



Adapting INTACT to analyse cell-type-specific transcriptomes and nucleocytoplasmic mRNA dynamics in the *Arabidopsis* embryo

Joakim Palovaara^{1,2} · Dolf Weijers¹

Received: 5 October 2018 / Accepted: 31 October 2018 / Published online: 15 November 2018
© The Author(s) 2019

Abstract

In the early embryo of vascular plants, the different cell types and stem cells of the seedling are specified as the embryo develops from a zygote towards maturity. How the key steps in cell and tissue specification are instructed by genome-wide transcriptional activity is poorly understood. Progress in defining transcriptional regulation at the genome-wide level in plant embryos has been hampered by difficulties associated with capturing cell-type-specific transcriptomes in this small and inaccessible structure. We recently adapted a two-component genetic nucleus labelling system called INTACT to isolate nuclei from distinct cell types at different stages of *Arabidopsis thaliana* embryogenesis. We have used these to generate a transcriptomic atlas of embryo development following microarray-based expression profiling. Here, we present a general description of the adapted INTACT procedure, including the two-component labelling system, seed isolation, nuclei preparation and purification, as well as transcriptomic profiling. We also compare nuclear and cellular transcriptomes from the early *Arabidopsis* embryo to assess nucleocytoplasmic differences and discuss how these differences can be used to infer regulation of gene activity.

Keywords Nuclear isolation · Cell-type-specific · Embryo · Transcriptome analysis

Introduction

With the advent of transcriptomics, plant research has gained important insights into the genetic regulatory mechanisms that underlie cell fate determination, pattern formation and cell–cell communication during plant development. This framework of developmental processes, crucial for the continuous formation of plant structures from stem cells, is first established in the early plant embryo (reviewed in Palovaara

et al. 2016). As such, much effort has been made in recent years to adapt transcriptomic approaches to this tissue at a cellular resolution (e.g. Belmonte et al. 2013; Casson et al. 2005; Slane et al. 2014). Recently, we adapted one such approach, INTACT (isolation of nuclei tagged in specific cell types), to isolate cell-type-specific nuclei from the early embryo of the flowering plant *Arabidopsis thaliana* for transcriptomic profiling (Palovaara et al. 2017).

Enrichment methods to isolate nuclei or cells based on the expression of specific promoters are very powerful, as the selection of cells for transcriptomics does not require manual selection and cells are by definition united by the expression of at least one marker gene. For the embryo, however, there is an important difficulty: genes expressed in the tissue precursors of the embryo are usually also expressed in the corresponding tissue of the seed that the embryo resides in (Belmonte et al. 2013; Palovaara et al. 2017). Thus, when using only expression of a marker gene as a selection step, transcriptomes of embryo cells would be overshadowed by the seed cells if entire seeds are used as starting material. Rather than using dissected embryos, we therefore resorted to a two-component genetic labelling system, INTACT (Deal and Henikoff 2010). INTACT is based on the selective biotinylation of a unique target peptide (biotin ligase recognition—BLRP) by the BirA biotin ligase

A contribution to the special issue ‘Cellular Omics Methods in Plant Reproduction Research’.

Communicated by Thomas Dresselhaus.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00497-018-0347-0>) contains supplementary material, which is available to authorized users.

✉ Dolf Weijers
dolf.weijers@wur.nl

¹ Laboratory of Biochemistry, Wageningen University, 6708 WE Wageningen, The Netherlands

² Present Address: Molecular Genetics, University of Bremen, 28359 Bremen, Germany

enzyme (from *Escherichia coli*). The BLRP is integrated in a nuclear targeting fusion (NTF) protein, which in addition carries a green fluorescent protein (GFP) and a nuclear lamina localization domain. Only when BirA and NTF are co-expressed in the same cells, will the latter be biotinylated. Since streptavidin has a high affinity for biotin, biotin-tagged nuclei can next be isolated from crude nuclear preparations using streptavidin-coated beads. This circumvents the need for BirA and NTF promoter expression being exclusive to specific cells or tissues in the embryo. Consequently, specificity increases since more markers become available for use. We exploited this combinatorial logic by expressing BirA from an embryo-enriched promoter and driving NTF expression from a large range of cell-type-specific promoters. This allowed us to generate a transcriptome atlas of early *Arabidopsis* embryo development using nuclei isolated from cell types necessary for root stem cell niche establishment (Palovaara et al. 2017). We identified shifts in cell-type-specific gene expression associated with the developmental stage of the embryo, and enrichment of transcription factors and biological processes important for cell fate determination. Our work provides a resource for further exploration into how gene activity shapes the formation of the first plant tissues.

Here, we present the INTACT-based approach used in Palovaara et al. (2017). We discuss two-component labeling and describe the adapted INTACT procedure, including generating microarray-based transcriptomes. In addition, we compare previously published nuclear and cellular transcriptomes from the early *Arabidopsis* embryo to illustrate how INTACT-generated data can be used to investigate nucleocytoplasmic differences in a cell-type-specific manner.

INTACT on *Arabidopsis* embryos

INTACT was initially developed for use on roots from *Arabidopsis* (Deal and Henikoff 2010, 2011), a popular model plant to study cell specification processes due to its highly invariant cell division patterns and the ease to genetically manipulate. Compared to roots, *Arabidopsis* embryos are small, contain few cells and are surrounded by the endosperm and seed coat. This affects the final yield and purity of INTACT when performed according to the original protocol (Palovaara et al. 2017). Thus, it was necessary to adapt the INTACT protocol when isolating cell-type-specific nuclei from the early *Arabidopsis* embryo.

First, we generated a codon-optimized BirA fragment to facilitate translation in *Arabidopsis*, and a collection of “gold-standard” INTACT plant lines that mark the major cell and tissue types of the early embryo. In these lines, two separate constructs express the modified BirA (mBirA) from an embryo-enriched promoter (either *pWOX2* or *pSCR*; Palovaara et al. 2017) and NTF from various cell-type-specific promoters (Fig. 1). Embryogenesis is a progressive process,

with a defined start point (fertilization). However, embryos within the same silique can differ in exact stage, which makes it difficult to capture temporal dynamics in gene expression. Therefore, we performed manual pollination to better synchronize embryo development within each silique. Optimal time intervals were determined for 16-cell (72 h), early globular (81 h) and late globular (100 h) embryo stages. Function of the modified BirA, NTF biotinylation, specificity of INTACT markers, and cell-type- and stage-specific nuclei isolation was experimentally verified. Secondly, we optimized several steps in the original INTACT protocol to improve the recovery efficiency (estimated to be 20–50%) and purity ($86.2\% \pm 6.6\%$, $n = 50$) of isolated biotin-tagged nuclei. With these changes, we could isolate 1000–5000 cell-type-specific nuclei from 100 to 200 siliques per biological replicate depending on the INTACT line used and embryonic stage.

