E. A. (1989). Major nucleolar proteins shuttle between nucleus and cytoplasm. *Cell* 56, 379–390. doi: 10.1016/0092-8674(89)90241-9
- Boudonck, K., Dolan, L., and Shaw, P. J. (1999). The movement of coiled bodies visualized in living plant cells by the green fluorescent protein. *Mol. Biol. Cell* 10, 2297–2307. doi: 10.1091/mbc.10.7.2297
- Bove, J., Kim, C. Y., Gibson, C. A., and Assmann, S. M. (2008). Characterization of wound-responsive RNA-binding proteins and their splice variants in *Arabidopsis*. *Plant Mol. Biol.* 67, 71–88. doi: 10.1007/s11103-008-9302-z
- Campbell, B. R., Song, Y., Posch, T. E., Cullis, C. A., and Town, C. D. (1992). Sequence and organization of 5S ribosomal RNA-encoding genes of *Arabidopsis thaliana*. *Gene* 112, 225–228. doi: 10.1016/0378-1119(92)90380-8
- Cao, X., Li, C., Xiao, S., Tang, Y., Huang, J., Zhao, S., et al. (2017). Acetylation promotes TyrRS nuclear translocation to prevent oxidative damage. *Proc. Natl. Acad. Sci. U.S.A.* 114, 687–692. doi: 10.1073/pnas.1608488114
- Castel, B., and Chae, E. (2021). Nucleocytoplasmic trafficking during immunity. *Mol. Plant* 14, 1612–1614. doi: 10.1016/j.molp.2021.08.008
- Caudron-Herger, M., Pankert, T., Seiler, J., Németh, A., Voit, R., Grummt, I., et al. (2015). Alu element-containing RNAs maintain nucleolar structure and function. *EMBO J.* 34, 2758–2774. doi: 10.15252/embj.201591458
- Chen, L., Ma, Y., Qian, L., and Wang, J. (2013). Sumoylation regulates nuclear localization and function of zinc finger transcription factor ZIC3. *Biochim. Biophys. Acta Mol. Cell Res.* 1833, 2725–2733. doi: 10.1016/j.bbarmac.2013.07.009
- Chen, M., Schwab, R., and Chory, J. (2003). Characterization of the requirements for localization of phytochrome B to nuclear bodies. *Proc. Natl. Acad. Sci. U.S.A.* 100, 14493–14498. doi: 10.1073/pnas.1935989100
- Chen, M., Tao, Y., Lim, J., Shaw, A., and Chory, J. (2005). Regulation of phytochrome B nuclear localization through light-dependent unmasking of nuclear-localization signals. *Curr. Biol.* 15, 637–642. doi: 10.1016/j.cub.2005.02.028
- Cho, H. S., Shimazu, T., Toyokawa, G., Daigo, Y., Maehara, Y., Hayami, S., et al. (2012). Enhanced HSP70 lysine methylation promotes proliferation of cancer cells through activation of Aurora kinase B. *Nat. Commun.* 3:1072. doi: 10.1038/ncomms2074
- Choudhary, C., Kumar, C., Gnäd, F., Nielsen, M. L., Rehman, M., Walther, T. C., et al. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834–840. doi: 10.1126/science.1175371
- Clack, T., Shokry, A., Moffett, M., Liu, P., Faul, M., and Sharrock, R. A. (2009). Obligate heterodimerization of *Arabidopsis* phytochromes C and E and interaction with the PIF3 basic helix-loop-helix transcription factor. *Plant Cell* 21, 786–799. doi: 10.1105/tpc.108.065227
- Cokol, M., Nair, R., and Rost, B. (2000). Finding nuclear localization signals. *EMBO Rep.* 1, 411–415. doi: 10.1093/embo-reports/kvd092
- Collier, S., Pendle, A., Boudonck, K., van Rij, T., Dolan, L., Shaw, P., et al. (2006). A distant coilin homologue is required for the formation of Cajal bodies in *Arabidopsis*. *Mol. Biol. Cell* 17, 2942–2951. doi: 10.1091/mbc.e05-12-1157
- Comella, P., Pontvianne, F., Lahmy, S., Vignols, F., Barbezier, N., DeBures, A., et al. (2008). Characterization of a ribonuclease III-like protein required for cleavage of the pre-rRNA in the 3'ETS in *Arabidopsis*. *Nucleic Acids Res.* 36, 1163–1175. doi: 10.1093/nar/gkm1130
- Contreras, R., Kallem, P., González-García, M. P., Lazarova, A., Sánchez-Serrano, J. J., Sanmartín, M., et al. (2019). Identification of domains and factors involved in MINIYO nuclear import. *Front. Plant Sci.* 10:1044. doi: 10.3389/fpls.2019.01044
- Créancier, L., Prats, H., Zanibellato, C., Amalric, F., and Bugler, B. (1993). Determination of the functional domains involved in nucleolar targeting of nucleolin. *Mol. Biol. Cell* 4, 1239–1250. doi: 10.1091/mbc.4.12.1239
- Darriere, T., Jobet, E., Zavala, D., Escande, M. L., Durut, N., De Bures, A., et al. (2022). Upon heat stress processing of ribosomal RNA precursors into mature rRNAs is compromised after cleavage at primary P site in *Arabidopsis thaliana*. *RNA Biol.* 19, 719–734. doi: 10.1080/15476286.2022.2071517
- Datta, D., Anbarasu, K., Rajabather, S., Priya, R. S., Desai, P., and Mahalingam, S. (2015). Nucleolar GTP-binding protein-1 (NGP-1) promotes G1 to S phase transition by activating cyclin-dependent kinase inhibitor p21 Cip1/Waf1. *J. Biol. Chem.* 290, 21536–21552. doi: 10.1074/jbc.M115.637280
- Degenhardt, R. F., and Bonham-Smith, P. C. (2008). *Arabidopsis* ribosomal proteins RPL23aA and RPL23aB are differentially targeted to the nucleolus and are separately required for normal development. *Plant Physiol.* 147, 128–142. doi: 10.1104/pp.107.111799
- di Bari, M. G., Ciuffini, L., Mingardi, M., Testi, R., Soddu, S., and Barilà, D. (2006). c-Abl acetylation by histone acetyltransferases regulates its nuclear-cytoplasmic localization. *EMBO Rep.* 7, 727–733. doi: 10.1038/sj.embor.7400700
- Dingwall, C., and Laskey, R. A. (1991). Nuclear target sequences – a consensus? *Trends Biochem. Sci.* 16, 478–481. doi: 10.1016/0968-0004(91)90184-w
- Dingwall, C., Robbins, J., Dilworth, S. M., Roberts, B., and Richardson, W. D. (1988). The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *J. Cell Biol.* 107, 841–849. doi: 10.1083/jcb.107.3.841
- Doron-Mandel, E., Koppel, I., Abraham, O., Rishal, I., Smith, T. P., Buchanan, C. N., et al. (2021). The glycine arginine-rich domain of the RNA-binding protein nucleolin regulates its subcellular localization. *EMBO J.* 40:e107158. doi: 10.15252/embj.2020107158
- Durut, N., and Sáez-Vásquez, J. (2015). Nucleolin: Dual roles in rDNA chromatin transcription. *Gene* 556, 7–12. doi: 10.1016/j.gene.2014.09.023
- Emenecker, R. J., Holehouse, A. S., and Strader, L. C. (2020). Emerging roles for phase separation in plants. *Dev. Cell* 55, 69–83. doi: 10.1016/j.devcel.2020.09.010
- Evraud, A., Kumar, M., Lecourieux, D., Lucks, J., von Koskull-Döring, P., and Hirt, H. (2013). Regulation of the heat stress response in *Arabidopsis* by MPK6-targeted phosphorylation of the heat stress factor HsfA2. *PeerJ* 1:e59. doi: 10.7717/peerj.59
- Fang, Y., and Spector, D. L. (2007). Identification of nuclear dicing bodies containing proteins for MicroRNA biogenesis in living *Arabidopsis* plants. *Curr. Biol.* 17, 818–823. doi: 10.1016/j.cub.2007.04.005
- Fang, Y., Hearn, S., and Spector, D. L. (2004). Tissue-specific expression and dynamic organization of SR splicing factors in *Arabidopsis*. *Mol. Biol. Cell* 15, 2664–2673. doi: 10.1091/mbc.e04-02-0100
- Fankhauser, C., and Chen, M. (2008). Transposing phytochrome into the nucleus. *Trends Plant Sci.* 13, 596–601. doi: 10.1016/j.tplants.2008.08.007
- Fernández-Tornero, C., Moreno-Morcillo, M., Rashid, U. J., Taylor, N. M. I., Ruiz, F. M., Gruene, T., et al. (2013). Crystal structure of the 14-subunit RNA polymerase I. *Nature* 502, 644–649. doi: 10.1038/nature12636
- Fischer, U., Huber, J., Boelens, W. C., Mattaj, I. W., and Lührmann, R. (1995). The HIV-1 rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* 82, 475–483. doi: 10.1016/0092-8674(95)90436-0
- Fong, K. W., Li, Y., Wang, W., Ma, W., Li, K., Qi, R. Z., et al. (2013). Whole-genome screening identifies proteins localized to distinct nuclear bodies. *J. Cell Biol.* 203, 149–164. doi: 10.1083/jcb.201303145
- Fornerod, M., and Ohno, M. (1997). CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* 90, 1051–1060. doi: 10.1016/S0092-8674(00)80371-2
- Frey, M. R., and Matera, A. G. (1995). Coiled bodies contain U7 small nuclear RNA and associate with specific DNA sequences in interphase human cells. *Proc. Natl. Acad. Sci. U.S.A.* 92, 5915–5919. doi: 10.1073/pnas.92.13.5915
- Fromont-Racine, M., Senger, B., Saveanu, C., and Fasiolo, F. (2003). Ribosome assembly in eukaryotes. *Gene* 313, 17–42. doi: 10.1016/S0378-1119(03)00629-2
- Fujioka, Y., Utsumi, M., Ohba, Y., and Watanabe, Y. (2007). Location of a possible miRNA processing site in SmD3/SmB nuclear bodies in *Arabidopsis*. *Plant Cell Physiol.* 48, 1243–1253. doi: 10.1093/pcp/pcm099
- Fukada, T., Shibata, S., Ueda, T., Sasaki, K., Shimoda, Y., Senda-Murata, K., et al. (2019). Characterization of nucleolar localization and exclusion signals in terminal deoxynucleotidyltransferase interacting factor 2/estrogen receptor α -binding protein. *Biosci. Biotechnol. Biochem.* 83, 1255–1262. doi: 10.1080/09168451.2019.1591265
- Füßl, M., Lassowskat, I., Née, G., Koskela, M. M., Brünje, A., Tilak, P., et al. (2018). Beyond histones: New substrate proteins of lysine deacetylases in *Arabidopsis* nuclei. *Front. Plant Sci.* 9:461. doi: 10.3389/fpls.2018.00461
- Gaete-Argel, A., Velásquez, F., Márquez, C. L., Rojas-Araya, B., Bueno-Nieto, C., Marín-Rojas, J., et al. (2021). Tellurite promotes stress granules and nuclear SG-like assembly in response to oxidative stress and DNA damage. *Front. Cell Dev. Biol.* 9:622057. doi: 10.3389/fcell.2021.622057
- Gallardo, P., Real-Calderón, P., Flor-Parra, I., Salas-Pino, S., and Daga, R. R. (2020). Acute heat stress leads to reversible aggregation of nuclear proteins into nucleolar rings in fission yeast. *Cell Rep.* 33:108377. doi: 10.1016/j.celrep.2020.108377
- Gallardo, P., Salas-Pino, S., and Daga, R. R. (2021). Reversible protein aggregation as cytoprotective mechanism against heat stress. *Curr. Genet.* 67, 849–855. doi: 10.1007/s00294-021-01191-2
- Ghule, P. N., Dominski, Z., Yang, X. C., Marzluff, W. F., Becker, K. A., Harper, J. W., et al. (2008). Staged assembly of histone gene expression machinery at

- subnuclear foci in the abbreviated cell cycle of human embryonic stem cells. *Proc Natl. Acad. Sci. U.S.A.* 105, 16964–16969. doi: 10.1073/pnas.0809273105
- Ginisty, H., Sicard, H., Roger, B., and Bouvet, P. (1999). Structure and functions of nucleolin. *J. Cell Sci.* 112, 761–772. doi: 10.1242/jcs.112.6.761
- Guddat, U., Bakken, A. H., and Pieler, T. (1990). Protein-mediated nuclear export of RNA: 5S rRNA containing small RNPs in *Xenopus oocytes*. *Cell* 60, 619–628. doi: 10.1016/0092-8674(90)90665-2
- Guilfoyle, T. J., and Dietrich, M. A. (1987). “Plant RNA polymerases: Structures, regulation, and genes,” in *Tailoring genes for crop improvement. Basic life sciences*, eds G. Bruening, J. Harada, T. Kosuge, A. Hollaender, G. Kuny, and C. M. Wilson (Berlin: Springer), 87–100. doi: 10.1007/978-1-4684-5329-4_8
- Guo, M., Liu, J. H., Ma, X., Luo, D. X., Gong, Z. H., and Lu, M. H. (2016). The plant heat stress transcription factors (HSFs): Structure, regulation, and function in response to abiotic stresses. *Front. Plant Sci.* 7:114. doi: 10.3389/fpls.2016.00114
- Guo, T., and Fang, Y. (2014). Functional organization and dynamics of the cell nucleus. *Front. Plant Sci.* 5:378. doi: 10.3389/fpls.2014.00378
- Haag, J. R., and Pikaard, C. S. (2007). RNA polymerase I: A multifunctional molecular machine. *Cell* 131, 1224–1225. doi: 10.1016/j.cell.2007.12.005
- Haasen, D., Ko, C., Hler, È, Neuhaus, G., and Merkle, T. (1999). Nuclear export of proteins in plants: AtXPO1 is the export receptor for leucine-rich nuclear export signals in *Arabidopsis thaliana*. *Plant J.* 20, 695–705. doi: 10.1046/j.1365-313x.1999.00644.x
- Hahn, J., Kim, K., Qiu, Y., and Chen, M. (2020). Increasing ambient temperature progressively disassemble *Arabidopsis* phytochrome B from individual photobodies with distinct thermostabilities. *Nat. Commun.* 11:1660. doi: 10.1038/s41467-020-15526-z
- Hall, M. N., Hereford, L., and Herskowitz, I. (1984). Targeting of E. coli β -galactosidase to the nucleus in yeast. *Cell* 36, 1057–1065. doi: 10.1016/0092-8674(84)90055-2
- Hannoun, Z., Greenhough, S., Jaffray, E., Hay, R. T., and Hay, D. C. (2010). Post-translational modification by SUMO. *Toxicology* 278, 288–293. doi: 10.1016/j.tox.2010.07.013
- Hayashi, K., and Matsunaga, S. (2019). Heat and chilling stress induce nucleolus morphological changes. *J. Plant Res.* 132, 395–403. doi: 10.1007/s10265-019-01096-9
- Hebert, M. D., and Matera, A. G. (2000). Self-association of coilin reveals a common theme in nuclear body localization. *Mol. Biol. Cell* 11, 4159–4171. doi: 10.1091/mbc.11.12.4159
- Hicks, C. R., Smith, H. M. S., Shieh, M., and Raikhel, N. V. (1995). Three classes of nuclear import signals bind to plant nuclei. *Plant Physiol.* 107, 1055–1058. doi: 10.1104/pp.107.4.1055
- Hicks, G. R. (2013). *Nuclear import of plant proteins*, in *madame curie bioscience database*. Austin, TX: Landes Bioscience.
- Hirakawa, T., and Matsunaga, S. (2019). Characterization of DNA repair foci in root cells of *Arabidopsis* in response to DNA damage. *Front. Plant Sci.* 10:990. doi: 10.3389/fpls.2019.00990
- Hirayama, S., Sugihara, M., Morito, D., Iemura, S. I., Natsume, T., Murata, S., et al. (2018). Nuclear export of ubiquitinated proteins via the UBIN-POST system. *Proc. Natl. Acad. Sci. U.S.A.* 115, E4199–E4208. doi: 10.1073/pnas.1711017115
- Hsu, S. F., Lai, H. C., and Jinn, T. L. (2010). Cytosol-localized heat shock factor-binding protein, AtHSBP, Functions as a negative regulator of heat shock response by translocation to the nucleus and is required for seed development in *Arabidopsis*. *Plant Physiol.* 153, 773–784. doi: 10.1104/pp.109.151225
- Huang, X., Zhang, Q., Jiang, Y., Yang, C., and Wang, Q. (2018). Shade-induced nuclear localization of PIF7 is regulated by phosphorylation and 14-3-3 proteins in *Arabidopsis*. *Elife* 7:e31636. doi: 10.7554/eLife.31636
- Hübner, S., Smith, H. M., Hu, W., Kai Chan, C., Rihs, H. P., Paschal, B. M., et al. (1999). Plant importin binds nuclear localization sequences with high affinity and can mediate nuclear import independent of importin β . *J. Biol. Chem.* 274, 22610–22617. doi: 10.1074/jbc.274.32.22610
- Jäkel, S., and Görlich, D. (1998). Importin β , transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. *EMBO J.* 17, 4491–4502. doi: 10.1093/emboj/17.15.4491
- Jordan, E. G. (1984). Nucleolar nomenclature. *J. Cell Sci.* 67, 217–220. doi: 10.1242/jcs.67.1.217
- Jung, J. H., Barbosa, A. D., Hutin, S., Kumita, J. R., Gao, M., Derwort, D., et al. (2020). A prion-like domain in ELF3 functions as a thermosensor in *Arabidopsis*. *Nature* 585, 256–260. doi: 10.1038/s41586-020-2644-7
- Kaganovich, D., Kopito, R., and Frydman, J. (2008). Misfolded proteins partition between two distinct quality control compartments. *Nature* 454, 1088–1095. doi: 10.1038/nature07195
- Kaiser, T. E., Intine, R. V., and Dundr, M. (2008). De novo formation of a subnuclear body. *Science* 322, 1713–1717. doi: 10.1126/science.1165216
- Kalderon, D., Roberts, B. L., Richardson, W. D., and Smith, A. E. (1984). A short amino acid sequence able to specify nuclear location. *Cell* 39, 499–509. doi: 10.1016/0092-8674(84)90457-4
- Kannan, K., Nelson, A. D. L., and Shippen, D. E. (2008). Dyskerin is a component of the *Arabidopsis* telomerase RNP required for telomere maintenance. *Mol. Cell Biol.* 28, 2332–2341. doi: 10.1128/mcb.01490-07
- Kanno, T., Lin, W. D., Fu, J. L., Wu, M. T., Yang, H. W., Lin, S. S., et al. (2016). Identification of coilin mutants in a screen for enhanced expression of an alternatively spliced GFP reporter gene in *Arabidopsis thaliana*. *Genetics* 203, 1709–1720. doi: 10.1534/genetics.116.190751
- Karpen, G. H., Schaefer, J. E., and Laird, C. D. (1988). A *Drosophila* rRNA gene located in euchromatin is active in transcription and nucleolus formation. *Genes Dev.* 2, 1745–1763. doi: 10.1101/gad.2.12b.1745
- Kehlenbach, R. H., Dickmanns, A., and Gerace, L. (1998). Nucleocytoplasmic shuttling factors including ran and CRM1 Mediate nuclear export of NFAT in vitro. *J. Cell Biol.* 141, 863–874. doi: 10.1083/jcb.141.4.863
- Kim, K., Dimitrova, D. D., Carta, K. M., Saxena, A., Daras, M., and Borowiec, J. A. (2005). Novel checkpoint response to genotoxic stress mediated by nucleolin-replication protein a complex formation. *Mol. Cell Biol.* 25, 2463–2474. doi: 10.1128/mcb.25.6.2463-2474.2005
- Kim, L., Kircher, S., Toth, R., Adam, E., Schäfer, E., and Nagy, F. (2000). Light-induced nuclear import of phytochrome-A:GFP fusion proteins is differentially regulated in transgenic tobacco and *Arabidopsis*. *Plant J.* 22, 125–133. doi: 10.1046/j.1365-313x.2000.00729.x
- Kim, S. H., Spensley, M., Choi, S. K., Calixto, C. P. G., Pendle, A. F., Koroleva, O., et al. (2010). Plant U13 orthologues and orphan snoRNAs identified by RNomics of RNA from *Arabidopsis* nucleoli. *Nucleic Acids Res.* 38, 3054–3067. doi: 10.1093/nar/gkp1241
- Kircher, S., Kozma-Bognar, L., Kim, L., Adam, E., Harter, K., Schäfer, E., et al. (1999). Light Quality-dependent nuclear import of the plant photoreceptors phytochrome A and B. *Plant Cell* 11, 1445–1456. doi: 10.1105/tpc.118.1445
- Kitsios, G., Alexiou, K. G., Bush, M., Shaw, P., and Doonan, J. H. (2008). A cyclin-dependent protein kinase, CDK2, colocalizes with and modulates the distribution of spliceosomal components in *Arabidopsis*. *Plant J.* 54, 220–235. doi: 10.1111/j.1365-313x.2008.03414.x
- Kondoh, K., Terasawa, K., Morimoto, H., and Nishida, E. (2006). Regulation of Nuclear translocation of extracellular signal-regulated kinase 5 by active nuclear import and export mechanisms. *Mol. Cell Biol.* 26, 1679–1690. doi: 10.1128/MCB.26.5.1679-1690.2006
- Korgaonkar, C., Hagen, J., Tompkins, V., Frazier, A. A., Allamargot, C., Quelle, F. W., et al. (2005). Nucleophosmin (B23) targets ARF to nucleoli and inhibits its function. *Mol. Cell Biol.* 25, 1258–1271. doi: 10.1128/mcb.25.4.1258-1271.2005
- Korostelev, A., and Noller, H. F. (2007). The ribosome in focus: New structures bring new insights. *Trends Biochem. Sci.* 32, 434–441. doi: 10.1016/j.tibs.2007.08.002
- Kosugi, S., Yanagawa, H., Terauchi, R., and Tabata, S. (2014). NESmapper: Accurate Prediction of leucine-rich nuclear export signals using activity-based profiles. *PLoS Comput. Biol.* 10:e1003841. doi: 10.1371/journal.pcbi.1003841
- Kurzbaue, M. T., Uanschou, C., Chen, D., and Schlögelhofer, P. (2012). The recombinases DMC1 and RAD51 are functionally and spatially separated during meiosis in *Arabidopsis*. *Plant Cell* 24, 2058–2070. doi: 10.1105/tpc.112.098459
- Kutay, U., and Bischoff, F. R. (1997). Export of importin from the nucleus is mediated by a specific nuclear transport factor. *Cell* 90, 1061–1071. doi: 10.1016/S0092-8674(00)80372-4
- Lafontaine, D. L. J., Riback, J. A., Bascetin, R., and Brangwynne, C. P. (2021). The nucleolus as a multiphase liquid condensate. *Nat. Rev. Mol. Cell Biol.* 22, 165–182. doi: 10.1038/s41580-020-0272-6
- Lamond, A. I., and Spector, D. L. (2003). Nuclear speckles: A model for nuclear organelles. *Nat. Rev. Mol. Cell Biol.* 4, 605–612. doi: 10.1038/nrm1172
- Lang, J., Smetana, O., Sanchez-Calderon, L., Lincker, F., Genestier, J., Schmit, A. C., et al. (2012). Plant γ H2AX foci are required for proper DNA DSB repair responses and colocalize with E2F factors. *New Phytol.* 194, 353–363. doi: 10.1111/j.1469-8137.2012.04062.x
- Lange, A., McLane, L. M., Mills, R. E., Devine, S. E., and Corbett, A. H. (2010). Expanding the definition of the classical bipartite nuclear localization signal. *Traffic* 11, 311–323. doi: 10.1111/j.1600-0854.2009.01028.x
- Lecker, S. H., Goldberg, A. L., and Mitch, W. E. (2006). Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *J. Am. Soc. Nephrol.* 17, 1807–1819. doi: 10.1681/ASN.2006010083

- Legris, M., Klose, C., Burgie, E. S., Rojas, C. C., Neme, M., Hiltbrunner, A., et al. (2016). Phytochrome B integrates light and temperature signals in *Arabidopsis*. *Science* 354, 897–900. doi: 10.1126/science.aaf5656
- Leung, A. K. L., Andersen, J. S., Mann, M., and Lamond, A. I. (2003). Bioinformatic analysis of the nucleolus. *Biochem. J.* 376, 553–569. doi: 10.1042/BJ20031169
- Li, C. F., Henderson, I. R., Song, L., Fedoroff, N., Lagrange, T., and Jacobsen, S. E. (2008). Dynamic regulation of ARGONAUTE4 within multiple nuclear bodies in *Arabidopsis thaliana*. *PLoS Genet.* 4:e27. doi: 10.1371/journal.pgen.0040027
- Li, C. F., Pontes, O., El-Shami, M., Henderson, I. R., Bernatavichute, Y. V., Chan, S. W. L., et al. (2006). An ARGONAUTE4-containing nuclear processing center colocalized with cajal bodies in *Arabidopsis thaliana*. *Cell* 126, 93–106. doi: 10.1016/j.cell.2006.05.032
- Li, J., Kinoshita, T., Pandey, S., Ng, C. K.-Y., Gygi, S. P., Shimazaki, K., et al. (2002). Modulation of an RNA-binding protein by abscisic-acid-activated protein kinase. *Nature* 418, 793–797. doi: 10.1038/nature00936
- Lin, Y. T., and Yen, P. H. (2006). A novel nucleocytoplasmic shuttling sequence of DAZAP1, a testis-abundant RNA-binding protein. *RNA* 12, 1486–1493. doi: 10.1261/rna.42206
- Liu, J. L., Murphy, C., Buszczak, M., Clatterbuck, S., Goodman, R., and Gall, J. G. (2006). The *Drosophila melanogaster* Cajal body. *J. Cell Biol.* 172, 875–884. doi: 10.1083/jcb.200511038
- Liu, J.-L., Wu, Z., Nizami, Z., Deryusheva, S., Rajendra, T. K., Beumer, K. J., et al. (2009). Coilin is essential for cajal body organization in *Drosophila melanogaster*. *Mol. Biol. Cell* 20, 1661–1670. doi: 10.1091/mbc.E08-05-0525
- Liu, Q., Shi, L., and Fang, Y. (2012). Dicing bodies. *Plant Physiol.* 158, 61–66. doi: 10.1104/pp.111.186734
- Liu, Y., Lin, G., Yin, C., and Fang, Y. (2020). B-box transcription factor 28 regulates flowering by interacting with constans. *Sci. Rep.* 10:11789. doi: 10.1038/s41598-020-74445-7
- Liu, Y., Liu, Q., Yan, Q., Shi, L., and Fang, Y. (2014). Nucleolus-tethering system (NoTS) reveals that assembly of photobodies follows a self-organization model. *Mol. Biol. Cell* 25, 1366–1373. doi: 10.1091/mbc.E13-09-0527
- Liu, Z., Jia, Y., Ding, Y., Shi, Y., Li, Z., Guo, Y., et al. (2017). Plasma membrane CRPK1-mediated phosphorylation of 14-3-3 proteins induces their nuclear import to fine-tune CBF signaling during cold response. *Mol. Cell* 66, 117.e–128.e. doi: 10.1016/j.molcel.2017.02.016
- Löbrich, M., Shibata, A., Beucher, A., Fisher, A., Ensminger, M., Goodarzi, A. A., et al. (2010). gammaH2AX foci analysis for monitoring DNA double-strand break repair: Strengths, limitations and optimization. *Cell Cycle* 9, 662–669. doi: 10.4161/cc.9.4.10764
- Lorković, Z. J., Hilscher, J., and Barta, A. (2008). Co-localisation studies of *Arabidopsis* SR splicing factors reveal different types of speckles in plant cell nuclei. *Exp. Cell Res.* 314, 3175–3186. doi: 10.1016/j.yexcr.2008.06.020
- Lorković, Z. J., Lopato, S., Pexa, M., Lehner, R., and Barta, A. (2004). Interactions of *Arabidopsis* RS domain containing cyclophilins with SR proteins and U1 and U11 small nuclear ribonucleoprotein-specific proteins suggest their involvement in pre-mRNA splicing. *J. Biol. Chem.* 279, 33890–33898. doi: 10.1074/jbc.M400270200
- Love, A. J., Yu, C., Petukhova, N. V., Kalinina, N. O., Chen, J., and Taliansky, M. E. (2017). Cajal bodies and their role in plant stress and disease responses. *RNA Biol.* 14, 779–790. doi: 10.1080/15476286.2016.1243650
- Lu, J., Wu, T., Zhang, B., Liu, S., Song, W., Qiao, J., et al. (2021). Types of nuclear localization signals and mechanisms of protein import into the nucleus. *Cell Commun. Signal.* 19:60. doi: 10.1186/s12964-021-00741-y
- Lunn, J. E. (2007). Compartmentation in plant metabolism. *J. Exp. Bot.* 58, 35–47. doi: 10.1093/jxb/erl134
- Maceluch, J., Kmiecik, M., Szwejkowska-Kulińska, Z., and Jarmowski, A. (2001). Cloning and characterization of *Arabidopsis thaliana* AtNAP57 — a homologue of yeast pseudouridine synthase Cbf5p. *Acta Biochim. Pol.* 48, 699–709.
- Makarov, V., Rakitina, D., Protopopova, A., Yaminsky, I., Arutunian, A., Love, A. J., et al. (2013). Plant coilin: Structural characteristics and RNA-binding properties. *PLoS One* 8:e53571. doi: 10.1371/journal.pone.0053571
- Mallet, P. L., and Bachand, F. (2013). A Proline-Tyrosine nuclear localization signal (PY-NLS) is required for the nuclear import of fission yeast PAB2, but not of human PABPN1. *Traffic* 14, 282–294. doi: 10.1111/tra.12036
- Mao, Y. S., Zhang, B., and Spector, D. L. (2011). Biogenesis and function of nuclear bodies. *Trends Genet.* 27, 295–306. doi: 10.1016/j.tig.2011.05.006
- Martin, R. M., Ter-Avetisyan, G., Herce, H. D., Ludwig, A. K., Lüttig- Tünnemann, G., and Cardoso, M. C. (2015). Principles of protein targeting to the nucleolus. *Nucleus* 6, 314–325. doi: 10.1080/19491034.2015.1079680
- Martoglio, B., and Dobberstein, B. (1998). Signal sequences: More than just greasy peptides. *Trends Cell Biol.* 8, 410–415. doi: 10.1016/S0962-8924(98)01360-9
- Maruri-López, I., Figueroa, N. E., Hernández-Sánchez, I. E., and Chodasiewicz, M. (2021). Plant stress granules: Trends and beyond. *Front. Plant Sci.* 12:722643. doi: 10.3389/fpls.2021.722643
- Matera, A. G., Izaguirre-sierra, M., Praveen, K., and Rajendra, T. K. (2011). Nuclear bodies: Random aggregates of sticky proteins or crucibles of macromolecular assembly? *Dev. Cell* 17, 639–647. doi: 10.1016/j.devcel.2009.10.017
- Matsuki, R., Iwasaki, T., Shoji, K., Jiang, C.-J., and Yamamoto, N. (1998). Isolation and characterization of two importin β genes from rice. *Plant Cell Physiol.* 39, 879–884. doi: 10.1093/oxfordjournals.pcp.a029448
- Matsushita, T., Mochizuki, N., and Nagatani, A. (2003). Dimers of the N-terminal domain of phytochrome B are functional in the nucleus. *Nature* 424, 571–574. doi: 10.1038/nature01837
- Matsuura, Y., and Stewart, M. (2005). Nup50/Np60 function in nuclear protein import complex disassembly and importin recycling. *EMBO J.* 24, 3681–3689. doi: 10.1038/sj.emboj.7600843
- Mekhalil, K., Rivero-Lopez, L., Al-Masri, A., Brandon, C., Khacho, M., and Lee, S. (2007). Identification of a common subnuclear localization signal. *Mol. Biol. Cell* 18, 3966–3977. doi: 10.1091/mbc.E07-03-0295
- Mélèse, T., and Xue, Z. (1995). The nucleolus: An organelle formed by the act of building a ribosome. *Curr. Opin. Cell Biol.* 7, 319–324. doi: 10.1016/0955-0674(95)80085-9
- Meng, L., Zhu, Q., and Tsai, R. Y. L. (2007). Nucleolar trafficking of nucleostemin family proteins: Common versus protein-specific mechanisms. *Mol. Cell Biol.* 27, 8670–8682. doi: 10.1128/mcb.00635-07
- Michael, W. M. (2000). Nucleocytoplasmic shuttling signals: Two for the price of one. *Trends Cell Biol.* 10, 46–50. doi: 10.1016/S0962-8924(99)01695-5
- Montacé, C., Durut, N., Opsomer, A., Palm, D., Comella, P., Picart, C., et al. (2017). Nucleolar proteome analysis and proteasomal activity assays reveal a link between nucleolus and 26S proteasome in *A. thaliana*. *Front. Plant Sci.* 8:1815. doi: 10.3389/fpls.2017.01815
- Nagatani, A. (2010). Phytochrome: Structural basis for its functions. *Curr. Opin. Plant Biol.* 13, 565–570. doi: 10.1016/j.pbi.2010.07.002
- Nardozi, J. D., Lott, K., and Cingolani, G. (2010). Phosphorylation meets nuclear import: A review. *Cell Commun. Signal.* 8:32. doi: 10.1186/1478-811X-8-32
- Németh, A., and Längst, G. (2011). Genome organization in and around the nucleolus. *Trends Genet.* 27, 149–156. doi: 10.1016/j.tig.2011.01.002
- Neville, M., Stutz, F., Lee, L., Davis, L. I., and Rosbash, M. (1997). The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr. Biol.* 7, 767–775. doi: 10.1016/S0960-9822(06)00335-6
- Nguyen Ba, A. N., Pogoutse, A., Provart, N., and Moses, A. M. (2009). NLStradamus: A simple Hidden markov model for nuclear localization signal prediction. *BMC Bioinformatics* 10:202. doi: 10.1186/1471-2105-10-202
- Niedojadlo, J., Kubicka, E., Kalich, B., and Smoliński, D. J. (2014). Poly(A) RNAs including coding proteins RNAs occur in plant cajal bodies. *PLoS One* 9:e111780. doi: 10.1371/journal.pone.0111780
- Nigg, E. A. (1997). Nucleocytoplasmic transport: Signals, mechanisms and regulation. *Nature* 386, 779–787. doi: 10.1038/386779a0
- Nisa, M. U., Huang, Y., Benhamed, M., and Raynaud, C. (2019). The plant DNA damage response: Signaling pathways leading to growth inhibition and putative role in response to stress conditions. *Front. Plant Sci.* 10:653. doi: 10.3389/fpls.2019.00653
- Nizami, Z., Deryusheva, S., and Gall, J. G. (2010). The cajal body and histone locus body. *Cold Spring Harb. Perspect. Biol.* 2:a000653. doi: 10.1101/cshperspect.a000653
- Ochs, R. L., Lischwe, M. A., Shen, E., Carroll, R. E., and Busch, H. (1985). Nucleologenesis: Composition and fate of prenucleolar bodies. *Chromosoma* 92, 330–336. doi: 10.1007/BF00327463
- Ogg, S. C., and Lamond, A. I. (2002). Cajal bodies and coilin – moving towards function. *J. Cell Biol.* 159, 17–21. doi: 10.1083/jcb.200206111
- Okuwaki, M., Saotome-Yoshimura, A., Masashi, N., Saito, S., Sekiya, H. H.-M. T., and Nagata, K. (2021). RNA-recognition motifs and glycine and arginine-rich region cooperatively regulate the nucleolar localization of nucleolin. *J. Biochem.* 169, 87–100. doi: 10.1093/jb/mvaa095

- Ossareh-Nazari, B., Bacherie, F., and Dargemont, C. (1997). Evidence for a role of CRM1 in signal-mediated nuclear protein export. *Science* 278, 141–144. doi: 10.1126/science.278.5335.141
- Oudhoff, M. J., Freeman, S. A., Couzens, A. L., Antignano, F., Kuznetsova, E., Min, P. H., et al. (2013). Control of the hippo pathway by Set7-dependent methylation of yap. *Dev. Cell* 26, 188–194. doi: 10.1016/j.devcel.2013.05.025
- Padeken, J., and Heun, P. (2014). Nucleolus and nuclear periphery: Velcro for heterochromatin. *Curr. Opin. Cell Biol.* 28, 54–60. doi: 10.1016/j.ceb.2014.03.001
- Palm, D., Simm, S., Darm, K., Weis, B. L., Ruprecht, M., Schleiff, E., et al. (2016). Proteome distribution between nucleoplasm and nucleolus and its relation to ribosome biogenesis in *Arabidopsis thaliana*. *RNA Biol.* 13, 441–454. doi: 10.1080/15476286.2016.1154252
- Pellar, G. J., and DiMario, P. J. (2003). Deletion and site-specific mutagenesis of nucleolin's carboxy GAR domain. *Chromosoma* 111, 461–469. doi: 10.1007/s00412-003-0231-y
- Pendle, A. F., Clark, G. P., Boon, R., Lewandowska, D., Lam, Y. W., Andersen, J., et al. (2005). Proteomic analysis of the *Arabidopsis* nucleolus suggests novel nucleolar functions. *Mol. Biol. Cell* 16, 260–269. doi: 10.1091/mbc.e04-09-0791
- Petrovská, B., Šebela, M., and Doležal, J. (2015). Inside a plant nucleus: Discovering the proteins. *J. Exp. Bot.* 66, 1627–1640. doi: 10.1093/jxb/erv041
- Picart-Piccolo, A., Picart, C., Picault, N., and Pontvianne, F. (2020). Nucleolus-associated chromatin domains are maintained under heat stress, despite nucleolar reorganization in *Arabidopsis thaliana*. *J. Plant Res.* 133, 463–470. doi: 10.1007/s10265-020-01201-3
- Pih, K. T., Yi, M. J., Liang, Y. S., Shin, B. J., Cho, M. J., Hwang, I., et al. (2000). Molecular cloning and targeting of a fibrillarin homolog from *Arabidopsis*. *Plant Physiol.* 123, 51–58. doi: 10.1104/pp.123.1.51
- Pontes, O., and Pikaard, C. S. (2008). siRNA and miRNA processing: New functions for Cajal bodies. *Curr. Opin. Genet. Dev.* 18, 197–203. doi: 10.1016/j.gde.2008.01.008
- Pontvianne, F., Carpentier, M. C., Durut, N., Pavlišťová, V., Jaške, K., Schořová, Š, et al. (2016). Identification of nucleolus-associated chromatin domains reveals a role for the nucleolus in 3D organization of the *A. thaliana* genome. *Cell Rep.* 16, 1574–1587. doi: 10.1016/j.celrep.2016.07.016
- Pontvianne, F., Matia, I., Douet, J., Sylvestre Tourmente, F. J. M., Echeverria, M., and Sáez-Vásquez, J. (2007). Characterization of AtNUC-L1 reveals a central role of nucleolin in nucleolus organization and silencing of AtNUC-L2 Gene in *Arabidopsis*. *Mol. Biol. Cell* 18, 369–379. doi: 10.1091/mbc.E06-08-0751
- Porter, F. W., and Palmenberg, A. C. (2009). Leader-induced phosphorylation of nucleoporins correlates with nuclear trafficking inhibition by cardioviruses. *J. Virol.* 83, 1941–1951. doi: 10.1128/jvi.01752-08
- Ramazi, S., and Zahiri, J. (2021). Post-translational modifications in proteins: Resources, tools and prediction methods. *Database (Oxford)*. 2021:baab012. doi: 10.1093/database/baab012
- Rausin, G., Tillemans, V., Stankovic, N., Hanikenne, M., and Motte, P. (2010). Dynamic Nucleocytoplasmic shuttling of an *Arabidopsis* Sr splicing factor: Role of the RNA-binding domains. *Plant Physiol.* 153, 273–284. doi: 10.1104/pp.110.154740
- Ream, T. S., Haag, J. R., Pontvianne, F., Nicora, C. D., Norbeck, A. D., Pašatolj, L., et al. (2015). Subunit compositions of *Arabidopsis* RNA polymerases I and III reveal Pol I- and Pol III-specific forms of the AC40 subunit and alternative forms of the C53 subunit. *Nucleic Acids Res.* 43, 4163–4178. doi: 10.1093/nar/gkv247
- Reddy, A. S. N., Day, I. S., Göhring, J., and Barta, A. (2012). Localization and dynamics of nuclear speckles in plants. *Plant Physiol.* 158, 67–77. doi: 10.1104/pp.111.186700
- Ren, Y., Li, Y., Jiang, Y., Wu, B., and Miao, Y. (2017). Phosphorylation of WHIRLY1 by CIPK14 shifts its localization and dual functions in *Arabidopsis*. *Mol. Plant* 10, 749–763. doi: 10.1016/j.molp.2017.03.011
- Ribbeck, K., Lipowsky, G., Kent, H. M., Stewart, M., and Görlch, D. (1998). NTF2 mediates nuclear import of Ran. *EMBO J.* 17, 6587–6598. doi: 10.1093/emboj/17.22.6587
- Rockwell, N. C., Su, Y. S., and Lagarias, J. C. (2006). Phytochrome structure and signaling mechanisms. *Annu. Rev. Plant Biol.* 57, 837–858. doi: 10.1146/annurev.arplant.56.032604.144208
- Roger, B., Moisan, A., Amalric, F., and Bouvet, P. (2003). Nucleolin provides a link between RNA polymerase I transcription and pre-ribosome assembly. *Chromosoma* 111, 399–407. doi: 10.1007/s00412-002-0221-5
- Rothenbusch, U., Sawatzki, M., Chang, Y., Caesar, S., and Schlenstedt, G. (2012). Sumoylation regulates Kap114-mediated nuclear transport. *EMBO J.* 31, 2461–2472. doi: 10.1038/emboj.2012.102
- Rowland, R. R. R., and Yoo, D. (2003). Nucleolar-cytoplasmic shuttling of PRRSV nucleocapsid protein: A simple case of molecular mimicry or the complex regulation by nuclear import, nucleolar localization and nuclear export signal sequences. *Virus Res.* 95, 23–33. doi: 10.1016/s0168-1702(03)00161-8
- Ryu, H., Kim, K., Cho, H., Park, J., Choe, S., and Hwang, I. (2007). Nucleocytoplasmic shuttling of BZR1 Mediated by phosphorylation is essential in *Arabidopsis* brassinosteroid signaling. *Plant Cell* 19, 2749–2762. doi: 10.1105/tpc.107.053728
- Sáez-Vásquez, J., and Delseny, M. (2019). Ribosome biogenesis in plants: From functional 45S ribosomal DNA organization to ribosome assembly factors. *Plant Cell* 31, 1945–1967. doi: 10.1105/tpc.18.00874
- Sáez-Vásquez, J., and Gadal, O. (2010). Genome organization and function: A view from Yeast and *Arabidopsis*. *Mol. Plant* 3, 678–690. doi: 10.1093/mp/ssq034
- Sáez-Vásquez, J., and Medina, F. J. (2008). “The plant nucleolus,” in *Advances in botanical research*, eds J. C. Kader and M. Delseny (San Diego, CA: Elsevier), 1–46. doi: 10.1016/S0065-2296(08)00001-3
- Sáez-Vásquez, J., Gallois, P., and Delseny, M. (2000). Accumulation and nuclear targeting of BnC24, a *Brassica napus* ribosomal protein corresponding to a mRNA accumulating in response to cold treatment. *Plant Sci.* 156, 35–46. doi: 10.1016/s0168-9452(00)00229-6
- Saracco, S. A., Miller, M. J., Kurepa, J., and Vierstra, R. D. (2007). Genetic analysis of SUMOylation in *Arabidopsis*: Conjugation of SUMO1 and SUMO2 to nuclear proteins is essential. *Plant Physiol.* 145, 119–134. doi: 10.1104/pp.107.102285
- Sazer, S., and Dasso, M. (2000). The ran decathlon: Multiple roles of ran. *J. Cell Sci.* 113, 1111–1118. doi: 10.1242/jcs.113.7.1111
- Schäfer, E., Marchal, B., and Marmé, D. (1972). In vivo measurements of the phytochrome photostationary state in far red light. *Photochem. Photobiol.* 15, 457–464. doi: 10.1111/j.1751-1097.1972.tb06257.x
- Schmid, F. X. (1995). Protein folding. Prolyl isomerases join the fold. *Curr. Biol.* 5, 993–994. doi: 10.1016/S0960-9822(95)00197-7
- Schmidt-Zachmann, M. S., and Nigg, E. A. (1993). Protein localization to the nucleolus: A search for targeting domains in nucleolin. *J. Cell Sci.* 105, 799–806. doi: 10.1242/jcs.105.3.799
- Scott, M. S., Boisvert, F. M., McDowall, M. D., Lamond, A. I., and Barton, G. J. (2010). Characterization and prediction of protein nucleolar localization sequences. *Nucleic Acids Res.* 38, 7388–7399. doi: 10.1093/nar/gkq653
- Scott, M. S., Troshin, P. V., and Barton, G. J. (2011). NoD: A Nucleolar localization sequence detector for eukaryotic and viral proteins. *BMC Bioinformatics* 12:317. doi: 10.1186/1471-2105-12-317
- Shieh, M. W., Wessler, S. R., and Raikhel, N. V. (1993). Nuclear Targeting of the maize r protein requires two nuclear localization sequences. *Plant Physiol.* 101, 353–361. doi: 10.1104/pp.101.2.353
- Shubina, M. Y., Arifulin, E. A., Sorokin, D. V., Sosina, M. A., Tikhomirova, M. A., Serebryakova, M. V., et al. (2020). The GAR domain integrates functions that are necessary for the proper localization of fibrillarin (FBL) inside eukaryotic cells. *PeerJ* 8:e9029. doi: 10.7717/peerj.9029
- Singh, G., Batzenschlager, M., Tomkova, D., Herzog, E., Hoffmann, E., Houlné, G., et al. (2022). GIP1 and GIP2 contribute to the maintenance of genome stability at the nuclear periphery. *Front. Plant Sci.* 12:804928. doi: 10.3389/fpls.2021.804928
- Singh, H., Kaur, K., Singh, M., and Kaur, G. (2020). Plant cyclophilins: Multifaceted proteins with versatile roles. *Front. Plant Sci.* 11:585212. doi: 10.3389/fpls.2020.585212
- Siomi, H., and Dreyfuss, G. (1995). A nuclear localization domain in the hnRNP A1 protein. *J. Cell Biol.* 129, 551–560. doi: 10.1083/jcb.129.3.551
- Sirri, V., Urcuqui-Inchima, S., Roussel, P., and Hernandez-Verdun, D. (2008). Nucleolus: The fascinating nuclear body. *Histochem. Cell Biol.* 129, 13–31. doi: 10.1007/s00418-007-0359-6
- Sleeman, J. E., and Trinkle-Mulcahy, L. (2014). Nuclear bodies: New insights into assembly/dynamics and disease relevance. *Curr. Opin. Cell Biol.* 28, 76–83. doi: 10.1016/j.ceb.2014.03.004
- Sloan, K. E., Warda, A. S., Sharma, S., Entian, K. D., Lafontaine, D. L. J., and Bohnsack, M. T. (2017). Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biol.* 14, 1138–1152. doi: 10.1080/15476286.2016.1259781
- Smith, A., Brownawell, A., and Macara, I. G. (1998). Nuclear import of ran is mediated by the transport factor NTF2. *Curr. Biol.* 8, 1403–1406. doi: 10.1016/s0960-9822(98)00023-2
- Smith, H. M., Hicks, G. R., and Raikhel, N. V. (1995). Importin a from *Arabidopsis thaliana* is a nuclear import receptor that recognizes three classes of import signals. *Plant Physiol.* 114, 411–417. doi: 10.1104/pp.114.2.411

- Snaar, S., Wiesmeijer, K., Jochemsen, A. G., Tanke, H. J., and Dirks, R. W. (2000). Mutational analysis of fibrillarin and its mobility in living human cells. *J. Cell Biol.* 151, 653–662. doi: 10.1083/jcb.151.3.653
- Solymosi, K., and Schoefs, B. (2019). Plant cell compartments. *Bot. Lett.* 166, 269–273. doi: 10.1080/23818107.2019.1652851
- Song, L., Han, M. H., Lesicka, J., and Fedoroff, N. (2007). *Arabidopsis* primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. *Proc. Natl. Acad. Sci. U.S.A.* 104, 5437–5442. doi: 10.1073/pnas.0701061104
- Spector, D. L., and Lamond, A. I. (2011). Nuclear speckles. *Cold Spring Harb. Perspect. Biol.* 3:a000646. doi: 10.1101/cshperspect.a000646
- Stacey, M. G., and Von Arnim, A. G. (1999). A novel motif mediates the targeting of the *Arabidopsis* COP1 protein to subnuclear foci. *J. Biol. Chem.* 274, 27231–27236. doi: 10.1074/jbc.274.38.27231
- Stewart, M. (2007). Molecular mechanism of the nuclear protein import cycle. *Nat. Rev. Mol. Cell Biol.* 8, 195–208. doi: 10.1038/nrm2114
- Strzelecka, M., Trowitzsch, S., Weber, G., Lüthmann, R., Oates, A. C., and Neugebauer, K. M. (2010). Coilin-dependent snRNP assembly is essential for zebrafish embryogenesis. *Nat. Struct. Mol. Biol.* 17, 403–409. doi: 10.1038/nsmb.1783
- Stutz, F., Izaurralde, E., Mattaj, I. W., and Rosbash, A. M. (1996). A role for nucleoporin FG repeat domains in export of human immunodeficiency virus type 1 rev protein and RNA from the nucleus. *Mol. Cell. Biol.* 16, 7144–7150. doi: 10.1128/mcb.16.12.7144
- Suzuki, T., Izumi, H., and Ohno, M. (2010). Cajal body surveillance of U snRNA export complex assembly. *J. Cell Biol.* 190, 603–612. doi: 10.1083/jcb.201004109
- Taddei, A., Hediger, F., Neumann, F. R., and Gasser, S. M. (2004). The function of nuclear architecture: A genetic approach. *Annu. Rev. Genet.* 38, 305–345. doi: 10.1146/annurev.genet.37.110801.142705
- Tajirishi, M. M., Tuteja, R., and Tuteja, N. (2011). Nucleolin. *Commun. Integr. Biol.* 4, 267–275. doi: 10.4161/cib.4.3.14884
- Takagi, M., Absalon, M. J., McLure, K. G., and Kastan, M. B. (2005). Regulation of p53 translation and induction after DNA damage by ribosomal protein L26 and nucleolin. *Cell* 123, 49–63. doi: 10.1016/j.cell.2005.07.034
- Tamura, K., and Hara-Nishimura, I. (2014). Functional insights of nucleocytoplasmic transport in plants. *Front. Plant Sci.* 5:118. doi: 10.3389/fpls.2014.00118
- Tang, H., McDonald, D., Middlesworth, T., Hope, T. J., and Wong-Staal, F. (1999). The carboxyl terminus of RNA helicase A contains a bidirectional nuclear transport domain. *Mol. Cell. Biol.* 19, 3540–3550. doi: 10.1128/MCB.19.3540
- Taneva, S. G., Lee, J. M. C., and Cornell, R. B. (2012). The amphipathic helix of an enzyme that regulates phosphatidylcholine synthesis remodels membranes into highly curved nanotubules. *Biochim. Biophys. Acta – Biomembr.* 1818, 1173–1186. doi: 10.1016/j.bbamem.2012.01.006
- Taoka, M., Nobe, Y., Yamaki, Y., Sato, K., Ishikawa, H., Izumikawa, K., et al. (2018). Landscape of the complete RNA chemical modifications in the human 80S ribosome. *Nucleic Acids Res.* 46, 9289–9298. doi: 10.1093/nar/gky811
- Tillemans, V., Dispa, L., Remacle, C., Collinge, M., and Motte, P. (2005). Functional distribution and dynamics of *Arabidopsis* SR splicing factors in living plant cells. *Plant J.* 41, 567–582. doi: 10.1111/j.1365-3113X.2004.02321.x
- Timney, B. L., Raveh, B., Mironska, R., Trivedi, J. M., Kim, S. J., Russel, D., et al. (2016). Simple rules for passive diffusion through the nuclear pore complex. *J. Cell Biol.* 215, 57–76. doi: 10.1083/jcb.201601004
- Trinkle-Mulcahy, L. (2009). Aberrant mRNA transcripts and nonsense-mediated decay. *F1000 Biol. Rep.* 1:93. doi: 10.3410/b1-93
- Trotman, L. C., Wang, X., Alimonti, A., Chen, Z., Yang, J. T. F. H., Pavletich, N. P., et al. (2007). Ubiquitination regulates PTEN nuclear import and tumor suppression. *Cell* 128, 141–156. doi: 10.1016/j.cell.2006.11.040
- Tucker, K. E., Berciano, M. T., Jacobs, E. Y., LePage, D. F., Shpargel, K. B., Rossire, J. J., et al. (2001). Residual Cajal bodies in coilin knockout mice fail to recruit Sm snRNPs and SMN, the spinal muscular atrophy gene product. *J. Cell Biol.* 154, 293–307. doi: 10.1083/jcb.200104083
- van Buskirk, E. K., Decker, P. V., and Chen, M. (2012). Photobodies in light signaling. *Plant Physiol.* 158, 52–60. doi: 10.1104/pp.111.186411
- Walker, M. P., Tian, L., and Matera, G. A. (2009). Reduced viability, fertility and fecundity in mice lacking the Cajal body marker protein, coilin. *PLoS One* 4:e6171. doi: 10.1371/journal.pone.0006171
- Wang, L. C., Wu, J. R., Hsu, Y. J., and Wu, S. J. (2015). *Arabidopsis* HIT4, a regulator involved in heat-triggered reorganization of chromatin and release of transcriptional gene silencing, relocates from chromocenters to the nucleolus in response to heat stress. *New Phytol.* 205, 544–554. doi: 10.1111/nph.13088
- Wang, L., Li, M., Cai, M., Xing, J., Wang, S., and Zheng, C. (2012). A PY-nuclear localization signal is required for nuclear accumulation of HCMV UL79 protein. *Med. Microbiol. Immunol.* 201, 381–387. doi: 10.1007/s00430-012-0243-4
- Weis, B. L., Kovacevic, J., Missbach, S., and Schleiff, E. (2015). Plant-specific features of ribosome biogenesis. *Trends Plant Sci.* 20, 729–740. doi: 10.1016/j.tplants.2015.07.003
- Wen, D., Wu, J., Wang, L., and Fu, Z. (2017). SUMOylation promotes nuclear import and stabilization of polo-like kinase 1 to support its mitotic function. *Cell Rep.* 21, 2147–2159. doi: 10.1016/j.celrep.2017.10.085
- Wen, W., Meinkoth, J. L., Tsien, R. Y., and Taylor, S. S. (1995). Identification of a signal for rapid export of proteins from the nucleus. *Cell* 82, 463–473. doi: 10.1016/0092-8674(95)90435-2
- Wiese, A. J., Steinbachová, L., Timofejeva, L., Ěermák, V., Klodová, B., Ganji, R. S., et al. (2021). *Arabidopsis* bZIP18 and bZIP52 accumulate in nuclei following heat stress where they regulate the expression of a similar set of genes. *Int. J. Mol. Sci.* 22:530. doi: 10.3390/ijms22020530
- Xia, C., Tao, Y., Li, M., Che, T., and Qu, J. (2020). Protein acetylation and deacetylation: An important regulatory modification in gene transcription (Review). *Exp. Ther. Med.* 20, 2923–2940. doi: 10.3892/etm.2020.9073
- Xiao, C., Hu, S., and Jans, D. A. (1997). SV40 large tumor antigen nuclear import is regulated by the double-stranded DNA-dependent protein kinase site (Serine 120) flanking the nuclear localization sequence. *J. Biol. Chem.* 272, 22191–22198. doi: 10.1074/jbc.272.35.22191
- Xie, D., Chen, M., Niu, J., Wang, L., Li, Y., Fang, X., et al. (2021). Phase separation of SERRATE drives dicing body assembly and promotes miRNA processing in *Arabidopsis*. *Nat. Cell Biol.* 23, 32–39. doi: 10.1038/s41556-020-00606-5
- Yamaguchi, R., Nakamura, M., Mochizuki, N., Kay, S. A., and Nagatani, A. (1999). Light-dependent translocation of a phytochrome B-GFP fusion protein to the nucleus in transgenic *Arabidopsis*. *J. Cell Biol.* 145, 437–445. doi: 10.1083/jcb.145.3.437
- Yang, X. J., and Seto, E. (2008). Lysine acetylation: Codified crosstalk with other posttranslational modifications. *Mol. Cell* 31, 449–461. doi: 10.1016/j.molcel.2008.07.002
- Zhao, L. J., Subramanian, T., Zhou, Y., and Chinnadurai, G. (2006). Acetylation by p300 regulates nuclear localization and function of the transcriptional corepressor CtBP2. *J. Biol. Chem.* 281, 4183–4189. doi: 10.1074/jbc.M509051200

1.4. The nucleolus: structure and function

The nucleolus is the most prominent sub-nuclear domain in eukaryotic cells. This membrane-less structure has been attributed to ribosome biogenesis, often referred as the “ribosome factory”. Nevertheless, this idea has evolved over the years, and the nucleolus is currently conceived a multifunctional entity. Structurally, it possesses generally a tripartite composition: the fibrillar centre (FC), the dense fibrillar component (DFC) and the granular component (GC). These three components are distributed in a vectorial fashion, where the FCs are surrounded by the DFC, being embedded in the GC (Goessens, 1984; Mélése and Xue, 1995; Shaw and Jordan, 1995). The structure of the nucleolus is tightly correlated with its role in ribosome biogenesis.

1.4.1. Structure of the nucleolus

The interphase nucleolus is a nuclear body composed of three different components, distinguishable by electronic microscopy (EM): FC, DFC and GC. Some exceptions are found in insects, such as *Drosophila melanogaster*, and some Reptilia (lizards, crocodiles and snakes), where nucleoli exhibit a bipartite organisation, lacking FCs (Knibiehler et al., 1982, 1984; Thiry et al., 2011).

The FC are low-density areas located in the inner part of the nucleolus. The state of the cell influences the size and number of FCs. Cells with a high rate of ribosome production possess loads of small FCs, whereas one large FC is characteristic of cells with low metabolic activity (reviewed by Sirri et al., 2008; reviewed by Stępiński, 2014). Interestingly, two different types of FCs can be distinguished in plant cells: homogeneous FCs, similar to those present in animal cells, and heterogeneous FCs. Homogeneous FCs are composed of loose material, which are small and numerous in cells with high ribosome production. On the other hand, heterogeneous FCs have small, condensed chromatin inclusions. These are characteristic of cells with low metabolism, where they tend to be large and scarce (Sato, 1985; Sáez-Vásquez and Medina, 2008).

The FCs are surrounded by the DFC, which appears as a denser area. It consists of short fibres which correspond to nascent rRNA molecules. In plants, the DFC occupies the majority of the nucleolar space (40 – 70% approximately). It is the site of early pre-rRNA processing, as pre-rRNA molecules have been observed in the DFC (Shaw and Jordan, 1995; Sáez-Vásquez and Medina, 2008; reviewed by Sirri et al., 2008; reviewed by Stępiński, 2014). For instance, fibrillarin is one of the most employed markers of the DFC (Cerdido et al., 1995). In fact, the filamentous structures composed of the FC and the DFC is termed the nucleolonema, which has differences between animal and plant cells (Yano and Sato, 2000; reviewed by Deltour et al., 1990). The densest region of the nucleolus is formed by the GC, where the FCs and the DFC are embedded. The processing of the

pre-rRNA molecules extends from the DFC to the GC, where the late stages take place. The GC is also the site of the assembly of the pre-ribosomal particles (Shaw and Jordan, 1995; reviewed by Stępiński, 2014).

One nucleolar sub-structure characteristic of plant nucleoli is the nucleolar cavity (also termed as nucleolar vacuole). These sub-structures are localised in the centre of actively transcribing nucleoli, and they are not common in animal cells. The function of this sub-structure is not completely known, but it has been hypothesised that it may mediate the temporary sequestration of certain molecules. Other nucleolar cavities have been observed to contain nucleolar associated chromatin (Figure 1.3; reviewed by Stępiński, 2014).

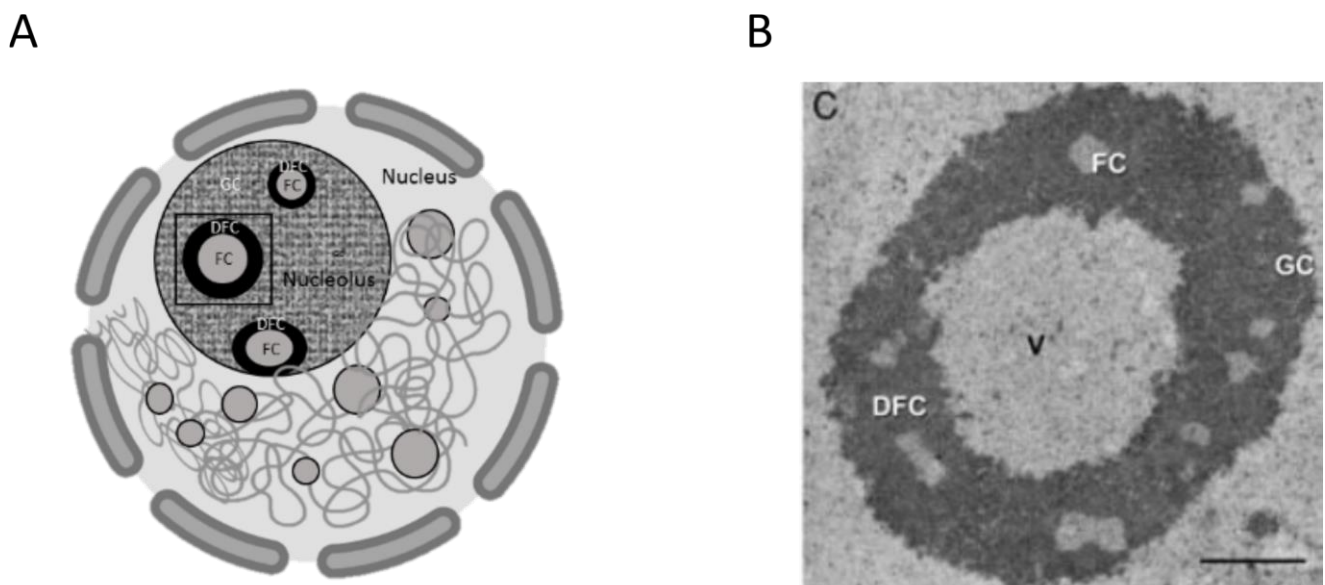


Figure 1.3. Representation of the plant nucleus and nucleolus. (A) The nucleus contains the majority of the genetic information along with certain types of nuclear bodies (grey circles). The most prominent nuclear body is the nucleolus. This part of the figure was partially generated using BioRender.com. (B) Nucleolus from Arabidopsis cells FC: fibrillar centre; DFC: dense fibrillar component; GC: granular component; V: nucleolar cavity. Scale bar set at 1 μm . This part of the figure was taken from Sáez-Vásquez and Medina, (2008).

1.4.2. Nucleolar-Associated Chromatin Domains

The nucleolus is mainly composed of proteins and RNA molecules (up to 90%; Stępiński, 2014). Nevertheless, the existence of chromatin in the nucleolus has been characterised in human and plant cells. Several megabases of human chromatin form the nucleolar associated chromatin domains (NADs) in human cells. Different genes involved in response to the presence of other organisms, odour perception, and tissue development, as well as satellite repeats are the main components of the human NADs (Németh et al., 2010; Van Koningsbruggen et al., 2010).

NADs were subsequently discovered and characterised in *Arabidopsis*. They contain transposable elements (TEs), sub-telomeric regions and inactive protein-coding. In fact, some heterochromatic marks have been found to be abundant, including CG methylation, H3K9me2 and H3K27me1 (associated with TEs and intergenic regions), and H3K27me3 (Pontvianne et al., 2012). All these marks seem to be correlated with genomic regions without transcription activity (Sequeira-Mendes et al., 2014). In contrast, *Arabidopsis* NADs also contain the entire left arm of chromosome 4, which contain active 45S rDNA copies (Pontvianne et al., 2016). The composition of *Arabidopsis* NADs is maintained during heat stress (Picart-Piccolo et al., 2020). In contrast, *Arabidopsis nuc1* null mutants display a completely different composition of the NADs compared to wild-type plants, which portrays the crucial function of NUC1 in chromosome organisation (Pontvianne et al., 2016; reviewed by Picart and Pontvianne, 2017).

1.4.3. Functions of the nucleolus

1.4.3.1. Ribosome biogenesis

Eukaryotic ribosomes (80S) are composed of two different particles: the small and the large ribosomal particles (40S and 60S, respectively; Figure 1.4). In *Arabidopsis*, the 60S ribosomal particle is formed of around 47 ribosomal proteins along with the 5S, 5.8S and 25S (28S in mammals) rRNAs. This particle is responsible of the peptidyl transferase reaction, in other words, the formation of peptide bonds. In contrast, the 40S ribosomal particle is only constituted of the 18S rRNA as well as 33 ribosomal proteins, approx., whose role is the mRNA decoding (Wilson and Cate, 2012; reviewed by Martinez-Seidel et al., 2020). The process of ribosome biogenesis can be divided into three different stages: (i) transcription of rRNA genes, (ii) processing of pre-rRNA molecules and (iii) formation of pre-ribosomal particles. The different stages of the ribosome biogenesis take place in different structures of the nucleolus: the transcription of the rRNA genes occurs in the transition between the FCs and the DFC, the early processing events of the pre-rRNA molecules happen in the DFC, and the late processing steps are located in the GC (Hozák et al., 1994; Puvion-Dutilleul et al., 1997; reviewed by Nazar, 2004). It must be noted that the transcription and processing of rRNA genes has been compared to a Christmas tree, based on the visualization of eukaryotic chromatin under electron microscopy. The “trunk” corresponds to the rDNA copies, with “branches” of nascent transcripts (pre-rRNAs), which are decorated with “balls” (90S/SSU processome and pre-60S ribosomal particles; see below; Mougey et al., 1993).

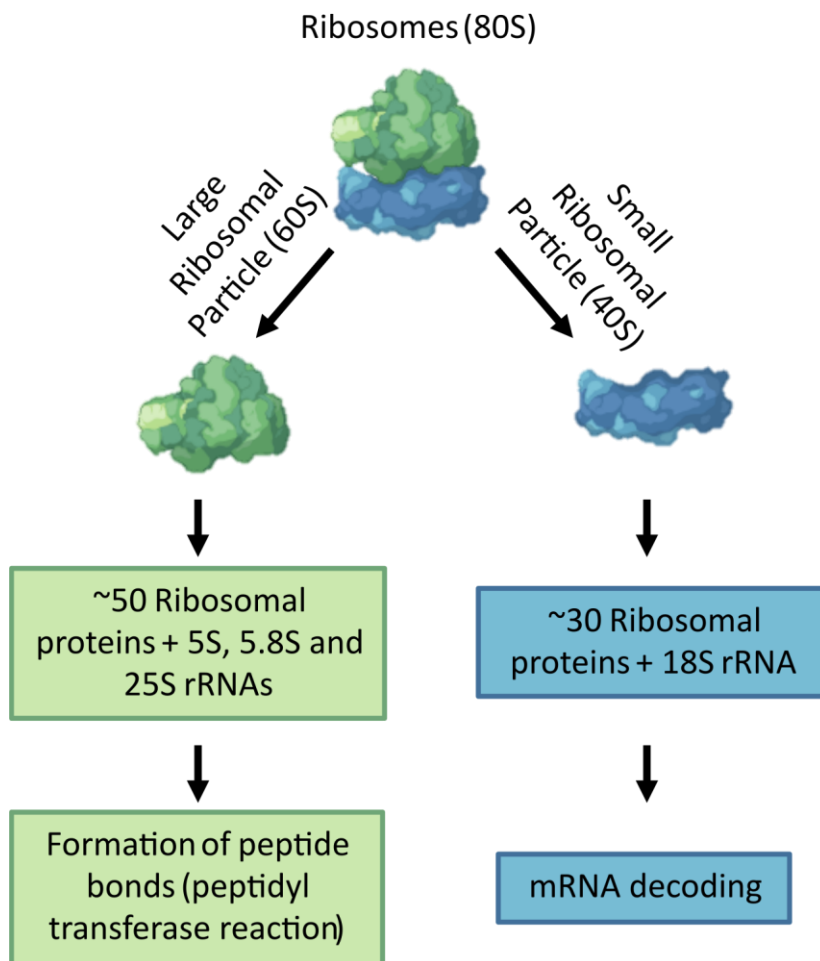


Figure 1.4. Structure and function of Arabidopsis ribosomes. On the one hand, the large ribosomal particle (60S) is composed of three rRNAs (5.8S, 18S and 25S rRNAs) and around 50 ribosomal proteins, approximately. The function of this particle is the formation of the peptide bonds during translation. On the other hand, the small ribosomal particle (40S) is formed of the 18S rRNA along with 30 ribosomal proteins, approximately. The role of this particle is the association with the tRNA to carry the mRNA decoding. This figure was partially generated using BioRender.com.

1.4.3.1.1. Transcription of rRNA genes

In Arabidopsis, three of the four rRNA genes are located in the same transcriptional unit, called the 45S rDNA (47S and 35S rDNA in mammals and yeast, respectively). This is the case of the 18S, 5.8S and 25S rRNA genes. Around 750 copies (per haploid genome) of the 45S rDNA are present in tandem, named Nucleolar Organiser Regions (NORs), located in the ends of the short arm of chromosomes 2 and 4 (NOR2 and NOR4, respectively) in Arabidopsis (Figure 1.5A; Copenhaver et al., 1995; Copenhaver and Pikaard, 1996; reviewed by Sáez-Vásquez and Delseny, 2019). The 45S rDNA is comprised of the 18S, 5.8S and 25S rRNAs genes separated by two internal transcribed spacers (ITS1 and 2), and flanked by the external transcribed spacers (5' and 3'ETS; Figure 1.5B; Unfried et al., 1989; Unfried and Gruendler, 1990; Gruendler et al., 1991). In fact, characterization of the sequence of the 3'ETS in Arabidopsis revealed the existence of five different variants of 45S rDNA (VAR1 – VAR5). VAR1 represents the most abundant variant (50%), being mostly present in the NOR2, while some VAR3 have been also observed in the NOR2. In contrast, VAR2 and VAR4 are found in the NOR4. VAR5 was later characterized as a VAR3 subtype (Abou-Ellail et al., 2011; Havlová et al., 2016). Under normal conditions, the

NOR2 is transcriptionally silent, whereas the NOR4 is active in *Arabidopsis* Col-0 (Pontvianne et al., 2012; Chandrasekhara et al., 2016). What is more, not all active NOR4 copies are transcribed. However, VAR1 units are transcribed in *Arabidopsis nuc1* mutants, demonstrating the importance of NUC1 in the maintenance of silenced 45S rDNA copies (Figure 1.5C; Pontvianne et al., 2010). A recent study reveals the existence of several subtypes of NORs in *Arabidopsis* as well as NOR2 copies which are not transcriptionally silent (Fultz et al., 2023). In addition, two different variants of the 5'ETS have been designated in *Arabidopsis*. The most abundant 5'ETS has been termed as varA, which happens to be longer. In contrast, the shorter and rarer 5'ETS variant is called varB (Gruendler et al., 1991; Havlová et al., 2016). The variation of the 5' and 3'ETS in a natural sample of *Arabidopsis* plants, as well as other *Brassicaceae* members, has been recently addressed (Delorme-Hinoux et al., 2023).

In addition, a sequence known as the intergenic spacer (IGS) is located between the 3'ETS of one 45S rDNA and the 5'ETS of the next 45S rDNA. Similar to animals, the length of this sequence is highly heterogeneous in plants (Wellauer et al., 1976; Boncinelli et al., 1983; Kuehn and Arnheim, 1983; reviewed by Rogers and Bendich, 1987). Sequence analysis of the IGS sequence in *Arabidopsis* revealed the presence of some "repetitive blocks" of different length: B, A₁, A₂, A₃, D₁, D₂ and C from the 5' to the 3' of the IGS (Figure 1.6A). The central region (A₁ – A₃) is characterised by the presence of a 20-bp sequence rich in *SalI* restriction sites. Consequently, the reference sequence of the IGS in *Arabidopsis* consists of three *SalI* boxes, two spacer promoters (SP) and the gene promoter (GP; Figure 1.5B; Gruendler et al., 1991). Nevertheless, further sequence analysis revealed the existence of several IGS variants in *Arabidopsis*, varying in the length of the repetitive block and the number of spacer promoters (Figure 1.6B; Havlová et al., 2016). Whereas the GP is responsible of the transcription of the 45S rDNA via RNA Pol I, the role of the spacer promoters is not certainly known. The sequence TATATAGGG (where the transcription initiation site is underlined) has been observed in the GP and SPs, even though the transcription signal is weaker in the SPs. Functionally, it has been hypothesised that certain regions of the IGS may act as enhancers of the transcription from the GP (Doelling et al., 1993).

Outside the nucleolus, the fourth rRNA gene (5S rRNA gene or 5S rDNA) is present as thousands of copies in the pericentromeric regions of chromosomes 3, 4 and 5 (5S rDNA clusters; Figure 1.5A). Each 5S rDNA unit consists of the transcribed region with an internal promoter, as well as an intergenic spacer. Contrary to the 45S rDNA, these rRNA genes are transcribed by RNA Pol III (Figure 1.7). In addition, the transcription factor IIIA (TFIII A) binds to the 5S rDNA promoter for its correct transcription (Campell et al., 1992; Cloix et al., 2000).

Transcriptionally active 45S rDNA units are transcribed by RNA Pol I from the GP located in the IGS (Doelling et al., 1993; Saez-Vasquez and Pikaard, 1997; reviewed by Raška et al., 2004). In *Arabidopsis*, RNA Pol I is a multimeric protein complex composed of 15 subunits (Ream et al., 2015). Transcription of the 45S rDNA is also achieved by the presence of GTFs. Little is known about plant GTFs involved in the transcription of the 45S

rDNA. An RNA-polymerase-I holoenzyme containing CKII and histone deacetylase has been reported in *Brassica olearcea* (Sáez-Vasquez et al., 2004). Others include plant-specific TFIIB-related protein pBrp (Imamura et al., 2008) or the serine/threonine-protein kinase TOR (Ren et al., 2011). In addition, it has been characterized the role of NUC1 in the transcription of the 45S rDNA, as it is able to bind the GP (Cong et al., 2012). This initial step in the ribosome biogenesis takes place in the transition area FC–DFC (Hozák et al., 1994).

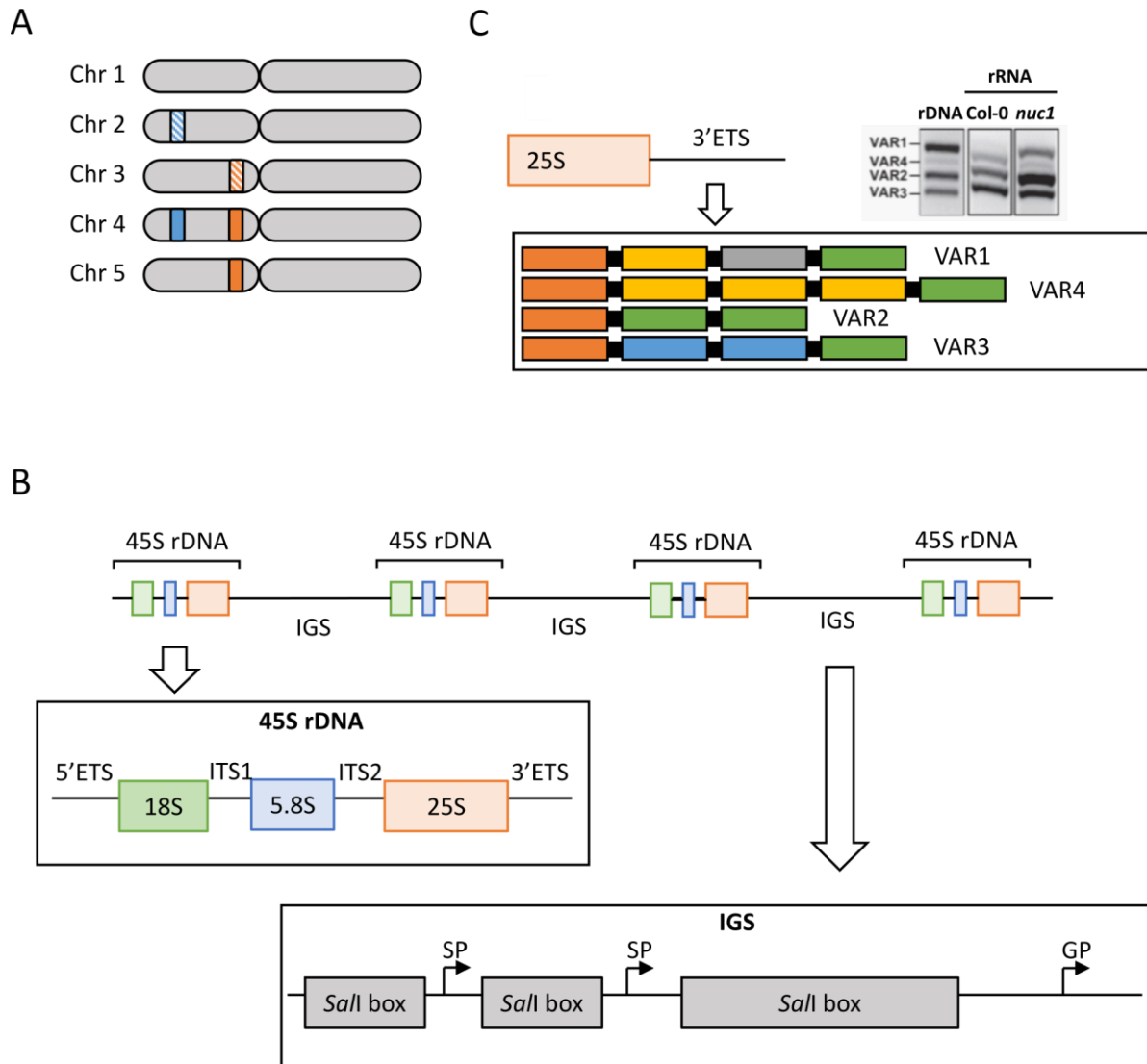


Figure 1.5. 45S rDNA organisation and expression in Arabidopsis. (A) Localisation of the NORs (blue) and 5S rDNA clusters (orange) in Arabidopsis. NOR2 and 5S rDNA clusters present in chromosome 3 have been coloured in stripes to indicate that they are transcriptionally silent in Col-0 plants under normal conditions. (B) Structural organisation of the NORs in Arabidopsis. The tandem copies of the 45S rDNA are separated by the IGS. The 45S rDNA includes the 18S, 5.8S and 25S genes as well as the 5' and 3'ETS, and the ITS1 and 2. On the other hand, the reference IGS in Arabidopsis is composed of three *SalI* boxes, two spacer promoters (SP) and the gene promoter (GP), from which transcription of the 45S rDNA via RNA Pol I is initiated. (C) The different variants of the 3'ETS are represented. The colour blocks portray the different structural units of each variant. The transcription of the different variants is shown for Col-0 and *nuc1* mutants. This part of the figure was modified from Mohannath et al., (2016).

A



B

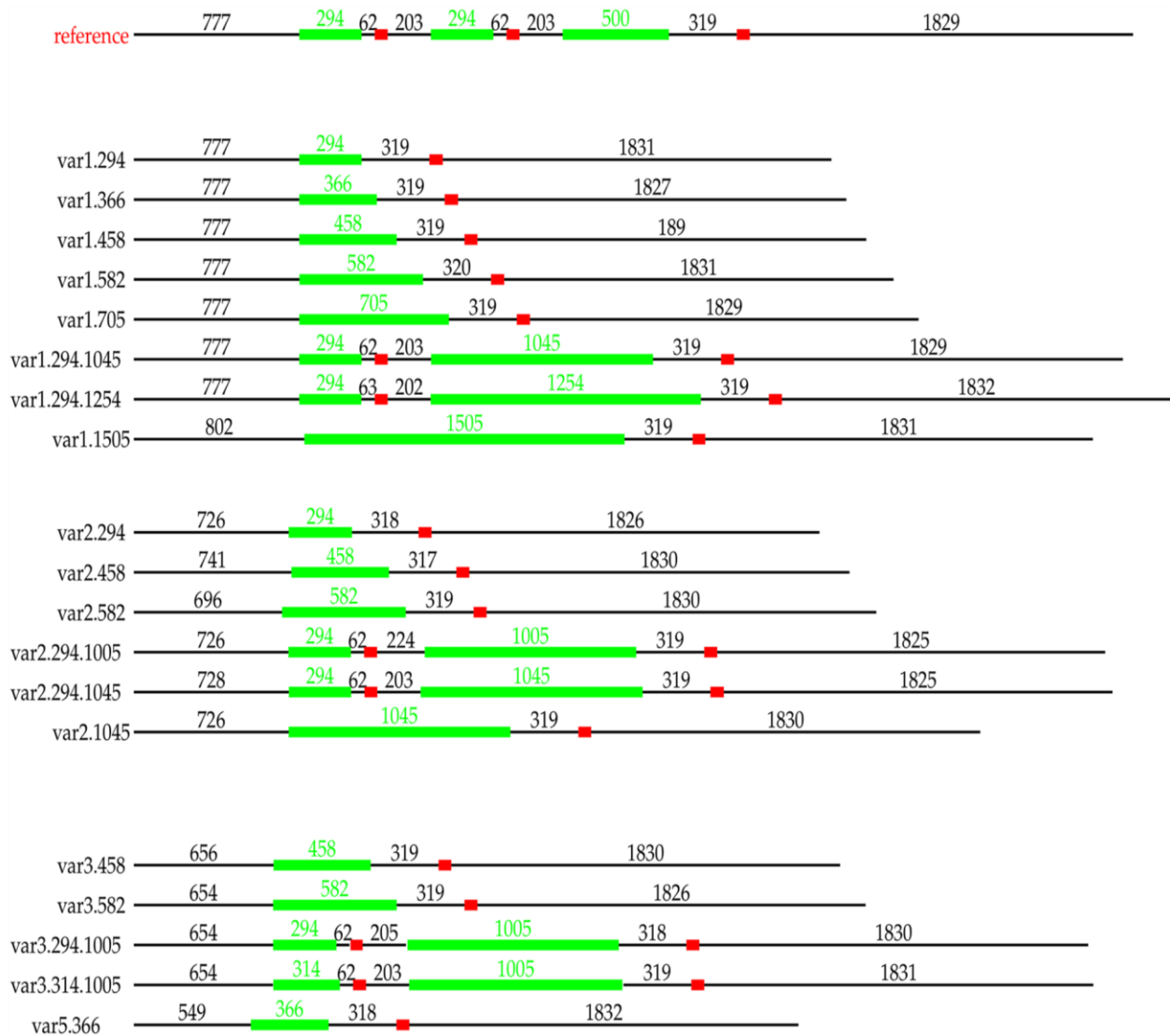


Figure 1.6. Structure of Arabidopsis IGS. (A) Representation of the repetitive regions of the IGS in Arabidopsis, taken from Gruendler et al., (1991). SP: spacer promoter; GP: gene promoter. (B) IGS variants described in Arabidopsis, taken from Havlová et al., (2016). These IGS variants have been found associated with different types of variants of the 3'ETS. For instance, each “var1” (var1.294, var1.366, var1.458, var1.582, var1.705, var1.294.1045, var2.294.1254 and var1.1505) has been observed in VAR1 45S rDNA. The *SalI* boxes and the promoters (SP or GP) are represented in green and red, respectively. The numbers indicate the length (nucleotides) of the corresponding part of the sequence.

1.4.3.1.2. Processing of pre-rRNA molecules

After transcription of the 45S rDNA, the 45S pre-rRNAs experiment early processing steps in the DFC (Figure 1.7; Puvion-Dutilleul et al., 1997). 45S pre-rRNA molecules are subjected to certain sugar and base modifications. Two main modifications are 2'-O-methylation and pseudouridylation. The former consists of the addition of a methyl group to the 2'-O position of the ribose of any of the four ribonucleosides. In contrast, pseudouridylation occurs through uridine isomerization, which translates into a 180° rotation of the pyrimidine ring (Sharma et al., 2013; reviewed by Sáez-Vásquez and Delseny, 2019). 2'-O-methylation and pseudouridylation are guided by C/D box and H/ACA box snoRNAs, respectively. C/D box snoRNAs possess the conserved C, C', D and D' boxes, which allow import into the nucleolus. The specificity towards the target pre-rRNA molecule is achieved by one or two antisense sequences present upstream the D box. These C/D box snoRNAs associate with certain proteins to form small nucleolar ribonucleoprotein particles (snoRNPs), which include fibrillarin, NOP56, NOP58 and 15.5K/L7Ae. On the other hand, H/ACA box snoRNAs contain the H and ACA boxes, which are necessary for nucleolar import. Similar to the C/D box snoRNAs, H/ACA box snoRNAs form snoRNPs, where the pseudouridine synthase dyskerin (also termed NAP57) is found (Figure 1.8A; Kiss, 2002; Bertrand and Fournier, 2004; reviewed by Streit and Schleiff, 2021).

In fact, upon the transcription of the 45S rDNA, a large macromolecular complex is formed in the nucleolus of eukaryotes: the 90S/SSU processome. This U3 snoRNP complex, involved in the processing of the 18S rRNA, will form the 40S ribosome subunit. U3 is a C/D box snoRNA associated with certain proteins, including Utp proteins or fibrillarin, to form the U3 snoRNP (Gallagher et al., 2004; Phipps et al., 2011). Even though the 90S/SSU processome has not been structurally characterized in Arabidopsis, most of the components of the complex have orthologues in the plant genome (reviewed by Sáez-Vásquez and Delseny, 2019). A possible 90S/SSU processome model has been created in Arabidopsis (Sáez-Vásquez and Delseny, 2019). This hypothetical complex possesses similar functions as in mammals or yeast, but it also exhibits some plant-specific features. In contrast to human and yeast, some of the genes encoding components of U3 snoRNPs are duplicated in Arabidopsis. For instance, the Arabidopsis genome possesses two genes encoding for two fibrillarin protein (FIB1 and FIB2; Barneche et al., 2000; Pih et al., 2000). Interestingly, a U3 snoRNP has been identified in *Brassica oleracea*, involved in many processing steps of the pre-rRNA (Sáez-Vasquez et al., 2004). This complex is associated with proteins such as nucleolin, fibrillarin and ribosomal proteins from the small ribosomal particle (Samaha et al., 2010). The pre-40S ribosomal particle is released upon SSU processome disassembly. Similarly, several protein factors are involved in the processing of the pre-60S ribosomal subunit. These ribosome biogenesis factors share homologues in mammals, yeast and plants (reviewed by Sáez-Vásquez and Delseny, 2019). For instance, the guanine nucleotide-binding protein nucleostamin is required for the pre-60S ribosomal particle processing (Romanova et al., 2009; Jeon et al., 2015). Moreover, the Arabidopsis genome

possesses two orthologues of the yeast ribosome biogenesis protein Brx1 (BRX1-1 and BRX1-2), which have redundant functions in Arabidopsis (Shimoji et al., 2012; reviewed by Weis et al., 2015).

Along with pre-rRNA modifications, pre-rRNAs suffer several processing events carried out by exo- and endonucleases to remove the external (5' and 3'ETS) and internal (ITS1 and 2) transcribed spacers (Figure 1.8B; reviewed by Tomecki et al., 2017). In Arabidopsis, the first processing steps include the trimming of the 5' end by the 5'-3' exoribonuclease XRN2, as well as the processing at the P site in the 5'ETS by the U3 snoRNP, rendering the 35S(P) pre-rRNA (Zakrzewska-Placzek et al., 2010; reviewed by Sáez-Vásquez and Delseny, 2019). Subsequently, two different processing pathways co-exist in Arabidopsis: the major ITS1-first pathway and the minor 5'ETS-first pathway (reviewed by Weis et al., 2015; reviewed by Tomecki et al., 2017; reviewed by Sáez-Vásquez and Delseny, 2019). Precursors obtained from both pathways suffer common processing steps, giving the mature 5.8S, 18S, and 25S rRNAs (Figure 1.8B). The protein factors responsible of the pre-rRNA processing are more or less well characterized in plants. For instance, the RNA helicase MTR4 is involved in the removal of the 3' end of the pre-5.8S rRNA (Lange et al., 2011). Moreover, the RNA-binding protein NOB1 participates in the cleavage at the 5'ETS and ITS1 (Missbach et al., 2013). The Arabidopsis Ribonuclease 3-like protein RTL2 mediates the processing of the 3'ETS, even though the precise site has not been mapped (Comella et al., 2008). Another example is the protein complex composed of the Pumilio protein APUM24, Las1-like family protein and the polynucleotide 5'-hydroxyl-kinase GRC3, which takes part in the processing of the ITS2 (Maekawa et al., 2018).

1.4.3.1.3. Formation of pre-ribosomal particles

The last steps of ribosome biogenesis consists of the (i) late processing events of the pre-rRNA molecules and (ii) the formation of pre-ribosomal particles: pre-60S and pre-40S ribosomal particles, which occurs in the GC, (Figure 1.7; reviewed by Nazar, 2004). The processing of the pre-rRNAs culminates with the mature 5S, 5.8S, 18S and 25S rRNAs. Whereas the 5S, 5.8S and 25S form the pre-60S ribosomal particle along with RPLs, the 18S rRNA and RPSs form the pre-40S ribosomal particle (reviewed by Sáez-Vásquez and Delseny, 2019). Then, both pre-40S and pre-60S ribosomal particle are subsequently translocated to the cytosol (Figure 1.7). The latter is further processed by other factors, including the cytoplasmic 60S subunit biogenesis factors REIL1 and REIL2 (Beine-Golovchuk et al., 2018), the eukaryotic translation initiation factor AtelF-6;1 (reviewed by Miluzio et al., 2009) and the GTPase LSG1-2 (Weis et al., 2014).

1.4.3.2. Other nucleolar functions in plants

For many years, the nucleolus was exclusively seen as a ribosome production factory. Nevertheless, this sub-nuclear structure is currently conceived as a multifunctional entity due to (i) the presence of proteins involved in diverse functions in the nucleolus, and (ii) the involvement of nucleolar proteins in other roles rather than ribosome biogenesis (Goessens, 1984; Mélése and Xue, 1995; Shaw and Jordan, 1995; Scheer and Hock, 1999). More specifically, there are certain functions that have been attributed to the plant nucleolus: surveillance of mRNA, assembly and transport of small RNAs, plant growth and development, the course of the cell cycle, interaction with viral material and response to stress scenarios (reviewed by Kalinina et al., 2018).

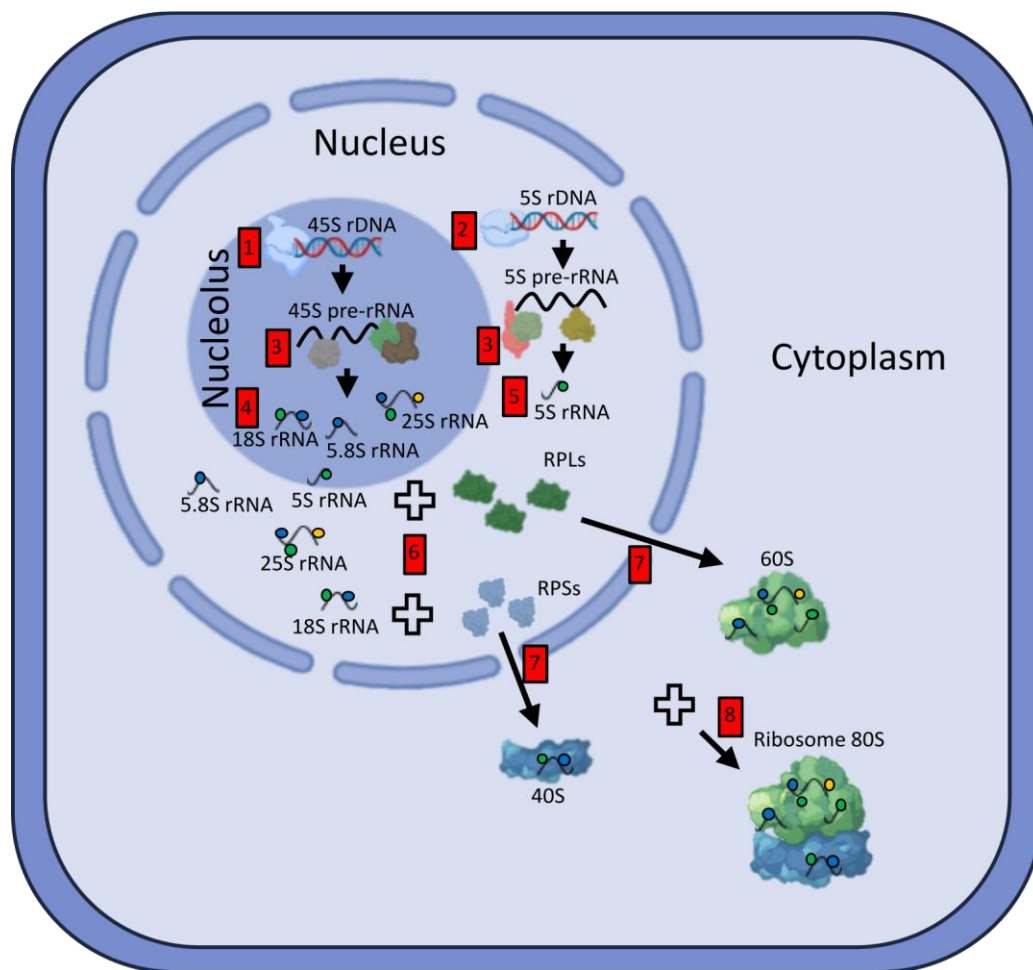
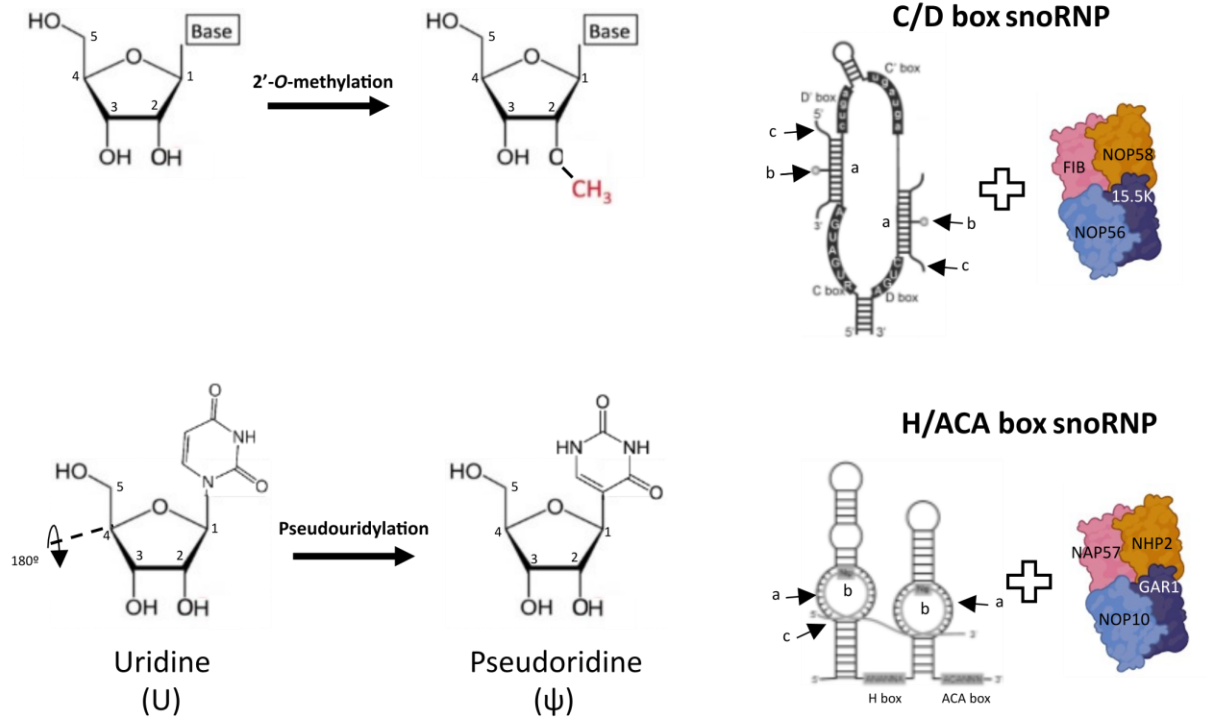


Figure 1.7. Representation of the different steps of ribosome biogenesis in Arabidopsis. In the nucleolus, the 45S rDNA is transcribed by RNA Pol I (1), whereas the 5S rDNA units are transcribed by RNA Pol III in the nucleoplasm (2). The pre-rRNA species are processed by several enzymes and protein complexes (3) to produce the mature 18S, 5.8S and 25S rRNAs (4), as well as the mature 5S rRNA (5), which is then imported into the nucleolus. Whereas the 5S, 5.8S and 25S rRNAs form the pre-large ribosomal particle along with ~47 ribosomal proteins (RPLs), the 18S rRNA and ~33 ribosomal proteins (RPSs) compose the pre-small ribosomal particle (6). The two pre-subunits are exported to the cytoplasm (7), where they are further processed into the 40S and 60S ribosomal particles. Together, they form the 80S ribosome in the cytoplasm (8). The figure was partially designed using BioRender.com.