The adapted INTACT procedure and the approach used to generate transcriptomic profiles are described below. In theory, the procedure can be applied to any *Arabidopsis* embryo cell type, using our INTACT lines or other lines with a suitable cell-type-specific promoter, and can be combined with several downstream applications (Fig. 1). A detailed step-by-step protocol will be published separately.

Adapted INTACT procedure

The workflow can be divided into three separate sections: isolation of seeds with embryos of a known developmental stage, crude preparation of nuclei and purification of biotin-tagged nuclei. Recipes for buffers are presented in Supplementary Table 1.

For seed isolation, a set number of flowers from an appropriate NTF/BirA transgenic plant line are emasculated based on expected yield of bead-bound nuclei. Manual pollination of the exposed stigmas is performed the next day in less than 1 h to synchronize embryo development. We have consistently performed these steps with multiple people to minimize the time between the first and last pollination. Siliques are collected after a defined time interval at which they contain seeds with embryos at the development stage of interest. To avoid circadian rhythm effects, collection should be performed at the same time point of the day for each experiment. Siliques are adhered to slides with double-sided tape and seeds are exposed by opening the silique with a needle. Slides are then transferred to a 120 × 120 mm square Petri dish and placed under a stereo microscope. Next, seeds are collected from siliques onto a 20- μ m nylon net filter using a suctioning apparatus consisting of a Pasteur pipette, a 25-mm filter holder and a vacuum pump. To avoid damage and dehydration of the seeds, a vacuum pressure at or under 40 mbar is used and 1 × PBS (pH 7.0) is continuously added

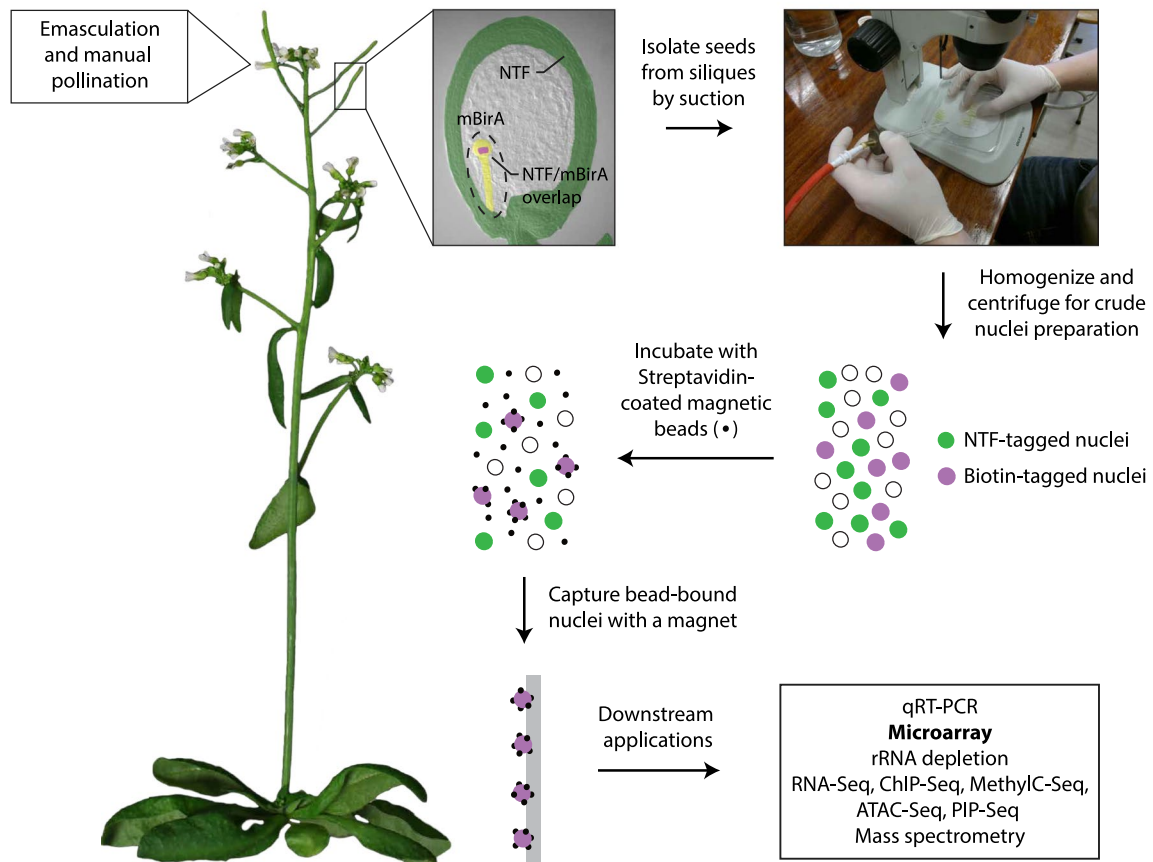


Fig. 1 Schematic diagram of adapted INTACT procedure used to isolate cell-type-specific nuclei and generate microarray-based transcriptomic profiles in the early *Arabidopsis* embryo. Microarray and other downstream applications that can be combined with the INTACT pro-

cedure are listed. This includes ChIP-Seq and PIP-Seq, which both require cross-linking of fresh tissue before INTACT purification of nuclei

to the Petri dish. Using this setup, it is possible to process siliques at a rate of 10 siliques per minute.

To prepare a crude nuclear extract, seeds are washed off the filter in 2-ml tubes containing $1 \times$ PBS. The buffer is removed with a pipette, and a 4.8-mm stainless steel ball is added to each tube before flash-freezing them in liquid nitrogen. Seeds are homogenized in a mixer mill with pre-cooled adaptor racks (30 Hertz for 2×30 s) and resuspended with 10 ml ice-cold nuclear purification buffer (NPB) and 200 units of a RNase inhibitor in a 15-ml tube. The suspension is filtered through a 40- μ m cell strainer and centrifuged at 1200g at 4 °C for 7 min to pellet nuclei. Nuclei are resuspended with 1 ml of NPB and transferred to a 1.5-ml tube. Ten microlitres of washed M-280 streptavidin-coated Dynabeads is added to the nuclei and incubated for 30 min at 4 °C with end-over-end rotation to bind beads to the nuclei.

Purification is performed by using a column system where the bead-bound nuclei are captured when they flow past a strong magnet. Columns are prepared by treating 1-ml pipette tips with NPB containing 1% (w/v) Casein for

20 min. Casein treatment prevents adhesion of non-bead-bound nuclei to the tip wall. A tip is inserted into an OctoMACS separator magnet placed vertically in a 4 °C cold room and then rinsed with ice-cold NPB containing 0.1% (v/v) Triton X-100 (NPBt). A two-way stopcock is attached to the narrow end of the pipette tip to control liquid flow rate. The bead and nuclei mixture is diluted with 9 ml NPBt and drawn into a plastic 10-ml serological pipette with a Parafilm-wrapped tip. The 10-ml pipette is fastened to the top of the 1-ml tip, and the mixture is flowed past the magnet at a rate of 0.75 ml min^{-1} . This allows for efficient capture of bead-bound nuclei to the tip wall. The 1-ml tip is removed when empty, and the inner tip wall is washed by drawing NPBt in and out without disturbing the attached beads. The bead-bound nuclei are rinsed with 1 ml NPBt and diluted with 9 ml NPBt for a second round of purification, which is necessary for achieving maximum purity. After washing and rinsing the tip a second time, the bead-bound nuclei are collected by centrifugation (1000g at 4 °C for 5 min) and resuspended in 25 μ l NPB. This solution contains the

purified biotin-tagged nuclei, which can be directly used in downstream applications.

The purity of biotin-tagged nuclei is determined by the ratio of DAPI-stained ($2 \mu\text{g ml}^{-1}$) bead-bound (biotinylated) and unbound (non-biotinylated) nuclei. Recovery efficiency is estimated by the ratio of bead-bound nuclei to the expected number of biotin-tagged nuclei in the starting material.

Generating transcriptomic profiles

One downstream application of the purified biotin-tagged nuclei is to generate and compare whole-genome expression data from various cell types. In Palovaara et al. (2017), this was performed using a microarray-based approach.

RNA is obtained from bead-bound nuclei isolated from different cell types and embryonic stages following TRIzol-based RNA extraction, DNase I treatment and RNA purification and concentration (Palovaara et al. 2017). To increase RNA amounts, reduce technical variances and avoid batch-to-batch effects (Clément-Ziza et al. 2009; Morse et al. 2010), RNA from biological replicates was pooled to 3–4 samples and then simultaneously amplified with an oligo(dT)/random primer mixture using a kit designed for small RNA quantities (Ovation Pico WTA System V2). This resulted in the amplification of both nascent transcripts and mature messenger RNA (mRNA). After labelling, amplified cDNA was hybridized to Arabidopsis Gene 1.1 ST 24-Array plates, which covers 84.2% (28,501) of all annotated genes (TAIR10). Analysis of the plates was performed using the MADMAX pipeline (Lin et al. 2011), with values normalized by RMA (Irizarry et al. 2003), to determine gene expression and differentially expressed genes. Further analyses were performed to determine dominant gene expression patterns by manual selection or hierarchical clustering (Orlando et al. 2009), biological significance of co-expressed genes (agriGO; Du et al. 2010; Tian et al. 2017) and enrichment of transcription factors (AtTFDB; Palaniswamy et al. 2006). Finally, a web-based tool was established at <http://www.albertodb.org> (ALBERTO) to display, compare and share the transcriptomic profiles.

In addition to our work, there are other publications concerning nuclear transcriptomics where RNA isolation, amplification methods and sequencing technologies differ (e.g. Deal and Henikoff 2010; Reynoso et al. 2018a, b; Slane et al. 2014; Zhang et al. 2008). For example, in Reynoso et al. (2018b) they demonstrate a method to remove ribosomal RNA (rRNA) from RNA isolated from INTACT-purified nuclei to produce a sample optimal for RNA-Seq. The preferred amplification method of nuclear RNA is oligo(dT)/random primer-based since it amplifies mRNA transcripts at different stages of processing. However, the highly abundant rRNAs are also amplified, which has a negative impact on

sequencing coverage. In theory, this method could be combined with our approach to generate high-quality RNA-Seq data.

Comparing nuclear and cellular transcriptomes

INTACT relies on profiling RNA not yet exported to the cytosol, which raises the question if nuclear RNA is representative for the transcriptomic output of a cell. Several studies, including our own, have established that nuclear RNA is a reasonable proxy for steady-state transcript levels regardless of the experimental approach (Deal and Henikoff 2010; Lake et al. 2017; Palovaara et al. 2017; Slane et al. 2015; Zhang et al. 2008). However, there are differences between nuclear and total transcriptomes, which suggest selective compartmentalization of RNA in the nucleus and the cytosol. This may impact cell fate during development since post-transcriptional regulation, critical for protein translation and expression level, typically takes place in the cytoplasm. Despite this, relatively few studies have investigated the differences between nuclear, cytosolic and total transcriptomes in detail, especially in plants (see e.g. Barthelson et al. 2007; Chen and van Steensel 2017; Djebali et al. 2012; Bahar Halpern et al. 2015; Reynoso et al. 2018a; Solnestam et al. 2012). Until recently, the primary reason for this has been the lack of technologies that facilitate such a comparison, particularly at the tissue or cellular level.

In Palovaara et al. (2017), we used INTACT and manual isolation to generate microarray-based nuclear and cellular RNA transcriptomes from whole *Arabidopsis* embryos at 16-cell stage to assess the reliability of nuclear RNA. Here, we perform a more comprehensive analysis of the data set to evaluate notable distinctions between the two RNA populations. When compared, a total of 3011 genes were significantly enriched in one or the other population, with 2180 genes being enriched in the cellular RNA and 831 genes being enriched in the nuclear RNA (Fig. 2, Supplementary Table 2a, b). Gene ontology (GO) enrichment analysis of the cellular population included “ribonucleoprotein complex biogenesis”, “ribosome biogenesis”, “translation” and “RNA processing”, among others (Fig. 2, Supplementary Table 2c). Ribonucleoprotein complexes consist of a diverse set of RNA-binding proteins (RBPs) that interact with RNAs to regulate post-transcriptional expression by influencing RNA biogenesis, stability, translation and subcellular localization (reviewed in Dreyfuss et al. 2002). In contrast, the nuclear population included “transmembrane transport”, “cell wall organization”, “signalling”, and terms associated with responses to various stimuli (Fig. 2, Supplementary Table 2c).