A



B

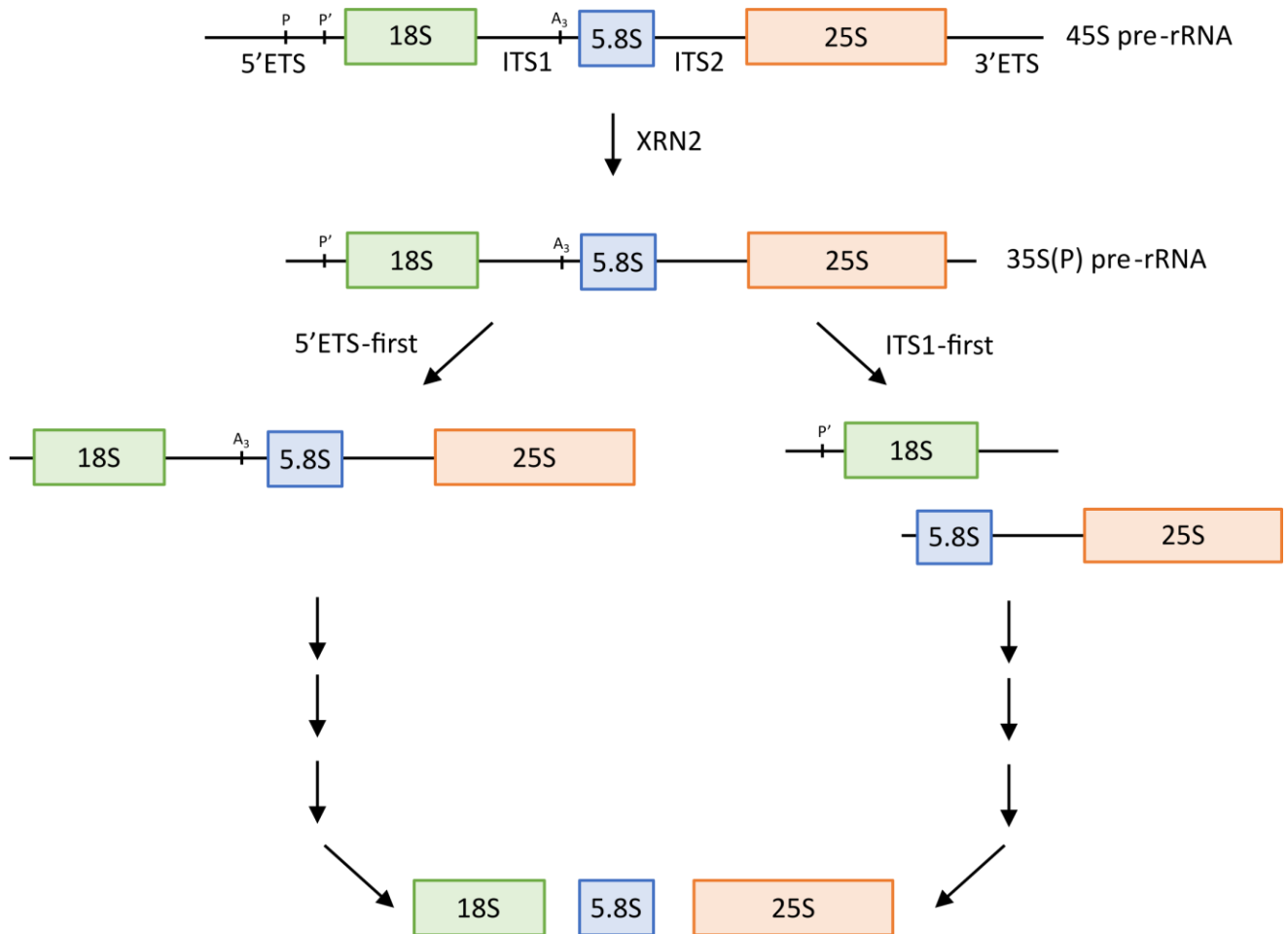


Figure 1.8. Processing steps of pre-rRNA molecules. (A) Main modifications of rRNA. Top: 2'-*O*-methylation consists of the addition of a methyl group to the oxygen of the carbon 2 of the ribose. The reaction is carried out by the C/D box snoRNP, which is composed by C/D box snoRNAs as well as a protein complex. C/D box snoRNAs possess the conserved C, C', D and D' boxes and a sequence complementary to the rRNA (a). The addition of the methyl group and the target rRNA molecule are marked with (b) and (c), respectively. The protein complex is formed of fibrillarin (FIB), NOP56, NOP58 and 15.5K/L7Ae. Bottom: The conversion of uridine into pseudouridine conforms the pseudouridylation, which is catalysed by the H/ACA snoRNP. These RNA and protein complexes contain the H/ACA snoRNAs, with the conserved H and ACA boxes, and the sequence complementary to the target rRNA (a). The conversion to pseudouridine and the rRNA are marked as (b) and (c), respectively. In addition, the protein included in this snoRNP comprise dyskerin (NAP57), NHP2, NOP10 and GAR1, Both C/D box and H/ACA box snoRNAs were modified from Thorenoor and Slaby, (2015). The protein complex were generated using BioRender.com. **(B)** Cleavage events of the 45S pre-rRNA. The first step is the cleavage by XRN2 in the P site to produce the 35S(P) rRNA. After that, two different pathways co-exist in Arabidopsis: the 5'ETS-first (minor) and the ITS1-first (major). The P' and A₃ sites represent the first cleavage site in the 3'ETS-first and ITS1-first pathways, respectively. Both of them conclude with the mature 18S, 5.8S and 25S rRNAs.

1.4.3.3. Nucleolar sequestration of proteins under abiotic stresses in human cells

Studies in human cells revealed the existence of a new post-translational mechanism involving the nucleolus. Upon some types of abiotic stress, such as acidosis or heat shock, certain proteins are retained in the nucleolus, preventing them from exerting their normal function (Figure 1.9A). Three different elements govern the nucleolar sequestration of proteins under abiotic stresses: (i) the nucleolar detention signal (NoDS), (ii) the IGS-derived lncRNAs and (iii) the detention centre (reviewed by Lam and Trinkle-Mulcahy, 2015).

1.4.3.3.1. The Nucleolar Detention Signal (NoDS)

The NoDS is the peptide sequence responsible of the nucleolar sequestration in response to abiotic stress. Initially, the von Hippel-Lindau (VHL) disease tumour suppressor was observed to be reversibly retained in the nucleolus upon acidosis (pH 5.8 – 6.0; Mekhail et al., 2004). Other proteins accumulating in the nucleolus upon low pH include 26S proteasome regulatory subunit SUG1, DNA (cytosine-5)-methyltransferase 1 (DNMT1), and the delta catalytic subunit of DNA polymerase (POLD1; Audas et al., 2012b, 2012a). Sequence analysis aimed to find the region of VHL responsible of the pH-dependent translocation to the nucleolus. The sequence was subsequently named [H⁺]-responsive nucleolar detention signal (NoDS^{H+}), characterised by the presence of, (i) at least, one arginine motif, (ii) two or more hydrophobic triplets, and (iii) arginine motifs located in a low disordered region (¹⁰⁷**RR**HSYRGHLWLFRDAGTHDGLLV¹³⁰, where the arginine motif and the hydrophobic triplets are in bold and underlined, respectively). These three criteria guarantee a high affinity with the nucleolar components (Mekhail et al., 2005, 2007).

Interestingly, other nucleolar sequestration events were previously described in response to other stressors. For instance, the sequestration of the heat shock protein 70 (Hsp70) upon heat shock (42°C) has been described in mammals (Pelham, 1984; Welch and Feramisco, 1984). Subsequently, the subcellular localisation of a synthetic NoDS^{H+}, consisting in one arginine motif and two hydrophobic triplets (**RR**HSYRLLVNQTELFV,

where the arginine motif and the hydrophobic triplets are shown in bold and underlined, respectively), was analysed under acidosis and heat shock (Audas et al., 2012a). The peptide was retained in the nucleolus in all of the stress scenarios. Thus, the motif known as NoDS^{H+} became NoDS. Several nucleolar proteins have been observed to be retained in the nucleolus upon heat stress or acidosis in human cells. These proteins include the RNA Pol I subunits RPA40 and RPA16, the tRNA (cytosine(72)-C(5))-methyltransferase NOL1, the nucleolar MIF4G domain-containing protein NOM1, the protease SENP3, and the ribosomal RNA processing proteins NOP52, PESCADILLO and RRP1B (Jacob et al., 2012).

1.4.3.3.2. IGS-derived lncRNAs

The translocation of NoDS-containing proteins into the nucleolus upon the corresponding stressor implies the interaction of these proteins with particular nucleolar components. Chromatin immunoprecipitation (ChIP) experiments carried out scanning the rDNA units revealed the association of proteins retained in the nucleolus under acidosis, such as VHL, POLD1 or DNMT1, with the 28 kb locus of the IGS, referred as IGS₂₈RNA. Moreover, Hsp70, which is retained in the nucleolus upon heat shock, interacts with the 16 and 22 kb loci of the IGS in heat-shocked cells, which corresponds to lncRNAs IGS₁₆RNA and IGS₂₂RNA, respectively. Thus, the interaction of the NoDS with these IGS-derived lncRNAs was established under the corresponding abiotic stress (Audas et al., 2012a, 2012b).

The IGS-derived lncRNAs are transcribed by RNA Pol I in the same strand as rRNA. The transcripts are post-transcriptionally processed to a smaller form, which accumulates under the corresponding stressor. Furthermore, each of the IGS loci are transcribed specifically upon every type of stress. In other words, the lncRNAs IGS₁₆RNA and IGS₂₂RNA are observed in cells exposed to heat shock, whereas IGS₂₈RNA is only accumulated upon acidosis (Figure 1.9B). Moreover, disruption of one of the loci does not affect the functionality of the other loci (Audas et al., 2012a).

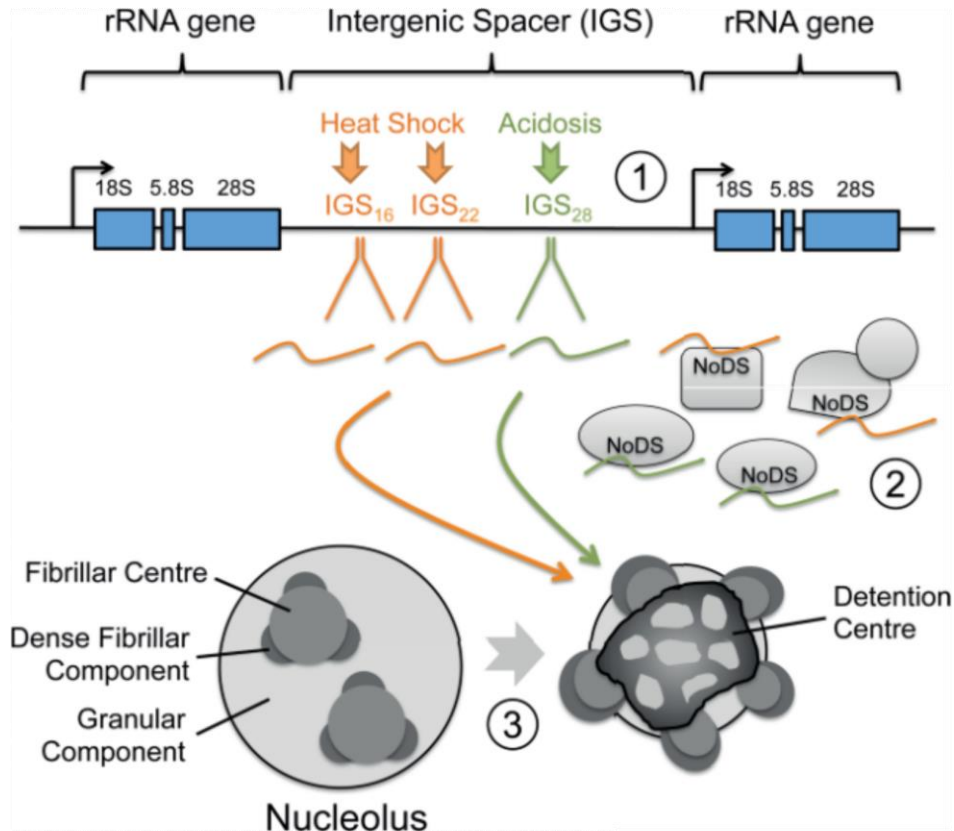
1.4.3.3.3. The detention centre

The retention of NoDS-containing proteins in the nucleolus upon adverse cellular conditions led to the analysis of the nucleolar architecture in these conditions. The proteins B23, fibrillarin and a subunit of the RNA Pol I (RPA194), used as markers for the GC, DFC and FC, respectively, move to the periphery of the nucleolus in response to heat shock or acidosis. On the other hand, NoDS-containing proteins were retained in the nucleolus forming an irregular structure. In contrast to the nucleolar markers, the NoDS-containing proteins are not mobile in the nucleolus under heat shock or acidosis. Thus, the nucleolar sequestration of these proteins leads to the reorganization of the regular nucleolar architecture. This rearrangement of the components of the

nucleolus is the consequence of the formation of the so-called detention centre, where NoDS-containing protein and IGS-derived lncRNAs are immobilised (Jacob et al., 2013). This change in the nucleolar distribution is reversible, as the nucleolar components return to their regular localisation when normal conditions are re-stored.

The detention centre is characterised for being hydrophobic, as its components interact via hydrophobic interactions. The nucleolar sequestration of some proteins involved in ribosome biogenesis leads to a reversible inhibition of ribosome biogenesis in heat-shocked and acidified cells. However, transcription inhibition does not lead to the formation of the detention centre. RNA FISH experiments showed the presence of IGS-derived lncRNAs in the detention centres, as they are fundamental for the structural and functional changes in the nucleolus under heat stress or acidosis (Jacob et al., 2013).

A



B

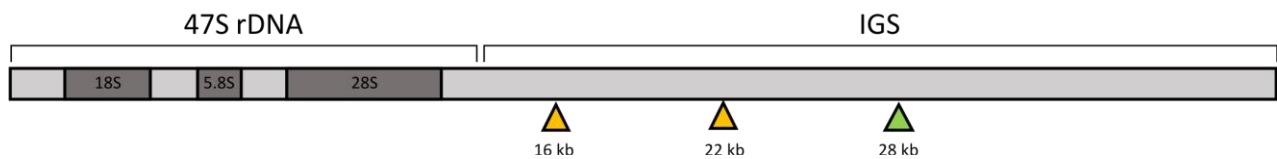


Figure 1.9. Nucleolar sequestration of proteins in human cells. (A) Upon certain abiotic stimuli, such as acidosis or heat shock, certain proteins are sequestered in the nucleolus. (1) Upon heat shock, two IGS-derived lncRNAs, named IGS₁₆RNA and IGS₂₂RNA (orange), are specifically transcribed. Acidosis triggers the transcription of a different IGS-derived lncRNA called IGS₂₈RNA (green). (2) The sequestered proteins possess a specific peptide sequence called NoDS. The IGS-derived lncRNAs interact with the NoDS of the sequestered proteins. (3) The sequestration of proteins in the nucleolus leads to the reorganisation of the nucleolar architecture, forming the detention centre. Figure taken from Lam and Trinkle-Mulcahy, (2015). (B) Scheme of the human 47S rDNA and the IGS. The positions 16 and 22 kb of the pre-rRNA denotes the formation of IGS₁₆RNA and IGS₂₂RNA under heat shock, whereas the position 28 kb marks the generation of IGS₂₈RNA under acidosis.

2. MATERIALS AND METHODS

2.1. Acronyms, abbreviations and chemical formulas

All the chemical compounds named in this thesis follow the nomenclature established by the International Union of Pure and Applied Chemistry (IUPAC). Acronyms and abbreviations correspond to the International System of Units.

2.2. Genetic nomenclature

Regarding gene, mutation, and genotype nomenclature used in this thesis, the guidelines established by Meinke and Koornneef, (1997) were followed. The gene alleles are indicated by their abbreviations in italics. On top of that, the wild-type allele is shown in capital letters whereas the mutants receive lower case letters. Protein names are indicated in normal typography. Different mutant alleles from the same gene are represented by a number separated with a hyphen from the gene abbreviation. In this thesis, the mutant individuals are homozygous for their mutations, unless specified.

Genetic fusions between a promoter and a coding region of a gene are indicated with a single colon ":". Promoters are shown with the abbreviation of the gene followed by the syllable "pro" in subscript. Genetic fusions between the coding regions of two genes are indicated with a hyphen "-". For instance, the expression $35S_{pro}:NoDS_{LAS1}-GFP$ denotes the fusion of the Green Fluorescent Protein (GFP) with the motif NoDS of LAS1 under the control of the 35S promoter from the plant pathogen Cauliflower Mosaic Virus (CaMV).

2.3. Biological material

Arabidopsis thaliana ecotype Columbia-0 (Col-0) was chosen as wild-type. The complementation lines $NRPA3_{pro}:NRPA3^m-FLAG-HA$ (*nrpa3*) and $NUC2_{pro}:NUC2-FLAG-HA$ (*nuc2*), as well as the overexpression lines $35S_{pro}:FIB2-YFP$ (Col-0) and $35S_{pro}:GFP-NoDS_{LAS1}$ (Col-0) had been previously obtained in the team. It must be mentioned that there is a modified version of NRPA3 in the complementation line $NRPA3_{pro}:NRPA3^m-FLAG-HA$ (*nrpa3*), where three cysteines have been replaced by serines (C317S, C320S and C323S). Moreover, the overexpression line $35S_{pro}:GFP$ (Col-0) and the complementation line $OLI2_{pro}:GFP-OLI2$ (*oli2-1*) were kindly provided by Mr. Guillaume Moissiard and Mr. Gorou Horiguchi, respectively. On the other hand, the lines $35S_{pro}:NoDS_{LAS1}-GFP$ (Col-0), $LAS1_{pro}:LAS1-GFP$ (Col-0) and $35S_{pro}:NoDS_{LAS1}-GFP-NLS$ (Col-0) were generated for this project.

As bacterial strains, *Escherichia coli* (*E. coli*) strain DH5 α was used in cloning steps. *Agrobacterium tumefaciens* strain GV3121 was also employed in this project for stable transformation of Arabidopsis. This strain possesses resistance genes for rifampicin and gentamycin.

2.4. Cultures, media and growth conditions

All media for Arabidopsis and bacteria were prepared in distilled H₂O and sterilised by autoclaving (121°C, 101 kPa for 20 min). Thermolabile compounds (e.g., antibiotics) were sterilised by filtration (0.22- or 0.45- μ m pore diameter). These compounds were added after cooling down of an autoclaved medium.

2.4.1. Cultures, media and growth conditions for Arabidopsis

Germination medium contained 1% (w/v) sucrose, 4.4 g/l basal Murashige Skoog (MS) supplemented with vitamin B5 (Duchefa Biochemie M0231), and 0.05% (w/v) 2-(N-Morpholino)ethanesulfonic acid (MES) hydrate (Sigma Aldrich). After dissolving the compounds, the pH was adjusted to 5.7 with 1 M KOH. For MS plates, 1% (w/v) plant agar (Duchefa Biochemie) was added.

Arabidopsis seeds were sown on MS plates. The seeds were sterilised by immersion and mixing with 1 ml of 70% (v/v) ethanol + 0.05% (v/v) SDS for 3 min (Grant Bio PTR-35 360° Vertical Multi-Function Tube Rotator). Residual agents were removed by washing three times with 95% (v/v) ethanol. Then, they were air-dried for a minimum of 2 h. After sowing, the plates were sealed with special air-permeable tape (3M Micropore) and kept at 4°C in the dark for two days for stratification to synchronise germination. Plants were grown in a growth chamber AR-41L2 or E-36L (CLF Plant Climatics) at 22°C under long-day conditions (16/8 h light/darkness) at a light intensity of 180 μ E·m⁻²·s⁻¹.

For seeds sown on soil, a mixture of soil and vermiculite in a volumetric proportion of 5:1 was used. To kill insects, a solution of 0.5 g/l of biological fungicide (Prestop) was used for 10 pots. These pots were kept on trays covered with a plastic layer for seven days, so that plants were able to get used to lower humidity. The trays were kept in climate chambers at 22°C, 60% humidity and long-day conditions. To facilitate seed recollection, plants were covered with a paper bag when starting flowering.

For heat stress, seeds were sterilized and sown on MS plates. After two days on darkness at 4°C, they were transferred to the growth chamber AR-41L2 or E-36L (CLF Plant Climatics) at 22°C under long-day conditions at 180 μ E·m⁻²·s⁻¹ for certain days. Then, plates were transferred to another growth chamber AR-41L2 or E-36L (CLF Plant Climatics) at 37°C under long-day conditions at 180 μ E·m⁻²·s⁻¹.

2.4.2. Cultures, media and growth conditions for bacteria

For *E. coli*, Luria-Bertani (LB) broth high salt [25 g/l of LB powder (Duchefa Biochemie)] were dissolved in water. For bacterial agar plates, LB agar high salt [35 g/l of LB powder (Duchefa Biochemie)] were dissolved in water. For *Agrobacterium tumefaciens*, Yeast Extract Peptone (YEP) medium was prepared using 10 g/l yeast extract (Duchefa Biochemie), 10 g/l bacto peptone (Duchefa Biochemie) and 5 g/l NaCl. The pH was adjusted to 7.0 using HCl. If required, antibiotics were added to a defined concentration (100 µg/ml ampicillin, spectinomycin and cefotaxime; and 50 µg/ml kanamycin, gentamycin, and rifampicin) using appropriate stock solutions (generally 1,000x), stored at -20°C.

For cloning steps, bacterial liquid cultures were prepared in tubes whose capacity exceeded, at least, twice the volume used, so that the aeration during the incubation was guaranteed. These liquid cultures were incubated at 150 rpm, 28°C or 37°C for *Agrobacterium tumefaciens* or *E. coli*, respectively.

Liquid LB or YEP inoculated with bacteria were spread on LB plates by a sterile bacterial spreader. The plates were sealed with parafilm tape. LB plates containing *E. coli* were incubated at 37°C for 24 h, whereas plates with *Agrobacterium tumefaciens* were kept at 28°C for 48 h.

2.5. Primers used in this project

All primers used in this project, synthesised by Sigma-Aldrich, are listed in Table 2.1.

2.6. DNA analysis techniques

2.6.1. Plasmid DNA isolation

Plasmid DNA was isolated from 10-ml bacterial liquid cultures using the Wizard Plus SV Minipreps DNA Purification System (Promega) following the manufacturer's instructions. Typical yields and purity were 200 – 500 ng/µl.

Table 2.1. Primers used in this project¹The *attB* sites are underlined.²The 3' end of these primers was modified with Cyanine 3 (Sigma-Aldrich).³The 5' end of these primers was modified with Cyanine 3 (sigma-Aldrich).⁴*Bam*HI restriction site is in bold.⁵*Sac*I restriction site is in bold, the stop codon is in red, and the NLS of SV40 is underlined.

Name	Purpose	Sequence (5' – 3')
p35	Forward primer for PCR of IGS cDNA	CCCAACTTTACACGAGCTCG
p145	Reverse primer for PCR of IGS cDNA	CAAGCAAGCCCATTCTCCTC
p146	Forward primer for PCR of IGS cDNA	CAACTAGACCATGAAAATCC
p150	Reverse primer for PCR of IGS cDNA	CTAGGTCATTTGACCTGATAC AACATC
p151	Forward primer for PCR of IGS cDNA	AGGACGGCGGAACCCTCGA
p1550	Forward primer with <i>attB</i> sites to amplify <i>LAS1</i>	¹ <u>GGGGACAAGTTTGTACAAAAAGCAGGC</u> <u>TCTGCAATGGGTATATGATCT</u>
p1551	Reverse primer with <i>attB</i> sites to amplify <i>LAS1</i>	¹ <u>GGGGACCACTTTGTACAAGAAAGCTGGG</u> <u>TCAACACAAATTGTAACACTAG</u>
p1583	Forward primer for RNA FISH of IGS transcripts	² GAGACACCGGTTCCAAAATA
p1584	Forward primer for RNA FISH of IGS transcripts	² CCCTCTTGGCACCAACATAC
p1585	Forward primer for RNA FISH of IGS transcripts	² GGCAAAATTCCCACTACACA
p1586	Forward primer for RNA FISH of IGS transcripts	² TCGACGATTCCCTCGGACCCG
p1587	Forward primer for RNA FISH of IGS transcripts	² TCGATGAAGTCTCGGACCTG
p1588	Forward primer for RNA FISH of IGS transcripts	² ATCAAATACTTGGACCCGGT
p1589	Forward primer for RNA FISH of IGS transcripts	² CTATATAGCTTTATAGCCCT
p1590	Forward primer for RNA FISH of IGS transcripts	² ACTCTATGGAGTTTTTGT
p1591	Forward primer for RNA FISH of IGS transcripts	² CACTCCTGGTCGACAAATTT
p1592	Forward primer for RNA FISH of IGS transcripts	² TGGTCGACAAATCCTCGGAC
p1593	Forward primer for RNA FISH of IGS transcripts	² ATCGACCATTCATCGTCCTG
p1594	Forward primer for RNA FISH of IGS transcripts	² CTCGGACACGGTCGATGAAG

p1595	Forward primer for RNA FISH of IGS transcripts	² TCCTGGTCGACGATTTCAAC
p1596	Forward primer for RNA FISH of IGS transcripts	² ATTCCTCGAACCCGATAGAT
p1597	Forward primer for RNA FISH of IGS transcripts	² TTTTTTTTCTCGGACATCCG
p1661	Forward primer for RNA FISH of U3B and U3D snoRNAs	³ TCTAGCCGCACGGTCATGGTTCATCAACCAGGTAA
p1662	Forward primer for RNA FISH of At5gCDbox137.1 (SnoR29-1)	³ AACGCCTAATTCCGGTCTGTTGAGCTTGA AATTTCC
p1663	Forward primer for RNA FISH of U31b and U32b snoRNAs	³ CGTAATGGATGCAVATTAAGCCTGGGGC GACAATTT
p1664	Forward primer for RNA FISH of At1gCDbox33.1 (AtsnoR65)	³ TCTCAATCATGTGATAAAAGGTCATTTTT TCTTCAT
p1665	Forward primer for RNA FISH of At5gCDbox141.1	³ TACCTCAGCGTCAAGATTGTAATCATGTA TATCAGA
p1666	Forward primer for RNA FISH of At4gCDbox117.1	³ GATATCAGTCACTACTTTTAACATTTTTA ATCTTAA
p1667	Forward primer for RNA FISH of At1gCDbox18.1	³ CAGAGATTGGTAGTCAACACTGATACAT CTGTGTGT
p1675	Forward primer to amplify NoDSL _{LAS1} -GFP	⁴ GGATCC CATGCTTGTTCTTGAATATGGACT TGCTGGACTTGTTCTTTGTAGAAGAATT
p1676	Reverse primer to amplify NoDSL _{LAS1} -GFP	⁵ GAGCTCTT <u>AGACCTTCTCTTCTCTTTG GAGCCTTGTACAGCTCGTCCAT</u>
p1677	Forward primer to amplify YLS8	TTACTGTTTCGGTTGTTCTCCATTT
p1678	Reverse primer to amplify YLS8	CACTGAATCATGTTCTGAAGCAAGT

2.6.2. Genomic DNA isolation from Arabidopsis

Arabidopsis genomic DNA was isolated from 100 mg of wild-type leaves using the DNeasy® Plant Mini Kit (Qiagen) following the manufacturer's instructions. Alternatively, a quick and simple method to obtain genomic DNA was followed, generally used in genotyping assays. For that, a 2-mm diameter leaf disk was introduced into an Eppendorf tube containing 150 µl of extraction buffer [200 mM Tris, 250 mM NaCl, 25 mM EDTA and 0.5% (v/v) SDS] at 99°C for 10 min. Samples were centrifuged at maximum speed for 8 min, discarding the pellet. 100 µl of supernatant were mixed with 100 µl of isopropanol, kept at RT for 15 min. After centrifugation at maximum speed for 15 min, the supernatant was discarded. The pellet was washed with 150 µl of 70% (v/v) ethanol. The pellet was then dried at 65°C for 10 min. Finally, 50 µl of H₂O were added, keeping the samples at 65°C for 10 min to enhance re-suspension. DNA must be kept overnight at 4°C before use.

2.6.3. DNA synthesis by Polymerase Chain Reaction (PCR)

PCR amplifications were carried out for different purposes. In most of the amplifications, the GoTaq® DNA polymerase (Promega) was used. Alternatively, the Phusion DNA polymerase (Thermo Fisher Scientific) was required for high fidelity amplifications since its error rate is 50 times lower than the regular Taq DNA polymerase.

The conditions of each amplification depended on the template and DNA polymerase used. The reactions were prepared in 0.2-ml tubes, having a final volume of 25 and 50 µl using the GoTaq® and Phusion DNA polymerase, respectively. In amplifications using the Phusion DNA polymerase, the manufacturer's instructions were followed. For that, the 5x HF Clear Buffer (1x final concentration) was employed, and the final concentration of the deoxyribonucleotides (dNTPs) and both primers were 0.2 mM and 0.5 µM, respectively. Moreover, the amount of template used was 50 ng. For amplifications in which the GoTaq® DNA polymerase was used, the 5x GoTaq® Green Buffer (1x final concentration) was employed, as well as 10 mM dNTPs (0.2 mM final concentration), and 0.7% (v/v) GoTaq® DNA polymerase. The final concentration of both primers was 0.4 µM. Then, up to 4 µl of template were added to the 21-µl reaction mix.

Colony PCRs used a template consisting of a LB liquid culture inoculated with *E. coli* transformants. To that end, the GoTaq® DNA polymerase was employed with the same conditions indicated above. 4 µl of the corresponding liquid culture were added to the PCR tube containing the master mix.

Reactions took place in the Tprofessional Basic thermocycler (Biometra). The melting temperature (T_m) was calculated according to the T_m calculator by Thermo Fisher Scientific. The extension temperature was set according to the guidelines for each DNA polymerase, which normally considered the size of the amplicon.

2.6.4. Restriction of DNA molecules

Analytical digestion was mostly employed with two purposes: (i) to confirm the presence of the correct construct or plasmid, and (ii) to sub-clone the fusion *NoDS_{LAS1}-GFP-NLS* into pCambia1300/35S (2.8.2. Generation of pCambia1300/35S-35S_{pro}:NoDS_{LAS1}-GFP-NLS by restriction enzyme digestion). For analytical digestions, 50-µl reactions were prepared including 1 µg of DNA template, 1 µl of the corresponding 20,000 units/ml restriction enzyme (NEB), 5 µl of 10x Cutsmart buffer (NEB), and water. The reactions were kept at 37°C for 30 min, followed by a thermal inactivation whose duration and intensity depended on the restriction enzyme used.

Occasionally, plasmids were digested during Gateway cloning prior to the BP and LR reactions. These restrictions led to the linearization of the plasmid, which ensured a more efficient recombination (2.8.1. Gateway technology). Likewise, the same digestion conditions were applied as above.

2.6.5. Quantification of nucleic acids

The concentration of both DNA and RNA was determined using a spectrophotometer (Nanodrop 2000, Thermo Fisher Scientific). The absorbance at 260 nm was measured using 1.3 µl of nucleic acid sample and was proportional to the DNA concentration. In addition, the absorbance ratios 260/280 and 260/230 nm were also calculated to determine protein and salt contamination, respectively. Typical A_{260}/A_{280} and A_{260}/A_{230} ratios were 1.7 - 2.0.

2.6.6. Nucleic acids electrophoresis

In general, 1% (w/v) agarose gels were used to visualize DNA molecules. The agarose was dissolved in 1x TAE (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) by heating in a microwave. Electrophoresis were carried out in horizontal gel chambers in 1x TAE. 0.01% (v/v) 5,000x Gel Red (Interchim) were added to the agarose solution prior to solidification. Electrophoresis were carried out at a constant voltage of 100 or 135 V. The visualization of the molecules was achieved by UV light at 312 nm using a gel documentation system (Vilber E-Box CX5 TS Edge).

2.6.7. DNA purification from agarose gels

In some cases, DNA molecules were purified from an agarose gel. For that, the ReliaPrep™ DNA Clean-Up and Concentration System (Promega) was used, following the manufacturer's instructions.

2.6.8. DNA sequencing

Sequencing of DNA samples was carried out by Eurofins using the Sanger sequencing method. DNA samples were diluted to a concentration of 80 - 100 ng/µl in a final volume of 5 µl, to which 5 µl of the sequencing primer (5 µM final concentration) were added.

2.7. RNA analyses techniques

2.7.1. Isolation of total RNA

For the isolation of total RNA, two different strategies were followed. On the one hand, 100 mg of plant tissues (generally whole seedlings) were ground in liquid nitrogen until a fine powder was obtained. The powder was dissolved in 1 ml of TRI Reagent® (MRC) using a vortex, followed by a 5-min incubation at RT. Then, 200 µl of chloroform were added, vortexing for 15 s and keeping the samples at RT for 2 – 3 min. After centrifugation (12,000 rpm for 15 min at 4°C), two different phases were formed in the tube, with a thin layer of DNA separating them. The upper (aqueous) phase, containing the RNA, was taken carefully, and transferred to a new Eppendorf tube. To precipitate the RNA, 500 µl of isopropanol, and 5 µl of 5 mg/ml glycogen (Invitrogen) were added, keeping the samples at RT for 10 min. For maximum precipitation, samples were kept at -20°C overnight. Then, samples were centrifuged at 12,000 rpm for 15 min at 4°C. The supernatant containing cellular debris was discarded, and the pellet was washed with 75% (v/v) ethanol. Finally, the pellet was dried at RT for 15 min, and it was dissolved in 10 - 30 µl, approx., of RNase-free H₂O. As a final step, RNA samples were treated with RQ1 RNase-Free DNase (Promega) in order to remove residual DNA molecules. For that, the purified-RNA sample was incubated with 1 µl of RQ1 DNase, and 1 µl of RQ1 DNase 10x Reaction Buffer at 37°C for 30 min, followed by inactivation with 1 µl of RQ1 DNase Stop Solution for 10 min at 65°C. RNA samples were kept at -80°C. On the other hand, the Monarch® Total RNA Miniprep Kit Protocol (NEB) was also used to extract total RNA, using the manufacturer's instructions. Likewise, isolated RNA was stored at -80°C.

2.7.2. cDNA synthesis

cDNA was synthesised from DNase-treated RNA samples using the Promega RT Kit. To start with, 800 ng of total RNA sample were used to perform the cDNA synthesis, whereas 200 ng of total RNA were used as a control for genomic DNA contamination. For that, 1 µl of oligo (dT) 15 primer and 1 µl of Random Primers were added. Samples were incubated at 70°C for 5 min to avoid the formation of secondary structures. Afterwards, 4 µl of GoScript™ 5x Reaction Buffer (1x final concentration), 2.5 µl of 25 mM MgCl₂, 1 µl of 10 mM dNTP mix (1 mM final concentration), and 0.5 µl of Recombinant RNasin® Ribonuclease Inhibitor were added. Finally, 1 µl of GoScript™ Reverse Transcriptase was added to the cDNA synthesis sample, making a final volume of 19 µl. cDNA synthesis occurred at 25°C for 20 min and 42°C for 60 min, followed by the inactivation of the reverse transcriptase at 70°C for 15 min. Finally, 180 µl of H₂O were added to the cDNA samples, being stored at -20°C.

2.7.3. Real-Time PCR (qPCR)

The accumulation levels of IGS transcripts in wild-type (Col-0) seedlings were quantified by qPCR in a 384-well plate using the PowerTrack™ SYBR™ Green Master Mix (Thermo Fisher Scientific). SYBR™ Green is a dye capable of binding DNA molecules, having an absorption peak at 497 nm. To this end, a mix containing 5 µl of the 2x PowerUp™ SYBR™ Green Master Mix (1x final concentration), and 0.4 µl of 10 µM forward and reverse primers (0.4 µM final concentrations) were pipetted in a final volume of 6 µl. Then, 4 µl of cDNA sample were added to the mix. Two different negative controls were included in each qPCR, those being water, and the equivalent RNA sample to the cDNA sample. Analysis of the RNA sample verified the lack of residual genomic DNA. The endogenous control used was Glyceraldehyde-3-phosphate dehydrogenase GAPC2 (GAPDH, At1g13440), which is a housekeeping gene whose expression levels are constant in different conditions.

The 384-well plate was inserted in a LightCycler 480® (Roche) using the following programme: 95°C, 5 min; 40x (95°C, 15 s; 60°C 1 min); 72°C + 0.04°C/s, 10 min; 40°C, 30 s. Results were analysed using the LightCycler® 480 Software.

An approximation of the Livak method ($2^{-\Delta\Delta C_T}$; Livak and Schmittgen, 2001) was followed to calculate the relative accumulation. The C_T value is the PCR cycle in which the fluorescence is detected by the device. The higher the C_T value, the lower amount of transcripts, since it requires a large number of cycles to start detecting fluorescence. First, the ΔC_T was calculating by subtracting the C_T value of the endogenous control (GAPDH) to the C_T value of the target locus (IGS).

$$\Delta C_T = C_T(IGS) - C_T(GAPDH)$$

Then, the $\Delta\Delta C_T$ value was obtained subtracting the ΔC_T value of the control sample (22°C) to the ΔC_T value of the tested conditions (37°C 4 h, 37°C 24 h and R22°C).

$$\Delta\Delta C_T = \Delta C_T(\text{tested condition}) - \Delta C_T(\text{control condition})$$

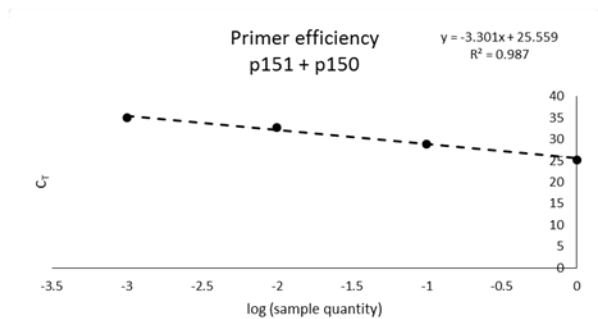
The relative quantification ($2^{-\Delta\Delta C_T}$) is the fold change of the IGS transcript accumulation of the tested condition compared to the control condition. For instance, a $2^{-\Delta\Delta C_T}$ value of 10 indicates that the accumulation of the IGS transcripts in the tested condition is ten times higher than the control condition.

Another important parameter in qPCR assays is the primer efficiency. Ideally, the number of molecules of the target sequence should double during each cycle, corresponding to a 100% amplification efficiency. To calculate the primer efficiency, serial dilutions of the target gene are made, and a qPCR is performed. The C_T values are plotted versus the corresponding concentrations on logarithmic scale. After generating a linear regression curve, the primer efficiency is calculated with the slope of the linear regression curve (m).

$$\text{Primer efficiency (\%)} = (10^{-1/m} - 1) \times 100$$

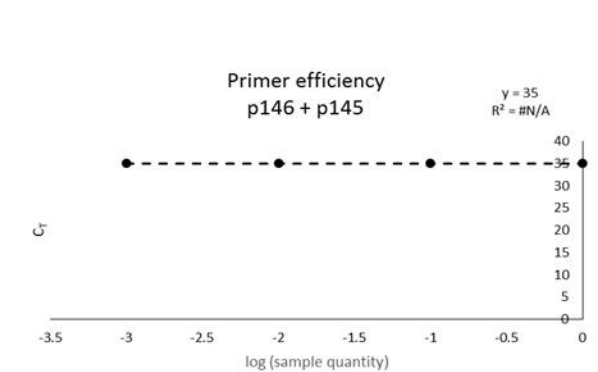
Typically, desired amplification efficiencies range from 90% to 110%. The primer efficiency of the GAPDH primers was 99%, whereas p151 + p150 was 101%. The primer efficiency of the primer pair 146 + p145 could not be calculated (Figure 2.1A - C).

A

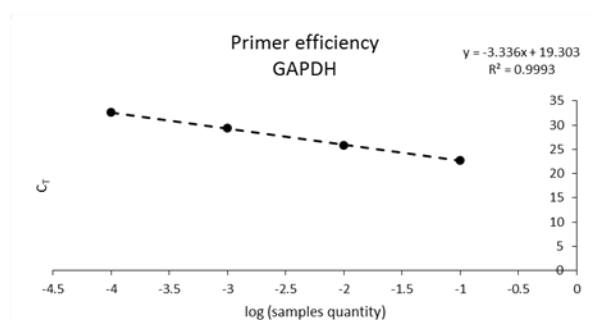


$$\text{Primer efficiency (\%)} = (10^{-1/-3.301} - 1) \times 100 \approx 101\%$$

B



C



$$\text{Primer efficiency (\%)} = (10^{-1/-3.336} - 1) \times 100 \approx 99\%$$

Figure 2.1. Primer efficiency for IGS accumulation. The log of the sample quantity vs the C_T value was represented for p151 + p150 (A), p146 + p145 (B) and GAPDH primers (C). The primer efficiency was calculated for p151 + p150 and GAPDH primers.

2.8. Cloning methods

2.8.1. Gateway technology

The Gateway technology (Thermo Fisher Scientific) was employed to generate the expression constructs pK7FWG2-35S_{pro}:NoDS_{LAS1}-GFP and pGWB404-LAS1_{pro}:LAS1-GFP used in this project. Concerning pK7FWG2-35S_{pro}:NoDS_{LAS1}-GFP, the construct pBlueScript-NoDS_{LAS1} (GeneCust) was already available in our team. A start codon (ATG) was introduced upstream NoDS_{LAS1} to ensure a correct expression. Moreover, an extra guanine was added downstream NoDS_{LAS1} to guarantee the correct frame in the expression construct. This sequence was flanked by *attB* sites. Next, pBS-NoDS_{LAS1} was linearized by digestion with *HindIII* and *EcoRI* prior to the BP reaction (2.6.4. Restriction of DNA molecules). The digestion took place at 37°C for 3 h, followed by thermal inactivation at 80°C for 15 min. The *attB*-flanked NoDS_{LAS1} was inserted into pDONR221 by a BP reaction, which occurred overnight at RT. pDONR221 contains the toxic gene *ccdB* between the *attP* sites, which recombine with the *attB* sites forming *attL* sites (Figure 2.2A). By the replacement of the toxic gene *ccdB*, NoDS_{LAS1} was located within pDONR221, now pENTRY221 (pENTRY221-NoDS_{LAS1}), flanked by *attL* sites.

Afterwards, the LR reaction allowed the transfer of NoDS_{LAS1} to the destination vector pK7FWG2. This vector does contain the *attR* sites, which are able to recombine with *attL* sites in the entry clone (pENTRY221-NoDS_{LAS1}). pK7FWG2 also has the left and right borders (LB and RB, respectively), delimiting the T-DNA, which is transferred to Arabidopsis genome by *Agrobacterium tumefaciens*. Moreover, it possesses both bacterial and plant resistance markers: spectinomycin/streptomycin and kanamycin resistance genes, respectively. On top of that, the sequence of GFP is located upstream the *attR2* site, and the 35S promoter is situated downstream the *attR1* site (Figure 2.2B). Prior to the LR reaction, the entry clone (pENTRY221-NoDS_{LAS1}) was linearized using *HpaI* (2.6.4. Restriction of DNA molecules) at 37°C overnight. Then, the LR reaction occurred at 25°C for 3 h. 0.5 µl of 20 mg/ml proteinase K (Invitrogen) were added to stop the reaction, being kept at 37°C for 10 min.

For the construct pGWB404-LAS1_{pro}:LAS1-GFP, the first step was to amplify the genomic region of *LAS1*, including the promoter and excluding the stop codon, from Col-0 using primers p1550 and p1551 flanked by the *attB* sequences (Table 2.1). Similarly, the *attB*-flanked *LAS1* was inserted into pDONR221 by a BP reaction, which occurred overnight at RT. Then, *LAS1* became flanked with *attL* sites in the entry clone pENTRY221-LAS1_{pro}:LAS1. After that, the LR reaction was performed. The chosen destination vector was pGWB404, which possesses the *attR* sites, LB and RB, bacterial and plant resistance markers (spectinomycin/streptomycin and kanamycin resistance genes, respectively), and the sequence of GFP upstream *attR2* (Figure 2.2C). The LR reaction occurred at 25°C for 3 h, followed to the addition of 0.5 µl of proteinase K to stop the reaction, incubating at 37°C for 10 min.

2.8.2. Generation of pCambia1300/35S-35S_{pro}:NoDS_{LAS1}-GFP-NLS by restriction enzyme digestion

The construct pCambia1300/35S-35S_{pro}:NoDS_{LAS1}-GFP-NLS was generated using the restriction digestion followed by ligation. After amplification of the sequence *NoDS_{LAS1}-GFP* from pK7FWG2-35S_{pro}:NoDS_{LAS1}-GFP using the primer pair p1675 + p1676 and the Phusion DNA Polymerase (Table 2.1; 2.6.3. DNA synthesis by Polymerase Chain Reaction (PCR)), the resulting amplicon (*NoDS_{LAS1}-GFP-NLS*) was purified using the ReliaPrep™ DNA Clean-Up and Concentration System (Promega). For cloning into pGEM-T-Easy (Figure 2.2D), the purified amplicon was subjected to A-tailing. For that, a 10-μl reaction containing 1 μl of purified amplicon, 1 μl of 5x GoTaq® Colourless Buffer, 0.2 μl of 10 mM dATP and 0.2 μl of GoTaq® DNA Polymerase was kept at 70°C for 25 minutes. After subcloning into pGEM-T-Easy (2.8.3. Subcloning into pGEM-T-Easy), the construct pGEM-T-Easy-NoDS_{LAS1}-GFP-NLS and the plasmid pCambia1300/35S (Figure 2.2E) were digested with *Bam*HI and *Sac*I. For that, 1 μg of DNA, 5 μl of 10x Cutsmart buffer, 1 μl of *Bam*HI-HF and 1 μl *Sac*I-HF were mixed in a 50-μl reaction. The reactions were kept at 37°C for 15 minutes followed by thermal inactivation at 65°C for 20 minutes. The desired fragment from the digestion of pGEM-T-Easy-NoDS_{LAS1}-GFP-NLS was purified from the agarose gel (2.6.7. DNA purification from agarose gels). The ligation of the restricted products was performed using 100 ng of digested pCambia1300/35S_{pro}, 27 ng of the digested NoDS_{LAS1}-GFP-NLS, 1 μl of 10x T4 Ligase buffer and 0.3 μl of T4 Ligase in a 10-μl reaction. Ligation reactions were kept at 4°C overnight, and the construct pCambia1300/35S_{pro}:NoDS_{LAS1}-GFP-NLS was obtained.

2.8.3. Subcloning into pGEM-T-Easy

Occasionally, the DNA molecules were subsequently sub-cloned into pGEM-T-Easy (Promega) for further sequencing analysis. For that, 1.5 μl of the DNA molecules were mixed with 2.5 μl of 2x Rapid Ligation Buffer, 0.5 μl of pGEM-T-Easy Vector, and 0.5 μl of T4 DNA Ligase (3 Weiss units/μl). The reactions were left at 4°C overnight. The ligation products were used to transform *E. coli* DH5α cells (2.10.1.1. Transformation of *E. coli* by heat shock). Transformants were selected according to blue-white screening. For this purpose, 50 μl of 20 mg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) were added to the antibiotic-supplemented LB plates right before bacterial inoculation. After growth at 37°C, white colonies were further analysed by colony PCR (2.6.2. Genomic DNA isolation from Arabidopsis). Plasmid isolation of positive white colonies (2.6.1. Plasmid DNA isolation) was followed by sequencing of the inserts of the plasmids (2.6.8. DNA sequencing).

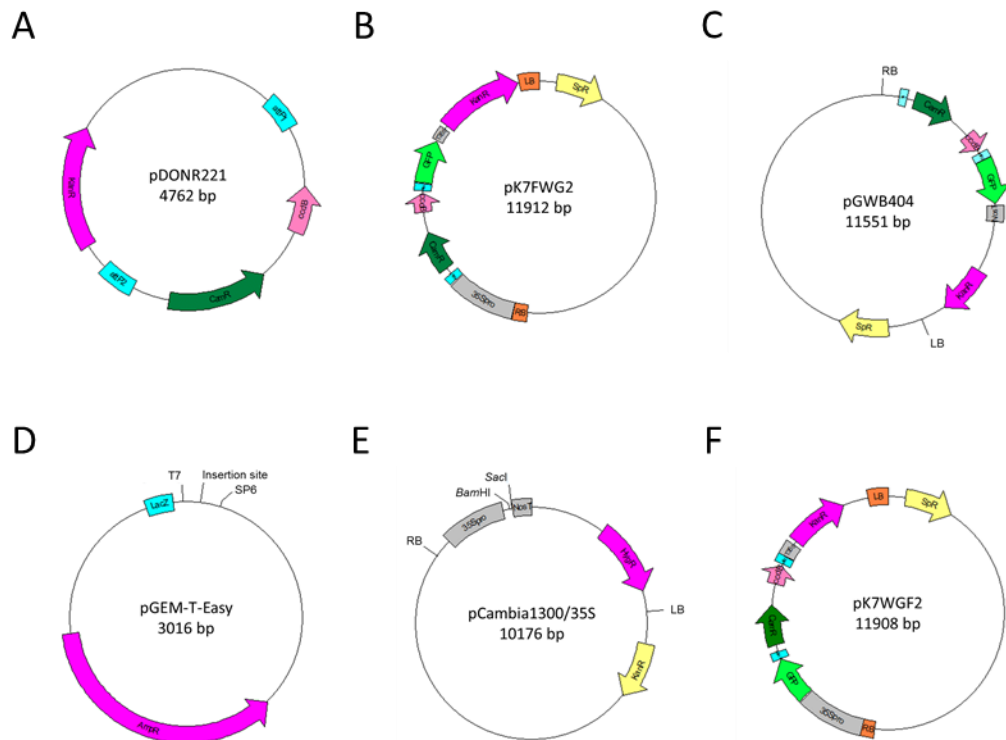


Figure 2.2. Plasmids employed in this project. (A) pDONR221 as donor vector for the BP reaction in Gateway cloning. (B) pK7FWG2 as destination vector for the generation of $35S_{pro}:NoDS_{LAS1}-GFP$. (C) pGWB404 as the destination vector for the generation of $LAS1_{pro}:LAS1-GFP$ plants. (D) pGEM-T-Easy as vector employed for the ligation with A-tailed fragments. (E) pCambia1300/35S for the generation of $35S_{pro}:NoDS_{LAS1}-GFP-NLS$ lines. (F) pK7WGF2 as destination vector for the generation of $35S_{pro}:GFP-NoDS_{LAS1}$ plants. RB: Right Border; LB: Light Border; HygR: resistance to hygromycin; KanR: resistance to kanamycin; CamR: cam repressor; AmpR: resistance to ampicillin; SpR: resistance to spectinomycin.

2.8.4. *In-silico* cloning strategies

Geneious (version 11.1.5) was used to generate all the *in-silico* versions of the constructs created in this project. This software was also employed with other purposes, such as to find restriction sites or check the correct primer design, among others.

2.9. Protein methods

2.9.1. Total protein extraction from Arabidopsis

The extraction of total protein began with the grinding of 100 mg of plant material previously frozen in liquid nitrogen. The fine powder was dissolved in 300 μ l of 2x Laemmli [62.5mM Tris-HCl pH 6.7, 2% (v/v) SDS,

10% (v/v) glycerol, 0.05% (v/v) bromophenol blue, 0.1M DTT], keeping the samples at 65°C for 10 min. After centrifugation at maximum speed for 15 min at 4°C, the supernatant was collected and stored at -20°C.

2.9.2. Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE gels were prepared to separate proteins whose mass is higher than 10 kDa. The upper stacking gel serves to concentrate into a thick band before they move through the second and lower section, the separating gel (Laemmli, 1970). In this project, 7.5, 10 and 12.5% (v/v) acrylamide + bisacrylamide (T) separating gels, and 3% (v/v) T stacking gels were prepared (Table 2.2).

The separating gel was pipetted into a vertical 8.5 x 7.5 x 0.1 cm (width x height x depth) gel caster (Mini-PROTEAN®, Bio-Rad), leaving 1 cm approx. in the upper part. 70% (v/v) ethanol was added on top to form a smooth surface. Once it was polymerized, the ethanol was removed, and the stacking gel was added on top of the separating gel. A 10-well comb was placed into the stacking gel right after.

Gels were placed into a Mini-PROTEAN® Tetra System (Bio-Rad), filled with 1x electrophoresis buffer [2.5 mM Tris base, 19.2 mM glycine, 0.1% (v/v) SDS]. 10% (v/v) β-mercaptoethanol was added to protein samples, heated at 95°C for 2 min in a Thermoblock (88870001 - Digital Dry Bath/Block Heater, Thermo Fisher Scientific) prior to loading to ensure denaturation. 5 µl, approx., of PageRuler Plus Prestained Protein Ladder (Thermo Fisher Scientific) were used as marker. After loading the samples, the gel was run at 20 mA for 1 h, approx., until the dye front reached the bottom of the gel.

2.9.3. Protein detection by immunoblot

The presence of certain proteins was detected in an SDS-gel by immunoblotting, which uses labelled antibodies specific for the protein of interest (Mahmood and Yang, 2012).

An 8.5 x 7 cm 0.45 µm nitrocellulose membrane (Bio-Rad) was used to perform a wet transfer assay. The membrane and SDS-gel as well as blotting paper were soaked in cold 1x transfer buffer [25 mM Tris base, 192 mM glycine, 20% (v/v) of 100% (v/v) ethanol]. Then, a sandwich-like system was assembled, removing any air bubble formed, and placed into a Mini Trans-Blot® Cell (Bio-Rad). The chamber was filled with 1 litre, approx., of cold transfer buffer, and the transfer happened at a constant voltage of 100 V for 90 min.

Once the transfer was stopped, the membrane was incubated with Ponceau Red stain [0.5% (w/v) Ponceau S, 1% (v/v) acetic acid] for 5 min at RT. The membrane was washed with water until the Ponceau Red stain was completely removed. To reduce non-specific binding, the membrane was incubated in blocking

solution [5% (w/v) dry milk powder (Régilait) in PBS-T buffer (20 mM Tris base, 136 mM NaCl, 0.1% (v/v) Tween-20, pH 7.6 with HCl] on an orbital shaker, at RT for 1 h.

To remove non-specific binders, the membrane was washed three times with PBS-T (10 min each) on an orbital shaker at RT. Then, it was incubated with a 1:5,000 dilution (in cold blocking solution) of a primary antibody, at 4°C, overnight. Alternatively, a 1:100,000 dilution of anti-TUB was employed, being incubated at RT for 1 h. This primary antibody solution could be re-used once more.

The membrane was washed again three times with PBS-T, as described above. Then, the membrane was incubated with a 1:10,000 dilution of a secondary antibody [Goat Anti-Rabbit IgG (H+L)-HRP Complex Conjugate, or Goat Anti-Mouse IgG (H+L)-HRP Conjugate; Bio-Rad]. The secondary antibody is able to bind the first antibody. On top of that, this secondary antibody is bound to the horseradish peroxidase, which allows the detection of its binding position in the immunoblot by chemiluminescence. The incubation with the secondary antibody occurred at RT, on an orbital shaker for 1 h. Finally, the membrane was washed three times with PBS-T (see above) prior to the detection.

Table 2.2. Separating and stacking gels preparation

	7.5% T separating gel	10% T separating gel	12.5% T separating gel	3% T stacking gel
30% (v/v) Acrylamide/Bisacrylamide 29:1 solution (Euromedex) (ml)	2.5	3.33	4.2	0.55
1M Tris-HCl pH 8.8 (ml)	3.74	3.74	3.74	-
1M Tris-HCl pH 6.8 (ml)	-	-	-	0.625
H ₂ O (ml)	3.66	2.78	1.91	3.9
20% (v/v) SDS (Euromedex) (μl)	50	50	50	25
10% (w/v) Ammonium persulfate (APS) (μl)	40	40	40	50
Tetramethylethylenediamine (TEMED; Euromedex) (μl)	6	6	6	10

For detection, the membrane was exposed to a 1:1 volumetric mixture of luminol and peroxide (Immobilon Western Chemiluminescent HRP Substrate, Millipore) for a few minutes. Images were acquired using a Vilber Fusion solo 6S Edge V0 70.

Eventually, it was necessary to de-hybridise the membrane to perform detection with another antibody. For that, the membrane was kept in 1x Re-Blot Plus Mild Solution (Millipore) at RT for 15 min in an orbital shaker. After washing three times with 1x PBS-T, the membrane was blocked again, following the same procedure as explain above.

2.9.4. FLAG immunoprecipitation

For FLAG immunoprecipitation, 150 – 300 mg of plant material, *NRPA3_{pro}:NRPA3^m-FLAG-HA (nrpa3)* seedlings in this case, were ground in liquid nitrogen. The powder was re-suspended in two volumes of cold Extraction Buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% (v/v) glycerol, 5 mM MgCl₂, 0.2% (v/v) NP40 and 10 µM of MG132 added prior to use). The mixture was incubated at 4°C for 1 h under rotation (Grant Bio PTR-35 360° Vertical Multi-Function Tube Rotator). Then, it was centrifuged twice at max. speed for 15 min at 4°C, discarding the pellet at each time. A fraction of the supernant (~300 µl) was kept as the “Input”.

The anti-FLAG magnetic beads (Sigma Aldrich) needed to be equilibrated prior to be used. For that, 25 – 50 µl of beads were mixed with two volumes of 0.1N glycine. The tube was placed in a magnetic rack, discarding the supernatant. 10 – 15 volumes of 1x PBS were added to the beads, placing the tube in a magnetic rack to discard the supernatant. The beads were then re-suspended in 25 – 50 µl of Extraction Buffer.

The equilibrated beads were added to the rest of the supernant obtained from the plant material. They were incubated under rotation (Grant Bio PTR-35 360° Vertical Multi-Function Tube Rotator) at 4°C for 90 minutes. Then, the tubes were placed in a magnetic rack, and ~300 µl of the supernant were kept as the “Unbound” fraction. The beads were washed three times with Extraction Buffer at 4°C for 10 minutes under rotation (Grant Bio PTR-35 360° Vertical Multi-Function Tube Rotator). The beads were recovered using the magnetic rack (“IP” fraction). Finally, total RNA extraction or total protein isolation was performed using the “Input”, “Unbound” and “IP” fractions.

2.10. Transformation methods

2.10.1. Bacterial transformation

2.10.1.1. Transformation of *E. coli* by heat shock

E. coli DH5α cells were transformed with different constructs and vectors by heat shock transformation (Chung et al., 1989). For this purpose, a mixture of 5 µl of the corresponding construct, 10 µl of DH5α buffer (0.5 M KCl, 0.15 M CaCl₂, 0.25 M MgCl₂), and 35 µl of water was added to 50 µl of cells. After 30 min on ice, cells

went through the actual heat shock by keeping them at 42°C in a Thermoblock for 40 s, followed by quick transference to ice.

Cells were incubated with 0.5 ml of liquid LB at 37°C, 150 rpm for 1 h. Then, they were centrifuged at 5,000 rpm for 5 min. The supernatant was discarded, and the pellet was re-suspended in 100 µl of liquid LB. Finally, half of the transformation volume was plated on selective LB plates with a sterile spreader. These plates were kept at 37°C overnight to allow the growth of colonies.

2.10.1.2. Transformation of *Agrobacterium tumefaciens* by heat shock

Agrobacterium tumefaciens GV3121 cells were transformed with different destination constructs by heat shock. To that end, 1,000 ng of the destination construct was pipetted into 100 µl of cells. They were kept at 37°C for 5 min, transferring them to ice right after. Then, 1 ml of YEP medium was added, and cells were kept at 28°C for 2 – 4 h at 150 rpm.

Cells were centrifuged at 12,000 rpm for 2 minutes, discarding the supernatant. The pellet was re-suspended in 100 µl of YEP. Finally, 50 and 50 µl of the transformation volume were plated on selective LB plates with a sterile spreader. These plates were kept at 28°C for two days to allow the growth of colonies.

2.10.2. Stable *Arabidopsis* transformation by floral dipping

To create stable *Arabidopsis* transformants, *Arabidopsis* was transformed with *Agrobacterium tumefaciens* by floral dipping. To achieve it, Col-0 seeds were sown on soil until plants started to flower. Then, a 10-ml overnight culture of *Agrobacterium tumefaciens* was prepared with YEP and cells carrying the desired construct overnight at 28°C, 150 rpm. The next day, 200 µl of the overnight culture were added to 200 ml of YEP and the corresponding antibiotics. It was kept at 28°C, 150 rpm overnight. The 200-ml overnight culture was centrifuged at 4,000 rpm for 20 min at 4°C. The supernatant was discarded, and pellet was re-suspended with 150 ml of freshly prepared 5% (w/v) sucrose, 0.05% (v/v) Silwet L-77. At this point, the OD₆₀₀ must be around 0.8.

The transformation solution was poured into an aluminium tray, and Col-0 *Arabidopsis* flowers were dipped into the transformation solution for some seconds. They were placed into a plastic bag, and they were kept on darkness for 24 h. Then, plants were transferred to climate chambers at 22°C, 60% humidity and long-day conditions (2.4.1. Cultures, media and growth conditions for *Arabidopsis*), until siliques were formed.

2.11. Confocal microscopy

To observe the subcellular localisation of different molecules, a LSM 700 confocal microscope (Zeiss) was used. As a nuclear marker, the fluorescent stain 4',6-diamidino-2-phenylindole (DAPI, Sigma Aldrich) was used. This blue-fluorescent stain (405 nm) binds AT regions of dsDNA, being useful in marking the nucleus. A 1:1,000 DAPI:VECTASHIELD® (Antifade Mounting Medium, Vector Laboratories) was added to the samples to allow this marking.

2.11.1. Observation of fluorophore-tagged molecules

This protocol was employed to visualise the subcellular localisation of fluorophore-tagged molecules, such as NoDS_{LAS1}-GFP. For that, stable Arabidopsis transformants were required, so that they possessed the transgene of interest. 5-day-old MS-grown seedlings were fixed in a solution of 2% (v/v) paraformaldehyde [PFM; 32% (v/v) PFM in 1x MTSB (see below)] + 0.2% (v/v) Triton X-100. Vacuum was applied for 2 min, followed by incubation at RT for 5 min in an orbital shaker. Seedlings were washed three times (5 min each) with 1x MTSB (50 mM PIPES, 5 mM EGTA, and 5 mM MgSO₄, pH to 7.0 with KOH) at RT.

The root apex of fixed seedlings was excised, and 5 – 7 apices were placed in 15 µl of 1x MTSB on a microscope slide. Then, 30 µl of 1:1,000 DAPI:VECTASHIELD® was added on top. A coverslip was placed on top, and the excess liquid was removed. The intersection coverslip – slide was sealed using nail polish. The slides were stored at 4°C.

2.11.2. Observation of proteins by immunofluorescence

The visualization of certain proteins was possible via immunofluorescence, when a specific antibody was available for that protein, i.e., nucleolin-1 (NUC1), or when the protein was FLAG-HA-tagged, such as NRPA3^m-FLAG-HA. 5-day seedlings grown on solid MS were fixed in 4% (v/v) PFM [32% (v/v) PFM in 1x PME (see below)] at RT for 45 min. Then, seedlings were washed with 1x PME (50 mM PIPES, 1 mM EGTA, and 5 mM MgSO₄, pH to 6.9 with KOH) three times (5 min each) at RT in an orbital shaker. The root apices were excised, and 8 – 10 apices were placed in 50 µl of digestion solution [1% (w/v) cellulase, 0.5% (w/v) cytohelicase, 1% (v/v) pectolyase, in 1x PME] on a microscope slide. The slides were kept in a humid chamber for 90 min at 37°C. The digestion solution was removed, and the apices were washed three times with 50 µl of 1x PME. Then, the liquid was removed, and a coverslip was placed on top. The apices were crashed by applying pressure for some seconds. The slides were dipped into liquid nitrogen for 10 s, approx., and the coverslip was taken out. The slides could be stored at -80°C several months.

The slides were placed in a coplin jar containing 1x PME at RT for 5 min. The 1x PME was replaced by 1x PME + 0.5% (v/v) Triton X-100, kept at RT for 15 min on an orbital shaker. After washing three times (5 min each) with 1x PME at RT, the slides were immersed in 100% (v/v) ethanol for 10 min at RT. After that, they were submerged in 1x PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄) at RT for 10 min on an orbital shaker. The slides were washed three times (5 min each) with 1x PBS at RT on an orbital shaker. 150 µl of 1:1,000 primary antibody:antibody solution [0.03% (w/v) Bovine Serum Albumin, 0.05% Tween20, in 1x PBS] were added to each slide. They were placed in a humid chamber overnight at 4°C.

The primary antibody solution was removed, and the slides were washed three times (5 min each) with 1x PBS at RT on an orbital shaker. Then, 150 µl of 1:1,000 secondary antibody:antibody solution were added to each slide. They were placed in a humid chamber for 2 – 3 h at RT. After removal of the secondary antibody solution, the slides were washed three times (5 min each) at RT on an orbital shaker. 40 µl of 1:1,000 DAPI:VECTASHIELD® were added to each slide, placing a coverslip on top. The excess liquid was removed, and the joint coverslip – slide was sealed with nail polish. The slides were kept at 4°C.

2.11.3. Observation of RNA molecules by Fluorescence *In-Situ* Hybridization (FISH)

RNA molecules were visualised via RNA FISH using the protocol established by Duncan et al., (2016). For that, fluorophore-marked probes, complementary to the RNA of interest, were generated (Table 2.1).

7-day-old seedlings grown on solid MS were fixed in 2% (v/v) PFM [32% (v/v) PFM in water] at RT for 30 min on an orbital shaker. Seedlings were washed three times (5 min each) with 1x PBS at RT. The root apices were excised, and 5 – 7 apices were placed in 30 µl of 1x PBS on a microscopic slide. Then, the apices were crashed with a coverslip, applying pressure. After submerging the slides in liquid nitrogen for 10 s, the coverslip was removed, and the slides were kept at RT for a minimum of 30 min. Then, they were placed in a coplin jar containing 70% (v/v) ethanol on an orbital shaker for a minimum of 1 h.

An RNase control was established to ensure that the detected molecules were RNA molecules. For that, 100 µl of 100 µg/ml RNase A (Sigma Aldrich) were added to the slide. It was incubated in a humid chamber at 37°C for 1 h. The RNase solution was rinsed in 10 mM HCl for 5 min, and washed twice with 2× (v/v) SSC for 5 min.

RNase- and non-treated slides were washed twice with wash buffer [2x SSC, 10% (v/v) nuclease-free deionized Formamide] for 2 min. 100 µl of hybridization solution [100 mg/ml dextran sulfate, 10 % (v/v) formamide, 2× SSC] with probes (250 nM final concentration, minimum) were then added to each slide. Coverslips were laid over the samples, and the probes were left to hybridize in a humid chamber at 37°C overnight in the dark.

After removal of the hybridization solution, the slides were then washed twice with 200 µl wash buffer. Then, they were immersed in coplin jars containing wash buffer for 30 min at 37°C. 100 µl of 1:1,000 DAPI:VECTASHIELD® were added and left at 37°C for 30 min. The DAPI slution was removed, and 100 µl of 2× SSC were added and removed right after. 100 µl GLOX buffer [0.4 % (w/v) glucose, 10 mM Tris-HCl pH 8.0, 2× SSC] was added to the samples and left to equilibrate for 2 min. Then, it was replaced with 100 µl of anti-fade GLOX buffer containing enzymes [1% (v/v) of 15 mg/ml glucose oxidase (dissolved in 50 mM), 1% (v/v) of Catalase from Bovine Liver, in GLOX buffer]. The samples were covered by coverslips, removing the excess liquid. After sealing the slides with nail polish, the samples were immediately screened under the microscope.

2.12. Image processing

The images acquired from the confocal microscope were further processed using the programme ImageJ. This programme was also employed with other purposes, such as quantifying band intensity from a gel, among others. Moreover, some of the graphs and statistical analysis presented in this thesis have been obtained using RStudio.

3. RESULTS

3.1. Quantitative analysis of the nuclear proteome and transcriptome under heat stress in Arabidopsis

3.1.1. The results of the first section have been included in an article titled “Quantitative proteomic profiling of Arabidopsis nuclei reveals distinct protein accumulation kinetics upon heat stress”, which has been submitted to The Plant Journal. Only supplemental tables 5 and 6 (Tables S5 and S6) are shown, as Tables S1 – S4 contain a considerable amount of data.

3.1.2. Additionally, a comparative analysis with the transcriptome of Arabidopsis under heat stress is detailed right after. This comparative analysis was performed in order to determine any correlation between changes in the proteomic and transcriptomic profiles.

3.1.1. Quantitative proteomic profiling of Arabidopsis nuclei reveals distinct protein accumulation kinetics upon heat stress

Article submitted to The Plant Journal

Quantitative proteomic profiling of Arabidopsis nuclei reveals distinct protein accumulation kinetics upon heat stress

Eduardo Muñoz-Díaz^{1,2}, Isabel Fuenzalida-Valdivia^{3,4,5}, Tommy Darriere^{1,2}, Anne DeBures^{1,2}, Francisca Blanco-Herrera^{3,4,5}, Magali Rompais⁶, Christine Carapito⁶ and Julio Sáez-Vásquez^{1,2*}

¹CNRS, Laboratoire Génome et Développement des Plantes (LGDP), UMR 5096, 66860 Perpignan, France

²Univ. Perpignan Via Domitia, LGDP, UMR 5096, 66860 Perpignan, France

³Centro de Biotecnología Vegetal, Facultad de Ciencias de la Vida, Universidad Andres Bello, Santiago, RM 837-0146, Chile.

⁴ANID - Millennium Institute for Integrative Biology (IBio), Santiago, Chile.

⁵ANID - Millennium Science Initiative Program - Millennium Nucleus for the Development of Super Adaptable Plants (MN-SAP), Santiago 8331150, Chile.

⁶Laboratoire de Spectrométrie de Masse BioOrganique, IPHC UMR 7178, CNRS, Université de Strasbourg, Infrastructure Nationale de Protéomique ProFI - FR2048, Strasbourg, France.

Running head: Nuclei protein profiling of Arabidopsis upon heat

SUMMARY

Heat stress (HS) impacts the nuclear proteome and, subsequently, protein activities in different nuclear compartments. In *Arabidopsis thaliana*, a short exposure to 37 °C leads to loss of the standard tripartite architecture of the nucleolus, the most prominent nuclear substructure, and, consequently, affects the assembly of ribosomes. Here, we report a quantitative label-free LC–MS/MS (Liquid Chromatography coupled to tandem Mass Spectrometry) analysis to determine the nuclear proteome of *Arabidopsis* at 22 °C, HS (37 °C for 4 and 24 hours), and a recovery phase. This analysis identified ten distinct groups of proteins based on relative abundance changes in the nucleus before, during and after HS: Early, Late, Transient, Early Persistent, Late Persistent, Recovery, Early-Like, Late-Like, Transient-Like and Continuous Groups (EG, LG, TG, EPG, LPG, RG, ELG, LLG, TLG and CG, respectively). Interestingly, the RNA Pol I subunit NRPA3 and other main nucleolar proteins, including Nucleolin and Fibrillarin, were detected in RG and CG, suggesting that plants require increased nucleolar activity and likely ribosome assembly to restore protein synthesis after HS.

Keywords: nanoLC–MS/MS, heat stress, nucleus, nucleolus, differential proteomics

INTRODUCTION

As sessile organisms, plants continuously face diverse biotic and abiotic stresses. Global warming and the greenhouse effect are leading to an increase in the average temperature worldwide, and exposure to a temperature 10-15 °C higher than optimal is considered heat stress (HS). HS negatively affects plant growth and development, limiting the photosynthetic rate and germination efficiency (Firmansyah, 2020). This HS condition threatens the yield and production of crop species worldwide. However, plants have developed a series of response mechanisms defining the heat stress response (HSR) (Kotak *et al.*, 2007, Qu *et al.*, 2013, Zhao *et al.*, 2020).

One of the main events of HSR is accumulation of heat shock proteins (HSPs), and transcription of these genes are activated by heat shock factors (HSFs) (Baniwal *et al.*, 2004). HSPs act as molecular chaperones, preventing unfolding of proteins, particularly during HS. They are classified according to their molecular weight as HSP100, HSP90, HSP70, HSP60 and small HSPs (Baniwal *et al.*, 2004, Wang *et al.*, 2004). The most studied HSFs in *Arabidopsis thaliana* are class A HSFs, including HSFA1a or HSFA2, which are critical players in the response to HS. For instance, *Arabidopsis* HSFA2 shuttles from the cytoplasm to accumulate in the nucleus in response to HS (Evrard *et al.*, 2013). In contrast, class B and C HSFs are less well characterized (Nover *et al.*, 2001, Guo *et al.*, 2016). Similarly, the transcription factors bZIP18 and bZIP52, as well as heat shock factor-binding protein (HSBP), translocate from the cytoplasm to the nucleus in response to heat stress in plants (Wiese *et al.*, 2021).

In addition to proteins, noncoding RNAs (ncRNAs) are involved in the response to HS (Zhao *et al.*, 2020). In *Arabidopsis*, HS rapidly induces accumulation of the microRNA miR398, as well as increasing HSFs and HSP levels, whereas miR156 operates in recovery from HS (Guan *et al.*, 2013, Stief *et al.*, 2014). Moreover, HTI1 and HTI2, two small-interfering RNAs (siRNAs), are involved in thermotolerance in *Arabidopsis* (Li *et al.*, 2014). In addition, certain epigenetic marks seem to be involved in the response and adaptation to HS. For instance, HS increases global methylation in the *Arabidopsis* genome (Boyko *et al.*, 2010). *Arabidopsis* with mutation in genes, such as mutants in dicer-like 3 (*dcl3*) or argonaute 4 (*ago4*), among others, involved in RNA-directed DNA methylation show increased sensitivity to heat stress (Popova *et al.*, 2013). Histone modifications are another

factor involved in the response to HS, as some of these modifications are affected by HS in plants (Pecinka *et al.*, 2010, Folsom *et al.*, 2014). Apart from the changes in the transcriptome, proteome and epigenetic marks mentioned above, several effects have been described with regard to the shape, number and composition of different nuclear bodies in plants (Munoz-Diaz and Saez-Vasquez, 2022). Specifically, the effect of HS on the nucleolus in plants has been explored (Hayashi and Matsunaga, 2019, Darriere *et al.*, 2022).

The nucleolus is the most prominent and best-characterized nuclear body in structure and function (Stepinski, 2014, Lafontaine *et al.*, 2021). Typically, it shows a tripartite composition distributed in a vectorial fashion: the fibrillar centre (FC) is surrounded by the dense fibrillar component (DFC), which is embedded in the granular component (GC) (Saez-Vasquez and Medina, 2008, Thiry *et al.*, 2011). A nucleolar cavity at the centre characterizes plant nucleoli (Saez-Vasquez and Medina, 2008). We previously showed that *Arabidopsis* nucleoli rapidly disorganize and disassemble in response to HS. Nevertheless, this disruption is reversible, as the structure of *Arabidopsis* nucleoli slowly recovers when the plants are returned to standard growing conditions (Darriere *et al.*, 2022). In this context, disruption and reassembly of the nucleolus might affect the nuclear proteome. Here, we present a quantitative analysis of the *Arabidopsis* nuclear proteome under and after HS. We describe ten groups of proteins according to their nuclear abundance changes during and after HS. This analysis allowed us to identify significant accumulation of major nucleolar proteins, such as fibrillarin 2 (FIB2), nucleolin 1 (NUC1) and a subunit of RNA polymerase I (RNA Pol I), in the recovery phase after HS. This striking phenomenon suggests the requirement of the nucleolar machinery to resume rRNA transcription and ribosome biogenesis after HS.

RESULTS

Overall nuclear proteome analysis and individual comparisons.

Fifteen-day-old *Arabidopsis* seedlings were exposed to 37 °C for 4 and 24 h and then returned to optimal growing conditions (R22 °C); seedlings maintained at 22 °C were used as untreated control plants (Figure 1A). The heat-treated seedlings (37 °C for 4 h and 37 °C for 24 h) did not show any particular phenotype during the heat treatment or recovery period compared to the untreated mock plants. The integrity of the purified nuclear fraction was verified by light microscopy, and the proteins extracted were analysed by denaturing PAGE before nano LC–MSMS (Figure 1B).

A total of 2837, 2770, 2626 and 3064 *Arabidopsis* protein accessions were detected by nano LC–MSMS in nuclear fractions from plants at treated 22 °C (mock), 37 °C for 4 h, 37 °C for 24 h and R22 °C, respectively (Table S3). Principal component analysis (PCA) showed that the three replicates of each condition grouped optimally, revealing a major contribution of the temperature change (22 °C vs. 37 °C) to the sample variability (Figure 2A). Since the HSR involves activation of HSPs and HSFs (Guo *et al.*, 2016, Ul Haq *et al.*, 2019), accumulation of these proteins was finely examined in the nuclear proteome under different conditions. The heat shock transcription factor A2 (HSFA2) and several HS accumulated upon HS, and for some of them, their level remains higher at R22 °C compared to the untreated plants (Figure 2B).

Differential analysis was conducted to compare the four conditions (22 °C, 37 °C 4 h, 37 °C 24 h and R22 °C), and six pairwise comparisons were made: 37 °C 4 h vs. 22 °C, 37 °C 24 h vs. 22 °C, R22 °C vs. 22 °C, 37 °C 24 h vs. 37 °C 4 h, R22 °C vs. 37 °C 4 h, and R22 °C vs. 37 °C 24 h (Table S1 and S2). The fold change (FC) was

used as a quantitative marker of each comparison. Then, the logarithm of the FC (\log_2 FC) was calculated. For instance, when comparing 37 °C for 4 h vs. 22 °C, a negative \log_2 FC indicates higher accumulation of proteins at 22 °C than at 37 °C for 4 h. On the other hand, a positive \log_2 FC corresponds to higher abundance at 37 °C for 4 h than at 22 °C. A p value was associated with each \log_2 FC to conclude the significance of the change in abundance (Table S2).

Considering this, 126 proteins were significantly differentially accumulated in the nucleus when comparing 37 °C for 4 h vs. 22 °C. Among these 126 proteins, the abundance of 107 increased at 37 °C for 4 h, whereas the abundance of 19 decreased at 37 °C for 4 h (Figure 2C Panel a). Similarly, 344 (228 up and 116 down), 248 (182 up and 66 down), 89 (53 up and 36 down), 178 (103 up and 75 down), and 115 (59 up and 56 down) proteins were differentially accumulated in the nucleus when comparing 37 °C 24 h vs. 22 °C, R22 °C vs. 22 °C, 37 °C 24 h vs. 37 °C 4 h, R22 °C vs. 37 °C 4 h and R22 °C vs. 37 °C 24 h, respectively (Figure 2C Panels b-f and Table S2).

Altogether, this analysis allowed identification of 522 proteins with nuclear abundance that significantly changed in heat-treated (37 °C for 4 h and 37 °C for 24 h) or recovered (R22 °C) plants.

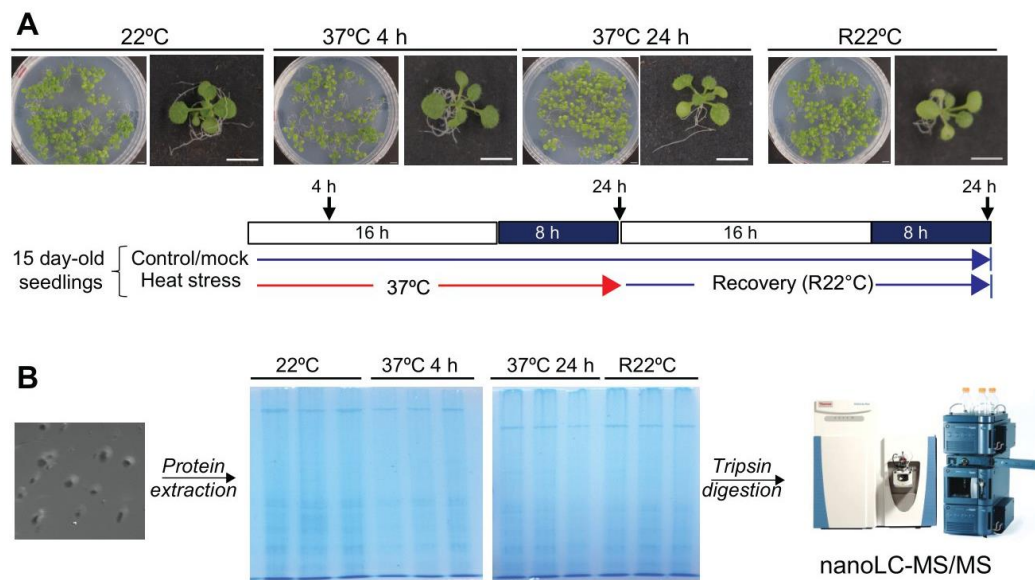


Figure 1. Plant growth conditions, nuclear protein extraction and nanoLC-MS/MS. **A)** Top, 15-day-old *Arabidopsis thaliana* (Col-0) plants at 22 °C (mock), heat-treated (37 °C 4 h and 37 °C 24 h) and then returned to 22 °C for 24 h (R22 °C). The phenotypes of the seedlings are shown for each condition. Scale bar = 0.5 cm. The bottom scheme shows light/dark growing conditions and time collection points of untreated (mock), heat-treated (37 °C 4 h and 37 °C 24 h), and recovered (R22 °C) plant samples **B)** Left, isolated nuclei, where nucleoli appear as dark spherical bodies; middle, SDS-PAGE and Coomassie blue staining of nuclear protein samples collected (as triplicates) at 22 °C, 37 °C 4 h, 37 °C 24 h and R22 °C; Right, nanoLC-MS/MS analysed trypsin digestion protein samples.

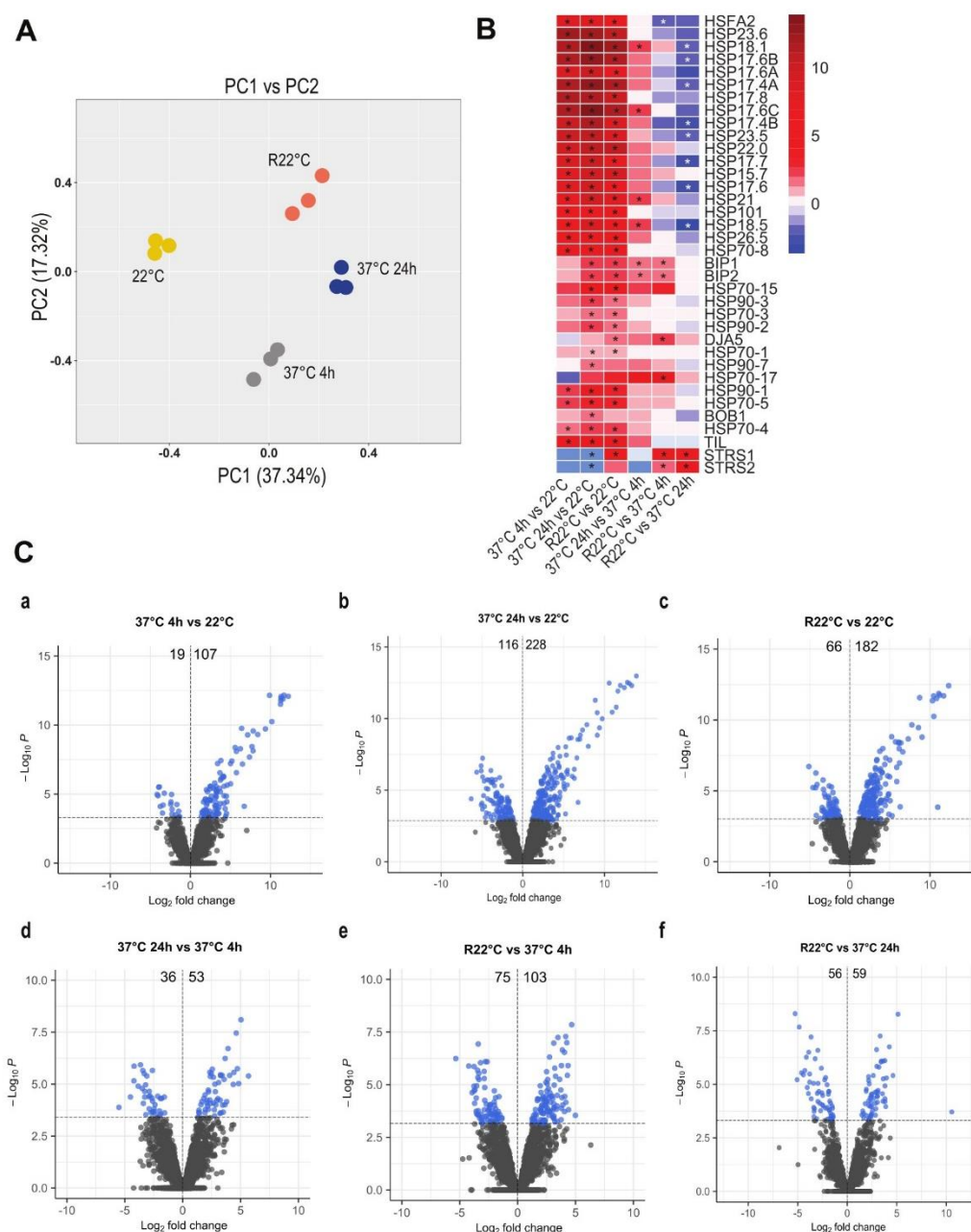


Figure 2. Analysis of the nuclear proteome of Arabidopsis plants under HS. A) Principal component analysis (PCA) of nuclear proteins extracted from three biological replicates of nontreated (22 °C), heat-treated (37 °C for 4 h and 37 °C for 24 h) and recovered (R22 °C) plants. **B)** Heatmaps of log FC of the heat shock factor A2 (HSFA2), heat shock proteins (HSPs), heat-induced TIL, STRS1 and STRS2 proteins at 22 °C, 37 °C 4 h, 37 °C 24 h and R22 °C. Asterisk indicates a significant difference calculated by “Student’s t test” **C)** Volcano plots with the differentially accumulated proteins in comparisons: 37 °C 4 h vs. 22 °C (a), 37 °C 24 h vs. 22 °C (b), R22 °C vs. 22 °C (c), 37 °C 24 h vs. 37 °C 4 h (d), R22 °C vs. 37 °C 4 h (e), and R22 °C vs. 37 °C 24 h (f). The number of proteins that increased (left) or decreased (right) in each comparison is indicated. Tables S1 and 2 provide the numbers of specific peptides used to generate PCA, boxplots and volcano plots.

Kinetics of nuclear accumulation under HS.

The 522 proteins differentially accumulated in the nucleus were grouped into ten clusters according to their abundance (see section Principal component analyses, clustering and GO annotation; Figure S2A). Some of the generated clusters were not conceived as a satisfactory representation of every cluster member, as a more homogeneous expression was pursued for the members of each cluster. Thus, they were not considered in further analyses. Then, the significant differences between the six different comparisons (37 °C 4 h vs. 22 °C, 37 °C 24 h vs. 22 °C, R22 °C vs. 22 °C, 37 °C 24 h vs. 37 °C 4 h, R22 °C vs. 37 °C 4 h, and R22 °C vs. 37 °C 24 h) were considered to group the 522 proteins. In this way, different kinetics were built regarding accumulation of proteins in the nucleus before, during and after HS (Table S5 sheet Statistical). This approach classified 149 proteins (out of 522) into six different major groups: Early Group (EG), Late Group (LG), Transient Group (TG), Early Persistent Group (EPG), Late Persistent Group (LPG), and Recovery Group (RG; Figure 3A and S2B).

Proteins in EG follow “expected” kinetics regarding protein abundance throughout the four conditions (22 °C, 37 °C for 4 h, 37 °C for 24 h and R22 °C). Their protein abundance significantly changed at 37 °C for 4 h and remained stable up to 24 h of heat treatment. Once plants were re-exposed to 22 °C (R22 °C), the level of protein abundance was restored to that before HS (22 °C). Fourteen proteins belonged to EG, including eleven and three proteins with accumulation in the nucleus that increased and decreased during HS in comparison to 22 °C, respectively (Figure 3A panel EG). Then, we performed a heatmap analysis to portray the log₂ FC (up or down) of the six comparisons (37 °C 4 h vs. 22 °C, 37 °C 24 h vs. 22 °C, R22 °C vs. 22 °C, 37 °C 24 h vs. 37 °C 4 h, R22 °C vs. 37 °C 4 h, and R22 °C vs. 37 °C 24 h; Figure 3B). The GO annotations reveal that the majority of the EG members are involved in “Protein binding”, “Hydrolyse and catalytic activities” and “RNA binding”, according to Molecular Function (Figure S3A panel EG). Members of EG include small ubiquitin-related modifier 1 (SUM1; At4g26840), a ubiquitin-like protein involved in stress responses (Kurepa *et al.*, 2003), and mitogen-activated protein kinase 3 (MPK3; At3g45640), which participates in the MAP kinase (MAPK) cascade against pathogens (Asai *et al.*, 2002). The only nucleolar protein found in EG was RNA-binding protein Y14 (At1g51510), which is a core protein of the exon junction complex (EJC) responsible for mRNA splicing (Mufarrege *et al.*, 2011) (Figure 3B panel EG, and Table S5 sheet EG).

Members of LG included specific proteins, the relative abundance of which in the nucleus remained significantly stable from 22 °C to 37 °C for 4 h but significantly changed up to 37 °C for 24 h. As in EG, the abundance of LG proteins in the nucleus returned to normal (22 °C) in the recovery phase (R22 °C). A total of 24 proteins belonged to LG: 19 and five proteins with increases and decreases in their protein abundance, respectively, at 37 °C for 24 h (Figure 3A panel LG). GO analysis of this group mainly revealed proteins involved in “protein binding” and/or “RNA binding” (Figure S3A panel LG), including the transcription initiation factor IIB-1 (TFIIB; At2g41630) (Pan *et al.*, 2000), the splicing factor for HSF and HSP mRNAs STABILIZED1 (STA1) (Kim *et al.*, 2017) and two nucleolar proteins CB-located ribonucleoprotein IMP4 (At1g63780) and cyclophilin 18-1 (At1g01940) (Figure 3B panel LG and Table S5 sheet LG).

The relative abundance of specific proteins from TG significantly changed from 22 °C to 37 °C for 4 h, returning to values observed before HS at 37 °C for 24 h. TG only included nine proteins. The abundance of seven of these proteins increased at 37 °C for 4 h, while the relative abundance of the other two decreased (Figure 3A panel TG). In addition to “Response to chemicals”, GO analysis did not highlight any particular group regarding

Molecular Function or Biological Process. However, most TG proteins are annotated as chloroplastic (Figure S3A panel TG and Table S5 sheet TG). Among TG proteins, we found GUN4 (GENOMES UNCOUPLED 4), which is needed for synthesis of chlorophyll and plastid-nucleus communication (Larkin *et al.*, 2003, Richter *et al.*, 2023), and the 60S acidic ribosomal protein P2-2 (At2g27710), which also mediates protein elongation during translation (Barakat *et al.*, 2001) (Figure 3B panel TG).

The kinetics of EPG resemble those of EG, showing a significant change in protein abundance at 37 °C for 4 h and 37 °C for 24 h compared to 22 °C. However, in contrast to EG, the change in relative abundance of these proteins persisted in the recovery phase (Figure 3A panel EPG). Interestingly, this was the largest group, with 44 proteins, 34 and ten proteins with abundance increases and decreases at 37 °C and R22 °C, respectively. Several HSPs belonged to EPG, including the mitochondrial 23.6 kDa heat shock protein (HSP23.6-MITO; At4g25200), the 17.6 kDa class I heat shock protein 1 (HSP17.6A; At1g59860), and the peroxisomal 15.7 kDa heat shock protein (HSP15.7; At5g37670) (Figure 3B panel EPG and Table S5). In addition to HSPs, the TIL (temperature-induced lipocalin-1) protein, which is involved in basal (BT) and acquired thermotolerance (AT) (Chi *et al.*, 2009) and translocates from the cell membrane to the cytoplasm upon salinity stress (Abo-Ogiala *et al.*, 2014), was identified (Figure 3B panel EPG; Table S3 sheet EPG). Overall, this is consistent with the fact that the major GO annotations were “Response to stress” and “Protein binding” as Biological Process and Molecular Function, respectively (Figure S3B panel EPG).

As with LG, LPG corresponded to a group of proteins with significant changes in nuclear abundance at 37 °C for 24 h compared to 22 °C and 37 °C for 4 h. However, in contrast to LG, the change in protein abundance in LPG was maintained after HS (R22 °C). LPG contained 24 proteins, 17 with increasing protein abundance from 22 °C and 37 °C 4 h to 37 °C 24 h and seven with decreasing protein abundance (Figure 3A panel LPG). The most common GO molecular functions were “protein binding” and “catalytic activity”, and “response to stress” was the main Biological Process (Figure S3B panel LPG). Similar to EPG, chloroplastic proteins were enriched in LPG (Figure S3B panel LPG Cellular Component). Among LPG, we found zeaxanthin epoxidase (ZEP; At5g67030), which is involved in ABA biosynthesis, and the protein export trigger factor-like protein (TIG1; At5g55220). The THO complex subunit ALY4 (ALWAYS EARLY 4), which participates in export of RNA molecules from the nucleus (Pfaff *et al.*, 2018), was also found in LPG (Figure 3B panel LPG, and Table S5 sheet LPG).