RNA processing and maturation is a unidirectional process. Pre-mRNAs are spliced in the nucleus, capped and poly-adenylated, and exported to the cytosol for translation.

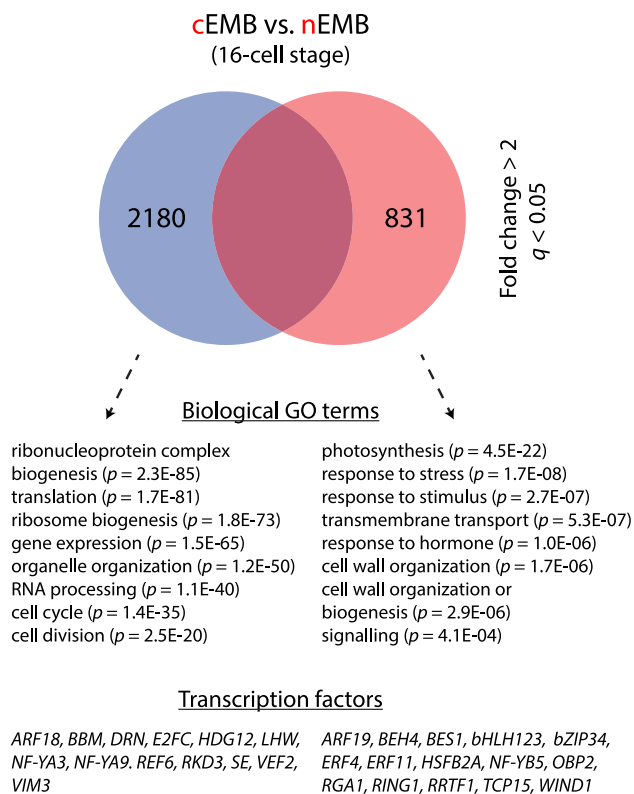


Fig. 2 Microarray-based comparison of whole embryo nuclear (INTACT; nEMB) and cellular (manual isolation; cEMB) RNA at 16-cell stage. Enriched genes (fold change > 2, $q < 0.05$) and selected enriched biological GO terms ($p \leq 0.001$) and transcription factors are presented. Note that only transcription factors that do not show tissue-specific enrichment, or are not present in the Belmonte et al. (2013) data set (see Fig. 4) are shown

Differences in nuclear/cytosolic abundance ratios between mRNAs may therefore derive from selective nuclear export, from stabilization or degradation, among others. A simple explanation for enrichment of a transcript in nuclear RNA pools would be that the mature mRNA is unstable and would quickly disappear from the cytosolic pool. Conversely, if an mRNA is exceptionally stable in the cytosol, its cytosolic levels will tend to be enriched. To determine if mRNA (in) stability could be a dominant cause of the observed enrichment, we investigated mRNA decay rates among cytosol- or nucleus-enriched transcripts using a mRNA decay data set from 5-day-old seedlings (Sorenson et al. 2018). We found that the cellular population had longer median mRNA half-life (142 min) compared to the nuclear population (92 min), with no apparent correlation to UTR lengths or coding region length (Supplementary Table 2c). This was especially evident when only transcription factors were analysed (101 vs. 41 min). This suggests that transcript enrichment in the cellular population may indeed reflect a relatively low decay rate. Earlier analyses have shown that high-flux RNAs associate with rapid signalling response pathways,

including communication, hormone response and biotic/abiotic stress-related response (Narsai et al. 2007; Sorenson et al. 2018). Together, this demonstrates genome-wide differences between nuclear and cellular RNA that yield distinct functional characteristics.

Given that different nuclear/cytoplasm mRNA partitioning and different mRNA stabilities may well translate to altered tissue specificities in the growing embryo, we next addressed the behaviour of cellular- and nuclear-enriched transcripts in relation to the atlas data set published in Palovaara et al. (2017). This revealed that median expression for genes enriched in the nuclear population was notably higher when compared to genes enriched in the cellular population in the atlas (18 vs. 10 signal intensity) (Supplementary Table 2d). This may therefore reflect a comparatively high transcription rate of this set of genes. Furthermore, when performing correlation and hierarchical clustering analysis, the atlas samples clustered according to developmental stage for nuclear but not for cellular-enriched genes (Fig. 3, Supplementary Table 2d), even though both sets of genes generally clustered according to cell type (Supplementary Fig. 1). Thus, cytoplasmically enriched transcripts do not appear to reflect the stage of the embryo as defined by stage of isolation. Increased mRNA stability is a likely cause for this observation.

The stage-specific clustering seen for the nuclear-enriched transcripts appears to be the result of rapidly changing expression levels between developmental stages: overall gene expression decreased from one stage (16-cell, early globular) to the next (late globular) as revealed by the expression heat map in Fig. 3b and a lower median expression value (21 vs. 14 signal intensity; Supplementary Table 2d). This, together with the low overall expression of the cellular-enriched genes in the atlas, supports the conclusion that RNAs with shorter half-lives and higher transcription rates are more common in the nuclear than the cellular RNA population. A list of “nuclear” or “cellular” genes that were enriched in the various cell types of the atlas is presented in Supplementary Table 2e.

To further refine the data set, we combined it with transcriptomic data from Belmonte et al. (2013) where they isolated the embryo proper and suspensor from globular embryos of *Arabidopsis* using laser capture microdissection (Supplementary Table 2f). This revealed tissue-specific patterns of GO terms and transcription factors enriched in the nuclear and cellular RNA (Fig. 3, Supplementary Table 2g, h). All mentioned GO terms enriched in the cellular population were also enriched in the embryo proper, which supports previous data indicating that RBRs are tissue specific in plants (Maronedez et al. 2016). Many transcription factors crucial for cell specification in the early embryo, such as *WOX2*, *HAN*, *ARF5/MP*, *ATML1*, *PHB* and *REV*, were similarly enriched, while *WOX8*, *WOX9*, *WRKY2* and *LEC1*

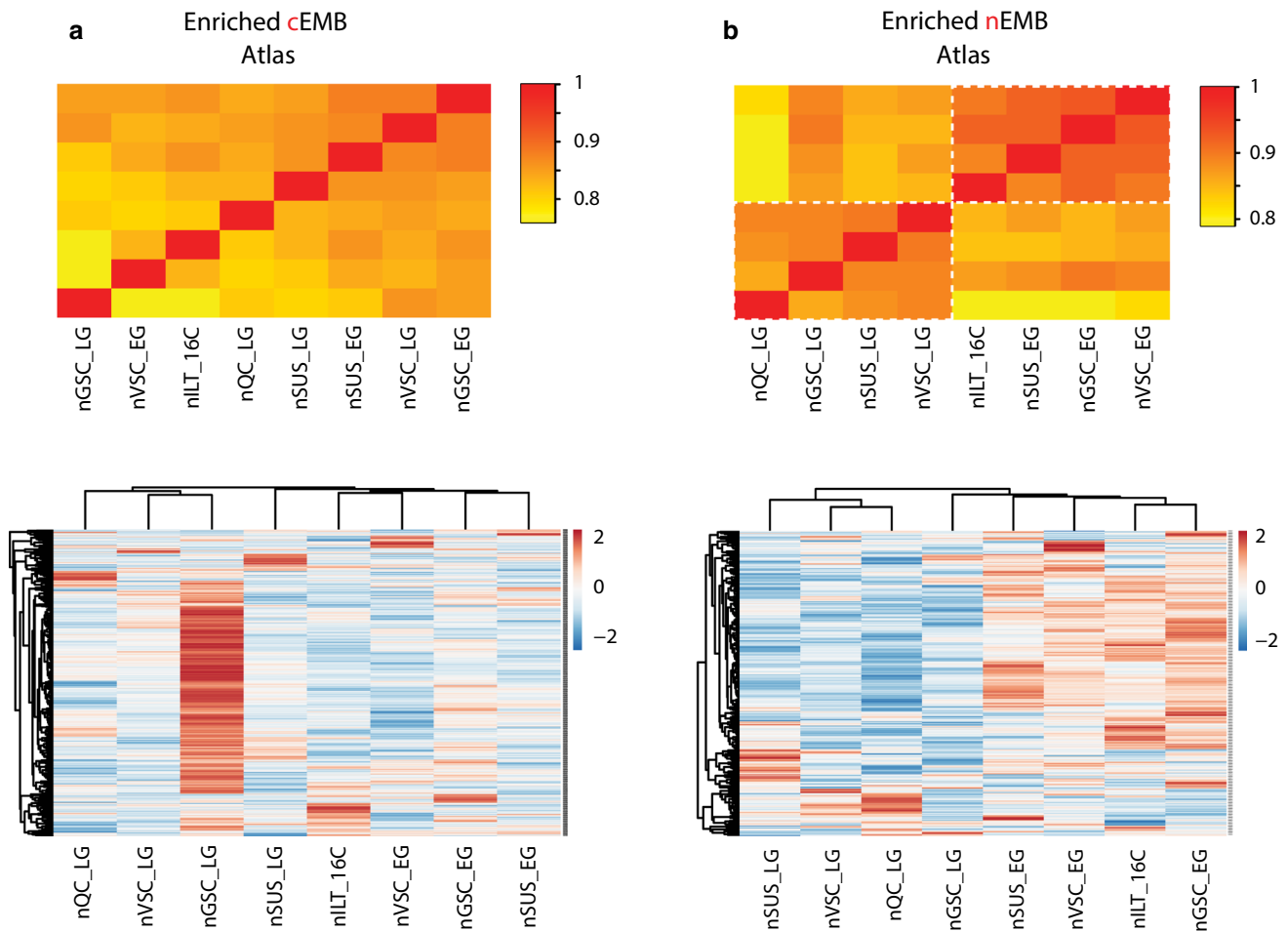


Fig. 3 Enriched cEMB (a) and nEMB (b) genes in cell types of the atlas data set presented in Palovaara et al. (2017). At top, heat map of Pearson's correlation coefficients (r) in pairwise comparisons using \log_2 -values where low correlation is in yellow and high correlation is in red. Samples are sorted according to the amount of correlation (low to high: left to right, top to bottom), which in b show a pattern specific to developmental stages (white rectangles). At bottom, hierarchical clustering and expression heat map of the 50% most vari-

able genes where values are \log -transformed, rows are centred, and both rows and columns are clustered using correlation distance and average linkage (Metsalu and Vilo 2015). Abbreviations: *c* cellular, *n* nuclear, *EMB* whole embryo, *16C* 16-cell stage, *EG* early globular stage, *LG* late globular stage, *ILT* inner lower tier, *VSC* vascular tissue precursors, *GSC* ground tissue precursors, *SUS* suspensor, *QC* quiescent centre precursor

were enriched in the suspensor (Lotan et al. 1998; Palovaara et al. 2016) (Fig. 4a). In the nuclear population, “cell wall organization”, “cell morphogenesis” and comparable GO terms were enriched in the embryo proper (Fig. 4b), suggesting that genes associated with these processes are actively and highly transcribed at this stage of development.

From these results, we propose that selective nucleocytoplasmic enrichment of RNAs, through nuclear retention and post-transcriptional regulation, is a tissue- or even a cell-type-specific characteristic that directly impacts cell fates in the early plant embryo. Indeed, it is already known that microRNAs are involved in the establishment and maintenance of the shoot apical meristem in the *Arabidopsis* embryo (Palovaara et al. 2016; Takanashi et al. 2018) and that targeted RNA degradation influences embryonic

stem cell fate in mammals (Li et al. 2015; Lou et al. 2014, 2016). However, to confirm our hypothesis, future work should compare the nuclear to the cytosolic RNA at a cellular resolution in wild-type embryos and embryos defective in post-transcriptional regulation. This could be achieved by using INTACT and another method, TRAP (translating ribosome affinity purification), to target the same tissue or cell type. TRAP allows for the isolation of actively translated transcripts after affinity purification of tagged ribosomal protein (Mustroph et al. 2009). Such a study would generate a comparative read-out of pre-processed (nuclear) and translated (cytosolic) mRNA, revealing the levels of regulation that direct gene expression in an individual cell. This read-out includes transcript sequence (size, codons, UTR regions) and isoforms, both of which can influence