Finally, in RG, the abundance of specific proteins in the nucleus did not vary significantly during HS (37 °C for 4 and 24 hours), but there was a significant change during the recovery period (R22 °C). Thirty-two of 34 proteins showed increased protein abundance at R22 °C, whereas the protein abundance of the other two decreased (Figure 3A panel RG). The most common GO Biological Process terms were “nucleobase-containing compound” and “response to stress”; the most abundant Molecular Function terms were “RNA binding” and “catalytic activity”. Interestingly, GO Cellular Component included a considerable number of nucleolar proteins (Figure S3B panel RG), such as the ribosome biogenesis proteins atBRX1-2 (BRX1 homologue 2) (Weis *et al.*, 2015), APPAN (Arabidopsis PETER PAN-LIKE PROTEIN) (Choi *et al.*, 2023), NSN1 (Nucleostemin-like 1) (Jeon *et al.*, 2015), OLI2 (OLIGOCELLULA 2) (Kojima *et al.*, 2017), ATNUG2 (NUCLEAR/NUCLEOLAR GTPASE 2) (Im *et al.*, 2011), and AC40/NRPA3 RNA Pol I subunit (Ream *et al.*, 2015) (Figure 3B panel RG and Table S5 sheet RG).

In conclusion, we identified six groups of proteins showing nuclear accumulation profile changes in an early (EG, EPG and TG) or late (LG and LPG) manner during exposure to 37 °C or after HS (RG). Although there were 29

of 149 proteins (approx.. 20%) with nuclear abundance decreases during the HS and/or recovery phase, the majority (120 of 149 proteins or approx.. 80%) listed in these six groups showed an increase in nuclear contents during and/or after HS.

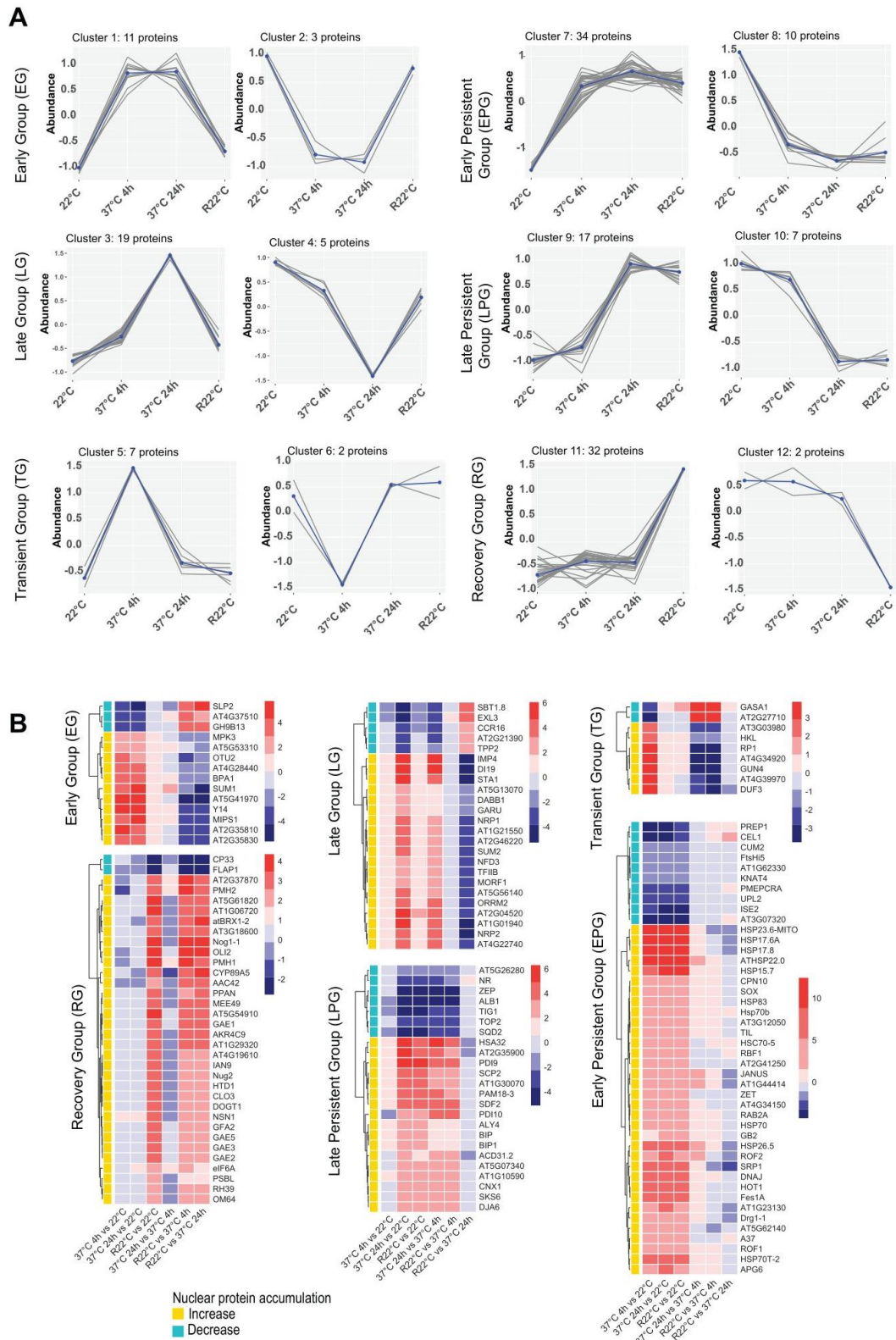


Figure 3. Nuclear protein accumulation under HS and recovery conditions. **A)** Line charts show the dynamics of the 149 proteins at 22 °C, 37 °C and R22 °C in six groups: Early Group (EG), Late Group (LG), Transient Group (TG), Early Persistent Group (EPG), Late Persistent Group (LPG) and Recovery Group (RG). For each group, the left and right charts represent the amount of proteins with increasing and decreasing accumulation, respectively. The blue lines represent the overall tendency, whereas the grey lines consist of the trajectory of each member of group. **B)** Heatmaps show the log2 FC for every protein from each group in each temperature/time comparison. The *p* values (in parenthesis) are for 37 °C 4 h vs. 22 °C (0.0005); 37 °C 24 h vs. 22 °C (0.001); R22 °C vs. 22 °C (0.001), 37 °C 24 h vs. 37 °C 4 h (0.0004), R22 °C vs. 37 °C 4 h (0.0007), and R22 °C vs. 37 °C 24 h (0.0005). FDR = 0.98-1.14%. Proteins with nuclear abundance increases or decreases during exposure to 37 °C are shown in yellow and blue, respectively.

Identification of Early-, Late-, Transient-Like and Continuously Induced Groups.

As mentioned above, 522 proteins were differentially accumulated in the nucleus in at least one of the six comparisons detailed above (Tables S2 and S4). However, only 149 of 522 proteins were classified as EG, EPG, LG, LPG, TG and RG (Figures 3 and S3). Therefore, the 373 remaining proteins were reclustered based only on their protein abundances in the nucleus at 22 °C, 37 °C and R22 °C without considering significant changes for each comparison (see section Principal component analyses, clustering and GO annotation). Eight clusters (C1-C8) were generated (Figure 4A), which were grouped according to their similarity with EG, EPG, LG, LPG, TG and RG. As a result, four additional groups were obtained: the Early-Like Group (ELG), comprising Clusters C1, C4 and C8; the Late-Like Group (LLG), comprising Clusters C2, C3 and C7; the Transient-Like Group (TLG), comprising Cluster C6; and the Continuous Group (CG), comprising Cluster C5 (Figure 4B).

Nuclear protein accumulation at 22 °C, 37 °C and R22 °C in ELG was similar to that in EG or EPG. ELG included 158 proteins (Figures 4B panel ELG and S4A, and Table S6 sheets ELG C1, ELG C4, and ELG C8). The main GO Molecular Function terms were “Protein binding” and “Catalytic activity” (Figures 4C and S5A). Some HSPs are present in ELG, similar to EPG. In addition, nucleolar proteins were abundant, including IRP8 (INVOLVED IN rRNA PROCESSING 8) (Palm *et al.*, 2019), PCP1 (PLANT-SPECIFIC COMPONENT OF THE PRE- rRNA PROCESSING COMPLEX1) (Ishida *et al.*, 2016), and the small nuclear ribonucleoprotein SMD1B involved in RNA splicing, RNA quality control (RQC), and posttranscriptional gene silencing (PTGS) (Elvira-Matelot *et al.*, 2016) (Figures 4C and S5A, and Table S6 sheets ELG C1, ELG C4 and ELG C8).

For LLG, accumulation of specific proteins in the nucleus throughout the four different conditions was similar to that in LG and LPG (Figure 4B panel LLG). We identified 94 proteins in this group (Figure S4B and Table S6 sheets LLG C2, LLG C3 and LLG C7). The main GO Molecular Function terms were “Protein binding”, “RNA binding” and “Catalytic activity” (Figures 4C and S5B). Among them, we distinguished the nucleolar, nucleoplasm, and chromenter-localized proteins STRS1 and SRTS2 (STRESS RESPONSE SUPPRESSORS 1 and 2). Interestingly, STRS1 and 2 are RNA helicases that attenuate the abiotic stress response (Khan *et al.*, 2014). LLG also contained nucleolar proteins, PONTIN/RIN1, a plant orthologue involved in telomerase assembly in the nucleolus (Schorova *et al.*, 2019), and RNA Pol I, II, III and IV common subunit NRPB5 (Ream *et al.*, 2015) (Table S6 sheets LLG C2, LLG C3 and LLG C7).

TLG comprised 49 proteins from Cluster C6 (Figures 4B panel TLG and S4C, and Table S6 sheet TLG C6). This group displayed kinetics of nuclear protein accumulation similar to that of TG (Figure 4B panel TLG). TLG contained chloroplastic, cytoplasmic, and nuclear proteins (GO Cellular Component) and various “Response to stress” or “Biosynthetic processes” (GO Biological Process) proteins; Figures 4C and S5C).

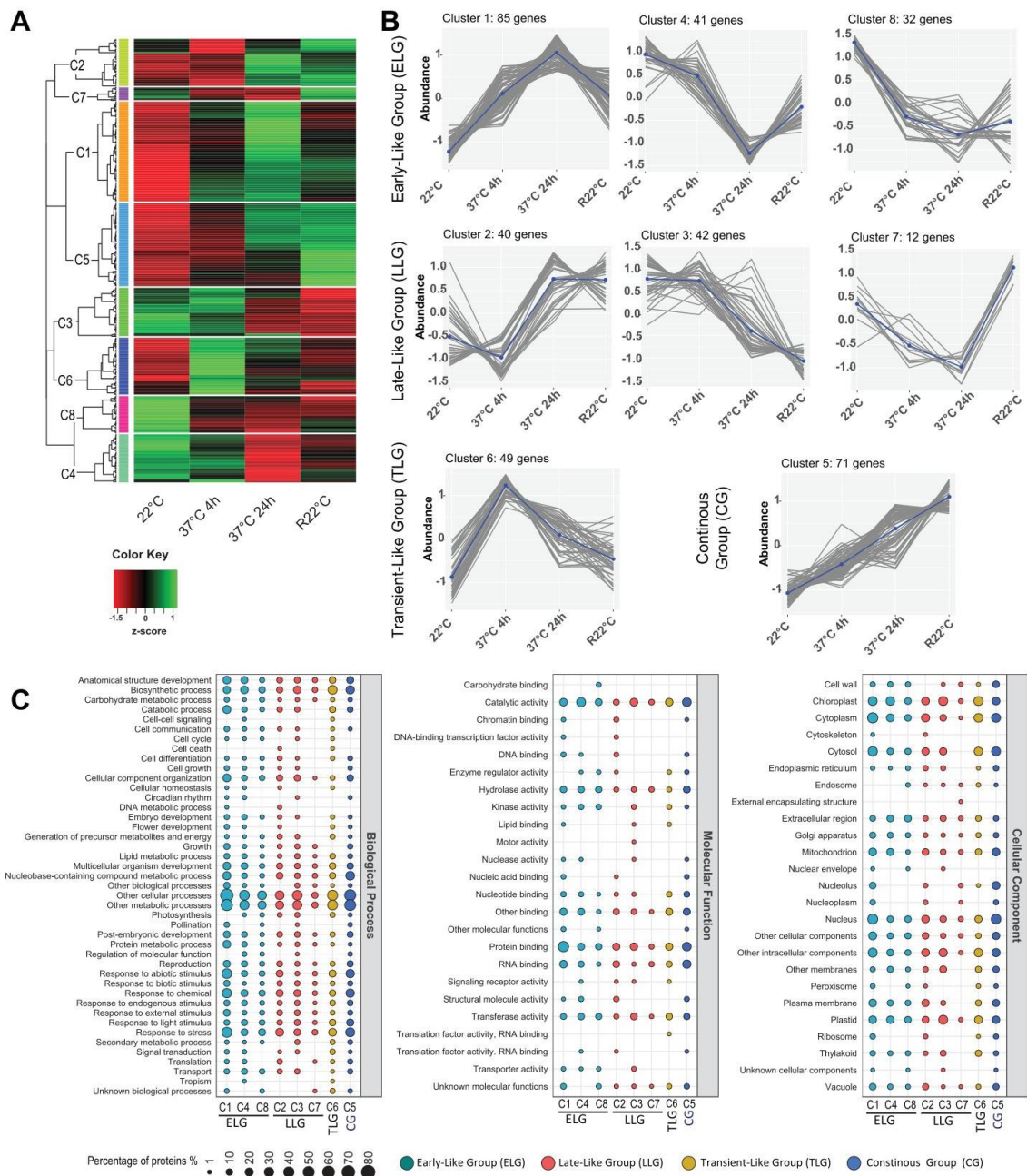


Figure 4. Protein accumulation in the nucleus of Early-Like (ELG), Late-Like (LLG), Transient-Like (TLG) and Continuous (CG) Groups **A)** Heatmap and clustering (C1-C8) of the 373 proteins differentially accumulated in response to HS but not classified in EG, LG, TG, EPG, LPG or RG. **B)** Line charts of Clusters C1, C4 and C8, from early-like (ELG); Clusters C2, C3 and C7, from late-like (LLG); Cluster 6, from transient-like (TLG); and Cluster C5, from continuous (CG) groups are shown. Coloured squares for each cluster are given in the heatmap and line charts. **C)** GO annotation of Clusters 1-8 in ELG (cyan), LLG (red), TLG (yellow) and CG (blue).

CG did not resemble any of the groups described above (EG, LG, TG, EPG, LPG or RG). In CG, the abundance of 71 proteins showed continuous accumulation during the four conditions (Figures 4B panel CG and S4D, and Table S6 sheet CG C5). The GO annotation revealed enrichment of “Response to stress”, “Response to chemical” and “Response to abiotic stimuli” as biological processes. Moreover, Molecular Function terms included “Protein binding” and “RNA binding”; Cellular Component terms included predominantly nucleus and cytosol/cytoplasm (Figures 4C and S5D). Notably, CG contained numerous nucleolar proteins involved in rRNA processing and ribosome assembly (Table S6 sheet CG C5), including RNA Pol I subunit A43/NRPA7 (Ream *et al.*, 2015), ARPF2 (Arabidopsis Ribosome Production Factor 2) (Choi *et al.*, 2020), FIB1/2 (Barneche *et al.*, 2000), NUC1 (Pontvianne *et al.*, 2010), ATFKBP53 (FK506 BINDING PROTEIN 53) (Li and Luan, 2010), C/D snoRNP subunit NOP56 (NUCLEOLAR PROTEIN 56) (Guo *et al.*, 2023), GAR2 and ATRH7 (Arabidopsis RNA HELICASE 75)(Huang *et al.*, 2016). Uncharacterized nucleolar proteins, such as Gar1/Naf1 (At3g03920), a subunit of the H/ACA RNP subunit, and U3-containing 90S preribosomal complex subunit (At2g43110), were also found in CG (Table S6 sheet CG C5).

Taking these results into consideration, this analysis allowed us to identify three novel groups of proteins – ELG, LLG and TLG – displaying protein accumulation kinetics similar to those of EG, EPG, LG, LPG and TG. Furthermore, we identified an additional group (CG) showing novel kinetics of protein accumulation in the nucleus in response to and after HS.

Western blot analysis of nucleolar protein from RG and CG

We addressed whether the changes in nuclear accumulation observed in the HS and recovery periods were due to an increased or reduced amount of total cellular protein. We focused on nucleolar proteins belonging to RG and CG. This is mainly due to the significant occurrence of nucleolar proteins in these two groups (Figures S3B panel RG and S5D) and because the nucleolus, the most prominent and multifunctional nuclear structure, becomes disorganized during HS and reorganizes during the recovery period (Darriere *et al.*, 2022).

Nucleolus assembly greatly depends on rRNA synthesis; therefore, we examined protein accumulation of three major factors involved in the transcription and processing of rRNA (Figure 5). NRPA3 (AAC42/ATRPAC42), one of the subunits of RNA Pol I (Ream *et al.*, 2015) detected in RG, and NUC1 (Kojima *et al.*, 2007, Pontvianne *et al.*, 2007, Pontvianne *et al.*, 2010) and FIB2 (Barneche *et al.*, 2000), both detected in CG (Tables S5 sheet RG and S6 sheet CG). In addition, fibrillarin interacts with nucleolin in a U3 snoRNP complex needed for the first processing step of the 45S pre-rRNA (Saez-Vasquez *et al.*, 2004, Samaha *et al.*, 2010).

First, we performed western blotting using total protein extracts from *NRPA3^{pro}:NRPA3^m-FLAG-HA* (*nrpa3*) plants (Figure S1) and antibodies against the epitope tag HA. At R22 °C, the NRPA3^m-FLAG-HA signal was substantially stronger (1.7X) than the signals detected at 22 °C. Western blot analysis also detected a slight reduction in NRPA3^m-FLAG-HA (0.7X - 0.8X) during HS compared to 22 °C (Figure 5 panel NRPA3).

Second, to determine the amount of CG proteins NUC1 and FIB2 at 22 °C, 37 °C 4 h, 37 °C 24 h and R22 °C, western blot analyses were performed using either Col-0 Arabidopsis extracts and a specific NUC1 antibody or Arabidopsis *35S_{pro}:FIB2-YFP* (Col-0) plants and an antibody against GFP. On the one hand, NUC1 protein levels increased at 37 °C for 24 h (1.5X) and during the recovery period (1.8X), in agreement with previous reports

showing accumulation of NUC1 at 37 °C for 24 h (Darriere *et al.*, 2022). On the other hand, FIB2-YFP accumulated at 4 h (1.5X) and 24 h (6X) at 37 °C and at R22 °C (9X) in *35S_{pro}:FIB2-YFP* Arabidopsis plants (Figure 5 panels NUC1 and FIB2). The increase in accumulation of FIB2-YFP was unexpectedly high (9X) in comparison to other nucleolar proteins, which might be due to the increased stability of chimeric FIB2-YFP.

Finally, we determined accumulation of the EPG protein TIL in Col-0 protein extracts. Western blot analysis shows that TIL rapidly accumulated after 4 h (3X) and 24 h (4X) at 37 °C. Moreover, the abundance of TIL in the nucleus at R22 °C remained higher than that at 22 °C (3.3X; Figure 5 panel TIL).

These results indicate that nuclear accumulation of the RG protein NRPA3, CG proteins NUC1 and FIB2 and EPG protein TIL is likely due to increased expression in response to HS and the recovery period. Finally, our western blot results agreed with the observed accumulation of specific NRPA3, NUC1, FIB2, and TIL peptides detected by LC–MS/MS analysis (Tables S1 and S2).

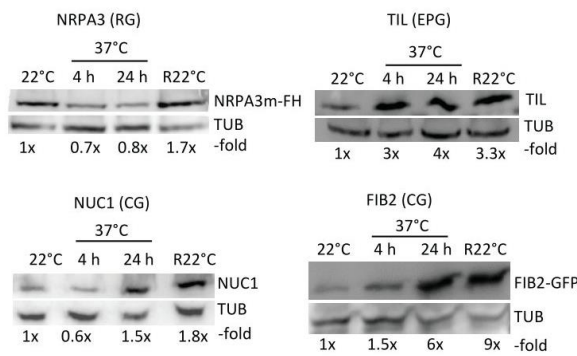


Figure 5. Nuclear protein accumulation of NRPA3, NUC1, FIB2 and TILT. A) Western blot using total protein extracts from Col-0, *NRPA3_{pro}:NRPA3^m-FLAG-HA* (*nrpa3*), and *35S_{pro}:FIB2-YFP* (Col-0) plants at 22 °C, 37 °C (4 and 24 h) and R22 °C. Western blots were performed using α-HA, α-NUC1, α-GFP, and α-TIL to detect NRPA3^m-FLAG-HA (RG), NUC1, FIB2-YFP (CG) and TIL (EPG) proteins, respectively. α-TUBULIN was used to verify protein loading and quantification (using ImageJ software) of the NRPA3^m-FLAG-HA, NUC1, FIB2-YFP, and TIL band intensities. The relative amount of each protein (x-fold) compared to 22 °C is given for each temperature (22 °C, 37 °C 4 h, 37 °C 24 h, and R22 °C). Expected sizes for NRPA3m-Flag-HA ~50 kDa, NUC1: ~60 kDa; FIB2-YFP: ~65 kDa; TIL: ~21 kDa and TUB: ~50 kDa are indicated.

DISCUSSION

In this study, we performed a quantitative nanoLC–MSMS analysis of the nuclear proteome of Arabidopsis at 22 °C, 37 °C for 4 h, 37 °C for 24 h and R22 °C. Our analysis revealed pronounced and relatively rapid changes in the proteomic profile of the nucleus upon HS. To our knowledge, this is the first label-free quantitative analysis of the Arabidopsis nuclear proteome under sustained HS. The proteome of whole-cell extracts of Arabidopsis plants exposed to extreme and short HS stress leading to seedling survival or death has been examined

(Echevarria-Zomeno *et al.*, 2016). Other proteomic studies using whole-cell extracts have also been carried out under HS in other plants, including spinach (Zhao *et al.*, 2018), tomato (Keller *et al.*, 2018) and *Clematis florida* (Jiang *et al.*, 2020).

In our investigation, ~20% of proteins detected by nanoLC–MSMS (522 of 2,629) exhibited differential nuclear accumulation during HS and/or the recovery stage (Figure 6). Notably, in response to HS and/or recovery conditions, most proteins (366 proteins) increased in nuclear accumulation, whereas the others (156 proteins) showed decreases. Consequently, these changes might induce or inhibit nuclear protein activities in response to HS. Remarkably, most of these proteins are related to diverse stress responses, including light intensity and redox (Figure S6), which are interrelated with HS (Karayekov *et al.*, 2013, Dard *et al.*, 2023).

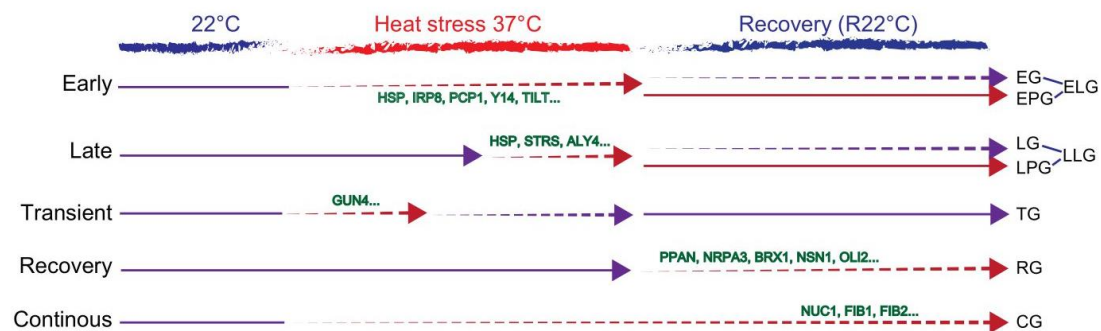


Figure 6. Illustration of the five tendencies of protein accumulation upon heat stress (37 °C) and recovery (R22 °C) conditions. Early (EG, EPG and ELG), late (LG, LPG and LLG), transient (TG and TLG), recovery (RG) and continuous (CG) tendencies are represented. Arrows with continuous lines indicate no changes in protein accumulation, and arrows with dashed lines indicate significant changes (increase or decrease) in protein abundance. Red arrows show changes in protein abundance compared to 22 °C (before stress), while purple arrows show protein abundances similar to 22 °C.

Most HSPs were found in EPG or ELG, as they strongly accumulated in the nucleus upon HS and remained relatively stable during recovery. Furthermore, other proteins that accumulated in the nucleus under HS were detected, including STRS2 (Khan *et al.*, 2014) and TIL (Chi *et al.*, 2009), which are involved in HSR and thermotolerance, respectively. In addition to HSFs, HSPs, and other heat-induced proteins, we did not identify robust groups of proteins associated with specific molecular functions and/or biological processes according to their GO annotations. Nevertheless, the categories “Response to stress” (Biological Process) and “Protein binding”, “RNA binding” and “Catalytic activity” (Molecular Function) were abundant in most of the groups (EG/ELG, LG/LLG, TG/TLG, EPG and CG).

Among the 522 differentially accumulated proteins, several are annotated as chloroplast-located and/or functionally related to the chloroplast, including AGY1 (Liu *et al.*, 2010) and GUN4 (Larkin *et al.*, 2003). Interestingly, both AGY1 (LLG) and GUN4 (TG) trigger retrograde signalling, indicating that HS affects the

information flux from plastids to the nucleus. In agreement, retrograde signalling GUN4 contributes to acquisition of BT in Arabidopsis (Lasorella *et al.*, 2022). In contrast to AGY1 and GUN4, nuclear abundance of the ribonuclease III family protein RNC1 (At4G37510) rapidly decreased upon HS (EG). RNC1 lacks endonuclease activity, but its RNA binding activity is needed for splicing in the chloroplast (Watkins *et al.*, 2007). The potential role of RNC1 in the nucleus remains uncertain. Only a few differentially accumulated proteins in our analysis are annotated as mitochondrial proteins. These included two HSPs (EPG) and at least four RNA binding proteins (two mRNA editing activities in LG and two helicases in RG). Whether these proteins are involved in a mitochondrial retrograde pathway remains unknown. However, it has been reported that specific mitochondrial unfolded and degraded proteins enter the nucleus and regulate transcription of genes related to mitochondrial protein homeostasis under stress conditions (Ma *et al.*, 2023). Moreover, other cytoplasmic proteins accumulated in the nucleus upon HS or during the recovery period, indicating that cytosolic and/or organelle activities are inhibited or activated. These include RNA and/or protein binding activities and other metabolic processes. Further and precise characterization of these proteins should be carried out for a better understanding of the molecular clues and mechanisms controlling protein nuclear translocation in response to HS.

A total of 149 of 522 proteins were assigned to EG, LG, TG, EPG, LPG or RG; the other proteins (373 accessions) were classified in ELG, LLG, TLG or CG. All groups, except for RG, included proteins with changes in nuclear abundance upon HS, either in an early, transient, or persistent manner. In contrast, proteins included in RG showed accumulation profile changes exclusively after HS, in other words, during the recovery period. Notably, we observed that this recovery period (R22 °C) allowed for partial reestablishment of the nuclear proteomic profile. The existence of EPG/ELG, LPG/LLG, RG and CG portrays significant accumulation of proteins at R22 °C. In addition, these groups were more abundant in comparison to EG, LG or TG. This suggests that the “recovery” phase established as 22 °C for 24 hours after HS did not fully restore levels observed during standard conditions (22 °C). This was also evident in PCA (Figure 2A), where the R22 °C and 37 °C 4 h replicates were practically equidistant from the 22 °C replicates. Thus, a longer recovery period, i.e., 22 °C for 48 or 60 h after HS, may achieve substantial recovery. Overall, the amount of proteins exhibiting significant nuclear accumulation in the new recovery phase must decrease dramatically.

Notably, several nucleolar proteins were found in RG, including NRPA3 (Ream *et al.*, 2015), BRX1 (Weis *et al.*, 2015), NSN1 (Jeon *et al.*, 2015), OLI2 (Kojima *et al.*, 2017), APPAN (Choi *et al.*, 2023), and NUCLEAR/NUCLEOLAR GTPASE 2 (ATNUG2) (Im *et al.*, 2011). Similarly, accumulation of other nucleolar proteins, including NUC1 (Kojima *et al.*, 2007, Pontvianne *et al.*, 2007, Pontvianne *et al.*, 2010) and FIB2 (Barneche *et al.*, 2000), was observed during the recovery period as part of CG. All these proteins are involved in rRNA synthesis and/or ribosome biogenesis. As mentioned above, the nucleolus is the most prominent subnuclear structure, and its assembly results in the transcription and processing of rRNA and assembly of ribosome particles (Sáez-Vásquez and Delseny, 2019). In Arabidopsis, pre-rRNAs transcription and processing, ribosome profiles, and functional nucleolar structures are disrupted upon HS and restored gradually after the plants are returned to optimal growing conditions (Darriere *et al.*, 2022). Therefore, increased accumulation of nucleolin, fibrillarin, and other ribosome biogenesis factors (RBFs) during the recovery period might be required to fully restore the assembly and activity of the nucleolus after HS. This phenomenon can be comparable to HSR. As with HSPs or HSFs during HSR, there is strong accumulation of RBFs at R22 °C to promote “nucleolar

recovery". It would be necessary to examine the nuclear abundance of these RBFs in a shorter time point in the recovery phase (22 °C for 4 hours after HS) to evaluate how rapid this response is. Subsequently, this "burst" of RBFs aims to restart ribosome biogenesis. The function of the nucleolus is strongly related to its tripartite structure (Saez-Vasquez and Medina, 2008). Thus, the spike of RBFs during the recovery phase may occur to reestablish the regular nucleolar architecture after HS. Remarkably, the nucleolus is a multiphase liquid condensate (Lafontaine *et al.*, 2021) that requires fibrillarin and the scaffold proteins of the granular component nucleophosmin (NPM1). Both proteins contain IDR (Intrinsically Disordered Region) and GAR (Gly–Arg-rich domain) domains. To our knowledge, an NPM1 orthologue has not been reported in *Arabidopsis*; however, NUC1 also contains IDRs, GAR domains, and acidic stretches (Pontvianne *et al.*, 2007), which might contribute to the formation of condensates governed by liquid–liquid phase separation (LLPS) (Figure S7). Furthermore, GAR1 accumulates in the nucleus during the recovery period and might also drive LLPS (Guillen-Chable *et al.*, 2021). In addition to protein composition, protein concentration, protein posttranslational modifications, and the presence of an RNA seed to promote nucleation are essential to drive LLPS (Guillen-Chable *et al.*, 2021). In this context, increasing the level of IDR/GAR-containing nucleolar proteins during the recovery period might contribute to effectively reassembling the nucleolus after HS.

EXPERIMENTAL PROCEDURES

Plant materials and growth conditions

All lines derived from the *Arabidopsis thaliana* Columbia-0 (Col-0) ecotype. *Arabidopsis* 35S_{pro}:FIB2-YFP plants were described in (Montacie *et al.*, 2017, Azevedo-Favory *et al.*, 2021). The NRPA3_{pro}:NRPA3^m-FLAG-HA (*nrpa3*) lines are detailed below. After sterilization, seeds were sown on 1X Murashige and Skoog (MS) medium (Duchefa Biochemie M0231), including Gamborg B5 vitamins, and supplemented with 1% (w/v) sucrose, 0.05% (w/v) 2-(N-morpholino)ethanesulfonic acid (MES), and 1% (w/v) plant agar (pH 5.7). After two days at 4 °C, the plants were grown for 15 days under a 16 h/8 h photoperiod (light/dark) in Percival growth chambers set at light intensity 180 $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ and hygrometry 55%/60% and temperature 22 °C/19 °C for light/dark, respectively.

For HS, 15-day-old seedlings were transferred to Percival chambers set at 37 °C for 4 h (during the light cycle) and 24 h (16 h light/8 h dark). For recovery experiments, seedlings treated for 24 h at 37 °C were returned to Percival chambers set at 22 °C for 24 h (light/dark, 22 °C/19 °C).

NRPA3_{pro}:NRPA3^m-FLAG-HA plant lines

Arabidopsis plants containing a T-DNA insertion in the fifth exon of NRPA3 (At1g60850) were obtained from The Salk Institute (Salk_seq 088247). Heterozygous NRPA3:*nrpa3* mutant plants were then transformed with a custom-made (GeneCust, BOYNES - FRANCE) NRPA3 gene sequence containing the ~1.2 kb sequence upstream from the ATG start codon and the ~2.1 kb sequence (introns and exons) downstream of the ATG. The NRPA3 sequence contains three missense mutations to replace cysteines C317, C320, and C323 with serines and the sequence FLAGFLAGHAHA at the C-terminus (Figure S1).

Nuclear protein extracts

For nuclear proteomic analysis, nuclear proteins were extracted from 15-day-old nontreated (22 °C), heat-treated (37 °C, 4 h and 37 °C 24 h), and recovered (R22 °C) seedlings (Figure 1A). The nuclear protein extracts were

prepared independently to generate three biological replicates per sample. Ground fine powder from approximately 2 g of seedlings was suspended in 20 ml of cold (4 °C) 1X extraction buffer EB1X (0.5 M hexylene glycol, 0.05 M MOPS, 0.01 M MgCl₂ x 6H₂O) supplemented with 20 mM β-mercaptoethanol and EDTA-free Roche Protease Inhibitor. The suspension was filtered through a 60-μm nylon membrane, and 2 ml of 10% Triton X-100 was added (1% final). After 10 min of incubation, the sample was loaded onto four cushions composed of 35% (4 ml) and 80% (3 ml) Percoll and centrifuged in a JA-20 rotor at 3,000 rpm (706 x g) for 30 min. Nuclei (white interphase between the 35% and 80% Percoll cushions) were recovered (~ 2 ml), diluted 1:2 with EB2X supplemented with 1% Triton X-100 and centrifuged for 10 min at 706 x g. The pellet was collected, suspended in EB1X with 1% Triton X-100 and centrifuged for 10 min at 706 x g; this wash step was performed twice. The final pellet was suspended in EB1X supplemented with 1% Triton X-100 and stored at -20 °C. For polyacrylamide gel electrophoresis (PAGE) analysis, 1 ml of nuclear fraction was centrifuged for 10 min at 706 x g and suspended in 600 μl of 4X Laemmli buffer [62.5 mM Tris-HCl pH 6.7, 2% (v/v) SDS, 10% (v/v) glycerol, 0.05% (v/v) bromophenol blue, 0.1 M DTT]. Then, 8 μl of the sample was used for proteomics analysis.

Quantitative proteomic analysis

For liquid chromatography coupled to tandem mass spectrometry analysis (LC–MS/MS), the nuclear protein samples were diluted in 1X Laemmli buffer supplemented with 10 mM DTT before being loaded onto an in-house 4% (v/v) acrylamide stacking gel. The gels were stained with Coomassie blue, and the stacking bands were manually excised. The proteins were then reduced, alkylated and digested overnight at 37 °C with modified trypsin in a 1:100 enzyme:protein ratio (Promega, Madison, USA). Peptides were extracted for 1 h with 80 μl of 80% acetonitrile and 0.1% formic acid before being dried and resuspended in water acidified with 0.1% formic acid prior to nanoLC–MS/MS analysis (Figure 1B).

The LC–MS/MS analyses were performed using a NanoAcquity LC-system (Waters, Milford, MA, USA) coupled to a Q Exactive Plus Orbitrap (Thermo Fisher Scientific, Waltham, USA) mass spectrometer operated in Data-Dependent Acquisition mode, as previously described (Montacie *et al.*, 2017). Peptides/proteins were identified using the Mascot search engine (version 2.5.1, MatrixScience, London, UK) against an *Arabidopsis thaliana* protein sequence database downloaded from The Arabidopsis Information Resource (TAIR) website (TAIR10 version gene model), to which common contaminants and decoy sequences were added (total of 2 x 27,534 protein entries). Identifications were validated, and label-free extracted ion chromatogram-based quantification was performed using the Proline software suite (Bouyssie *et al.*, 2020). The false discovery rate was optimized to be below 1% of the PSM level using the Mascot adjusted E-value and below 1% at the protein level using the Mascot Mudpit score. Differential statistical analysis was performed on protein abundances (sum of peptide abundances) using the Prostar software suite (version 1.12.11) (Wieczorek *et al.*, 2017). Pairwise Limma t tests were performed. P value calibration was corrected using the adapted Benjamini-Hochsberg method, and FDR was set to <1-2%. (For more details, see Supplementary Material and Method 1; Tables S1 and S2).

Principal component analyses, clustering and GO annotations

Principal component analysis (PCA) was performed with the `prcomp()` function and visualized using the `autoplot()` function, both from the package `ggfortify` v0.4.15 (Tang Y *et al.*, 2016). Hierarchical clustering analysis was carried out with the `hclust()` function from the `stats` package v4.2.0 (<https://stat.ethz.ch/R-manual/R->

devel/library/stats/html/00Index.html) using the Pearson and complete methods as correlation and clustering methods, respectively. The gap statistic method was used to assess the optimal number of clusters using the `fviz_nbclust()` function from the `factoextra` package v1.0.7 (<https://cran.r-project.org/package=factoextra>). Heatmaps for cluster visualization were created with the `pheatmap()` function of the `pheatmap` package v1.0.12 (<https://CRAN.R-project.org/package=pheatmap>) using $\log_2(\text{FC})$ values. Line charts and boxplots were generated with the `ggplot2` package v3.4.0 (<https://ggplot2.tidyverse.org>) using protein abundance values. All these analyses were performed in R v4.2.0 using RStudio v2022.07.2 (<https://www.R-project.org>). Gene Ontology (GO) annotations of protein groups were carried out with the Gene Annotation tool from TAIR (<http://www.arabidopsis.org>).

Western blotting

Total proteins were extracted from 100 mg of 15-day-old nontreated (22 °C), heat-treated (37 °C 4 h and 37 °C 24 h) and recovered (R22 °C) seedlings. Then, plant material was ground to a fine powder in a mortar with liquid nitrogen. The ground samples were suspended in 300 μl of 2X Laemmli buffer and incubated at 65 °C for 10 min. After centrifugation at 14,000 $\times g$ for 15 min at 4 °C, the supernatant was retained as the total protein extract. Fifteen microlitres of total protein extract separated by SDS–PAGE, followed by western blot analyses as previously described (Durut *et al.*, 2014). Membranes were hybridized either with a 1:5,000 dilution of α -NUC1 (Pontvianne *et al.*, 2010) to detect NUC1, a 1:5,000 dilution of α -GFP (Rockland) to detect FIB2-YFP in Col-0-transformed *35S_{pro}:FIB2-YFP* (Montacie *et al.*, 2017), a 1:10,000 dilution of α -HA-HRP to detect NRPA3^m-FLAG-HA in *NRPA3_{pro}:NRPA3^m-FLAG-HA* (*nrpa3*) plants, or with a 1:100,000 dilution of α -Tubulin (Agrisera). The western blot bands were analysed with ImageJ.

ACKNOWLEDGEMENTS

The authors thank Jacinthe Azevedo-Favory for critical reading of the manuscript and Mr. Gorou Horiguchi for kindly providing *OLI2_{pro}:GFP-OLI2* (*oli2-1*) Arabidopsis seeds. This work was supported by the CNRS (fellowships to EMD; UMR5096-JULSAE-004) and by grants from the ANR, RiboStress ANR-17-CE12-0026-01 and MetRibo ANR-20-CE12-0024-01 and Program Ecos-Sud (to JSV and FBH) C21B02. This study is set within the framework of the “Laboratoires d’Excellences (LABEX)” TULIP (ANR-10-LABX-41) and of the “Ecole Universitaire de Recherche (EUR)” TULP-GS (ANR-18-EURE -00019). The proteomics experiments were supported by the French Proteomic Infrastructure (ProFI FR2048, ANR-10-INBS-08-03).

AUTHOR CONTRIBUTIONS

MDE, DT, and deBA performed the experiments. FI performed the bioinformatics analysis, RM and CC performed the LC–MS/MS analyses and interpreted the proteomics results, and JSV and FBH supervised the experiments. MDE and SVJ wrote the manuscript. All authors approved the final manuscript.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest in relation to this work

DATA AVAILABILITY

The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE (Perez-Riverol *et al.*, 2022) partner repository with the dataset identifiers PXD045038 and 10.6019/PXD045038.

SUPPORTING INFORMATION

Figure S1. NRPA3mFLAG-HA protein expression in NRPA3pro:NRPA3m-FLAG-HA (nrpa3) lines.

Figure S2. Analysis of the nuclear proteome of Arabidopsis plants under HS.

Figure S3. Gene Ontology (GO) analysis.

Figure S4. Heat map.

Figure S5. Gene Ontology (GO) analysis of the four additional groups.

Figure S7. Intrinsically Disorder Regions (IDR) *in silico* of NRPA3, NUC1, FIB2, NOP56 and GAR1.

Table S1: LC_MSMS raw data

Table S2: Abundance of peptides from 2629 Arabidopsis accessions

Table S3: List of proteins detected by LC-MSMS

Table S4: List of 522 proteins differentially accumulated proteins

Table S5: List of proteins from the EG, LG, TG, LPG and BG.

Table S6: List of proteins from clusters C1, C2, C3, C4, C5, C6, C7 and C8.

Supplementary Material & Method: LC-MS/MS

REFERENCES

- Abo-Ogiala, A., Carsjens, C., Diekmann, H., Fayyaz, P., Herrfurth, C., Feussner, I. and Polle, A. (2014) Temperature-induced lipocalin (TIL) is translocated under salt stress and protects chloroplasts from ion toxicity. *J Plant Physiol*, **171**, 250-259.
- Asai, T., Tena, G., Plotnikova, J., Willmann, M.R., Chiu, W.L., Gomez-Gomez, L., Boller, T., Ausubel, F.M. and Sheen, J. (2002) MAP kinase signalling cascade in Arabidopsis innate immunity. *Nature*, **415**, 977-983.
- Azevedo-Favory, J., Gaspin, C., Ayadi, L., Montacie, C., Marchand, V., Jobet, E., Rompais, M., Carapito, C., Motorin, Y. and Saez-Vasquez, J. (2021) Mapping rRNA 2'-O-methylations and identification of C/D snoRNAs in Arabidopsis thaliana plants. *RNA Biology*, **18**, 1760-1777.
- Baniwal, S.K., Bharti, K., Chan, K.Y., Fauth, M., Ganguli, A., Kotak, S., Mishra, S.K., Nover, L., Port, M., Scharf, K.D., Tripp, J., Weber, C., Zielinski, D. and von Koskull-Doring, P. (2004) Heat stress response in plants: a complex game with chaperones and more than twenty heat stress transcription factors. *J Biosci*, **29**, 471-487.
- Barakat, A., Szick-Miranda, K., Chang, I.F., Guyot, R., Blanc, G., Cooke, R., Delseny, M. and Bailey-Serres, J. (2001) The organization of cytoplasmic ribosomal protein genes in the Arabidopsis genome. *Plant Physiol*, **127**, 398-415.
- Barneche, F., Steinmetz, F. and Echeverría, M. (2000) Fibrillarin Genes Encode Both a Conserved Nucleolar Protein and a Novel Small Nucleolar RNA Involved in Ribosomal RNA Methylation in Arabidopsis thaliana. *Journal of Biological Chemistry*, **275**, 27212-27220.

- Bouyssie, D., Hesse, A.M., Mouton-Barbosa, E., Rompais, M., Macron, C., Carapito, C., Gonzalez de Peredo, A., Coute, Y., Dupieris, V., Burel, A., Menetrey, J.P., Kalaitzakis, A., Poisat, J., Romdhani, A., Burret-Schiltz, O., Cianferani, S., Garin, J. and Bruley, C. (2020) Proline: an efficient and user-friendly software suite for large-scale proteomics. *Bioinformatics*, **36**, 3148-3155.
- Boyko, A., Blevins, T., Yao, Y., Golubov, A., Bilichak, A., Ilnytsky, Y., Hollunder, J., Meins, F., Jr. and Kovalchuk, I. (2010) Transgenerational adaptation of Arabidopsis to stress requires DNA methylation and the function of Dicer-like proteins. *PLoS One*, **5**, e9514.
- Chi, W.T., Fung, R.W., Liu, H.C., Hsu, C.C. and Charng, Y.Y. (2009) Temperature-induced lipocalin is required for basal and acquired thermotolerance in Arabidopsis. *Plant Cell Environ*, **32**, 917-927.
- Choi, I., Jeon, Y. and Pai, H.S. (2023) Brix protein APPAN plays a role in ribosomal RNA processing in Arabidopsis. *Plant Sci*, **333**, 111721.
- Choi, I., Jeon, Y., Yoo, Y., Cho, H.S. and Pai, H.S. (2020) The in vivo functions of ARPF2 and ARRS1 in ribosomal RNA processing and ribosome biogenesis in Arabidopsis. *J Exp Bot*, **71**, 2596-2611.
- Dard, A., Weiss, A., Bariat, L., Auverlot, J., Fontaine, V., Picault, N., Pontvianne, F., Riondet, C. and Reichheld, J.P. (2023) Glutathione-mediated thermomorphogenesis and heat stress responses in Arabidopsis thaliana. *J Exp Bot*, **74**, 2707-2725.
- Darriere, T., Jobet, E., Zavala, D., Escande, M.L., Durut, N., de Bures, A., Blanco-Herrera, F., Vidal, E.A., Rompais, M., Carapito, C., Gourbiere, S. and Saez-Vasquez, J. (2022) Upon heat stress processing of ribosomal RNA precursors into mature rRNAs is compromised after cleavage at primary P site in Arabidopsis thaliana. *RNA Biol*, **19**, 719-734.
- Durut, N., Abou-Ellail, M., Pontvianne, F., Das, S., Kojima, H., Ukai, S., de Bures, A., Comella, P., Nidelet, S., Rialle, S., Merret, R., Echeverria, M., Bouvet, P., Nakamura, K. and Saez-Vasquez, J. (2014) A duplicated NUCLEOLIN gene with antagonistic activity is required for chromatin organization of silent 45S rDNA in Arabidopsis. *The Plant cell*, **26**, 1330-1344.
- Echevarria-Zomeno, S., Fernandez-Calvino, L., Castro-Sanz, A.B., Lopez, J.A., Vazquez, J. and Castellano, M.M. (2016) Dissecting the proteome dynamics of the early heat stress response leading to plant survival or death in Arabidopsis. *Plant Cell Environ*, **39**, 1264-1278.
- Elvira-Matlot, E., Bardou, F., Ariel, F., Jauvion, V., Bouteiller, N., Le Masson, I., Cao, J., Crespi, M.D. and Vaucheret, H. (2016) The Nuclear Ribonucleoprotein SmD1 Interplays with Splicing, RNA Quality Control, and Posttranscriptional Gene Silencing in Arabidopsis. *Plant Cell*, **28**, 426-438.
- Evrard, A., Kumar, M., Lecourieux, D., Lucks, J., von Koskull-Doring, P. and Hirt, H. (2013) Regulation of the heat stress response in Arabidopsis by MPK6-targeted phosphorylation of the heat stress factor HsfA2. *PeerJ*, **1**, e59.
- Firmansyah (2020) A review of heat stress signaling in plants. *IOP Conf. Series: Earth and Environmental Science*, **484**.
- Folsom, J.J., Begcy, K., Hao, X., Wang, D. and Walia, H. (2014) Rice fertilization-Independent Endosperm1 regulates seed size under heat stress by controlling early endosperm development. *Plant Physiol*, **165**, 238-248.
- Guan, Q., Lu, X., Zeng, H., Zhang, Y. and Zhu, J. (2013) Heat stress induction of miR398 triggers a regulatory loop that is critical for thermotolerance in Arabidopsis. *Plant J*, **74**, 840-851.
- Guillen-Chable, F., Bayona, A., Rodriguez-Zapata, L.C. and Castano, E. (2021) Phase Separation of Intrinsically Disordered Nucleolar Proteins Relate to Localization and Function. *Int J Mol Sci*, **22**.
- Guo, M., Liu, J.H., Ma, X., Luo, D.X., Gong, Z.H. and Lu, M.H. (2016) The Plant Heat Stress Transcription Factors (HSFs): Structure, Regulation, and Function in Response to Abiotic Stresses. *Front Plant Sci*, **7**, 114.

- Guo, Z., Wang, X., Li, Y., Xing, A., Wu, C., Li, D., Wang, C., de Bures, A., Zhang, Y., Guo, S., Saez-Vasquez, J., Shen, Z. and Hu, Z. (2023) Arabidopsis SMO2 modulates ribosome biogenesis by maintaining the RID2 abundance during organ growth. *Plant J.*
- Hayashi, K. and Matsunaga, S. (2019) Heat and chilling stress induce nucleolus morphological changes. *J Plant Res*, **132**, 395-403.
- Huang, C.K., Shen, Y.L., Huang, L.F., Wu, S.J., Yeh, C.H. and Lu, C.A. (2016) The DEAD-Box RNA Helicase AtRH7/PRH75 Participates in Pre-rRNA Processing, Plant Development and Cold Tolerance in Arabidopsis. *Plant Cell Physiol*, **57**, 174-191.
- Im, C.H., Hwang, S.M., Son, Y.S., Heo, J.B., Bang, W.Y., Suwastika, I.N., Shiina, T. and Bahk, J.D. (2011) Nuclear/nucleolar GTPase 2 proteins as a subfamily of YlqF/YawG GTPases function in pre-60S ribosomal subunit maturation of mono- and dicotyledonous plants. *J Biol Chem*, **286**, 8620-8632.
- Ishida, T., Maekawa, S. and Yanagisawa, S. (2016) The Pre-rRNA Processing Complex in Arabidopsis Includes Two WD40-Domain-Containing Proteins Encoded by Glucose-Inducible Genes and Plant-Specific Proteins. *Mol Plant*, **9**, 312-315.
- Jeon, Y., Park, Y.J., Cho, H.K., Jung, H.J., Ahn, T.K., Kang, H. and Pai, H.S. (2015) The nucleolar GTPase nucleostemin-like 1 plays a role in plant growth and senescence by modulating ribosome biogenesis. *J Exp Bot*, **66**, 6297-6310.
- Jiang, C., Bi, Y., Mo, J., Zhang, R., Qu, M., Feng, S. and Essemine, J. (2020) Proteome and transcriptome reveal the involvement of heat shock proteins and antioxidant system in thermotolerance of Clematis florida. *Sci Rep*, **10**, 8883.
- Karayekov, E., Sellaro, R., Legris, M., Yanovsky, M.J. and Casal, J.J. (2013) Heat shock-induced fluctuations in clock and light signaling enhance phytochrome B-mediated Arabidopsis deetiolation. *Plant Cell*, **25**, 2892-2906.
- Keller, M., Consortium, S.-I. and Simm, S. (2018) The coupling of transcriptome and proteome adaptation during development and heat stress response of tomato pollen. *BMC Genomics*, **19**, 447.
- Khan, A., Garbelli, A., Grossi, S., Florentin, A., Batelli, G., Acuna, T., Zolla, G., Kaye, Y., Paul, L.K., Zhu, J.K., Maga, G., Grafi, G. and Barak, S. (2014) The Arabidopsis STRESS RESPONSE SUPPRESSOR DEAD-box RNA helicases are nucleolar- and chromocenter-localized proteins that undergo stress-mediated relocalization and are involved in epigenetic gene silencing. *Plant J*, **79**, 28-43.
- Kim, G.D., Cho, Y.H., Lee, B.H. and Yoo, S.D. (2017) STABILIZED1 Modulates Pre-mRNA Splicing for Thermotolerance. *Plant Physiol*, **173**, 2370-2382.
- Kojima, H., Suzuki, T., Kato, T., Enomoto, K., Sato, S., Kato, T., Tabata, S., Saez-Vasquez, J., Echeverria, M., Nakagawa, T., Ishiguro, S. and Nakamura, K. (2007) Sugar-inducible expression of the nucleolin-1 gene of Arabidopsis thaliana and its role in ribosome synthesis, growth and development. *Plant J*, **49**, 1053-1063.
- Kojima, K., Tamura, J., Chiba, H., Fukada, K., Tsukaya, H. and Horiguchi, G. (2017) Two Nucleolar Proteins, GDP1 and OLI2, Function As Ribosome Biogenesis Factors and Are Preferentially Involved in Promotion of Leaf Cell Proliferation without Strongly Affecting Leaf Adaxial-Abaxial Patterning in Arabidopsis thaliana. *Front Plant Sci*, **8**, 2240.
- Kotak, S., Larkindale, J., Lee, U., von Koskull-Doring, P., Vierling, E. and Scharf, K.D. (2007) Complexity of the heat stress response in plants. *Curr Opin Plant Biol*, **10**, 310-316.
- Kurepa, J., Walker, J.M., Smalle, J., Gosink, M.M., Davis, S.J., Durham, T.L., Sung, D.Y. and Vierstra, R.D. (2003) The small ubiquitin-like modifier (SUMO) protein modification system in Arabidopsis. Accumulation of SUMO1 and -2 conjugates is increased by stress. *J Biol Chem*, **278**, 6862-6872.

- Lafontaine, D.L.J., Riback, J.A., Bascetin, R. and Brangwynne, C.P.** (2021) The nucleolus as a multiphase liquid condensate. *Nature Reviews Molecular Cell Biology*, **22**, 165-182.
- Larkin, R.M., Alonso, J.M., Ecker, J.R. and Chory, J.** (2003) GUN4, a regulator of chlorophyll synthesis and intracellular signaling. *Science*, **299**, 902-906.
- Lasorella, C., Fortunato, S., Dipierro, N., Jeran, N., Tadini, L., Vita, F., Pesaresi, P. and de Pinto, M.C.** (2022) Chloroplast-localized GUN1 contributes to the acquisition of basal thermotolerance in *Arabidopsis thaliana*. *Front Plant Sci*, **13**, 1058831.
- Li, H. and Luan, S.** (2010) AtFKBP53 is a histone chaperone required for repression of ribosomal RNA gene expression in *Arabidopsis*. *Cell Res*, **20**, 357-366.
- Li, S., Liu, J., Liu, Z., Li, X., Wu, F. and He, Y.** (2014) HEAT-INDUCED TAS1 TARGET1 Mediates Thermotolerance via HEAT STRESS TRANSCRIPTION FACTOR A1a-Directed Pathways in *Arabidopsis*. *Plant Cell*, **26**, 1764-1780.
- Liu, D., Gong, Q., Ma, Y., Li, P., Li, J., Yang, S., Yuan, L., Yu, Y., Pan, D., Xu, F. and Wang, N.N.** (2010) cpSecA, a thylakoid protein translocase subunit, is essential for photosynthetic development in *Arabidopsis*. *J Exp Bot*, **61**, 1655-1669.
- Ma, J., Sun, L., Gao, W., Li, Y. and Dong, D.** (2023) RNA binding protein: coordinated expression between the nuclear and mitochondrial genomes in tumors. *J Transl Med*, **21**, 512.
- Montacie, C., Durut, N., Opsomer, A., Palm, D., Comella, P., Picart, C., Carpentier, M.C., Pontvianne, F., Carapito, C., Schleiff, E. and Saez-Vasquez, J.** (2017) Nucleolar Proteome Analysis and Proteasomal Activity Assays Reveal a Link between Nucleolus and 26S Proteasome in *A. thaliana*. *Front Plant Sci*, **8**, 1815.
- Mufarrege, E.F., Gonzalez, D.H. and Curi, G.C.** (2011) Functional interconnections of *Arabidopsis* exon junction complex proteins and genes at multiple steps of gene expression. *J Exp Bot*, **62**, 5025-5036.
- Munoz-Diaz, E. and Saez-Vasquez, J.** (2022) Nuclear dynamics: Formation of bodies and trafficking in plant nuclei. *Front Plant Sci*, **13**, 984163.
- Nover, L., Bharti, K., Doring, P., Mishra, S.K., Ganguli, A. and Scharf, K.D.** (2001) *Arabidopsis* and the heat stress transcription factor world: how many heat stress transcription factors do we need? *Cell Stress Chaperones*, **6**, 177-189.
- Palm, D., Streit, D., Shanmugam, T., Weis, B.L., Ruprecht, M., Simm, S. and Schleiff, E.** (2019) Plant-specific ribosome biogenesis factors in *Arabidopsis thaliana* with essential function in rRNA processing. *Nucleic Acids Res*, **47**, 1880-1895.
- Pan, S., Czarnecka-Verner, E. and Gurley, W.B.** (2000) Role of the TATA binding protein-transcription factor IIB interaction in supporting basal and activated transcription in plant cells. *Plant Cell*, **12**, 125-136.
- Pecinka, A., Dinh, H.Q., Baubec, T., Rosa, M., Lettner, N. and Mittelsten Scheid, O.** (2010) Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in *Arabidopsis*. *Plant Cell*, **22**, 3118-3129.
- Perez-Riverol, Y., Bai, J., Bandla, C., Garcia-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., Walzer, M., Wang, S., Brazma, A. and Vizcaino, J.A.** (2022) The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res*, **50**, D543-D552.
- Pfaff, C., Ehrnsberger, H.F., Flores-Tornero, M., Sorensen, B.B., Schubert, T., Langst, G., Griesenbeck, J., Sprunck, S., Grasser, M. and Grasser, K.D.** (2018) ALY RNA-Binding Proteins Are Required for Nucleocytoplasmic mRNA Transport and Modulate Plant Growth and Development. *Plant Physiol*, **177**, 226-240.

- Pontvianne, F., Abou-Ellail, M., Douet, J., Comella, P., Matia, I., Chandrasekhara, C., Debures, A., Blevins, T., Cooke, R., Medina, F.J., Tourmente, S., Pikaard, C.S. and Saez-Vasquez, J. (2010) Nucleolin is required for DNA methylation state and the expression of rRNA gene variants in *Arabidopsis thaliana*. *PLoS genetics*, **6**, e1001225.
- Pontvianne, F., Matia, I., Douet, J., Tourmente, S., Medina, F.J., Echeverria, M. and Saez-Vasquez, J. (2007) Characterization of AtNUC-L1 reveals a central role of nucleolin in nucleolus organization and silencing of AtNUC-L2 gene in *Arabidopsis*. *Molecular biology of the cell*, **18**, 369-379.
- Popova, O.V., Dinh, H.Q., Aufsatz, W. and Jonak, C. (2013) The RdDM pathway is required for basal heat tolerance in *Arabidopsis*. *Mol Plant*, **6**, 396-410.
- Qu, A.L., Ding, Y.F., Jiang, Q. and Zhu, C. (2013) Molecular mechanisms of the plant heat stress response. *Biochem Biophys Res Commun*, **432**, 203-207.
- Ream, T.S., Haag, J.R., Pontvianne, F., Nicora, C.D., Norbeck, A.D., Pasa-Tolic, L. and Pikaard, C.S. (2015) Subunit compositions of *Arabidopsis* RNA polymerases I and III reveal Pol I- and Pol III-specific forms of the AC40 subunit and alternative forms of the C53 subunit. *Nucleic Acids Res*, **43**, 4163-4178.
- Richter, A.S., Nagele, T., Grimm, B., Kaufmann, K., Schroda, M., Leister, D. and Kleine, T. (2023) Retrograde signaling in plants: A critical review focusing on the GUN pathway and beyond. *Plant Commun*, **4**, 100511.
- Saez-Vasquez, J., Caparros-Ruiz, D., Barneche, F. and Echeverria, M. (2004) A plant snoRNP complex containing snoRNAs, fibrillarin, and nucleolin-like proteins is competent for both rRNA gene binding and pre-rRNA processing in vitro. *Molecular and cellular biology*, **24**, 7284-7297.
- Sáez-Vásquez, J. and Delseny, M. (2019) Ribosome Biogenesis in Plants: from Functional 45S Ribosomal DNA Organization to Ribosome Assembly Factors. *The Plant Cell* **31**, 1945–1967.
- Saez-Vasquez, J. and Medina, F.J. (2008) The plant nucleolus. In *Botanical Research: Incorporating Advances in Plant Pathology* (Kader, J.-C. and Delseny, M. eds). San Diego: Elsevier Academic Press Inc, pp. 1-46.
- Samaha, H., Delorme, V., Pontvianne, F., Cooke, R., Delalande, F., Van Dorselaer, A., Echeverria, M. and Saez-Vasquez, J. (2010) Identification of protein factors and U3 snoRNAs from a *Brassica oleracea* RNP complex involved in the processing of pre-rRNA. *Plant J*, **61**, 383-398.
- Schorova, S., Fajkus, J., Zaveska Drabkova, L., Honys, D. and Schrumpfova, P.P. (2019) The plant Pontin and Reptin homologues, RuvBL1 and RuvBL2a, colocalize with TERT and TRB proteins in vivo, and participate in telomerase biogenesis. *Plant J*, **98**, 195-212.
- Stepinski, D. (2014) Functional ultrastructure of the plant nucleolus. *Protoplasma*, **251**, 1285-1306.
- Stief, A., Altmann, S., Hoffmann, K., Pant, B.D., Scheible, W.R. and Baurle, I. (2014) *Arabidopsis* miR156 Regulates Tolerance to Recurring Environmental Stress through SPL Transcription Factors. *Plant Cell*, **26**, 1792-1807.
- Tang Y, Horikoshi M and W., L. (2016) Ggfortify: unified interface to visualize statistical results of popular r packages. *The R Journal*, **8**, 474-485.
- Thiry, M., Lamaye, F. and Lafontaine, D.L. (2011) The nucleolus: when 2 became 3. *Nucleus*, **2**, 289-293.
- Ul Haq, S., Khan, A., Ali, M., Khattak, A.M., Gai, W.X., Zhang, H.X., Wei, A.M. and Gong, Z.H. (2019) Heat Shock Proteins: Dynamic Biomolecules to Counter Plant Biotic and Abiotic Stresses. *Int J Mol Sci*, **20**.
- Wang, W., Vinocur, B., Shoseyov, O. and Altman, A. (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci*, **9**, 244-252.
- Watkins, K.P., Kroeger, T.S., Cooke, A.M., Williams-Carrier, R.E., Friso, G., Belcher, S.E., van Wijk, K.J. and Barkan, A. (2007) A ribonuclease III domain protein functions in group II intron splicing in maize chloroplasts. *Plant Cell*, **19**, 2606-2623.

- Weis, B.L., Palm, D., Missbach, S., Bohnsack, M.T. and Schleiff, E.** (2015) atBRX1-1 and atBRX1-2 are involved in an alternative rRNA processing pathway in *Arabidopsis thaliana*. *RNA*, **21**, 415-425.
- Wieczorek, S., Combes, F., Lazar, C., Gai Gianetto, Q., Gatto, L., Dorffer, A., Hesse, A.M., Coute, Y., Ferro, M., Bruley, C. and Burger, T.** (2017) DAPAR & ProStaR: software to perform statistical analyses in quantitative discovery proteomics. *Bioinformatics*, **33**, 135-136.
- Wiese, A.J., Steinbachova, L., Timofejeva, L., Cermak, V., Klodova, B., Ganji, R.S., Limones-Mendez, M., Bokvaj, P., Hafidh, S., Potesil, D. and Honys, D.** (2021) Arabidopsis bZIP18 and bZIP52 Accumulate in Nuclei Following Heat Stress where They Regulate the Expression of a Similar Set of Genes. *Int J Mol Sci*, **22**.
- Zhao, J., Lu, Z., Wang, L. and Jin, B.** (2020) Plant Responses to Heat Stress: Physiology, Transcription, Noncoding RNAs, and Epigenetics. *Int J Mol Sci*, **22**.
- Zhao, Q., Chen, W., Bian, J., Xie, H., Li, Y., Xu, C., Ma, J., Guo, S., Chen, J., Cai, X., Wang, X., Wang, Q., She, Y., Chen, S., Zhou, Z. and Dai, S.** (2018) Proteomics and Phosphoproteomics of Heat Stress-Responsive Mechanisms in Spinach. *Front Plant Sci*, **9**, 800.

Supplemental Figures and Tables

Figure S1. NRPA3^mFLAG-HA protein expression in NRPA3_{pro}:NRPA3^m-FLAG-HA (*nrpa3*) lines. **A)** Top, scheme of NRPA3 (At1g60850) gene shows T-DNA insertion in SALKseq_088247 line and NRPA3 genomic sequence cloned into a pCambia vector to obtain NRPA3^m fused to FlagFlag and HAHA sequences. Bottom, NRPA3-FLAG-HA protein sequence. The three cysteine residues mutated into serine and the 2X FlagHA sequence in the C-terminus sequence are shown **B)** Western blot analysis of Col-0 (lane 1) and four *nrpa3* plant lines transformed with the plasmid construct NRPA3^m-FLAG-HA (lanes 2-5).

Figure S2. Heat map of 522 and 149 differentially accumulated proteins in the nuclear proteome Heat map show clustering (C1-C10) of the 522 **A)** or 149 **B)** differentially accumulated proteins in the nucleus at 22°C, 37°C (4 and 24 h) and R22°C.

Figure S3. Gene Ontology (GO) analysis. GO annotation with the Biological Process (red bars), Molecular Function (green bars) and Cellular Component (blue bars) of **A)** Early (EG), Late (EG) and Transient (TG) Groups and **B)** Early Persistent (EPG), Late Persistent (LPG), and Recovery (BG) Groups.

Figure S4. Heat map. Heat maps show the log₂ FC for every protein from **A)** Clusters C1, C4 and C8 from Early-Like Group (ELG) **B)** Clusters C2, C3 and C7 from Late-Like Group **C)** Cluster C6 from Transient-Like Groups (TLG) and **D)** Cluster C5 from Continuous Group (CIG) in indicated temperature/time comparisons. *p* values for 37°C 4 h vs 22°C=0.0005; 37°C 24 h vs 22°C=0.001; R22°C vs 22°C=0.001, 37°C 24 h vs 37°C 4 h=0.0004, R22°C vs 37°C 4 h = 0.0007 and R22°C vs 37°C 24 h= 0.0005. FDR = 0.98-1.14%. Proteins whose nuclear abundance increases or decreases during exposure to 37°C are shown in yellow and blue respectively.

Figure S5. Gene Ontology (GO) analysis of the four additional groups. GO annotation with the Biological Process (red bars), Molecular Function (green bars) and Cellular Component (blue bars) of **A)** Early-Like (ELG) clusters 1, 4 and 8 **B)** Late-Like (LLG) clusters 2, 3 and 7 **C)** Transient-Like (TG) cluster 6 and **D)** Continuous (CG) cluster 5.

Figure S6. Gene Ontology (GO) analysis. Biological Process of the 522 differentially accumulated proteins (Table S4). The most relevant Biological Process are zoomed in a, b and c.

Figure S7. Intrinsically Disorder Regions (IDR) *in silico* of NRPA3, NUC1, FIB2, NOP56 and GAR1. IDRs have been described to drive the LLPS behavior of the nucleolus. FIB2 from the C/D Box complex displays a well-documented disordered region in its N-terminal domain. The major nucleolar factor NUC1 have tendency to disorder in its N-terminal, but also the C-terminal (GAR) domain. Similarly NOP56 found in the C/D box has region disorder in N- and -C terminal domains. GAR1 protein found in the H/ACA complexes has region disorder in C-terminal domains. In contrast, NRPA3 has not disordered region. The intrinsically disordered tendency was calculated by IUPred algorithm (www.iupred2a.elte.hu, accessed on August, 4 2023), a tendency to disorder is predicted when values per residue exceed the 0.5 IUPred scores.

Table S1: LC_MSMS raw data

Table S2: Abundance of peptides from 2629 Arabidopsis accessions with the fold changes and p_values in the six different comparisons: 22°C vs 37°C 4 h; 22°C vs 37°C 24 h; 22°C vs R22°C; 37°C 4 h vs 37°C 24 h; 37°C 4 h vs R22°C and; 37°C 24 h vs R22°C.

Table S3: List of proteins detected by LC-MSMS at 22°C (2,837 accessions), 37°C 4h (2,770 accessions), 37°C 24 h (2,626 accessions) and R22°C (3,064 accessions).

Table S4: List of 522 proteins differentially accumulated proteins at 22°C, 37°C (4 and 24 h) and R22°C.

Table S5: List of proteins from the EG (14 accessions), LG (24 accessions), TG (9 accessions), LPG (24 accessions) and BG (34 accessions). Proteins whose abundance decreases in the corresponding comparisons are marked in grey. The sheet "statistical-criteria" shows the statistical requirement (p-values) applied to perform this classification. "ns" stands for non-significant.

Table S6: List of proteins from clusters C1 (85 accessions), C2 (40 accessions) C3 (42 accessions), C4 (41 accessions), C5 (72 accessions), C6 (49 accessions), C7 (12 accessions) and C8 (32 accessions).

Figure S1

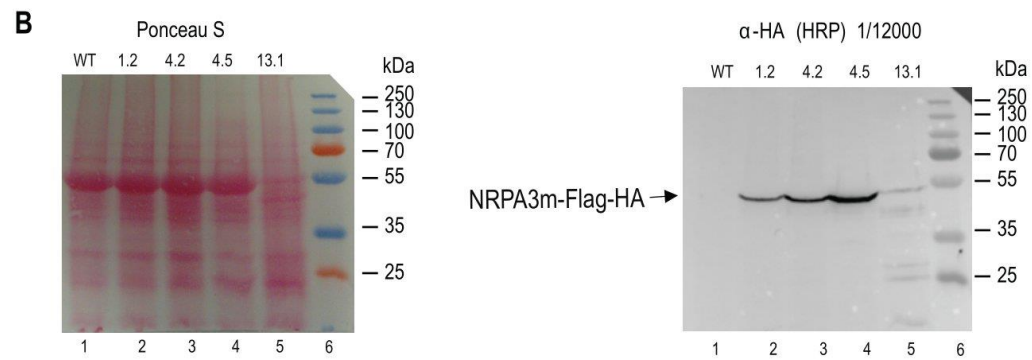
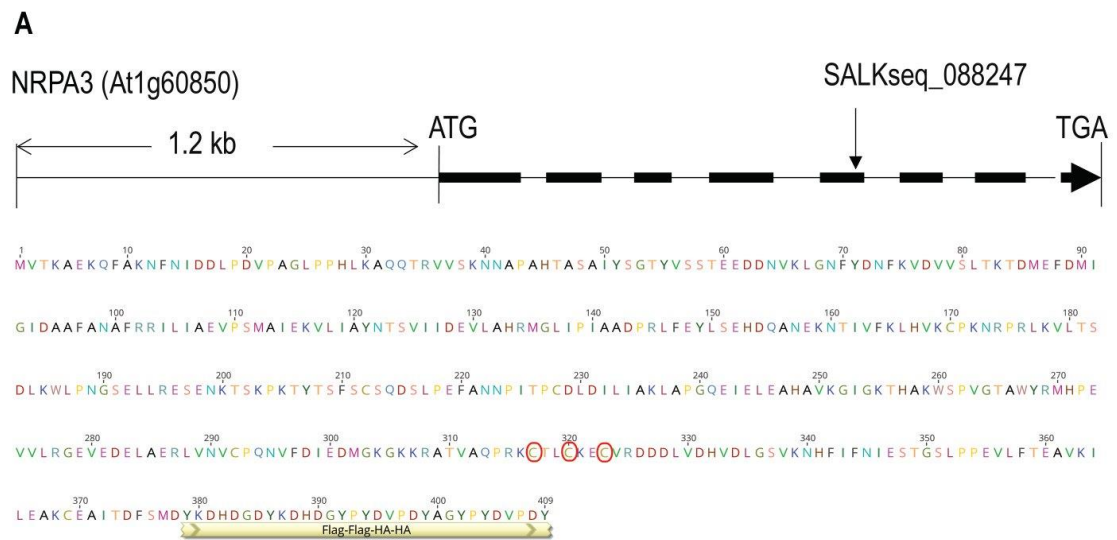


Figure S2

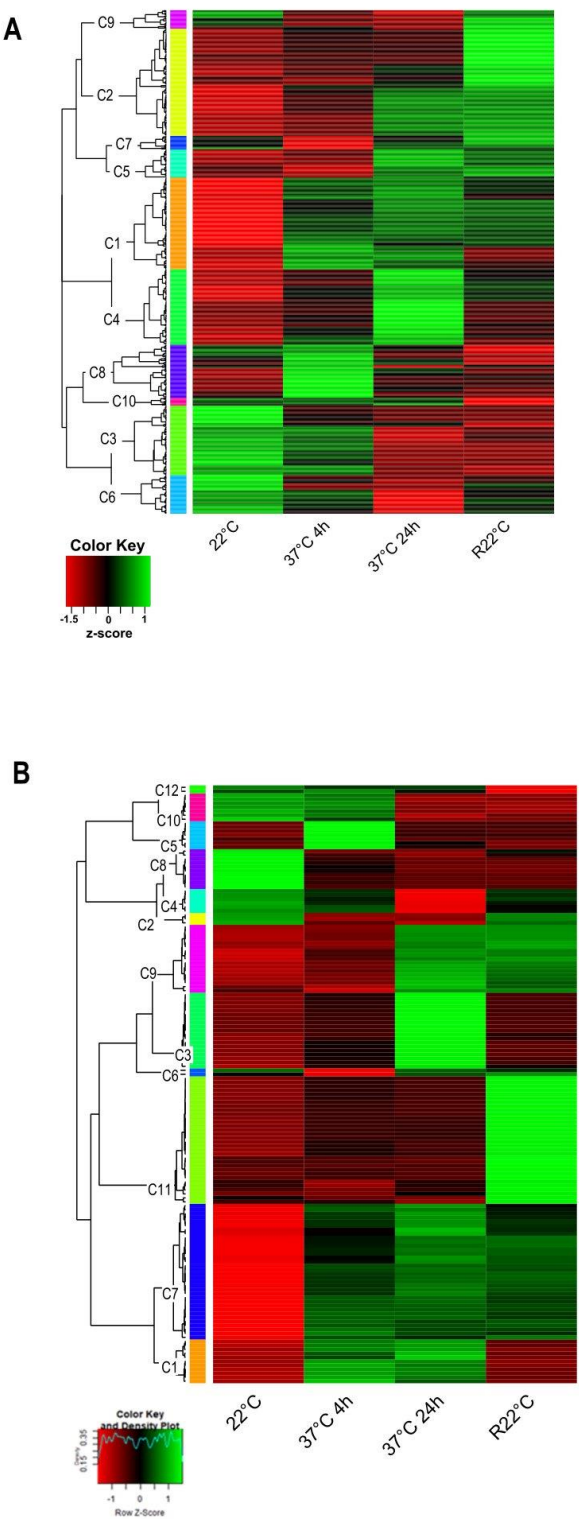
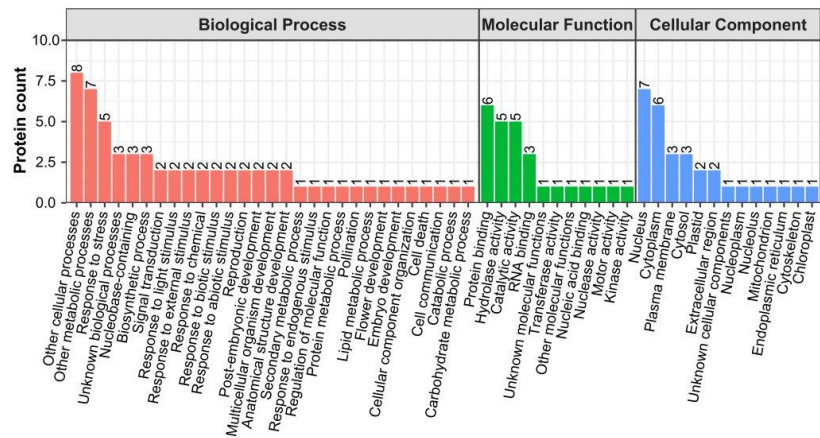
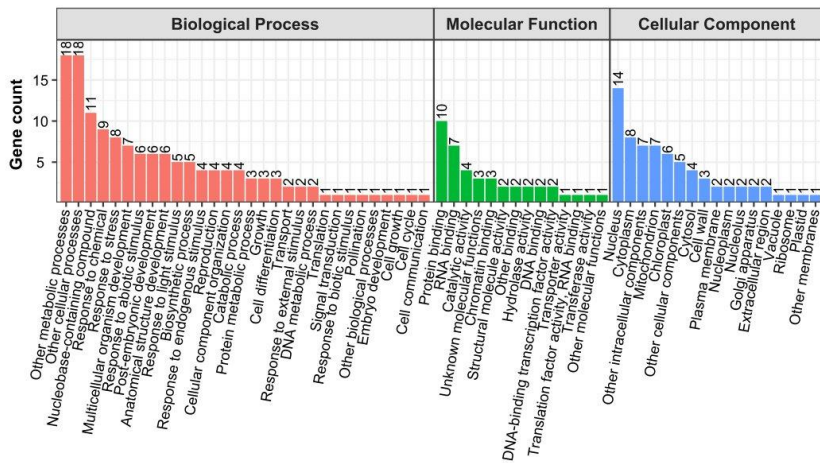


Figure S3A

Early Group (EG)



Late Group (LG)



Transient Group (TG)

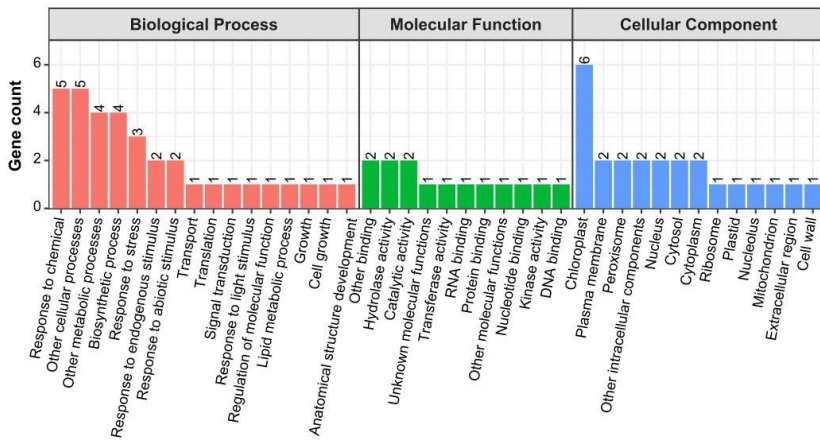


Figure S3B

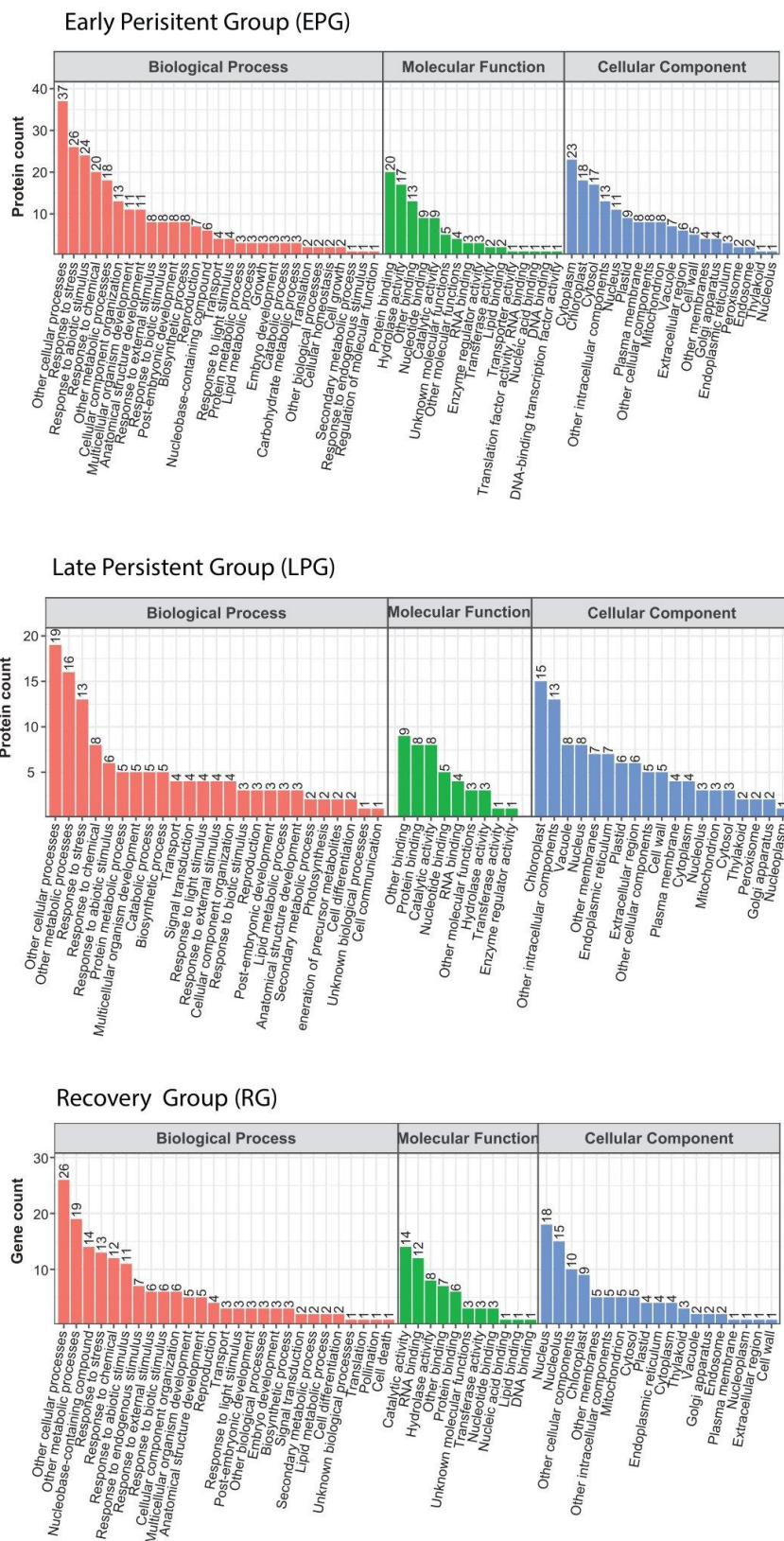


Figure S4A

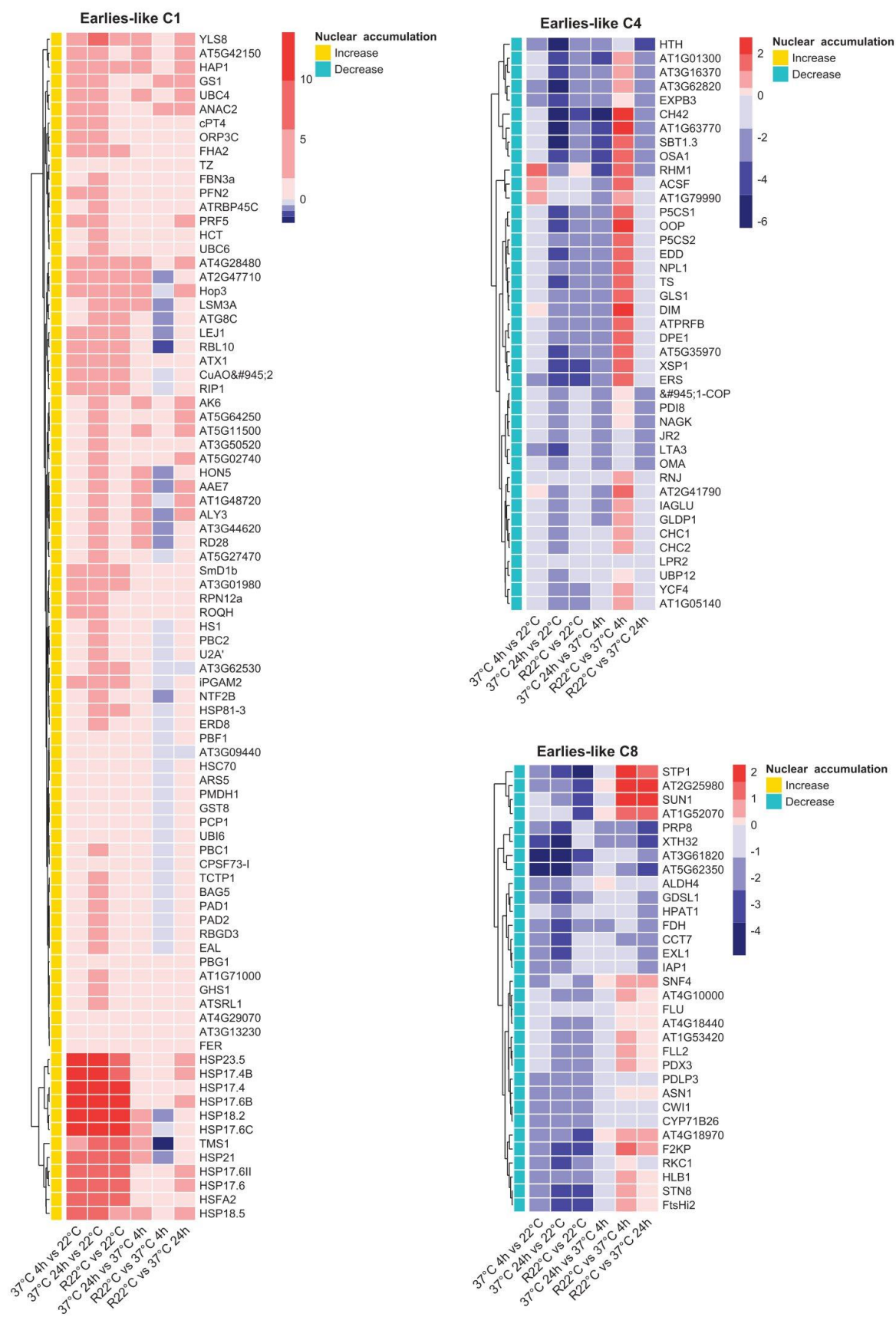


Figure S4B

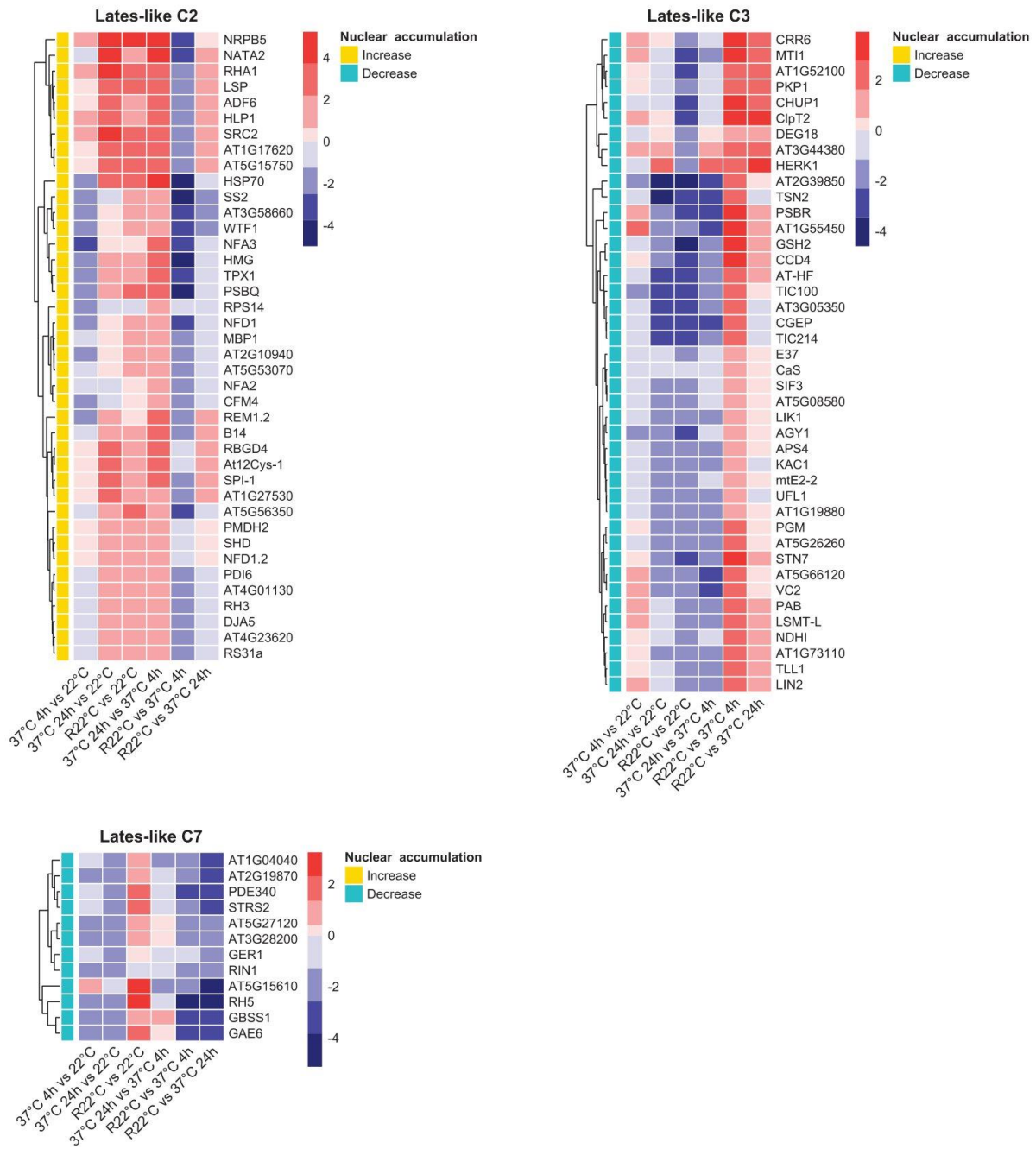


Figure S4C

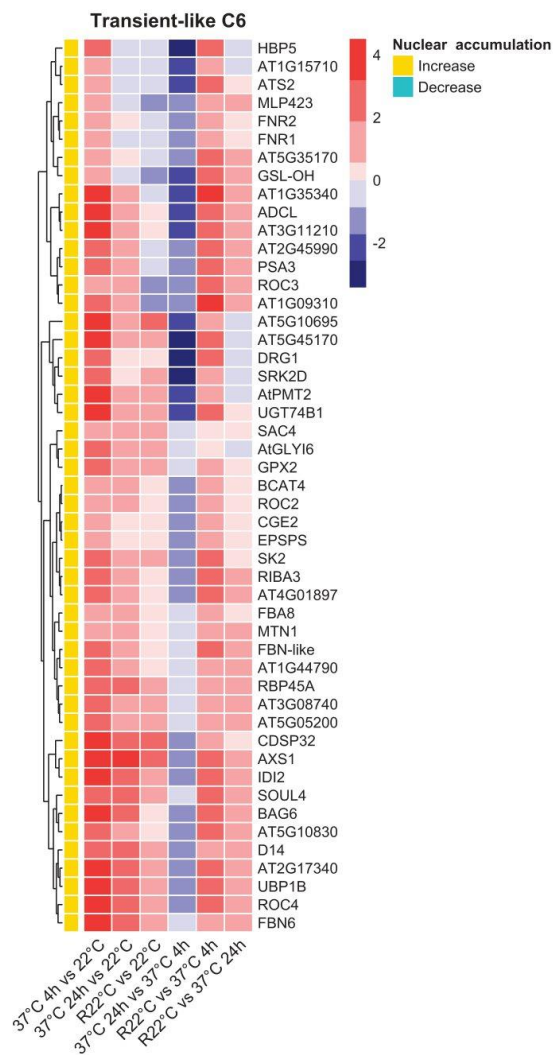
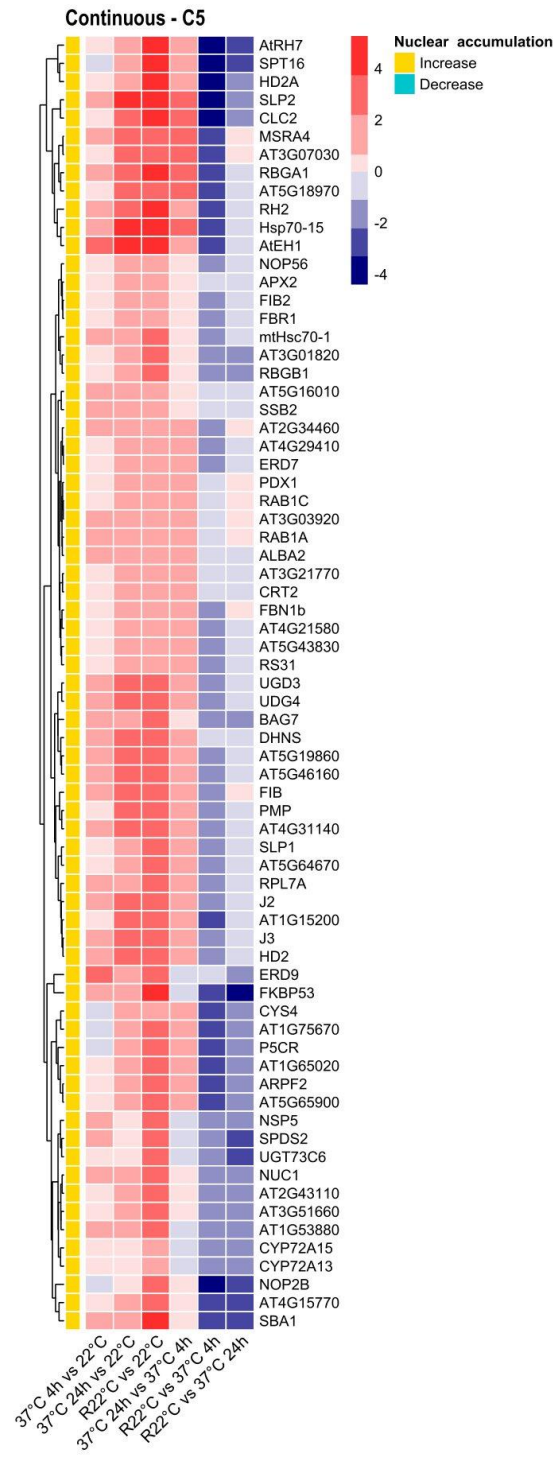


Figure S4D

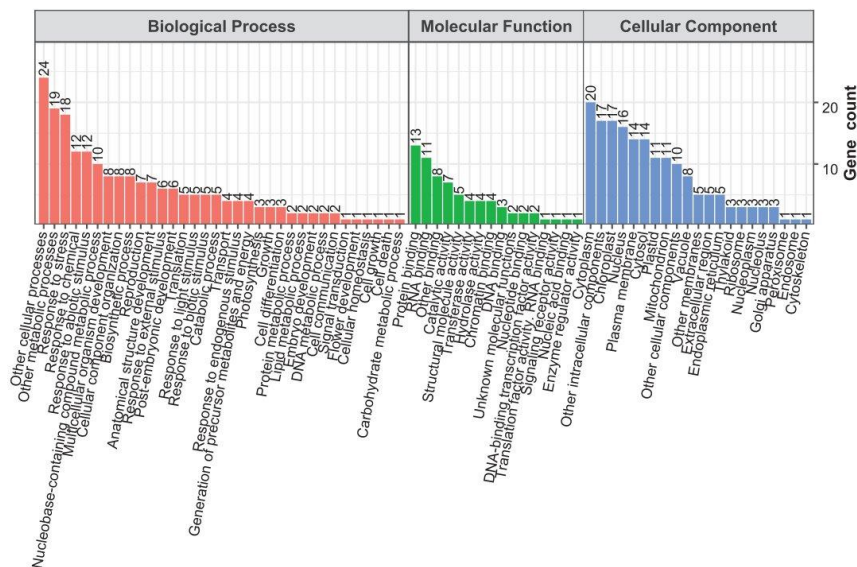


Early-Like Group (ELG)

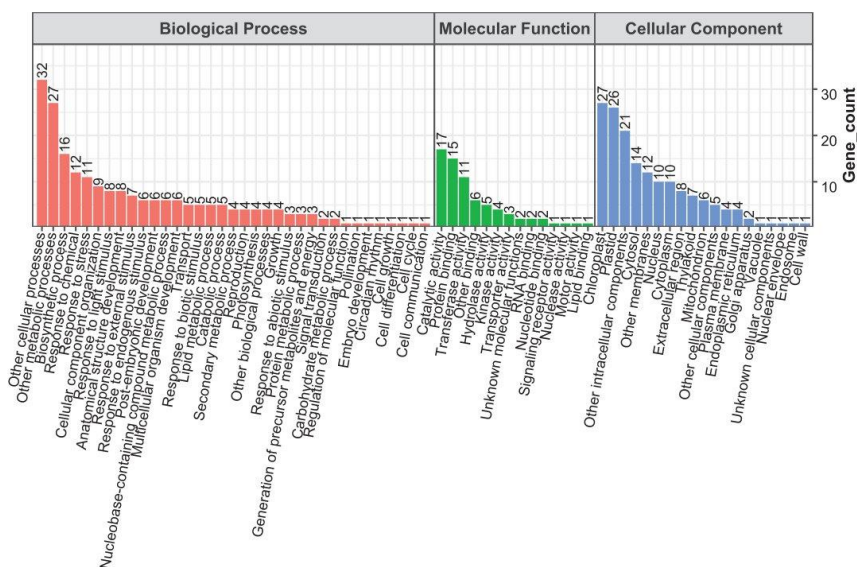


Late-Like Group (LLG)

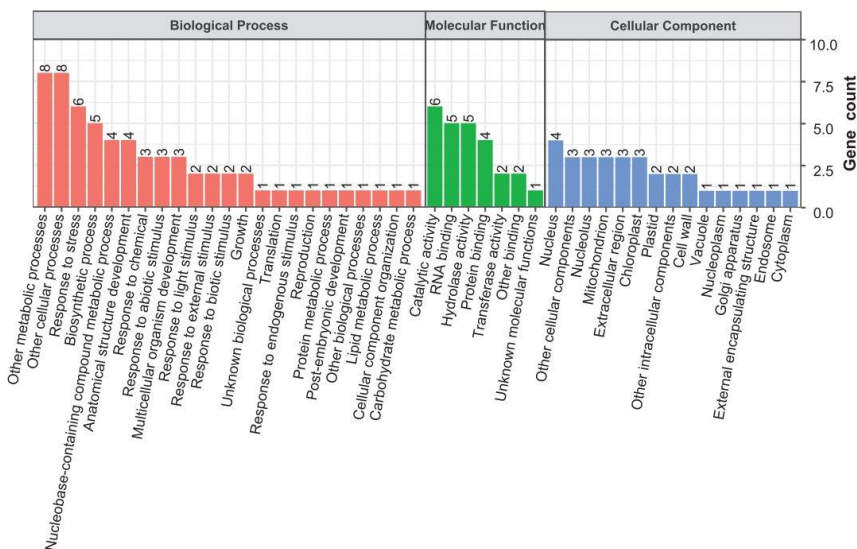
Cluster 2



Cluster 3



Cluster 7



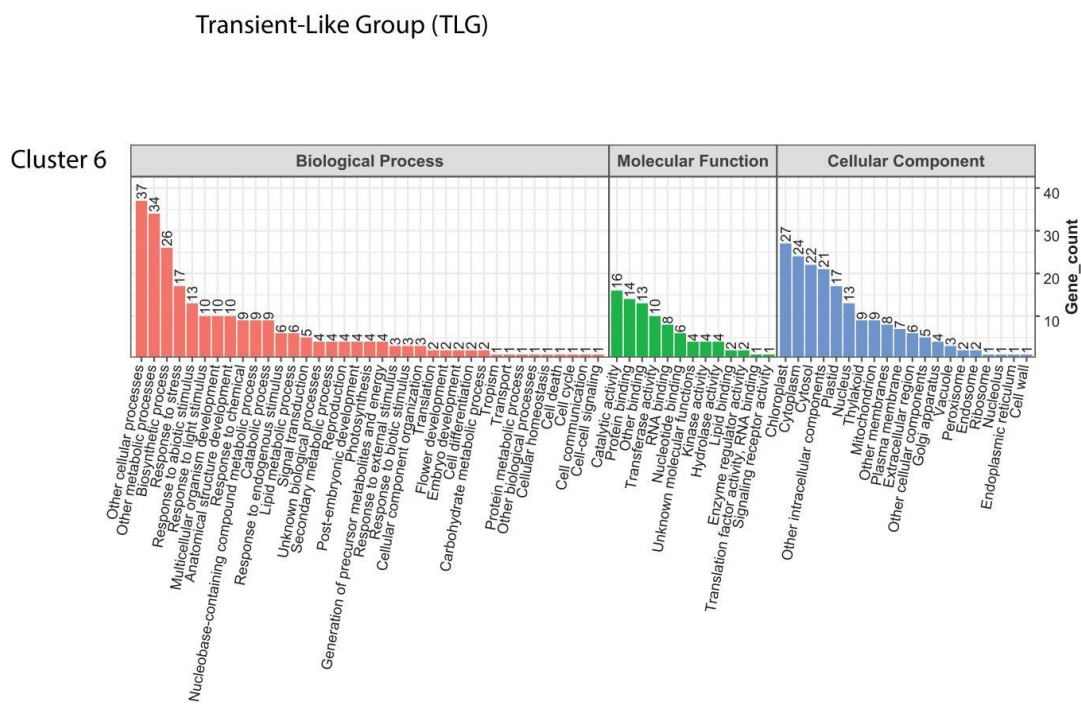
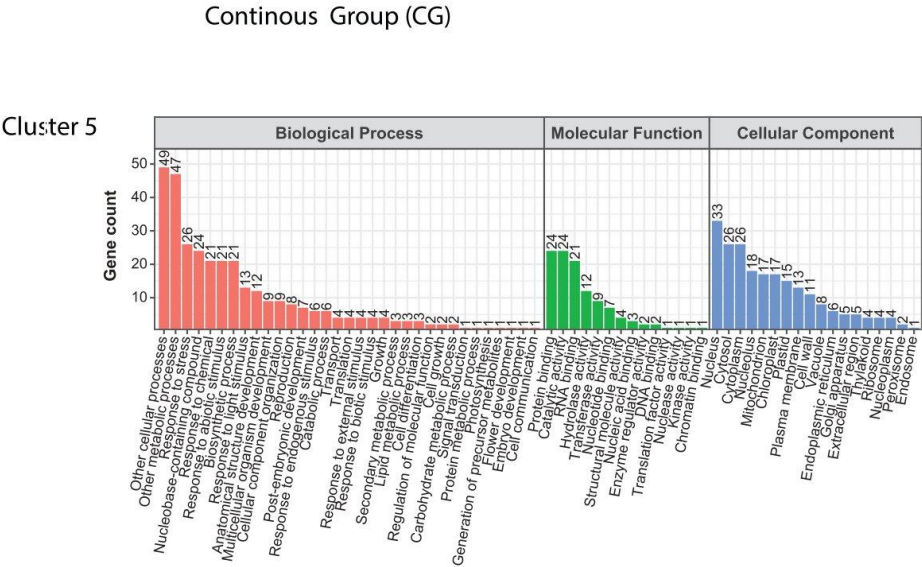


Figure S5D



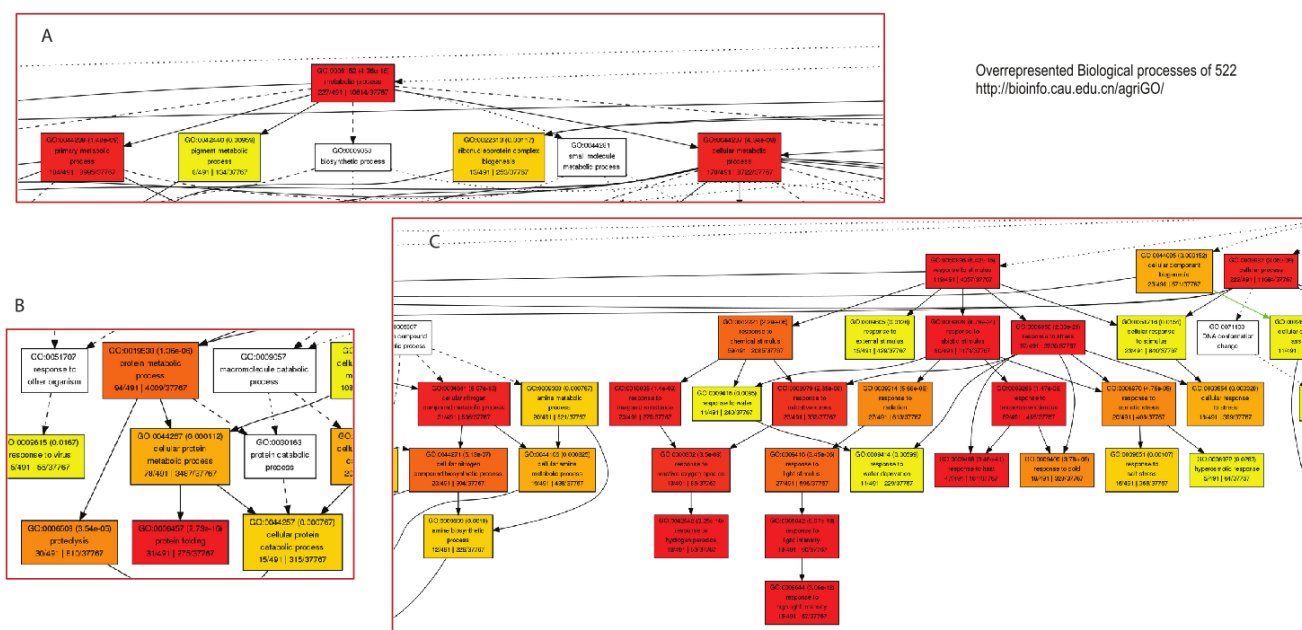


Figure S7

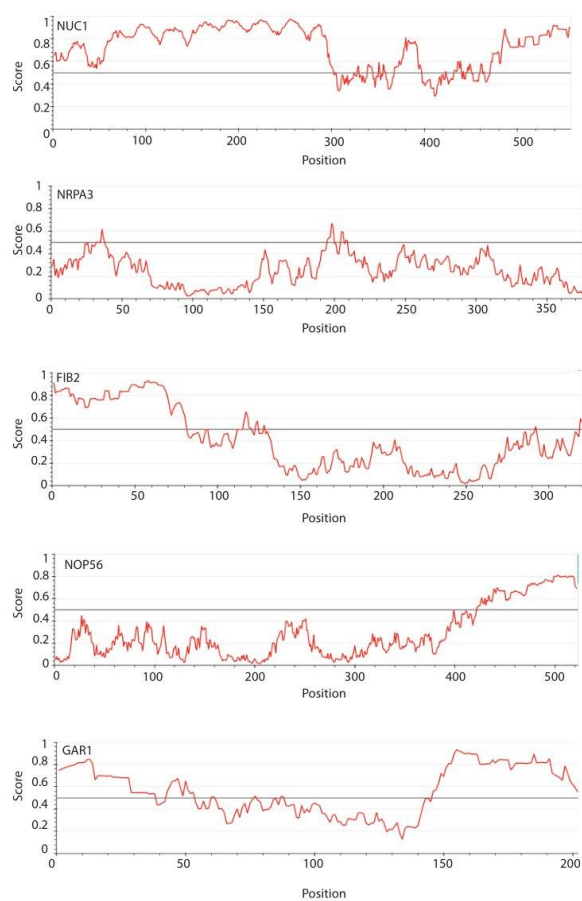


Table S5 (statistical criteria)

	37°C 4 h vs. 22°C	37°C 24 h vs. 22°C	R22°C vs. 22°C	37°C 24 h vs. 37°C 4 h	R22°C vs. 37°C 4 h	R22°C vs. 37°C 24 h
Early Group (EG)	< 0.00501	< 0.00132	ns	ns	< 0.00692	< 0.00479
Late Group (LG)	ns	< 0.00132	ns	< 0.00389	ns	< 0.00479
Transient Group (TG)	< 0.00501	ns	ns	< 0.00389	< 0.00692	ns
Early Persistent Group (EPG)	< 0.00501	< 0.00132	< 0.00955	ns	ns	ns
Late Persistent Group (LPG)	ns	< 0.00132	< 0.00955	< 0.00389	ns	ns
Recovery Group (RG)	ns	ns	< 0.00955	ns	< 0.00692	< 0.00479

Table S5 (EG)

Gene Model Name	Gene Model Description
AT2G35810.1	ureidoglycolate hydrolase;(source:Araport11)
AT5G41970.1	Metal-dependent protein hydrolase;(source:Araport11)
AT5G53310.1	myosin heavy chain-like protein;(source:Araport11)
AT2G35830.2	ureidoglycolate hydrolase;(source:Araport11)
AT4G28440.1	Nucleic acid-binding, OB-fold-like protein;(source:Araport11)
AT1G51510.1	This gene is predicted to encode a protein involved in the exon junction complex. Though there is a predicted RNA binding motif, in the Drosophila ortholog (33% identity), this motif mediates interactions with Mago and is not available for RNA binding. The Arabidopsis Y14 protein appears to be predominantly nucleolar, but there is also some evidence for its presence in the cytoplasm.
AT5G16840.2	Binds to ACD11 and fungal elicitor RxLR207. Regulates ROS mediated defense response.
AT4G39800.1	** Referred to as MIPS2 in Mitsuhashi et al 2008. myo-inositol-1-phosphate synthase isoform 1.Expressed in leaf, root and silique. Immunolocalization experiments with an antibody recognizing MIPS1, MIPS2, and MIPS3 showed endosperm localization.
AT3G45640.1	Encodes a mitogen-activated kinase whose mRNA levels increase in response to touch, cold, salinity stress and chitin oligomers.Also functions in ovule development. Heterozygous MPK3 mutants in a homozygous MPK6 background are female sterile due to defects in integument development. MPK3 can be dephosphorylated by MKP2 in vitro. The mRNA is cell-to-cell mobile.
AT1G50670.1	OTU-like cysteine protease family protein;(source:Araport11)

AT4G26840.1	Encodes a small ubiquitin-like modifier (SUMO) polypeptide that becomes covalently attached to various intracellular protein targets, much like ubiquitination, leading to post-translational modification of those targets.
AT4G37510.1	Ribonuclease III family protein;(source:Araport11)
AT4G02290.1	glycosyl hydrolase 9B13;(source:Araport11)
AT4G34980.1	Serine protease similar to subtilisin.

Table S5 (LG)

Gene Model Name	Gene Model Description
AT5G13070.1	MSF1-like family protein;(source:Araport11)
AT4G22740.1	glycine-rich protein;(source:Araport11)
AT2G04520.1	Nucleic acid-binding, OB-fold-like protein;(source:Araport11)
AT1G21550.1	Calcium-binding EF-hand family protein;(source:Araport11)
AT2G46220.1	DUF2358 family protein (DUF2358);(source:Araport11)
AT1G63780.1	Small nucleolar ribonucleoprotein protein involved in ribosomal RNA processing. Located in nucleolus and cajal bodies.
AT5G56140.1	Member of the K-homologous (KH) family, members have a type of nucleic acid-binding protein containing the KH domain and has been found to affect splicing and transcriptional regulation.
AT5G54580.1	Encodes an RNA recognition motif (RRM) and is involved in C-> U RNA editing in mitochondria.
AT1G01940.1	Heat stress induced cyclophilin involved in splicing of retained introns.Modulates phosphorylation status of PRP18 to modulate the spliceosome activity under heat stress.
AT1G51360.1	Involved in defense against fungal pathogens and located in cytosol.
AT1G56280.2	Encodes a gene whose transcript level in root and leaves increases to progressive drought stress. The increase in transcript level is independent from abscisic acid level. Sequence is not similar to any protein of known function. It appears to be a member of plant-specific gene family. It's phosphorylated by AtCPK11 in a Ca(2+)-dependent manner at Thr105 and Ser107 within the AtDi19 bipartite nuclear localization signal
AT1G20696.2	Encodes a protein belonging to the subgroup of HMGB (high mobility group B) proteins that have a distinctive DNA-binding motif, the HMG-box domain. The motif confers non-sequence specific interaction with linear DNA and structure-specific binding to distorted DNA sites. The HMGB proteins are involved in the assembly of nucleoprotein complexes. Can be phosphorylated by CK2alpha. The mRNA is cell-to-cell mobile.
AT4G20020.1	Encodes a protein involved in RNA editing in mitochondria. Member of MORF family consisting of nine full-length proteins encoded in the nuclear genome. MORF proteins are required for all RNA editing events in plastids and for many, possibly also all, sites in mitochondria. Potential link between the RNA binding PPR protein and the protein contributing the enzymatic activity in RNA editing.

AT1G74560.3	Double nrp1-1 nrp2-1 mutants show arrest of cell cycle progression at G2/M and disordered cellular organization occurred in root tips. Localize in the nucleus and can form homomeric and heteromeric protein complexes with NRP2. Bind histones Histone2A and Histone2B and associate with chromatin in vivo. Plant mutated in both NRP1 and NRP2 genes show hypersensitivity to genotoxic stresses including UV and DSB-inducing agent Bleomycin. NRP genes act synergistically with NAP1 genes in promoting somatic homologous recombination.
AT1G18800.1	Double nrp1-1 nrp2-1 mutants show arrest of cell cycle progression at G2/M and disordered cellular organization occurred in root tips. Localize in the nucleus and can form homomeric and heteromeric protein complexes with NRP1. Bind histones Histone2A and Histone2B and associate with chromatin in vivo. Plant mutated in both NRP1 and NRP2 genes show hypersensitivity to genotoxic stresses including UV and DSB-inducing agent Bleomycin. NRP genes act synergistically with NAP1 genes in promoting somatic homologous recombination.
AT2G40830.1	Encodes an E3 ubiquitin ligase for the GA-receptor GID1 that functions as a negative regulator of GA signaling in seedlings and seeds by inducing ubiquitin-dependent proteolysis of GID1s. Tyr321 phosphorylation of GARU by TAGK2 inactivates GARU.
AT5G55160.2	Encodes a small ubiquitin-like modifier (SUMO) polypeptide that becomes covalently attached to various intracellular protein targets, much like ubiquitination, leading to post-translational modification of those targets. SUMO2 can form SUMO chains through lysine residue 10 during in vitro assays.
AT4G03430.1	Encodes a nuclear protein similar to the human U5 small ribonucleoprotein-associated 102-kD protein and to the yeast pre-mRNA splicing factors Prp1p and Prp6p. STA1 expression is upregulated by cold stress, and the sta1-1 mutant is defective in the splicing of the cold-induced COR15A gene. Luciferase imaging was used to isolate a recessive mutant, sta1-1, with enhanced stability of the normally unstable luciferase transcript. This mutation also causes the stabilization of some endogenous gene transcripts and has a range of developmental and stress response phenotypes.
AT2G41630.1	Encodes a transcription factor, TFIIB1, that plays important roles in pollen tube growth, guidance, and reception as well as endosperm development and is partially functionally different from AtTFIIB2 and AtTFIIB3/AtpBRP2.
AT2G21390.1	Coatomer subunit alpha-2. Part of endomembrane trafficking system. Interacts with SINAT1.
AT4G20850.1	Tripeptidyl Peptidase II. Ser protease that assembles into a large oligomeric complex containing two proteins of 153 and 142 kD that are derived from a single TPP2 gene, with the smaller version missing part of the C-terminal end. Not essential, based on the lack of phenotype of a T-DNA disruption mutant.
AT2G05920.1	Subtilase family protein;(source:Araport11)
AT5G51550.1	EXORDIUM like 3;(source:Araport11)
AT1G02150.1	Tetratricopeptide repeat (TPR)-like superfamily protein;(source:Araport11)

Table S5 (TG)

Gene Model Name	Gene Model Description
AT3G03980.1	NAD(P)-binding Rossmann-fold superfamily protein;(source:Araport11)
AT4G34920.1	PLC-like phosphodiesterases superfamily protein;(source:Araport11)
AT4G39970.1	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein;(source:Araport11)

AT2G26540.1	Encodes a uroporphyrinogen-III synthase involved in tetrapyrrole biosynthesis. The protein localizes to the chloroplast. Member of the plant-specific DUF724 protein family. Arabidopsis has 10 DUF724 proteins. Loss of function mutant has a WT phenotype
AT3G59400.1	GUN, genomes uncoupled, is necessary for coupling the expression of some nuclear genes to the functional state of the chloroplast. Binds to the magnesium chelatase complex and promotes formation of the substrate, a tetrapyrrole signaling molecule. Porphyrin-binding protein that enhances the activity of Mg-chelatase. Although required for chlorophyll accumulation under normal growth conditions, GUN4 is not essential for chlorophyll synthesis.
AT1G66080.1	Encodes a glucose-regulated protein that binds to the promoters of glucose-regulated heat shock responsive genes and promotes chromatin acetylation. HLP1 is required in maintaining histone H3K acetylation and H3K4 methylation marks at the promoters of heat shock protein genes in providing thermotolerance/thermomemory response.
AT4G21210.1	Encodes a PPK regulatory protein that has both protein kinase and protein phosphatase activities towards PPK (pyruvate orthophosphate dikinase).
AT1G75750.1	GA-responsive GAST1 protein homolog regulated by BR and GA antagonistically. Possibly involved in cell elongation based on expression data The mRNA is cell-to-cell mobile.
AT2G27710.1	60S acidic ribosomal protein family;(source:Araport11)

Table S5 (EPG)

Gene Model Name	Gene Model Description
AT1G23130.1	Polyketide cyclase/dehydrase and lipid transport superfamily protein;(source:Araport11)
AT4G34150.1	Calcium-dependent lipid-binding (CaLB domain) family protein;(source:Araport11)
AT2G41250.1	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein;(source:Araport11)
AT3G12050.1	Aha1 domain-containing protein;(source:Araport11)
AT5G62140.1	ATP-dependent Clp protease ATP-binding subunit;(source:Araport11)
AT1G44414.1	zinc-ribbon domain protein;(source:Araport11)
AT4G10250.1	Columbia endomembrane-localized small heat shock protein
AT4G39520.1	Encodes a member of the DRG (developmentally regulated G-protein) family. Has GTPase activity.
AT5G37670.1	HSP20-like chaperones superfamily protein;(source:Araport11)
AT1G59860.1	HSP20-like chaperones superfamily protein;(source:Araport11)
AT1G07400.1	HSP20-like chaperones superfamily protein;(source:Araport11)
AT1G52560.1	HSP20-like chaperones superfamily protein;(source:Araport11)

AT5G48570.1	Encodes one of the 36 carboxylate clamp (CC)-tetra-tricopeptide repeat (TPR) proteins (Prasad 2010, Pubmed ID: 20856808) with potential to interact with Hsp90/Hsp70 as co-chaperones.
AT5G15450.1	Encodes a chloroplast-targeted Hsp101 homologue. Functions as a molecular chaperone involved in plastid differentiation mediating internal thylakoid membrane formation and conferring thermotolerance to chloroplasts during heat stress. APG6 is constitutively expressed in the root tips, the organ boundary region, the reproductive tissues of mature plants where plastids exist as proplastids, and slightly in the stems and leaves. APG6 expression is upregulated in response to heat shock in various organs, but not in response to other abiotic stresses. Apg6 mutants have a pale-green phenotype.
AT2G44650.1	Encodes a chloroplast-localized chaperonin 10 whose mRNA is expressed in leaves and stems but not roots.
AT2G20560.1	DNAJ heat shock family protein;(source:Araport11)
AT2G18510.1	Essential gene (embryo lethal) that is similar to component of spliceosome. Regulates embryonic pattern formation through Pol II-Mediated transcription of WOX2 and PIN7 (DOI:10.1016/j.jsci.2019.09.004). JANUS positively regulates PLT1 expression in the root meristem by recruiting RNA polymerase II (Pol II) to PLT1 and by interacting with PLT1. Nuclear accumulation of JANUS in root meristem depends on IMB4. (DOI:10.1105/tpc.20.00108)
AT3G09350.1	Encodes one of the Arabidopsis orthologs of the human Hsp70-binding protein 1 (HspBP-1) and yeast Fes1p: Fes1A (AT3G09350), Fes1B (AT3G53800), Fes1C (AT5G02150). Fes1A is cytosolic and associates with cytosolic Hsp70. Mutants showed increased heat-sensitive phenotype suggesting the involvement of Fes1A in acquired thermotolerance. Does not have nucleotide exchange factor activity in vitro.
AT4G35860.1	GTP-binding protein ATGB2
AT3G12580.1	Cytoplasmically localized member of the heat shock protein 70 family. Forms complex(heat shock granule) with J3. Mutants are heat sensitive and show increased accumulation of insoluble proteins suggesting a role in thermotolerance.
AT1G74310.1	Encodes ClpB1, which belongs to the Casein lytic proteinase/heat shock protein 100 (Clp/Hsp100) family. Involved in refolding of proteins which form aggregates under heat stress. Also known as AtHsp101. AtHsp101 is a cytosolic heat shock protein required for acclimation to high temperature.
AT1G16030.1	heat shock protein 70B;(source:Araport11)
AT5G52640.1	Encodes a cytosolic heat shock protein AtHSP90.1. AtHSP90.1 interacts with disease resistance signaling components SGT1b and RAR1 and is required for RPS2-mediated resistance. The mRNA is cell-to-cell mobile.
AT2G32120.1	heat-shock protein 70T-2;(source:Araport11)
AT1G67360.1	Encodes a small rubber particle protein homolog. Plays dual roles as positive factors for tissue growth and development and in drought stress responses.
AT5G09590.1	heat shock protein 70 (Hsc70-5); nuclear
AT4G25200.1	AtHSP23.6-mito mRNA, nuclear gene encoding mitochondrial
AT3G16050.1	Encodes a protein with pyridoxal phosphate synthase activity whose transcripts were detected mostly in roots and accumulate during senescence. The protein was found in very low abundance, which prevented a specific localisation.
AT4G17170.1	member of RAB gene family
AT4G34730.1	ribosome-binding factor A family protein;(source:Araport11)
AT3G25230.2	Encodes a a high molecular weight member of the FK506 binding protein (FKBP) family. It has three FKBP12-like domains, tetra-tricopeptide repeats, and a putative calmodulin binding domain. Modulates thermotolerance by interacting with HSP90.1 and affecting the accumulation of HsfA2-regulated sHSPs. Belongs to one of the 36 carboxylate clamp (CC)-tetra-tricopeptide repeat (TPR) proteins (Prasad 2010, Pubmed ID: 20856808) with potential to interact with Hsp90/Hsp70 as co-chaperones.

AT3G01910.1	Encodes a homodimeric Mo-enzyme with molybdopterin as organic component of the molybdenum cofactor. It lacks the heme domain that other eukaryotic Mo-enzymes possess and has no redox-active centers other than the molybdenum. SO protein has been found in all parts of the plant. The plant SO combines its enzymatic sulfite oxidation with a subsequent nonenzymatic step using its reaction product H ₂ O ₂ as intermediate for oxidizing another molecule of sulfite.
AT5G58070.1	Encodes a temperature-induced lipocalin that is involved in thermotolerance and is required for retinal-mediated de novo root organogenesis. Peripherally associated with plasma membrane.
AT1G64760.1	ZERZAUST is an atypical α -1,3 glucanase. The protein is localized to punctate regions of the apoplast, near cellular junctions. Mutants in Ler background display aberrant floral morphology and twisted siliques and stems. Biochemical analysis of mutant cell wall composition indicates cell wall defects. However, in Col background, there is no phenotype due to compensatory effect of ZETH gene expression.
AT1G62330.1	O-fucosyltransferase family protein;(source:Araport11)
AT3G04340.1	Functions in maintaining the cellular redox balance and regulates photorespiratory metabolism.Strong interaction with TIC inner envelope protein translocon which consists of Tic20/Tic56/Tic100/Tic214(Ycf1)(DOI:10.1105/tpc.18.00357).
AT5G11060.1	A member of Class II KN1-like homeodomain transcription factors (together with KNAT3 and KNAT5), with greatest homology to the maize knox1 homeobox protein. Expression regulated by light. Detected in all tissues examined, but most prominent in leaves and young siliques. Transient expression of GFP translational fusion protein suggests bipartite localization in nucleus and cytoplasm. KNAT4 promoter activity showed cell-type specific pattern along longitudinal root axis; GUS expression pattern started at the elongation zone, predominantly in the phloem and pericycle cells, extending to endodermis toward the base of the root.
AT3G60240.4	protein synthesis initiation factor 4G (EIF4G). A mutation in this gene (cum2-1) results in decreased accumulation of CMV coat protein in upper, uninoculated leaves. Likely affects cell-to-cell movement of the virus, also affects TCV multiplication.
AT1G11580.2	Pectin methylesterase involved in root growth, plant-pathogen interactions.
AT1G70320.1	encodes a ubiquitin-protein ligase-like protein containing a HECT domain. There are six other HECT-domain UPLs in Arabidopsis.
AT1G70070.1	Allelic to ISE2(increased size exclusion limit of plasmodesmata 2). Mutants maintain dilated plasmodesmata at the embryonic torpedo stage.
AT3G07320.1	O-Glycosyl hydrolases family 17 protein;(source:Araport11)
AT3G19170.1	Zinc metalloprotease pitrilysin subfamily A. Signal peptide degrading enzyme targeted to mitochondria and chloroplasts. Expressed only in siliques and flowers
AT1G70710.1	endo-1,4-beta-glucanase. Involved in cell elongation.

Table S5 (LPG)

Gene Model Name	Gene Model Description
AT2G35900.1	Mal d 1-associated protein;(source:Araport11)
AT1G10590.3	Nucleic acid-binding, OB-fold-like protein;(source:Araport11)
AT5G07340.2	Calreticulin family protein;(source:Araport11)

AT1G30070.2	SGS domain-containing protein;(source:Araport11)
AT5G37720.1	Encodes a member of a family of proteins involved in export of RNAs from the nucleus and transcriptional coactivation.
AT5G28540.1	Encodes the luminal binding protein BiP, an ER-localized member of the HSP70 family. BiP is composed of an N-terminal ATP binding domain and a C-terminal domain that binds to hydrophobic patches on improperly/incompletely folded proteins in an ATP-dependent manner. Involved in polar nuclei fusion during proliferation of endosperm nuclei.
AT5G42020.3	Luminal binding protein (BiP2) involved in polar nuclei fusion during proliferation of endosperm nuclei.
AT5G03030.2	Chaperone DnaJ-domain superfamily protein;(source:Araport11)
AT1G06460.1	ACD32.1 encodes an alpha-crystallin domain containing protein with homology to small heat shock proteins.
AT5G61790.1	Calnexin is a conserved endoplasmic reticulum (ER) chaperone protein.
AT2G22360.1	Essential for chloroplast iron-sulfur cluster biogenesis.
AT4G21320.1	Encodes heat-stress-associated 32-kD protein. Up-regulated by heat shock. Thermotolerance in a knockout mutant was compromised following a long recovery period (> 24 h) after acclimation heat shock treatment.
AT1G04980.1	Encodes a protein disulfide isomerase-like (PDIL) protein, a member of a multigene family within the thioredoxin (TRX) superfamily. Transcript levels for this gene are up-regulated in response to three different chemical inducers of ER stress (dithiothreitol, beta-mercaptoethanol, and tunicamycin). AtIRE1-2 does not appear to be required for this response, but the atbzp60 mutant has a diminished response.
AT2G32920.1	Encodes a protein disulfide isomerase-like (PDIL) protein, a member of a multigene family within the thioredoxin (TRX) superfamily. Transcript levels for this gene are up-regulated in response to three different chemical inducers of ER stress (dithiothreitol, beta-mercaptoethanol, and tunicamycin). AtIRE1-2 does not appear to be required for this response, but the atbzp60 mutant has a diminished response.
AT1G41830.1	SKU5-similar 6;(source:Araport11)
AT5G42890.1	sterol carrier protein 2;(source:Araport11)
AT2G25110.1	Encodes an endoplasmic reticulum protein SDF2 (stromal-derived factor-2). Forms a complex SDF2-ERdj3B-BiP that is required for the proper accumulation of the surface-exposed leucine-rich repeat receptor kinases EFR. EFR is involved in PAMP (pathogen associated molecular patterns) triggered immunity.
AT5G26280.3	TRAF-like family protein;(source:Araport11)
AT1G37130.1	Identified as a mutant resistant to chlorate. Encodes nitrate reductase structural gene. Involved in nitrate assimilation. Has nitrate reductase activity. Up-regulated by the fungus <i>P. indica</i> . Binds transcription factor At2g35940. The mRNA is cell-to-cell mobile.
AT5G10540.1	Zincin-like metalloproteases family protein;(source:Araport11)
AT5G01220.1	Encodes a UDP-sulfoquinovose:DAG sulfoquinovosyltransferase that is involved in sulfolipid biosynthesis and whose expression is responsive to both phosphate (Pi) and phosphite (Phi) in both roots and shoots.
AT5G55220.1	Contains with HP22 a protein that is related to the bacterial trigger factor chaperone. Plants depleted of either HP22 or HP65b or even both were increasingly delayed in leaf senescence and retained much longer stromal chloroplast constituents than wild-type plants.
AT5G67030.1	Encodes a single copy zeaxanthin epoxidase gene that functions in first step of the biosynthesis of the abiotic stress hormone abscisic acid (ABA). Mutants in this gene are unable to express female sterility in response to beta-aminobutyric acid, as wild type plants do.

AT1G08520.1	Encodes the CHLD subunit of the Mg-chelatase enzyme involved in chlorophyll biosynthesis. Lines carrying recessive mutations of this locus are white and seedling lethal.
-------------	---

Table S5 (RG)

Gene Model Name	Gene Model Description
AT3G18600.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein;(source:Araport11)
AT5G61820.1	stress up-regulated Nod 19 protein;(source:Araport11)
AT1G29320.2	Transducin/WD40 repeat-like superfamily protein;(source:Araport11)
AT2G37870.1	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein;(source:Araport11)
AT4G19610.1	nucleotide/nucleic acid binding protein;(source:Araport11)
AT1G06720.2	P-loop containing nucleoside triphosphate hydrolases superfamily protein;(source:Araport11)
AT5G54910.1	DEA(D/H)-box RNA helicase family protein;(source:Araport11)
AT1G60850.1	DNA-directed RNA polymerase family protein;(source:Araport11)
AT1G50920.1	Putative GTPase involved in HA - and ABA-mediated signaling pathways, particularly during defense respnses to pathogens. Has paralog NOG1-2.
AT1G52930.1	Encodes one of two Arabidopsis orthologs of yeast BRX1, a protein involved in maturation of the large ribosomal subunit. The proteins are mainly localized in nucleolus. Mutant plants are affected in pre-rRNA processing.
AT2G37770.2	Encodes an NADPH-dependent aldo-keto reductase that can act on a wide variety of substrates in vitro including saturated and unsaturated aldehydes, steroids, and sugars. GFP-tagged AKR4C9 localizes to the chloroplast where it may play a role in detoxifying reactive carbonyl compounds that threaten to impair the photosynthetic process. Transcript levels for this gene are up-regulated in response to cold, salt, and drought stress.
AT1G64950.1	member of CYP89A The mRNA is cell-to-cell mobile.
AT2G36800.1	Encodes a DON-Glucosyltransferase. The UGT73C5 glucosylates both brassinolide and castasterone in the 23-O position. The enzyme is presumably involved in the homeostasis of those steroid hormones hence regulating BR activity. Transgenic plants overexpressing UGT73C5 show a typical BR-deficient phenotype.
AT3G55620.1	Translation initiation factor IF6;(source:Araport11)
AT5G48030.1	encodes a mitochondrially targeted DNAJ protein involved in female gametophyte development.
AT2G19540.1	Transducin family protein / WD-40 repeat family protein;(source:Araport11)
AT1G33970.1	IAN9 is a member of a small family of proteins. It's expression is repressed upon pathogen infection and loss of function mutants show increased resistance to bacterial pathogens.

AT4G01560.1	Ribosomal RNA processing Brix domain protein;(source:Araport11)
AT1G52980.1	Encodes a GTPase that belongs to the subfamily of YlqF/YawG GTPases. Functions in Pre-60S ribosomal subunit maturation. The mRNA is cell-to-cell mobile.
AT3G07050.1	Arabidopsis NSN1 encodes a nucleolar GTP- binding protein and is required for maintenance of inflorescence meristem identity, floral organ development and megaspore mother cell specification. Functions downstream of BZR1 family transcription factors to restrict MMC specification to a single cell.
AT5G55920.1	Encodes a homolog of the <i>S. cerevisiae</i> Nop2 that is involved in ribosome biogenesis and plays a role on organ size control by promoting cell proliferation and preventing compensation in normal leaf development.
AT5G61770.2	A single-copy gene encoding a 346 aa protein with a single Brix domain. Similar to yeast ribosome biogenesis proteins Ssf1/2.
ATCG00560.1	PSII L protein
AT3G22310.1	Sequence similarity of DEAD-box RNA helicases. Binds RNA and DNA. Involved in drought, salt and cold stress responses. The mRNA is cell-to-cell mobile.
AT3G22330.1	DEAD-box protein required for efficient group II intron splicing in mitochondria.
AT2G33380.1	Encodes a calcium binding protein whose mRNA is induced upon treatment with NaCl, ABA and in response to desiccation. mRNA expression under drought conditions is apparent particularly in leaves and flowers. Isoform of caleosin with a role as a peroxygenase involved in oxylipin metabolism during biotic and abiotic stress. Involved in the production of 2-hydroxy-octadecatrienoic acid. The peroxygenase has a narrow substrate specificity thus acting as a fatty acid hydroperoxide reductase in vivo.
AT4G09730.1	Encodes RH39, a DEAD-box protein involved in the introduction of the hidden break into the 23S rRNA in the chloroplasts. Recombinant RH39 binds to the 23S rRNA in a segment adjacent to the stem-loop creating the hidden break target loop in a sequence-dependent manner. Has ATP-hydrolyzing activity at a Kcat of 5.3 /min in the presence of rRNA sequence. Mutants have drastically reduced level of level of ribulose 1,5-bisphosphate carboxylase/oxygenase. The mRNA is cell-to-cell mobile.
AT5G09420.1	Encodes one of the 36 carboxylate clamp (CC)-tetra tricopeptide repeat (TPR) proteins (Prasad 2010, Pubmed ID: 20856808) with potential to interact with Hsp90/Hsp70 as co-chaperones.
AT4G30440.1	Encodes a UDP-D-glucuronate 4-epimerase involved in pectin biosynthesis in the cell wall and affects cell wall integrity and immunity to fungi and bacteria.
AT1G02000.1	UDP-D-glucuronate 4-epimerase The mRNA is cell-to-cell mobile.
AT4G00110.1	Encodes a putative membrane-anchored UDP-D-glucuronate 4-epimerase.
AT4G12250.1	UDP-D-glucuronate 4-epimerase
AT3G52380.1	Encodes a chloroplast RNA-binding protein that stabilizes chloroplast RNAs as evidenced by analyses of transcript accumulation in null mutants. Essential for seedling development (albino, strongly retarded growth even on sucrose-containing medium).
AT1G54520.1	FLAP1 is a chloroplast membrane protein of unknown function. When grown in variable light mutants have reduced chloroplast number and size. NPQ is increased relative to wild type under low light.

Table S6 (ELG_C1)

Gene Model Name	Gene Model Description
AT1G71000.1	Chaperone DnaJ-domain superfamily protein;(source:Araport11)
AT5G11500.1	coiled-coil protein;(source:Araport11)
AT5G27470.1	seryl-tRNA synthetase / serine-tRNA ligase;(source:Araport11)
AT1G48720.1	Copia-like polyprotein/retrotransposon;(source:Araport11)
AT5G64250.2	Aldolase-type TIM barrel family protein;(source:Araport11)
AT3G50520.1	Phosphoglycerate mutase family protein;(source:Araport11)
AT5G42150.1	Glutathione S-transferase family protein;(source:Araport11)
AT2G47710.1	Adenine nucleotide alpha hydrolases-like superfamily protein;(source:Araport11)
AT3G01980.3	NAD(P)-binding Rossmann-fold superfamily protein;(source:Araport11)
AT4G29070.1	Phospholipase A2 family protein;(source:Araport11)
AT5G02740.1	Ribosomal protein S24e family protein;(source:Araport11)
AT4G28480.1	DNAJ heat shock family protein;(source:Araport11)
AT2G46030.3	Ubiquitin conjugating enzyme E2
AT3G44620.2	protein-tyrosine phosphatase;(source:Araport11)
AT1G66260.1	RNA-binding (RRM/RBD/RNP motifs) family protein;(source:Araport11)
AT1G01720.1	Belongs to a large family of putative transcriptional activators with NAC domain. Transcript level increases in response to wounding and abscisic acid. ATAF1 attenuates ABA signaling and synthesis. Mutants are hyposensitive to ABA. The mRNA is cell-to-cell mobile.
AT3G08970.1	J domain protein localized in ER lumen. Can compensate for the growth defect in jem1 scj1 mutant yeast. Also shows similarity to HSP40 proteins and is induced by heat stress. At high temperatures, mutant alleles are not transmitted through the pollen due to defects in pollen tube growth.
AT3G13230.1	RNA-binding KH domain-containing protein;(source:Araport11)
AT4G27000.1	RNA-binding (RRM/RBD/RNP motifs) family protein;(source:Araport11)
AT5G37370.1	encodes a putative splicing factor. Over-expression in yeast and Arabidopsis result in increased tolerance to high salt.
AT5G57440.1	A member of haloacid dehalogenase-like hydrolase family, HAD-type phosphosugar phosphatase.
AT1G54050.1	HSP20-like chaperones superfamily protein;(source:Araport11)
AT2G29500.1	HSP20-like chaperones superfamily protein.

AT1G53540.1	Member of the class I small heat-shock protein (sHSP) family, which accounts for the majority of sHSPs in maturing seeds
AT2G19310.1	HSP20-like chaperones superfamily protein;(source:Araport11)
AT5G51440.1	HSP20-like chaperones superfamily protein;(source:Araport11)
AT3G60820.1	Encodes 20S proteasome beta subunit PBF1 (PBF1).
AT3G27160.2	GHS1 encodes plastid ribosomal protein S21The mRNA is cell-to-cell mobile. Required for photosynthesis and C/N balance.
AT4G02840.2	SmD1b is one of two Yeast SmD1 orthologs, the other being SmD1a.SmD1b accumulates to higher levels than SmD1a.It is localized to the nucleolus and nuclear speckles and appears to have a role in RNA splicing and indirectly facilitating PTGS.
AT5G54770.1	Encodes a thiamine biosynthetic gene that has a dual function in thiamine biosynthesis and mitochondrial DNA damage tolerance . It appears to be involved in producing the thiazole portion of thiamine (vitamin B1). A crystal structure of the protein reveals that it forms a 2-ring homo-octamer. The mRNA is cell-to-cell mobile.
AT5G12020.1	17.6 kDa class II heat shock protein;(source:Araport11)
AT3G08590.1	Encodes a 2,3-biphosphoglycerate-independent phosphoglycerate mutase that is involved in pollen development and stomatal movement.
AT3G51260.1	20S proteosomal alpha subunits. Interacts with SnRK, SKP1/ASK1 during proteasomal binding of an SCF ubiquitin ligase.
AT1G77440.1	Encodes beta subunit of 20s proteasome complex which is involved in protein degradation.
AT1G56450.1	20S proteasome beta subunit PBG1 (PBG1) mRNA, complete cds
AT3G16910.1	Encodes a peroxisomal protein with acetyl-CoA synthetase activity that is responsible for the activation of acetate for entry into the glyoxylate cycle.
AT3G62530.1	ARM-repeat superfamily protein containing a mitochondria -targeting peptide; pollen killer antidote of the pollen killer activity of locus PK3.
AT5G60340.1	Encodes a nuclear adenylate kinase that interacts with a putative homolog of Rps14, AtRPS14-1 and affects the elongation of cells in the stem.
AT1G62040.1	Autophagy protein.
AT1G12060.1	A member of Arabidopsis BAG (Bcl-2-associated athanogene) proteins, plant homologs of mammalian regulators of apoptosis. Plant BAG proteins are multi-functional and remarkably similar to their animal counterparts, as they regulate apoptotic-like processes ranging from pathogen attack, to abiotic stress, to plant development.
AT5G53400.1	Encodes BOBBER1 (BOB1), a non-canonical small heat shock protein required for both development and thermotolerance. BOB1 is cytoplasmic at basal temperatures but forms heat shock granules containing canonical small heat shock proteins at high temperatures. The mRNA is cell-to-cell mobile.
AT5G58770.1	AtCPT7 synthesizes medium-chain polyprenols of approximately 55 carbons in length. The enzyme utilizes geranylgeranyl pyrophosphate (GGPP) and isopentenyl pyrophosphate (IPP) as substrates. The enzymatic product accumulates into plastidial membranes (DOI:10.1105/tpc.16.00796).
AT1G61010.1	cleavage and polyadenylation specificity factor 73-I;(source:Araport11)
AT1G31690.1	Copper amine oxidase family protein;(source:Araport11)
AT3G51550.1	Encodes a synergid-expressed, plasma-membrane localized receptor-like kinase that accumulates asymmetrically in the synergid membrane at the filiform apparatus and mediates male-female gametophyte interactions during pollen tube reception. Also involved in powdery mildew infection. Mutants show faster root elongation under dim light, the protein is required for intracellular accumulation of AHA2 under dim -light growth conditions. Positively regulates flowering by modulating the transcript accumulation and mRNA alternative splicing of certain flowering-related genes, including

	FLOWERING LOCUS C (FLC) and its homolog MADS AFFECTING FLOWERING (MAF). However, the RALF1 ligand negatively regulates flowering compared with FER.
AT3G26070.1	PAP/fibrillin (ECM1) localized to chloroplasts, involved in response to H ₂ O ₂ .
AT3G07220.1	Forkhead domain protein that is a subunit of ISWI chromatin remodeling complex. Interacts with histones and regulates the expression of genes involved in stamen filament elongation.
AT1G78380.1	Encodes a glutathione transferase that is a member of Tau GST gene family. Expression is induced by drought stress, oxidative stress, and high doses of auxin and cytokinin. naming convention according to Wagner et al. (2002) The expression of this gene is upregulated by herbicide safeners such as benoxacor and fenclorim.
AT5G02500.1	Encodes a member of heat shock protein 70 family. Hsc70-1 negatively regulates the expression of Hsp101 through HsfA1d, HsfA1e and HsfA2. During non-HS condition, Hsc70-1 attenuates the activity of HsfAs and finally affects the expression of HsfA2 and Hsp101 genes. hsc70-1 mutant showed thermotolerance phenotype due to higher expression of Hsp101 and other HS inducible genes.
AT3G46230.1	Member of the class I small heat-shock protein (sHSP) family, which accounts for the majority of sHSPs in maturing seeds. Induced by heat, cold, salt, drought and high-light.
AT5G12030.1	Encodes a cytosolic small heat shock protein with chaperone activity that is induced by heat and osmotic stress and is also expressed late in seed development.
AT5G59720.1	encodes a low molecular weight heat shock protein that contains the heat shock element in the promoter region. Expression is induced in response to heat shock.
AT4G27670.1	Encodes Hsp21, a chloroplast located small heat shock protein. A structure model of Hsp21, obtained by homology modeling, single-particle electron microscopy, and lysine-specific chemical crosslinking, shows that the Hsp21 subunits are arranged in two hexameric discs, rotated by 25 degree in relation to each other.
AT3G09440.1	Heat shock protein 70 isoform.
AT5G56030.2	A member of heat shock protein 90 (HSP90) gene family. Expressed in all tissues and abundant in root apical meristem, pollen and tapetum. Expression is NOT heat-induced but induced by IAA and NaCl. Interacts with HsfA1d in the cytosol and the nucleus and negatively regulates HsfA1d. Did not bind to AtHsfA4c. The mRNA is cell-to-cell mobile.
AT5G56010.1	A member of heat shock protein 90 (HSP90) gene family. Expressed in all tissues and abundant in root apical meristem, pollen and tapetum. Expression is NOT heat-induced but induced by IAA and NaCl. Overexpression reduced tolerance to heat and conferred higher tolerance to calcium. The mRNA is cell-to-cell mobile.
AT2G26150.1	member of Heat Stress Transcription Factor (Hsf) family. Involved in response to misfolded protein accumulation in the cytosol. Regulated by alternative splicing and non-sense-mediated decay.
AT3G17210.1	Encodes a heat stable protein with antimicrobial and antifungal activity.
AT1G48620.1	This gene is predicted to encode a histone H1/H5 family member. A plant line expressing an RNAi construct targeted against HON5 shows a reduced level of agrobacterium-mediated root transformation.
AT1G66240.1	homolog of anti-oxidant 1; (source: Araport11)
AT4G12400.2	Encodes one of the 36 carboxylate clamp (CC)-tetra tricopeptide repeat (TPR) proteins (Prasad 2010, Pubmed ID: 20856808) with potential to interact with Hsp90/Hsp70 as co-chaperones.

AT5G48930.1	At5g48930 has been shown to encode for the hydroxycinnamoyl-Coenzyme A shikimate/quinate hydroxycinnamoyltransferase (HCT) both synthesizing and catabolizing the hydroxycinnamoyl esters (coumaroyl/caffeoyl shikimate and quinate) involved in the phenylpropanoid pathway. Influence on the accumulation of flavonoids which in turn inhibit auxin transport and reduce plant growth. The mRNA is cell-to-cell mobile.
AT1G15420.1	Encodes a novel plant specific protein that is co-expressed with components of pre-rRNA processing complex. Co-localizes with NuGWD1 and SWA1.
AT4G34120.1	Encodes a single cystathionine beta-Synthase domain-containing protein. Modulates development by regulating the thioredoxin system. The mRNA is cell-to-cell mobile.
AT1G02140.1	Encodes a homologue of the exon junction complex (EJC) component MAGO that participates in intron-mediated enhancement of gene expression.
AT1G27970.2	Encodes an ortholog of yeast NTF2, a nuclear envelope transport protein that functions as the nuclear import receptor for RanGDP, an essential player in nucleocytoplasmic transport. The mRNA is cell-to-cell mobile.
AT5G59420.1	OSBP(oxysterol binding protein)-related protein 3C;(source:Araport11)
AT2G22780.1	encodes an peroxisomal NAD-malate dehydrogenase that is involved in fatty acid beta-oxidation through providing NAD to the process of converting fatty acyl CoA to acetyl CoA.
AT4G29350.1	Encodes profilin2, a low-molecular weight, actin monomer-binding protein that regulates the organization of actin cytoskeleton. Expressed in vegetative organs. The first intron of PRF2 enhances gene expression. The mRNA is cell-to-cell mobile.
AT2G19770.1	Encodes profilin 5, originally named profilin 4 (PRO4/PFN4). Low-molecular weight, actin monomer-binding protein that regulates the organization of actin cytoskeleton. Pollen-specific plant profilin present predominantly in mature pollen and growing pollen tubes.
AT5G66140.1	Encodes alpha5 subunit of 20S proteasome complex involved in protein degradation.
AT5G42790.1	encodes a protein with extensive homology to the largest subunit of the multicatalytic proteinase complex (proteasome). Negatively regulates thiol biosynthesis and arsenic tolerance.
AT1G21720.1	20S proteasome beta subunit PBC1 truncated protein (PBC1)
AT1G64520.1	regulatory particle non-ATPase 12A;(source:Araport11)
AT4G31530.2	Atypical short-chain dehydrogenase-reductase that functions as a qH-relaxation factor.
AT2G37180.1	a member of the plasma membrane intrinsic protein PIP2. functions as aquaporin and is involved in desiccation.
AT3G17611.1	RHOMBOID-like protein 14;(source:Araport11)
AT3G13224.2	Belongs to a member of the RNA-binding glycine-rich (RBG) gene superfamily.
AT3G15000.1	Encodes RIP1 (RNA-editing factor interacting protein 1). Involved in chloroplast and mitochondrial RNA editing. The mRNA is cell-to-cell mobile.
AT1G21190.1	Small nuclear ribonucleoprotein family protein;(source:Araport11)
AT3G16640.1	Encodes a protein homologous to translationally controlled tumor protein (TCTP) from Drosophila. In flies, TCTP functions guanine nucleotide exchange factor in the TOR signaling pathway. TCTP is expressed throughout the plant with highest levels seen in meristematic regions of the shoot and root. Loss of function alleles are not transmitted through the male gametophyte due to defects in pollen tube growth. Hypomorphs, generated through RNAi, are dwarf and have smaller cells. These plants also have defects in lateral and primary root growth as well as root hair growth. The phenotypes are similar to TOR mutants suggesting that TCTP functions in the is pathway in Arabidopsis as well.
AT1G09760.1	U2 small nuclear ribonucleoprotein A;(source:Araport11)

AT2G47110.1	polyubiquitin gene The mRNA is cell-to-cell mobile.
AT5G41340.1	Belongs to Ubiquitin conjugating enzyme family. Gene expression is developmentally regulated.
AT5G08290.1	Encodes Dim1 homolog.

Table S6 (ELG_C4)

Gene Model Name	Gene Model Description
AT1G63770.3	Peptidase M1 family protein;(source:Araport11)
AT1G79990.1	coatomer subunit beta-2;(source:Araport11)
AT1G01300.1	Eukaryotic aspartyl protease family protein;(source:Araport11)
AT2G41790.1	Insulinase (Peptidase family M16) family protein;(source:Araport11)
AT1G05140.1	Peptidase M50 family protein;(source:Araport11)
AT5G35970.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein;(source:Araport11)
AT5G65620.2	Zincin-like metalloproteases family protein;(source:Araport11)
AT4G18480.1	Encodes the CHLI subunit of magnesium chelatase which is required for chlorophyll biosynthesis. All four cysteine residues of the protein form two disulfide bonds (Cys102-Cys193 and Cys354-Cys396) under oxidized conditions but are fully reduced by reduction. It was suggested that the redox state of CHLI is regulated in vivo by the change of the redox environment in the chloroplasts probably via the Trx system.
AT3G16370.1	GDSL-motif esterase/acyltransferase/lipase. Enzyme group with broad substrate specificity that may catalyze acyltransfer or hydrolase reactions with lipid and non-lipid substrates. The mRNA is cell-to-cell mobile.
AT3G52200.2	Encodes a dihydrolipoamide S-acetyltransferase, a subunit of the mitochondrial pyruvate dehydrogenase complex.
ATCG00520.1	Encodes a protein required for photosystem I assembly and stability. In cyanobacteria, loss of function mutation in this gene increases PSII/PSI ratio without any influence on photoautotrophic growth.
AT5G64940.1	ABC1K8 is a member of an atypical protein kinase family that is induced by heavy metals. Loss of function mutations affect the metabolic profile of chloroplast lipids. It appears to function along with ABC1K7 in mediating lipid membrane changes in response to stress. The mRNA is cell-to-cell mobile.
AT1G62020.1	Member of the Coat Protein I (COPI) complex is a seven-subunit coatomer complex consisting of the α, β, β′, γ, δ, ε, and ζ proteins. COPI is required for retrograde transport from the Golgi to the endoplasmic reticulum, Golgi maintenance, and cell plate formation. Required for the acceptance of compatible pollen.

AT3G11130.1	CHC1 heavy chain subunit of clathrin. Involved in vesicle mediated trafficking. Mutants show reduced rates of endocytosis and defects clathrin mediated exocytosis. Mutants also have increased dehydration tolerance which may be related to the overall slower stomatal aperture dynamics. Overall growth is affected.
AT3G08530.1	CHC2 heavy chain subunit of clathrin. Involved in vesicle mediated trafficking. Mutants show reduced rates of endocytosis and defects clathrin mediated exocytosis. Mutants have increased drought tolerance due to defects in stomatal movement.
AT3G56940.1	Encodes a putative ZIP protein with varying mRNA accumulation in leaves, stems and roots. Has a consensus carboxylate-bridged di-iron binding site. The mRNA is cell-to-cell mobile.
AT4G23600.1	Encodes cystine lyase which is expected to be involved in amino acid metabolism, providing the plant with cysteine and the generation of precursors of ethylene biosynthesis. mRNA levels are elevated in response to wounding.
AT3G55610.1	encodes delta 1-pyrroline-5-carboxylate synthetase B. Gene expression is induced by dehydration, high salt and ABA. Knock-out mutations in P5CS2 are embryo-lethal. P5CS2 appears to be present in different cells and/or different subcellular locations from P5CS1 in a tissue-dependent manner. Mutants are defective in pollen development.
AT2G39800.1	encodes a delta 1-pyrroline-5-carboxylate synthase that catalyzes the rate-limiting enzyme in the biosynthesis of proline. Gene is expressed in reproductive organs and tissues under non-stress conditions but in the whole plant under water-limiting condition. Expression is also induced by abscisic acid and salt stress in a light-dependent manner. encodes a delta 1-pyrroline-5-carboxylate synthase that catalyzes the rate-limiting enzyme in the biosynthesis of proline. Gene is expressed in reproductive organs and tissues under non-stress conditions but in the whole plant under water-limiting condition. Expression is also induced by abscisic acid and salt stress in a light-dependent manner. P5CS1 appears to be involved in salt stress responses related to proline accumulation, including protection from reactive oxidative species. P5CS1 appears to be present in different cells and/or different subcellular locations from P5CS2 in a tissue-dependent manner.
AT5G64860.1	Encodes a maltotriose-metabolizing enzyme with chloroplastic α-1,4-glucanotransferase activity. Mutant has altered starch degradation.
AT3G19820.1	Involved in the conversion of the early brassinosteroid precursor 24-methylecholesterol to campesterol. Brassinosteroids affect cellular elongation. Mutants have dwarf phenotype. DWF1 is a Ca ²⁺ -dependent calmodulin-binding protein.
AT5G63420.1	Encodes a member of the metallo-beta-lactamase protein family that plays a vital role in embryo morphogenesis and apical-basal pattern formation by regulating chloroplast development. In bacteria, RNase J plays an important role in rRNA maturation and in the 5'′ stability of mRNA.
AT3G48110.1	glycine-tRNA ligase
AT4G28250.1	putative beta-expansin/allergen protein. Naming convention from the Expansin Working Group (Kende et al, 2004. Plant Mol Bio). Involved in the formation of nematode-induced syncytia in roots of <i>Arabidopsis thaliana</i> .
AT5G04140.2	Encodes a gene whose sequence is similar to ferredoxin dependent glutamate synthase (Fd-GOGAT). Expression in leaves is induced by light and sucrose. Proposed to be involved in photorespiration and nitrogen assimilation. The mRNA is cell-to-cell mobile.
AT5G64050.1	Glutamate-tRNA ligase. Targeted to mitochondria and chloroplast. Its inactivation causes developmental arrest of chloroplasts and mitochondria in <i>Nicotiana benthamiana</i> .
AT4G33010.1	Glycine decarboxylase P-protein 1, catalyzes the degradation of glycine.
AT5G36170.1	Required for normal processing of polycistronic plastidial transcripts
AT1G72970.1	Originally identified as a mutation that causes floral organs to fuse together. About 10-20% of mutants also have defects in ovules. Mutants have reduced fertility most likely as because of fusions that pistil emergence. The protein has similarity to the mandelonitrile lyase family of FAD containing

	oxidoreductases and is predicted to be secreted (SignalP).It is expressed in all tissue layers of roots, inflorescences, stems, leaves, and flowers and is also expressed in siliques. Expression is highest in inflorescence and flower tissue.Transmission of mutant alleles to the progeny shows non mendelian segregation- a percentage of mutant alleles revert back to a previous parental (e.g. grandparental) wild type allele. It has been suggested that an RNA template driven or other extra-DNA genomic mechanism may be responsible for the non-mendelian inheritance of HTH. Reversion events in alleles at other loci have also been observed to occur in plants with an hth mutant background indicating a genome wide effect.
AT4G15550.1	UDP-glucose:indole-3-acetate beta-D-glucosyltransferase
AT1G71040.1	Encodes LPR2. Function together with LPR1 (AT1G23010) and a P5-type ATPase (At5g23630/PDR2) in a common pathway that adjusts root meristem activity to Pi (inorganic phosphate) availability.
AT4G29840.1	Encodes threonine synthase, an enzyme that performs the last step in threonine biosynthesis.
AT3G57560.1	encodes a N-acetylglutamate kinase, involved in arginine biosynthesis
AT1G80410.2	Encodes the catalytic subunit of a N-terminal acetyltransferase.
AT1G35620.1	Encodes a protein disulfide isomerase-like (PDIL) protein, a member of a multigene family within the thioredoxin (TRX) superfamily. Unlike several other PDI family members, transcript levels for this gene are not up-regulated in response to three different chemical inducers of ER stress (dithiothreitol, beta-mercaptoethanol, and tunicamycin). However, the level of transcripts for this gene is slightly elevated in atbzip60 mutants. The mRNA is cell-to-cell mobile.
AT3G62820.1	Functions in stomatal dynamics and stomatal dimension. Interacts with PME31 in stomatal dynamics. Overexpression of PME18 reduced pectin demethylesterification and increased pectin degradation, causing more rapid stomatal dynamics.
AT5G58140.1	Membrane-bound protein serine/threonine kinase that functions as blue light photoreceptor in redundancy with PHO1. Involved in stomatal opening, chloroplast movement and phototropism. Mediates blue light-induced growth enhancements. PHOT1 and PHOT2 mediate blue light-dependent activation of the plasma membrane H ⁺ -ATPase in guard cell protoplasts. PHOT2 possesses two LOV (LOV1 and LOV2, for light-oxygen-voltage-sensing) domains involved in FMN-binding and a C-terminus forming a serine/threonine kinase domain. LOV2 acts as an inhibitor of phototropin kinase in the dark, and light cancels the inhibition through cysteine-FMN adduct formation. LOV1 in contrast acts as an attenuator of photoactivation. Localized to the Golgi apparatus under the induction of blue light. The mRNA is cell-to-cell mobile.
AT1G78570.1	Encodes a UDP-L-Rhamnose synthase involved in the biosynthesis of rhamnose, a major monosaccharide component of pectin. Catalyzes the conversion of UDP-D-Glc to UDP-L-Rha. The dehydrogenase domain of RHM1 was shown to catalyze the conversion of UDP-D-Glc to the reaction intermediate UDP-4-keto-6-deoxy-D-Glc using recombinant protein assay but the activity of the full-length protein was not determined as it could not be expressed in <i>E. coli</i> .
AT5G51750.1	subtilase 1.3;(source:Araport11)
AT5G06600.1	Encodes a ubiquitin-specific protease which together with UBP13 deubiquitinates DA1, DAR1 and DAR2, hence reducing their peptidase activity. Works upstream of DA1, DAR1 and DAR2 to restrict their protease activity and hence fine-tune plant growth and development.
AT4G00230.1	xylem serine peptidase 1;(source:Araport11)

Table S6 (ELG_C8)

Gene Model Name	Gene Model Description
AT4G18440.1	L-Aspartase-like family protein;(source:Araport11)
AT4G10000.1	Thioredoxin family protein;(source:Araport11)
AT1G53420.1	Leucine-rich repeat transmembrane protein kinase;(source:Araport11)
AT2G25980.1	Mannose-binding lectin superfamily protein;(source:Araport11)
AT5G62350.1	Plant invertase/pectin methylesterase inhibitor superfamily protein;(source:Araport11)
AT1G52070.1	Mannose-binding lectin superfamily protein;(source:Araport11)
AT3G26290.2	putative cytochrome P450
AT3G61820.1	Eukaryotic aspartyl protease family protein;(source:Araport11)
AT3G13790.1	Encodes a protein with invertase activity.
AT1G29670.2	GDSL-motif esterase/acyltransferase/lipase. Enzyme group with broad substrate specificity that may catalyze acyltransfer or hydrolase reactions with lipid and non-lipid substrates.The mRNA is cell-to-cell mobile.
AT4G18970.2	GDSL-motif esterase/acyltransferase/lipase. Enzyme group with broad substrate specificity that may catalyze acyltransfer or hydrolase reactions with lipid and non-lipid substrates.
AT2G26250.1	epidermis-specific, encodes KCS10, a putative 3-ketoacyl-CoA synthase. probably involved in the synthesis of long-chain lipids found in the cuticle.
AT1G80070.1	Encodes a factor that influences pre-mRNA splicing and is required for embryonic development. Mutations result in an abnormal suspensor and embryo lethality. The mRNA is cell-to-cell mobile.
AT1G44170.1	Encodes an aldehyde dehydrogenase induced by ABA and dehydration that can oxidize saturated aliphatic aldehydes. It is also able to oxidize beta-unsaturated aldehydes, but not aromatic aldehydes. Activity of ALDH3H1 is NAD ⁺ -dependent.
AT3G11830.1	TCP-1/cpn60 chaperonin family protein;(source:Araport11)
AT4G23250.2	cysteine-rich receptor-like protein kinase 17;(source:Araport11)
AT3G16290.1	Strong interaction with TIC inner envelope protein translocon which consists of Tic20/Tic56/Tic100/Tic214(Ycf1)(DOI:10.1105/tpc.18.00357).
AT3G14110.3	Encodes a novel coiled-coil, TPR domain containing protein that is localized to the chloroplast membrane and is involved in chlorophyll biosynthesis. Mutants accumulate protochlorophyllide, an intermediate in the chlorophyll biosynthesis pathway, in dark and release singlet oxygen in plastids in a controlled and non-invasive manner upon a dark/light shift.
AT1G07110.1	Encodes the bifunctional enzyme fructose-6-phosphate 2-kinase/fructose-2,6-bisphosphatase.
AT3G47340.1	encodes a glutamine-dependent asparagine synthetase, the predicted ASN1 peptide contains a purF-type glutamine-binding domain, and is expressed predominantly in shoot tissues, where light has a negative effect on its mRNA accumulation. Expression is induced within 3 hours of dark treatment, in senescing leaves and treatment with exogenous photosynthesis inhibitor. Induction of gene expression was suppressed in excised leaves supplied with sugar. The authors suggest that the gene's expression pattern is responding to the level of sugar in the cell.

AT1G09020.1	Component of the regulatory subunit of SNF1-related protein kinase. As part of the regulatory complex it binds maltose which promotes kinase activity.
AT5G25265.1	Hyp O-arabinosyltransferase-like protein;(source:Araport11)
AT5G41950.1	Tetratricopeptide repeat (TPR)-like superfamily protein;(source:Araport11)
AT1G18660.4	Membrane localized protein of unknown function. Involved in negative regulation of immune response. Mutants have increased resistance to pathogens.
AT1G35140.1	EXL1 is involved in the C-starvation response. Phenotypic changes of an exl1 loss of function mutant became evident only under corresponding experimental conditions. For example, the mutant showed diminished biomass production in a short-day/low light growth regime, impaired survival during extended night, and impaired survival of anoxia stress.
AT2G33330.1	Encodes a plasmodesmal protein that affects the intercellular movement of molecules through the plasmodesmata. The protein has two DUF26 domains and a single transmembrane domain.
AT5G49970.1	encodes the bifunctional pyridoxine (pyridoxamine) 5'-phosphate oxidase (PPOX)(EC 1.4.3.5) that is involved in the formation of pyridoxal 5'-phosphate (member of the vitamin B6 group). NAD(P)HX epimerase (AT5G49970) interconverts the two epimers of NAD(P)HX.
AT1G01320.1	Encodes REDUCED CHLOROPLAST COVERAGE 1 (REC1) a protein with similarity to the FLOURY locus in maize. Located in the nucleus and cytosol. Contributes to establishing the size of the chloroplast compartment.
AT5G04990.1	Encodes a member of the Sad1/UNC-84 (SUN)-domain proteins: AtSUN1(At5g04990), AtSUN2(AT3G10730). SUN domain proteins are part of the cytoskeletal-nucleoskeletal bridging complexes. AtSUN1 and AtSUN2 are localized to the nuclear envelope and are present as homomers and heteromers in vivo. Encodes an outer nuclear membrane protein that anchors RanGAP1 to the nuclear envelope. It interacts with WPP domain interacting-proteins (WIPs). It is involved in maintaining the elongated nuclear shape of epidermal cells.
AT5G01920.1	Chloroplast thylakoid protein kinase STN8 is specific in phosphorylation of N-terminal threonine residues in D1, D2 and CP43 proteins, and Thr-4 in PsbH protein of photosystem II. Phosphorylation of Thr-4 in the wild type required both light and prior phosphorylation at Thr-2.
AT1G11260.1	Encodes a H ⁺ /hexose cotransporter. The mRNA is cell-to-cell mobile.
AT2G36870.1	Encodes a xyloglucan endotransglycosylase/hydrolase. Protein sequence and phylogenetic analysis indicates that this enzyme resides in Group III-A of the XTH family, with high similarity to Tropaeolum majus (nasturtium) xyloglucanase 1 (TmNXG1). By sequence similarity to XTH 31 (At3g44990) and in vivo analysis, likely to exhibit predominant xyloglucan endo-hydrolase activity (EC 3.2.1.151) with only limited potential to act as a xyloglucan endo-transglycosylase (EC 2.4.1.207).

Table S6 (LLG_C2)

Gene Model Name	Gene Model Description
AT5G15750.1	Alpha-L RNA-binding motif/Ribosomal protein S4 family protein;(source:Araport11)
AT3G58660.1	Ribosomal protein L1p/L10e family;(source:Araport11)
AT5G56350.1	Pyruvate kinase family protein;(source:Araport11)

AT2G10940.1	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein;(source:Araport11)
AT1G17620.1	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family;(source:Araport11)
AT4G01130.1	GDSL-motif esterase/acyltransferase/lipase. Enzyme group with broad substrate specificity that may catalyze acyltransfer or hydrolase reactions with lipid and non-lipid substrates.
AT5G45130.2	small GTP binding protein. The mRNA is cell-to-cell mobile.
AT1G27530.1	ubiquitin-fold modifier-conjugating enzyme;(source:Araport11)
AT5G64400.2	CHCH domain protein;(source:Araport11) involved in mechanotransduction. Loss of both At12cys-1 and At12cys-2 lead to enhanced tolerance to drought and light stress and increased anti-oxidant capacity.
AT5G40490.1	HLP1 is a member of the conserved hnRNP A/B family and contains RNA Recognition Motifs (RRM).It binds mRNA and appears to be involved in targeting alternative polyadenylation (APA). APA targets include genes involved in flowering. Loss of HLP1 function results causes late flowering under long and short day conditions. This phenotype is suppressed by loss of FLC.
AT2G39020.1	Although this locus shares considerable sequence similarity with the adjacent NATA1 gene (At2g39030), they appear to encode genes with different functions. NATA1 is involved in the production of N-delta-acetylornithine, but, overexpression of At2g39020 in tobacco does not lead to the formation of this defense compound. The mRNA is cell-to-cell mobile.
AT3G12260.1	LYR family of Fe/S cluster biogenesis protein;(source:Araport11)
AT3G22320.1	Non-catalytic subunit common to DNA-dependent RNA polymerases I, II, III and IV; homologous to budding yeast RPB5.
AT2G31200.1	Encodes actin depolymerizing factor 6 (ADF6). The mRNA is cell-to-cell mobile.
AT2G46610.1	Barta et al (2010) have proposed a nomenclature for Serine/Arginine-Rich Protein Splicing Factors (SR proteins): Plant Cell. 2010, 22:2926.
ATCG00330.1	30S chloroplast ribosomal protein S14
AT4G39040.1	RNA-binding CRS1 / YhbY (CRM) domain protein;(source:Araport11)
AT4G39960.1	Essential for chloroplast iron?sulfur cluster biogenesis.
AT5G26742.2	DEAD box RNA helicase (RH3);(source:Araport11)
AT4G16660.2	Heat shock protein 70 (Hsp 70) family protein.
AT3G28730.1	encodes a component of the FAcilitates Chromatin Transcription (FACT) complex, SSRP1. Along with STP16 binds to the promoter of FLC.
AT3G51880.2	Encodes a protein belonging to the subgroup of HMGB (high mobility group B) proteins that have a distinctive DNA-binding motif, the HMG-box domain. The motif confers non-sequence specific interaction with linear DNA and structure-specific binding to distorted DNA sites. The HMGB proteins are involved in the assembly of nucleoprotein complexes. Can be phosphorylated by CK2alpha. In interphase cells, HMGB1 is found throughout the nucleus, whereas in mitotic cells it is not chromatin-associated.
AT5G35620.1	Cap-binding protein, binds to the 5' cap structure of nuclear-encoded mRNAs. Mutant is resistant to potyvirus infection.
AT4G30930.1	Encodes a ribosomal RPL21M protein that is localized to the mitochondrion and is involved in karyogamy during female gametophyte development and fertilization. Mutants display defects in both male and female gametophyte development (i.e.collapsed pollen and female gametophytes with unfused central cells).

AT2G19480.1	This gene is predicted to encode a nucleosome assembly protein. Plant lines expressing an RNAi construct directed against this gene show a reduction in agrobacterium-mediated root transformation. The mRNA is cell-to-cell mobile. Plants mutated in three ubiquitously expressed NAP1 genes (NAP1;1~NAP1;3) and organ-specifically expressed NAP1;4 gene show hypersensitivity to genotoxic stresses including UV and DSB-inducing agent Bleomycin. The NAP1 genes act synergistically with NRP genes in promoting somatic homologous recombination.
AT5G56950.1	Encodes a member of a small gene family of proteins with similarity to nucleosome assembly proteins. May function in nucleotide excision repair. Loss of function mutations have no obvious visible phenotypes but do seem to affect transcription of NER related genes. Plants mutated in three ubiquitously expressed NAP1 genes (NAP1;1~NAP1;3) and organ-specifically expressed NAP1;4 gene show hypersensitivity to genotoxic stresses including UV and DSB-inducing agent Bleomycin. The NAP1 genes act synergistically with NRP genes in promoting somatic homologous recombination.
AT1G77510.1	Encodes a protein disulfide isomerase-like (PDIL) protein, a member of a multigene family within the thioredoxin (TRX) superfamily. Transcript levels for this gene are up-regulated in response to three different chemical inducers of ER stress (dithiothreitol, beta-mercaptoethanol, and tunicamycin). AtIRE1-2 does not appear to be required for this response, but the atbzip60 mutant has a diminished response. This protein has been shown to be an attenuator of D1 synthesis, modulating photoinhibition in a light-regulated manner.
AT5G09660.4	encodes a microbody NAD-dependent malate dehydrogenase encodes an peroxisomal NAD-malate dehydrogenase that is involved in fatty acid beta-oxidation through providing NAD to the process of converting fatty acyl CoA to acetyl CoA.
AT4G21280.2	Encodes the PsbQ subunit of the oxygen evolving complex of photosystem II.
AT4G38630.1	Regulatory particle non-ATPase subunit of the 26S proteasome with multiubiquitin-chain-binding capabilities
AT3G61260.1	Lipid raft regulatory protein, crucial for plasma membrane nanodomain assembly to control plasmodesmata aperture and functionality. Negatively regulates the cell-to-cell movement of TuMV via competition with PCaP1 for binding actin filaments.
AT4G23620.2	Ribosomal protein L25/Gln-tRNA synthetase, anti-codon-binding domain-containing protein;(source:Araport11)
AT5G53070.1	Ribosomal protein L9/RNase H1;(source:Araport11)
AT4G14300.1	Belongs to a member of the RNA-binding glycine-rich (RBG) gene superfamily.
AT4G24190.1	encodes an ortholog of GRP94, an ER-resident HSP90-like protein and is involved in regulation of meristem size and organization. Single and double mutant analyses suggest that SHD may be required for the correct folding and/or complex formation of CLV proteins. Lines carrying recessive mutations in this locus exhibits expanded shoot meristems, disorganized root meristems, and defective pollen tube elongation. Transcript is detected in all tissues examined and is not induced by heat. Endoplasmic reticulum supports the protein secretory pathway and has a role in proliferating tissues.
AT1G09070.1	SRC2 specifically binds the peptide PIEPPPHH, and moves from ER to a vacuole fraction where it gets internalized. Involved in Protein Storage Vacuole targeting. The mRNA is cell-to-cell mobile.
AT1G74020.1	Encodes AtSS-2 strictosidine synthase.
AT1G71950.1	SPI-1 is a member of the I9 inhibitor family. It is an inhibitor of SBT4.13 subtilase.
AT1G65980.1	thioredoxin-dependent peroxidase
AT4G01037.1	Ubiquitin carboxyl-terminal hydrolase family protein;(source:Araport11)

Table S6 (LLG_C3)

Gene Model Name	Gene Model Description
AT2G39850.1	Subtilisin-like serine endopeptidase family protein;(source:Araport11)
AT1G45201.3	Target of AtGRP7 regulation.
AT5G66120.2	3-dehydroquinate synthase;(source:Araport11)
AT1G73110.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein;(source:Araport11)
AT3G05350.1	Metallopeptidase M24 family protein;(source:Araport11)
AT1G55450.2	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein;(source:Araport11)
AT5G26260.1	TRAF-like family protein;(source:Araport11)
AT3G44380.1	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family;(source:Araport11)
AT5G08580.1	Calcium-binding EF hand family protein;(source:Araport11)
AT1G52100.1	Mannose-binding lectin superfamily protein;(source:Araport11)
AT5G43780.1	sulfate adenylyltransferase, ATP sulfurylase
AT4G12060.1	Protein involved in chloroplast proteostasis which was highly enriched in the CLPC1-TRAP.
ATCG01090.1	Encodes subunit of the chloroplast NAD(P)H dehydrogenase complex
AT3G22960.1	encodes a chloroplast pyruvate kinase alpha subunit. Important for seed oil biosynthesis. Ubiquitously expressed, with significantly increased expression in maturing seeds. The mRNA is cell-to-cell mobile.
AT3G46220.2	E3 ligase that mediates ufmylation. Part of complex with C53 and the ER-resident adaptor protein DDRGK1. Involved in the pathway that links ribosome-associated quality control with selective autophagy at the ER.
ATCG01130.1	Ycf1 protein;(source:Araport11)
AT2G05830.1	Encodes a 5-methylthioribose-1-phosphate isomerase.
AT4G01800.2	Encodes the ATPase subunit of the chloroplast Sec translocation machinery which plays an essential role in chloroplast biogenesis and the regulation of photosynthesis, the absence of which triggers a retrograde signal, eventually leading to a reprogramming of chloroplast and mitochondrial gene expression.
AT3G63410.1	Encodes a MPBQ/MSBQ methyltransferase located in the chloroplast inner envelope membrane. Mutant plants lack plastoquinone (PQ), suggesting that the APG1 protein is involved in the methylation step of PQ biosynthesis. The gene product is also involved in tocopherol (vitamin E) biosynthesis.
AT4G12980.1	Activated by OXS2 under the treatment of salt.
AT5G23060.1	Encodes a chloroplast-localized protein that modulates cytoplasmic Ca ²⁺ concentration and is crucial for proper stomatal regulation in response to elevations of external Ca ²⁺ . Phosphorylation of this protein is dependent on calcium.
AT2G47390.1	Chloroplast stroma localized glutamyl peptidase.

AT3G25690.5	Actin binding protein required for normal chloroplast positioning. The mRNA is cell-to-cell mobile.
AT2G47910.1	Encodes a chloroplast thylakoid membrane protein. Required for the assembly/accumulation of the NAD(P)H dehydrogenase complex of the photosynthetic electron transport chain.
AT4G24510.1	Encodes a component of the fatty acid elongation machinery required for C28 to C30 fatty acid elongation. It does not require the acyltransferase catalytic site for biological function.
AT5G22640.1	EMB1211 is a MORN (multiple membrane occupation and recognition nexus) motif containing protein involved in embryo development and chloroplast biogenesis. The mRNA is cell-to-cell mobile.
AT5G27380.1	Encodes a protein with similarity to glutathione synthetases, which catalyzes one of the early steps in glutathione biosynthesis. Two transcripts have been detected; the longer transcript is less abundant and the protein is localized to the chloroplast. The smaller transcript, in which the transit peptide is truncated, is localized to the cytosol. Increased glutathione accumulation in response to cesium stress.
AT3G46290.1	Encodes HERCULES1 (HERK1), a receptor kinase regulated by brassinosteroids and required for cell elongation during vegetative growth. Along with ANJ functions as the female (embryo sac) determinant of pollen tube guidance.
AT4G26900.1	encodes a glutamine amidotransferase and cyclase, catalyzes the fifth and sixth steps of the histidine biosynthetic pathway
AT5G10470.2	Kinesin that binds cyclin-dependent kinase CDKA;1 as homodimer or as heterodimer with KCA2. Demarcates the division site in plant cells.
AT1G03475.1	Encodes coproporphyrinogen III oxidase, a key enzyme in the biosynthetic pathway of chlorophyll and heme, a tetrapyrrole pathway. Mutants express cytological and molecular markers associated with the defense responses, usually activated by pathogen infection.
AT1G14030.1	Encodes a lysine methyltransferase whose main soluble physiological substrates are chloroplastic fructose 1,6-bisphosphate aldolases, FBA1, FBA2, and FBA3. Lysines near the C-terminal end of the target proteins are trimethylated.
AT3G14840.2	Encodes LRR-RLK protein that is localized to the plasma membrane and is involved in regulation of plant innate immunity to microbes. LIK1 is phosphorylated by CERK1, a kinase involved in chitin perception. The mRNA is cell-to-cell mobile.
AT3G13930.1	Encodes a subunit of the mitochondrial pyruvate dehydrogenase complex.
AT4G19170.1	Encodes a chloroplast-targeted member of a family of enzymes similar to nine-cis-epoxycarotenoid dioxygenase that acts as a major regulator of carotenoid degradation during dark-induced leaf senescence.. The mRNA is cell-to-cell mobile.
AT5G51820.1	Encodes a plastid isoform of the enzyme phosphoglucosmutase involved in controlling photosynthetic carbon flow. Effective petiole movement against the direction of the gravity requires functional PGM activity that is required for full development of amyloplasts.
AT1G79040.1	Encodes for the 10 kDa PsbR subunit of photosystem II (PSII). This subunit appears to be involved in the stable assembly of PSII, particularly that of the oxygen-evolving complex subunit PsbP. Mutants defective in this gene have reduced amounts of subunits PsbP and PsbQ in PSII. In turn, assembly of PsbR is dependent on the presence of PsbJ.
AT1G19880.1	Encodes a regulator of chromatin condensation 1 (RCC1) family protein; confers plasticity of rosette diameter in response to changes in N availability.
AT4G34090.3	cyclin delta-3;(source:Araport11)
AT1G51805.1	Leucine-rich repeat protein kinase family protein;(source:Araport11)
AT1G68830.1	STN7 protein kinase; required for state transitions, phosphorylation of the major antenna complex (LHCII) between PSII and PSI, and light adaptation. STN7 is involved in state transitions.
AT5G61780.1	Involved in the regulation of AtGA20ox3 expression, as well as seed germination. The mRNA is cell-to-cell mobile.

Table S6 (LLG_C7)

Gene Model Name	Gene Model Description
AT5G27120.1	SAR DNA-binding protein, putative, strong similarity to SAR DNA-binding protein-1 (Pisum sativum) GI:3132696; contains Pfam profile PF01798: Putative snoRNA binding domain; has similarity to MAR binding NOP58 protein The mRNA is cell-to-cell mobile.
AT3G28200.1	Peroxidase superfamily protein;(source:Araport11)
AT5G15610.1	Proteasome component (PCI) domain protein;(source:Araport11)
AT1G04040.1	HAD superfamily, subfamily IIIB acid phosphatase;(source:Araport11)
AT2G19870.1	tRNA/rRNA methyltransferase (SpoU) family protein;(source:Araport11)
AT1G72610.1	germin-like protein (GLP1)
AT1G32900.1	Glucosyltransferase specifically responsible for elongating amylose polymers.
AT5G08610.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein;(source:Araport11)
AT5G22330.1	Encodes a plant orthologue of Pontin that associates with the catalytic subunit of telomerase (AtTERT) in the nucleolus and is involved in telomerase assembly.
AT1G31970.1	DEA(D/H)-box RNA helicase family protein;(source:Araport11)
AT5G08620.1	Similar in sequence to DEAD-box RNA helicases. Binds RNA. Involved in drought, salt and cold stress responses.
AT3G23820.1	Encodes a UDP-D-glucuronate 4-epimerase involved in pectin biosynthesis in the cell wall and affects cell wall integrity and immunity to fungi and bacteria. The mRNA is cell-to-cell mobile.

Table S6 (TLG_C6)

Gene Model Name	Gene Model Description
AT5G35170.1	adenylate kinase family protein;(source:Araport11)
AT1G44790.1	ChaC-like family protein;(source:Araport11)
AT5G10830.1	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein;(source:Araport11)

AT4G01897.1	dihydroorotate dehydrogenase;(source:Araport11)
AT5G05200.1	ABC1- atypical kinase
AT2G45990.1	ribosomal RNA small subunit methyltransferase G;(source:Araport11)
AT2G17340.1	pantothenate kinase;(source:Araport11)
AT5G10695.1	methionyl-tRNA synthetase;(source:Araport11)
AT3G08740.1	elongation factor P (EF-P) family protein;(source:Araport11)
AT5G45170.1	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein;(source:Araport11)
AT1G35340.1	ATP-dependent protease La (LON) domain protein;(source:Araport11)
AT4G30580.1	Encodes a plastidic lysophosphatidic acid acyltransferase (LPAAT). Is critical for chloroplasts phosphatidic acid biosynthesis. The null allele is embryo lethal.
AT1G18060.1	localized to chloroplasts
AT3G11210.1	SGNH hydrolase-type esterase superfamily protein;(source:Araport11)
AT1G09310.1	ABA responsive trichome formation regulator.
AT5G57850.1	ADCL encodes a protein that acts as a 4-amino-4-deoxychorismate lyase. It catalyzes the production 4-aminobenzoate (pABA) production which is required for folate biosynthesis. The enzyme localizes to chloroplasts based on an import assay and GFP localization experiments. Involved in D-Amino Acid Stimulated Ethylene Production.
AT1G48860.1	5-enolpyruvylshikimate-3-phosphate synthase involved in shikimic acid biosynthesis.
AT2G46240.1	A member of Arabidopsis BAG (Bcl-2-associated athanogene) proteins, plant homologs of mammalian regulators of apoptosis. Expression of BAG6 in leaves was strongly induced by heat stress. Knockout mutants exhibited enhanced susceptibility to fungal pathogen Botrytis cinerea. Plant BAG proteins are multi-functional and remarkably similar to their animal counterparts, as they regulate apoptotic-like processes ranging from pathogen attack, to abiotic stress, to plant development. The mRNA is cell-to-cell mobile.
AT3G19710.1	Belongs to the branched-chain amino acid aminotransferase gene family. Encodes a methionine-oxo-acid transaminase. Involved in the methionine chain elongation pathway that leads to the ultimate biosynthesis of methionine-derived glucosinolates.
AT1G36390.1	Chloroplast GrpE protein.
AT1G76080.1	Encodes a thioredoxin like protein. Localizes to the chloroplast and is redistributed to the chloroplast envelope under heat stress. It is involved in non host resistance and thermotolerance.
AT1G17470.1	Encodes a member of the DRG (developmentally regulated G-protein) family expressed throughout the plant, with highest expression in actively growing tissues. Has GTPase activity.
AT3G03990.1	Encodes an alpha/beta hydrolase essential for strigolactone signaling. Degradation of the protein is promoted by strigolactone. The mRNA is cell-to-cell mobile.
AT5G66190.1	Encodes a leaf-type ferredoxin:NADP(H) oxidoreductase. It is present in both chloroplast stroma and thylakoid membranes but is more abundant in the thylakoid. The affinity of this enzyme for ferredoxin is slightly, but significantly, higher than AtLFNR2, an isoform of the same enzyme. AtLFNR1 forms a heterodimer with AtFNR2 and is also a prerequisite to attach AtFNR2 to the thylakoid membrane.

AT1G20020.1	Encodes a leaf-type ferredoxin:NADP(H) oxidoreductase. It is present in both chloroplast stroma and thylakoid membranes but is more abundant in the stroma. The mRNA is cell-to-cell mobile.
AT5G19940.1	Putative fibrillin; enables plants to cope with moderate light stress and affects cadmium tolerance.
AT3G52930.1	Fructose 1,6-bisphosphate aldolase.
AT2G25450.1	Encodes a 2-oxoacid-dependent dioxygenase involved in the production of 2-hydroxybut-3-enyl glucosinolate.
AT2G31570.1	glutathione peroxidase GPx
AT1G67280.1	Encodes a Ni ²⁺ dependent glyoxylase.
AT5G20140.2	Encodes a haem-binding protein, HBP5. HBP5 binds haem and interacts with the haem oxygenase, HY1. Disrupting the binding of HBP5 to HY1 leads to oxidative stress.
AT5G59750.2	monofunctional riboflavin biosynthesis protein RIBA 3;(source:Araport11)
AT3G02780.1	Encodes a protein with isopentenyl diphosphate:dimethylallyl diphosphate isomerase activity. There is genetic evidence that it functions in the mevalonate, but not the MEP biosynthetic pathway.
AT4G38800.1	Encodes one of the 5'-methylthioadenosine nucleosidases (AT4G38800/MTN1; AT4G34840/MTN2). Double mutant, mtn1-1 mtn2-1, retains approximately 14% of the MTN enzyme activity present in the wild type and displays a pleiotropic phenotype that includes altered vasculature and impaired fertility.
AT1G24020.1	MLP-like protein 423;(source:Araport11)
AT1G17370.1	Encodes an RNA-binding protein involved in stress granule formation. Regulated by a transposable element small RNA.
AT1G48600.2	Encodes a phosphoethanolamine N-methyltransferase that catalyses the last two methylation steps of the three sequential methylations of phosphoethanolamine (PEA) that are required for the synthesis of phosphocholine (PCho) in plants.
AT3G55250.1	Encodes a nucleus-encoded protein, Photosystem I Assembly 3 (PSA3), that is required for PSI accumulation.
AT5G54900.1	RNA-binding protein 45A;(source:Araport11)
AT3G56070.1	rotamase cyclophilin 2 (ROC2) exhibiting peptidyl-prolyl cis-trans isomerase activity involved in signal transduction.
AT2G16600.1	Encodes cytosolic cyclophilin ROC3. The mRNA is cell-to-cell mobile.
AT3G62030.2	nuclear-encoded chloroplast stromal cyclophilin CYP20-3 (also known as ROC4). Protein is tyrosine-phosphorylated and its phosphorylation state is modulated in response to ABA in <i>Arabidopsis thaliana</i> seeds.
AT4G39540.3	Encodes a shikimate kinase. Its transcripts appear to be expressed in vegetative tissues and developing embryos. SK2 transcript levels rise in response to <i>Phytophthora infestans</i> spores. SK2 is believed to be localized to the chloroplast.
AT3G50500.2	encodes a member of SNF1-related protein kinases (SnRK2) whose activity is activated by ionic (salt) and non-ionic (mannitol) osmotic stress. Enzyme involved in the ABA signaling during seed germination, dormancy and seedling growth.
AT3G10130.1	Chloroplast plastoglobule localized heme binding protein.
AT5G20840.1	Phosphoinositide phosphatase family protein;(source:Araport11)
AT1G15710.1	Plastidic argonate dehydrogenase with minimal prephenate dehydrogenase activity. Contains a single catalytic domain.

AT2G27860.1	Encodes UDP-d-apirose/UDP-d-xylose synthase that requires NAD ⁺ for enzymatic activity and is strongly inhibited by UDP-d-galacturonate.
AT1G24100.1	Encodes a UDP-glucose:thiohydroximate S-glucosyltransferase, involved in glucosinolate biosynthesis

Table S6 (CG_C5)

Gene Model Name	Gene Model Description
AT5G19860.1	transmembrane protein, putative (Protein of unknown function, DUF538);(source:Araport11)
AT5G16010.1	3-oxo-5-alpha-steroid 4-dehydrogenase family protein;(source:Araport11)
AT5G18970.1	AWPM-19-like family protein;(source:Araport11)
AT3G01820.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein;(source:Araport11)
AT4G21580.1	oxidoreductase, zinc-binding dehydrogenase family protein;(source:Araport11)
AT1G65020.1	plasma protein;(source:Araport11)
AT2G43110.1	U3 containing 90S pre-ribosomal complex subunit;(source:Araport11)
AT1G53880.1	translation initiation factor eIF-2B subunit alpha;(source:Araport11)
AT4G15770.1	RNA binding protein;(source:Araport11)
AT3G21770.1	Peroxidase superfamily protein;(source:Araport11)
AT3G03920.1	H/ACA ribonucleoprotein complex, subunit Gar1/Naf1 protein;(source:Araport11)
AT3G14690.2	putative cytochrome P450The mRNA is cell-to-cell mobile.
AT5G43830.1	aluminum induced protein with YGL and LRDR motifs;(source:Araport11)
AT1G75670.1	DNA-directed RNA polymerase;(source:Araport11)
AT4G31140.1	O-Glycosyl hydrolases family 17 protein;(source:Araport11)
AT2G34460.1	Rossmann-fold protein located in chloroplast.
AT3G07030.5	Together with other ALBA proteins phase separates into stress granules and processing bodies under heat stress and directly binds to and recruits mRNAs to protect from degradation. Confers thermotolerance through stabilizing HSF messenger RNAs in cytoplasmic granules.
AT1G20220.1	Together with other ALBA proteins phase separates into stress granules and processing bodies under heat stress and directly binds to and recruits mRNAs to protect from degradation. Confers thermotolerance through stabilizing HSF messenger RNAs in cytoplasmic granules.
AT3G44110.1	Homologous to the co-chaperon DNAJ protein from E coli. Member of the HSP40 family. Interacts with HSP70-4(AT3G12580) to mediate heat shock response. Formation of HSGs is dependent on the farnylation status of J3.