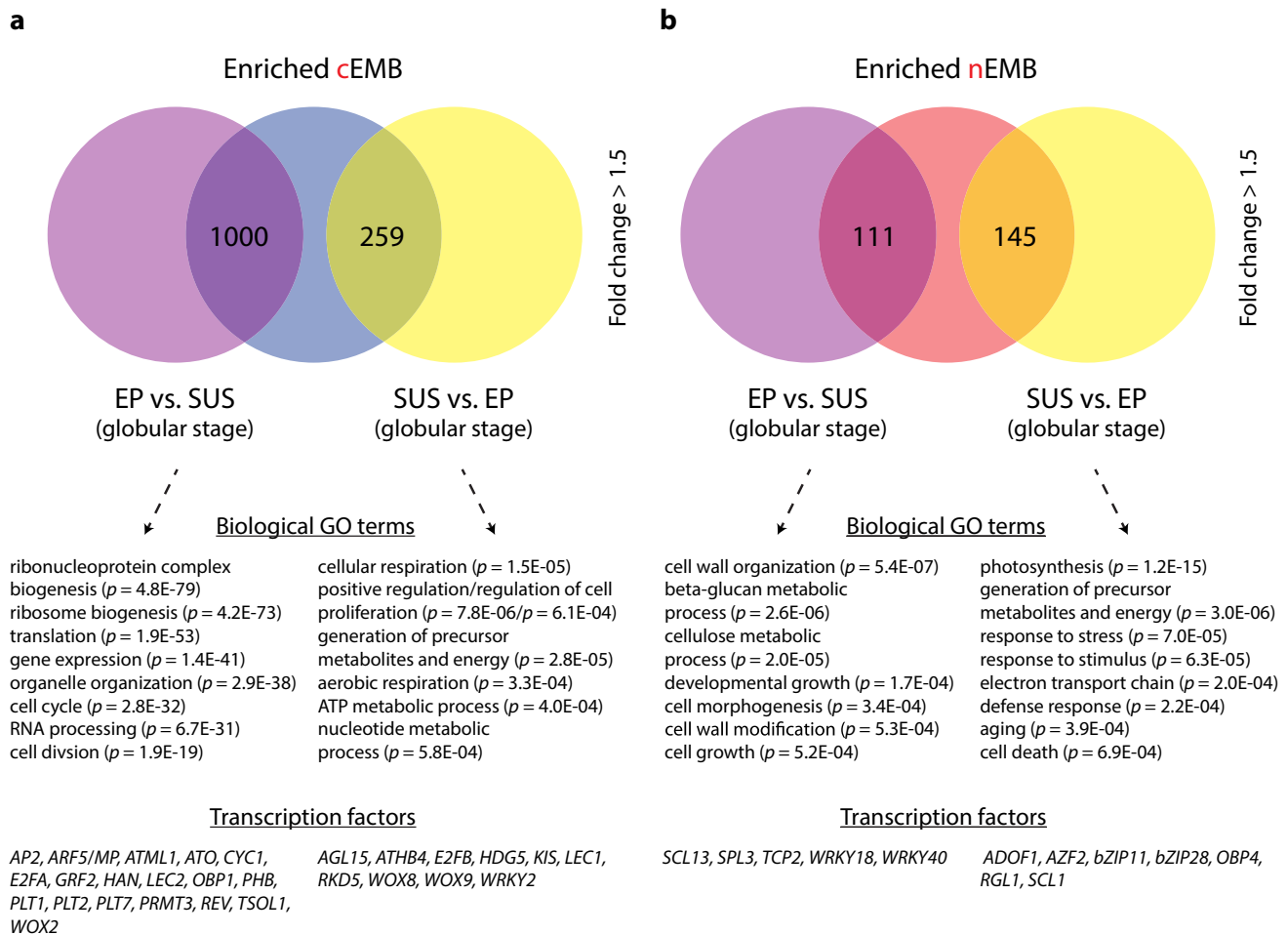


Fig. 4 Refinement of enriched cEMB and nEMB data sets based on globular stage embryo proper (EP) and suspensor (SUS) transcriptomic data from Belmonte et al. (2013). Fold change > 1.5 denotes enriched genes in the Belmonte et al. (2013) data set. Overlapping

genes and selected enriched biological GO terms ($p \leq 0.001$) and transcription factors are presented. Abbreviations: *c* cellular, *n* nuclear, *EMB* whole embryo

post-transcriptional regulation (e.g. Chen 2010; Hartmann et al. 2018; Theil et al. 2018), if RNA-Seq is used as the primary sequencing platform.

Outlook

INTACT is a versatile tool that has been used for transcriptomic, epigenomic and proteomic studies of tissue- and cell-type-specific nuclei in plants and animals (Amin et al. 2014; Foley et al. 2017; Henry et al. 2012; Moreno-Romero et al. 2017; Park et al. 2016; Reynoso et al. 2018a; Ron et al. 2014; Steiner et al. 2012). In fact, a recent publication showed that several of these -omics studies can be performed on the same pool of INTACT-isolated nuclei (Mo et al. 2015). Here, we have presented how INTACT can be used to provide a snapshot of the nuclear transcriptome at a cellular resolution in the early *Arabidopsis* embryo. This has provided valuable information regarding mRNA synthesis. However, to fully

explore the transcriptional networks that govern cell fate in the plant embryo we need to capture the full dynamics of the mRNA life cycle, especially at the cellular resolution. Therefore, the next logical step is to use INTACT in conjunction with a method such as TRAP and other -omics approaches to evaluate multiple levels of mRNA regulation. This would be a powerful addition to current research in plant embryos, where focus is on how cell fates are reprogrammed to establish the first tissues of the plant.

Acknowledgements The authors thank Yanbo Mao and Tatyana Radoeva for providing the *Arabidopsis* plant and seed image, respectively, in Fig. 1. This work was supported by the Federation of European Biochemical Societies (FEBS) to J.P. and by the European Research Council (ERC; Starting Grant “CELLPATTERN”; Contract number 281573) and ERA-CAPS (EURO-PEC; 849.13.006) to D.W.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use,

distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Amin NM, Greco TM, Kuchenbrod LM, Rigney MM, Chung M-I, Wallingford JB, Cristea IM, Conlon FL (2014) Proteomic profiling of cardiac tissue by isolation of nuclei tagged in specific cell types (INTACT). *Development* 141:962–973. <https://doi.org/10.1242/dev.098327>
- Bahar Halpern K, Caspi I, Lemze D, Levy M, Landen S, Elinav E, Ulitsky I, Itzkovitz S (2015) Nuclear retention of mRNA in mammalian tissues. *Cell Rep* 13:2653–2662. <https://doi.org/10.1016/j.celrep.2015.11.036>
- Barthelson RA, Lambert GM, Vanier C, Lynch RM, Galbraith DW (2007) Comparison of the contributions of the nuclear and cytoplasmic compartments to global gene expression in human cells. *BMC Genom* 8:340. <https://doi.org/10.1186/1471-2164-8-340>
- Belmonte MF, Kirkbride RC, Stone SL, Pelletier JM, Bui AQ, Yeung EC, Hashimoto M, Fei J, Harada CM, Munoz MD, Le BH, Drews GN, Brady SM, Goldberg RB, Harada JJ (2013) Comprehensive developmental profiles of gene activity in regions and subregions of the *Arabidopsis* seed. *PNAS* 110:E435–E444. <https://doi.org/10.1073/pnas.1222061110>
- Casson S, Spencer M, Walker K, Lindsey K (2005) Laser capture microdissection for the analysis of gene expression during embryogenesis of *Arabidopsis*. *Plant J* 42:111–123. <https://doi.org/10.1111/j.1365-313X.2005.02355.x>
- Chen L (2010) A global comparison between nuclear and cytosolic transcriptomes reveals differential compartmentalization of alternative transcript isoforms. *Nucleic Acids Res* 38:1086–1097. <https://doi.org/10.1093/nar/gkp1136>
- Chen T, van Steensel B (2017) Comprehensive analysis of nucleocytoplasmic dynamics of mRNA in *Drosophila* cells. *PLoS Genet* 13:e1006929. <https://doi.org/10.1371/journal.pgen.1006929>
- Clément-Ziza M, Gentien D, Lyonnet S, Thiery J-P, Besmond C, Decraene C (2009) Evaluation of methods for amplification of picogram amounts of total RNA for whole genome expression profiling. *BMC Genom* 10:246. <https://doi.org/10.1186/1471-2164-10-246>
- Deal RB, Henikoff S (2010) A simple method for gene expression and chromatin profiling of individual cell types within a tissue. *Dev Cell* 18:1030–1040. <https://doi.org/10.1016/j.devcel.2010.05.013>
- Deal RB, Henikoff S (2011) The INTACT method for cell type-specific gene expression and chromatin profiling in *Arabidopsis thaliana*. *Nat Protoc* 6:56–68. <https://doi.org/10.1038/nprot.2010.175>
- Djebali S et al (2012) Landscape of transcription in human cells. *Nature* 489:101–108. <https://doi.org/10.1038/nature11233>
- Dreyfuss G, Kim VN, Kataoka N (2002) Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol* 3:195–205. <https://doi.org/10.1038/nrm760>
- Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) Agrigo: a go analysis toolkit for the agricultural community. *Nucleic Acids Res* 38:W64–W70. <https://doi.org/10.1093/nar/gkq310>
- Foley SW, Gosai SJ, Wang D, Selamoglu N, Solitti AC, Köster T, Steffen A, Lyons E, Daldal F, Garcia BA, Staiger D, Deal RB, Gregory BD (2017) A global view of RNA-protein interactions reveals novel root hair cell fate regulators. *Dev Cell* 41:204–220. <https://doi.org/10.1016/j.devcel.2017.03.018>
- Hartmann L, Wießner T, Wachter A (2018) Subcellular compartmentation of alternatively-spliced transcripts defines SERINE/ARGININE-RICH PROTEIN 30 expression. *Plant Physiol* 176:2886–2903. <https://doi.org/10.1104/pp.17.01260>
- Henry GL, Davis FP, Picard S, Eddy SR (2012) Cell type-specific genomics of *Drosophila* neurons. *Nucleic Acids Res* 40:9691–9704. <https://doi.org/10.1093/nar/gks671>
- Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4:249–264. <https://doi.org/10.1093/biostatistics/4.2.249>
- Lake BB, Codeluppi S, Yung YC, Gao D, Chun J, Kharchenko PV, Linnarsson S, Zhang K (2017) A comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA. *Sci Rep* 7:6031. <https://doi.org/10.1038/s41598-017-04426-w>
- Li T, Shi Y, Wang P, Guachalla LM, Sun B, Joeris T, Chen Y-S, Groth M, Krueger A, Platzer M, Yang Y-G, Rudolph KL, Wang Z-Q (2015) Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. *EMBO J* 34:1630–1647. <https://doi.org/10.15252/emboj.201489947>
- Lin K et al (2011) MADMAX—Management and analysis database for multiple -omics experiments. *J Integr Bioinform* 8:160
- Lotan T, M-a Ohto, Yee KM, West MAL, Lo R, Kwong RW, Yamagishi K, Fischer RL, Goldberg RB, Harada JJ (1998) *Arabidopsis* LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* 93:1195–1205. [https://doi.org/10.1016/S0092-8674\(00\)81463-4](https://doi.org/10.1016/S0092-8674(00)81463-4)
- Lou CH, Shao A, Shum EY, Espinoza JL, Huang L, Karam R, Wilkinson MF (2014) Posttranscriptional control of the stem cell and neurogenic programs by the nonsense-mediated RNA decay pathway. *Cell Rep* 6:748–764. <https://doi.org/10.1016/j.celrep.2014.01.028>
- Lou C-H, Dumdie J, Goetz A, Shum EY, Brafman D, Liao X, Moracastilla S, Ramaiah M, Cook-Andersen H, Laurent L, Wilkinson MF (2016) Nonsense-mediated RNA decay influences human embryonic stem cell fate. *Stem Cell Rep* 6:844–857. <https://doi.org/10.1016/j.stemcr.2016.05.008>
- Marondedze C, Thomas L, Serrano NL, Lilley KS, Gehring C (2016) The RNA-binding protein repertoire of *Arabidopsis thaliana*. *Sci Rep* 6:29766. <https://doi.org/10.1038/srep29766>
- Metsalu T, Vilo J (2015) Clustvis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic Acids Res* 43:W566–W570. <https://doi.org/10.1093/nar/gkv468>
- Mo A, Mukamel EA, Davis FP, Luo C, Henry GL, Picard S, Urich MA, Nery JR, Sejnowski TJ, Lister R, Eddy SR, Ecker JR, Nathans J (2015) Epigenomic signatures of neuronal diversity in the mammalian brain. *Neuron* 86:1369–1384. <https://doi.org/10.1016/j.neuron.2015.05.018>
- Moreno-Romero J, Santos-González J, Hennig L, Köhler C (2017) Applying the INTACT method to purify endosperm nuclei and to generate parental-specific epigenome profiles. *Nat Protoc* 12:238–254. <https://doi.org/10.1038/nprot.2016.167>
- Morse AM, Carballo V, Baldwin DA, Taylor CG, McIntyre LM (2010) Comparison between NuGEN's WT-Ovation Pico and One-Direct amplification systems. *J Biomol Tech* 21:141–147
- Mustroph A, Zanetti ME, Jang CJH, Holtan HE, Repetti PP, Galbraith DW, Girke T, Bailey-Serres J (2009) Profiling transcriptomes of discrete cell populations resolves altered cellular priorities during hypoxia in *Arabidopsis*. *PNAS* 106:18843–18848. <https://doi.org/10.1073/pnas.0906131106>
- Narsai R, Howell KA, Millar AH, O'Toole N, Small I, Whelan J (2007) Genome-wide analysis of mRNA decay rates and their