AT4G26600.5	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein;(source:Araport11)
AT1G80750.1	Cytosolic ribosomal 60S subunit protein.
AT3G23620.1	BRIX domain containing protein, similar to RNA biogenesis factors in yeast. Binds rRNA and likely also functions in RNA biogenesis in Arabidopsis. Essential gene, mutants are embryo lethal and does not transmit well through the gametophyte.
AT3G61860.1	encodes an arginine/serine-rich splicing factor. transcript is alternatively spliced and is differentially expressed in different tissues (flowers, roots, stems, and leaves) examined. Barta et al (2010) have proposed a nomenclature for Serine/Arginine-Rich Protein Splicing Factors (SR proteins): Plant Cell. 2010, 22:2926.
AT3G09640.1	Encodes a cytosolic ascorbate peroxidase APX2. Ascorbate peroxidases are enzymes that scavenge hydrogen peroxide in plant cells. Eight types of APX have been described for Arabidopsis: three cytosolic (APX1, APX2, APX6), two chloroplastic types (stromal sAPX, thylakoid tAPX), and three microsomal (APX3, APX4, APX5) isoforms.
AT5G62390.1	A member of Arabidopsis BAG (Bcl-2-associated athanogene) proteins, plant homologs of mammalian regulators of apoptosis. Plant BAG proteins are multi-functional and remarkably similar to their animal counterparts, as they regulate apoptotic-like processes ranging from pathogen attack, to abiotic stress, to plant development. Localized to the ER. Necessary for the proper maintenance of the unfolded protein response during heat and cold tolerance.
AT1G09210.1	Encodes one of three Arabidopsis calreticulins. Post-transcriptionally regulates together with CRT1 VAMP721/722 levels under ER stress.
AT2G40060.1	Encodes a clathrin that is localized to the cortical division zone and the cell plate and colocalizes with TPLATE during cell plate anchoring. The mRNA is cell-to-cell mobile.
AT3G14660.1	putative cytochrome P450 The mRNA is cell-to-cell mobile.
AT5G22060.1	Co-chaperonin similar to E. coli DnaJ
AT2G17840.1	Identified as drought-inducible gene by differential hybridization. Upregulated by high light, drought, cold and salt stress determined by microarray analysis.
AT1G10370.1	Encodes GSTU17 (Glutathione S-Transferase U17). Functions as a negative component of stress-mediated signal transduction pathways in drought and salt stress responses.
AT1G60550.1	enoyl-CoA hydratase/isomerase D;(source:Araport11)
AT3G05600.1	Encodes a cytosolic epoxide hydrolase capable of acting on 9,10-epoxystearic acid and on 12,13-epoxyoctadec-9-enoic acid that is involved in the synthesis of poly-hydroxylated cutin monomers.
AT3G19760.1	Encodes an RNA helicase that may be a component of the Exon Junction Complex. Subcellular localization is modulated by stress. Under normal conditions it is localized to the nucleoplasm but under hypoxic conditions it localizes to the nucleolus and splicing speckles.
AT5G52470.1	Encodes a fibrillarin, a key nucleolar protein in eukaryotes which associates with box C/D small nucleolar RNAs (snoRNAs) directing 2'-O-ribose methylation of the rRNA. This gene also encodes a novel box C/D snoRNA, U60.1f in its fifth intron that accumulates in seedlings and that their targeted residue on the 25 S rRNA is methylated.
AT4G25630.1	encodes a fibrillarin, a key nucleolar protein in eukaryotes which associates with box C/D small nucleolar RNAs (snoRNAs) directing 2'-O-ribose methylation of the rRNA. This gene also encodes a novel box C/D snoRNA, U60.2f in its fifth intron that accumulates in seedlings and that their targeted residue on the 25 S rRNA is methylated. The mRNA is cell-to-cell mobile.
AT4G04020.1	Fibrillin precursor protein. The fibrillin preprotein, but not the mature protein interacts with ABI2. Regulated by abscisic acid response regulators. Involved in abscisic acid-mediated photoprotection. The mRNA is cell-to-cell mobile.

AT4G22240.1	Involved in photoprotection of photosystem II. The RVSI and twin-positive motifs in the transit peptide are necessary for efficient leucoplast import of prFB.
AT4G25340.1	Encodes a member of the FKBP-type immunophilin family that functions as a histone chaparone. Binds to 18S rDNA and represses its expression. The N-terminal nucleoplasmin domain interacts with H2A/H2B and H3/H4 histone oligomers, individually, as well as simultaneously, suggesting two different binding sites for H2A/H2B and H3/H4.
AT4G10710.1	encodes a component of the FAcilitates Chromatin Transcription (FACT) complex, SPT16. Along with SSRP1 binds to the promoter of FLC.
AT1G79920.1	Heat shock protein 70 (Hsp 70) family protein;(source:Araport11)
AT5G22650.1	Encodes a member of a plant-specific class of histone deacetylases. Controls the development of adaxial/abaxial leaf polarity. Its mRNA is widely expressed in stems, leaves, flowers and young siliques. Plant lines expressing RNAi constructs directed against this gene showed a marked reduction in agrobacterium-mediated root transformation.
AT3G44750.1	Encodes a histone deacetylase. Controls the development of adaxial/abaxial leaf polarity. Two lines with RNAi-directed against this gene show reduced Agrobacterium-mediated DNA transformation of the roots. Involved in development of the vascular tissue of the stem by affecting cell proliferation and differentiation.
AT1G56110.1	NOP56-like protein
AT3G51660.1	Chemokine-like MDL protein; modulate flowering time and innate immunity in plants.
AT4G37910.1	Mitochondrial chaperone that interacts with COX2 and functions in cytochrome oxidase assembly. Mutants have pleiotropic defects in growth and development including defects in auxin gradients that may be secondary effects related to mitochondrial metabolism.
AT3G18580.1	Member of the family of canonical mitochondrial DNA binding proteins. Single-stranded binding protein which does not interfere with MMEJ.
AT5G48180.1	Encodes a nitrile-specifier protein NSP5. NSP5 is one out of five (At3g16400/NSP1, At2g33070/NSP2, At3g16390/NSP3, At3g16410/NSP4 and At5g48180/NSP5) <i>A. thaliana</i> epithiospecifier protein (ESP) homologues that promote simple nitrile, but not epithionitrile or thiocyanate formation.
AT1G48920.1	Encodes ATNUC-L1 (NUCLEOLIN LIKE 1), the predominant form of the two nucleolin proteins found in Arabidopsis. This protein is involved in rRNA processing, ribosome biosynthesis, and vascular pattern formation. PARL1 localizes to the nucleolus and parl1 mutants accumulate elevated levels of the unspliced 35S pre-rRNA. parl1 mutants also have defects in cotyledon, leaf, sepal, and petal vein patterning and have reduced stature, reduced fertility, increased bushiness, and reduced root length. The sugar-induced expression of ribosome proteins is also reduced in parl1 mutants. The mRNA is cell-to-cell mobile.
AT4G25130.1	Encodes a chloroplast-localized methionine sulfoxide reductase that is a member of the MSRA family. Involved in protection of chloroplasts from oxidative stress.
AT4G16500.1	Cystatin/monellin superfamily protein;(source:Araport11)
AT1G15200.3	protein-protein interaction regulator family protein;(source:Araport11)
AT5G62190.1	Encodes a ATP-dependent RNA unwinding protein targeted to the nucleolus and presumably involved in translation by assisting ribosome maturation. DEAD/DEAH box RNA helicase PRH75
AT3G24160.1	Encodes a putative Type 1 membrane protein (PMP).

AT5G14800.1	Delta 1-pyrroline-5-carboxylate reductase, catalyzes the final step in proline biosynthesis from glutamate and ornithine. In situ hybridization indicated that under normal growth conditions, the highest concentration of P5CR transcripts occurs in the cortical parenchyma, phloem, vascular cambium and pith parenchyma in the vicinity of the protoxylem. Single gene in Arabidopsis.
AT5G47200.1	AtRabD2b encodes a Rab GTPase, which plays important roles in pollen development, germination and tube elongation. The mRNA is cell-to-cell mobile.
AT4G17530.1	AtRabD2c encodes a Rab GTPase, which plays important roles in pollen development, germination and tube elongation.
AT5G01410.1	Encodes a protein predicted to function in tandem with PDX2 to form glutamine amidotransferase complex with involved in vitamin B6 biosynthesis.
AT4G29410.1	Ribosomal L28e protein family; (source: Araport11)
AT5G46160.1	Ribosomal protein L14p/L23e family protein; (source: Araport11)
AT5G64670.1	Ribosomal protein L18e/L15 superfamily protein; (source: Araport11)
AT5G65900.1	DEA(D/H)-box RNA helicase family protein; (source: Araport11)
AT1G18630.1	encodes a glycine-rich RNA binding protein.
AT1G60650.1	Encodes one of the zinc finger-containing glycine-rich RNA-binding proteins involved in cold tolerance: AT3G26420 (ATRZ-1A), AT1G60650 (AtRZ-1b), AT5G04280 (AtRZ-1c). It also, along with AtRZ-1c, plays important roles in plant development, pre-mRNA splicing, and general gene expression.
AT3G03060.1	Homologue of animal ATPase Family AAA Domain-Containing Protein 3 (ATAD3), which is involved in mitochondrial nucleoid organization; interacts with SHOT1.
AT1G70310.1	Spermidine synthase.
AT4G27585.1	Encodes a stomatin-like protein that is present in a mitochondrial membrane-bound 3 MDa protein complex and is involved in the assembly of mitochondrial respiratory supercomplexes. There is an observed increase in abundance of respiratory complex III2 in SLP1 single and double knockout mutants. The protein is not redundant with Arabidopsis SLP2 (At5g54100).
AT5G54100.1	SPFH/Band 7/PHB domain-containing membrane-associated protein family; (source: Araport11)
AT5G15490.1	Encodes one of four UDP-glucose dehydrogenase (UGD) genes. Mutation of this gene in combination with UGD2 leads to swollen plant cell walls and severe developmental defects associated with changes in pectic polysaccharides.
AT5G39320.1	UDP-glucose 6-dehydrogenase family protein; (source: Araport11)
AT2G36790.1	The At2g36790 gene encodes a UDP-glucose:flavonol-3-O-glycoside-7-O-glucosyltransferase (UGT73C6) attaching a glucosyl residue to the 7-O-position of the flavonols kaempferol, quercetin and their 3-O-glycoside derivatives.
AT2G01210.1	ZAR1 encodes a plasma membrane localized leucine-rich repeat receptor-like kinase (LRR-RLK) that contains a putative CaM-binding domain and a Gβ-binding motif within its intracellular kinase region. Homozygous of function mutations are embryo-lethal and fail to properly make the first asymmetric division of the zygote. ZAR1 interacts with both CaM and Gβ in vivo and that interaction activates ZAR1 kinase activity.

Supplementary Material & Method 1

Electrophoresis and in gel trypsin digestion

Purified nuclei were diluted four times in water to reach 1X Laemmli buffer. 32 µl were loaded on an in-house poured 4% acrylamide stacking gel. Gel was stained with Coomassie Blue and the stacking bands were manually excised. Proteins were then reduced, alkylated and digested overnight at room temperature with modified trypsin at a 1:100 enzyme:protein ratio (Promega, Madison, USA). Peptides were extracted during 1 hour with 90 µL of 80% acetonitrile, 0.1% formic acid. Peptide mixtures were then dried and resuspended in water acidified with 0.1% formic acid.

Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) Analyses

LC-MS/MS analyses of peptide extracts were performed on a NanoAcquity LC-system (Waters, Milford, MA, USA) coupled to a Q-Exactive plus Orbitrap (Thermo Fisher Scientific, Waltham, MA, USA) mass spectrometer equipped with a nanoelectrospray ion source. Mobile phase A (99.9% water and 0.1% FA) and mobile phase B (99.9% acetonitrile and 0.1%FA) were delivered at 400 nl/min. Samples were loaded into a Symmetry C18 precolumn (0.18 x 20 mm, 5 µm particle size, Waters) over 3 minutes in 1% buffer B at a flow rate of 5 µL/min. This step was followed by reverse-phase separation at a flow rate of 400 nl/min using an ACQUITY UPLC® BEH130 C18 separation column (250 mm x 75 µm id, 1.7 µm particle size, Waters). 1 µl of peptide mixtures were eluted using a gradient from 1% to 35% B in 79 minutes, from 35% B to 90% B in 1 minute, maintained at 90% B for 5 minutes and the column was reconditioned at 1% B for 20 minutes.

The Q-Exactive Plus Orbitrap instrument was operated in data dependent acquisition mode by automatically switching between full MS and consecutive MS/MS acquisitions. Survey full scan MS spectra (mass range 300-1800) were acquired with a resolution of 70,000 at 200 m/z with an automatic gain control (AGC) fixed at 3×10^6 ions and a maximum injection time set at 50 ms. The ten most intense peptide ions in each survey scan with a charge state ≥ 2 were selected for MS/MS fragmentation. MS/MS scans were performed at 17,500 resolution at 200 m/z with a fixed first mass at 100 m/z, AGC was fixed at 1×10^5 and the maximum injection time was set to 100 ms. Peptides were fragmented by higher-energy collisional dissociation (HCD) with a normalised collision energy set to 27. Peaks selected for fragmentation were automatically put on a dynamic exclusion list for 60 s and peptide match selection was turned on. MS data were saved in .raw file format (Thermo Fisher Scientific) using XCalibur.

LC-MS/MS data interpretation and validation

Raw files were converted to .mgf peaklists using MsConvert (using MSAngel) and were submitted to Mascot database searches (version 2.6.2, MatrixScience, London, UK) against an *Arabidopsis thaliana* protein sequences database downloaded from *The Arabidopsis Information Resource TAIR* site (TAIR10 version gene model), common contaminants and decoy sequences were added. The concatenated database contains 2 x 27 534 protein entries. Spectra were searched with a mass tolerance of 5 ppm in MS mode and 0.07 Da in MS/MS mode. One trypsin missed cleavages was tolerated. Carbamidomethylation of cysteine residues was set as a fixed modification. Oxidation of methionine residues and acetylation of protein n-termini were set as variable modifications. Identification results were imported into Proline software (<http://proline.profi-proteomics.fr/>) for

validation. Peptide Spectrum Matches (PSM) with pretty rank equal to one were retained. False Discovery Rate was then optimized to be below 1% at PSM level using Mascot Adjusted E-value and below 1% at Protein Level using Mascot Mudpit score. Only protein sets with at least a specific peptide were retained ¹.

Label Free Quantification and Statistical Analysis

Label Free Quantification

Peptides Abundances were extracted thanks to Proline software version 2.0 (<http://proline.proteomics.fr/>) using a m/z tolerance of 5 ppm. Alignment of the LC-MS runs was performed using Loess smoothing. No Cross Assignment was performed. The best 2+ or 3+ peptide ion was used to assign an abundance to a peptide. Peptides abundances were normalized using the median across the 12 runs. Protein Abundance was then computed by summing the abundance of all the peptides except the one with a doubtful peak assignation (different peptides identification sharing the same speak).

Statistical Analysis

Protein Abundances were loaded into Prostar software version 1.18.6 (<http://www.prostar-proteomics.org/>) and associated to their conditions (22°C, 37°C 4h, 37°C 24h, R22°C). Proteins with at least 3 values in at least one condition (3/3 replicates) were kept for further statistical analysis. Contaminants were removed. Residual Missing Values were imputed in a conservative way (quantile 2.5%). Pairwised Limma t-tests were performed.

P-values calibration was corrected using adapted Benjamini-Hochsberg method, and FDR was set to ~1%. More precisely, the comparison of 22°C vs 37°C 24h lead to 1.00% FDR using p-values below 0.00132, the comparison of 22°C vs 37°C 4h lead to 1.03% FDR using p-values below 0.00501, the comparison of 22°C vs R22°C lead to 1.01% FDR using p-values below 0.00955, the comparison of 37°C 4h vs 37°C 24h lead to 1.14% FDR using p-values below 0.00389, the comparison of 37°C 4h vs R22°C lead to 1.02% FDR using p-values below 0.00692, and finally the comparison of 37°C 24h vs R22°C lead to 0.98% FDR using p-values below 0.00479.

- 1 Bouyssie, D. *et al.* Proline: an efficient and user-friendly software suite for large-scale proteomics. *Bioinformatics* (Oxford, England) **36**, 3148-3155 (2020).
<https://doi.org:10.1093/bioinformatics/btaa118>

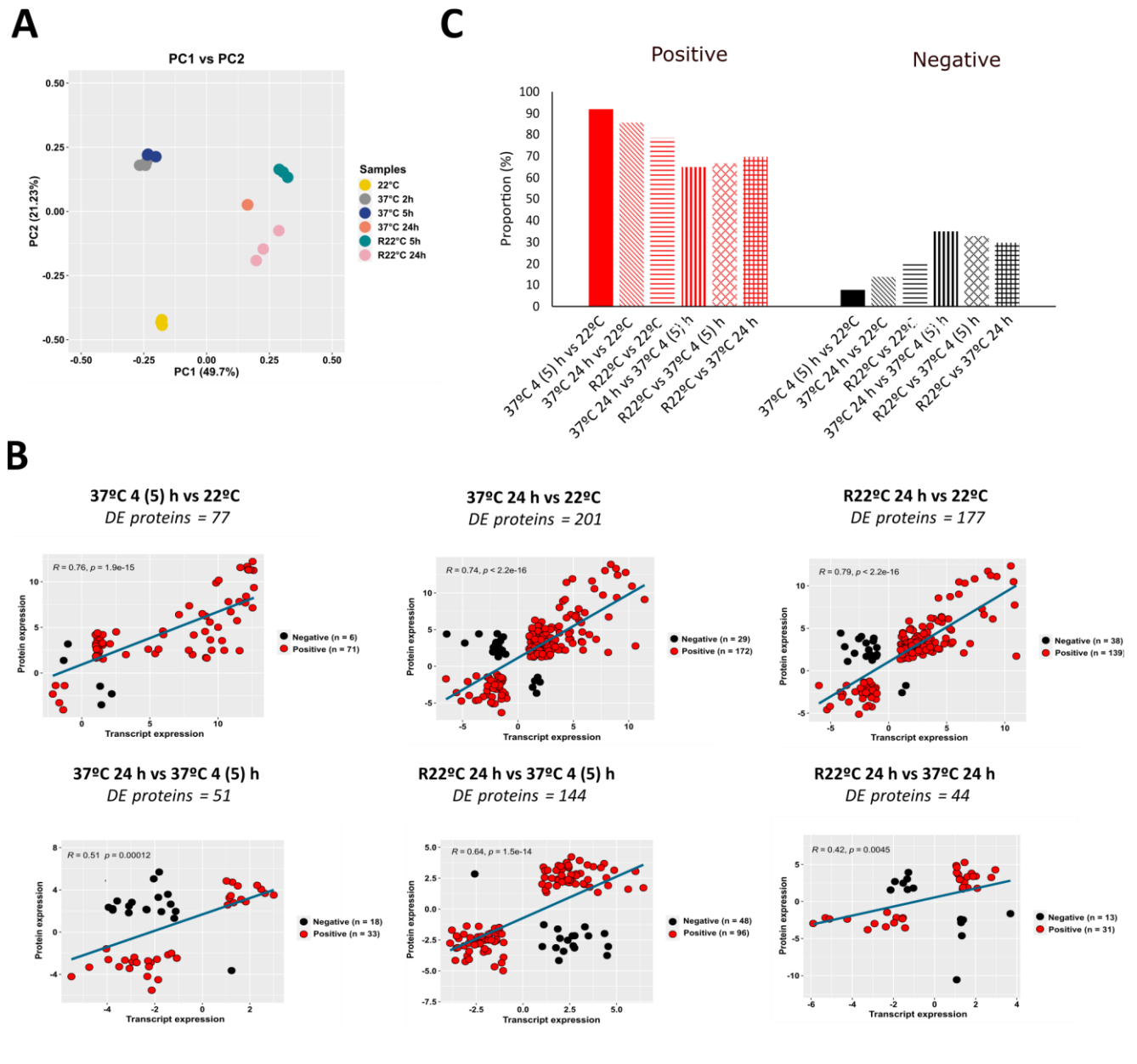
3.1.2. Comparative analysis of the transcriptome and nuclear proteome of Arabidopsis under heat stress

Previously in our group, the transcriptome of Arabidopsis had been obtained from 15-day-old Col-0 seedlings under normal conditions (22°C), heat stress (37°C 2, 5 and 24 h) and two recovery time periods, consisting of 5 and 24 hours at 22°C after 24 hours at 37°C, (R22°C 5 h and R22°C 24 h, respectively). The total amount of loci detected in the transcriptomic analysis is 23,158 at 22°C, 22,127 at 37°C 2 h, 22,383 at 37°C 5 h, 23,937 at 37°C 24 h, 24,971 at R22°C 5 h and 24,554 at R22°C 24 h. The PCA portrayed in Figure 3.1A shows how the three replicates of each condition are grouped. The “37°C 24 h”, “R22°C 5 h” and “R22°C 24 h” samples show the highest difference when compared to “22°C” samples.

Given the analysis of the nuclear proteome of Arabidopsis under and after heat stress, a comparative study was conducted with the transcriptome of Arabidopsis due to the similarity in the conditions tested (from now on, R22°C 24 h will be referred as R22°C in the transcriptomic analysis). The first comparative analysis was carried out using the data from the individual comparisons: 37°C 4 (5) h vs 22°C, 37°C 24 h vs 22°C, R22°C vs 22°C, 37°C 24 h vs 37°C 4 (5) h, R22°C vs 37°C 4 (5) h and R22°C vs 37°C 24 h. Out of the 126 differentially accumulated proteins between 22°C and 37°C 4 h (37°C 4 h vs 22°C), 77 are also differently expressed on a transcriptomic level (37°C 5 h vs 22°C), which corresponds to 61%, approximately. The majority (71 proteins or 92% approx.) exhibit a positive correlation between the proteomic and transcriptomic analysis. Similarly, the amount of differentially expressed proteins in the transcriptome and nuclear proteome in the other five comparisons have been obtained (Figure 3.1B). The comparison R22°C vs 22°C exhibits the highest percentage of correlation between the transcriptome and nuclear proteome (71% of the differentially accumulated proteins are also differentially regulated in a transcriptional level in this comparison). This means that a significant change in the relative protein abundance in the nucleus corresponds to a significant change in the transcriptome. However, the comparison 37°C 4 (5) h vs 22°C shows the highest percentage of positive correlation between the nuclear proteome and the transcriptome. In other words, an increase (or decrease) in the relative protein concentration in the nucleus is accompanied by an increase (or decrease) in transcript levels (Figure 3.1C).

Next, the protein abundances of the members of each group described in the analysis of the nuclear proteome (EG, LG, TG, EPG, LPG, RG, ELG, LLG, TLG and CG) were compared to their transcriptomic profiles. In general, the transcript levels of the members of the groups are more heterogeneous and do not resemble their protein abundances (Figure 3.2). There are three cases where the transcriptomic profile and the proteomic profile in the nucleus behave similarly: the cluster 5 in the Transient Group (TG), the cluster 10 in the Late Persistent Group (LPG) and the cluster 12 in the Recovery Group (RG). In the rest of the cases, the transcriptomic and nuclear proteomic profiles do not resemble during and after heat stress. In other words, the transcript

levels of the members of each group (EG, LG, TG, EPG, LPG, RG, ELG, LLG, TLG and CG) do not fully correspond to their relative protein levels in the nucleus.



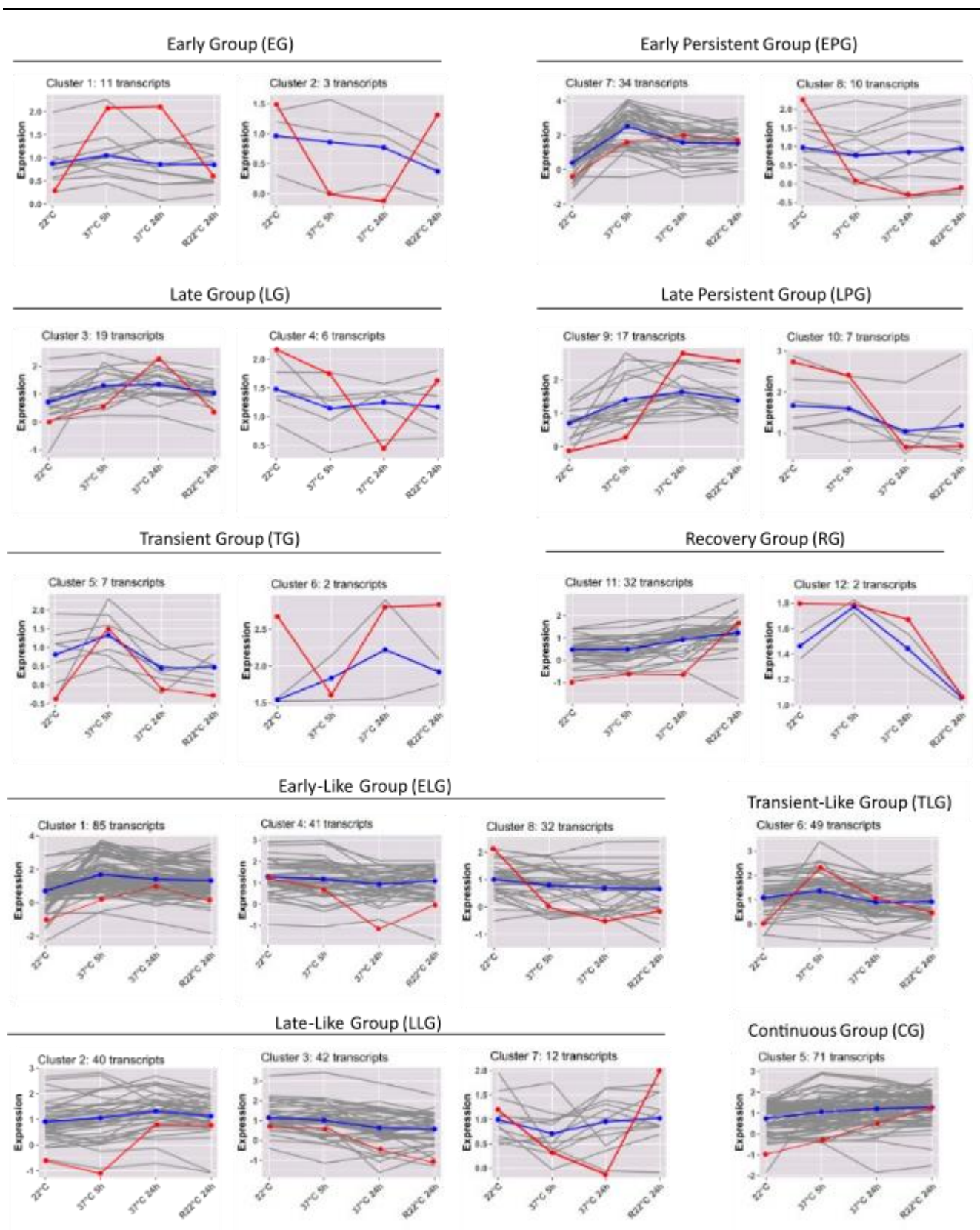


Figure 3.2. Transcriptomic profile under heat stress of the ten established groups. Transcriptome levels are represented for the ten identified groups in the analysis of the nuclear proteome (EG, LG, TG, EPG, LPG, RG, ELG, LLG, TLG and CG). The grey lines show the transcriptomic profiles of each member of the corresponding group, whereas the blue line represents the overall tendency of the transcriptomic profile. The red line shows an approximate representation of the proteomic profile in the nucleus for each group.

3.1.3. Overall results

In the first section of this PhD project, the nuclear proteome of *Arabidopsis* seedlings has been analysed under normal conditions (22°C), heat stress (37°C 4 and 24 hours) and recovery period (R22°C; 37°C 24 h + 22°C 24 h). Being a quantitative proteomic analysis, the amount of proteins exerting differential abundance in the nucleus has been addressed in the following comparisons: 37°C 4 h vs 22°C, 37°C 24 h vs 22°C, R22°C vs 22°C, 37°C 24 h vs 37°C 4 h, R22°C vs 37°C 4 h, and R22°C vs 37°C 24 h. Besides that, we identified 522 proteins whose relative abundance in the nucleus changes during and/or after heat stress. Thus, these 522 proteins have been classified in ten groups according to the kinetics of relative protein abundance in the nucleus throughout the four conditions tested: EG, LG, TG, EPG, LPG, RG, ELG, LLG, TLG and CG. Interestingly, several major nucleolar proteins have been found as members of the RG and CG, such as the RNA Pol I subunit NRPA3 (RG) or FIB1 and 2 (CG), among others. Both RG and CG are characterised by a pronounced protein accumulation in the recovery period (R22°C) compared to normal (22°C) and heat stress (37°C 4 and 24 hours) conditions.

Moreover, the transcriptome of *Arabidopsis* seedlings had been previously obtained in our group at 22°C, 37°C 2,5 and 24 hours, and two recovery periods, being R22°C 5 h (37°C 24 h + 22°C 5 h) and R22°C 24 h (37°C 24 h + 22°C 24 h). Due to the similar experimental conditions of the analysed transcriptome and nuclear proteome, a comparative analysis has been performed. Overall, this comparative analysis reveals how the transcriptomic levels and the relative protein abundances in the nucleus behave differently during and after heat stress. In other words, the transcriptomic profiles of the members of the ten groups established in the analysis of the nuclear proteome (EG, LG, TG, EPG, LPG, RG, ELG, LLG, TLG and CG) do not fully match the relative protein abundances in the nucleus.

3.2. Behaviour of nucleolar components to heat stress in Arabidopsis

3.2.1. The nucleolus and heat stress in Arabidopsis

Besides being a “ribosome factory”, the nucleolus has been labelled as a stress sensor over the years. This link between the nucleolus and abiotic stress has been well described in animals (reviewed by Boulon et al., 2010). Abiotic stressors such as, osmotic stress, heat stress, cold stress or hypoxia induce morphological changes in animal nucleoli (reviewed by Yang et al., 2018; reviewed by Hayashi and Matsunaga, 2019). Besides that, the tumor suppressor protein p53 is a key component in the defence response in stress scenarios in mammals. p53 is activated upon nucleolar stress, resulting in cell cycle arrest or apoptosis (reviewed by James et al., 2014). However, this relationship is not as well studied in plants. In contrast to mammals, plants do not possess an orthologue of p53. The plant equivalent would be the transcription factor family NAC, which participate in certain stress responses (reviewed by Ohbayashi and Sugiyama, 2018). More specifically, ANAC082 acts as a participant of the ribosomal stress response in plants (Ohbayashi et al., 2017).

The effect of abiotic stress, especially heat stress, on the nucleolar proteome has been addressed in Arabidopsis. Certain proteins accumulate or abandon the nucleolus in response to heat stress. For instance, Arabidopsis HIT4 is involved in the decondensation of chromocentres upon heat stress. Normally present in the chromocentres, HIT4 moves to the nucleolus in response to heat stress (Wang et al., 2015). Similarly, the component of the exon junction complex eIF4A-III migrates from the nucleoplasm to the nucleolus and splicing speckles in response to hypoxia. This may lead to the retention of certain mRNA species in the nucleolus, preventing their translation (Koroleva et al., 2009). In contrast, the Arabidopsis DEAD-box RNA helicases STRESS RESPONSE SUPPRESSOR1 and 2 (STRS1 and STRS2, respectively) experience a reversible re-localisation event from the nucleolus to the nucleoplasm under some stresses, including heat stress. Interestingly, these proteins positively regulate the heat stress response, as Arabidopsis *STRS1* and *STRS2* overexpression lines exhibit enhanced tolerance (Khan et al., 2014).

Previously in our group, a major change in the nucleolar architecture in Arabidopsis upon short and long exposure to heat stress (37°C 5, 8 and 24 hours) had been addressed. Analysis by transmission electronic microscopy (TEM) revealed three different classes of nucleoli before, during and after heat stress: regular nucleoli (class I), mild-affected nucleoli (class II) and severely affected nucleoli (class III; Figure 3.3A Left). The relative amount of each class of nucleoli was determined at each condition. Class I nucleoli are predominant at 22°C. However, the amount of class II and class III nucleoli dramatically increases upon heat stress (37°C 5, 8 and 24 hours). This disruption in the nucleolar structure is reversible, as class I nucleoli becomes the major class once the heat stress halts (Figure 3.3A Right; Darriere et al., 2022). In addition, the morphology of Arabidopsis nucleoli under different stressors has also been analysed by the use of 5-ethynyl uridine, which an analogue of uridine (Hayashi and Matsunaga, 2019). That way, the distribution of rRNA was examined under heat stress by

confocal microscopy. The authors describe the formation of speckled structures in the nucleolus upon exposure to 37°C for 2 h (Figure 3.3B).

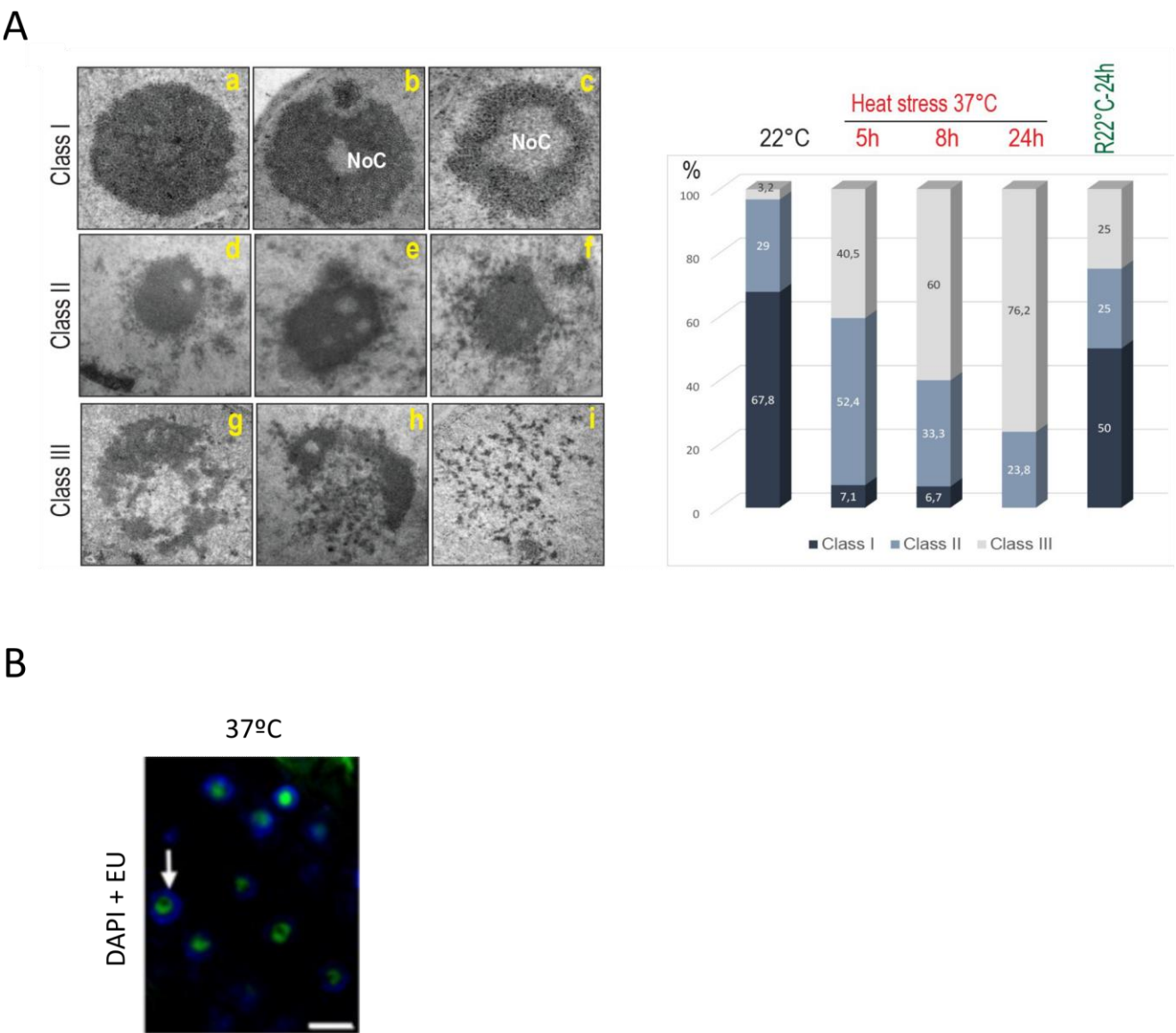


Figure 3.3. Effect of high temperatures in Arabidopsis nucleoli. (A) Left, Visualisation of Arabidopsis nucleoli under TEM. Three classes of nucleoli were designed according to the effects of heat stress: class I (non-affected), class II (mildly affected) and class III (severely affected). Right. Relative quantification of each class of nucleoli in the different stages. From Darriere et al., (2022). (B) Formation of rRNA speckled structures (white arrow) in the nucleolus of roots of 5-day-old Col-0 seedlings exposed to 37°C for 2 hours. Scale bar set at 10 µm. Figure modified from Hayashi and Matsunaga, (2019).

3.2.2. Goals and experimental approach

The effects of heat stress on the overall architecture of Arabidopsis nucleoli have been described in Arabidopsis. Nevertheless, the distribution of the major nucleolar components in Arabidopsis under heat stress has not been addressed yet. Whether these components stay in the nucleolar space or not remains unanswered. Thus, the main goal of this section is the analysis of the subcellular distribution of main nucleolar components before, during and after heat stress. Notably, the proteomic analysis described in the previous section (3.1.1. Quantitative proteomic profiling of Arabidopsis nuclei reveals distinct protein accumulation kinetics upon heat stress) revealed increasing abundance in the nucleus of major nucleolar proteins during heat stress and/or recovery period. These observations led us to investigate more precisely the subcellular localisation of major nucleolar components: the RNA Pol I subunit NRPA3, NUC1, the U3 snoRNP constituents FIB2 and C/D box snoRNAs, the pre-rRNA processing factors OLI2 and LAS1, and the heat activated NUC2.

For that, the subcellular localisation of these components has been analysed under normal conditions (22°C), heat stress (37°C 4 and 24 hours) and a recovery phase consisting of 24 hours at 22°C after 24 hours at 37°C (R22°C). Three main techniques have been employed to deepen into the localisation of these components. First of all, the localisation of certain proteins has been tackled by immunofluorescence. This is the case of NUC1, where an antibody against Arabidopsis NUC1 is available. Moreover, an antibody against HA has been employed for NRPA3^m and NUC2, since these proteins are tagged with FLAG-HA. Besides that, other proteins are fused with commercial fluorophores, including FIB2-YFP, GFP-OLI2 and LAS1-GFP. Finally, RNA FISH experiments have been performed to detect the C/D box snoRNAs. For that, Cy3-tagged probes have been designed to target these RNA molecules.

3.2.3. Results

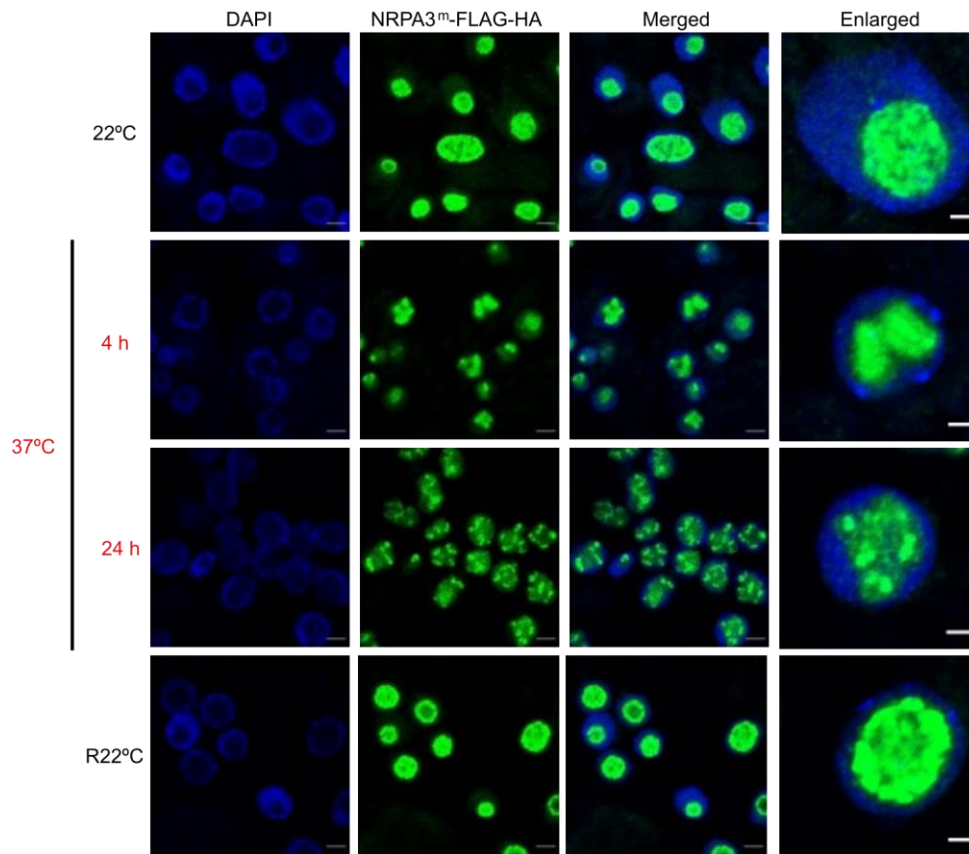
3.2.3.1. Subcellular distribution of NRPA3^m-FLAG-HA under heat stress

In Arabidopsis, RNA Pol I is a multimeric enzyme composed of several subunits (Table 1.1). One of them is NRPA3 or RPAC42 (At1g60850), which is present in the core region based on homology analysis (Ream et al., 2015). The subcellular localisation of NRPA3^m-FLAG-HA has been analysed in *NRPA3_{pro}:NRPA3^m-FLAG-HA* (*nrpa3*) seedlings using an antibody against HA. As expected, this protein is located in the nucleolus at 22°C, specifically in the FC and DFC. Upon heat stress (37°C 4 and 24 hours), NRPA3^m-FLAG-HA remains in the nucleolar space. Nevertheless, the protein seems to form foci-like structures (Figure 3.4A). In the recovery phase (R22°C), this protein stops forming foci, and the signal is present again in the FC and DFC (Figure 3.4A). A relative quantification of nucleoli displaying wild-type and foci-like signals was carried out at 22°C, 37°C 4 and 24 hours, and R22°C, to portray the behaviour of the distribution of NRPA3^m-FLAG-HA. Whereas no foci-like distribution is observed for NRPA3^m-FLAG-HA at 22°C, the amount of these foci increases to 74 and 95% at 37°C 4 and 24 hours, respectively. Nevertheless, when the temperature goes back to 22°C (R22°C), the number of nucleoli exhibiting foci-like signal is 4% (Figure 3.4B).

3.2.3.2. Response of NUC1, FIB2 and C/D box snoRNAs to high temperatures

As previously mentioned, the processing steps of the pre-rRNA species include the formation of a U3 snoRNP complex involved in the formation of the pre-40S particle. Even though this complex has not been structurally dissected in Arabidopsis, the components of the U3 snoRNP complex are found in Arabidopsis. For instance, NUC1 (At1g48920) is a major nucleolar protein involved in many steps of ribosome biogenesis (Kojima et al., 2007; Petricka and Nelson, 2007; Pontvianne et al., 2007). Besides that, rRNA in *nuc1* mutants is hypomethylated (2'-O-Methylation) in comparison to wild-type plants (Azevedo-Favory et al., 2021). The subcellular localisation of Arabidopsis NUC1 was analysed in Col-0 seedlings using an anti-NUC1. At 22°C, this protein is located in the periphery of the nucleolus. Upon exposure to short-termed heat stress (37°C 4 h), the distribution does not change in comparison to normal conditions. However, NUC1 forms ring-like structures in the nucleolar space at 24 hours at 37°C. Once the temperature decreases to 22°C (R22°C), these ring-like structures disappear, restoring the usual NUC1 distribution (Figure 3.5A). Once again, the percentage of nucleoli showing wild-type and ring-like distribution was quantified at every time point. Whereas most nucleoli exhibit wild-type distribution at 22°C and 37°C 4 hours (100 and 96%, respectively), 66% of the nucleoli show ring-like distribution at 37°C 24 h. Nevertheless, this fluorescence pattern changes in the recovery phase (R22°C), as 100% of the nucleoli exhibit the wild-type distribution of fluorescence (Figure 3.5B).

A



B

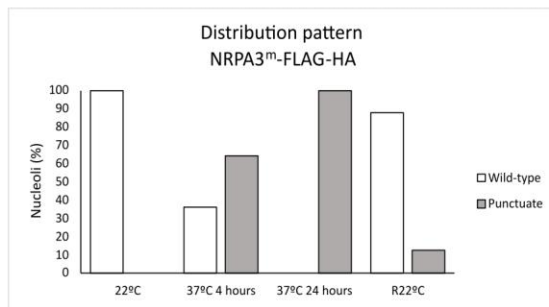
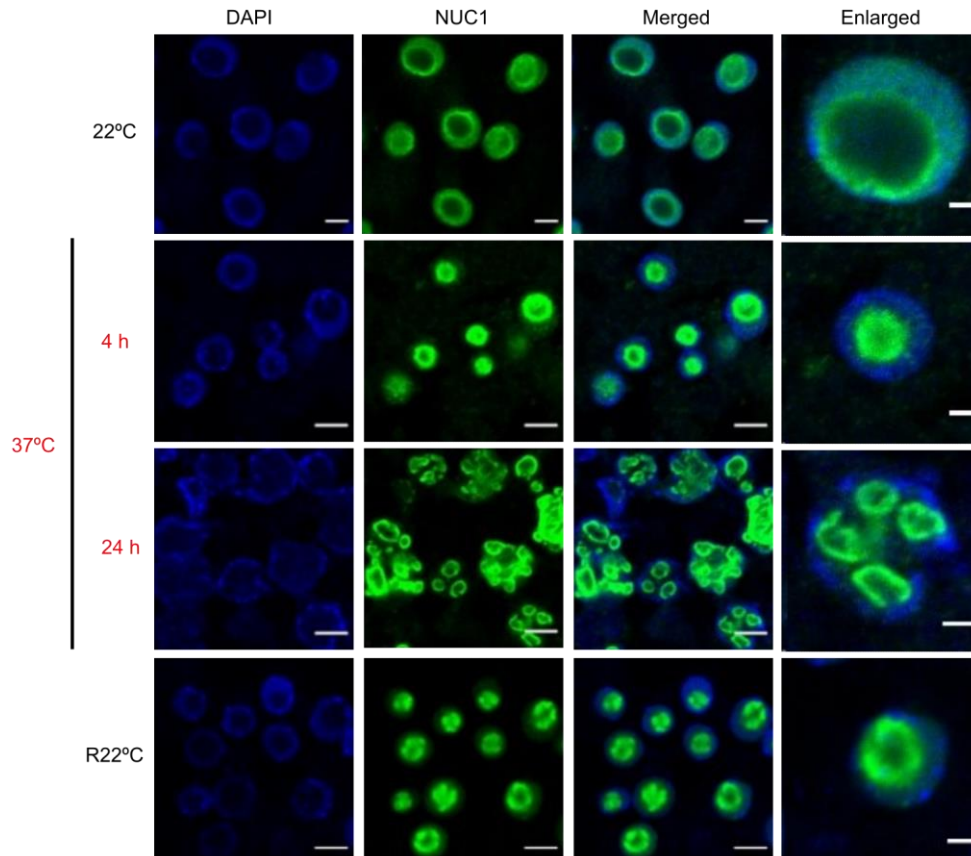


Figure 3.4. NRPA3^m-FLAG-HA forms punctuate structure under heat stress. (A) Subcellular distribution of the fusion protein NRPA3^m-FLAG-HA in roots of 5-day-old *NRPA3_{pro}:NRPA3^m-FLAG-HA* (*nrpa3*) seedlings exposed at 22°C, 37°C 4 h, 37°C 24 h, and R22°C. Scale bar set at 5 μm. On the right, one nucleus at each condition has been enlarged, where the scale is 2 μm. (B) Percentage of nucleoli exhibiting wild-type and punctuate (foci-like) distribution at the different time points. n = 194 for 22°C, n = 335 for 37°C 4h, n = 173 for 37°C 24 h and n = 254 for R22°C.

A



B

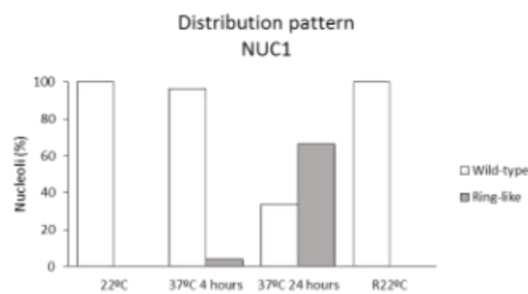


Figure 3.5. Subcellular distribution of NUC1 before, during and after heat stress. (A) Localisation of NUC1 in roots of 5-day-old Col-0 seedlings at 22°C, 37°C 4 h, 37°C 24 h, and R22°C. Scale bar set at 5 μ m. On the right, one nucleus at every time point has been enlarged, where the scale is 2 μ m. (B) Percentage of nucleoli exhibiting wild-type and ring-like pattern at the different time points. n = 139 for 22°C, n = 454 for 37°C 4 h, n = 229 for 37°C 24 h and n = 118 for R22°C.

Fibrillarin, which is involved in 2'-O-methylation, is one of the main proteins of the U3 snoRNP complex. As previously stated, the Arabidopsis genome possesses two genes encoding for fibrillarin: *FIB1* and *FIB2* (Barneche et al., 2000). Subcellular localisation analyses of *FIB2* (At4g25630) were performed using the overexpression line *35S_{pro}:FIB2-YFP* (Col-0). Under normal conditions, the *FIB2*-YFP fusion protein is located in

the nucleolus, specifically in the DFC. Upon heat stress (37°C 4 and 24 hours), FIB2-YFP forms some type of aggregates in the nucleolar space. At 37°C 4 h, these round structures are present in 33% of cells. In contrast, almost half of cells (44%) show these bodies at 37°C 24 h (Figure 3.6A and C). Moreover, few cells (4%) exhibit some type of ring-like distribution at 37°C 24 h (Figure 3.6B and C). In the recovery phase (R22°C), the distribution of FIB2-YFP is the same as at 22°C, being present in the FC of the nucleolus (Figure 3.6A and C).

The subcellular localisation of some C/D box snoRNAs has been analysed in Col-0 seedlings by RNA FISH using Cyanine-3 (Cy3)-tagged probes (Table 2.1). These snoRNAs include the major U3 snoRNAs [U3B (At5g53902) and U3D (At5g54075)], which are involved in cleavage of pre-rRNA. Moreover, other C/D box snoRNAs involved in methylation of rRNA are also included: At5gCDbox137.1 (At5g07725), At5gCDbox141.1 (At5g09000), U31b (At1g06263), U31a (At1g06233), At1gCDbox18.1 (At1g19373), At1gCDbox33.1 (At1g07393), and At4gCDbox117.1 (At4g08615; Table 3.1; Azevedo-Favory et al., 2021). Their subcellular localisation by RNA FISH reveals that they are uniformly distributed in the nucleolus at 22°C, 37°C 4 and 24 hours, and R22°C (Figure 3.7A). It must be remarked that the detection of the C/D box snoRNAs by RNA FISH has been contrasted using two different controls. The first one consisted of the lack of probes, so that background signal could be spotted. The second control is the treatment with RNase, so that the signal detected corresponds to RNA molecules (Figure 3.7B).

3.2.3.3. Behaviour of the fusion protein GFP-OLI2 under heat stress

OLI2 (also named *NOP2A*; At5g55920) is an orthologue of Nop2 in *Saccharomyces cerevisiae*, which is involved in the maturation of the 25S rRNA, methylating the C5 of cytosine 2870 of 25S rRNA in yeast (Sharma et al., 2013). In Arabidopsis, *OLI2* plays a role in the establishment of leaf abaxial-adaxial patterning (Maekawa and Yanagisawa, 2021). Besides that, the disruption of *OLI2* in Arabidopsis leads to defects in the pre-rRNA processing (Kojima et al., 2018). In contrast, its role in rRNA methylation has not been characterised yet (Burgess et al., 2015). The distribution of GFP-*OLI2* was achieved using an anti-GFP in the complementation line *OLI2_{pro}::GFP-OLI2 (oli2-1)*. The distribution of GFP-*OLI2* is similar at 22°C and 37°C 4 hours, being present in the nucleolar space in 100 and 80% of the cells, respectively. In contrast, 20% of cells at 37°C 4 h exhibit a ring-like distribution in the DFC and GC. The wild-type distribution decreases at 37°C 24 hours, where the ring-like structures are visible in the majority of cells (89%). However, this change in GFP-*OLI2* distribution is mostly reversible, as the ring-like distribution disappears in the recovery phase (R22°C; Figure 3.8A and C). In certain cases (17%), the signal was not completely uniform in the nucleolus at R22°C, showing some transition between the ring-like structures and the uniform distribution (Figure 3.8B and C).

A

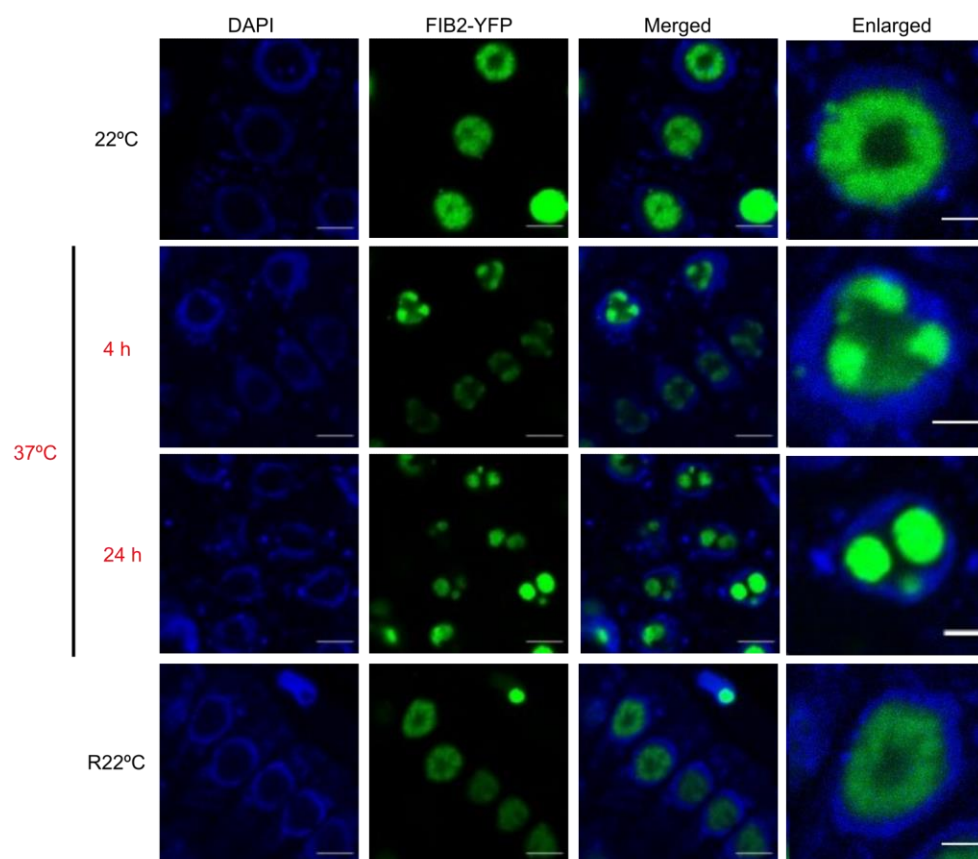
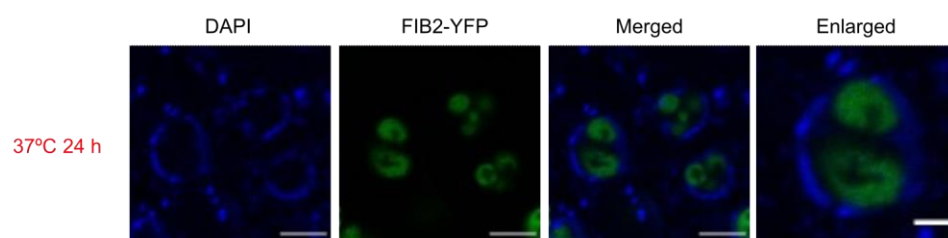


Figure 3.6. Behaviour of FIB2-YFP before, during and after heat stress in Arabidopsis.

(A) Subcellular distribution of the fusion protein FIB2-YFP in the roots of 5-day-old *35S_{pro}:FIB2-YFP* (Col-0) seedlings exposed at 22°C, 37°C 4 h, 37°C 24 h, and R22°C. Scale bar set at 5 µm. On the right, one nucleus at each condition has been enlarged, where the scale is 2 µm. (B) 5-day-old *35S_{pro}:FIB2-YFP* overexpression plants exhibit ring-like structures in the nucleolar space at 37°C 24 h. Scale bar set at 5 µm. On the right, one nucleus has been enlarged at 37°C 24 h, where the scale is 2 µm. (C) Percentage of nucleoli exhibiting the different distributions of FIB2-YFP at the different time points. n = 226 for 22°C, n = 249 for 37°C 4 h, n = 142 for 37°C 24 h and n = 193 for R22°C.

B



C

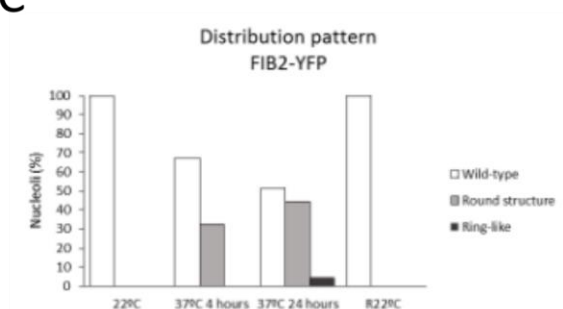


Table 3.1. List of the different C/D box snoRNAs analysed by RNA FISH

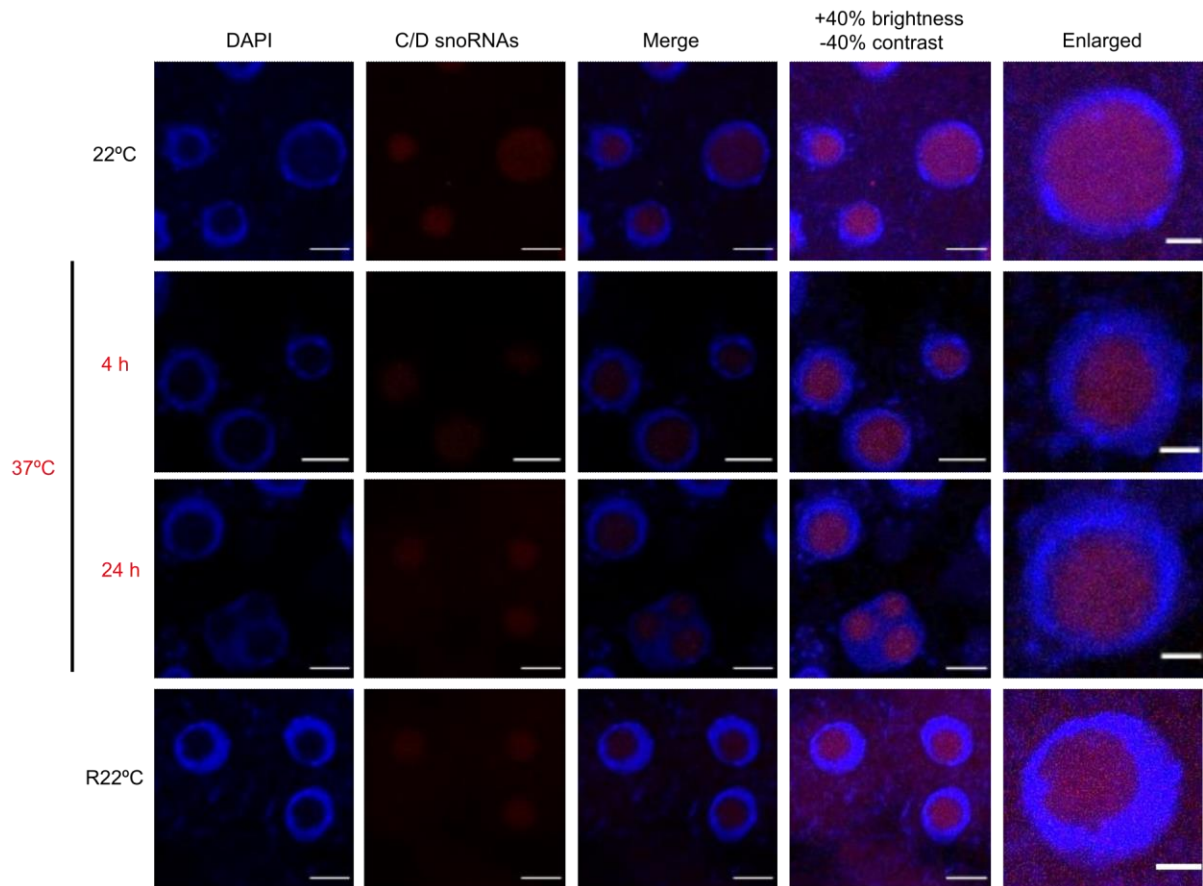
Nomenclature based on Azevedo-Favory et al., (2021)

Function	Name	Chromosome
Cleavage of pre-rRNA	U3B	5
	U3D	5
Methylation of rRNA	At5gCDbox137.1 (SnoR29-1)	5
	At5gCDbox141.1	5
	At1gCDbox25.2 (U31b)	1
	U31a	1
	At1gCDbox18.1	1
	At1gCDbox33.1 (AtsnoR65)	1
	At4gCDbox117.1	4

3.2.3.4. LAS1-GFP stays in the nucleolus under heat stress

LAS1 (At5g12220) is involved in the processing of the ITS2 site of the 45S pre-rRNA. The subcellular distribution of Arabidopsis LAS1 had been already performed in transiently transformed *Nicotiana benthamiana* leaves, with the construct pGWB5-35S_{pro}:AtLAS1-GFP. Arabidopsis LAS1 is located in the nucleolus, as well as its interactors AtGrc3 and APUM24 (Maekawa et al., 2018). Nevertheless, it was decided to generate Arabidopsis stable transformants with a construct containing the genomic sequence of Arabidopsis *LAS1*. The distribution of LAS1-GFP under heat stress and recovery phase has been analysed. At 22°C, LAS1-GFP is located exclusively in the nucleolus, excluded from the FC. At 37°C 4 and 24 hours, LAS1-GFP continues in the nucleolar space with no remarkable changes in its distribution pattern. Nevertheless, a decrease in fluorescence is observed at 37°C 24 h. As expected, the fusion protein is present in the nucleolus in the recovery phase (Figure 3.9).

A



B

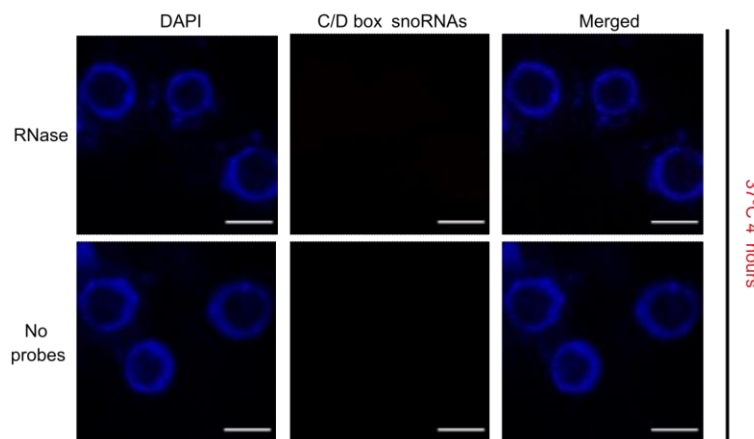


Figure 3.7. Distribution of C/D box snoRNAs under normal conditions, heat stress and recovery period in *Arabidopsis*. (A) RNA FISH was conducted in roots of 7-day-old Col-0 seedlings at 22°C, 37°C 4 h, 37°C 24 h, and R22°C. The “Merged” pictures have been modified so that the signals were more visible (+40% brightness and -40% contrast). Scale bar set at 5 μ m. On the right, one nucleus at each time point has been enlarged, where the scale is 2 μ m. (B) Controls for the RNA FISH experiments. The RNase and “no probes” controls have been performed at every time point (22°C, 37°C 4 and 24 hours, and R22°C). Nevertheless, the ones conducted at 37°C 4 h are shown. Scale set at 5 μ m.

A

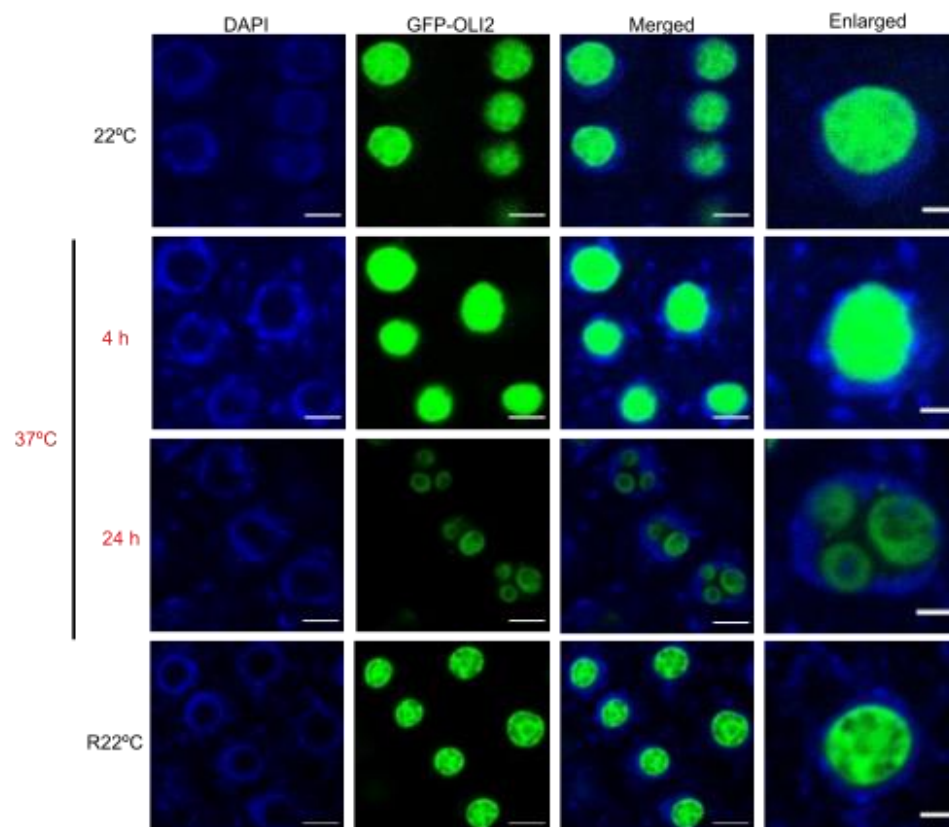
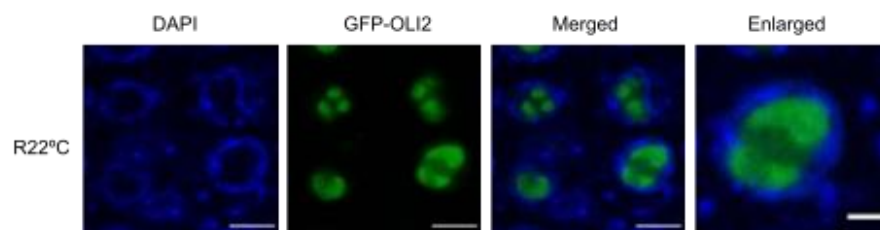
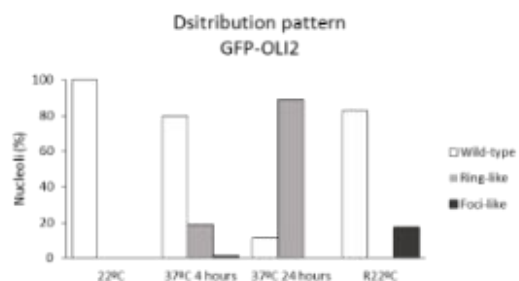


Figure 3.8. Behaviour of GFP-OLI2 under heat stress. (A) Subcellular distribution of GFP-OLI2 in roots of 5-day-old *OLI2_{pro}:GFP-OLI2* (*oli2-1*) seedlings at 22°C, 37°C 4 h, 37°C 24 h, and R22°C. Scale bar set at 5 μ m. On the right, one nucleus at every time point has been enlarged, where the scale is 2 μ m. (B) 5-day-old *OLI2_{pro}:GFP-OLI2* (*oli2-1*) seedlings showing foci-like aggregates at R22°C. Scale bar at 5 μ m. On the right, one nucleus has been enlarged, where the scale is 2 μ m. (C) Percentage of the different distribution patterns of GFP-OLI2 at the different time points. n = 211 for 22°C, n = 390 for 37°C 4 h, n = 168 for 37°C 24 h and n = 214 for R22°C.

B



C



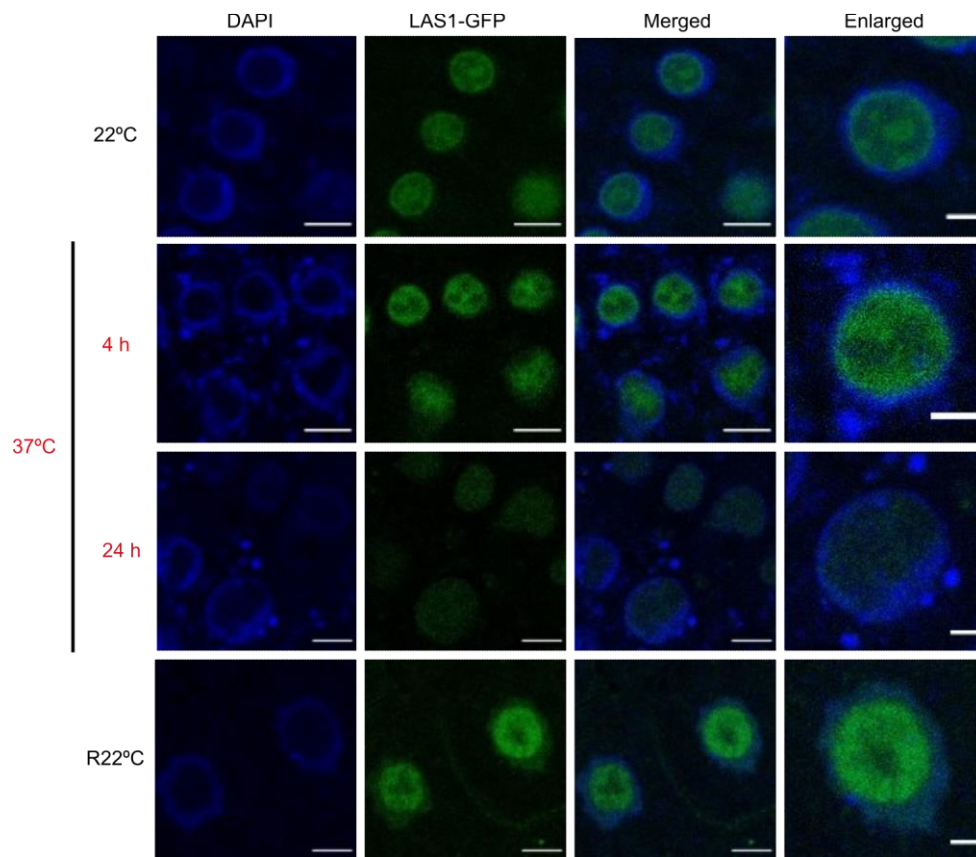
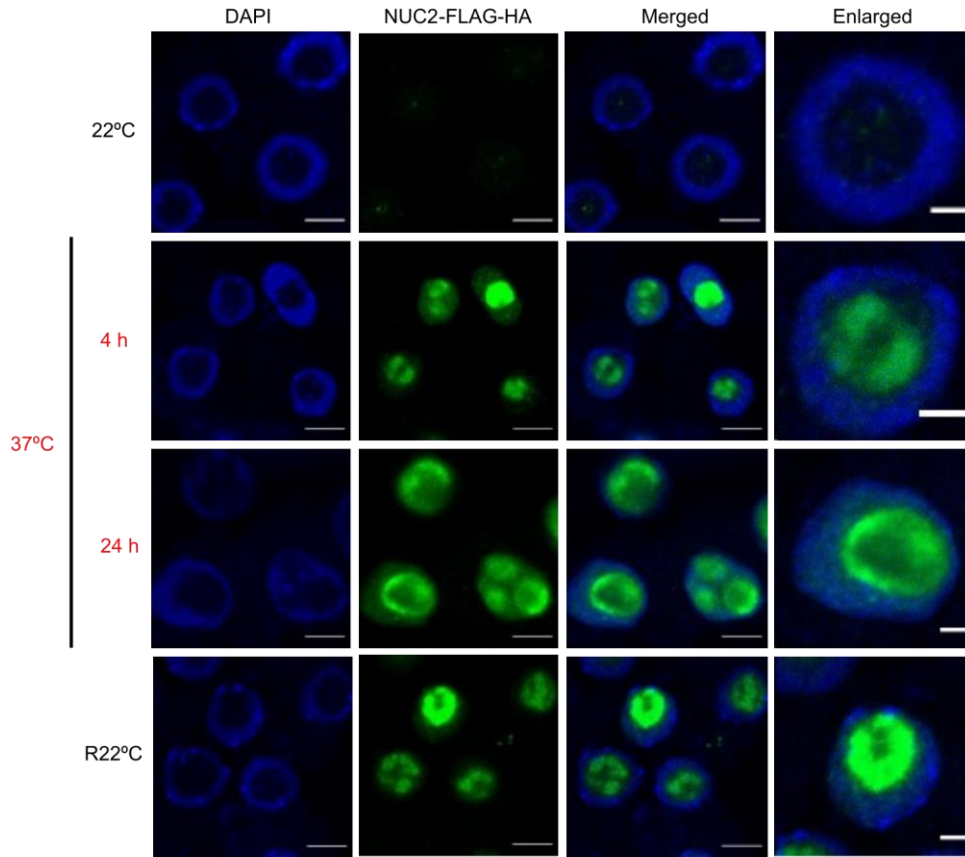


Figure 3.9. Nucleolar distribution of LAS1-GFP under heat stress in Arabidopsis. The analyses have been carried out in 5-day-old *LAS1_{pro}:LAS1-GFP* (Col-0) seedlings. Scale bar set at 5 μ m. On the right, one nucleus at every time point has been enlarged, where the scale is 2 μ m.

3.2.3.5. Arabidopsis NUC2 is induced by heat stress.

As previously mentioned, some of the genes involved in ribosome biogenesis have several paralogues in Arabidopsis. Apart from fibrillarin, Arabidopsis possesses two nucleolin genes: *NUC1* and *NUC2*. In contrast to *NUC1*, *NUC2* (At3g18610) is not detected in wild-type Col-0 plants. The expression of this gene has been observed in *nuc1* mutants (Durut et al., 2014). The complementation line *NUC2_{pro}:NUC2-FLAG-HA* (*nuc2*) was employed to analyse the subcellular localisation of the fusion protein NUC2-FLAG-HA at the same time points. As expected, NUC2-FLAG-HA is not detected at 22°C. At 37°C 4 h, the protein starts to accumulate in the nucleolar space in 22% of the cases, forming punctuate structures. Upon exposure to 37°C 24 h, the protein is present in the nucleolar space in 76% of cells, showing in some cases ring-like structures. Surprisingly, the protein is still accumulated and located in the nucleolus at R22°C in most cells (78%; Figure 3.10A and B).

A



B

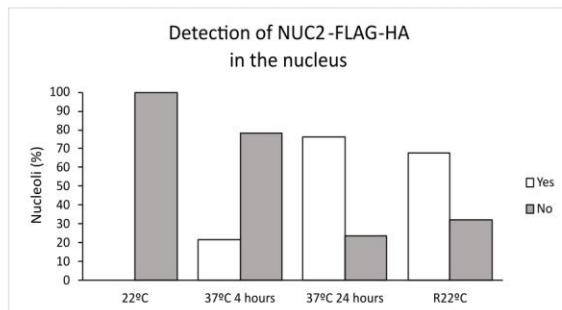


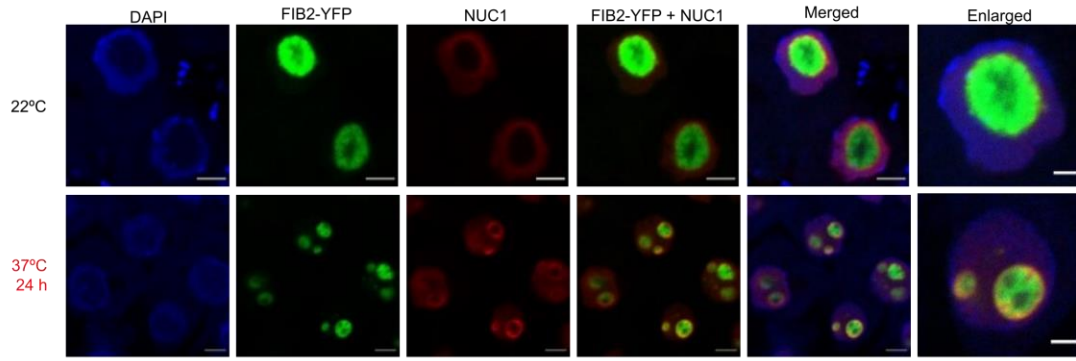
Figure 3.10. Heat stress promotes the expression and accumulation of NUC2-FLAG-HA. (A) Subcellular distribution of NUC2-FLAG-HA in roots of 5-day-old *NUC2_{pro}:NUC2-FLAG-HA (nuc2)* seedlings at 22°C, 37°C 4 h, 37°C 24 h, and R22°C. Scale bar set at 5 µm. On the right, one nucleus at every condition has been enlarged, where the scale is 2 µm. (B) Percentage of nucleoli exhibiting accumulation of NUC2-FLAG-HA at the different time points. n = 95 for 22°C, n = 138 for 37°C 4 h, n = 122 for 37°C 24 h and n = 259 for R22°C.

3.2.3.7. NUC1 surrounds FIB2-YFP and GFP-OLI2 in the nucleolus.

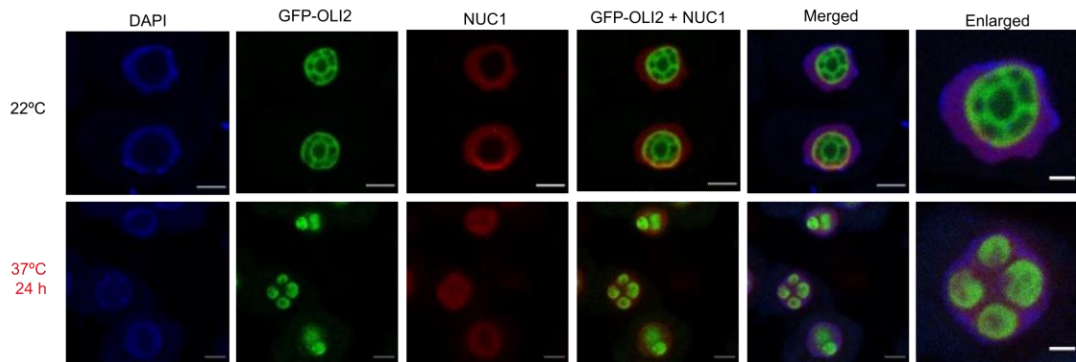
The subcellular localisation of NUC1 was analysed along with FIB2-YFP and GFP-OLI2 in *35S_{pro}:FIB2-YFP* (Col-0) and *OLI2_{pro}:GFP-OLI2 (oli2-1)* seedlings, respectively, at 22°C and 37°C 24 h. These two time points were considered critical in terms of change in the distribution pattern of the analysed nucleolar proteins. For that,

the antibody against NUC1 was employed in 5-day-old *35S_{pro}:FIB2-YFP* (Col-0) and in 5-day-old *OLI2_{pro}:GFP-OLI2* (*oli2-1*) seedlings. NUC1 is detected in the periphery of the nucleolus, surrounding FIB2-YFP and GFP-OLI2 under control conditions (22°C). FIB2-YFP is located in the DFC (Figure 3.11A), whereas GFP-OLI2 is mainly present in the DFC and GC (Figure 3.11B), as previously characterised. At 37°C 24 h, FIB2-YFP forms round structures in the nucleolar space, and NUC1 and GFP-OLI2 exhibit ring-like distribution. Interestingly, these NUC1 ring-like structures surround the FIB2-YFP round bodies and GFP-OLI2 ring-like structures under heat stress (Figure 3.11A - C).

A



B



C

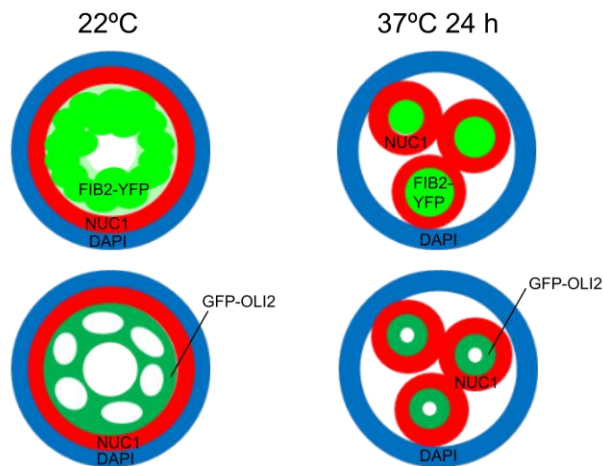


Figure 3.11. Co-localisation of NUC1, FIB2-YFP and GFP-OLI2 in $35S_{pro}:FIB2-YFP$ (Col-0) and $OLI2_{pro}:GFP-OLI2$ (*oli2-1*) seedlings. (A) Subcellular localisation of FIB2-YFP (green) and NUC1 (red) in 5-day-old $35S_{pro}:FIB2-YFP$ (Col-0) seedlings at 22°C and 37°C 24 h. DAPI is shown in blue. Scale bar set at 5 µm. On the right, one nucleus has been enlarged, where the scale is 2 µm. (B) Subcellular distribution of GFP-OLI2 (green) and NUC1 (red) in 5-day-old $OLI2_{pro}:GFP-OLI2$ (*oli2-1*) seedlings. DAPI is shown in blue. Scale bar set at 5 µm. On the right, one nucleus has been enlarged, where the scale is 2 µm. (C) Representation of the subcellular distribution of NUC1 (red), FIB2-YFP (fluorescent green) and GFP-OLI2 (dark green) in the nuclear space (DAPI marked in dark blue). At 22°C, FIB2-YFP is present in the DFC (top), whereas GFP-OLI2 is present in the DFC and GC (bottom). NUC1 is located in the nucleolar periphery. Under heat stress, NUC1 forms ring-like structures that surround FIB2-YFP round bodies and GFP-OLI2 ring-like structures.

3.2.4. Overall results

The subcellular distribution of nucleolar proteins (NRPA3^m-FLAG-HA, NUC1, FIB2-YFP, GFP-OLI2, LAS1-GFP and NUC2-FLAG-HA) and RNA molecules (C/D box snoRNAs) before, during and after heat stress reveals the formation of particular structures in the nucleolar space: foci-like particles for NRPA3^m-FLAG-HA, ring structures for NUC1, NUC2-FLAG-HA and GFP-OLI2, and round bodies for FIB2-YFP. Co-localisation experiments show how the ring-like structures from NUC1 surround the ring structures of GFP-OLI2 and the round bodies of FIB2-YFP at 37°C 24 h. It must be mentioned how NUC2-FLAG-HA is induced by exposure to high temperatures, as it is not detected at 22°C. In contrast, LAS1-GFP and the C/D box snoRNAs do not form any type of atypical structure under and/or after heat stress, as they are uniformly distributed in the nucleolus.

3.3. Addressing the nucleolar sequestration of proteins under heat stress in Arabidopsis

3.3.1. Nucleolar sequestration of proteins in Arabidopsis

3.3.1.1. Human-like NoDS in Arabidopsis

As previously explained, a post-translational mechanisms termed nucleolar sequestration of proteins has been described in human cells under certain abiotic stresses, such as acidosis or heat shock. The sequestered proteins are characterised by the presence of a sequence termed NoDS, which is responsible of the sequestration upon the corresponding stressor. To study whether the nucleolar sequestration of proteins via NoDS under heat stress is conserved in Arabidopsis, the consensus sequence of a human NoDS was searched in the Arabidopsis proteome. This human-like NoDS possessed the key features of a functional NoDS in human cells. It consisted of an arginine motif (R-R-L/I) and two hydrophobic triplets (L- Φ /N-L/V, where Φ represents any hydrophobic residue). Moreover, there were a maximum linker of ten residues between these traits: R-R-I/L-X₀₋₁₀-L- Φ /N-L/V-X₀₋₁₀-L- Φ /N-L/V. 118 proteins contained this potential NoDS in the Arabidopsis proteome (Table 3.2). To narrow down the number of candidates, proteins involved in DNA replication, transcription or RNA metabolism were selected. This filtering led to 15 proteins which possessed the human-like NoDS with the specified function (Table 3.3). Among the 15 proteins, Las1-like family protein (At5g12220; referred as LAS1) was selected (Figure 3.12A and B). As previously mentioned (3.2.3.4. LAS1-GFP stays in the nucleolus under heat stress), LAS1 is a nucleolar protein (Figure 3.12C) which forms a complex with APUM24 and GRC3 in Arabidopsis. Its function is the cleavage of the ITS2 site in the processing of the 45S pre-rRNA (Maekawa et al., 2018). Thus, the human-like NoDS of Arabidopsis LAS1 was sub-cloned, and GFP was fused to its N-terminus using the destination vector pK7WGF2 (Figure 2.2F), obtaining the construct pK7WGF-35S_{pro}:GFP-NoDS_{LAS1}. Then, Arabidopsis Col-0 plants were transformed by floral dipping, generating the 35S_{pro}:GFP-NoDS_{LAS1} line.

3.3.1.2. Transcription of Arabidopsis IGS under heat stress

The nucleolar sequestration of proteins is also characterised by the transcription of lncRNAs from the IGS: IGS₁₆RNA and IGS₂₂RNA under heat stress, and IGS₂₈RNA under acidosis. These lncRNAs interact with the NoDS of the sequestered proteins in human cells. Even though the IGS are conceived as untranscribed sequences under normal conditions in Arabidopsis, the accumulation of IGS transcripts during heat stress had been previously initiated in our team [Kawach, 2019 (Master's report)]. RT-PCR experiments were performed using different pairs of primers that target different regions of the reference IGS (Figure 3.13A). These experiments revealed an accumulation of IGS transcripts in the 3'ETS, between SP1 and SP2, upstream the GP, and downstream the GP at 37°C 24 hours (Figure 3.13B see pairs of primers 1, 2, 4 and 5, respectively). Nevertheless, those amplicons were not sequenced, and it is not known whether these transcripts come from the IGS.

Table 3.2. 118 NoDS-containing proteins (R-R-L/I-X₍₀₋₁₀₎-L- Φ -L/V-X₍₀₋₁₀₎-L- Φ -L/V) in Arabidopsis proteome

The arginine motif and the hydrophobic triplets are marked in bold and underlined, respectively.

Accession	Description	NoDS-like sequence	Accession	Description	NoDS-like sequence
AT1G22860.1	Vacuolar sorting protein 39	<u>RRIGKETLV</u> <u>LQILALL</u>	AT1G64900.1	Cytochrome P450, family 89, subfamily A, polypeptide 2	<u>LILASLSGS</u> <u>LLHLLRR</u>
AT1G68780.1	RNI-like superfamily protein	<u>LVVLENKLTG</u> <u>PLPVNLAKLTRL</u> <u>RRRL</u>	AT1G55265.1	Protein of unknown function, DUF538	<u>LALFSC</u> <u>LFSL</u> <u>SLSSSLN</u> <u>LRRL</u>
AT1G68795.1	CLAVATA3/ESR-RELATED 1	<u>VLWLSLFF</u> <u>LLHHL</u> <u>YSLNFRRL</u>	AT1G64170.1	Cation/H ⁺ exchanger 16	<u>LLLFLF</u> <u>LVGLEID</u> <u>LTSLRR</u>
AT1G71960.1	ATP-binding cassette family G25	<u>LFLMFFGYR</u> <u>VLAYL</u> <u>ARRI</u>	AT1G72590.1	3-oxo-5- α -steroid 4-dehydrogenase family protein	<u>VFL</u> <u>LLMEIH</u> <u>VLRR</u>
AT1G76270.1	O-fucosyltransferase family protein	<u>RRL</u> <u>LPL</u> <u>VIAVSL</u> <u>LLIL</u>	AT1G73080.1	PEP1 receptor 1	<u>LVL</u> <u>VVVL</u> <u>ALVFIC</u> <u>LRRL</u>
		<u>LRRL</u> <u>LPL</u> <u>VIAVSL</u> <u>LLIL</u>	AT1G66880.1	Protein kinase superfamily protein	<u>LL</u> <u>AGL</u> <u>FLC</u> <u>IRRL</u>
AT1G76250.1	Unknown protein	<u>RRISD</u> <u>VDRRL</u> <u>LLLL</u> <u>IIPSLTLL</u> <u>VL</u>	AT1G50270.1	Pentatricopeptide repeat (PPR) superfamily protein	<u>LLL</u> <u>TSPIFY</u> <u>TRRD</u> <u>FLSR</u> <u>LLRR</u>
AT1G50610.1	Leucine-rich repeat protein kinase family protein	<u>RRL</u> <u>GRLNHPN</u> <u>LPL</u> <u>VAYYYR</u> <u>REEK</u> <u>LLV</u>	AT1G05890.2	RING/U-box superfamily protein	<u>LL</u> <u>YSCNS</u> <u>NY</u> <u>VVL</u> <u>KEED</u> <u>IRRL</u>
AT1G80310.1	Sulfate transmembrane transporters	<u>RRL</u> <u>RLSSIPS</u> <u>ALIV</u> <u>FALGLVL</u>	AT2G27120.1	DNA polymerase epsilon catalytic subunit	<u>LML</u> <u>HKVMQK</u> <u>VFALL</u> <u>TDLRR</u>
AT1G33730.1	Cytochrome P450, family 76	<u>RRIC</u> <u>PGLPL</u> <u>AL</u> <u>KTVH</u> <u>LML</u>			<u>LML</u> <u>HKVMQK</u> <u>VFALL</u> <u>TDLRR</u>
AT1G33720.1	Cytochrome P450, family 76	<u>RRIC</u> <u>PGLPL</u> <u>AMKT</u> <u>VHLM</u>	AT2G40090.1	ABC2 homolog 9	<u>LRRL</u> <u>PRVILL</u> <u>ML</u>
AT1G05440.1	C-8 sterol isomerases	<u>RRL</u> <u>LIV</u> <u>VFLK</u> <u>VILVIG</u> <u>LVL</u>	AT2G23980.1	Cyclic nucleotide-gated channel 6	<u>LR</u> <u>RDIKRHL</u> <u>CL</u> <u>AL</u> <u>RRLPRVIL</u> <u>ML</u>
AT1G08260.1	DNA polymerase epsilon catalytic subunit	<u>LML</u> <u>HKVMQK</u> <u>VFALL</u> <u>TDLRR</u>	AT2G29990.1	Alternative NAD(P)H dehydrogenase 2	<u>IRRK</u> <u>LLNL</u> <u>ML</u>
		<u>LML</u> <u>HKVMQK</u> <u>VFALL</u> <u>TDLRR</u>	AT2G20710.1	Tetratricopeptide repeat (TPR)-like superfamily protein	<u>LLL</u> <u>LRVL</u> <u>PKPN</u> <u>YLRR</u>
AT1G26170.1	ARM repeat superfamily protein	<u>RRL</u> <u>QSAEIL</u> <u>AL</u> <u>KGS</u> <u>LLL</u>	AT2G16530.1	3-oxo-5- α -steroid 4-dehydrogenase family protein	<u>VFL</u> <u>LLMEIH</u> <u>VLRR</u>
AT1G26200.1	TRAM, LAG1 and CLN8 (TLC) lipid-sensing domain containing protein	<u>LFL</u> <u>RL</u> <u>L</u> <u>DRCIFER</u> <u>VARRL</u>	AT2G27260.1	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family	<u>IRRL</u> <u>FIVFTT</u> <u>FLLL</u> <u>GLIL</u>
AT1G16870.1	Mitochondrial 28S ribosomal protein S29-related	<u>LL</u> <u>VRKSYL</u> <u>AL</u> <u>RDNFRRI</u>			<u>RRL</u> <u>FIVFTT</u> <u>FLLL</u> <u>GLIL</u> <u>FIFF</u> <u>LIV</u>
AT1G12380.1	Unknown protein	<u>RR</u> <u>IEEEK</u> <u>LLVG</u> <u>QCPL</u>	AT2G44980.2	SNF2 domain-containing protein / helicase domain-containing protein	<u>LR</u> <u>RTKSL</u> <u>LIES</u> <u>GNL</u> <u>VL</u> <u>PPLTEL</u> <u>TVM</u> <u>VPL</u>
AT1G22510.2	RING/U-box protein with domain of unknown function (DUF 1232)	<u>RRL</u> <u>LRE</u> <u>MDPQRT</u> <u>LPL</u> <u>VIRAR</u> <u>VYIALIL</u>	AT2G07360.2	SH3 domain-containing protein	<u>LRRL</u> <u>PL</u> <u>DPGN</u> <u>PLFL</u>
AT2G03310.1	Unknown protein	<u>RRL</u> <u>HEEK</u> <u>LRFL</u> <u>LLL</u> <u>REKDQRIK</u> <u>ELMV</u>	AT3G29300.1	Unknown protein	<u>LVL</u> <u>VLL</u> <u>AEL</u> <u>YCS</u> <u>LLRR</u>
AT2G45550.1	Cytochrome P450, family 76, subfamily C, polypeptide 4	<u>RRIC</u> <u>PGLPL</u> <u>AVKTVS</u> <u>LML</u>	AT3G03570.1	Protein of unknown function, DUF538	<u>LR</u> <u>RL</u> <u>LL</u> <u>RFIFCRAV</u> <u>LAL</u>
AT2G45570.1	Cytochrome P450, family 76, subfamily C, polypeptide 2	<u>RRIC</u> <u>PGLPL</u> <u>AVKTVPL</u> <u>LML</u>	AT3G50140.1	Plant protein of unknown function (DUF247)	<u>IR</u> <u>RD</u> <u>ML</u> <u>ML</u> <u>ENQL</u> <u>PL</u>

AT2G45320.1	Unknown protein	<u>RRILLIPVLF</u>	AT3G50000.1	Casein kinase II, alpha chain 2	<u>LRRYLLLLCAIL</u>
AT2G30320.1	Pseudouridine synthase family protein	<u>LPLAPSEVLLRGNSFEVRR</u>	AT3G62210.1	Putative endonuclease or glycosyl hydrolase	<u>IRRHRQEKLERALPLL</u>
AT2G32360.1	Ubiquitin-like superfamily protein	<u>LPVTRLILVIGDEETRR</u>	AT3G05940.1	Protein of unknown function (DUF300)	<u>IRRCKQGCLQFVILKPIVAVTLVL</u>
AT2G20970.2	Unknown protein	<u>RRLEANGLFWTERLFLFLV</u>	AT3G25670.1	Leucine-rich repeat (LRR) family protein	<u>VILDLSKMGLRGEVPLGLTSLRR</u>
AT3G07330.1	Cellulose-synthase-like C6	<u>LFLIQSVDRLLVLGCFWIKLRR</u> <u>LFLIQSVDRLLVLGCFWIKLRR</u>	AT3G57160.1	Unknown protein	<u>LRRDILNKVILKVIL</u>
AT3G03305.1	Calcineurin-like metallo-phosphoesterase superfamily protein	<u>RRISFTILLLLL</u>	AT3G14630.1	Cytochrome P450, family 72, subfamily A, polypeptide 9	<u>VVLWCIWRILEWVWLKPKMLESYLRR</u>
AT3G59740.1	Concanavalin A-like lectin protein kinase family protein	<u>RRILAVCLTLAV</u>	AT3G63290.1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protei	<u>LRRKLLPMARKLALLDPDKRKL</u>
AT3G13380.1	BR11-like 3	<u>LILCLLVFLTVDSRGRR</u>	AT3G50190.1	Plant protein of unknown function (DUF247)	<u>IRRDMLMLENQLPL</u>
AT3G03770.1	Leucine-rich repeat protein kinase family protein	<u>LLAGALFVVLRR</u>	AT3G21470.1	Pentatricopeptide repeat (PPR-like) superfamily protein	<u>IRRRGVYFPGWVPLILRACACVVPRVVL</u>
AT3G05975.1	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family	<u>RRICCVSGIIFVLFVIFMTAL</u>	AT4G24950.1	Unknown protein	<u>IRRFVSVTMVLLSWVLVVL</u>
AT3G11960.1	Cleavage and polyadenylation specificity factor (CPSF) A subunit protein	<u>LPVNLLIATRRI</u>	AT4G29110.1	Unknown protein	<u>LMLDLNLMLKRGNKAITNLRR</u>
AT3G57140.1	Sugar-dependent 1-like	<u>LAVIKANFGIELALDECVTVLNMMRR</u>	AT4G06534.1	Unknown protein	<u>VVLSLLVVFLLRRWCCLRR</u>
AT3G08920.1	Rhodanese/Cell cycle control phosphatase superfamily protein	<u>RRLRQSYRLPVISAVSGKEL</u>	AT4G34320.1	Protein of unknown function (DUF677)	<u>LRRARDSHLLILVAL</u>
AT3G54100.1	O-fucosyltransferase family protein	<u>RRILGLMLLVV</u>	AT4G30560.1	Cyclic nucleotide gated channel 9	<u>LRRDIKRHLCLALVRRVPL</u>
AT3G19940.1	Major facilitator superfamily protein	<u>RRLLFLEGGIQMFICQLV</u>	AT4G12730.1	FASCICLIN-like arabinogalactan 2	<u>LRRATAALVLIQHLFL</u>
AT3G02260.	Auxin transport protein (BIG)	<u>LIVIRGLIVQKTKLINDCNRR</u>	AT4G01510.1	Arv1-like protein	<u>LVLAYLLLDTYRSLLLRR</u>
AT3G08870.1	Concanavalin A-like lectin protein kinase family protein	<u>LIVALSIVTLVLLVLFIFVMYKRRI</u>	AT4G38830.1	Cysteine-rich RLK (RECEPTOR-like protein kinase) 26	<u>VVLFIVLLVVFLKLRR</u>
AT3G50950.2	HOPZ-ACTIVATED RESISTANCE	<u>LPLTIKAVGGLLCKDHVYHEWRR</u>	AT4G01020.1	Helicase domain-containing protein / IBR domain-containing protein / zinc finger protein-related	<u>IRRVHLGVALLRMLAL</u>
AT3G05936.1	Unknown protein	<u>LVLKALALPRRI</u>	AT4G21740.1	Unknown protein	<u>IRRDHVMRHVFLAASIICGVSGVAL</u>
AT4G35170.1	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family	<u>RRITRFYSCLLFTLVLAFTLFCLIL</u>	AT4G32180.1	Pantothenate kinase 2	<u>VILGMLPLAREFLRR</u>
AT4G30510.1	Homolog of yeast autophagy 18 (ATG18) B	<u>RRLCLFKTTTGLPLRELNFLTSLAV</u>	AT5G22360.1	Vesicle-associated membrane protein 714	<u>LRRALWMKNAKLLVLLTCLIVFL</u>
AT4G01350.1	Cysteine/Histidine-rich C1 domain family protein	<u>LLVIEHIKTHKHPILFPRRI</u>	AT5G57940.1	Cyclic nucleotide gated channel 5	<u>LRRDIKRHLCLALVRRVPL</u>
			AT5G12220.1	Las1-like family protein	<u>IRRCLVLGALGYELVL</u>

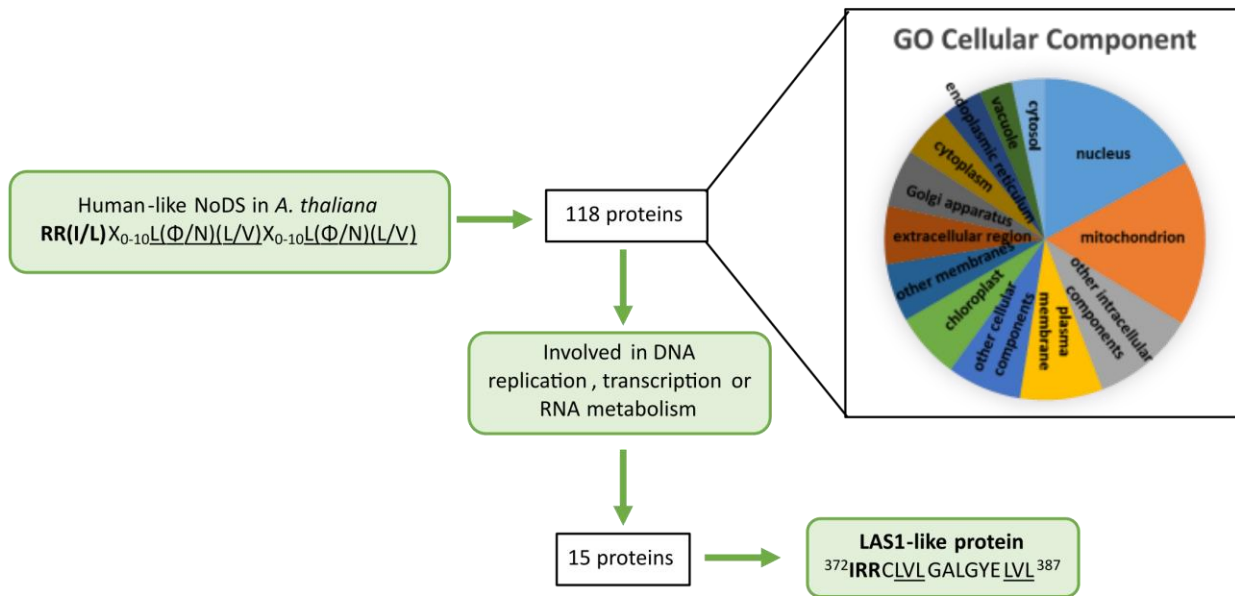
AT4G00450.1	RNA polymerase II transcription mediators	<u>RRL</u><u>LLRE</u><u>LLVE</u><u>KGYWNN</u><u>LVL</u>	AT1G07180.1	Alternative NAD(P)H dehydrogenase 1	<u>IRR</u><u>K</u><u>LL</u><u>N</u><u>M</u><u>L</u>
AT4G33210.1	F-box family protein	<u>RR</u><u>I</u><u>T</u><u>I</u><u>T</u><u>S</u><u>N</u><u>A</u><u>L</u><u>R</u><u>R</u><u>L</u><u>A</u><u>L</u><u>Q</u><u>K</u><u>Q</u><u>E</u><u>N</u><u>L</u><u>T</u><u>T</u><u>L</u><u>V</u><u>L</u>	AT5G26740.1	Protein of unknown function (DUF300)	<u>IRR</u><u>C</u><u>K</u><u>Q</u><u>G</u><u>C</u><u>L</u><u>Q</u><u>F</u><u>V</u><u>I</u><u>L</u><u>K</u><u>P</u><u>I</u><u>L</u><u>V</u><u>A</u><u>V</u><u>T</u><u>L</u><u>V</u><u>L</u>
		<u>L</u><u>RR</u><u>I</u><u>T</u><u>I</u><u>T</u><u>S</u><u>N</u><u>A</u><u>L</u><u>R</u><u>R</u><u>L</u><u>A</u><u>L</u><u>Q</u><u>K</u><u>Q</u><u>E</u><u>N</u><u>L</u><u>T</u><u>T</u><u>L</u><u>V</u><u>L</u>	AT5G02250.1	Ribonuclease II/R family protein	<u>L</u><u>RR</u><u>Q</u><u>E</u><u>K</u><u>G</u><u>K</u><u>K</u><u>Y</u><u>T</u><u>A</u><u>L</u><u>V</u><u>L</u><u>R</u><u>F</u><u>V</u><u>K</u><u>D</u><u>R</u><u>I</u><u>A</u><u>S</u><u>L</u><u>L</u><u>L</u>
AT4G33430.2	BR11-associated receptor kinase	<u>RRL</u><u>M</u><u>I</u><u>P</u><u>C</u><u>F</u><u>F</u><u>W</u><u>L</u><u>I</u><u>L</u><u>V</u><u>L</u><u>D</u><u>L</u><u>V</u><u>L</u>	AT5G04330.1	Cytochrome P450 superfamily protein	<u>V</u><u>P</u><u>L</u><u>L</u><u>L</u><u>F</u><u>L</u><u>F</u><u>P</u><u>H</u><u>L</u><u>L</u><u>L</u><u>R</u><u>R</u>
AT5G36870.1	Glucan synthase-like 9	<u>L</u><u>W</u><u>V</u><u>F</u><u>G</u><u>A</u><u>S</u><u>W</u><u>L</u><u>L</u><u>I</u><u>L</u><u>I</u><u>L</u><u>L</u><u>L</u><u>T</u><u>V</u><u>T</u><u>V</u><u>L</u><u>D</u><u>Y</u><u>A</u><u>R</u><u>R</u><u>R</u><u>L</u>	AT5G52540.1	Protein of unknown function (DUF819)	<u>V</u><u>V</u><u>L</u><u>N</u><u>F</u><u>L</u><u>L</u><u>P</u><u>L</u><u>A</u><u>V</u><u>P</u><u>L</u><u>L</u><u>L</u><u>F</u><u>R</u><u>A</u><u>D</u><u>L</u><u>R</u><u>R</u>
AT5G58480.1	O-Glycosyl hydrolases family 17 protein	<u>RRL</u><u>F</u><u>L</u><u>L</u><u>L</u><u>L</u><u>A</u><u>V</u>	AT5G20680.1	TRICHOME BIREFRINGENCE-LIKE 16	<u>L</u><u>RR</u><u>R</u><u>A</u><u>R</u><u>D</u><u>I</u><u>S</u><u>V</u><u>M</u><u>L</u><u>V</u><u>V</u><u>L</u>
AT5G51220.1	Ubiquinol-cytochrome C chaperone family protein	<u>L</u><u>L</u><u>V</u><u>L</u><u>H</u><u>M</u><u>W</u><u>L</u><u>V</u><u>L</u><u>R</u><u>R</u><u>L</u>	AT5G10370.1	Helicase domain-containing protein / IBR domain-containing protein / zinc finger protein-related	<u>IRR</u><u>V</u><u>H</u><u>L</u><u>G</u><u>V</u><u>A</u><u>L</u><u>L</u><u>R</u><u>M</u><u>L</u><u>A</u><u>L</u>
AT5G47560.1	Tonoplast dicarboxylate transporter	<u>L</u><u>V</u><u>L</u><u>G</u><u>S</u><u>F</u><u>I</u><u>L</u><u>A</u><u>L</u><u>A</u><u>V</u><u>E</u><u>H</u><u>Y</u><u>N</u><u>I</u><u>H</u><u>R</u><u>R</u><u>L</u>	AT5G49960.1	Unknown protein	<u>IRR</u><u>H</u><u>L</u><u>E</u><u>S</u><u>L</u><u>P</u><u>L</u><u>E</u><u>T</u><u>F</u><u>D</u><u>S</u><u>I</u><u>L</u><u>I</u><u>L</u>
AT5G04400.1	NAC domain containing protein 77	<u>L</u><u>P</u><u>L</u><u>C</u><u>V</u><u>L</u><u>N</u><u>K</u><u>E</u><u>A</u><u>P</u><u>L</u><u>P</u><u>L</u><u>I</u><u>Q</u><u>Y</u><u>K</u><u>R</u><u>K</u><u>R</u><u>R</u><u>I</u>	AT5G35980.1	Yeast YAK1-related gene 1	<u>L</u><u>F</u><u>L</u><u>G</u><u>L</u><u>P</u><u>L</u><u>F</u><u>P</u><u>G</u><u>G</u><u>S</u><u>E</u><u>F</u><u>D</u><u>I</u><u>L</u><u>R</u><u>R</u>
AT5G38640.1	NagB/RpiA/CoA transferase-like superfamily protein	<u>L</u><u>V</u><u>V</u><u>D</u><u>S</u><u>R</u><u>P</u><u>K</u><u>E</u><u>G</u><u>Q</u><u>L</u><u>L</u><u>L</u><u>R</u><u>R</u><u>L</u>	AT5G66440.1	Unknown protein	<u>L</u><u>R</u><u>R</u><u>I</u><u>A</u><u>A</u><u>V</u><u>I</u><u>S</u><u>T</u><u>P</u><u>L</u><u>F</u><u>V</u><u>F</u><u>L</u><u>L</u><u>G</u><u>N</u><u>S</u><u>I</u><u>V</u><u>V</u><u>L</u><u>L</u>
AT5G25820.1	Exostosin family protein	<u>RR</u><u>L</u><u>L</u><u>W</u><u>L</u><u>L</u><u>G</u><u>L</u><u>T</u><u>F</u><u>A</u><u>L</u><u>I</u><u>V</u>	AT5G64470.2	Plant protein of unknown function (DUF828)	<u>L</u><u>L</u><u>L</u><u>S</u><u>L</u><u>L</u><u>L</u><u>L</u><u>L</u><u>F</u><u>Y</u><u>S</u><u>L</u><u>I</u><u>L</u><u>R</u><u>R</u>
AT5G67380.1	Casein kinase alpha 1	<u>RRL</u><u>LL</u><u>L</u><u>L</u><u>C</u><u>A</u><u>V</u><u>L</u><u>A</u><u>L</u>	AT5G37000.1	Exostosin family protein	<u>IRR</u><u>LL</u><u>L</u><u>L</u><u>G</u><u>I</u><u>I</u><u>A</u><u>G</u><u>F</u><u>V</u><u>V</u><u>L</u><u>L</u>
		<u>L</u><u>RR</u><u>LL</u><u>L</u><u>L</u><u>C</u><u>A</u><u>V</u><u>L</u><u>A</u><u>L</u>	AT5G28463.1	Unknown protein	<u>IRR</u><u>A</u><u>G</u><u>R</u><u>P</u><u>V</u><u>F</u><u>S</u><u>W</u><u>L</u><u>F</u><u>L</u><u>V</u><u>L</u><u>L</u>
AT5G52980.1	Unknown protein	<u>L</u><u>I</u><u>L</u><u>S</u><u>A</u><u>T</u><u>E</u><u>P</u><u>L</u><u>R</u><u>S</u><u>F</u><u>L</u><u>A</u><u>L</u><u>A</u><u>S</u><u>G</u><u>D</u><u>R</u><u>R</u><u>L</u>	AT5G05690.1	Cytochrome P450 superfamily protein	<u>L</u><u>L</u><u>L</u><u>S</u><u>S</u><u>I</u><u>A</u><u>A</u><u>G</u><u>F</u><u>L</u><u>L</u><u>L</u><u>R</u><u>R</u>
AT5G15680.1	ARM repeat superfamily protein	<u>L</u><u>I</u><u>V</u><u>E</u><u>A</u><u>L</u><u>L</u><u>Q</u><u>R</u><u>F</u><u>L</u><u>P</u><u>L</u><u>A</u><u>R</u><u>R</u><u>R</u><u>I</u>	AT5G15581.1	Unknown protein	<u>L</u><u>L</u><u>L</u><u>L</u><u>L</u><u>C</u><u>L</u><u>L</u><u>N</u><u>S</u><u>G</u><u>L</u><u>R</u><u>R</u>
AT5G40240.2	Nodulin MtN21 /EamA-like transporter family protein	<u>L</u><u>L</u><u>L</u><u>L</u><u>P</u><u>L</u><u>S</u><u>V</u><u>I</u><u>F</u><u>G</u><u>R</u><u>S</u><u>R</u><u>R</u><u>L</u>	AT5G50020.2	DHHC-type zinc finger family protein	<u>L</u><u>R</u><u>R</u><u>E</u><u>L</u><u>L</u><u>P</u><u>N</u><u>N</u><u>A</u><u>G</u><u>H</u><u>V</u><u>F</u><u>L</u><u>V</u><u>A</u><u>G</u><u>V</u><u>L</u><u>F</u><u>T</u><u>V</u><u>F</u><u>V</u><u>L</u><u>I</u><u>L</u>
AT1G49730.1	Protein kinase superfamily protein	<u>V</u><u>A</u><u>L</u><u>T</u><u>M</u><u>L</u><u>V</u><u>V</u><u>L</u><u>V</u><u>I</u><u>L</u><u>I</u><u>R</u><u>R</u>	ATMG00830.1	Cytochrome C biogenesis 382	<u>L</u><u>L</u><u>L</u><u>R</u><u>S</u><u>N</u><u>R</u><u>S</u><u>P</u><u>L</u><u>M</u><u>L</u><u>L</u><u>R</u><u>R</u>
AT1G22540.1	Major facilitator superfamily protein	<u>V</u><u>A</u><u>L</u><u>V</u><u>V</u><u>L</u><u>L</u><u>G</u><u>T</u><u>C</u><u>T</u><u>Y</u><u>R</u><u>F</u><u>S</u><u>I</u><u>R</u><u>R</u>			

Table 3.3. 15 Arabidopsis proteins involved in DNA and/or RNA metabolism containing a human-like NoDS.

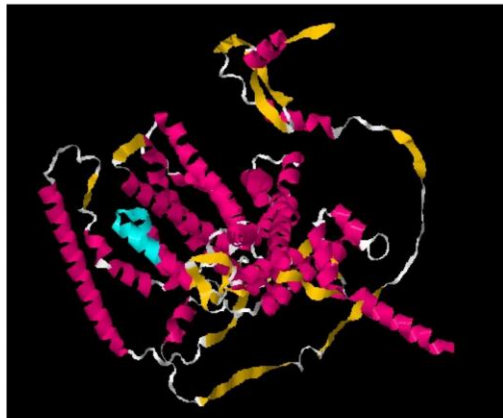
The arginine motif is shown in bold, whereas the hydrophobic triplets are underlined. The position of the NoDS in the protein has been indicated with the numbers as index. Table modified from Camuel, (2020) Master's report.

Accession	Putative function	Human-like NoDS
AT1G08260.1	DNA polymerase epsilon catalytic subunit	¹⁷⁸⁶ <u>LM</u> <u>LHKVMQKVFA</u> <u>LL</u> TDL RRL ¹⁸⁰⁶ ¹⁷⁸⁶ <u>LM</u> <u>LHKVMQKVFA</u> <u>LL</u> TDL RRL ¹⁸⁰⁵
AT1G26170.1	ARM repeat superfamily protein	⁷⁷⁶ RRL QSAEIL <u>ALKGS</u> <u>LLL</u> ⁷⁹⁰
AT1G16870.1	Mitochondrial 28S ribosomal protein S29-related	¹⁷⁴ <u>LLVRKSYL</u> <u>AL</u> RDN FRRI ¹⁹⁰
AT2G27120.1	DNA polymerase epsilon catalytic subunit	¹⁷⁶² <u>LM</u> <u>LHKVMQKVFA</u> <u>LL</u> TDL RRL ¹⁷⁸² ¹⁷⁶² <u>LM</u> <u>LHKVMQKVFA</u> <u>LL</u> TDL RRL ¹⁷⁸¹
AT2G30320.1	Pseudouridine synthase family protein	³⁹⁰ <u>LPLAPSEVL</u> <u>LILRGNSFEV</u> RRL ⁴¹⁰
AT2G32360.1	Ubiquitin-like superfamily protein	³⁶ <u>LPVTRL</u> <u>LILVIGDEET</u> RRL ⁵³
AT3G11960.1	Cleavage and polyadenylation specificity factor (CPSF) A subunit protein	⁸¹⁰ <u>LPVN</u> <u>LLLIAT</u> RRI ⁸²²
AT4G00450.1	RNA polymerase II transcription mediators	⁷⁴⁴ RRL LLRE <u>LLVEKGKYWNN</u> <u>LV</u> ⁷⁶⁴
AT5G04400.1	NAC domain containing protein 77	⁴⁸ <u>LPLC</u> <u>VLNKEAP</u> <u>LPLIQYKRK</u> RRI ⁷⁰
AT5G15680.1	ARM repeat superfamily protein	⁹ <u>LIVEALLQRFL</u> <u>LPLARR</u> RRI ²⁶
AT2G44980.2	SNF2 domain-containing protein / helicase domain-containing protein	²⁸¹ LRRTK SL LIESGN <u>LVLPPLTE</u> <u>LTVM</u> <u>VPL</u> ³⁰⁸
AT4G01020.1	Helicase / IBR domain-containing protein / zinc finger protein-related	⁶⁵⁷ IRRVHLG <u>VALLRML</u> <u>LAL</u> ⁶⁷²
AT5G02250.1	Ribonuclease II/R family protein	⁷³⁶ LRREQEKGK YTA <u>LVLRFVKDRIAS</u> <u>LLL</u> ⁷⁶²
AT5G10370.1	Helicase / IBR domain-containing protein / zinc finger protein-related	⁶⁶⁰ IRRVHLG <u>VALLRML</u> <u>LAL</u> ⁶⁷⁵
AT5G12220.1	Las1-like family protein	³⁷² IRRCL <u>VLGALGYEL</u> <u>VL</u> ³⁸⁷

A



B



C

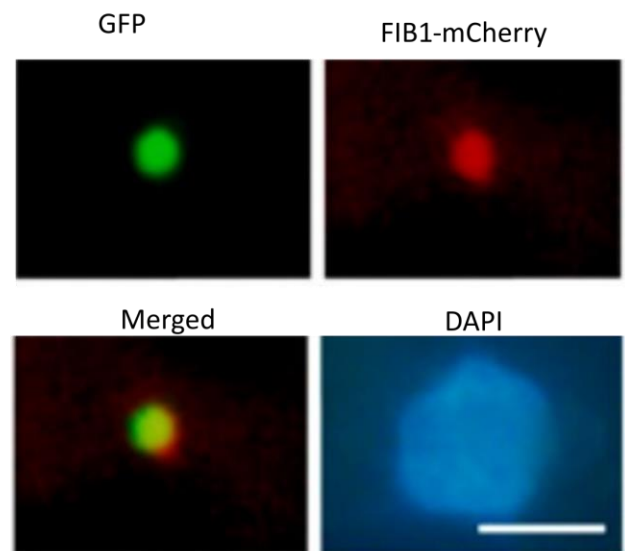
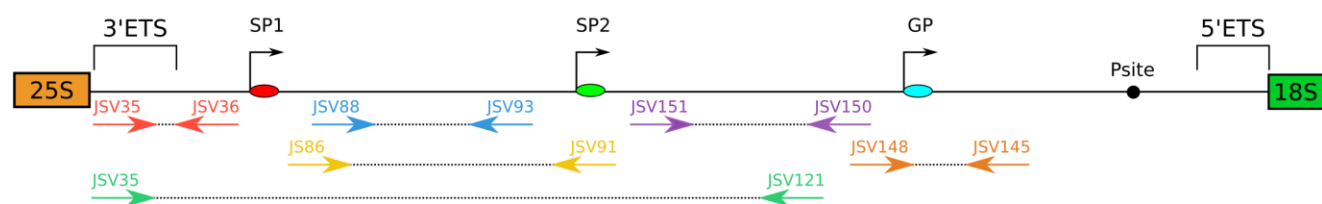


Figure 3.12. Possible NoDS-containing proteins in Arabidopsis. (A) *In-silico* strategy to find NoDS-containing proteins in Arabidopsis proteome. The consensus sequence of the human NoDS was $RR(I/L)X_{0-10}L(\Phi/N)(L/V)X_{0-10}L(\Phi/N)(L/V)$, where the arginine motif and hydrophobic triplets are marked in red and blue, respectively. Φ represents any hydrophobic residue. This search gave 118 proteins, whose subcellular localization according to the GO annotation is specified in the sector diagram on the right. Proteins participating in DNA and/or RNA metabolism were selected, resulting in 15 proteins. Among them, Arabidopsis LAS1 was chosen for experimental testing. (B) Top. Schematic representation of the sequence of Arabidopsis LAS1. The LAS1 domain was positioned after performing a BLAST with *Homo sapiens* LAS1. Bottom. Predicted 3D structure of Arabidopsis LAS1 (AlphaFold), where the alpha helices have been marked in magenta, the beta sheets are yellow, and the coiled-coil regions are represented in white. The NoDS have been coloured in cyan. (C) Nucleolar localisation of Arabidopsis LAS1 in transiently transformed *N. benthamiana* leaves. FIB1-mCherry was used as a nucleolar marker. Scale bar set at 10 μ m. This part of the figure was modified from Maekawa et al., (2018).

A



B

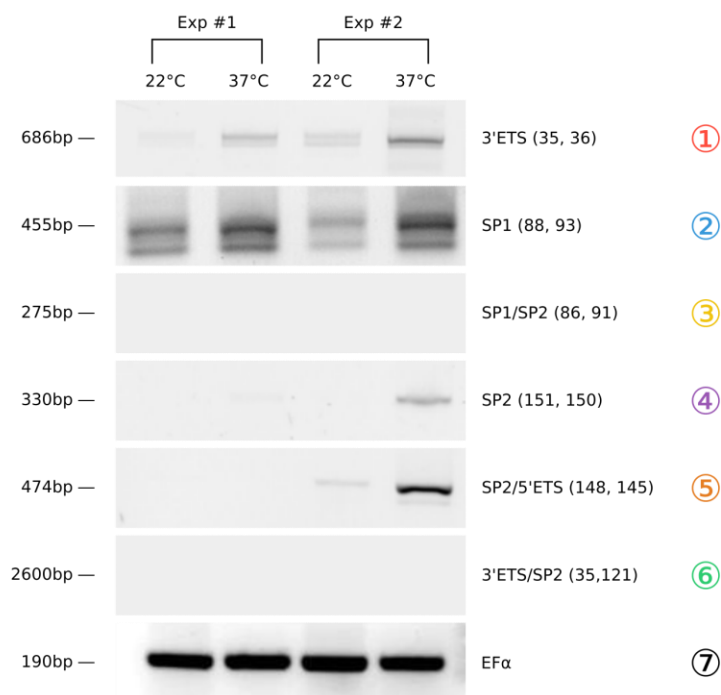


Figure 3.13. RT-PCR analyses of the accumulation of IGS transcripts under heat stress in Arabidopsis from Kawach, (2019) Master's report. (A) Schematic representation of the location of the different pair of primers on the reference IGS sequence of Arabidopsis. (B) RT-PCR products obtained from the different pair of primers. Elongation Factor 1-alpha (EF1 α) was employed as housekeeping gene.

3.3.2. Goals and experimental approach

The existence of a potential human-like NoDS has been tested *in silico* in Arabidopsis, and the NoDS from LAS1 (NoDS_{LAS1}) has been chosen as the candidate to perform the experimental validation. Moreover, the Arabidopsis line *35S_{pro}:GFP-NoDS_{LAS1}* had been obtained. Therefore, one of the goals of this section is to conduct the experimental validation of the NoDS_{LAS1} concerning its involvement in nucleolar sequestration in Arabidopsis under heat stress. To do so, the subcellular localisation of LAS1 has been examined at normal conditions and heat stress. To dive into the role of the NoDS_{LAS1}, the subcellular localisation of GFP-NoDS_{LAS1} and NoDS_{LAS1}-GFP has been analysed at 22°C and heat stress. Besides that, the retention capacity of NoDS_{LAS1} within the nucleolus has been tackled using the recombinant protein NoDS_{LAS1}-GFP-NLS. The subcellular distribution of this protein has been examined under heat stress (37°C 4 h).

To do so, the genetic fusion *LAS1-GFP* has been obtained under the control of its endogenous promoter, and the Arabidopsis line *LAS1_{pro}:LAS1-GFP* (Col-0) has been further generated. Apart from the N-terminal fusion, the C-terminal fusion NoDS_{LAS1}-GFP has been obtained, and *35S_{pro}:NoDS_{LAS1}-GFP* (Col-0) plants have been generated. In addition, the genetic fusion *NoDS_{LAS1}-GFP-NLS* has been created, transforming Arabidopsis Col-0 plants to generate the *35S_{pro}:NoDS_{LAS1}-GFP-NLS* (Col-0) line. Then, the subcellular localisation of these recombinant proteins (LAS1-GFP, NoDS_{LAS1}-GFP and NoDS_{LAS1}-GFP-NLS) as well as the already generated GFP-NoDS_{LAS1} has been analysed.

On the other hand, the accumulation of the IGS transcripts has been further analysed under heat stress. RT-PCR analyses have been performed using some of the available primers at different time points (22°C, 37°C 4 h, 37°C 24 h and R22°C). Moreover, RT-qPCR experiments have been conducted to complement the RT-PCR experiments. The subcellular localisation of the IGS transcripts has been also analysed under heat stress using RNA FISH. Finally, further analyses concerning the IGS transcripts have been performed, including a primer specific RT-PCR and the possible interaction of the IGS transcripts with RNA Pol I.

3.3.3. Results

3.3.3.1. Analysis of *LAS1_{pro}:LAS1-GFP* (Col-0) plants

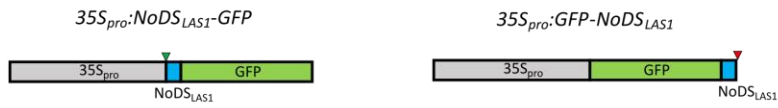
The overexpression line *LAS1_{pro}:LAS1-GFP* (Col-0) has been obtained, as detailed in 3.2.3.4. *LAS1-GFP* stays in the nucleolus under heat stress. The subcellular localisation was compared at 22°C and 37°C 4 hours, as already shown (Figure 3.9). *LAS1* is uniformly distributed in the nucleolus before, during and after heat stress.

3.3.3.2. Subcellular localisation of GFP-NoDS_{LAS1} and NoDS_{LAS1}-GFP at 22 and 37°C

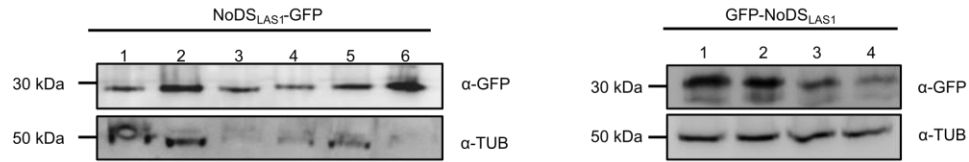
Since Arabidopsis *LAS1* is located in the nucleolus, the putative function of its human-like NoDS in nucleolar sequestration upon heat stress may become tricky to analyse. Thus, NoDS_{LAS1} was sub-cloned to generate the recombinant protein NoDS_{LAS1}-GFP, and stable *35S_{pro}:NoDS_{LAS1}-GFP* lines were subsequently obtained (Figure 3.14A). The expression of the recombinant proteins NoDS_{LAS1}-GFP and GFP-NoDS_{LAS1} was also tested in the *35S_{pro}:NoDS_{LAS1}-NoDS* and *35S_{pro}:GFP-NoDS_{LAS1}* lines, respectively. Among the six and four lines tested for *35S_{pro}:NoDS_{LAS1}-GFP* and *35S_{pro}:GFP-NoDS_{LAS1}*, respectively, line number 1 was selected for further analysis in both cases (Figure 3.14B). Similar to *35S_{pro}:GFP-NoDS_{LAS1}* (Col-0) seedlings, *35S_{pro}:NoDS_{LAS1}-GFP* (Col-0) seedlings do not seem to have a particular phenotype compared to Col-0 seedlings (results not shown).

The subcellular localization of the recombinant proteins GFP-NoDS_{LAS1} and NoDS_{LAS1}-GFP was analysed at 22°C and 37°C 4 hours. At 22°C, both GFP-NoDS_{LAS1} and NoDS_{LAS1}-GFP are distributed homogeneously in the cytoplasm and nucleus. On the other hand, both recombinant proteins concentrate in the periphery of the nucleolar space at 4 hours at 37°C. Notably, the recombinant proteins also accumulate in the cytoplasm under heat stress conditions, forming some type of cytoplasmic aggregates (Figure 3.14C). Thus, the subcellular localisation of NoDS_{LAS1}-GFP was analysed at different time points at 37°C (30 minutes, 2, 8 and 24 hours). As observed in Figure 3.14D, NoDS_{LAS1}-GFP is still present in the cytoplasm and nucleus at 30 minutes at 37°C. The change in subcellular localisation is observed from 2 hours at 37°C, where the recombinant protein starts to concentrate in the nucleolar space. However, this change in localisation is transient, as NoDS_{LAS1}-GFP is distributed uniformly in the cytoplasm and nucleus at 37°C 8 and 24 hours (Figure 3.14D). Following these results, *35S_{pro}:NoDS_{LAS1}-GFP* plants were exposed to 37°C for 4 hours followed by a recovery phase at 22°C for 24 hours. In that case, the nuclear and cytoplasmic distribution observed at 22°C is recovered (Figure 3.14E). Besides that, GFP alone does not generate these cytosolic aggregates at 37°C 4 h, meaning that NoDS_{LAS1} is the responsible of this particular localisation pattern (Figure 3.14F).

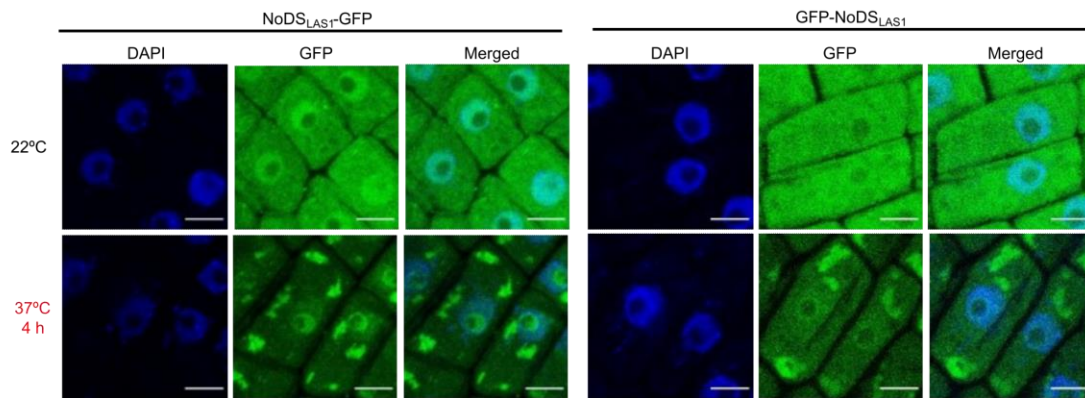
A



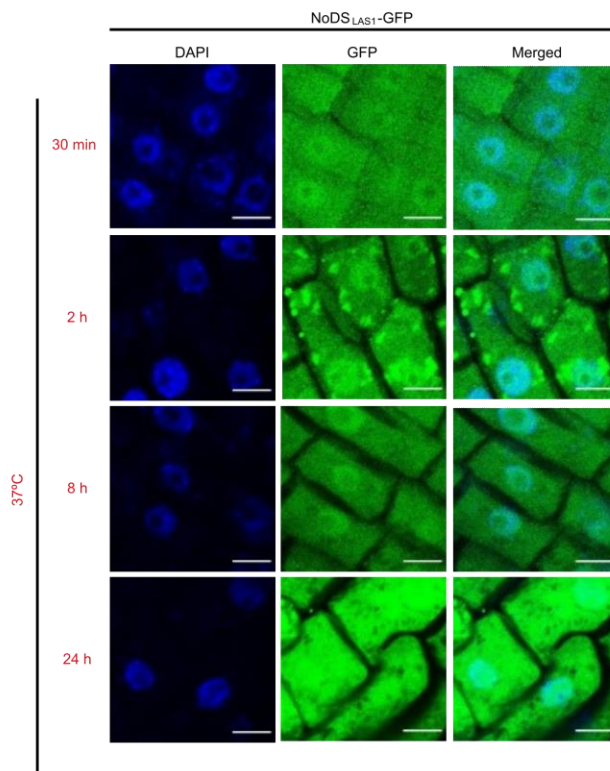
B



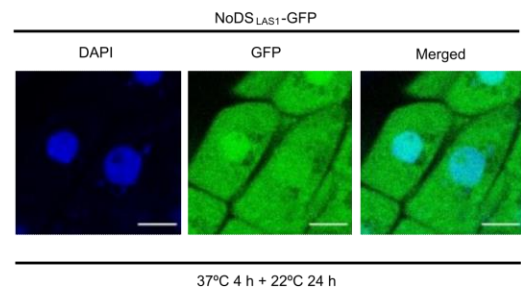
C



D



E



F

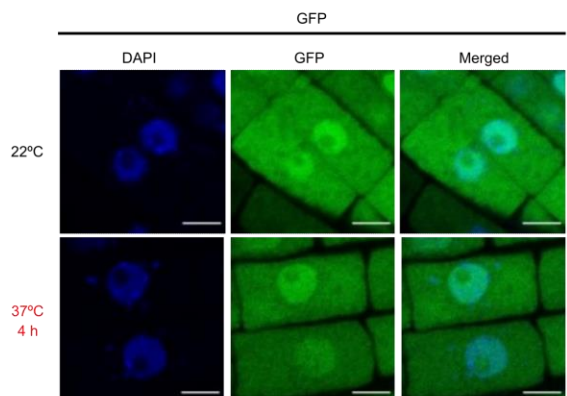


Figure 3.14. Subcellular distribution of NoDS_{LAS1}-GFP and GFP_NoDS_{LAS1}. (A) Schematic representation of the transgenes *35S_{pro}:NoDS_{LAS1}-GFP* (left) and *35S_{pro}:GFP-NoDS_{LAS1}* (right). A start codon (green triangle) was included before the NoDS_{LAS1} for the NoDS_{LAS1}-GFP fusion. In contrast, a stop codon (red triangle) was added as the last codon in the GFP-NoDS_{LAS1} fusion. (B) Immunoblot to detect the recombinant proteins NoDS_{LAS1}-GFP (left) and GFP-NoDS_{LAS1} (right) in six and four lines respectively. The expected size of NoDS_{LAS1}-GFP and GFP-NoDS_{LAS1} is 30 kDa, approx. Tubulin was employed as housekeeping control, whose size is 50 kDa, approx. (C) Subcellular localisation of the NoDS_{LAS1}-GFP (left) and GFP-NoDS_{LAS1} (right) proteins in roots of 5-day-old *35S_{pro}:NoDS_{LAS1}-GFP* and *35S_{pro}:GFP-NoDS_{LAS1}* seedlings, respectively, at 22°C and 37°C 4 h. (D) Kinetics of the change of fluorescence distribution of NoDS_{LAS1}-GFP at 37°C 30 min, 2, 8 and 24 hours. (E) Distribution of NoDS_{LAS1}-GFP at 22°C for 24 hours after 4 hours at 37°C. (F) The fluorescence pattern was analysed in 5-day-old *35S_{pro}:GFP* seedlings at 22°C and 37°C 4 hours. Scale bar set at 5 µm.

3.3.3.3. Generation and analysis of *35S_{pro}:NoDS_{LAS1}-GFP-NLS* plants

Another approach was also followed to analyse the putative role of the human-like NoDS in Arabidopsis. It was decided to include an NLS in the C-terminus of NoDS_{LAS1}-GFP. To do that, the reverse primer used in this amplification contained the sequence of a functional NLS from the SV40 T antigen (GCTCCAAAGAAGAAGAGAAAGGTC → APKKKRKV), so that the NLS was inserted downstream GFP (Figure 3.15A; 2.8.2. Generation of pCambia1300/35S-35S_{pro}:NoDS_{LAS1}-GFP-NLS by restriction enzyme digestion). Once the construct pCambia1300/35S-NoDS_{LAS1}-GFP-NLS was generated, Arabidopsis Col-0 plants were transformed by floral dipping, and *35S_{pro}:NoDS_{LAS1}-GFP-NLS* (Col-0) transformants were subsequently obtained. These seedlings do not show a particular phenotypic appearance compared to Col-0 seedlings (results not shown).

The subcellular localisation of the recombinant protein NoDS_{LAS1}-GFP-NLS was analysed at 22°C and 37°C 4 h. Whereas the recombinant protein is present in the nucleus at 22°C, it relocates in the nucleolar space at 37°C 4 h (Figure 3.15B). Moreover, NoDS_{LAS1}-GFP-NLS seems to be forming some type of bodies in the nucleolar space, which is consistent with the results shown in the previous section (3.2. Behaviour of nucleolar component to heat stress in Arabidopsis). It must be noted that no cytoplasmic aggregates are observed in these analyses. When the temperature decreases to 22°C after the stress period of 37°C for 4 hours (37°C 4 h + 22°C 24 h), the protein is present again in the nucleus (Figure 3.15C).

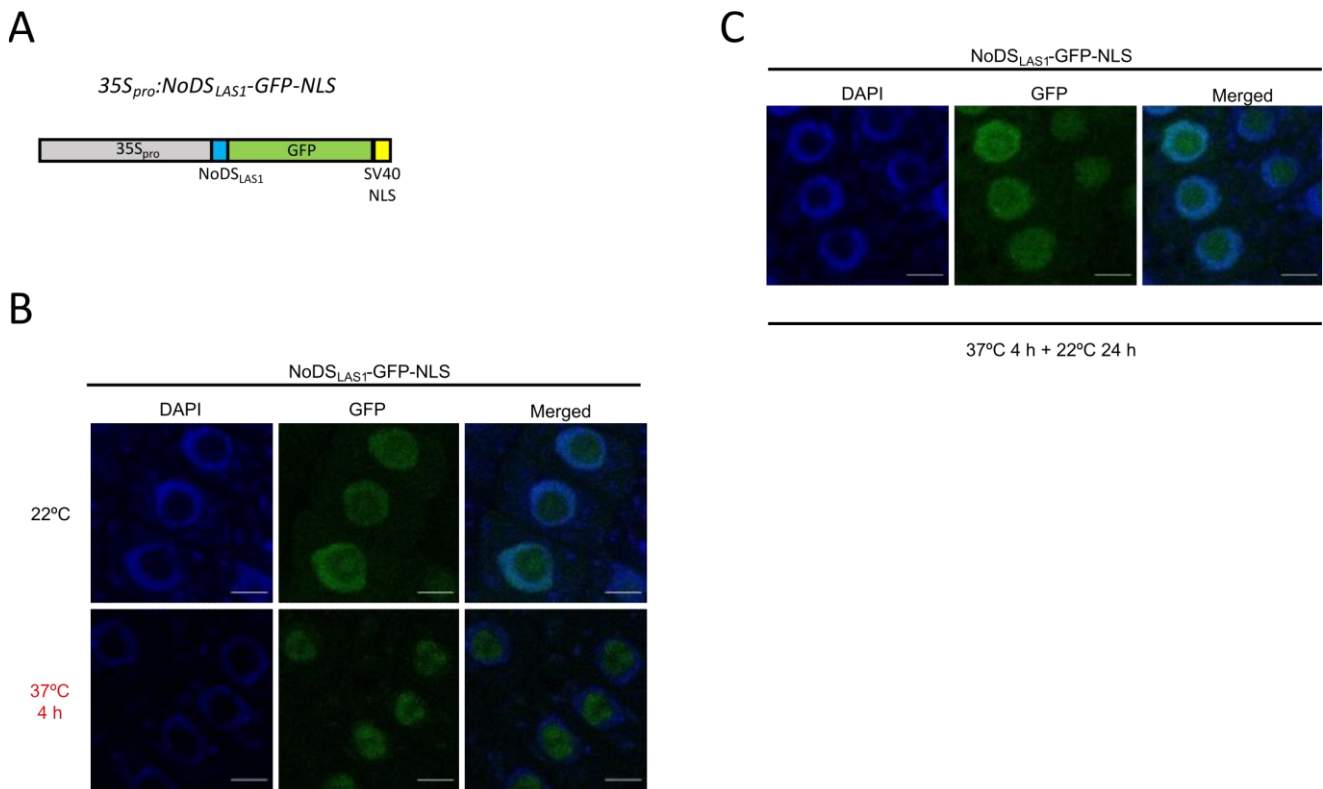


Figure 3.15. Subcellular distribution of NoDSLAS1-GFP-NLS in Arabidopsis. (A) Schematic representation of the transgene 35S_{pro}:NoDSLAS1-GFP-NLS. (B) The fusion protein is located in the nucleus of 5-day-old 35S_{pro}:NoDSLAS1-GFP-NLS (Col-0) seedlings at 22°C, whereas it is present in the nucleolar space at 37°C 4 hours. (C) The recombinant protein is re-located into the nucleus when temperatures return to 22°C. Scale bar set at 5 μm.

3.3.3.4. Arabidopsis IGS transcripts under heat stress

Based on the previous results concerning the transcription and accumulation of IGS in Arabidopsis under heat stress (Figure 3.13), two pairs of primers were employed in RT-PCR analysis: p151 + p150 and p145 + p146 (Table 2.1; Figure 3.16A). For that, cDNA from 15-day-old Col-0 seedlings exposed at 22°C, 37°C 4 and 24 hours, and R22°C were employed. IGS transcripts accumulate at 37°C, having a maximum accumulation at 24 hours. The amount of IGS transcripts decreases in the recovery phase (R22°C; Figure 3.16B). An additional pair of primers was tested to see if a longer transcript was also accumulated. For that, the primer p35 + p150 was also employed at 22°C and 37°C 4 h (Figure 3.16A). However, no accumulation was observed using this pair of primers at 37°C 4 h (Figure 3.16B). Besides that, RT-qPCR analyses were performed to complement these RT-PCR results. The primer efficiency of p146 + p145 is outside the established range of 90 – 110% while the primer efficiency using p151 + p150 is 101% (Figure 2.1; 2.7.3. Real-Time PCR (qPCR)). Thus, p151 + p150 was employed in qPCR analyses. Accordingly, the IGS transcripts accumulate during heat stress, reaching a peak after 24 hours of exposure. During the recovery, the amount of IGS transcripts decreases substantially (Figure 3.16C).

The IGS bands obtained at 37°C 4 h using the pairs of primers p151 + p150 and p146 + p145 were purified from the agarose gel. The purified products were then sequenced and aligned to the reference IGS (Figure 3.17A). The single band observed when amplifying with p146 + p145 corresponds to the expected amplicon (460 bp, approx.). Moreover, the pair p151 + p150 produced two bands at 37°C: one that correlates with the expected amplicon (332 bp approx.) and another band formed as an unspecific binding of p151 in the reference IGS (515 bp approx.). Besides that, p151 + p150 and p146 + p145 amplicons were aligned to the IGS variants described by Havlová et al., (2016). The amplicons from both pairs of primers exhibit a positive alignment with most of the “var2” IGS variants (Figure 3.17B). It must be mentioned that all “var1” IGS variants were marked as negative alignment (white circles) since the VAR1 is not transcribed in Col-0 under heat stress (“Nucleolus and rRNA genes” team unpublished data).

Finally, the subcellular localisation of the IGS transcripts was analysed at 22°C, 37°C 4 and 24 hours, and R22°C by RNA FISH (2.11.3. Observation of RNA molecules by Fluorescence *In-Situ* Hybridization (FISH)). To do that, several Cy3-tagged probes complementary to the IGS DNA sequence were designed (Table 2.1; Figure 3.18A). At 22°C, no signal is observed within the cell, which is consistent with the fact that the IGS are hardly accumulated under normal conditions. On the other hand, foci are visible in plant exposed to 37°C (4 and 24 hours) in the nucleolar space. Surprisingly, and in contrast to the RT-PCR and RT-qPCR results, these foci are still visible once the heat stress ceases (R22°C; Figure 3.18B). Two controls were performed to validate the signals observed in the RNA FISH experiments: the RNase control and the “no probe” control (Figure 3.18C).

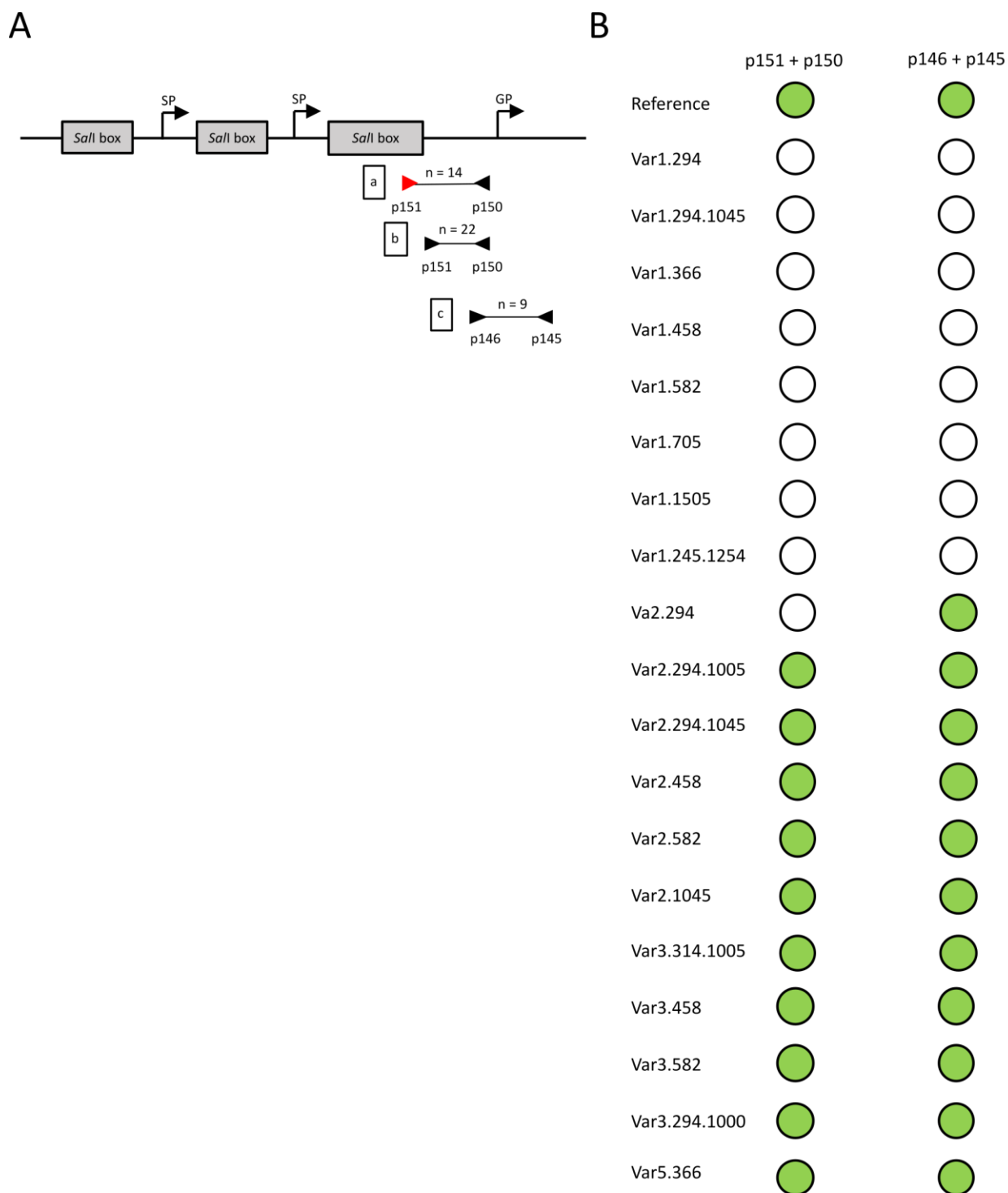


Figure 3.17. Sequence analysis of the IGS transcripts. (A) Alignment of the purified IGS amplicons at 37°C 4 h from p151 + p150 (a and b) and p146 + p145 (c) with the reference IGS. The number of clones with each amplicon are indicated (n). The red triangle portrays unspecific binding of the corresponding primer. (B) The purified amplicons from p151 + p150 and p146 + p145 at 37°C 4 h were also aligned with the different variants of the Arabidopsis IGS (var1.294, var1.294.1045, var1.366, var1.458, var1.582, var1.705, var1.1050, var1.245.1254, var2.294, var2.294.1005, var2.294.1045, var2.458, var2.582, var2.1045, var3.314.1005, var3.458, var3.582, var3.294.1000 and var5.366; Havlová et al., 2016). Green circles correspond to positive alignment, whereas white circles represent negative alignment.

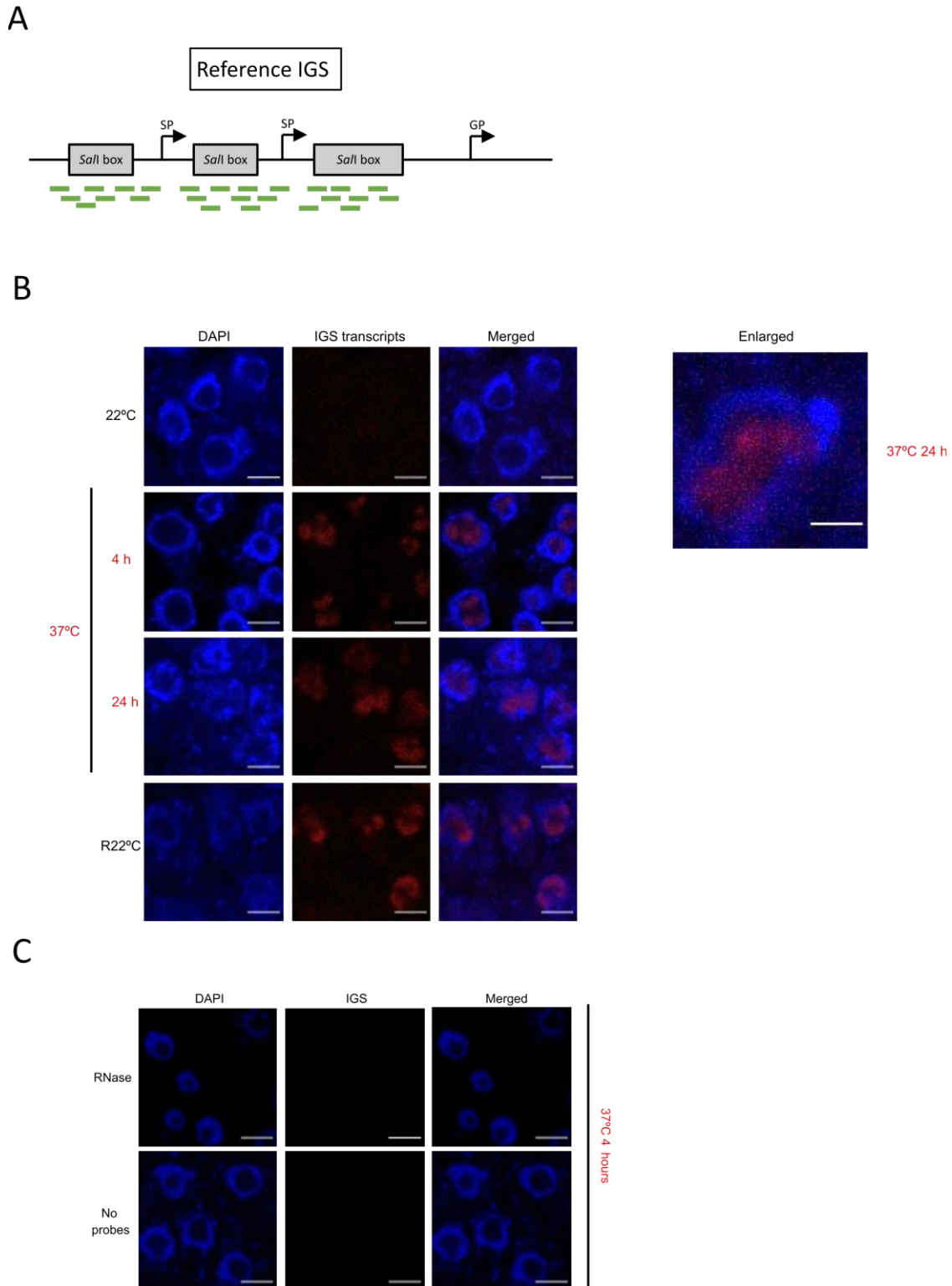


Figure 3.18. RNA FISH to detect IGS transcripts under heat stress. (A) Representation of the alignment of the 15 different Cy3-tagged probes (green bars) employed to perform the RNA FISH. The probes were aligned to the reference IGS. (B) RNA FISH portraying the subcellular localisation of the IGS transcripts in the nucleolar space of roots of 7-day-old Col-0 seedlings before, during and after heat stress. Scale bar set at 5 μ m. On the right, a nucleus at 37°C 24 h has been enlarged, where the scale is 2 μ m. (C) Controls for the RNA FISH experiments. The RNase and “no probes” controls were performed at every time point (22°C, 37°C 4 and 24 hours, and R22°C). Nevertheless, the ones conducted at 37°C 4 h are shown due to the accumulation of IGS transcripts under heat stress. Scale set at 5 μ m.

3.3.3.5. IGS-derived RNAs: sense or antisense transcripts

In humans, there are two ncRNAs that modulate rDNA transcription: the promoter-associated RNAs (pRNAs) and the promoter and pre-rRNA anti-sense (PAPAS) transcripts. The former are 150-300-nt sense transcripts generated by RNA Pol I, whereas PAPAS are longer (12-16kb) anti-sense ncRNAs transcribed by RNA Pol II. Both mediate the silencing of rDNA copies in human cells, being PAPAS the ones accumulated during abiotic stresses (reviewed by Vydzhak et al., 2020; reviewed by Pirogov et al., 2019). On the other hand, IGS₁₆RNA and IGS₂₂RNA, both transcribed under heat shock from the IGS in human cells, are sense transcripts (Audas et al., 2012a). Thus, the transcription sense of these transcripts was analysed in Arabidopsis. For that, a strand-specific RT-PCR was performed with the Arabidopsis IGS primers p151 + p150 (Figure 3.19A). Since there is amplification in both senses (Figure 3.19B and C), the amplicons were purified and sequenced. The amplicons obtained with the antisense cDNA do not correspond to IGS transcripts, as the sequence reveals unspecific binding of the primers to other transcripts. Therefore, the IGS transcripts detected under heat stress in Arabidopsis are sense transcripts.

3.3.3.6. RNA Pol I may interact with the IGS transcripts under heat stress in Arabidopsis

In human cells, the IGS-derived lncRNAs observed under abiotic stress are sense transcripts produced by RNA Pol I (Audas et al., 2012a). Whether the IGS transcripts are transcribed by RNA Pol I or not under heat stress in Arabidopsis remains uncertain. Based on previous results from our group (Sáez-Vásquez, J. personal communication) and due to the fact that both RNA Pol I and IGS are located in the nucleolus, our hypothesis supports the idea of the IGS being transcribed by RNA Pol I. FLAG immunoprecipitation experiments were conducted in *NRPA3_{pro}:NRPA3^m-FLAG-HA (nrpa3)* seedlings followed by RNA extraction and cDNA synthesis to detect the presence of IGS transcripts under heat stress (Figure 3.20A and B). Preliminary results suggest that there might be an interaction between the IGS transcripts and NRPA3^m-FLAG-HA at 37°C 4 h, as there is amplification with p151 + p150 after FLAG immunoprecipitation (Figure 3.20C). However, these experiments need to be repeated and verified.

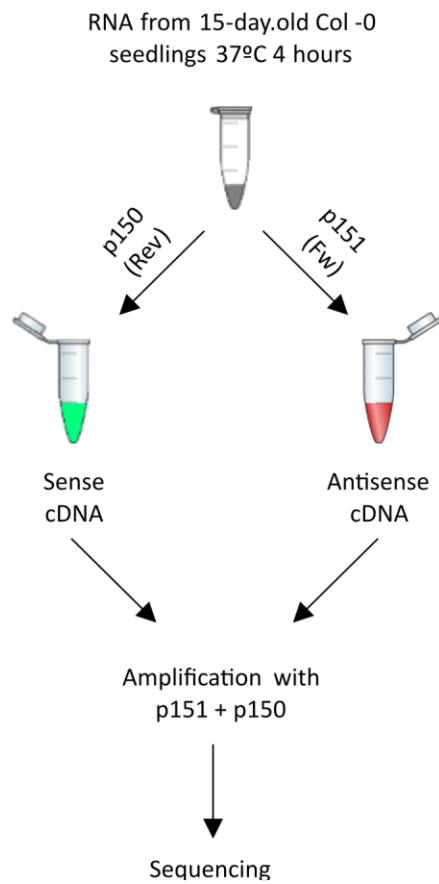
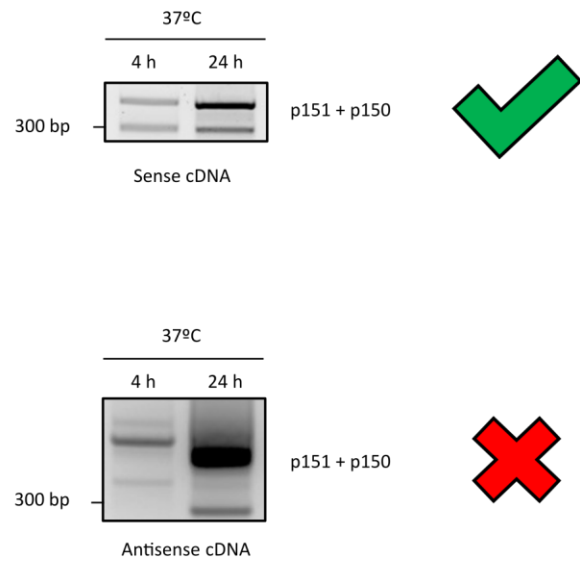
A**C**

Figure 3.19. Strand-specific RT-PCR of IGS transcripts in Arabidopsis under heat stress. (A) Strand-specific RT-PCR protocol to obtain sense and antisense cDNAs. (B) IGS transcript amplification from sense cDNA using p151 + p150. The green tick represents the validation of the amplicons after sequencing. (C) RT-PCR employing antisense cDNA and IGS primers p151 + p150 gives unspecific amplicons. The red cross corresponds to unspecific transcripts derived from the antisense cDNA that do not correspond to the IGS.

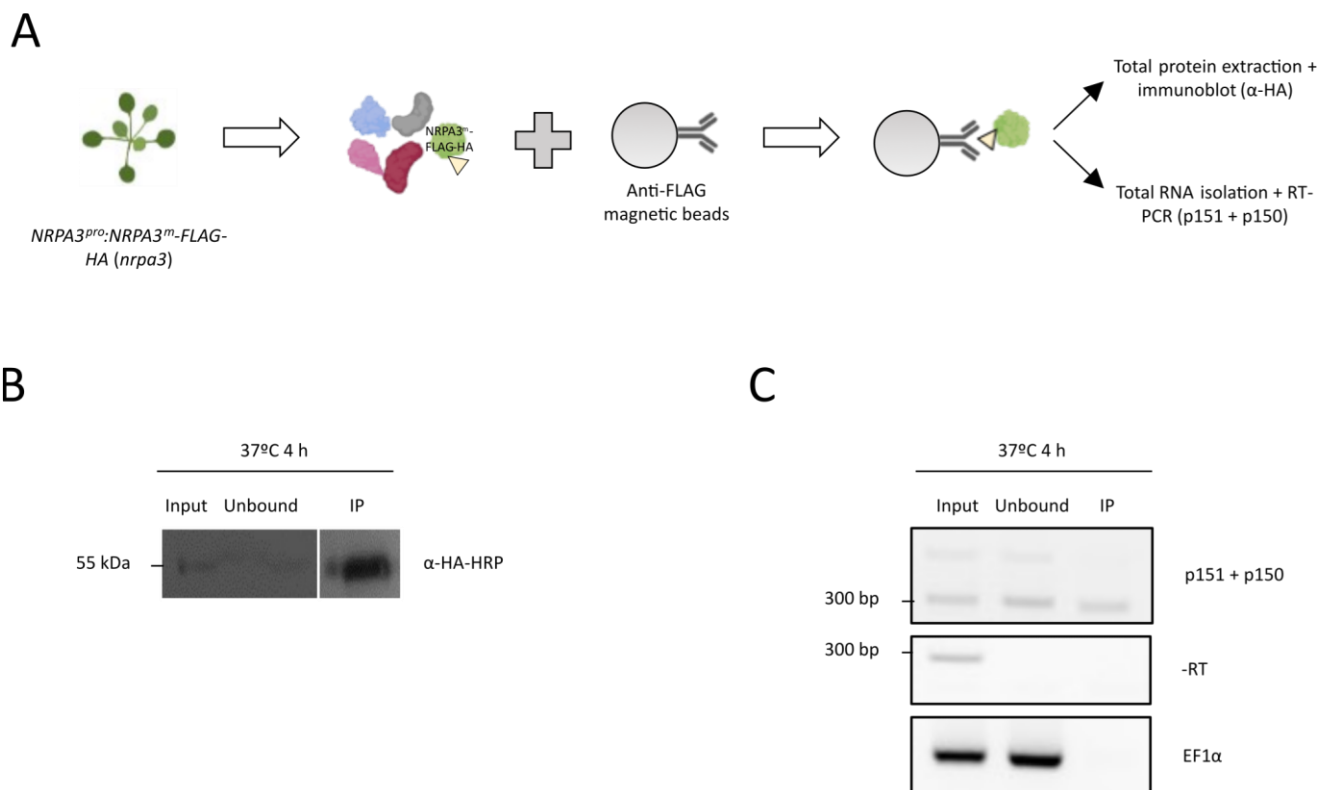


Figure 3.20. Detection of IGS transcripts after FLAG immunoprecipitation in *NRPA3_{pro}:NRPA3^m-FLAG-HA (nrpa3)* seedlings. (A) Strategy followed to identify IGS transcripts potentially associated with NRPA3^m-FLAG-HA at 37°C 4 h. Anti-FLAG magnetic beads were employed. (B) Detection of NRPA3^m-FLAG-HA after immunoprecipitation by immunoblot. The “Unbound” fraction represents the proteins which did not associate with the anti-FLAG magnetic beads. The “IP” fraction corresponds to the immunoprecipitated proteins. Expected size (NRPA3^m-FLAG-HA): ~50 kDa. (C) RT-PCR to detect IGS transcripts using p151 + p150 after immunoprecipitating NRPA3^m-FLAG-HA (top panel). The Elongation Factor 1-alpha (EF1α) was employed as a control for genomic DNA contamination (middle panel) as well as a positive control for cDNA synthesis (bottom panel).

3.3.4. Overall results

The existence of Arabidopsis proteins containing human-like NoDS had been demonstrated *in silico*. The NoDS from Arabidopsis LAS1 had been selected to perform the experimental characterisation of its putative nucleolar retention capacity. Since LAS1-GFP localizes in the nucleolus at 22°C, the role of its NoDS in protein sequestration to the nucleolus under heat stress seems tricky to analyse. Thus, the subcellular localisation of the recombinant proteins NoDS_{LAS1}-GFP and GFP-NoDS_{LAS1} was analysed under normal conditions and heat stress. Whereas both recombinant proteins are uniformly present in the cytosol and nucleus at 22°C, they re-locate to the periphery of the nucleolus at 37°C 4 h, as well as forming cytosolic aggregates. This change in localisation is transient, as the nucleolar localisation and cytosolic aggregates are lost after 4 hours at 37°C. Thus, the recombinant protein NoDS_{LAS1}-GFP-NLS was generated, so that the formation of cytoplasmic aggregates was avoided. Indeed, this recombinant protein is present in the nucleus at 22°C, shifting to the

nucleolus at 37°C 4 h. This proves the functionality of NoDS_{LAS1} in promoting nucleolar localisation under heat stress.

Moreover, the accumulation of IGS transcripts under heat stress in Arabidopsis has been determined by RT-PCR and RT-qPCR experiments. These transcripts start to accumulate from 4 to 24 hours at 37°C. However, the accumulation halts in the recovery period. RNA FISH experiments show the accumulation of these transcripts in the nucleolar space. Interestingly, these transcripts are still accumulated in the recovery phase. Strand-specific RT-PCR experiments conclude that the IGS transcripts are sense transcripts. Finally, preliminary FLAG immunoprecipitation experiments suggest a possible interaction of the IGS transcripts with the RNA Pol I subunit NRPA3^m-FLAG-HA.

4. DISCUSSION AND FUTURE PERSPECTIVES

4.1. Significant accumulation of major nucleolar proteins after heat stress in Arabidopsis

For the first section of my PhD project, a quantitative analysis of the nuclear proteome of Arabidopsis at normal conditions (22°C), heat stress (37°C 4 and 24 hours) and recovery phase (R22°C) has been presented. To my knowledge, there is a single article concerning heat stress and quantitative proteomics in Arabidopsis (Echevarría-Zomeño et al., 2016). However, the experimental conditions in the already-published article are different, defining three scenarios: the acclimation (Acc) treatment (38°C 1 h + 22°C 1 h), the committed to survival (CS) treatment (38°C 1 h + 45°C 3 h + 22°C 5 h) and the committed to death (CD) treatment (45°C 3 h + 22°C 5 h; Figure 4.1A). Our nuclear proteome at 37°C 4 h is the closest to the Acc treatment, whereas the CS treatment would be comparable to our nuclear proteome at 37°C 24 h. On the one hand, 32% of the induced or repressed proteins in the Acc treatment by Echevarría-Zomeño et al., (2016) are present as differentially accumulated proteins in the comparison 37°C 4 h vs 22°C of our nuclear proteome (Figure 4.1B). On the other hand, 29%, approx., of the induced and repressed proteins in the CS treatment are detected as differentially accumulated proteins in the comparison 37°C 24 h vs 22°C of our nuclear proteome (Figure 4.1C). It must be mentioned that the list of proteins provided in Echevarría-Zomeño et al., (2016) work is not the whole proteome detected in the three scenarios. It only contains a set of proteins involved in the heat stress response in Arabidopsis. On the other hand, proteomic analysis under heat stress have been performed in other plant species (Keller et al., 2018; Zhao et al., 2018; Jiang et al., 2020).

Besides that, the nuclear proteome of Arabidopsis under normal conditions has been already described in the literature. Two of these proteomes have been taken into account for comparative analysis: Goto et al., (2019) and Palm et al., (2016) (Figure 4.2A). Around 37% of the proteins detected in our proteomic analysis at 22°C has been detected in Palm et al., (2016) analysis. On the other hand, only 3% of our nuclear proteome at 22°C has been found in Goto et al., (2019) analysis. Both Goto et al., (2019) and Palm et al., (2016) proteomes have been obtained from Arabidopsis nuclei isolated from cultured cells, in contrast to our proteome, which has been obtained from nuclei isolated from Arabidopsis seedlings. Thus, the disparities among these nuclear proteomes may be explained by the experimental approach.

On the other hand, the nucleolar proteome of Arabidopsis was previously characterised at 22°C and 37°C 24 h in our team [Figure 4.2B; Montacié, 2019 (PhD thesis)]. At 22°C, 66% of the nucleolar proteome has been identified in the nuclear proteome. In contrast, it is surprising the fact that the nucleolar proteome is larger than the nuclear proteome at 37°C 24 h (3,188 and 2,627 proteins in the nucleolar and nuclear proteomes, respectively). Moreover, the nucleolar proteome at 37°C 24 h is more than four times larger than the nucleolar proteome at 22°C. These inconsistencies in the results could be explained by the experimental approach followed in the characterisation of the nucleolar proteomes. Arabidopsis nucleoli were purified by FACS (fluorescence-activated cell sorting) employing the fusion protein FIB2-YFP in Arabidopsis 35S_{pro}:FIB2-YFP

overexpression lines at 22°C and 37°C 24 h. Nevertheless, and as previously mentioned, the integrity of Arabidopsis nucleoli is lost under heat stress (Darriere et al., 2022). Thus, the “nucleoli” purified at 37°C 24 h may correspond to artefacts consisting of additional proteins, as a results of the disruption of the nucleolus under heat stress.

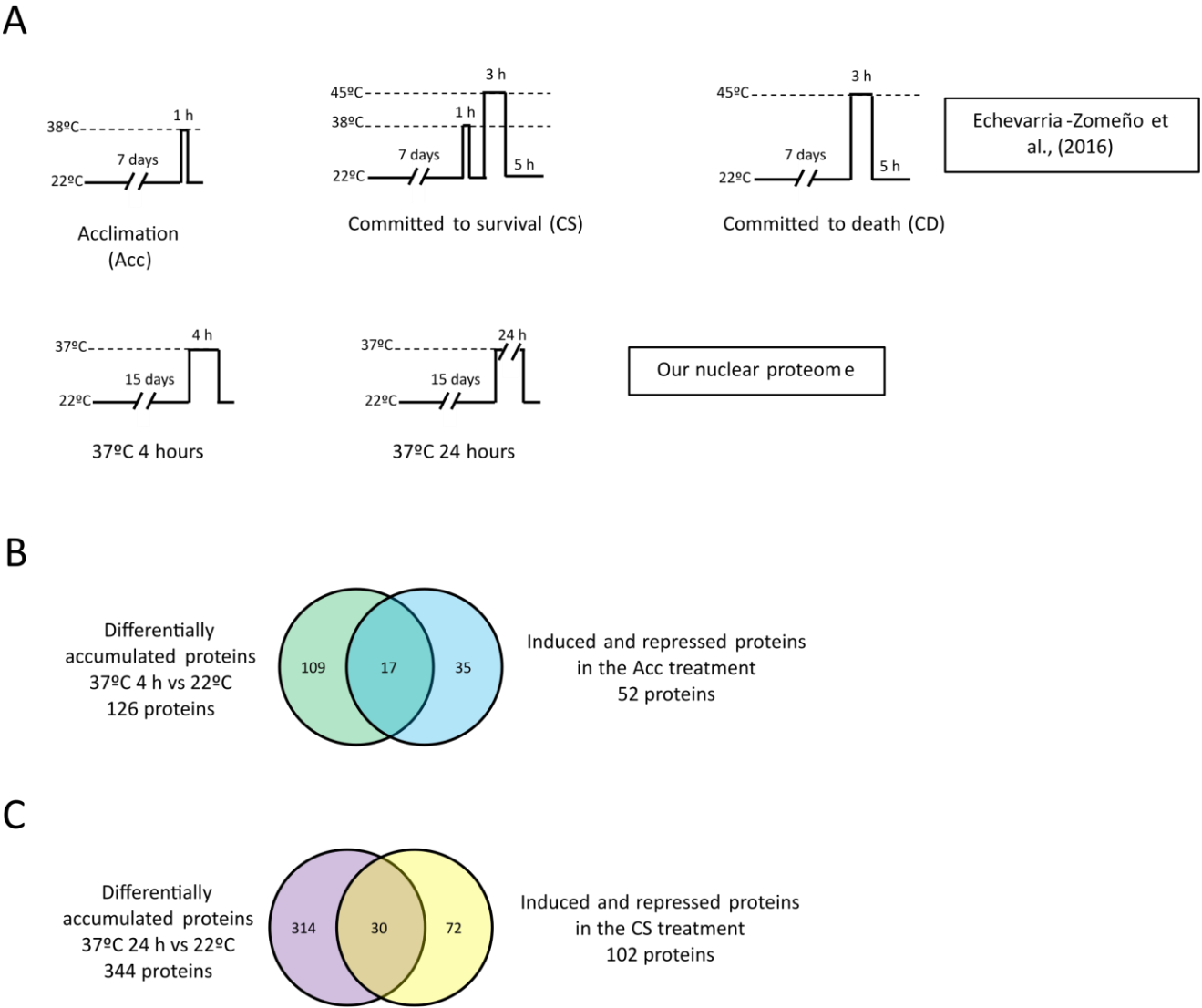


Figure 4.1. Echevarría-Zomeño et al., (2016) proteome and our nuclear proteome under high temperatures in Arabidopsis. (A) Three different scenarios were analysed in Echevarría-Zomeño et al., (2016) work: acclimation (left), committed to survival (middle) and committed to death (right). (B) Comparison of the differentially accumulated proteins at 37°C 4 h vs 22°C of our nuclear proteome (green) and the induced and repressed proteins in the Acc treatment in Echevarría-Zomeño et al., (2016) work (blue). (C) Venn diagram comparing the proteins differentially accumulated at 37°C 24 h vs 22°C in our analysis of the nuclear proteome (purple) and the induced and repressed proteins in the CS treatment in Echevarría-Zomeño et al., (2016) work (yellow).

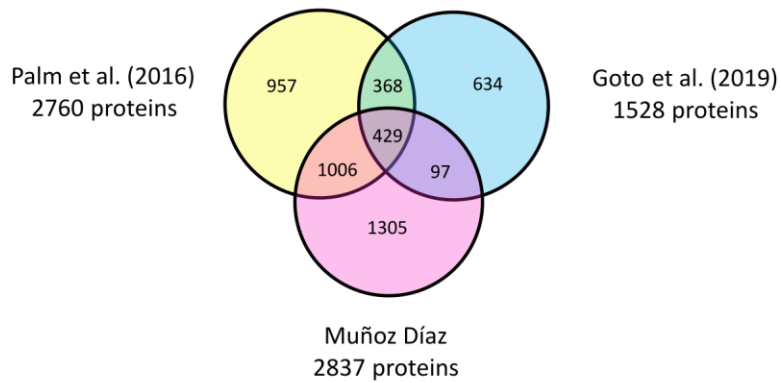
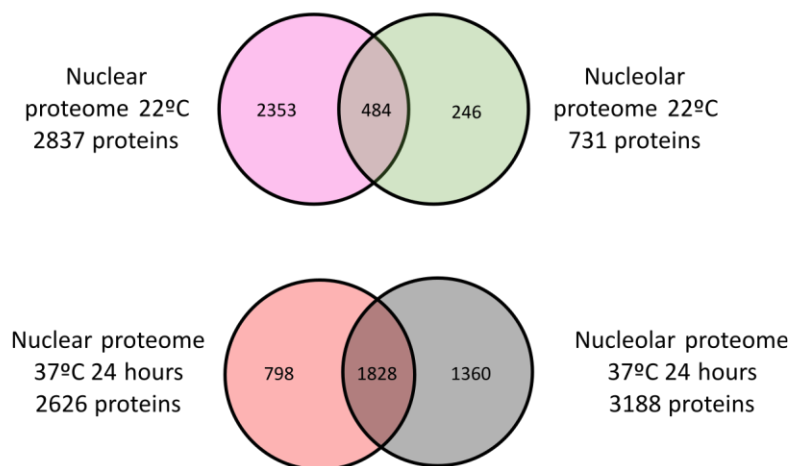
A

Figure 4.2. Arabidopsis nuclear and nucleolar proteomes. (A) The Venn diagram shows the comparison of three nuclear proteomes at 22°C in Arabidopsis: Goto et al., (2019) (cyan), Palm et al., (2016) (yellow) and our nuclear proteome (pink). **(B)** Our nuclear proteome and the nucleolar proteome [Montacié, 2019 (PhD thesis)] of Arabidopsis at 22°C (top) and 37°C 24 hours (bottom) are compared.

B

Even though heat-stressed (37°C 4 and 24 hours) and control (22°C) seedlings are phenotypically indistinguishable, the effects of heat stress were examined in the accumulation of certain proteins in the nucleus. The heat stress response is characterised by the expression of Heat HSFs and the accumulation of HSPs (reviewed by Kotak et al., 2007). In the current analysis of the nuclear proteome, several HSPs and Arabidopsis HsfA2 have been observed to increase their protein abundance in the nucleus upon heat stress. Besides that, a significant amount of chloroplastic proteins has been detected in the nuclear proteome. The absence of proteins belonging to the Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCo), one of the major chloroplastic proteins, indicates that the finding of chloroplastic proteins in the current analysis of the nuclear proteome may not be due to crossed contamination. One of these chloroplastic proteins is GENOMES UNCOUPLED 4 (GUN4; At3g59400), which differentially accumulates in the nucleus upon heat stress (member of the TG). GUN4 is involved in the chlorophyll biosynthesis pathway in Arabidopsis. In addition, this protein has

been described to potentially participate in the retrograde signalling of the chloroplast. This retrograde signalling allows a communication system between the chloroplasts, which are stress sensors, and the nucleus, so that changes in gene expression can take place. It has been hypothesised that GUN4 modulates the accumulation of magnesium-protoporphyrin IX, a molecule which is capable of leaving the chloroplast and mediate the plastid-to-nucleus signalling. Nevertheless, the exact mechanism is not fully understood (Larkin et al., 2003; reviewed by Sun and Guo, 2016; reviewed by Richter et al., 2023).

It must be mentioned that certain nuclear and/or nucleolar proteins have not been detected in our nuclear proteome. Some examples of nucleolar proteins include RTL2, LAS1 or the photobodies components phyA/B (Comella et al., 2008; Maekawa et al., 2018; reviewed by Muñoz-Díaz and Sáez-Vásquez, 2022). A possible explanation would be the scarce concentration of these proteins, so that they were not able to be detected. This is consistent with the fact that these proteins have not been found neither in Goto et al., (2019) nor Palm et al., (2016) nuclear proteomes. Even though several major nuclear proteins, such as histones, have been detected in our analysis of the nuclear proteome, others, including some nuclear envelope proteins (NUP93A, NDC1 or KAKU4), have not been detected in our nuclear proteome (Palm et al., 2016; Goto et al., 2019). The use of different filters during the MS/MS analysis as well as the statistical criteria applied during the analysis of our nuclear proteome may explain this phenomenon.

522 proteins (20% approx. of the nuclear proteome) exhibit a differential accumulation in the nucleus during and/or after heat stress. The next step was to find specific kinetics in terms of protein accumulation in the nucleus throughout the four conditions tested. Thus, these 522 proteins were initially clustered according to their protein abundance in the nucleus before, during and after heat stress. However, this initial approach was not further considered, as the clusterisation method does not take into consideration all the significant differences between the protein abundance in the nucleus (37°C 4 h vs 22°C, 37°C 24 h vs 22°C, R22°C vs 22°C, 37°C 24 h vs 37°C 4 h, R22°C 37°C 4 h and R22°C vs 37°C 24 h). Thus, these 522 proteins were classified according to these statistical differences. Accordingly, 149 proteins are included in the EG, LG, TG, EPG, LPG, and RG. That way, the members of each group possess the same significant differences in protein abundance in the nucleus (Table 4.1).

Initially, the EG was envisioned as the “expected” kinetics, as proteins differentially accumulate in the nucleus at 37°C 4 h up to 37°C 24 h. Then, protein levels decrease in the recovery phase. Similarly, the LG was also initially conceived as an “ideal” group, even though proteins significantly accumulate in the nucleus after 4 hours at 37°C. Nevertheless, it was surprising to observe that the EPG and LPG, which are characterised by a significant accumulation in the nucleus during the recovery phase compared to 22°C, are more numerous than the EG and LG. On top of that, the RG, which is composed of proteins whose nuclear accumulation significantly changes from the heat stress to the recovery phase, contains a significant number of proteins. As discussed

below, this tendency suggests that the recovery conditions established in the current analysis are not strictly recovering the normal-condition state in the nuclear proteome of Arabidopsis.

Table 4.1. Statistical criteria applied to create the EG, LG, TG, EPG, LPG and RG.

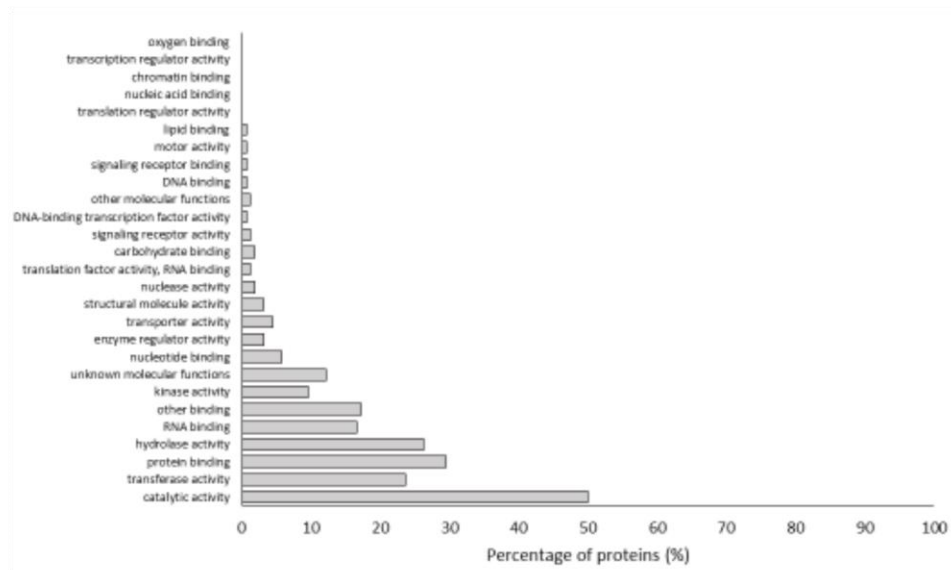
“ns” stands for not significant.

	37°C 4 h vs 22°C	37°C 24 h vs 22°C	R22°C vs 22°C	37°C 24 h vs 37°C 4 h	R22°C vs 37°C 4 h	R22°C vs 37°C 24 h
Early Group (EG)	p-value < 0.000501	p-value < 0.00132	ns	ns	p-value < 0.000692	p-value < 0.000479
Late Group (LG)	ns	p-value < 0.00132	ns	p-value < 0.000389	ns	p-value < 0.000479
Transient Group (TG)	p-value < 0.000501	ns	ns	p-value < 0.000389	p-value < 0.000692	ns
Early Persistent Group (EPG)	p-value < 0.000501	p-value < 0.00132	p-value < 0.000955	ns	ns	ns
Late Persistent Group (LPG)	ns	p-value < 0.00132	p-value < 0.000955	p-value < 0.000389	ns	ns
Recovery Group (RG)	ns	ns	p-value < 0.000955	ns	p-value < 0.000692	p-value < 0.000479

The 373 remaining proteins were clustered and grouped into the ELG, LLG, TLG and CG. The ELG, LLG and TLG were created based on similarities with the already-existing groups (EG, LG, TG, EPG, LPG, and RG). On the contrary, the CG possesses a particular kinetics where the nuclear abundance increases throughout the four different conditions. Nevertheless, and in contrast to the six initial groups, each member of these four new groups (also referred as pseudo-groups) do not have the same significant differences in protein abundance in the nucleus throughout the different conditions.

Generally talking, 366 of 522 differential accumulated proteins (70%) show an increased accumulation in the nucleus throughout and after heat stress, whereas the 156 proteins left (30%) show a decreasing kinetics in protein abundance in the nucleus. This dynamic suggests that heat stress leads to an overall movement of proteins towards the nucleus, and/or the expression of nuclear proteins during and after heat stress. Nevertheless, the functional implication of these two events is very heterogeneous, from the expression of HSFs and HSPs involved in the heat stress response, to the reestablishment of the nucleolar activity (nucleolar recovery). Analysis of the GO annotation of the 156 proteins with decreasing kinetics in protein abundance in the nucleus reveals the presence of many chloroplastic proteins (Cellular Component) and proteins with “catalytic activity” (Molecular Function; Figure 4.3). Once again, the prevalence of chloroplastic proteins suggests a bidirectional communication nucleus – chloroplast during and after heat stress.

GO MolecularFunction



GO Cellular Component

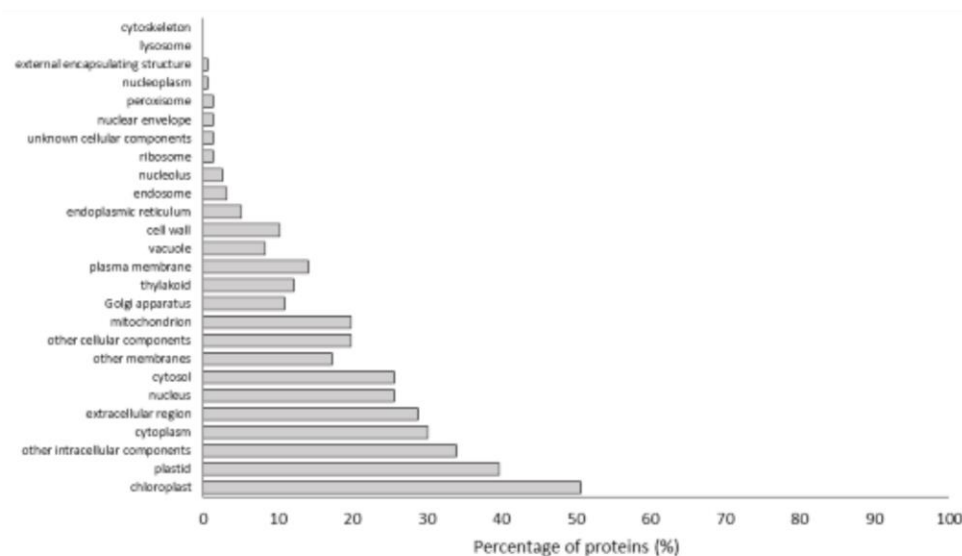


Figure 4.3. GO annotation analysis of the 156 proteins with decreasing protein abundance in the nucleus. The GO Molecular Function (top) and the GO Cellular Component (bottom) reveal a considerable amount of protein involved in “catalytic activity” and targeted to the chloroplast, respectively.

One striking finding was the considerable amount of major nucleolar proteins present in the RG and CG. These proteins include the RG members RNA Pol I subunit NRPA3 (also referred as RPAC42), OLI2, Nucleostamin-like 1 (NSN1) or BRX1 homolog 2 (BRX1-2), among others; and the CG members FIB1/2 or NUC1, among others. These two groups are characterised by a maximum of nuclear accumulation of their members in

the recovery phase. The biological explanation of this event has been termed as “nucleolar recovery”. As already mentioned, Arabidopsis nucleoli are severely affected by exposure to 37°C in terms of structure and function (Darriere et al., 2022). Since one of the functions of the nucleolus is the ribosome biogenesis, plus ribosomes are in charge of protein synthesis, the correct reassembly of the nucleolus is crucial after exposure to high temperatures. Thus, the nucleolar recovery consists of a peak in the expression of major nucleolar proteins, so that the structure and function of the nucleolus will be restored as soon as possible. Nevertheless, some of these proteins that accumulate in the recovery phase are involved in other roles other than ribosome biogenesis. For instance, the role of FIB2 in the immune response has been described in Arabidopsis (Seo et al., 2019). Another example is the involvement of NUC1 in rDNA chromatin organisation in Arabidopsis (Pontvianne et al., 2007). Thus, these proteins may be also required after heat stress to perform other roles rather than ribosome biogenesis.

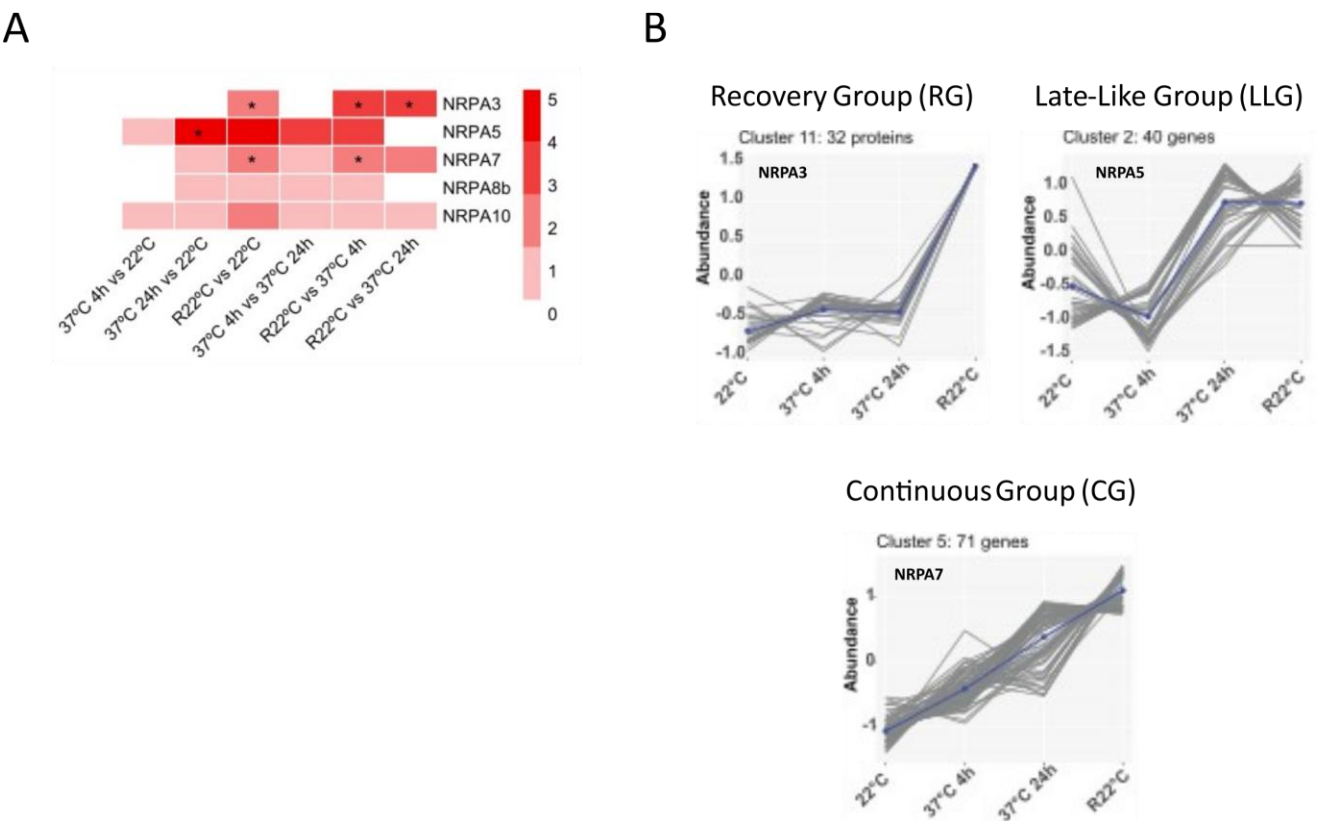
Arabidopsis RNA Pol I is a multimeric complex composed of 15 subunits (Table 1.1). Nevertheless, NRPA3 is the only RNA Pol I subunit detected in the RG. Thus, the behaviour of the rest of the RNA Pol I subunits was examined. Five RNA Pol I subunits have been detected in our nuclear proteome: NRPA3, NRPA5, NRPA7, NRPA8b and NRPA10. The significant differences in protein abundances in the six established comparisons were taken into consideration (Figure 4.4A). The nuclear abundance of NRPA8b and NRPA10 do not change significantly during and after heat stress. Regarding NRPA5, only one comparison exerts a significant difference (37°C 24 h vs 22°C), being excluded from the six initial groups (EG, LG, TG, EPG, LPG and RG). However, it has been clustered as part of the C2 of the LLG. NRPA7 could not be considered as a RG member since it must possess a significant difference in protein abundance between 37°C 24 h and R22°C (R22°C vs 37°C 24 h). Nevertheless, it has been clustered as a CG member (Figure 4.4B). NRPA3 and NRPA7 have similar behaviours in relative protein abundance in the nucleus under heat stress. It must be mentioned that these two subunits are exclusive of RNA Pol I. On the other hand, NRPA5 is also present in RNA Pol II, III and IV in Arabidopsis [named NRP(B/C/D)5 accordingly], and NRPA8b and NRPA10 are also subunits of RNA Pol II, III, IV and V in Arabidopsis [named NRP(B/C/D/E)8b and NRP(B/C/D/E)10 accordingly; Ream et al., 2015]. This may explain the different behaviour of the different RNA Pol I subunits in the nuclear proteome, as the other RNA polymerases in Arabidopsis may be regulated differently during and after heat stress. Besides that, NRPA3 alone may act in the regulation of the transcription by RNA Pol I during and after heat stress in Arabidopsis. This regulatory role exerted by RNA Pol I subunits has been described in yeast upon nutritional deprivation (Torreira et al., 2017). On the other hand, the Arabidopsis RNA polymerase V subunit NRPE3b, which is exclusively found in this RNA polymerase, has been detected in the analysis of the nuclear proteome. The nuclear abundance of this protein does not change significantly during and after heat stress (results not shown).

In Arabidopsis, many genes encoding for ribosome biogenesis factors possess several paralogues. Both Arabidopsis fibrillarin proteins (FIB1 and 2) have been detected in the nuclear proteome and classified as CG

members. Nevertheless, only NUC1 has been certainly detected in the nuclear proteome, being also a member of the CG. Even though NUC2 is not detected under normal conditions, it accumulates in the nucleus under heat stress. As the amino acid sequence of NUC1 and NUC2 shares similarity in certain regions, NUC2 has been detected as a subset. In other words, the peptide profile of NUC1 and NUC2 analysed in the MS/MS is highly similar. However, NUC1 was detected more confidently than NUC2 based on the peptide profile. The same scenario can be applied to BRX1-1 and BRX1-2. Whereas the latter is a member of the RG, the former has been considered as a subset. In Arabidopsis, OLI2 (NOP2A) possesses two other paralogues: NOP2B and NOP2C. On the one hand, NOP2C has not been found in the current nuclear proteome. On the other hand, NOP2B is not part of the RG due to its lack of a significant difference between 37°C 24 h and R22°C (R22°C vs 37°C 24 h). However, it has been grouped as a CG member, showing a similar pattern in significant differences in relative protein abundance in the nucleus as OLI2 (Figure 4.5). Nevertheless, these two paralogues have similar behaviours under heat stress.