- determinants in *Arabidopsis thaliana*. *Plant Cell* 19:3418–3436. <https://doi.org/10.1105/tpc.107.055046>
- Orlando DA, Brady SM, Koch JD, Dinnyen JR, Benfey PN (2009) Manipulating large-scale *Arabidopsis* microarray expression data: identifying dominant expression patterns and biological process enrichment. *Methods Mol Biol* 553:57–77. https://doi.org/10.1007/978-1-60327-563-7_4
- Palaniswamy SK, James S, Sun H, Lamb RS, Davuluri RV, Grotewold E (2006) AGRIS and AtRegNet. A platform to link cis-regulatory elements and transcription factors into regulatory networks. *Plant Physiol* 140:818–829. <https://doi.org/10.1104/pp.105.072280>
- Palovaara J, Zeeuw Td, Weijers D (2016) Tissue and organ initiation in the plant embryo: a first time for everything. *Annu Rev Cell Dev Biol* 32:47–75. <https://doi.org/10.1146/annurev-cellbio-111315-124929>
- Palovaara J, Saiga S, Wendrich JR, van 't Wout Hofland N, van Schayck JP, Hater F, Mutte S, Sjollem J, Boekschoten M, Hooiveld GJ, Weijers D (2017) Transcriptome dynamics revealed by a gene expression atlas of the early *Arabidopsis embryo*. *Nat Plants* 3:894–904. <https://doi.org/10.1038/s41477-017-0035-3>
- Park K, Kim MY, Vickers M, Park J-S, Hyun Y, Okamoto T, Zilberman D, Fischer RL, Feng X, Choi Y, Scholten S (2016) DNA demethylation is initiated in the central cells of *Arabidopsis* and rice. *PNAS* 113:15138–15143. <https://doi.org/10.1073/pnas.1619047114>
- Reynoso MA, Pauluzzi GC, Kajala K, Cabanlit S, Velasco J, Bazin J, Deal R, Sinha NR, Brady SM, Bailey-Serres J (2018a) Nuclear transcriptomes at high resolution using retooled INTACT. *Plant Physiol* 176:270–281. <https://doi.org/10.1104/pp.17.00688>
- Reynoso MA, Pauluzzi GC, Cabanlit S, Velasco J, Bazin J, Deal R, Brady S, Sinha N, Bailey-Serres J, Kajala K (2018b) Isolation of nuclei in tagged cell types (INTACT), RNA extraction and ribosomal RNA degradation to prepare material for RNA-Seq. *Bio Protoc* 8:2458. <https://doi.org/10.21769/BioProtoc.2458>
- Ron M, Kajala K, Pauluzzi G, Wang D, Reynoso MA, Zumstein K, Garcha J, Winte S, Masson H, Inagaki S, Federici F, Sinha N, Deal RB, Bailey-Serres J, Brady SM (2014) Hairy root transformation using *Agrobacterium rhizogenes* as a tool for exploring cell type-specific gene expression and function using tomato as a model. *Plant Physiol* 166:455–469. <https://doi.org/10.1104/pp.114.239392>
- Slane D, Kong J, Berendzen KW, Kilian J, Henschen A, Kolb M, Schmid M, Harter K, Mayer U, De Smet I, Bayer M, Jürgens G (2014) Cell type-specific transcriptome analysis in the early *Arabidopsis thaliana* embryo. *Development* 141:4831–4840. <https://doi.org/10.1242/dev.116459>
- Slane D, Kong J, Schmid M, Jürgens G, Bayer M (2015) Profiling of embryonic nuclear vs. cellular RNA in *Arabidopsis thaliana*. *Genom Data* 4:96–98. <https://doi.org/10.1016/j.gdata.2015.03.015>
- Solnestam BW, Stranneheim H, Hällman J, Källner M, Lundberg E, Lundeberg J, Akan P (2012) Comparison of total and cytoplasmic mRNA reveals global regulation by nuclear retention and miRNAs. *BMC Genom* 13:574. <https://doi.org/10.1186/1471-2164-13-574>
- Sorenson RS, Deshotel MJ, Johnson K, Adler FR, Sieburth LE (2018) *Arabidopsis* mRNA decay landscape arises from specialized RNA decay substrates, decapping-mediated feedback, and redundancy. *PNAS* 115:E1485–E1494. <https://doi.org/10.1073/pnas.1712312115>
- Steiner FA, Talbert PB, Kasinathan S, Deal RB, Henikoff S (2012) Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling. *Genome Res* 22:766–777. <https://doi.org/10.1101/gr.131748.111>
- Takanashi H, Sumiyoshi H, Mogi M, Hayashi Y, Ohnishi T, Tsutsumi N (2018) MiRNAs control HAM1 functions at the single-cell-layer level and are essential for normal embryogenesis in *Arabidopsis*. *Plant Mol Biol* 96:627–640. <https://doi.org/10.1007/s11103-018-0719-8>
- Theil K, Herzog M, Rajewsky N (2018) Post-transcriptional regulation by 3' UTRs can be masked by regulatory elements in 5' UTRs. *Cell Rep* 22:3217–3226. <https://doi.org/10.1016/j.celrep.2018.02.094>
- Tian T, Liu Y, Yan H, You Q, Yi X, Du Z, Xu W, Su Z (2017) Agrigo v2.0: a go analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res* 45:W122–W129. <https://doi.org/10.1093/nar/gkx382>
- Zhang C, Barthelson RA, Lambert GM, Galbraith DW (2008) Global characterization of cell-specific gene expression through fluorescence-activated sorting of nuclei. *Plant Physiol* 147:30–40. <https://doi.org/10.1104/pp.107.115246>