Additionally, the Arabidopsis transcriptome under normal conditions (22°C), heat stress (37°C 2, 5 and 24 hours) and recovery conditions (5 and 24 h at 22°C after 37°C 24 h) had been obtained in our group. A comparative analysis has been made due to the similarity in the experimental conditions. As expected, the size of the transcriptome is substantially higher than the size of the nuclear proteome since the transcriptome contains the whole set of cellular mRNAs. It would be interesting to select the 2,629 loci detected in the nuclear proteome under heat stress, and analyse the transcriptome levels of those 2,629 genes, performing similarly as with the nuclear proteome. Interestingly, the comparison 37°C 4 (5) h vs 22°C exerts the highest degree of positive correlation between the transcriptome and nuclear proteome. This positive correlation in the transcriptome and nuclear proteome may correspond to the activation of the heat stress response, promoting the expression of HSPs and HSFs, among others.

Whereas the kinetics in relative protein abundance in the nucleus of the ten groups (EG, LG, TG, EPG, LPG, RG, ELG, LLG, TLG and CG) are well defined, the transcriptomic profiles of the members of these groups behave more heterogeneously. One of the reasons of this disparity in the transcriptomic and proteomic profiles is the existence of post-transcriptional modifications. These modifications are going to modulate the fate of mRNA molecules, as they regulate the translation capability or RNA structure and stability, among others (reviewed by Nachtergaele and He, 2017). Besides that, the degradation of mRNAs upon exposure to heat stress in Arabidopsis has been described (Merret et al., 2015). Moreover, changes in the nuclear proteome from one condition to another can be also due to translocation of a protein from/to the nucleus and/or changes in stability of the proteins (Figure 4.6). Thus, a disparity between the transcriptome and proteome is created.



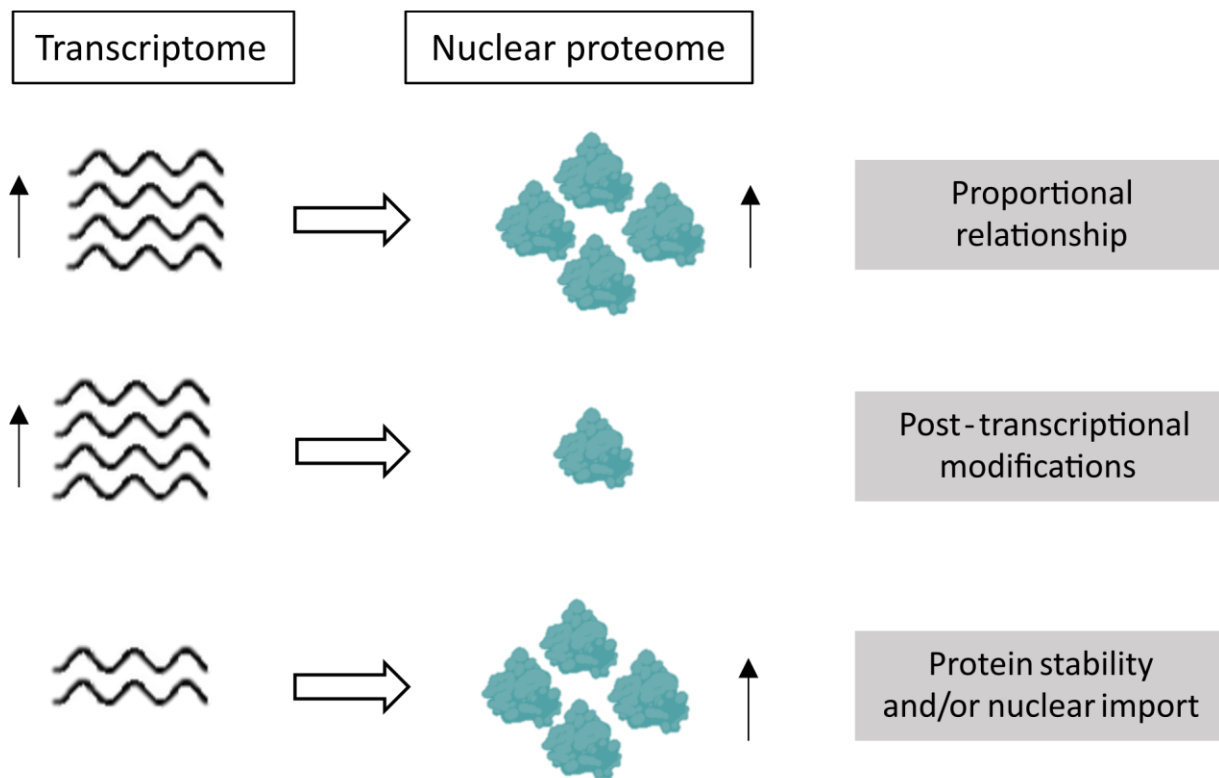


Figure 4.6. Relationship between changes in the transcriptome and nuclear proteome under heat stress. Three different scenarios are shown. Top. An increase in the transcriptome can be accompanied by an increase in the nuclear proteome. This is the case of HSFs and HSPs in the early stages of the heat stress response. Middle. An increase in the transcriptome could not be reflected in the nuclear proteome. The exposure to heat stress and the existence of post-transcriptional modification affect the fate of mRNA. Bottom. An increase in the nuclear proteome could not be a consequence of an increase in the transcriptome. This increase in protein abundance in the nucleus may come from an increase in protein stability and/or nuclear import during heat stress.

4.2. Atypical distribution patterns of nucleolar proteins during and after heat stress in Arabidopsis

The subcellular distribution of major nucleolar components has been addressed under and after heat stress. For that, the following components were chosen: the RNA Pol I subunit NRPA3, FIB2, C/D box snoRNAs), NUC1, the putative rRNA methyltransferase OLI2, the ITS2 processing factor LAS1, and heat induced NUC2. It has been previously reported the effects of prolonged heat stress in the structure and function of Arabidopsis nucleolus. Regarding the structure, the nucleolus disorganise and disassemble upon exposure to 37°C for up to 24 hours (Darriere et al., 2022). The lack of a normal tripartite nucleolus at 37°C led to speculate whether the distribution of nucleolar components (proteins and RNA species) changes during heat stress. Surprisingly, all of the analysed components are localised in the nucleus during and after heat stress. In addition, these components stay during heat stress in what it has been termed as the “nucleolar space”. The nucleolar space corresponds to the “hole” observed in the nucleus of heat-stressed cells marked with DAPI (see Figure 3.4 as an example). This term has been used instead of “nucleolus” due to the loss of the regular distribution upon heat stress.

The behaviour of the studied nucleolar components during and after heat stress can be classified according to two criteria. On the one hand, the response is either “early” or “late”. A change in the distribution pattern of the nucleolar components at 37°C 4 h in comparison to 22°C is considered as early. NRPA3^m-FLAG-HA mainly forms punctuated structures at 37°C 4 h, whereas FIB2-YFP shows round bodies in 30% of cells, approx., at the same time point. Moreover, the accumulation of NUC2-FLAG-HA is detected at 37°C 4 h in 20% of cells, approx., forming some sort of round bodies. The change in distribution of these three nucleolar components is maintained at 37°C 24 h. In contrast, a late response consists of a change in distribution at 37°C 24 h in comparison to 22°C and 37°C 4 h. NUC1 and GFP-OLI2 have a similar behaviour, as the distribution of both proteins stays practically uniform from 22°C to 37°C 4 h. However, there is a change in the distribution of these proteins at 37°C 24 h, where both form ring-like structure in most of the cases. On the other hand, the change of distribution during heat stress may disappear in the recovery phase. Thus, the response can be also termed “persistent” or “reversible”. NRPA3^m-FLAG-HA, FIB2-YFP, NUC1 and GFP-OLI2 display a reversible behaviour, as the distribution observed at 22°C is present in the recovery phase (R22°C). On the contrary, NUC2-FLAG-HA is considered to have a persistent response, as it is still expressed and present in the nucleolus at R22°C (Figure 4.7). It must be pointed out that the behaviour of NUC1, GFP-OLI2, NRPA3^m-FLAG-HA and FIB2-YFP regarding subcellular distribution under heat stress (early reversible for NRPA3^m-FLAG-HA and FIB2-YFP, and late persistent for NUC1 and GFP-OLI2) and the kinetics in relative protein abundance in the nucleus under heat stress described in “3.1.1. Quantitative proteomic profiling of Arabidopsis nuclei reveals distinct protein accumulation kinetics upon heat stress” (RG for NRPA3 and OLI2, and CG for FIB2 and NUC1) are two different

classifications. The relative abundance of these proteins in the nucleus before, during and after heat stress does not have to correlate with the formation of atypical distribution patterns at the same time points.

Analyses with NRPA3^m-FLAG-HA have been performed with a mutated version of NRPA3 where cysteins 317, 320 and 323, present in the C-terminus, have been substituted for serines (C317S, C320S and C323S, respectively). I strongly believe that these substitutions do not lead to a change in the subcellular pattern of NRPA3 during and after heat stress, since they are point mutations and do not imply the removal of any critical domain/sequence of the protein. In other words, the formation of foci-like structures in the nucleolar space during heat stress is not caused by these point mutations. However, subcellular localisation experiments with the standard NRPA3-FLAG-HA will be necessary.

The only nucleolar components whose distribution is unchanged during and after heat stress are the C/D box snoRNAs (U31a, U31b, At1gCDbox18.1, At1gCDbox33.1, At4gCDbox117.1, U3B, U3D, At5gCDbox137.1 and At5gCDbox141.1) and LAS1. These two nucleolar components are homogeneously distributed in the nucleolus at 22°C, and remain homogeneous during and after heat stress (Figure 4.7). The distribution of the U3 and At5gCDbox137.1 snoRNAs had been already characterised at 22°C (Streit et al., 2020), being homogeneously distributed in the nucleolus. We particularly chose to analyse these snoRNAs due to its nucleolar localisation. The employment of several probes for different C/D box snoRNAs makes it difficult to try to distinguish the formation of any particular distribution under heat stress. On the other hand, it is known that LAS1 is located in the nucleolus under normal conditions (Maekawa et al., 2018). Nevertheless, it does not form any type of distinctive distribution during or after heat stress. A possible explanation to this phenomenon would be the low expression of LAS1. For instance, even though it is known that LAS1 is a nucleolar protein, it was not detected neither in the current analysis of the nuclear proteome, neither in other nuclear proteomes of Arabidopsis (Palm et al., 2016; Goto et al., 2019) nor in the nucleolar proteome at 22°C [Montacié, 2019 (PhD thesis)].

During the recovery period (R22°C), the majority of the nucleoli are well-structured and functional (Darriere et al., 2022). In our group, the recovery period has been set as 24 hours at 22°C after exposure to heat stress (generally 37°C for 24 hours). The term “recovery” can be debated, as these conditions do not seem to fully mimic normal conditions. The existence of persistent responses of nucleolar components, such as NUC2-FLAG-HA, portrays the necessity of a longer exposure to 22°C after heat stress to reach the “22°C” status. This is also visible in the analysis of the nuclear proteome (3.1.1. Quantitative proteomic profiling of Arabidopsis nuclei reveals distinct protein accumulation kinetics upon heat stress), as the number of detected proteins at R22°C was not similar to 22°C. Besides that, the existence of the EPG, LPG, RG, ELG, LLG and CG shows the accumulation of certain proteins in the nucleus in the recovery phase (discussed above). Expanding the recovery phase to 36 or 48 hours at 22°C will be required to define a *bona fide* “R22°C”.

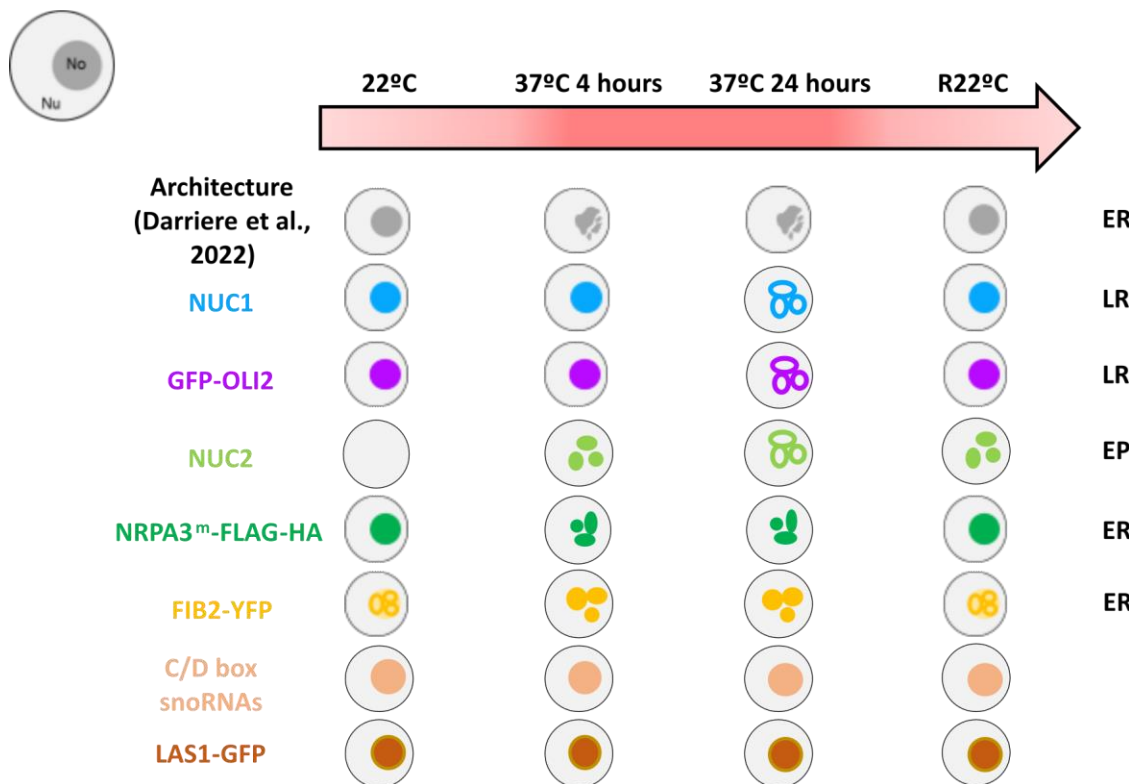


Figure 4.7. Schematic representation of the behaviour of nucleolar components under heat stress. The different distribution patterns have been depicted at the different time points. On the right side for each nucleolar component, the type of response is specified. Nu: nucleus; No: nucleolus; ER: early reversible; LR: late reversible; EP: early persistent.

Co-localisation experiments performed with FIB2-YFP and NUC1 in *35S_{pro}:FIB2-YFP* (Col-0) seedlings, and GFP-OLI2 and NUC1 in *OLI2_{pro}:GFP-OLI2* (*oli2-1*) seedlings show how the ring-like structures of NUC1 surround the FIB2-YFP and GFP-OLI2 structures at 37°C 24 h. Preliminary subcellular localisation results suggested that the NUC1 and GFP-OLI2 ring-like structures may co-localise based on their similar morphology and behaviour under heat stress. Nevertheless, NUC1 is always surrounding FIB2-YFP and GFP-OLI2 under normal conditions and heat stress. It must be mentioned that the subcellular localisation of NUC1 at 22°C does not fit with previous reports, where NUC1 and FIB2 co-localise in the nucleolus (Pontvianne et al., 2010). In fact, the peripheral localisation of NUC1 in the nucleolus overlaps with the DAPI signal in some cases. This inconsistency in the subcellular localisation of NUC1 may be attributed to an insufficient penetration of the anti-NUC1 into the nucleolus, as the same antibody had been previously employed (Pontvianne et al., 2010). Besides that, further co-localization experiments will be necessary to have a complete vision of how the different nucleolar components behave among them under heat stress. For instance, co-localisation experiments of NUC1 and NRPA3^m-FLAG-HA, or performing the crosses *NRPA3_{pro}:NRPA3^m-FLAG-HA* (*nrpa3*) x *35S_{pro}:FIB2-YFP* (Col-0) or

NRPA3_{pro}:NRPA3^m-FLAG-HA (nrpa3) x OLI2_{pro}:GFP-OLI2 (oli2-1). In addition, a suitable protocol to perform immunolocalisation and RNA FISH will be also required.

The formation of speckled structures has been documented in the nucleoli of *Arabidopsis* upon heat stress (Hayashi and Matsunaga, 2019). These speckles correspond to nascent rRNA, employing 5-ethynyl uridine, an analogue of uridine. It would be tricky to try to find a correlation between these speckles and the different structures described in this work, as the growing conditions are different. Besides that, it would be crucial to perform an RNA FISH employing probes against the 45S pre-rRNA, as these transcripts accumulate under heat stress.

It has been documented how the nucleolus is a multiphase condensate whose organisation is governed by liquid-liquid phase separation (LLPS; reviewed by Lafontaine et al., 2021). LLPS is the demix of a liquid solution into multiple phases. This immiscibility is promoted by the existence of multivalent molecules. These multivalent molecules are able to establish intra- and intermolecular interactions, promoting phase separation. Multivalent proteins are characterised by a modular architecture, the presence of intrinsically disordered domains (IDRs) and/or low complexity domains (LCDs). IDRs do not possess a defined structure, and are rich in certain residues, i.e., charged residues. On the other hand, LCDs are regions of the protein lacking amino acid heterogeneity. These features allow the establishment of weak interactions (reviewed by Banani et al., 2017; reviewed by Mittag and Parker, 2018). The presence of Gly-Arg-rich (GAR) domains and post-translational modifications has also been demonstrated to promote LLPS (reviewed by Bah and Forman-Kay, 2016; reviewed by Emenecker et al., 2020; reviewed by Guillen-Chable et al., 2021). Many of the nucleolar proteins are rich in IDRs, as well as exhibiting a modular structure (Figure 4.8). For instance, NUC1 possesses three different regions. The N-terminus is characterised by the presence of acidic residues, followed by two RRM, and a GAR domain in the C-terminus (Figure 4.8A; reviewed by Tajrishi et al., 2011). Similarly, the structure of FIB2 contains an N-terminal GAR domain, whose arginine residues are methylated (Liu and Dreyfuss, 1995), a central R region, and a C-terminal α -rich region. Both R and α -rich regions conform the conserved Ado-Met-dependent methyltransferase-like domain, which is involved in the methyltransferase activity of FIB2. Two RRM have been characterised in FIB2, being present in the R region and α -rich region (Figure 4.8B; Pih et al., 2000; Rakitina et al., 2011). NSN1 is a nucleolar protein involved in embryogenesis by modulating ribosome biogenesis, as it participates in the maturation of the pre-60S particle (Jeon et al., 2015). From the N- to the C-terminus, it contains a stretch of basic amino acids, the GTP-binding domain (GBD), an RRM, and a stretch of acid residues (Figure 4.8C; Wang). Another example is the ribosome biogenesis factor BRX1-2, which possesses a predicted Brix domain and a stretch of basic and acid residues in its N- and C-terminus, respectively (Figure 4.8D). These proteins are also predicted to have a certain degree of disordered regions. Thus, the modular architecture of these (and other) nucleolar proteins as well as the presence of IDRs promote the behaviour of the nucleolus as a molecular condensate. Multivalent nucleic acids include RNA molecules, which are also common drivers of

LLPS. Nevertheless, the role of RNA molecules in LLPS is less understood. The involvement of RNA post-transcriptional modifications in LLPS has been described, especially N⁶-methyladenosine (reviewed by (Kang and Xu, 2023)). Not only is the nucleolus separated from the nucleoplasm via LLPS, but the three nucleolar components (FC, DFC and GC) also behave as three different phases. A series of unfavourable interactions among the components of the three different nucleolar components leads to the co-existence of these three liquid phases (reviewed by Lafontaine et al., 2021).

The formation of biomolecular condensates is heavily influenced by cellular conditions, such as temperature, salt concentration or pH (Solis-Miranda et al., 2023). The effect of high temperatures has been addressed in biomolecular LLPS (Ruff et al., 2018; Cinar et al., 2019; Falahati and Haji-Akbari, 2019; reviewed by Londoño Vélez et al., 2022). Thus, it is expected that the tripartite structure of the nucleolus is sensitive to heat stress. The different types of structures formed by the different nucleolar proteins observed during heat stress, such as the ring-like structures of NUC1 or the punctuated formations of NRPA3^m-FLAG-HA, correspond to the effect of temperature in the LLPS, leading to a reorganisation of the nucleolar components and their subsequent redistribution.

Functionally, it is known that ribosome biogenesis is altered during heat stress (Darriere et al., 2022). Even though the transcription of the 45S rRNA genes is still occurring, the processing of the pre-rRNA molecules is inhibited, leading to the accumulation of 45S pre-rRNA molecules. As already mentioned, NRPA3^m-FLAG-HA forms punctuated structures under heat stress in the nucleolar space of Arabidopsis. Nevertheless, it is not known whether other subunits of RNA Pol I exert the same behaviour under heat stress. Whatever the case may be, NRPA3^m-FLAG-HA structures during heat stress may represent active sites of 45S rDNA transcription. On the other hand, the change in distribution of FIB2-YFP, NUC1 and GFP-OLI2 may have an impact on their function, inhibiting the processing of the pre-rRNA species. In this context of heat stress, NUC2-FLAG-HA is detected in the nucleolar space. It is known that NUC2 is also expressed in *nuc1* mutants, and its function is antagonistic to NUC1 under normal conditions (Durut et al., 2014). However, the role of NUC2 upon heat stress has not been entirely elucidated yet.

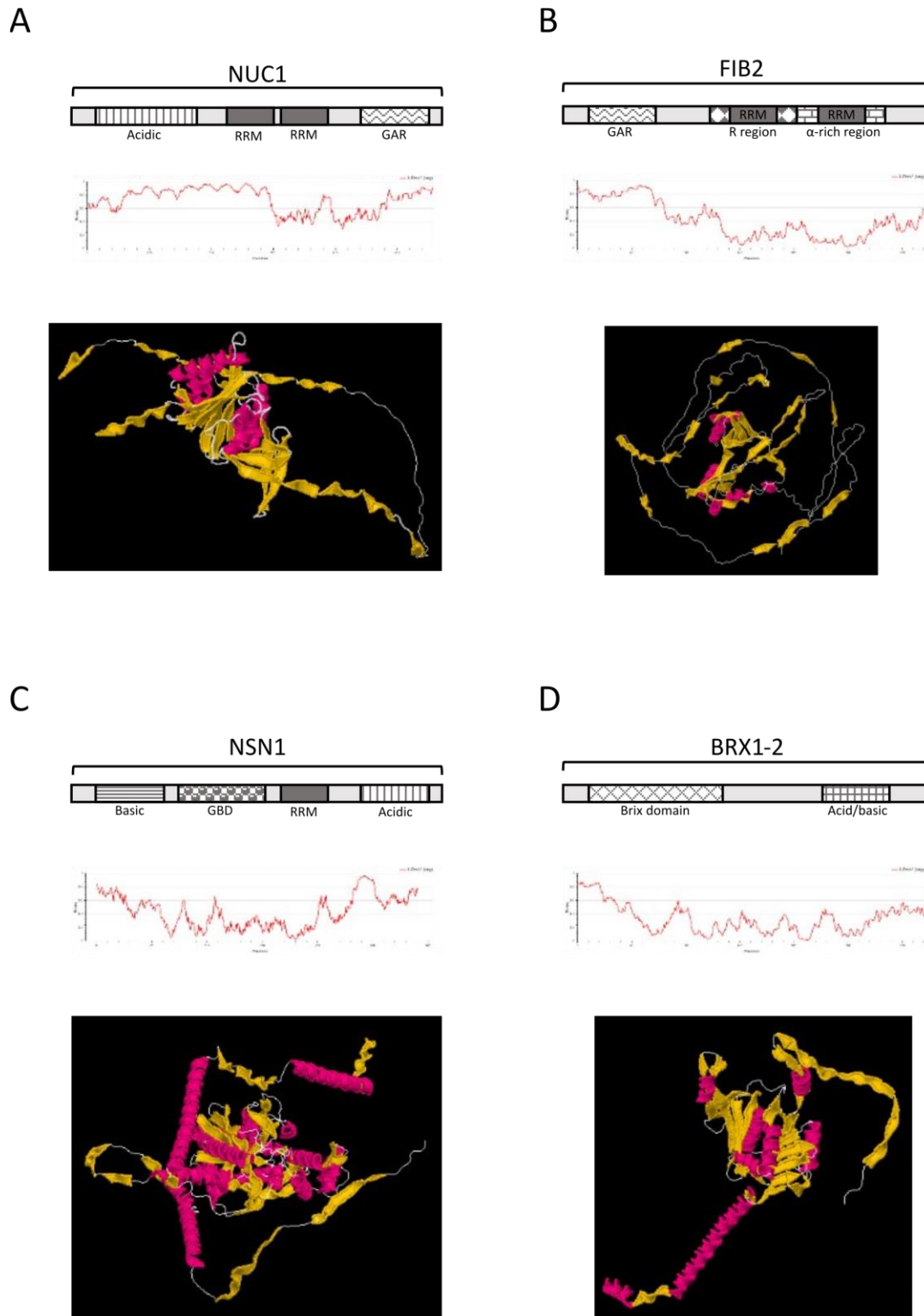


Figure 4.8. Domain organisation, IDRs analysis and 3D structure of major nucleolar proteins. The different domains (top) the IDRs prediction (middle) and the predicted 3D structure (bottom) are shown for NUC1 (A), FIB2 (B), NSN1 (C) and BRX1-2 (D) in Arabidopsis. The server IUPred2A (Mészáros et al., 2018) was employed to obtain the IDRs profiles, whereas AlphaFold was used to predict the 3D structures. The alpha helices, beta sheets and disordered regions are represented in magenta, yellow and white, respectively. RRM: RNA recognition motif; GAR: glycine-arginine-rich regions; Acid and Basic correspond to regions with high abundance of acid and basic residues, respectively; GBD: GTP binding domain.

4.3. Nucleolar targeting of NoDS_{LAS1} and accumulation of IGS transcripts under heat stress

The post-translational mechanism of nucleolar sequestration of proteins upon abiotic stresses, including high temperatures, has been described in human cells. Three different elements govern this phenomenon. Firstly, the sequestered proteins share a motif termed NoDS, which has been proven to be responsive to heat shock, acidosis, and transcriptional stress in human cells. Then, the transcription of the IGS leads to the accumulation of lncRNAs, which interact with the NoDS. Finally, the sequestered proteins are immobilised in the nucleolus forming the detention centre (Mekhail et al., 2007; Audas et al., 2012a; Lam and Trinkle-Mulcahy, 2015). However, the existence of this sequestration event in Arabidopsis has not been described in the literature.

Right off the bat, the structure of Arabidopsis nucleoli is compromised under heat stress (Darriere et al., 2022), which may affect the formation of the detention centre. Besides that, the formation of the detention centre is characterised by the movement of the markers of the different nucleolar components (FC, DFC and GC) to the periphery of the nucleolus (Jacob et al., 2013). This is not the case in Arabidopsis, as various nucleolar proteins (FIB2-YFP, NRPA3^m-FLAG-HA or GFP-OLI2, among others), do not translocate to the periphery of the nucleolus under heat stress (3.2. Behaviour of nucleolar components to heat stress in Arabidopsis).

On the one hand, the existence of human-like NoDS had been tested in Arabidopsis. As described in Camuel, (2020) Master's report, a preliminary search was conducted where the presence of arginine motifs (R-R-I/L) and hydrophobic triplets (L-Φ/N-L/V, where Φ represents any hydrophobic residue), was searched in the same protein in the Arabidopsis proteome. Nevertheless, this approach resulted in almost 4,000 proteins, with the presence of complex combinations of these features. Thus, a second attempt was performed employing the consensus sequence **RR(I/L)X₀₋₁₀L(Φ/N)(L/V)X₀₋₁₀L(Φ/N)(L/V)**, where the arginine motif and the hydrophobic triplets are shown in bold and underlined, respectively, X represents any residue and Φ references any hydrophobic residue. This second search reduced dramatically the number of candidates to 118 proteins with the potential human-like NoDS. Nevertheless, not all the proteins that re-localise to the nucleolar space under heat stress in Arabidopsis contain a human-like NoDS. For instance, HIT4, which moves to the nucleolus under heat stress (Wang et al., 2018), does not possess a human-like NoDS (Table 3.2). Only nine proteins with a putative human-like NoDS shown in Table 3.2 have been found in the nuclear proteome under heat stress (3.1.1. Quantitative proteomic profiling of Arabidopsis nuclei reveals distinct protein accumulation kinetics upon heat stress): the RING-type E3 ubiquitin transferase (At1g22510), the Ubiquinol-cytochrome C chaperone family protein (At5g51220), CASEIN KINASE ALPHA 1 (At5g67380), the Casein kinase II alpha chain 2 (At3g5000), the Rhodanese-like domain –containing protein STR10 (At3g08920), the mitochondrial ribosomal protein mS29 (At1g16870), the SH3 domain containing protein (At2g07360), the vesicle-associated membrane protein 714 (At5g22360) and the auxin transport protein BIG (At3g02260). Except the CASEIN KINASE ALPHA 1 and the

Casein kinase II alpha chain, which are annotated as nuclear proteins, and the SH3 domain containing protein, whose subcellular localisation is not fully known, the rest are proteins present outside the nucleus under normal conditions. In fact, the mitochondrial ribosomal protein mS29 appeared after the second filter, where NoDS-containing proteins involved in DNA and/or RNA metabolism were selected. Thus, it would be interesting to examine the role of the putative human-like NoDS of mS29 under heat stress, as this protein is present in the mitochondria under normal conditions. Then, proteins involved in DNA replication and/or RNA metabolism were selected from the 118 candidates, resulting in 15 proteins. Among them, LAS1 was selected for functional analysis. The selection of LAS1 was made due to (i) its role in rRNA processing (Maekawa et al., 2018) and (ii) the resemblance to previously characterised NoDS in human proteins. Thus, the fusion protein LAS1-GFP was generated as well as NoDS_{LAS1}-GFP and GFP-NoDS_{LAS1} recombinant proteins.

The fusion protein LAS1-GFP is located in the nucleolus at 22°C, during and after heat stress in *LAS1_{pro}:LAS1-GFP* (Col-0) seedlings. This is in agreement with a previous reports, where the subcellular localisation of Arabidopsis LAS1 was analysed in transiently transformed *Nicotiana benthamiana* plants using the fusion *35S_{pro}:LAS1-GFP* (Maekawa et al., 2018). As previously mentioned, LAS1-GFP stays in the nucleolar space at 37°C 4 h, without any major change in its distribution. However, the effect of NoDS_{LAS1} in targeting the nucleolus under heat stress would be hard to define, as this protein is already found in the nucleolus under normal conditions.

In contrast to 22°C, where NoDS_{LAS1}-GFP and GFP-NoDS_{LAS1} are homogeneously distributed in the cytoplasm and nucleus, these two recombinant proteins are located in the periphery of the nucleolus at 37°C 4 h. A similar tendency is observed in budding yeast (*Schizosaccharomyces pombe*), where certain nuclear and nucleolar proteins reversibly form nucleolar rings under acute heat stress (42°C; Gallardo et al., 2020). However, the existence of proteins forming nucleolar rings in plants remains uncharacterised. In addition, the formation of cytosolic aggregates has also been observed with NoDS_{LAS1}-GFP and GFP-NoDS_{LAS1} under heat stress. These cytosolic aggregates might be non-dynamic cellular bodies. In contrast to dynamic cellular bodies, such as the nucleolus, non-dynamic cellular bodies are formed of immobile proteins. These non-dynamic bodies have been well-described in vertebrates, such as the Balbiani Bodies (Wang et al., 2018; reviewed by Woodruff et al., 2018). A possible explanation to this phenomenon would be that heat stress enhances unspecific interactions of NoDS_{LAS1} in the cytosol, leading to the formation of these bodies (reviewed by Londoño Vélez et al., 2022). Besides that, those aggregates might be stress granules, as they appear in the cytosol under heat stress (reviewed by Maruri-López et al., 2021). However, whether these aggregates are composed exclusively of NoDS_{LAS1} or other proteins is not known.

This change in the localisation of NoDS_{LAS1}-GFP is transient, as the nucleolar targeting as well as the cytosolic aggregates disappears between 4 and 8 hours at 37°C. Nevertheless, the re-localisation of the

recombinant protein is “reversible”, as the normal distribution of NoDS_{LAS1}-GFP is observed upon exposure to 22°C for 24 hours after 4 hours at 37°C. Since NoDS_{LAS1}-GFP abandons the nucleolar localisation after 37°C 4 h, the idea of reversibility can be questioned. Thus, it would be necessary to perform a shorter recovery period, i.e., 37°C 2 h + 22°C 2 h or 37°C 4 h + 22°C 4 h, to fully address the reversibility of this change in subcellular localisation of NoDS_{LAS1}-GFP.

The recombinant protein NoDS_{LAS1}-GFP-NLS was generated to avoid the formation of cytosolic aggregates and to have a better understanding on the role of NoDS_{LAS1}. For that, the classical NLS from the simian virus protein SV40 (¹²⁶PKKKRKV¹³²) was chosen, as it had been already tested in Arabidopsis (Monsiard, G. personal communication). As expected, the recombinant protein NoDS_{LAS1}-GFP-NLS is located in the nucleus under normal conditions. Nevertheless, it is present in the nucleolar space under heat stress (37°C 4 h), which is consistent with the results observed with the recombinant proteins NoDS_{LAS1}-GFP and GFP-NoDS_{LAS1}. The recombinant protein NoDS_{LAS1}-GFP-NLS relocates in the nucleus once the temperature goes back to 22°C, similar to NoDS_{LAS1}-GFP.

Given these results, it can be stated that the human-like NoDS found in Arabidopsis LAS1 promotes localisation in the nucleolus under heat stress. However, it cannot be confirmed whether all the human-like NoDS found in the Arabidopsis proteome have the same role. Thus, the subcellular localisation of other Arabidopsis proteins containing a human-like NoDS in Arabidopsis must be tested in order to have a more solid understanding. In the case of Arabidopsis LAS1, the protein is already located in the nucleolus under normal conditions. Therefore, Arabidopsis proteins present in other cellular localisations, i.e., nucleoplasm or cytosol, must be selected for further analysis, so that the role of their corresponding NoDSs in nucleolar sequestration can be fully addressed. For that, nucleolar proteins must be excluded from the 118 Arabidopsis proteins containing human-like NoDS. For instance, the chloroplastic helicase DEAH12 (At5g10370) or the nuclear membrane protein atPOLLUX (At5g49960), whose predicted NoDS are shown in Table 3.3, could be used, among many others. Moreover, and as already mentioned, the mitochondrial ribosomal protein mS29 (At1g16870) possesses a human-like NoDS and it is involved in DNA and/or RNA metabolism (Table 3.2 and 3.3). Another approach would consist of the use of the recombinant NoDS (recNoDS), described by Audas et al., (2012a), in Arabidopsis. This recNoDS was synthesised based on the features of the consensus sequence of the NoDS (**RRI**HYSRLLVNQTE**LFV**, where the arginine motif is shown in bold and the hydrophobic triplets are underlined, respectively). The recombinant protein recNoDS-GFP was located in the nucleolus under heat shock, acidosis and transcriptional stress in human cells. Therefore, the same recNoDS could be tested in Arabidopsis under heat stress to see if it is targeted to the nucleolus.

The change in subcellular localisation carried out by NoDS_{LAS1} has been examined under heat stress (37°C 4 h). The authors of the nucleolar sequestration in human cells analysed the effect of the NoDS under

heat shock (Audas et al., 2012a). The first substantial difference resides in the heat treatment. In the case of human cells, 37°C are considered as normal conditions, and the heat treatment (heat shock) consisted in exposing cells to 42°C for up to 3 hours, followed by a recovery phase of 37°C for 30 min. In contrast, normal conditions in *Arabidopsis* are set at 22°C. Instead of heat shock, it was decided to apply heat stress (37°C 4 h). It would be interesting to see whether the same results are observed in *35S_{pro}:NoDS_{LAS1}-GFP* and *35S_{pro}:GFP-NoDS_{LAS1}* seedlings under heat shock. Besides that, human cells were also exposed to acidosis (1% oxygen) and transcriptional stress (4 mM actinomycin D and 8 mM MG132 for 3 h) to examine the role of the NoDS (Audas et al., 2012a). Neither acidosis nor transcriptional stress have been conducted to study the role of NoDS_{LAS1} in *Arabidopsis*.

Even though NoDS_{LAS1} promotes nucleolar localisation under at 37°C 4 h in *Arabidopsis*, it is not known whether this motif exerts a role in protein sequestration. Thus, the idea whether the human-like NoDS of *Arabidopsis* LAS1 may act as a Nucleolar Localisation Signal (NoLS) is also discussed. Compared to NoLSs, or even NLSs, the human-like NoDS is more complex in terms of amino acid composition. Generally talking and despite the different types of NLSs, these signals are mainly composed of arginine and lysine. Similarly, NoLSs are rich in hydrophobic residues such as lysine or arginine (reviewed by Muñoz-Díaz and Sáez-Vásquez, 2022). Some examples of these signals include the NLSs of HsfA2 and RTL2 in *Arabidopsis* (²³⁰KEKKSLFGLDVGRKRR²⁴⁵ and ³⁷¹KKAESSAYHMIRALRK³⁸⁷, respectively; Comella et al., 2008; Evrard et al., 2013), or the NoLS of *Arabidopsis* coilin (²⁰²KKKKKKK²⁰⁸; Makarov et al., 2013). Apart from the NoDS_{LAS1}, a NoLS has been predicted in the sequence of *Arabidopsis* LAS1 using the webserver “NoD” (Scott et al., 2011). Its putative NoLS is ⁵²³RNQLEKSPCKRARKSAGDSE⁵⁴², which is not close to NoDS_{LAS1} (Figure 4.9A and B). Even though it is a predicted NoLS, it must be experimentally validated to confirm its role. In addition, NoDS_{LAS1} cannot be considered as a NoLS since it drives localisation in the nucleolus only during heat stress. NoDS_{LAS1}-GFP and GFP-NoDS_{LAS1} are distributed in the cytoplasm and nucleus of *Arabidopsis* under normal conditions.

Several strategies could be followed to fully characterise the role of NoDS_{LAS1} in *Arabidopsis* (Figure 4.9C). First of all, a truncated version of LAS1 could be generated, where the NoDS_{LAS1} is removed. The subcellular localisation of the truncated version of LAS1 would be examined at normal conditions as well as 37°C, in order to analyse possible changes in distribution. Moreover, the different mutagenesis analyses could be performed to dive into the critical regions of NoDS_{LAS1} in nucleolar localisation under heat stress. Finally, a possible functional interaction between the NoDS and NoLS could be examined. On the one hand, its predicted NoLS must be experimentally validated. For that, the subcellular localisation of the C-terminal fragment would be tested under normal conditions. Moreover, a truncated version of LAS1 lacking its predicted NoLS could also be generated. On the other hand, the subcellular of a recombinant LAS1 lacking both NoDS and NoLS would also provide determinant information. Nevertheless, whether NoDS_{LAS1} can be considered as a human-like NoDS can be debated.

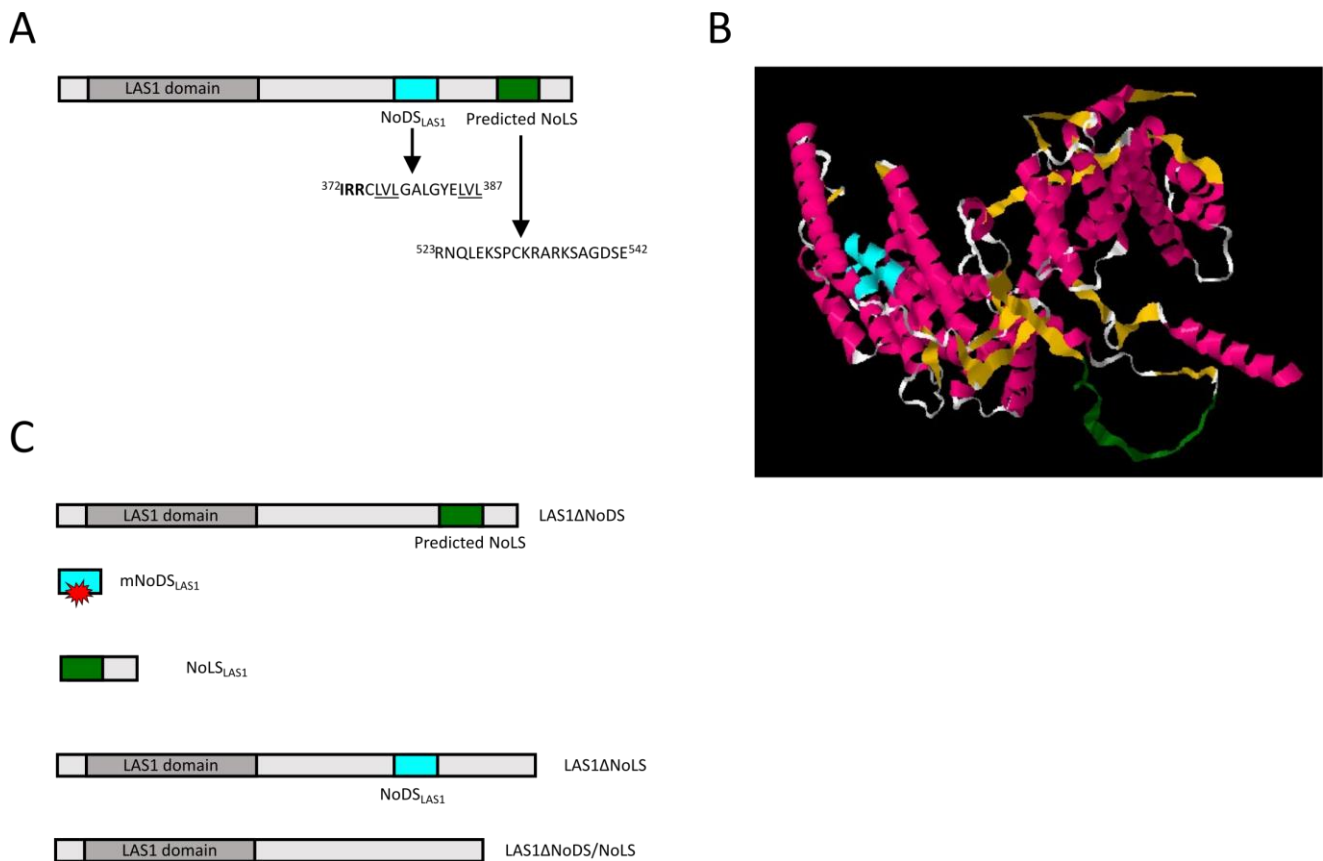


Figure 4.9. Analysis of the sequence and structure of Arabidopsis LAS1. (A) Schematic representation of the sequence of Arabidopsis LAS1. The LAS1 domain was positioned after performing a blastn with *Homo sapiens* LAS1 domain. The prediction of the NoLS was performed using NoD (Scott et al., 2011). (B) Predicted 3D structure of Arabidopsis LAS1 using AlphaFold. The NoDS and the predicted NoLS have been marked in cyan and green, respectively. (C) Strategies to characterise the role of NoDS_{LAS1}. LAS1ΔNoDS is the recombinant LAS1 protein lacking its NoDS, mNoDS_{LAS1} represents different versions of a mutagenized NoDS, NoLS_{LAS1} portrays the experimental validation of the predicted NoLS, LAS1ΔNoLS consists of the lack of LAS1 NoLS, and LAS1ΔNoDS/NoLS exemplifies the recombinant protein lacking NoDS_{LAS1} and NoLS_{LAS1}.

On the other hand, the IGS transcripts are part of the nucleolar sequestration of proteins under abiotic stresses in humans. These sense transcripts are transcribed by RNA Pol I, being further processed into mature lncRNAs. The accumulation of IGS transcripts under heat stress has been characterised in Arabidopsis, having a peak at 37°C 24 h. Nevertheless, the behaviour of this accumulation is rather particular. They accumulate at 37°C according to RT-PCR, RT-qPCR and RNA FISH analyses, forming some sort of aggregates at 37°C 4 and 24 hours. Thus, the IGS transcripts have an early response. Besides that, according to RT-PCR and RT-qPCR results, the IGS transcripts would have a reversible response, as their accumulation halts at 22°C (Figure 3.16B). Nevertheless, the response of the IGS could be considered as persistent if the results from the RNA FISH are taken into account (Figure 3.18B). This striking difference can be explained due to the methodology. In RT-PCR and RT-qPCR experiments, the pair of primers p151 + p150 was employed, targeting a specific region of the IGS

transcripts, which is upstream the GP. In contrast, a pool of diverse probes targeting different sites of the IGS has been employed in RNA FISH analysis. Thus, the signal observed in the RNA FISH experiments might come from different IGS transcripts, whereas RT-PCR and RT-qPCR signals correspond to a specific transcript (Figure 3.16A and 3.18A). Moreover, strand-specific RT-PCR analyses revealed that the IGS transcripts detected under heat stress are sense transcripts, similar to IGS₁₆RNA, IGS₂₂RNA and IGS₂₈RNA in humans.

The participation of the IGS transcripts in a possible nucleolar sequestration of proteins in *Arabidopsis* under heat stress cannot be confirmed. In other words, it is not known whether the foci of IGS transcripts observed in the nucleolar space under heat stress in RNA FISH experiments are related with the nucleolar sequestration of proteins. The composition of these foci remains uncharacterised. GFP immunoprecipitation experiments in *35S_{pro}:NoDS_{LAS1}-GFP* or *35S_{pro}:GFP-NoDS_{LAS1}* plants at 37°C 4 h followed by RT-PCR employing p151 + p150 would be required to establish a possible interaction of the IGS transcripts with the putative NoDS. In fact, a preliminary pull-down of the IGS transcripts at 22°C and 37°C 24 h using biotinylated probes in collaboration with Dr. Federico Ariel at the “Instituto de Agrobiotecnología del Litoral, CONICET-UNL” had been performed to identify protein interactors. Nevertheless, these experiments were not successful, as no specific protein interactors were identified under heat stress (Sáez-Vasquez, J. personal communication). However, it could be stated that these transcripts participate in the response of the nucleolus to heat stress. It is known that ncRNAs are involved in the response to abiotic stresses in *Arabidopsis*, including heat stress (reviewed by Ma et al., 2022; reviewed by Zhang et al., 2022). Besides that, the formation of cellular bodies by LLPS requires a stochastic first step: the nucleation. Whereas the transcription of the rDNA is the nucleation in the nucleologenesis (Falahati et al., 2016; Dash et al., 2023), the transcription of the IGS may act as the nucleation step in the formation of the atypical bodies of the nucleolar proteins under heat stress (3.2. Behaviour of nucleolar components to heat stress in *Arabidopsis*). Alternatively, the IGS transcripts may possess a chaperone role in the nucleolus under heat stress, gathering the nucleolar proteins in the nucleolar space. However, the actual function of the IGS transcripts remains unanswered.

Regarding the transcription of the IGS, preliminary FLAG immunoprecipitation experiments carried out using *NRPA3_{pro}:NRPA3^m-FLAG-HA (nrpa3)* seedlings suggest a possible interaction of the IGS transcripts with this RNA Pol I subunit at 37°C 4 h. Based on these preliminary results and previous results from our group (Sáez-Vásquez, J. personal communication), we strongly support the idea that the IGS are transcribed by RNA Pol I under heat stress in *Arabidopsis*. On the contrary, the presence of SPs in the IGS leads to consider whether the IGS is transcribed independently from the 45S rDNA, whose TATA box has been characterised. Nevertheless, these promoters have been suggested to be acting as enhancers of the transcription of the 45S rDNA (Doelling et al., 1993). Besides that, some of the IGS variants described by Havlová et al., (2016), do not possess any SP, fuelling the idea of transcription from the GP by RNA Pol I under heat stress. In addition, the formation of similar

foci-like structures for NRPA3^m-FLAG-HA and the IGS transcripts under heat stress suggests a possible physical and functional interaction between them.

Thus, the current model supports the idea of the IGS being transcribed by RNA Pol I. At 22°C, the 45S rDNA is transcribed by RNA Pol I from the GP, stopping after the 3'ETS. The 45S pre-rRNA is then processed to obtain the mature 5.8S, 18S and 25S rRNAs (Figure 4.10A). At 37°C, the 45S rDNA is transcribed by RNA Pol I from the GP. The IGS is also transcribed from the GP as a read through transcript. This transcript is quickly processed to obtain the 45S pre-rRNA and the pre-IGS transcript. Whereas the 45S pre-rRNA is not further processed and accumulates (Darriere et al., 2022), the pre-IGS transcript goes through a series of processing steps to render the mature IGS transcripts (Figure 4.10B). It must be mentioned that this model is hypothetical, as some experiments will be necessary to support or fully prove it. First of all, the RNA species comprising the 45S pre-rRNA and the pre-IGS has not been detected by Northern blot (Sáez-Vásquez, J. personal communication). Similarly, the pre-IGS transcript has not been observed in RT-PCR experiments (Figure 3.16). This could be explained by a fast cleavage events, limiting the possibility of detecting putative IGS precursors. Nevertheless, further experiments employing the RNA Pol I, II and III inhibitor actinomycin D as well as the RNA Pol II inhibitor α -amanitin will be required to fully address whether the IGS is transcribed by RNA Pol I in *Arabidopsis*.

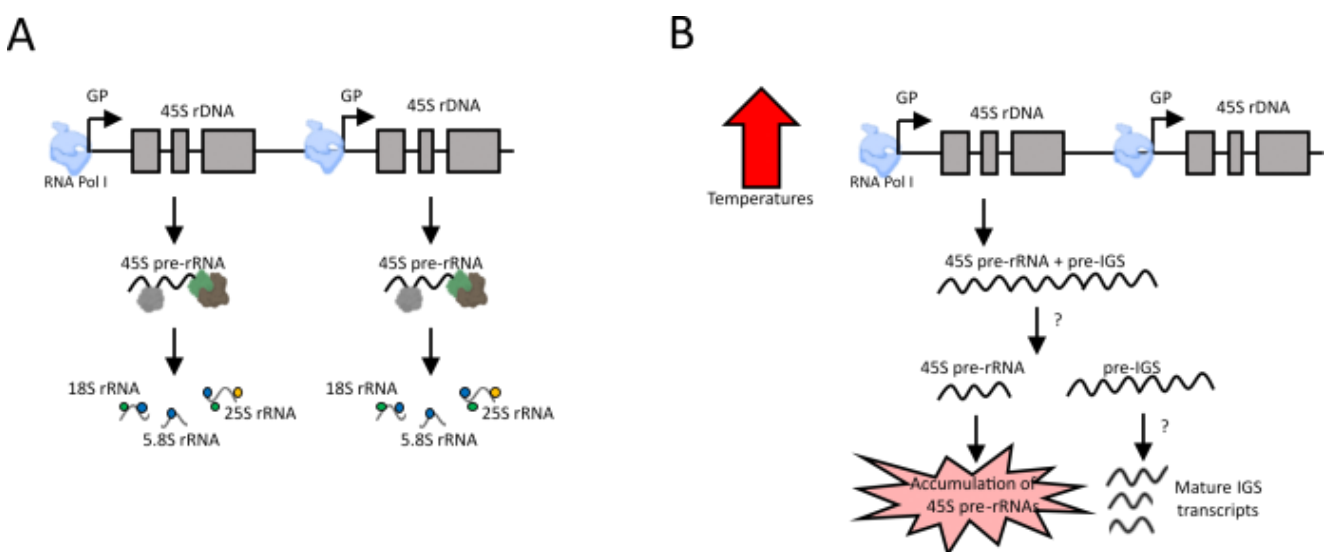


Figure 4.10. Model of rRNA IGS transcription at normal conditions and heat stress. (A) The 45S rDNA is transcribed by RNA Pol I from the GP at 22°C. After the processing of the 45S pre-rRNA, the mature 5.8S, 18S and 25S rRNAs are obtained. (B) Under heat stress, the 45S rDNA and the IGS are transcribed by RNA Pol I from the GP as a single transcript. The processing of the 45S pre-rRNAs halts under heat stress, leading of the accumulation of 45S pre-rRNAs. The pre-IGS transcripts are further processed to obtain the mature IGS transcripts, which accumulate. In this model, we propose read-through transcription from GP under heat stress, although transcription from IGS cannot be entirely excluded. This figure was partially generated using BioRender.com.

Even though the NoDS_{LAS1} drives localisation to the nucleolus under heat stress, plus IGS transcripts are accumulated and located in the nucleolar space under heat stress, the occurrence of the nucleolar sequestration of proteins in Arabidopsis remains elusive. Moreover, the formation of the detention centre in Arabidopsis can be questioned. It must be mentioned that the nucleolar sequestration of proteins under certain abiotic stresses is a controversial subject in the scientific community (Sáez-Vásquez, J. personal communication). Even though it has been described in the literature (Audas et al., 2012a), the authors later described a similar phenomenon, named the formation of Amyloid-bodies (A-bodies) in the nucleus (Audas et al., 2016). The formation of A-bodies is based on the amyloidogenic propensity of proteins under certain stimuli, i.e., acidosis or heat shock. This amyloidogenic behaviour is determined by a motif termed amyloid-converting motif (ACM), which is equivalent to the NoDS. The ACM has been described in proteins such as VHL, where the presence of the NoDS has been also described. Moreover, the ACM is characterised by the presence of an R/H-rich domain and an amyloidogenic region, which resembles the NoDS consensus sequence. As an example, the NoDS of the human VHL is ¹⁰⁷**RRI**HSYRGHLWLFRDAGTHDGLLV¹³⁰ (arginine motif and hydrophobic triples shown in bold and underlined respectively), whereas its ACM is ¹⁰⁴**GTG****RRI**HSYRGHLWLFRDAGTHDGLLVNQTELFVPS¹³⁹ (where the R/H-rich domain and the amyloidogenic region are highlighted in yellow and cyan, respectively, and the arginine motif and hydrophobic triplets of the NoDS are shown in bold and underlined, respectively). Nevertheless, Hsp70, which was initially catalogued as a NoDS-containing protein (Audas et al., 2012a), does not possess an ACM. In fact, HSPs seem to be involved in the process of assembly and disaggregation of A-bodies (Audas et al., 2016). On top of that, the interaction of IGS-derived lncRNAs with the ACM of the proteins forming A-bodies has also been described (Audas et al., 2016). These lncRNAs possess low-complexity regions, containing cytosine/uracil or adenosine/guanine repeats. These RNA species promote the interaction with the ACM of the proteins, seeding the formation of A-bodies upon certain stimuli (Wang et al., 2018). Nevertheless, the authors have not addressed anything regarding the nucleolar sequestration of proteins in the later published paper concerning A-bodies, generating a substantial degree of unreliability.

4.4. Concluding remarks

Overall, my PhD project deepened into several aspects of the response of the nucleus and nucleolus to high temperatures in Arabidopsis. First of all, a quantitative analysis of the nuclear proteome of Arabidopsis under normal conditions (22°C), heat stress (37°C 4 and 24 hours) and recovery period (R22°C; 37°C 24 h + 22°C 24 h) allowed the classification of nuclear proteins in ten different groups according to their relative protein abundance: EG, LG, TG, EPG, LPG, RG, ELG, LLG, TLG and CG. The presence of major nucleolar proteins, including FIB1/2 or NRPA3, has been observed in the RG and CG. Since these two groups are characterised by an increase in the relative protein abundance at R22°C, it is suggested that Ribosome Biogenesis Factors are required after prolonged exposure to heat stress, so that 45S rDNA transcription and pre-rRNA processing take place. Moreover, the comparison with the transcriptome under heat stress reveals substantial differences in the transcriptomic profiles compared to the proteomic profiles in the nucleus of the ten established groups. One of the reasons of this disparity is the existence of post-transcriptional modifications, nuclear import and export, and mRNA and protein stability under heat stress.

Besides that, the effects of heat stress towards Arabidopsis nucleoli have been addressed. The subcellular distribution of nucleolar proteins and RNA molecules under heat stress (37°C 4 and 24 hours) and recovery period (R22°C) reveals the formation of particular structures in the nucleolar space. These structures are formed due to the intrinsic nature of the nucleolus being a molecular condensate, whose formation is governed by LLPS. Nucleolar components act as multivalent molecules, and the exposure to high temperatures interferes in the normal organisation of the nucleolus as a molecular condensate.

Lastly, the role of the human-like NoDS has been addressed in Arabidopsis under heat stress. The NoDS_{LAS1} promotes nucleolar localisation under heat stress. Whether other human-like NoDS in Arabidopsis exert a similar role or not needs to be characterised. Besides that, the accumulation of IGS transcripts has been observed in the nucleolar space under heat stress. The implication of the human-like NoDS and the role of the IGS transcripts in the sequestration of proteins in the nucleolus under heat stress in Arabidopsis has not entirely been determined.

5. ANNEXES

5.1. Annexe 1. Impact of repeated heat/light stress exposure on nucleolus-based regulatory processes over generations (Ribostress)

The following annexe corresponds to a collaboration project carried out at CEREAP-Ecotron IleDeFrance as part of the ANR RiboStress research programme. This project aims to study the impact of high temperatures and intense light exposure, separately, over generations. For that, different Arabidopsis mutant (*det1-1*, *h1.1/h2.2*, *hda6*, *nuc1-2*, *nuc2-2*, *rh10-1* and *rtl2*) and control lines (Col-0F₃, Col-0F₄ and *35S_{pro}:FIB2-YFP*) were exposed to high temperatures (29°C/22°C for light/dark, respectively, 16 h light/8 h dark at 100 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) or increased light intensity (22°C/19°C for light/dark, respectively, 16 h light/8 h dark at 1000 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) up to the seventh generation. These mutants (*det1-1*, *hda6*, *nuc1-2*, *nuc2-2*, *rh10-1* and *rtl2*) exhibit abnormal pre-rRNA processing and/or impaired NOR expression (Earley et al., 2006; Kojima et al., 2007; Comella et al., 2008; Durut et al., 2014). Among the control lines, the *35S_{pro}:FIB2-YFP* overexpression line acts as a marker of the nucleolus for further analysis. The stress conditions are applied to 15-day-old seedlings until they bolted (Figure 5.1). Apart from the overall phenotypic aspect, other parameters, such as seed weight or bolting date, were measured and compared in different generations as well as between different lines and different conditions. Thus, the consequences of the exposure to high temperatures or intense light were evaluated from generation 1 to generation 5 (Run 1 to Run 5). It must be noted that the results from “Run 2 – Light stress” must not be considered due to experimental inconsistencies.

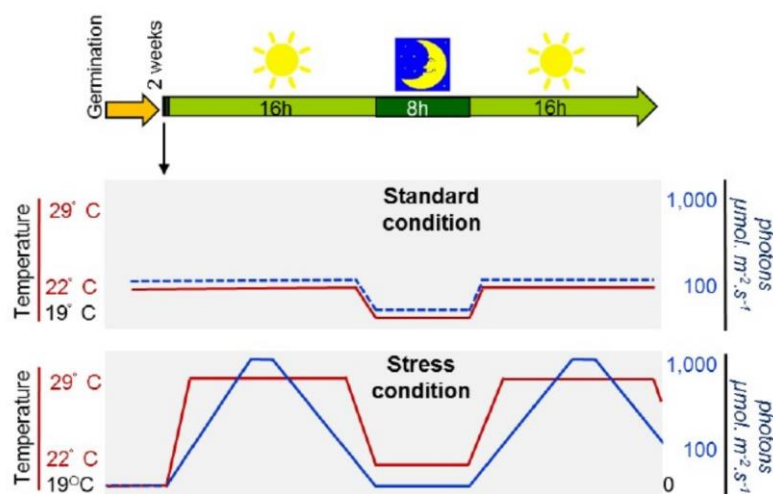


Figure 5.1. Experimental design for the evaluation of high temperatures and intense light over generations. The heat stress conditions are shown in red, whereas intense light conditions are represented in blue. Figure taken and modified from AAPG ANR 2017 “Genetic, epigenetic and protein-based crosstalks between nucleolus functional reorganization and response to stress”.

Regarding phenotypic appearance, *det1-1* and *nuc1-2* mutants are smaller than Col-0 plants under normal conditions (Figure 5.2A - C and G), which has been already mentioned in the literature (Chory and Peto, 1990; Durut et al., 2014; Gangappa and Kumar, 2018). Moreover, the size of *35S_{pro}:FIB2-YFP* overexpression

lines is slightly smaller than Col-0 (Figure 5.2A, B and D). Considering that this line was also envisioned as a control, it may be considered that the presence of the transgene and the subsequent overexpression of *FIB2* lead to the disparity in phenotypes. One of the main effects of the exposure to high temperatures is the presence of elongated petioles in comparison to control conditions (Figure 5.2A - J), as a result of the inhibition of cell growth (Saini et al., 2022). One of the most striking observations is the increased size of *nuc1-2* mutants grown under heat stress in comparison to the same mutants grown under normal conditions. In fact, this increase in size is more pronounced throughout the generations (Figure 5.2G). On the other hand, no apparent effects of exposure to increased light intensity have been observed in any of the control (Col-0 F₃ and F₄, and *35S_{pro}:FIB2-YFP* overexpression lines) or mutant plants, despite a slight decrease in size in some cases (Figure 5.2A - J).

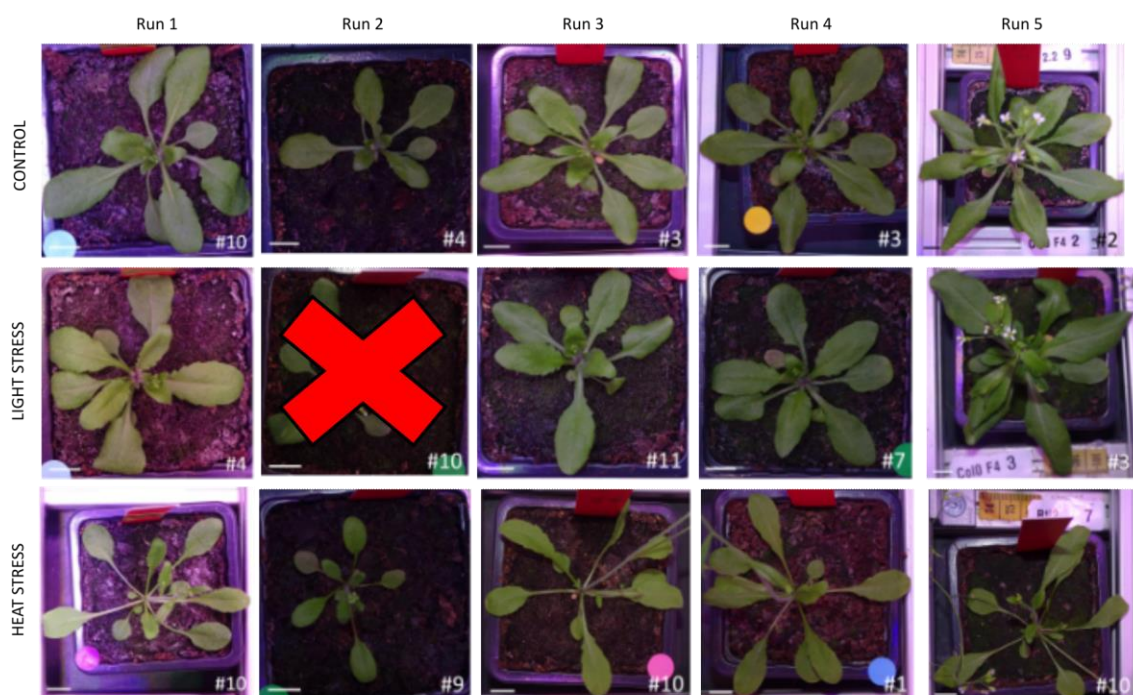
Analyses in the dry seed weight show a lower weight in *det1-1* and *nuc1-2* mutants in comparison to control plants in every run (generation) and at every condition. This is expected due to the smaller size of these mutants (Figure 5.3A). Once again, the *35S_{pro}:FIB2-YFP* overexpression line exhibits lower seed weight in comparison to Col-0 plants in many scenarios. These differences become quite dramatic in some cases (Figure 5.3A). When comparing the seed weight under heat stress throughout the five runs, there is a heterogeneous response in the behaviour of the different lines tested. *hda6*, *nuc1-2* and *nuc2-2* mutants display a very similar tendency, where the seed weight significantly increases in Run 5 compared of Run 1, after having decreased in Runs 2, 3 and 4 (Figure 5.3B). On the other hand, Col-0 (F₃ and F₄), *det1-1* mutants and *35S_{pro}:FIB2-YFP* overexpression line are characterised by a significant decrease in seed weight from Run 1 up to Run 5 (Figure 5.3B). Even though the seed weight is similar in control and heat stress conditions in Run 1, it significantly decreases in heat-stressed plants compared to control conditions in later runs (Run 2 – Run 5) as a general tendency (Figure 5.3C). As far as light stress is concerned, the seed weight behaves more randomly, and no particular behaviour could have been detected (Figure 5.3A - C).

Finally, the bolting times were measured as the number of days after sowing required for the plants to start flowering (Figure 5.4). The bolting times remain relatively stable in the initial runs (Run 1 – 3) among the different plant lines tested. Nevertheless, certain lines, such as *hda6* or *nuc1-2* mutants, seem to bolt later at the three conditions in Runs 4 and 5 (Figure 5.4A). Moreover, the tendency in bolting times looks very similar in all lines throughout the runs at every condition (Figure 5.4B). Whereas no particular conclusion has been extracted from light-stressed plants, heat-stressed plants bolt earlier than control plants over generations (Figure 5.4C).

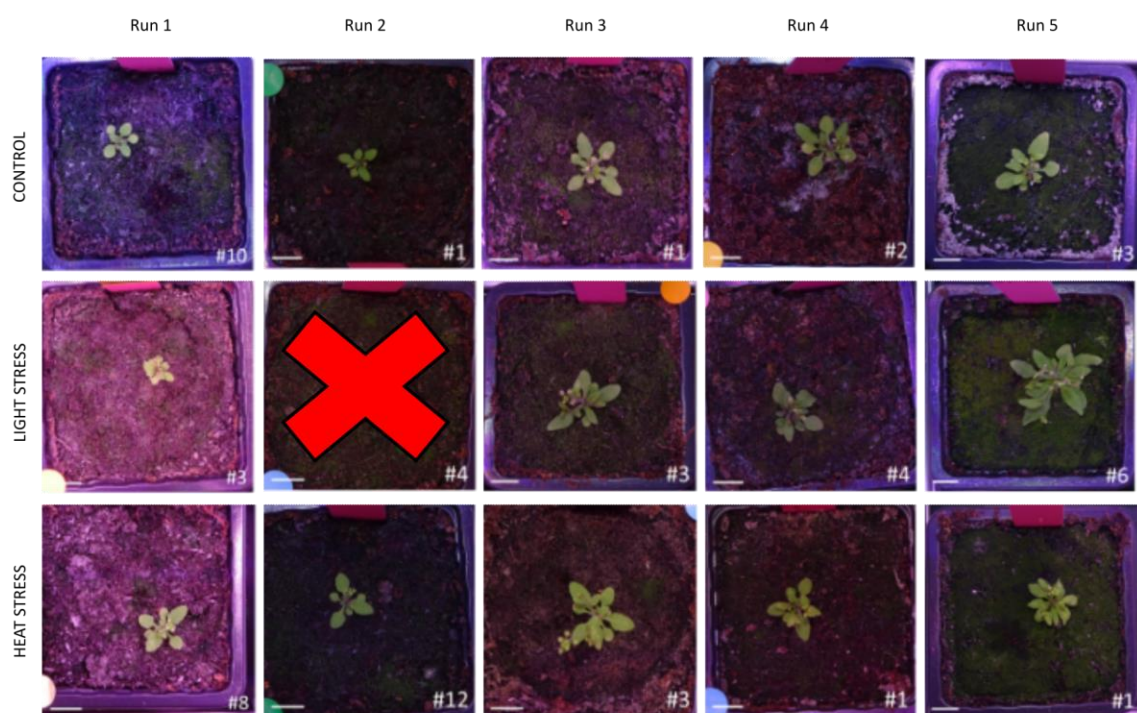
A



B

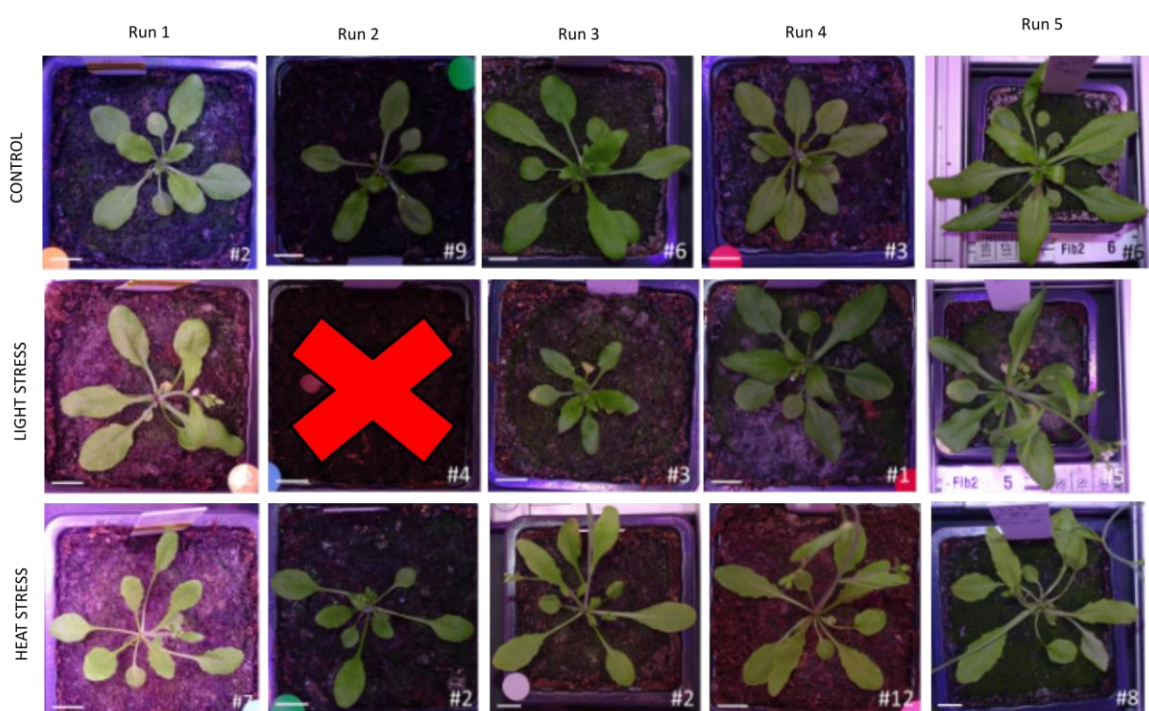


C



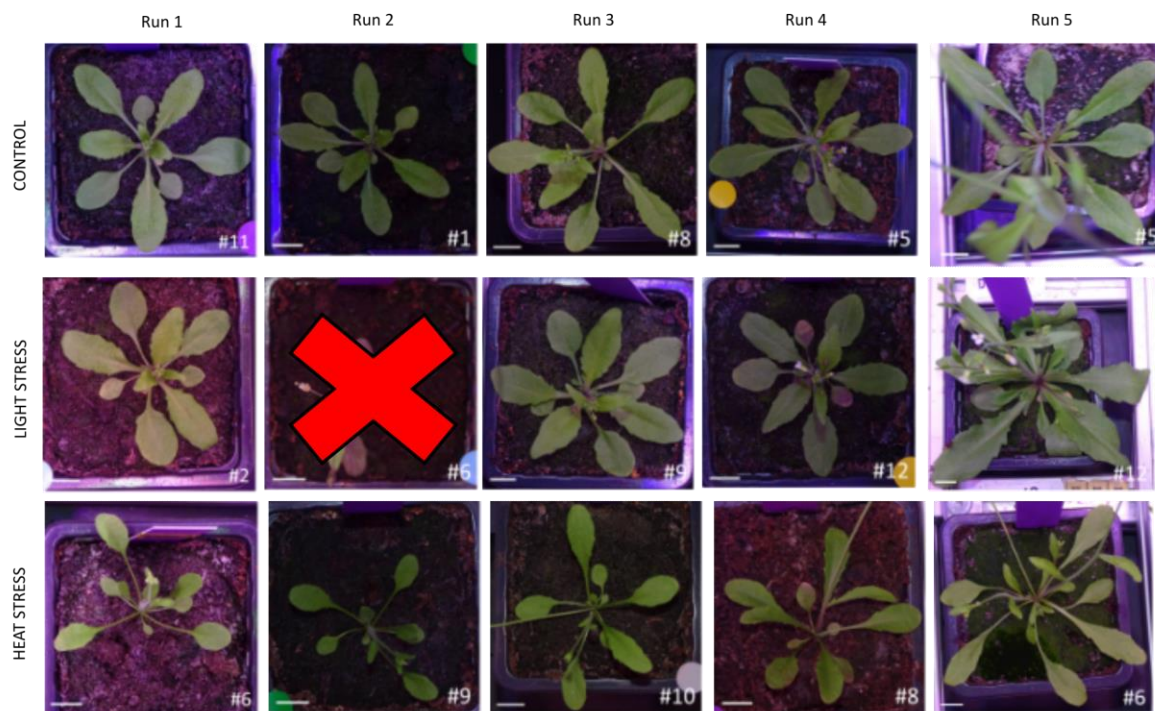
det1-1

D



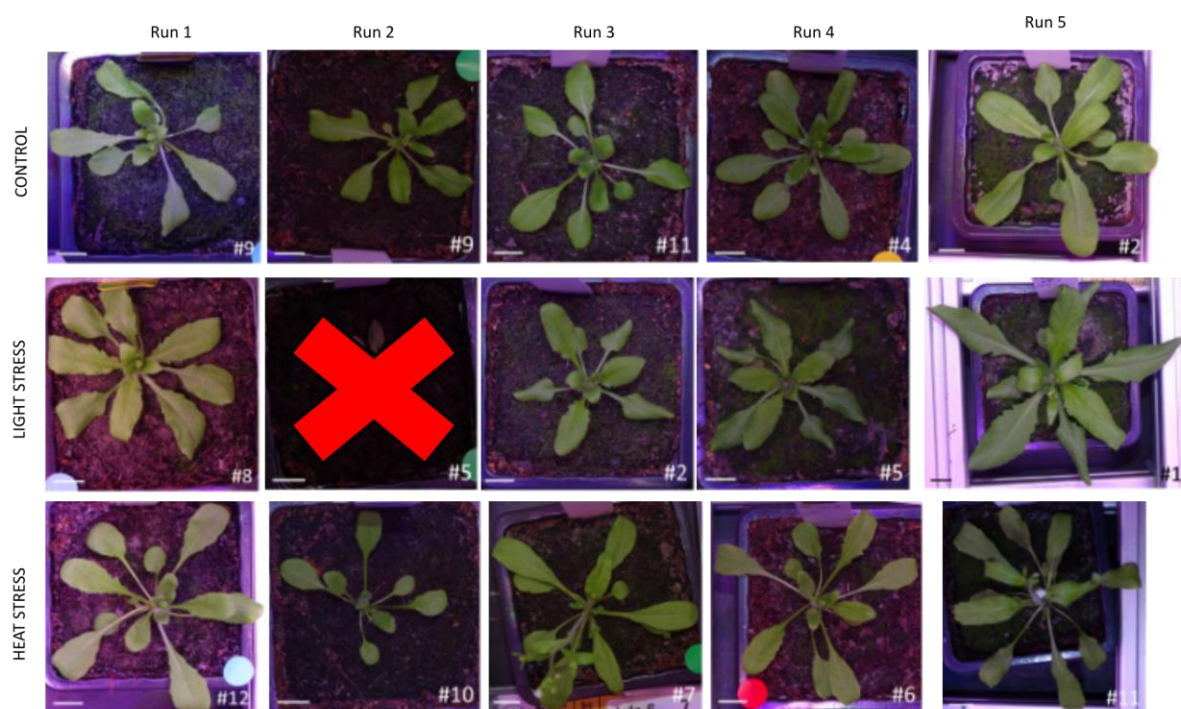
FIB2-YFP

E



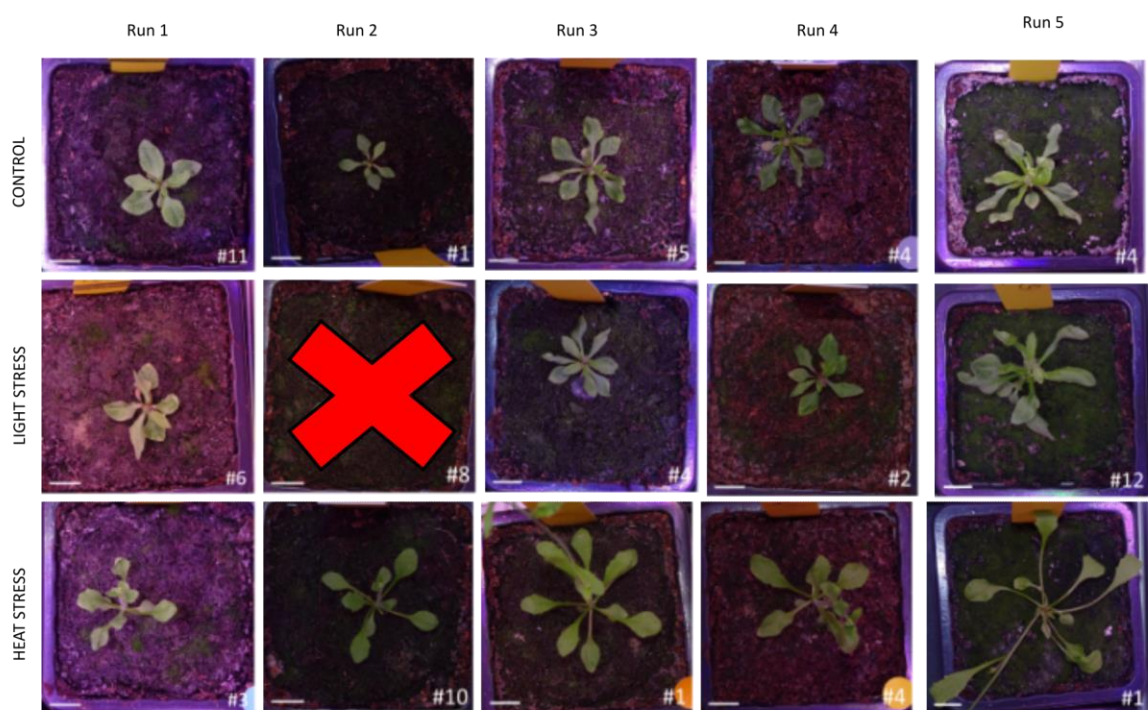
h1.1/h2.2

F



hda6

G



nuc1-2

H



nuc2-2

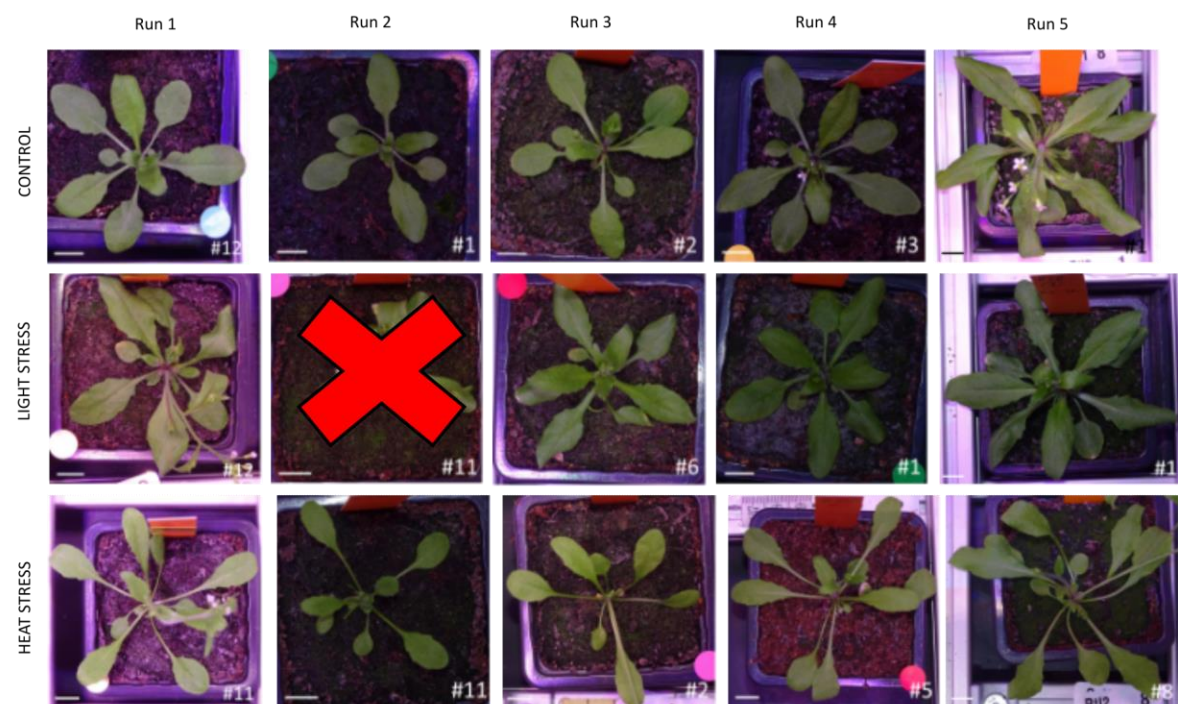
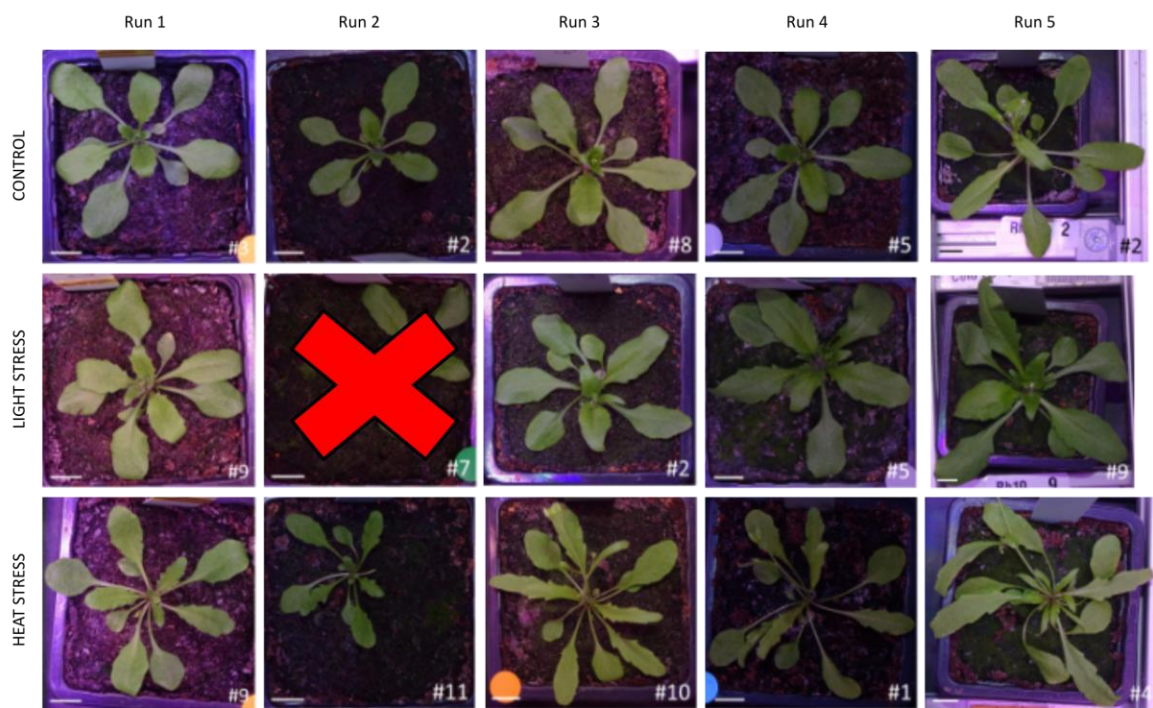
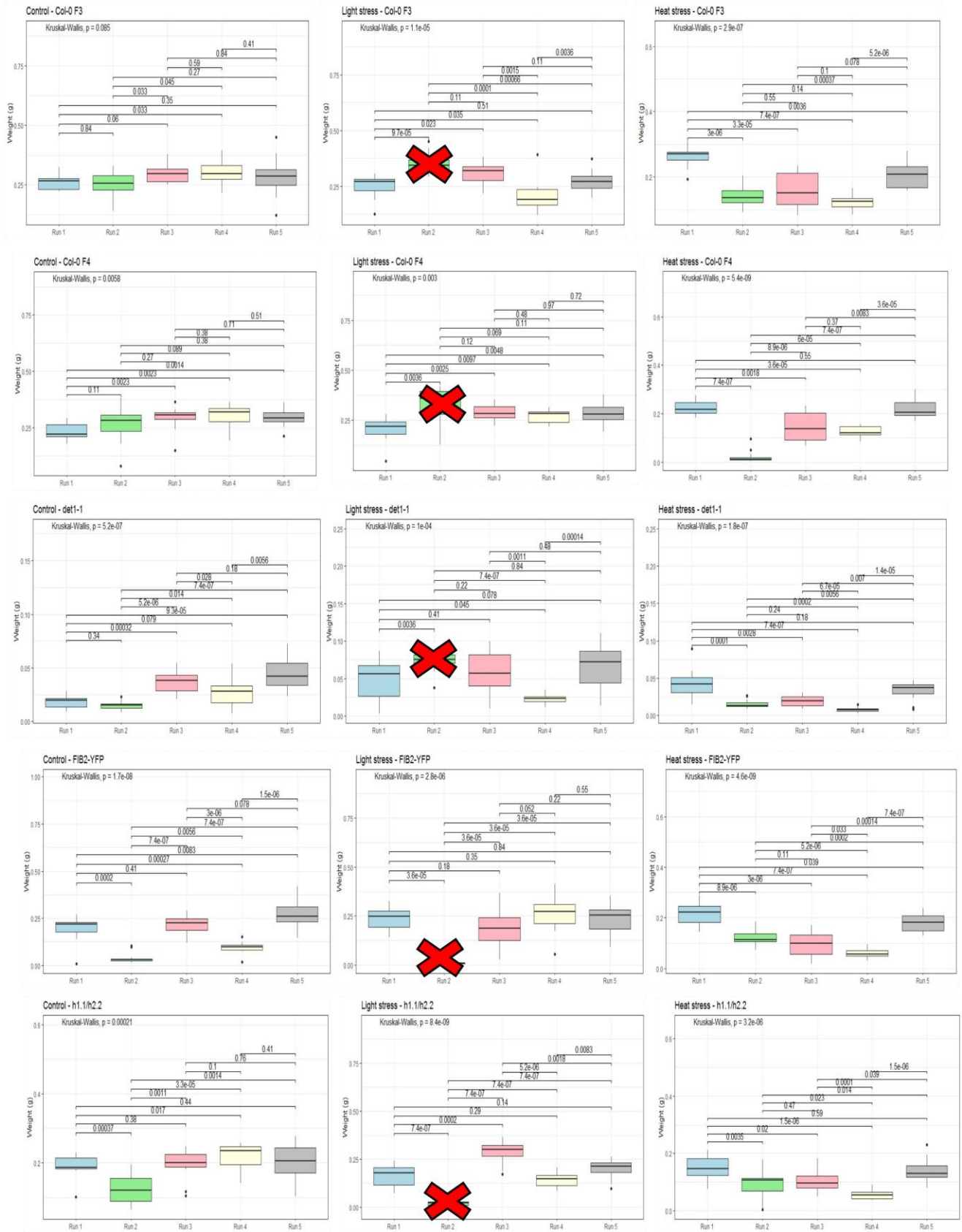


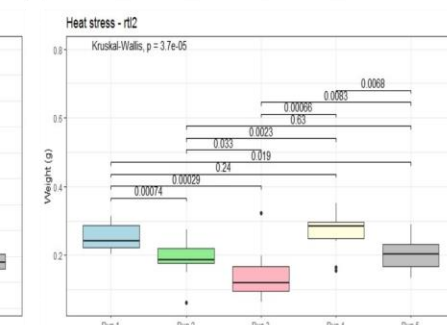
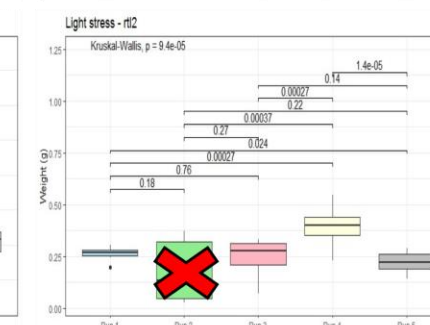
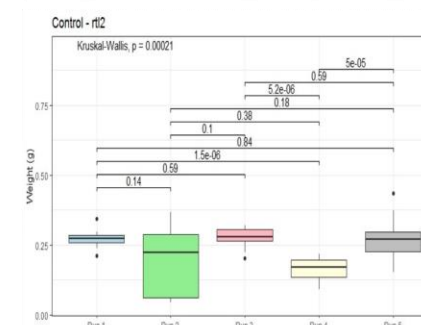
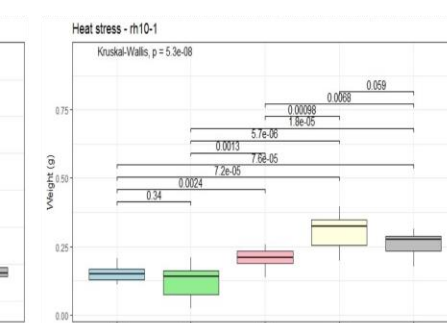
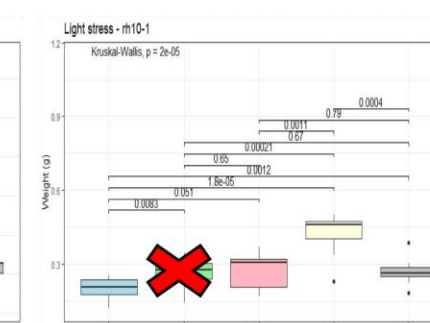
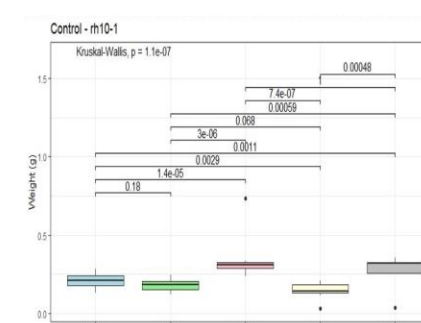
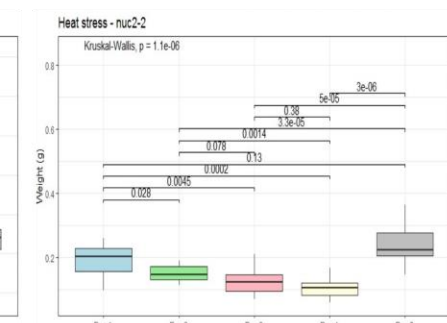
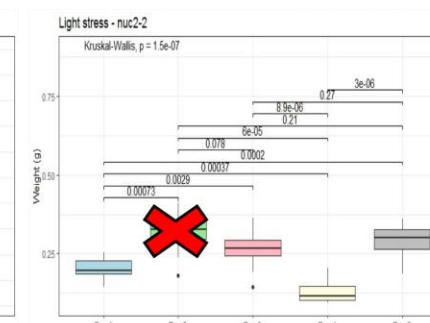
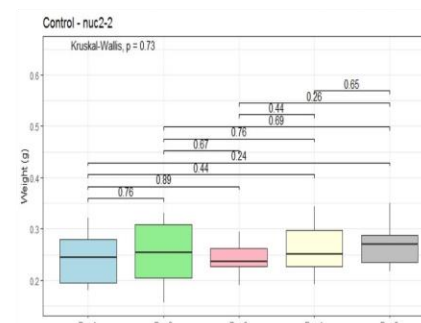
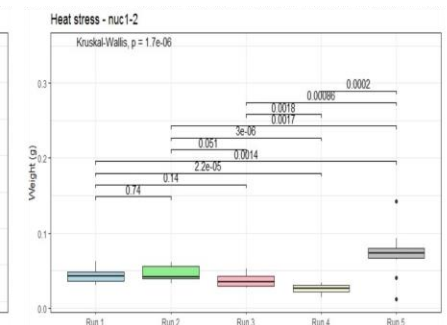
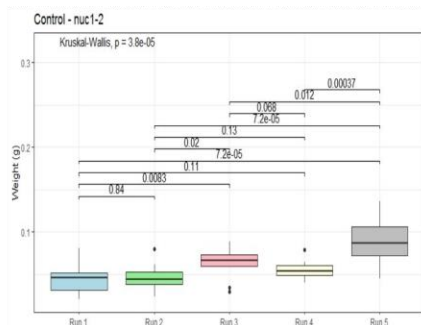
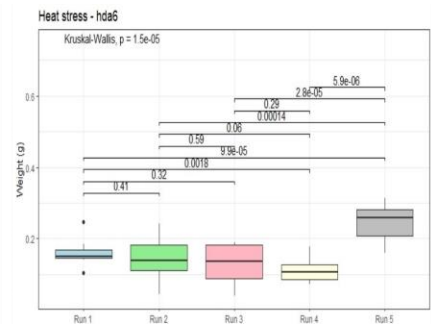
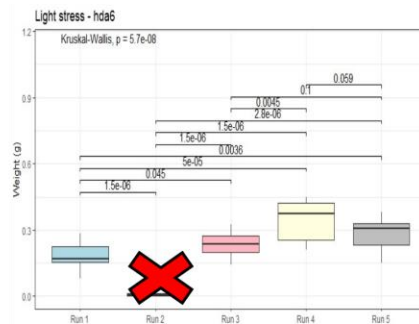
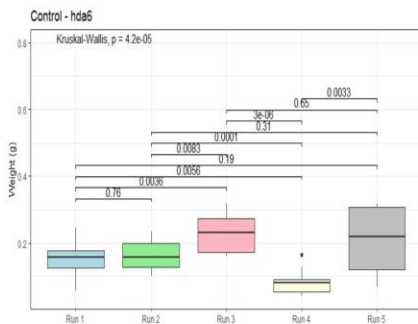
Figure 5.2. Phenotypic analysis of heat- or light-stressed plants. The effects of high temperatures and increased light intensity, separately, were evaluated in five generations (Run 1 – Run 5) in Col-0 F₃ (A), Col-0 F₄ (B), *det1-1* mutants (C), *35S_{pro}:FIB2-YFP* overexpression line (*FIB2-YFP*) (D), *h1.1/h2.2* double mutants (E), *hda6* (F), *nuc1-2* (G), *nu2-2* (H), *rh10-1* (I) and *rtl2* (J) mutants. Scale bar set at 1 cm. The pictures of “Run 2 – Light stress” have not been considered as for experimental inconsistencies.

A

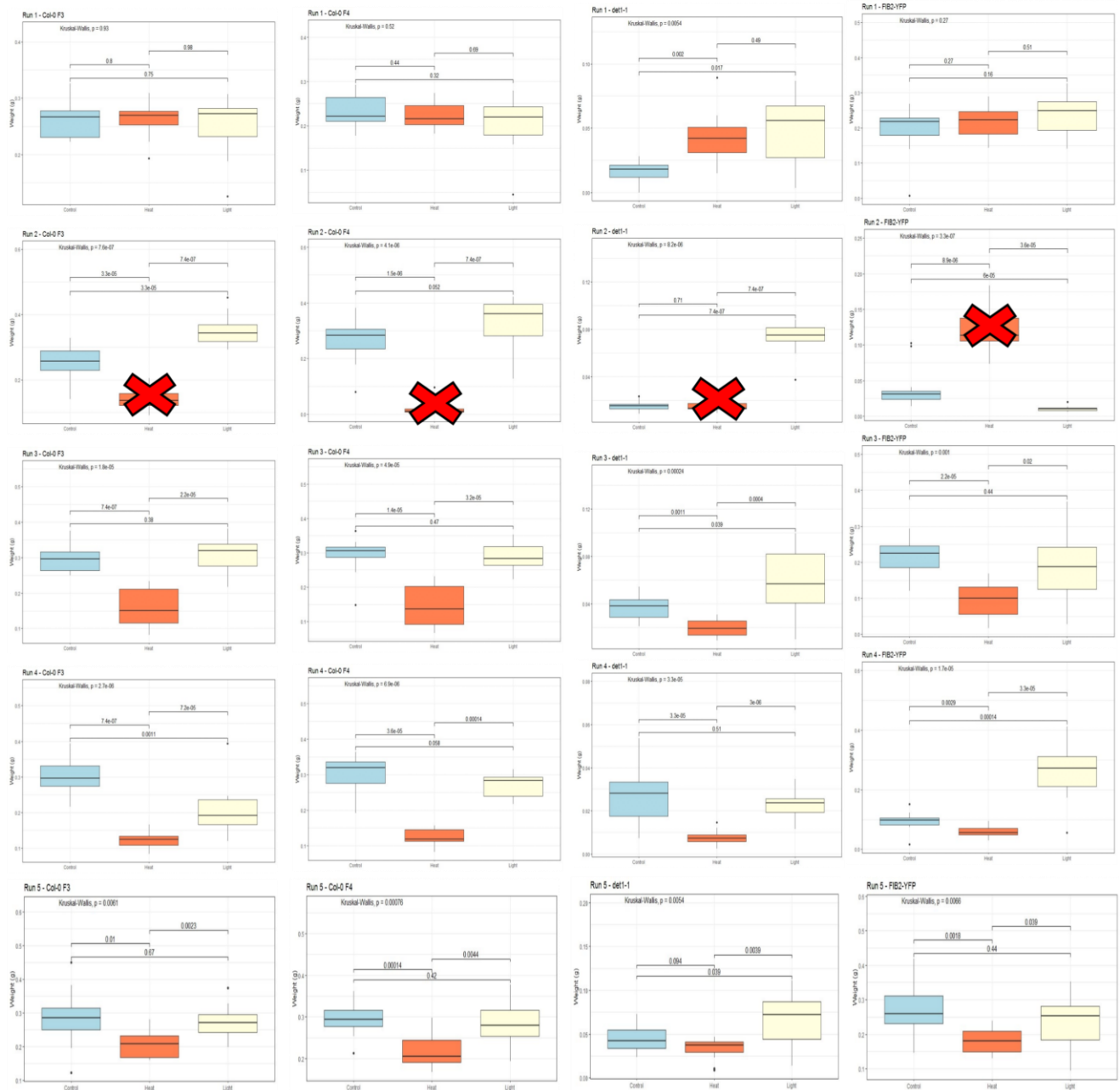


B





C





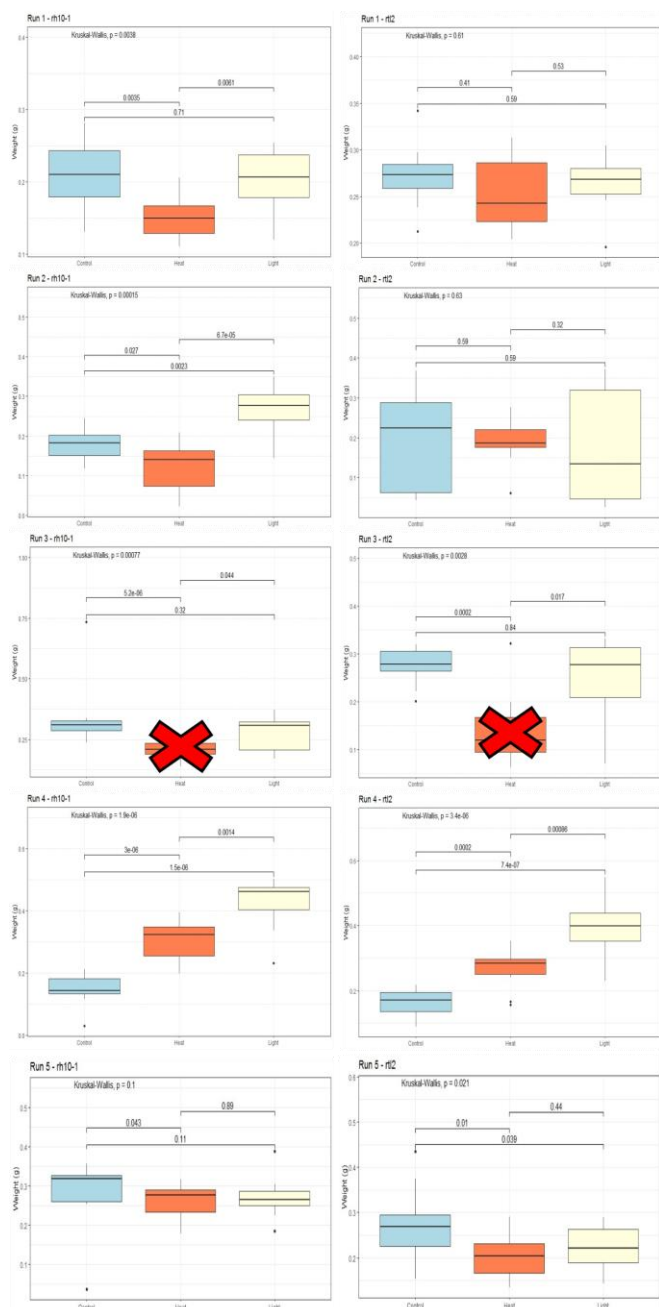
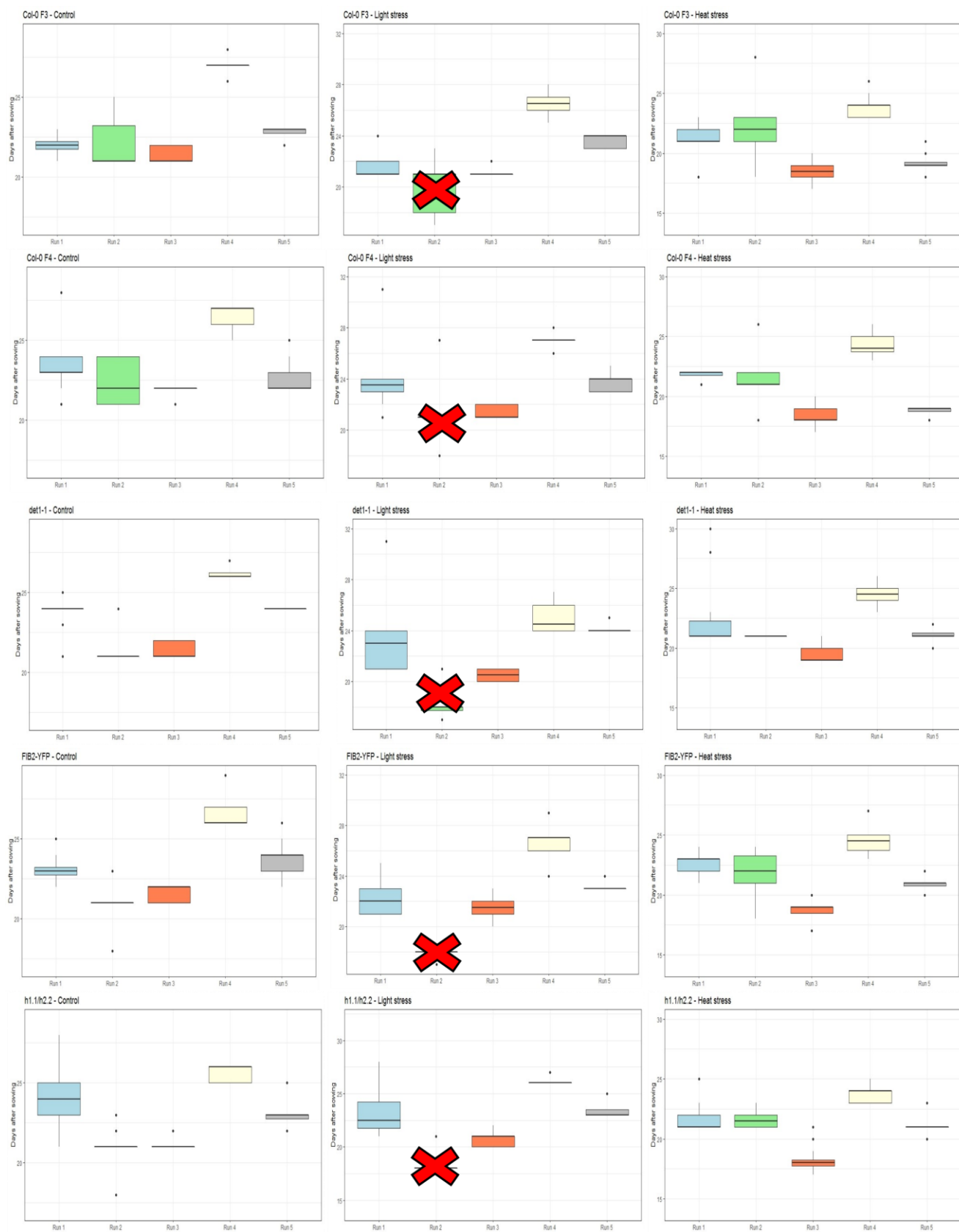


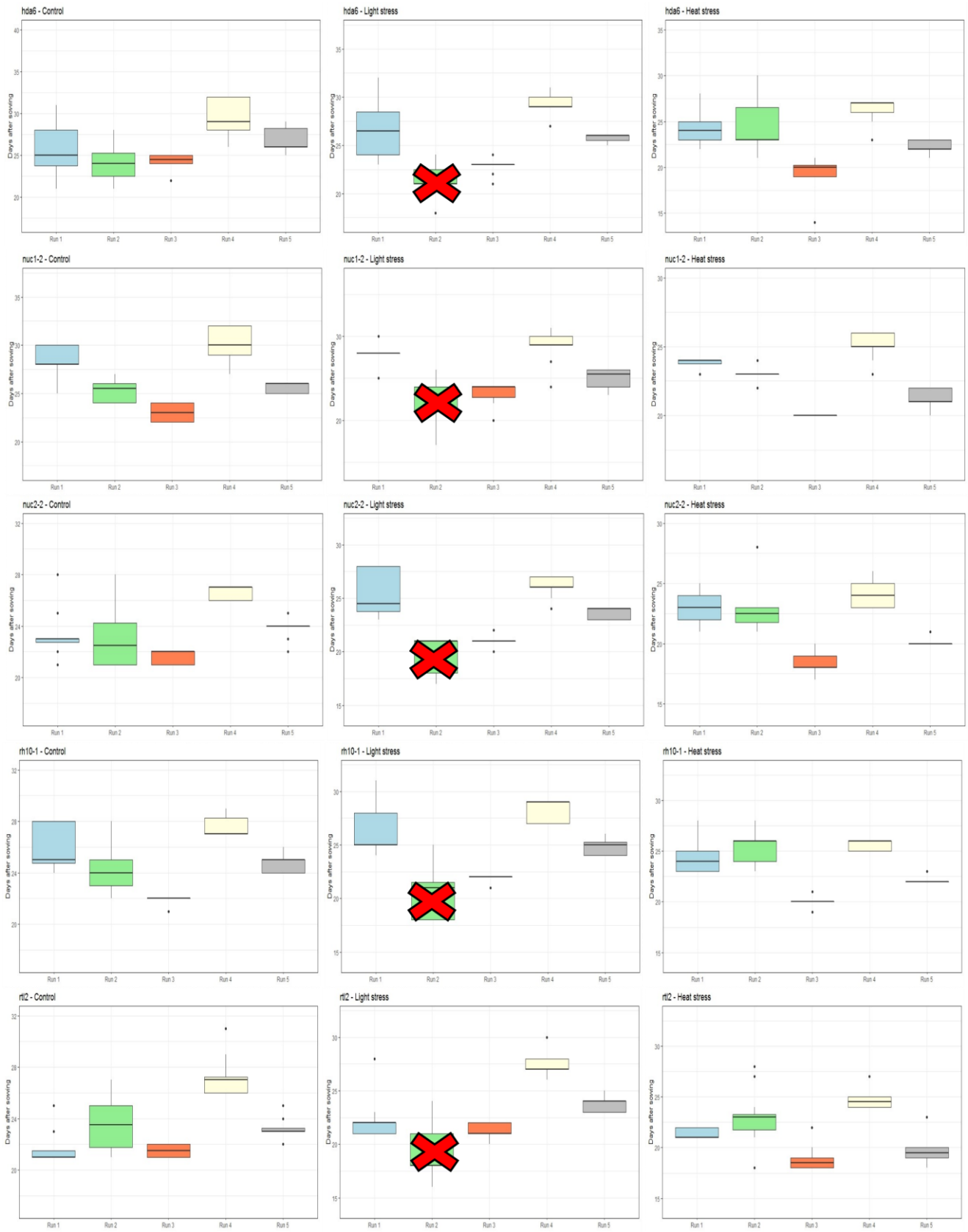
Figure 5.3. Comparative analysis of the dry seed weight. (A) Comparison of the seed weight of the different plant lines per run and condition. The statistical method applied is t-test, where “ns” stands for “not significant”, “*” p-value < 0.1, “**” p-value < 0.01, “***” p-value < 0.001, and “****” p-value < 0.0001. (B) Comparison of the seed weight of the different runs per condition and plant line. (C) Comparison of the seed weight of different conditions per run and line. The results from “Run 2 - Light stress” have not been considered due to experimental inconsistencies.

A

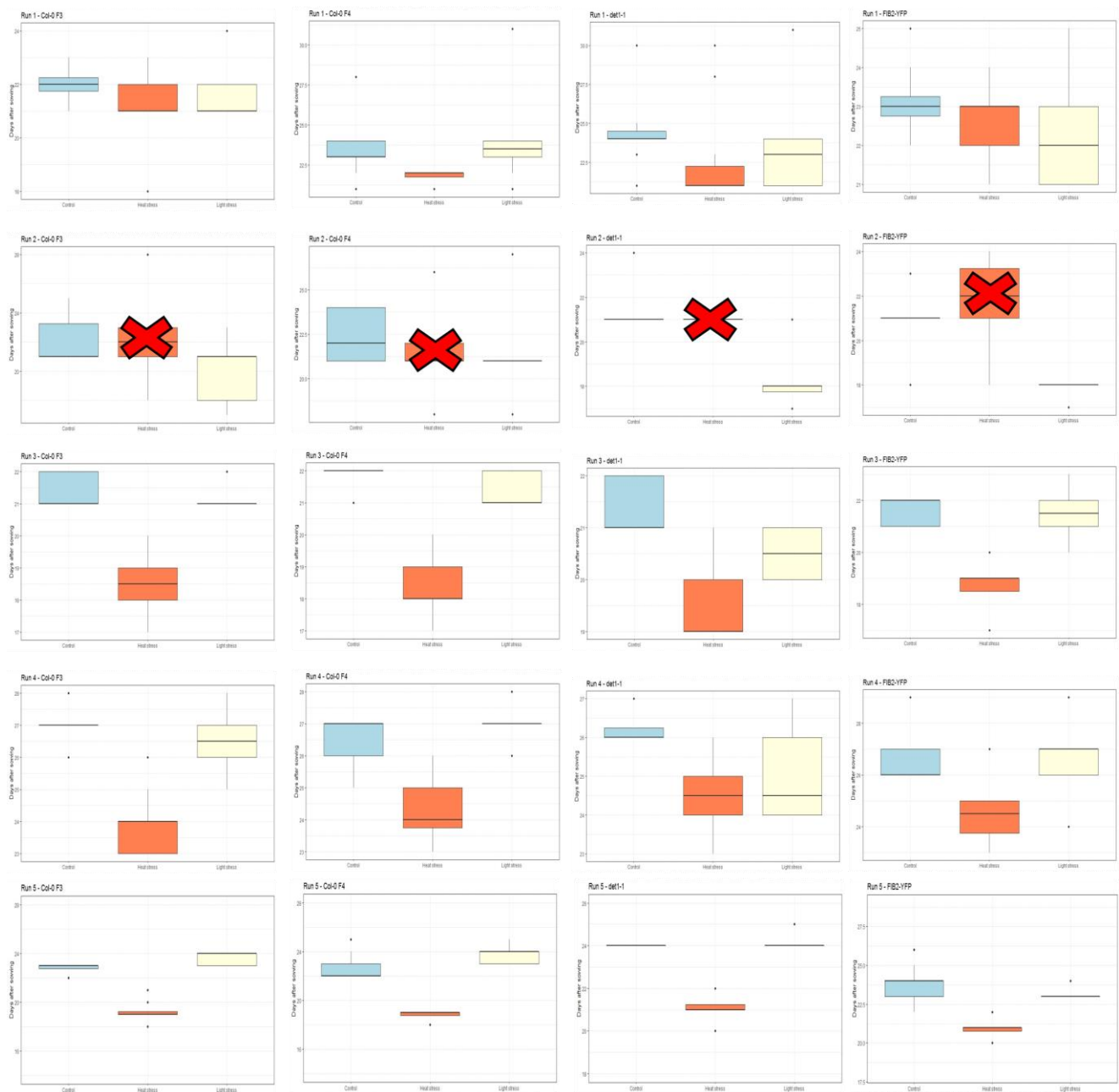


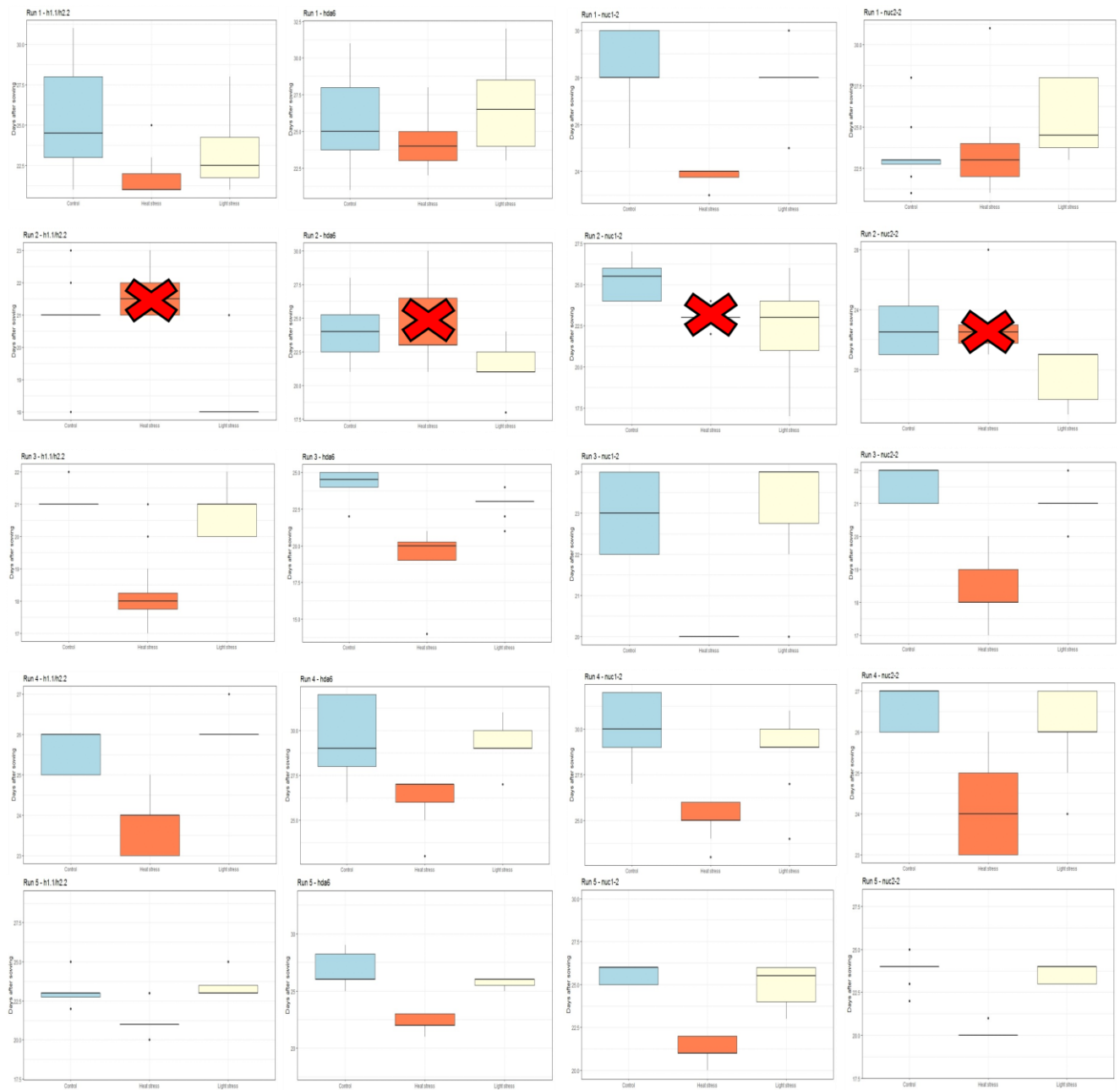
B





C





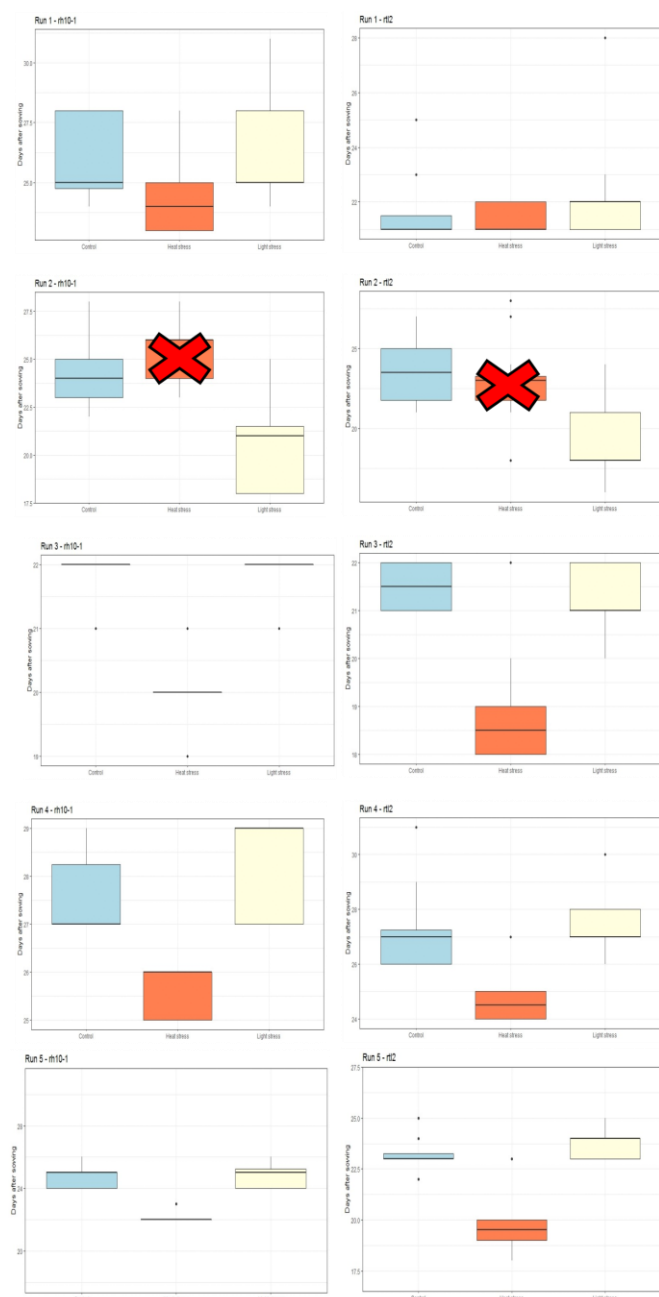


Figure 5.4. Comparative analysis of the bolting date. (A) Comparison of the days required to bolt of the different plant lines per run and condition. (B) Comparison of the days required to bolt of the different runs per condition and plant line. (C) Comparison of the days required to bolt under two conditions (light and temperature stresses) per run and plant line. The Y-axis represents the number of days from the sowing to the bolting. The results from “Run 2 - Light stress” have not been considered due to experimental inconsistencies.

5.2. Annexe 2. Further analysis regarding the IGS transcripts

In this thesis project, the accumulation of IGS transcripts under heat stress has been demonstrated by RT-PCR, RT-qPCR, and RNA FISH experiments (Figure 3.16 and 3.18). Nevertheless, this accumulation has been examined in 15-day-old Col-0 seedlings. Therefore, RT-PCR analyses were also performed using cDNA from other plant tissues and developmental stages. On the other hand, the question whether the IGS could be transcribed and accumulated upon biotic stresses in human cells or Arabidopsis plants remains elusive. Thus, the accumulation of IGS transcripts was analysed in Arabidopsis plants infected with *Myzus persicae* thanks to a collaboration with the Centro de Biotecnología Vegetal at Universidad Andrés Bello in Santiago de Chile.

5.2.1. Arabidopsis IGS transcripts in different tissues and developmental stages

The accumulation of IGS transcripts was preliminarily analysed in rosette leaves, 15-day-old seedlings, the aerial part of adult plants, siliques and flowers by RT-PCR using the pair of primers p151 + p150. As observed in Figure 5.5A, the IGS are hardly accumulated in rosette leaves, 15-day-old seedlings and flowers. There is basal accumulation of the IGS transcripts in the aerial part as well as in siliques. In addition, this accumulation was also examined in dry seeds, 2-, 4-, 7- and 10-day-old seedlings. Interestingly, the accumulation of IGS transcripts seems to be increasing from 4- to 10-day-old seedlings (Figure 5.5B). Nevertheless, a positive control (heat-stressed seedlings) must have been included in the same amplification, so that the accumulation levels could have been qualitatively compared. These preliminary results will need to be repeated, as some of the controls for genomic DNA contamination (-RT samples) are slightly positive.

5.2.2. Analysis of the accumulation of the IGS transcripts under biotic stress

M. persicae, commonly named green peach aphid (order Hemiptera), has been widely described as one of main contributors in crop pests. This aphid is also a vector of diverse plant viruses, including the Cauliflower Mosaic Virus (CMV), being a substantial threat for plant growth and development (Valenzuela and Hoffmann, 2015; Silva-Sanzana et al., 2023). To analyse the effect of *M. persicae* infection in the accumulation of IGS transcripts, 4-week-old Arabidopsis Col-0 plants were infected as described in Silva-Sanzana et al., (2019). Five different infection time points were tested, including the control: 0, 6, 12, 24 and 48 hours after infection. Then, total RNA was extracted, and cDNA was synthesised as previously described. The cDNAs were employed in a PCR using the pair of primers p151 + p150. The housekeeping gene used in this case is the thioredoxin-like protein YLS8 (Table 2.1), which has been in already published papers (Silva-Sanzana et al., 2019, 2022). Preliminary results suggest that infection with *M. persicae* do not trigger the accumulation of IGS transcripts (Figure 5.5C)

compared to the accumulation observed under heat stress (Figure 5.5D). Nevertheless, a positive control for *M. persicae* infection would be required to confirm a successful aphid infestation. For instance, it is known that the expression levels of ENHANCED DISEASE SUSCEPTIBILITY 1 (EDS) and PHYTOALEXIN DEFICIENT 4 (PAD4) are increased upon *M. persicae* infection (Louis et al., 2012). These two proteins are involved in the synthesis of salicylic acid, a molecule known to participate in the plant innate response against pathogens.

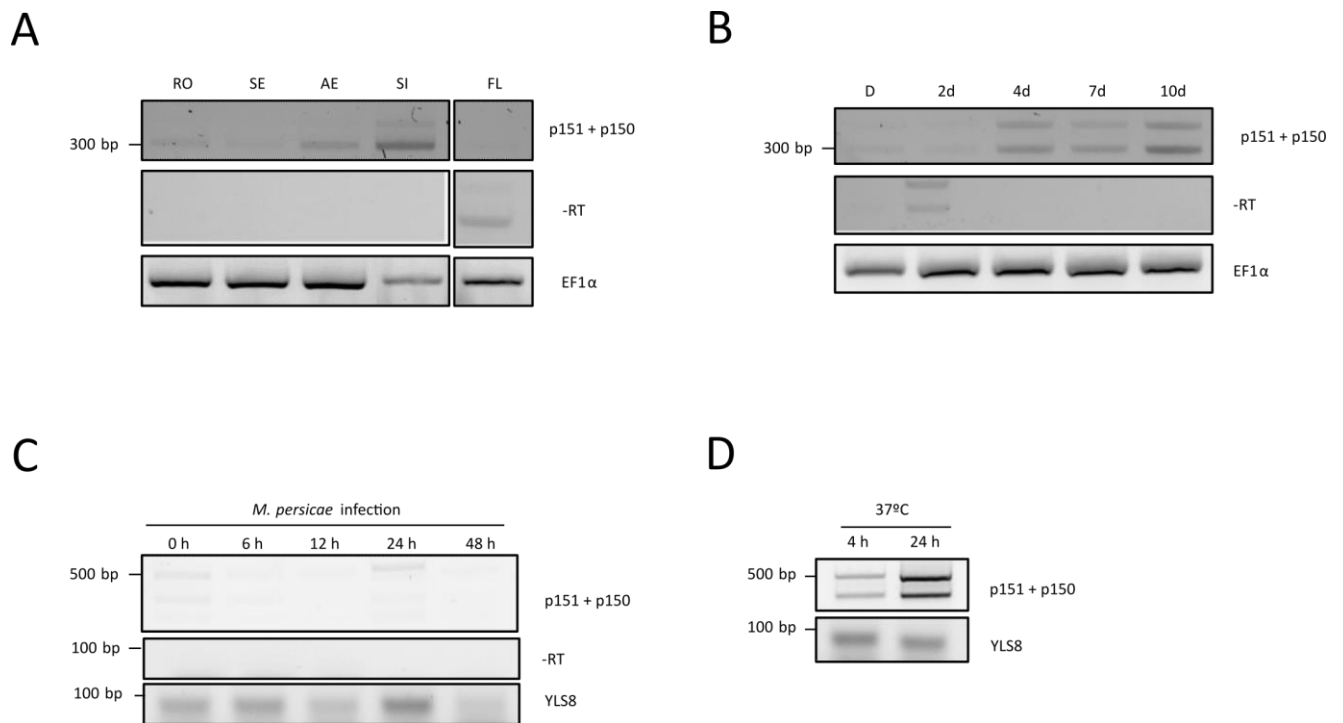


Figure 5.5. Further analysis with the IGS transcripts. (A) Accumulation of IGS transcripts in different tissues of Arabidopsis Col-0 plants. RO: rosette leaves; SE: 15-day-old seedlings; AE: aerial part; SI: siliques; FL: flowers. (B) RT-PCR analysis to detect IGS transcripts in different developmental stages of Arabidopsis Col-0 plants. D: dry seeds; 2d: 2-day-old seedlings; 4d: 4-day-old seedlings; 7d: 7-day-old seedlings; 10d: 10-day-old seedlings. Elongation Factor 1- α (EF1 α) was employed as housekeeping gene. (C) IGS accumulation under infection with *M. persicae*. RT-PCR shows employing cDNAs of infected plants after 0, 6, 12, 24 and 48 hours. (D) Accumulation of IGS transcripts under heat stress. YLS8 was used as a housekeeping gene.

6. REFERENCES

- Abou-Ellail, M., Cooke, R., and Sáez-Vásquez, J. (2011). Variations on a team: Major and minor variants of *Arabidopsis thaliana* rDNA genes. *Nucleus* 2, 294–299. doi: 10.4161/nucl.2.4.16561.
- Anderson, S. J., Sikes, M. L., Zhang, Y., French, S. L., Salgia, S., Beyer, A. L., et al. (2011). The transcription elongation factor Spt5 influences transcription by RNA polymerase I positively and negatively. *J. Biol. Chem.* 286, 18816–18824. doi: 10.1074/jbc.M110.202101.
- Audas, T. E., Audas, D. E., Jacob, M. D., Ho, J. J. D., Khacho, M., Wang, M., et al. (2016). Adaptation to Stressors by Systemic Protein Amyloidogenesis. *Dev. Cell* 39, 155–168. doi: 10.1016/j.devcel.2016.09.002.
- Audas, T. E., Jacob, M. D., and Lee, S. (2012a). Immobilization of Proteins in the Nucleolus by Ribosomal Intergenic Spacer Noncoding RNA. *Mol. Cell* 45, 147–157. doi: 10.1016/j.molcel.2011.12.012.
- Audas, T. E., Jacob, M. D., and Lee, S. (2012b). The nucleolar detention pathway: A cellular strategy for regulating molecular networks. *Cell Cycle* 11, 2059–2062. doi: 10.4161/cc.20140.
- Awasthi, R., Bhandari, K., and Nayyar, H. (2015). Temperature stress and redox homeostasis in agricultural crops. *Front. Environ. Sci.* 3, 1–24. doi: 10.3389/fenvs.2015.00011.
- Azevedo-Favory, J., Gaspin, C., Ayadi, L., Montacié, C., Marchand, V., Jobet, E., et al. (2021). Mapping rRNA 2'-O-methylations and identification of C/D snoRNAs in *Arabidopsis thaliana* plants. *RNA Biol.* 18, 1760–1777. doi: 10.1080/15476286.2020.1869892.
- Bah, A., and Forman-Kay, J. D. (2016). Modulation of intrinsically disordered protein function by post-translational modifications. *J. Biol. Chem.* 291, 6696–6705. doi: 10.1074/jbc.R115.695056.
- Banani, S. F., Lee, H. O., Hyman, A. A., and Rosen, M. K. (2017). *Biomolecular condensates: Organizers of cellular biochemistry*. doi: 10.1038/nrm.2017.7.
- Baniwal, S. K., Bharti, K., Chan, K. Y., Fauth, M., Arnab, G., Kotak, S., et al. (2004). Heat stress response in plants - a complex game with chaperones and more than twenty heat shock proteins. *J. Biosci* 29, 471–487.
- Barneche, F., Steinmetz, F., and Echeverría, M. (2000). Fibrillarin Genes Encode Both a Conserved Nucleolar Protein and a Novel Small Nucleolar RNA Involved in Ribosomal RNA Methylation in *Arabidopsis thaliana*. *J. Biol. Chem.* 275, 27212–27220. doi: 10.1074/jbc.M002996200.
- Beine-Golovchuk, O., Pereira Firmino, A. A., Dąbrowska, A., Stefanie, S., Erban, A., Walther, D., et al. (2018). Plant Temperature Acclimation and Growth Rely on Cytosolic Ribosome Biogenesis Factor Homologs. *Plant Physiol.* 176, 2251–2276. doi: 10.1104/pp.17.01448.
- Bertrand, E., and Fournier, M. J. (2004). The snoRNPs and Related Machines: Ancient Devices That Mediate Maturation of rRNA and Other RNAs. *Madame Curie Biosci. Database [Internet]*, 225–261.
- Bokszczanin, K. L., and Fragkostefanakis, S. (2013). Perspectives on deciphering mechanisms underlying plant heat stress response and thermotolerance. *Front. Plant Sci.* 4, 1–20. doi: 10.3389/fpls.2013.00315.
- Boncinelli, E., Borghese, A., Graziani, F., La Mantia, G., Manzi, A., Mariani, C., et al. (1983). Inheritance of the rDNA spacer in *D. melanogaster*. *MGG Mol. Gen. Genet.* 189, 370–374. doi: 10.1007/BF00325897.
- Borges, F., and Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nat Rev Mol Cell Biol.* 16, 727–741. doi: 10.1038/nrm4085.
- Boulon, S., Westman, B. J., Hutten, S., Boisvert, F. M., and Lamond, A. I. (2010). The Nucleolus under Stress. *Mol. Cell* 40, 216–227. doi: 10.1016/j.molcel.2010.09.024.
- Burgess, A. L., David, R., and Searle, I. R. (2015). Conservation of tRNA and rRNA 5-methylcytosine in the kingdom Plantae. *BMC Plant Biol.*, 1–17. doi: 10.1186/s12870-015-0580-8.
- Campbell, B. R., Song, Y., Posch, T. E., Cullis, C. A., and Town, C. D. (1992). Sequence and organization of 5S ribosomal RNA-encoding genes of *Arabidopsis thaliana*. *Gene* 112, 225–228. doi: 10.1016/0378-1119(92)90380-8.
- Camuel, A. (2020). Caractérisation Fonctionnelle des ARNr Non-Codant Impliqués dans la Rétention Nucléolaire de Protéines en Réponse au Stress chez *Arabidopsis thaliana*.
- Casañal, A., Kumar, A., Hill, C. H., Easter, A. D., Emsley, P., Degliesposti, G., et al. (2017). Architecture of eukaryotic mRNA 3'-end processing machinery. *Science (80-.)*. 358, 1056–1059. doi: 10.1126/science.aao6535.

- Cerdido, A., Medina, F. J., Csic, C. D. I. B., and Madrid, E.- (1995). Subnucleolar location of fibrillarin and variation in its levels during the cell cycle and during differentiation of plant cells. 36, 625–634.
- Chandrasekhara, C., Mohannath, G., Blevins, T., Pontvianne, F., and Pikaard, C. S. (2016). Chromosome-specific NOR inactivation explains selective rRNA gene silencing and dosage control in Arabidopsis. *Genes Dev.* 30, 177–190. doi: 10.1101/gad.273755.115.
- Chen, F. X., Smith, E. R., and Shilatifard, A. (2018). Born to run: Control of transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* 19, 464–478. doi: 10.1038/s41580-018-0010-5.
- Chory, J., and Peto, C. A. (1990). Mutations in the DET1 gene affect cell-type-specific expression of light-regulated genes and chloroplast development in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* 87, 8776–8780. doi: 10.1073/pnas.87.22.8776.
- Chung, C. T., Niemela, S. L., and Miller, R. H. (1989). One-step preparation of competent Escherichia coli: Transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. U. S. A.* 86, 2172–2175. doi: 10.1073/pnas.86.7.2172.
- Cinar, H., Fetahaj, Z., Cinar, S., Vernon, R. M., Chan, H. S., and Winter, R. H. A. (2019). Temperature, Hydrostatic Pressure, and Osmolyte Effects on Liquid–Liquid Phase Separation in Protein Condensates: Physical Chemistry and Biological Implications. *Chem. - A Eur. J.* 25, 13049–13069. doi: 10.1002/chem.201902210.
- Clarke, S. M., Mur, L. A. J., Wood, J. E., and Scott, I. M. (2004). Salicylic acid dependent signaling promotes basal thermotolerance but is not essential for acquired thermotolerance in Arabidopsis thaliana. *Plant J.* 38, 432–447. doi: 10.1111/j.1365-3113X.2004.02054.x.
- Cloix, C., Tutois, S., Mathieu, O., Cuvillier, C., Espagnol, M. C., Picard, G., et al. (2000). Analysis of 5S rDNA Arrays in Arabidopsis thaliana : Physical Mapping and Chromosome-Specific Polymorphisms. *Genome Res.* 10, 679–690. doi: 10.1101/gr.10.5.679.
- Cloix, C., Yukawa, Y., Tutois, S., Sugiura, M., Tourmente, S., Pascal, Â. B., et al. (2003). In vitro analysis of the sequences required for transcription of the Arabidopsis thaliana 5S rRNA genes. *Plant J.* 35, 251–261. doi: 10.1046/j.1365-3113X.2003.01793.x.
- Comella, P., Pontvianne, F., Lahmy, S., Vignols, F., Barbezier, N., DeBures, A., et al. (2008). Characterization of a ribonuclease III-like protein required for cleavage of the pre-rRNA in the 3'ETS in Arabidopsis. *Nucleic Acids Res.* 36, 1163–1175. doi: 10.1093/nar/gkm1130.
- Cong, R., Das, S., Ugrinova, I., Kumar, S., Mongelard, F., Wong, J., et al. (2012). Interaction of nucleolin with ribosomal RNA genes and its role in RNA polymerase I transcription. *Nucleic Acids Res.* 40, 9441–9454. doi: 10.1093/nar/gks720.
- Copenhaver, G. P., and Pikaard, C. S. (1996). RFLP and physical mapping with an rDNA-specific endonuclease reveals that nucleolus organizer regions of Arabidopsis thaliana adjoin the telomeres on chromosomes 2 and 4. *Plant J.* 9, 259–272. doi: 10.1046/j.1365-3113X.1996.09020259.x.
- Copenhaver, G. R., Doelling, J. H., Gens, J. S., and Pikaard, C. S. (1995). Use of RFLPs larger than 100 kbp to map the position and internal organization of the nucleolus organizer region on chromosome 2 in Arabidopsis thaliana. 7, 273–286.
- Cozzarelli, N. R., Gerrard, S. P., Schlissel, M., Brown, D. D., and Bogenhagen, D. F. (1983). Purified RNA polymerase III accurately and efficiently terminates transcription of 5S RNA genes. *Cell* 34, 829–835. doi: 10.1016/0092-8674(83)90540-8.
- Czarniecka-Verner, E., Pan, S., Salem, T., and Gurley, W. B. (2004). Plant class B HSFs inhibit transcription and exhibit affinity for TFIIB and TBP. *Plant Mol. Biol.* 56, 57–75. doi: 10.1007/s11103-004-2307-3.
- Darriere, T., Jobet, E., Zavala, D., Escande, M. L., Durut, N., De Bures, A., et al. (2022). Upon heat stress processing of ribosomal RNA precursors into mature rRNAs is compromised after cleavage at primary P site in Arabidopsis thaliana. *RNA Biol.* 19, 719–734. doi: 10.1080/15476286.2022.2071517.
- Dash, S., Lamb, M. C., Lange, J. J., McKinney, M. C., Tsuchiya, D., Guo, F., et al. (2023). rRNA transcription is integral to phase separation and maintenance of nucleolar structure. *PLoS Genet.* 19, 1–24. doi: 10.1371/journal.pgen.1010854.

- Delorme-Hinoux, V., Mbodj, A., Brando, S., De Bures, A., Llauro, C., Covato, F., et al. (2023). 45S rDNA Diversity In Natura as One Step towards Ribosomal Heterogeneity in *Arabidopsis thaliana*. *Plants* 12, 1–21. doi: 10.3390/plants12142722.
- Deltour, R., Motte, P., Liege, U. De, Vgdtale, S. D. M., Botanique, D. De, Bat, B., et al. (1990). The nucleolonema of plant and animal cells: a comparison. 5–11.
- Devireddy, A. R., Tschaplinski, T. J., Tuskan, G. A., Muchero, W., and Chen, J. G. (2021). Role of reactive oxygen species and hormones in plant responses to temperature changes. *Int. J. Mol. Sci.* 22. doi: 10.3390/ijms22168843.
- Doelling, J. E. D. H., Gaudino, R. J., Varner, C. J. E., and Pikaard, C. S. (1993). Functional analysis of *Arabidopsis thaliana* rRNA gene and spacer promoters in vivo and by transient expression. *Proc. Natl. Acad. Sci. USA* 90, 7528–7532. doi: 10.1073/pnas.90.16.7528.
- Duncan, S., Olsson, T. S. G., Hartley, M., Dean, C., and Rosa, S. (2016). A method for detecting single mRNA molecules in *Arabidopsis thaliana*. *Plant Methods* 12. doi: 10.1186/s13007-016-0114-x.
- Durut, N., Abou-Ellail, M., Pontvianne, F., Das, S., Kojima, H., Ukai, S., et al. (2014). A duplicated NUCLEOLIN gene with antagonistic activity is required for chromatin organization of silent 45S rDNA in *Arabidopsis*. *Plant Cell* 26, 1330–1344. doi: 10.1105/tpc.114.123893.
- Earley, K., Lawrence, R. J., Pontes, O., Reuther, R., Enciso, A. J., Silva, M., et al. (2006). Erasure of histone acetylation by *Arabidopsis* HDA6 mediates large-scale gene silencing in nucleolar dominance. *Genes Dev.* 20, 1283–1293. doi: 10.1101/gad.1417706.
- Echevarría-Zomeño, S., Fernández-Calvino, L., Castro-Sanz, A. B., López, J. A., Vázquez, J., and Castellano, M. M. (2016). Dissecting the proteome dynamics of the early heat stress response leading to plant survival or death in *Arabidopsis*. *Plant Cell Environ.* 39, 1264–1278. doi: 10.1111/pce.12664.
- Emenecker, R. J., Holehouse, A. S., and Strader, L. C. (2020). Emerging Roles for Phase Separation in Plants. *Dev. Cell* 55, 69–83. doi: 10.1016/j.devcel.2020.09.010.
- Evrard, A., Kumar, M., Lecourieux, D., Lucks, J., von Koskull-Döring, P., and Hirt, H. (2013). Regulation of the heat stress response in *arabidopsis* by MPK6-targeted phosphorylation of the heat stress factor HsfA2. *PeerJ* 1, e59. doi: 10.7717/peerj.59.
- Falahati, H., and Haji-Akbari, A. (2019). Thermodynamically driven assemblies and liquid-liquid phase separations in biology. *Soft Matter* 15, 1135–1154. doi: 10.1039/c8sm02285b.
- Falahati, H., Pelham-Webb, B., Blythe, S., and Wieschaus, E. (2016). Nucleation by rRNA dictates the precision of nucleolus assembly. *Curr. Biol.* 26, 277–285. doi: 10.1016/j.cub.2015.11.065.
- Firmansyah, and Argosubekti, N. (2020). A review of heat stress signaling in plants. in *IOP Conference Series: Earth and Environmental Science* doi: 10.1088/1755-1315/484/1/012041.
- Fortunato, S., Lasorella, C., Dipierro, N., Vita, F., and de Pinto, M. C. (2023). Redox Signaling in Plant Heat Stress Response. *Antioxidants* 12. doi: 10.3390/antiox12030605.
- Fukudome, A., Singh, J., Mishra, V., Reddem, E., Martinez-Marquez, F., Wenzel, S., et al. (2021). Structure and RNA template requirements of *Arabidopsis* RNA-DEPENDENT RNA POLYMERASE 2. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2115899118. doi: 10.1073/pnas.2115899118.
- Fultz, D., McKinlay, A., Enganti, R., and Pikaard, C. S. (2023). Sequence and epigenetic landscapes of active and silent nucleolus organizer regions in *Arabidopsis*. *Sci. Adv.* 9, eadj4509. doi: 10.1126/sciadv.adj4509.
- Gallagher, J. E. G., Dunbar, D. A., Granneman, S., Mitchell, B. M., Osheim, Y., Beyer, A. L., et al. (2004). RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components. *Genes Dev.* 18, 2507–2517. doi: 10.1101/gad.1226604.
- Gallardo, P., Real-Calderón, P., Flor-Parra, I., Salas-Pino, S., and Daga, R. R. (2020). Acute Heat Stress Leads to Reversible Aggregation of Nuclear Proteins into Nucleolar Rings in Fission Yeast. *Cell Rep.* 33, 108377. doi: 10.1016/j.celrep.2020.108377.
- Gangappa, S. N., and Kumar, S. V. (2018). DET1 and COP1 Modulate the Coordination of Growth and Immunity in Response to Key Seasonal Signals in *Arabidopsis*. *Cell Rep.* 25, 29–37.e3. doi: 10.1016/j.celrep.2018.08.096.

- Giesguth, M., Sahm, A., Simon, S., and Dietz, K. J. (2015). Redox-dependent translocation of the heat shock transcription factor AtHSFA8 from the cytosol to the nucleus in *Arabidopsis thaliana*. *FEBS Lett.* 589, 718–725. doi: 10.1016/j.febslet.2015.01.039.
- Girbig, M., Misiaszek, A. D., Vorländer, M. K., Lafita, A., Grötsch, H., Baudin, F., et al. (2021). Cryo-EM structures of human RNA polymerase III in its unbound and transcribing states. *Nat. Struct. Mol. Biol.* 28, 210–219. doi: 10.1038/s41594-020-00555-5.
- Goessens, G. (1984). *Nucleolar Structure*. doi: 10.1016/S0074-7696(08)62441-9.
- Goto, C., Hashizume, S., Fukao, Y., Hara-Nishimura, I., and Tamura, K. (2019). Comprehensive nuclear proteome of *Arabidopsis* obtained by sequential extraction. *Nucleus* 10, 81–92. doi: 10.1080/19491034.2019.1603093.
- Gruendler, P., Unfried, I., Pascher, K., and Schweizer, D. (1991). rDNA intergenic region from *Arabidopsis thaliana* structural analysis, intraspecific variation and functional implications. *J. Mol. Biol.* 221, 1209–1222. doi: 10.1016/0022-2836(91)90929-Z.
- Guan, Q., Lu, X., Zeng, H., Zhang, Y., and Zhu, J. (2013). Heat stress induction of miR398 triggers a regulatory loop that is critical for thermotolerance in *Arabidopsis*. *Plant J.* 74, 840–851. doi: 10.1111/tpj.12169.
- Guilfoyle, T. J., and Dietrich, M. A. (1987). “Plant RNA polymerases: structures, regulation, and genes,” in Bruening, G., Harada, J., Kosuge, T., Hollaender, A., Kuny, G., Wilson, C.M. (eds) *Tailoring Genes for Crop Improvement. Basic life sciences* (Springer), 87–100. doi: 10.1007/978-1-4684-5329-4_8.
- Guillen-Chable, F., Bayona, A., Rodríguez-Zapata, L. C., and Castano, E. (2021). Phase separation of intrinsically disordered nucleolar proteins relate to localization and function. *Int. J. Mol. Sci.* 22. doi: 10.3390/ijms222313095.
- Gull, A., Lone, A. A., and Wani, N. U. I. (2019). Biotic and Abiotic Stresses in Plants [Internet]. *Abiotic Biot. Stress Plants. toIntechOpen*. doi: 10.5772/intechopen.85832.
- Hammoudi, V., Beerens, B., Jonker, M. J., Helderma, T. A., Vlachakis, G., Giesbers, M., et al. (2021). The protein modifier SUMO is critical for integrity of the *Arabidopsis* shoot apex at warm ambient temperatures. *J. Exp. Bot.* 72, 7531–7548. doi: 10.1093/jxb/erab262.
- Han, D., Yu, Z., Lai, J., and Yang, C. (2022). Post-translational modification: a strategic response to high temperature in plants. *aBIOTECH* 3, 49–64. doi: 10.1007/s42994-021-00067-w.
- Havlová, K., Dvořáčková, M., Peiro, R., Abia, D., Mozgová, I., Vansáčová, L., et al. (2016). Variation of 45S rDNA intergenic spacers in *Arabidopsis thaliana*. *Plant Mol. Biol.* 92, 457–471. doi: 10.1007/s11103-016-0524-1.
- Hayashi, K., and Matsunaga, S. (2019). Heat and chilling stress induce nucleolus morphological changes. *J. Plant Res.* 132, 395–403. doi: 10.1007/s10265-019-01096-9.
- Hendrix, S., Dard, A., Meyer, A. J., and Reichheld, J. P. (2023). Redox-mediated responses to high temperature in plants. *J. Exp. Bot.* 74, 2489–2507. doi: 10.1093/jxb/erad053.
- Hewezi, T., Léger, M., and Gentzbittel, L. (2008). A comprehensive analysis of the combined effects of high light and high temperature stresses on gene expression in sunflower. *Ann. Bot.* 102, 127–140. doi: 10.1093/aob/mcn071.
- Hozák, P., Cook, P. R., Schöfer, C., Mosgöller, W., and Wachtler, F. (1994). Site of transcription of ribosomal RNA and intranucleolar structure in HeLa cells. *J. Cell Sci.* 107, 639–648. doi: 10.1242/jcs.107.2.639.
- Huang, J., Zhao, X., Bürger, M., Wang, Y., and Chory, J. (2021a). Two interacting ethylene response factors regulate heat stress response. *Plant Cell* 33, 338–357. doi: 10.1093/plcell/koaa026.
- Huang, K., Wu, X. X., Fang, C. L., Xu, Z. G., Zhang, H. W., Gao, J., et al. (2021b). Pol IV and RDR2: A two-RNA-polymerase machine that produces double-stranded RNA. *Science (80-.).* 374, 1579–1586. doi: 10.1126/science.abj9184.
- Huang, Y. C., Niu, C. Y., Yang, C. R., and Jinn, T. L. (2016). The heat stress factor HSFA6b connects ABA signaling and ABA-mediated heat responses. *Plant Physiol.* 172, 1182–1199. doi: 10.1104/pp.16.00860.
- Ikedo, M., Mitsuda, N., and Ohme-Takagi, M. (2011). *Arabidopsis* HsfB1 and HsfB2b act as repressors of the expression of heat-inducible Hsfs but positively regulate the acquired thermotolerance. *Plant Physiol.* 157, 1243–1254. doi: 10.1104/pp.111.179036.

- Imamura, S., Hanaoka, M., and Tanaka, K. (2008). The plant-specific TFIIIB-related protein, pBrp, is a general transcription factor for RNA polymerase I. *EMBO J.* 27, 2317–2327. doi: 10.1038/emboj.2008.151.
- Jacob, M. D., Audas, T. E., Mullineux, S. T., and Lee, S. (2012). Where no RNA polymerase has gone before: Novel functional transcripts derived from the ribosomal intergenic spacer. *Nucleus* 3, 315–319. doi: 10.4161/nucl.20585.
- Jacob, M. D., Audas, T. E., Uniacke, J., Trinkle-Mulcahy, L., and Lee, S. (2013). Environmental cues induce a long noncoding RNA-dependent remodeling of the nucleolus. *Mol. Biol. Cell* 24, 2943–2953. doi: 10.1091/mbc.E13-04-0223.
- James, A., Wang, Y., Raje, H., Rosby, R., and DiMario, P. (2014). Nucleolar stress with and without p53. *Nucleus* 5, 402–426. doi: 10.4161/nucl.32235.
- Janda, T., Prerostová, S., Vanková, R., and Darkó, É. (2021). Crosstalk between light- and temperature-mediated processes under cold and heat stress conditions in plants. *Int. J. Mol. Sci.* 22. doi: 10.3390/ijms22168602.
- Jeon, Y., Park, Y. J., Cho, H. K., Jung, H. J., Ahn, T. K., Kang, H., et al. (2015). The nucleolar GTPase nucleostemin-like 1 plays a role in plant growth and senescence by modulating ribosome biogenesis. *J. Exp. Bot.* 66, 6297–6310. doi: 10.1093/jxb/erv337.
- Jiang, C., Bi, Y., Mo, J., Zhang, R., Qu, M., Feng, S., et al. (2020). Proteome and transcriptome reveal the involvement of heat shock proteins and antioxidant system in thermotolerance of *Clematis florida*. *Sci. Rep.* 10, 1–13. doi: 10.1038/s41598-020-65699-2.
- Kalinina, N. O., Makarova, S., Makhotenko, A., Love, A. J., and Taliansky, M. (2018). The Multiple Functions of the Nucleolus in Plant Development, Disease and Stress Responses. *Front. Plant Sci.* 9, 132. doi: 10.3389/fpls.2018.00132.
- Kang, H., and Xu, T. (2023). N6-methyladenosine RNA methylation modulates liquid–liquid phase separation in plants. *Plant Cell* 35, 3205–3213. doi: <https://doi.org/10.1093/plcell/koad103>.
- Kawach, R. (2019). Characterization of IGS-rRNAs induced in response to heat stress.
- Kawauchi, J., Mischo, H., Braglia, P., Rondon, A., and Proudfoot, N. J. (2008). Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. *Genes Dev.* 22, 1082–1092. doi: 10.1101/gad.463408.
- Keller, M., Simm, S., Bokszezanin, K. L., Bostan, H., Bovy, A., Chaturvedi, P., et al. (2018). The coupling of transcriptome and proteome adaptation during development and heat stress response of tomato pollen. *BMC Genomics* 19, 1–20. doi: 10.1186/s12864-018-4824-5.
- Khan, A., Garbelli, A., Grossi, S., Florentin, A., Batelli, G., Acuna, T., et al. (2014). The Arabidopsis STRESS RESPONSE SUPPRESSOR DEAD-box RNA helicases are nucleolar- and chromocenter-localized proteins that undergo stress-mediated relocalization and are involved in epigenetic gene silencing. *Plant J.* 79, 28–43. doi: 10.1111/tpj.12533.
- Kiss, T. (2002). Small nucleolar RNAs: An abundant group of noncoding RNAs with diverse cellular functions. *Cell* 109, 145–148. doi: 10.1016/S0092-8674(02)00718-3.
- Knibiehler, B., Mirre, C., Navarro, A., and Rosset, R. (1984). Studies on chromatin organization in a nucleolus without fibrillar centres. Presence of a sub-nucleolar structure in KCo cells of *Drosophila*. *Cell Tissue Res* 236, 279–288. doi: 10.1007/BF00214228.
- Knibiehler, B., Mirre, C., and Rosset, R. (1982). Nucleolar organizer structure and activity in a nucleolus without fibrillar centres: the nucleolus in an established *Drosophila* cell line. *Sci. J. Cell* 57, 351–364. doi: 10.1242/jcs.57.1.351.
- Kojima, H., Suzuki, T., Kato, T., Enomoto, K. I., Sato, S., Kato, T., et al. (2007). Sugar-inducible expression of the nucleolin-1 gene of *Arabidopsis thaliana* and its role in ribosome synthesis, growth and development. *Plant J.* 49, 1053–1063. doi: 10.1111/j.1365-313X.2006.03016.x.
- Kojima, K., Tamura, J., Chiba, H., Fukada, K., Tsukaya, H., and Horiguchi, G. (2018). Two nucleolar proteins, GDP1 and OLI2, function as ribosome biogenesis factors and are preferentially involved in promotion of leaf cell proliferation without strongly affecting leaf adaxial–abaxial patterning in *Arabidopsis thaliana*. *Front. Plant Sci.* 8, 1–15. doi: 10.3389/fpls.2017.02240.
- Kolovos, P., Knoch, T. A., Grosveld, F. G., Cook, P. R., and Papantonis, A. (2012). Enhancers and silencers: An

- integrated and simple model for their function. *Epigenetics and Chromatin* 5, 1–8. doi: 10.1186/1756-8935-5-1.
- Koroleva, O. A., Brown, J. W. S., and Shaw, P. J. (2009). Localization of eIF4A-III in the nucleolus and splicing speckles is an indicator of plant stress. *Plant Signal. Behav.* 4, 1148–1151. doi: 10.1105/tpc.108.060434.he.
- Korotko, U., Chwiałkowska, K., Sańko-Sawcenko, I., and Kwasniewski, M. (2021). DNA demethylation in response to heat stress in arabidopsis thaliana. *Int. J. Mol. Sci.* 22, 1–20. doi: 10.3390/ijms22041555.
- Kotak, S., Larkindale, J., Lee, U., von Koskull-Döring, P., Vierling, E., and Scharf, K. D. (2007). Complexity of the heat stress response in plants. *Curr. Opin. Plant Biol.* 10, 310–316. doi: 10.1016/j.pbi.2007.04.011.
- Kuehn, M., and Arnheim, N. (1983). Nucleotide sequence of the genetically labile repeated elements 5' to the origin of mouse rRNA transcription. *Nucleic Acids Res.* 11, 211–224. doi: 10.1093/nar/11.1.211.
- Laemmli, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* 227, 680–685.
- Lafontaine, D. L. J., Riback, J. A., Bascetin, R., and Brangwynne, C. P. (2021). The nucleolus as a multiphase liquid condensate. *Nat. Rev. Mol. Cell Biol.* 22, 165–182. doi: 10.1038/s41580-020-0272-6.
- Lam, Y. W., and Trinkle-Mulcahy, L. (2015). New insights into nucleolar structure and function. *F1000Prime Rep.* 7, 48. doi: 10.12703/P7-48.
- Lämke, J., Brzezinka, K., Altmann, S., and Bäurle, I. (2016). A hit-and-run heat shock factor governs sustained histone methylation and transcriptional stress memory. *EMBO J.* 35, 162–175. doi: 10.15252/embj.201592593.
- Lange, H., Sement, F. M., and Gagliardi, D. (2011). MTR4, a putative RNA helicase and exosome co-factor, is required for proper rRNA biogenesis and development in Arabidopsis thaliana. *Plant J.* 68, 51–63. doi: 10.1111/j.1365-313X.2011.04675.x.
- Larkin, R. M., Alonso, J. M., Ecker, J. R., and Chory, J. (2003). GUN4, a regulator of chlorophyll synthesis and intracellular signaling. *Science (80-.)*. 299, 902–906. doi: 10.1126/science.1079978.
- Lewicki, M. C., Srikumar, T., Johnson, E., and Raught, B. (2015). The *S. cerevisiae* SUMO stress response is a conjugation-deconjugation cycle that targets the transcription machinery. *J. Proteomics* 118, 39–48. doi: 10.1016/j.jprot.2014.11.012.
- Li, J., Cao, Y., Zhang, J., Zhu, C., Tang, G., and Yan, J. (2023). The miR165/166–PHABULOSA module promotes thermotolerance by transcriptionally and posttranslationally regulating HSFA1. *Plant Cell* 35, 2952–2971. doi: 10.1093/plcell/koad121.
- Li, S., Liu, J., Liu, Z., Li, X., Wu, F., and He, Y. (2014). HEAT-INDUCED TAS1 TARGET1 mediates thermotolerance via heat stress transcription factor A1a-directed pathways in arabidopsis. *Plant Cell* 26, 1764–1780. doi: 10.1105/tpc.114.124883.
- Lippmann, R., Babben, S., Menger, A., Delker, C., and Quint, M. (2019). Development of Wild and Cultivated Plants under Global Warming Conditions. *Curr. Biol.* 29, R1326–R1338. doi: 10.1016/j.cub.2019.10.016.
- Liu, H. T., Gao, F., Li, G. L., Han, J. L., Liu, D. L., Sun, D. Y., et al. (2008). The calmodulin-binding protein kinase 3 is part of heat-shock signal transduction in Arabidopsis thaliana. *Plant J.* 55, 760–773. doi: 10.1111/j.1365-313X.2008.03544.x.
- Liu, Q., and Dreyfuss, G. (1995). In Vivo and In Vitro Arginine Methylation of RNA-Binding Proteins. *Mol. Cell. Biol.* 15, 2800–2808. doi: 10.1128/mcb.15.5.2800.
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262.
- Londoño Vélez, V., Alquraish, F., Tarbiyyah, I., Rafique, F., Mao, D., and Chodasiewicz, M. (2022). Landscape of biomolecular condensates in heat stress responses. *Front. Plant Sci.* 13, 1032045. doi: 10.3389/fpls.2022.1032045 OPEN.
- Louis, J., Mondal, H. A., and Shah, J. (2012). Green peach aphid infestation induces Arabidopsis PHYTOALEXIN-DEFICIENT4 expression at site of insect feeding. *Plant Signal. Behav.* 7, 1431–1433. doi: 10.4161/psb.22088.
- Ma, X., Zhao, F., and Zhou, B. (2022). The Characters of Non-Coding RNAs and Their Biological Roles in Plant

- Maekawa, S., Ishida, T., and Yanagisawa, S. (2018). Reduced expression of APUM24, encoding a novel rRNA processing factor, induces sugar-dependent nucleolar stress and altered sugar responses in *Arabidopsis thaliana*. *Plant Cell* 30, 209–227. doi: 10.1105/tpc.17.00778.
- Maekawa, S., and Yanagisawa, S. (2021). Ribosome biogenesis factor oli2 and its interactor brx1-2 are associated with morphogenesis and lifespan extension in *Arabidopsis thaliana*. *Plant Biotechnol.* 38, 117–125. doi: 10.5511/plantbiotechnology.20.1224a.
- Mahmood, T., and Yang, P. C. (2012). Western blot: Technique, theory, and trouble shooting. *N. Am. J. Med. Sci.* 4, 429–434. doi: 10.4103/1947-2714.100998.
- Makarov, V., Rakitina, D., Protopopova, A., Yaminsky, I., Arutiunian, A., Love, A. J., et al. (2013). Plant Coilin: Structural Characteristics and RNA-Binding Properties. *PLoS One* 8, e53571. doi: 10.1371/journal.pone.0053571.
- Marasco, M., Li, W., Lynch, M., and Pikaard, C. S. (2017). Catalytic properties of RNA polymerases IV and V: Accuracy, nucleotide incorporation and rNTP/dNTP discrimination. *Nucleic Acids Res.* 45, 11315–11326. doi: 10.1093/nar/gkx794.
- Martinez-Seidel, F., Beine-Golovchuk, O., Hsieh, Y. C., and Kopka, J. (2020). Systematic Review of Plant Ribosome Heterogeneity and Specialization. *Front. Plant Sci.* 11, 948. doi: 10.3389/fpls.2020.00948.
- Maruri-López, I., Figueroa, N. E., Hernández-Sánchez, I. E., and Chodasiewicz, M. (2021). Plant Stress Granules: Trends and Beyond. *Front. Plant Sci.* 12, 722643. doi: 10.3389/fpls.2021.722643.
- Meinke, D., and Koornneef, M. (1997). Community Standards - a New Series of Guidelines for Plant Science - Community Standards for *Arabidopsis* Genetics. *Plant J.* 12, 247–253.
- Mekhail, K., Gunaratnam, L., Bonicalzi, M. E., and Lee, S. (2004). HIF activation by pH-dependent nucleolar sequestration of VHL. *Nat. Cell Biol.* 6, 642–647. doi: 10.1038/ncb1144.
- Mekhail, K., Khacho, M., Carrigan, A., Hache, R. R. J., Gunaratnam, L., and Lee, S. (2005). Regulation of ubiquitin ligase dynamics by the nucleolus. *J. Cell Biol.* 170, 733–744. doi: 10.1083/jcb.200506030.
- Mekhail, K., Rivero-Lopez, L., Al-Masri, A., Brandon, C., Khacho, M., and Lee, S. (2007). Identification of a Common Subnuclear Localization Signal. *Mol. Biol. Cell* 18, 3966–3977. doi: 10.1091/mbc.E07-03-0295.
- Mélèse, T., and Xue, Z. (1995). The nucleolus: an organelle formed by the act of building a ribosome. *Curr. Opin. Cell Biol.* 7, 319–324. doi: 10.1016/0955-0674(95)80085-9.
- Merret, R., Nagarajan, V. K., Carpentier, M. C., Park, S., Favory, J. J., Descombin, J., et al. (2015). Heat-induced ribosome pausing triggers mRNA co-translational decay in *Arabidopsis thaliana*. *Nucleic Acids Res.* 43, 4121–4132. doi: 10.1093/nar/gkv234.
- Mészáros, B., Erdős, G., and Dosztányi, Z. (2018). IUPred2A: Context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* 46, W329–W337. doi: 10.1093/nar/gky384.
- Miluzio, A., Beugnet, A., Volta, V., and Biffo, S. (2009). Eukaryotic initiation factor 6 mediates a continuum between 60S ribosome biogenesis and translation. *EMBO Rep.* 10, 459–465. doi: 10.1038/embor.2009.70.
- Missbach, S., Weis, B. L., Martin, R., Simm, S., Bohnsack, M. T., and Schleiff, E. (2013). 40S Ribosome Biogenesis Co-Factors Are Essential for Gametophyte and Embryo Development. *PLoS One* 8, 1–19. doi: 10.1371/journal.pone.0054084.
- Mittag, T., and Parker, R. (2018). Multiple Modes of Protein–Protein Interactions Promote RNP Granule Assembly. *J. Mol. Biol.* 430, 4636–4649. doi: 10.1016/j.jmb.2018.08.005.
- Mohannath, G., Pontvianne, F., and Pikaard, C. S. (2016). Selective nucleolus organizer inactivation in *Arabidopsis* is a chromosome position-effect phenomenon. *Proc. Natl. Acad. Sci. U. S. A.* 113, 13426–13431. doi: 10.1073/pnas.1608140113.
- Montacié, C. (2019). Le protéasome et le fer : Rôles et/ou régulations dans le nucléole d’*Arabidopsis thaliana* Soutenue.
- Mougey, E. B., O’Reilly, M., Osheim, Y., Miller, O. L., Beyer, A., and Sollner-Webb, B. (1993). The terminal balls characteristic of eukaryotic rRNA transcription units in chromatin spreads are rRNA processing complexes. *Genes*

- Muñoz-Díaz, E., and Sáez-Vásquez, J. (2022). Nuclear dynamics: Formation of bodies and trafficking in plant nuclei. *Front. Plant Sci.* 13, 984163. doi: 10.3389/fpls.2022.984163.
- Nachtergaele, S., and He, C. (2017). The emerging biology of RNA post-transcriptional modifications. *RNA Biol.* 14, 156–163. doi: 10.1080/15476286.2016.1267096.
- Nardozi, J. D., Lott, K., and Cingolani, G. (2010). Phosphorylation meets nuclear import: a review. *Cell Commun. Signal.* 8, 32. doi: 10.1186/1478-811X-8-32.
- Nazar, R. N. (2004). Ribosomal RNA processing and ribosome biogenesis in eukaryotes. *IUBMB Life* 56, 457–465. doi: 10.1080/15216540400010867.
- Németh, A., Conesa, A., Santoyo-Lopez, J., Medina, I., Montaner, D., Péterfia, B., et al. (2010). Initial genomics of the human nucleolus. *PLoS Genet.* 6. doi: 10.1371/journal.pgen.1000889.
- Németh, A., Perez-Fernandez, J., Merkl, P., Hamperl, S., Gerber, J., Griesenbeck, J., et al. (2013). RNA polymerase I termination: Where is the end? *Biochim. Biophys. Acta - Gene Regul. Mech.* 1829, 306–317. doi: 10.1016/j.bbagr.2012.10.007.
- Nover, L., Bharti, K., Döring, P., Mishra, S. K., Ganguli, A., and Scharf, K. D. (2001). Arabidopsis and the heat stress transcription factor world: How many heat stress transcription factors do we need? *Cell Stress Chaperones* 6, 177–189. doi: 10.1379/1466-1268(2001)006<0177:aathst>2.0.co;2.
- Ohbayashi, I., Lin, C. Y., Shinohara, N., Matsumura, Y., Machida, Y., Horiguchi, G., et al. (2017). Evidence for a role of ANAC082 as a ribosomal stress response mediator leading to growth defects and developmental alterations in arabidopsis. *Plant Cell* 29, 2644–2660. doi: 10.1105/tpc.17.00255.
- Ohbayashi, I., and Sugiyama, M. (2018). Plant nucleolar stress response, a new face in the NAC-dependent cellular stress responses. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.02247.
- Palm, D., Simm, S., Darm, K., Weis, B. L., Ruprecht, M., Schleiff, E., et al. (2016). Proteome distribution between nucleoplasm and nucleolus and its relation to ribosome biogenesis in Arabidopsis thaliana. *RNA Biol.* 13, 441–454. doi: 10.1080/15476286.2016.1154252.
- Panda, K., McCue, A. D., and Slotkin, R. K. (2020). Arabidopsis RNA Polymerase IV generates 21–22 nucleotide small RNAs that can participate in RNA-directed DNA methylation and may regulate genes. *Philos. Trans. R. Soc. B Biol. Sci.* 375. doi: 10.1098/rstb.2019.0417.
- Pelham, H. R. B. (1984). Hsp70 accelerates the after heat shock. *EMBO J.* 3, 3095–3100.
- Pérez-Salamó, I., Papdi, C., Rigó, G., Zsigmond, L., Vilela, B., Lumbreras, V., et al. (2014). The heat shock factor A4A confers salt tolerance and is regulated by oxidative stress and the mitogen-activated protein kinases MPK3 and MPK6. *Plant Physiol.* 165, 319–334. doi: 10.1104/pp.114.237891.
- Petricka, J. J., and Nelson, T. M. (2007). Arabidopsis nucleolin affects plant development and patterning. *Plant Physiol.* 144, 173–186. doi: 10.1104/pp.106.093575.
- Phipps, K. R., Charette, J. M., and Baserga, S. J. (2011). The small subunit processome in ribosome biogenesis-progress and prospects. *Wiley Interdiscip. Rev. RNA* 2, 1–21. doi: 10.1002/wrna.57.
- Picart-Piccolo, A., Picart, C., Picault, N., and Pontvianne, F. (2020). Nucleolus-associated chromatin domains are maintained under heat stress, despite nucleolar reorganization in Arabidopsis thaliana. *J. Plant Res.* 133, 463–470. doi: 10.1007/s10265-020-01201-3.
- Picart, C., and Pontvianne, F. (2017). Plant nucleolar DNA : Green light shed on the role of Nucleolin in genome organization. *Nucleus* 8, 11–16. doi: 10.1080/19491034.2016.1236167.
- Pih, K. T., Yi, M. J., Liang, Y. S., Shin, B. J., Cho, M. J., Hwang, I., et al. (2000). Molecular Cloning and Targeting of a Fibrillar Homolog from Arabidopsis. *Plant Physiol.* 123, 51–58. doi: 10.1104/pp.123.1.51.
- Pirogov, S. A., Gvozdev, V. A., and Klenov, M. S. (2019). Long noncoding RNAs and stress response in the nucleolus. *Cells* 8. doi: 10.3390/cells8070668.
- Pontvianne, F., Abou-Ellail, M., Douet, J., Comella, P., Matia, I., Chandrasekhara, C., et al. (2010). Nucleolin is required

for DNA methylation state and the expression of rRNA gene variants in *Arabidopsis thaliana*. *PLoS Genet.* 6, 1–13. doi: 10.1371/journal.pgen.1001225.

- Pontvianne, F., Blevins, T., Chandrasekhara, C., Feng, W., Stroud, H., Jacobsen, S. E., et al. (2012). Histone methyltransferases regulating rRNA gene dose and dosage control in *Arabidopsis*. *Genes Dev.* 26, 945–957. doi: 10.1101/gad.182865.111.
- Pontvianne, F., Carpentier, M. C., Durut, N., Pavlišťová, V., Jaške, K., Schořová, Š., et al. (2016). Identification of Nucleolus-Associated Chromatin Domains Reveals a Role for the Nucleolus in 3D Organization of the *A. thaliana* Genome. *Cell Rep.* 16, 1574–1587. doi: 10.1016/j.celrep.2016.07.016.
- Pontvianne, F., Matía, I., Douet, J., Sylvette Tourmente, F. J. M., Echeverria, M., and Saéz-Vásquez, J. (2007). Characterization of AtNUC-L1 Reveals a Central Role of Nucleolin in Nucleolus Organization and Silencing of AtNUC-L2 Gene in *Arabidopsis*. *Mol. Biol. Cell* 18, 369–379. doi: 10.1091/mbc.E06-08-0751.
- Popova, O. V., Dinh, H. Q., Aufsatz, W., and Jonak, C. (2013). The RdDM pathway is required for basal heat tolerance in *Arabidopsis*. *Mol. Plant* 6, 396–410. doi: 10.1093/mp/sst023.
- Proudfoot, N. J. (2016). Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science* (80-.). 352. doi: 10.1126/science.aad9926.
- Puvion-Dutilleul, F., Puvion, E., and Bachellerie, J. P. (1997). Early stages of pre-rRNA formation within the nucleolar ultrastructure of mouse cells studied by in situ hybridization with a 5'ETS leader probe. *Chromosoma* 105, 496–505. doi: 10.1007/BF02510486.
- Qu, A. L., Ding, Y. F., Jiang, Q., and Zhu, C. (2013). Molecular mechanisms of the plant heat stress response. *Biochem. Biophys. Res. Commun.* 432, 203–207. doi: 10.1016/j.bbrc.2013.01.104.
- Rakitina, D. V., Taliansky, M., Brown, J. W. S., and Kalinina, N. O. (2011). Two RNA-binding sites in plant fibrillarin provide interactions with various RNA substrates. *Nucleic Acids Res.* 39, 8869–8880. doi: 10.1093/nar/gkr594.
- Raška, I., Koberna, K., Malinský, J., Fidlerová, H., and Mašata, M. (2004). The nucleolus and transcription of ribosomal genes. *Biol. Cell* 96, 579–594. doi: 10.1016/j.biocel.2004.04.015.
- Ream, T. S., Haag, J. R., Pontvianne, F., Nicora, C. D., Norbeck, A. D., Paša-Tolić, L., et al. (2015). Subunit compositions of *Arabidopsis* RNA polymerases I and III reveal Pol I- and Pol III-specific forms of the AC40 subunit and alternative forms of the C53 subunit. *Nucleic Acids Res.* 43, 4163–4178. doi: 10.1093/nar/gkv247.
- Ream, T. S., Haag, J. R., Wierzbicki, A. T., Nicora, C. D., Norbeck, A. D., Zhu, J. K., et al. (2009). Subunit Compositions of the RNA-Silencing Enzymes Pol IV and Pol V Reveal Their Origins as Specialized Forms of RNA Polymerase II. *Mol. Cell* 33, 192–203. doi: 10.1016/j.molcel.2008.12.015.
- Reindl, A., Schöffl, F., Schell, J., Koncz, C., and Bakó, L. (1997). Phosphorylation by a cyclin-dependent kinase modulates DNA binding of the *Arabidopsis* heat-shock transcription factor HSF1 in vitro. *Plant Physiol.* 115, 93–100. doi: 10.1104/pp.115.1.93.
- Ren, M., Qiu, S., Venglat, P., Xiang, D., Feng, L., Selvaraj, G., et al. (2011). Target of Rapamycin Regulates Development and Ribosomal RNA Expression through Kinase Domain in *Arabidopsis*. *Plant Physiol.* 155, 1367–1382. doi: 10.1104/pp.110.169045.
- Richter, A. S., Nägele, T., Grimm, B., Kaufmann, K., Schroda, M., Leister, D., et al. (2023). Retrograde signaling in plants: A critical review focusing on the GUN pathway and beyond. *Plant Commun.* 4. doi: 10.1016/j.xplc.2022.100511.
- Rogers, S. O., and Bendich, A. J. (1987). Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. *Plant Mol. Biol.* 9, 509–520. doi: 10.1007/BF00015882.
- Romanova, L., Grand, A., Zhang, L., Rayner, S., Katoku-Kikyo, N., Kellner, S., et al. (2009). Critical role of nucleostemin in pre-rRNA processing. *J. Biol. Chem.* 284, 4968–4977. doi: 10.1074/jbc.M804594200.
- Ruff, K. M., Roberts, S., Chilkoti, A., and Pappu, R. V. (2018). Advances in Understanding Stimulus-Responsive Phase Behavior of Intrinsically Disordered Protein Polymers. *J. Mol. Biol.* 430, 4619–4635. doi: 10.1016/j.jmb.2018.06.031.
- Sadowski, M., and Sarcevic, B. (2010). Mechanisms of mono- and poly-ubiquitination: Ubiquitination specificity depends

on compatibility between the E2 catalytic core and amino acid residues proximal to the lysine. *Cell Div.* 5, 1–5. doi: 10.1186/1747-1028-5-19.

- Saez-Vasquez, J., and Pikaard, C. S. (1997). Extensive purification of a putative RNA polymerase I holoenzyme from plants that accurately initiates rRNA gene transcription in vitro. *Proc. Natl. Acad. Sci. U. S. A.* 94, 11869–11874. doi: 10.1073/pnas.94.22.11869.
- Sáez-Vasquez, J., Caparros-Ruiz, D., Barneche, F., and Echeverría, M. (2004). A Plant snoRNP Complex Containing snoRNAs, Fibrillarin, and Nucleolin-Like Proteins Is Competent for both rRNA Gene Binding and Pre-rRNA Processing In Vitro. *Mol. Cell. Biol.* 24, 7284–7297. doi: 10.1128/mcb.24.16.7284-7297.2004.
- Sáez-Vásquez, J., and Delseny, M. (2019). Ribosome Biogenesis in Plants: From Functional 45S Ribosomal DNA Organization to Ribosome Assembly Factors. *Plant Cell* 31, 1945–1967. doi: 10.1105/tpc.18.00874.
- Sáez-Vásquez, J., and Medina, F. javier (2008). “The Plant Nucleolus,” in *Kader, J.C., Delseny, M., (Eds.). Adv. Bot. Res.* (San Diego (CA): Elsevier), 1–46. doi: 10.1016/S0065-2296(08)00001-3.
- Saini, K., Dwivedi, A., and Ranjan, A. (2022). High temperature restricts cell division and leaf size by coordination of PIF4 and TCP4 transcription factors. *Plant Physiol.* 190, 2380–2397. doi: 10.1093/plphys/kiac345.
- Samaha, H., Delorme, V., Pontvianne, F., Cooke, R., Delalande, F., Van Dorsselaer, A., et al. (2010). Identification of protein factors and U3 snoRNAs from a Brassica oleracea RNP complex involved in the processing of pre-rRNA. *Plant J.* 61, 383–398. doi: 10.1111/j.1365-313X.2009.04061.x.
- Sato, S. (1985). Ultrastructural localization of nucleolar material by a simple silver staining technique devised for plant cells. *J. Cell Sci.* 79, 259–269. doi: 10.1242/jcs.79.1.259.
- Scheer, U., and Hock, R. (1999). Structure and function of the nucleolus. *Curr. Opin. Cell Biol.* 11, 385–390. doi: 10.1016/S0955-0674(99)80054-4.
- Scott, M. S., Troshin, P. V., and Barton, G. J. (2011). NoD: A Nucleolar localization sequence detector for eukaryotic and viral proteins. *BMC Bioinformatics* 12, 317. doi: 10.1186/1471-2105-12-317.
- Seo, J. S., Diloknawarit, P., Park, B. S., and Chua, N. H. (2019). ELF18-INDUCED LONG NONCODING RNA 1 evicts fibrillarin from mediator subunit to enhance PATHOGENESIS-RELATED GENE 1 (PR1) expression. *New Phytol.* 221, 2067–2079. doi: 10.1111/nph.15530.
- Sequeira-mendes, J., Aragüez, I., Peiró, R., Mendez-giraldez, R., Zhang, X., Jacobsen, S. E., et al. (2014). The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of. 26, 2351–2366. doi: 10.1105/tpc.114.124578.
- Sharma, S., Yang, J., Watzinger, P., Kötter, P., and Entian, K. D. (2013). Yeast Nop2 and Rcm1 methylate C2870 and C2278 of the 25S rRNA, respectively. *Nucleic Acids Res.* 41, 9062–9076. doi: 10.1093/nar/gkt679.
- Shaw, P. J., and Jordan, E. G. (1995). The Nucleolus. *Annu. Rev. Cell Dev. Biol.* 11, 93–121. doi: 10.1146/annurev.cb.11.110195.000521.
- Shi, Y., Ke, X., Yang, X., Liu, Y., and Hou, X. (2022). Plants response to light stress. *J. Genet. Genomics* 49, 735–747. doi: 10.1016/j.jgg.2022.04.017.
- Shimoji, K., Jakovljevic, J., Tsuchihashi, K., Umeki, Y., Wan, K., Kawasaki, S., et al. (2012). Ebp2 and Brx1 function cooperatively in 60S ribosomal subunit assembly in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 40, 4574–4588. doi: 10.1093/nar/gks057.
- Silva-Sanzana, C., Celiz-Balboa, J., Garzo, E., Marcus, S. E., Parra-Rojas, J. P., Rojas, B., et al. (2019). Pectin methylesterases modulate plant homogalacturonan status in defenses against the aphid *myzus persicae*. *Plant Cell* 31, 1913–1929. doi: 10.1105/tpc.19.00136.
- Silva-Sanzana, C., Gangas, M. V., Zavala, D., and Blanco-Herrera, F. (2023). A Recipe for Success: Three Key Strategies Used by Aphids and *Pseudomonas syringae* to Colonize the Phyllosphere. *Microb. Ecol.* 85, 1–8. doi: <https://doi.org/10.1007/s00248-022-01965-2>.
- Silva-Sanzana, C., Zavala, D., Moraga, F., Herrera-Vásquez, A., and Blanco-Herrera, F. (2022). Oligogalacturonides Enhance Resistance against Aphids through Pattern-Triggered Immunity and Activation of Salicylic Acid Signaling. *Int. J. Mol. Sci.* 23. doi: 10.3390/ijms23179753.

- Sirri, V., Urcuqui-Inchima, S., Roussel, P., and Hernandez-Verdun, D. (2008). Nucleolus: the fascinating nuclear body. *Histochem. Cell Biol.* 129, 13–31. doi: 10.1007/s00418-007-0359-6.
- Solis-Miranda, J., Chodasiewicz, M., Skirycz, A., Fernie, A. R., Moschou, P. N., Bozhkov, P. V., et al. (2023). Stress-related biomolecular condensates in plants. *Plant Cell* 35, 3187–3204. doi: 10.1093/plcell/koad127.
- Song, Y., Chen, Q., Ci, D., Shao, X., and Zhang, D. (2014). Effects of high temperature on photosynthesis and related gene expression in poplar. *BMC Plant Biol.* 14, 1–20. doi: 10.1186/1471-2229-14-111.
- Stępiński, D. (2014). Functional ultrastructure of the plant nucleolus. *Protoplasma* 251, 1285–1306. doi: 10.1007/s00709-014-0648-6.
- Stief, A., Altmann, S., Hoffmann, K., Pant, B. D., Scheible, W. R., and Bäurle, I. (2014). Arabidopsis miR156 regulates tolerance to recurring environmental stress through SPL transcription factors. *Plant Cell* 26, 1792–1807. doi: 10.1105/tpc.114.123851.
- Streit, D., and Schleiff, E. (2021). The Arabidopsis 2'-O-Ribose-Methylation and Pseudouridylation Landscape of rRNA in Comparison to Human and Yeast. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.684626.
- Streit, D., Shanmugam, T., Garbelyanski, A., Simm, S., and Schleiff, E. (2020). The existence and localization of nuclear snoRNAs in arabidopsis thaliana revisited. *Plants* 9, 1–18. doi: 10.3390/plants9081016.
- Sun, A. Z., and Guo, F. Q. (2016). Chloroplast retrograde regulation of heat stress responses in plants. *Front. Plant Sci.* 7, 1–16. doi: 10.3389/fpls.2016.00398.
- Suzuki, N., Rivero, R. M., Shulaev, V., Blumwald, E., and Mittler, R. (2014). Abiotic and biotic stress combinations. *New Phytol.* 203, 32–43. doi: 10.1111/nph.12797.
- Tajrish, M. M., Tuteja, R., and Tuteja, N. (2011). Nucleolin. *Commun. Integr. Biol.* 4, 267–275. doi: 10.4161/cib.4.3.14884.
- Thiry, M., Lamaye, F., and Lafontaine, D. L. J. (2011). The nucleolus : When 2 became 3. *Nucleus* 2, 289–293. doi: 10.4161/nucl.2.4.16806.
- Thorenoor, N., and Slaby, O. (2015). Small nucleolar RNAs functioning and potential roles in cancer. *Tumor Biol.* 36, 41–53. doi: 10.1007/s13277-014-2818-8.
- Tomecki, R., Sikorski, P. J., and Zakrzewska-Placzek, M. (2017). Comparison of preribosomal RNA processing pathways in yeast, plant and human cells – focus on coordinated action of endo- and exoribonucleases. *FEBS Lett.* 591, 1801–1850. doi: 10.1002/1873-3468.12682.
- Torreira, E., Louro, J. A., Pazos, I., González-Polo, N., Gil-Carton, D., Duran, A. G., et al. (2017). The dynamic assembly of distinct RNA polymerase i complexes modulates rDNA transcription. *Elife* 6, 1–23. doi: 10.7554/eLife.20832.
- Tyagi, A. K. (2001). Plant genes and their expression. *Curr. Sci.* 80, 161–169.
- Umar, O. B., Ranti, L. A., Abdulbaki, A. S., Bola, A. L., Abdulhamid, A. K., Biola, M. R., et al. (2022). Stresses in Plants: Biotic and Abiotic [Internet]. *Curr. Trends Wheat Res. IntechOpen*. doi: 10.5772/intechopen.100501.
- Unfried, I., Stocker, U., and Gruendler, P. (1989). Nucleotide sequence of the 18S rRNA gene from Arabidopsis thaliana col0. *Nucleic Acids Res.* 17, 7513. doi: 10.1093/nar/17.18.7513.
- Unfried, I., and Gruendler, P. (1990). Nucleotide sequence of the 5.8S and 25S rRNA genes and of the internal transcribed spacers from Arabidopsis thaliana. *Nucleic Acids Res.* 18, 4011. doi: 10.1093/nar/18.13.4011.
- Valenzuela, I., and Hoffmann, A. A. (2015). Effects of aphid feeding and associated virus injury on grain crops in Australia. *Austral Entomol.* 54, 292–305. doi: 10.1111/aen.12122.
- Van Koningsbruggen, S., Gierliński, M., Schofield, P., Martin, D., Barton, G. J., Ariyurek, Y., et al. (2010). High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. *Mol. Biol. Cell* 21, 3735–3748. doi: 10.1091/mbc.E10-06-0508.
- Verma, S., Nizam, S., and Verma, P. K. (2013). Biotic and Abiotic Stress Signaling in Plants. *Stress Signal. Plants Genomics Proteomics Perspect. Vol. 1* 1, 25–50. doi: 10.1007/978-1-4614-6372-6_2.
- Vydzhak, O., Luke, B., and Schindler, N. (2020). Non-coding RNAs at the Eukaryotic rDNA Locus: RNA–DNA Hybrids

and Beyond. *J. Mol. Biol.* 432, 4287–4304. doi: 10.1016/j.jmb.2020.05.011.

- Wang, L. C., Wu, J. R., Hsu, Y. J., and Wu, S. J. (2015). Arabidopsis HIT4, a regulator involved in heat-triggered reorganization of chromatin and release of transcriptional gene silencing, relocates from chromocenters to the nucleolus in response to heat stress. *New Phytol.* 205, 544–554. doi: 10.1111/nph.13088.
- Wang, M., Tao, X., Jacob, M. D., Bennett, C. A., Ho, J. J. D., Gonzalgo, M. L., et al. (2018). Stress-Induced Low Complexity RNA Activates Physiological Amyloidogenesis. *Cell Rep.* 24, 1713–1721.e4. doi: 10.1016/j.celrep.2018.07.040.
- Wang, X., Tan, N. W. K., Chung, F. Y., Yamaguchi, N., Gan, E. S., and Ito, T. (2023). Transcriptional Regulators of Plant Adaptation to Heat Stress. *Int. J. Mol. Sci.* 24. doi: 10.3390/ijms241713297.
- Weis, B. L., Kovacevic, J., Missbach, S., and Schleiff, E. (2015). Plant-Specific Features of Ribosome Biogenesis. *Trends Plant Sci.* 20, 729–740. doi: 10.1016/j.tplants.2015.07.003.
- Weis, B. L., Missbach, S., Marzi, J., Bohnsack, M. T., and Schleiff, E. (2014). The 60S associated ribosome biogenesis factor LSG1-2 is required for 40S maturation in *Arabidopsis thaliana*. *Plant J.* 80, 1043–1056. doi: 10.1111/tbj.12703.
- Welch, W. J., and Feramisco, J. R. (1984). Nuclear and nucleolar localization of the 72,000-dalton heat shock protein in heat-shocked mammalian cells. *J. Biol. Chem.* 259, 4501–4513. doi: 10.1016/s0021-9258(17)43075-4.
- Wellauer, P. K., Reeder, R. H., Dawid, I. B., and Brown, D. D. (1976). The arrangement of length heterogeneity in repeating units of amplified and chromosomal ribosomal DNA from *Xenopus laevis*. *J. Mol. Biol.* 105, 487–505. doi: 10.1016/0022-2836(76)90230-8.
- Weng, M., Yang, Y., Feng, H., Pan, Z., Shen, W. H., Zhu, Y., et al. (2014). Histone chaperone ASF1 is involved in gene transcription activation in response to heat stress in *Arabidopsis thaliana*. *Plant, Cell Environ.* 37, 2128–2138. doi: 10.1111/pce.12299.
- Wierzbicki, A. T., Haag, J. R., and Pikaard, C. S. (2008). Noncoding Transcription by RNA Polymerase Pol IVb/Pol V Mediates Transcriptional Silencing of Overlapping and Adjacent Genes. *Cell* 135, 635–648. doi: 10.1016/j.cell.2008.09.035.
- Wilson, D. N., and Cate, J. H. D. (2012). The Structure and Function of the Eukaryotic Ribosome. *Cold Spring Harb. Perspect. Biol.* 4, a011536. doi: 10.1101/cshperspect.a011536.
- Woodruff, J. B., Hyman, A. A., and Boke, E. (2018). Organization and Function of Non-dynamic Biomolecular Condensates. *Trends Biochem. Sci.* 43, 81–94. doi: 10.1016/j.tibs.2017.11.005.
- Xu, Y., Bernecky, C., Lee, C. T., Maier, K. C., Schwalb, B., Tegunov, D., et al. (2017). Architecture of the RNA polymerase II-Paf1C-TFIIS transcription elongation complex. *Nat. Commun.* 8. doi: 10.1038/ncomms15741.
- Yang, D.-L., Huang, K., Deng, D., Zeng, Y., Wang, Z., and Zhang, Y. (2023). DNA-dependent RNA polymerases in plants. *Plant Cell* 35, 3641–3661. doi: 10.1093/plcell/koad195.
- Yang, K., Yang, J., and Yi, J. (2018). Nucleolar stress: Hallmarks, sensing mechanism and diseases. *Cell Stress* 2, 125–140. doi: 10.15698/cst2018.06.139.
- Yano, H., and Sato, S. (2000). Ultrastructural localization of transcription sites, DNA, and RNA reveals a concentric arrangement of structural and functional domains in plant nucleolonema. 129–140.
- Zakrzewska-Placzek, M., Souret, F. F., Sobczyk, G. J., Green, P. J., and Kufel, J. (2010). *Arabidopsis thaliana* XRN2 is required for primary cleavage in the pre-ribosomal RNA. *Nucleic Acids Res.* 38, 4487–4502. doi: 10.1093/nar/gkq172.
- Zhang, Y., Sikes, M. L., Beyer, A. L., and Schneider, D. A. (2009). The Paf1 complex is required for efficient transcription elongation by RNA polymerase I. *Proc. Natl. Acad. Sci. U. S. A.* 106, 2153–2158. doi: 10.1073/pnas.0812939106.
- Zhang, Y., Zhou, Y., Zhu, W., Liu, J., and Cheng, F. (2022). Non-coding RNAs fine-tune the balance between plant growth and abiotic stress tolerance. *Front. Plant Sci.* 13, 1–17. doi: 10.3389/fpls.2022.965745.
- Zhao, J., Lu, Z., Wang, L., and Jin, B. (2021). Plant responses to heat stress: Physiology, transcription, noncoding RNAs, and epigenetics. *Int. J. Mol. Sci.* 22, 1–14. doi: 10.3390/ijms22010117.

Zhao, Q., Chen, W., Bian, J., Xie, H., Li, Y., Xu, C., et al. (2018). Proteomics and phosphoproteomics of heat stress-responsive mechanisms in Spinach. *Front. Plant Sci.* 9, 1–22. doi: 10.3389/fpls.2018.00800.